



UNIVERSIDADE CATÓLICA PORTUGUESA

“THE ROLE OF AI IN BATTLING DISINFORMATION AND FAKE NEWS
ON SOCIAL MEDIA: HOW CAN AI BE USED TO IDENTIFY FAKE NEWS
ON SOCIAL MEDIA FROM THE USERS PERSPECTIVE?”

DISSERTATION

Dissertation to Universidade Católica Portuguesa to obtain a
Master's Degree in Marketing, Communication and Advertising

By

Cecilie Katharina Mössle

Universidade Católica Portuguesa

September 2024



UNIVERSIDADE CATÓLICA PORTUGUESA

“THE ROLE OF AI IN BATTLING DISINFORMATION AND FAKE NEWS
ON SOCIAL MEDIA: HOW CAN AI BE USED TO IDENTIFY FAKE NEWS
ON SOCIAL MEDIA FROM A USERS PERSPECTIVE?”

DISSERTATION

Dissertation to Universidade Católica Portuguesa to obtain a
Master’s Degree in Master’s Degree in Marketing, Communication
and Advertising

Cecilie Katharina Mössle

Universidade Católica Portuguesa

Under the supervision of Prof. Fernando Ilharco

September 2024

Abstract

This dissertation examines the role of artificial intelligence in the fight against disinformation and fake news on social media. Over the past 15 years, social media have become an important media that have revolutionized the way information is disseminated. Not only is information disseminated in a matter of seconds, but it is as well published and distributed by anyone with access to social media. And precisely because the boundaries between information producers and information consumers are blurred by social media, there is an increase in fake news on social media. This also gave rise to the central question of this research: “The role of AI in battling disinformation and fake news on social media: How can AI be used to identify fake news on social media from a users perspective?”

This work provides insights into the challenge of using AI as well as analyzes the risks that arise both from the social media user and from the use of AI. The findings of this research show that the use of AI is promising and indispensable today, as it is a great help in the fight against fake news. However, they also show that AI still has a considerable need for development so that it may become more reliable.

Key Words: Artificial Intelligence (AI), Fake News, Disinformation, Social Media, Algorithms, Filter Bubbles, Fact-Checking

Resumo

Esta dissertação analisa o papel da inteligência artificial na luta contra a desinformação e as notícias falsas nas redes sociais. Nos últimos 15 anos, as redes sociais tornaram-se um media que revolucionou a forma como a informação é divulgada. Não só a informação pode ser divulgada numa questão de segundos, como pode também ser publicada e distribuída por qualquer pessoa com acesso às redes sociais. Precisamente porque as fronteiras entre produtores e consumidores de informação são esbatidas pelas redes sociais, existe um aumento de notícias falsas. Este facto deu também origem à questão central desta investigação: “O papel da IA no combate à desinformação e às notícias falsas nas redes sociais: Como pode a IA ser utilizada para identificar notícias falsas nas redes sociais?”

Esta investigação fornece uma visão detalhada do desafio da utilização da IA, analisando os riscos que surgem tanto pelo simples uso das redes sociais como os que podem surgir na utilização da IA. O resultado deste estudo mostra que a utilização da IA é promissora e indispensável atualmente, uma vez que é uma ajuda importante na luta contra as notícias falsas. No entanto, mostramos igualmente que a IA ainda tem uma necessidade considerável de desenvolvimento para se tornar mais fiável.

Palavras-chave: Inteligência Artificial (IA), Notícias Falsas, Desinformação, Redes Sociais, Algoritmos, Bolhas de Filtro, Verificação de Factos

Table of Content

| | |
|--|-----------|
| ABSTRACT | 3 |
| RESUMO | 3 |
| TABLE OF FIGURES | 8 |
| INTRODUCTION | 9 |
| 1. SOCIAL MEDIA | 12 |
| 1.1. INTRODUCTION TO SOCIAL MEDIA | 12 |
| 1.1.1. DEFINITION AND OVERVIEW OF SOCIAL MEDIA | 12 |
| 1.1.2. HISTORICAL DEVELOPMENT AND EVOLUTION OF SOCIAL MEDIA PLATFORMS | 14 |
| 1.2. IMPACT OF SOCIAL MEDIA ON COMMUNICATION | 15 |
| 1.2.1. CHANGES IN COMMUNICATION PATTERNS AND STYLES | 15 |
| 1.2.2. THE ROLE OF SOCIAL MEDIA IN INTERPERSONAL RELATIONSHIPS | 16 |
| 1.3. SOCIAL MEDIA AND INFORMATION DISSEMINATION | 17 |
| 1.3.1. HOW SOCIAL MEDIA HAS TRANSFORMED INFORMATION SHARING | 17 |
| 1.3.2. INFLUENCE OF SOCIAL MEDIA ON NEWS DISTRIBUTION AND CONSUMPTION | 19 |
| 1.4. USER BEHAVIOR AND PSYCHOLOGY ON SOCIAL MEDIA | 20 |
| 1.4.1. USER MOTIVATIONS FOR USING SOCIAL MEDIA | 20 |
| 1.4.2. PSYCHOLOGICAL EFFECTS OF SOCIAL MEDIA USE | 21 |
| 1.5. SOCIAL MEDIA AND SOCIETY | 23 |
| 1.5.1. INFLUENCE OF SOCIAL MEDIA ON POLITICS, ACTIVISM, AND PUBLIC DISCOURSE | 23 |
| 1.5.2. ETHICAL AND PRIVACY CONCERNS RELATED TO SOCIAL MEDIA | 25 |
| 1.6. SOCIAL MEDIA MARKETING AND BUSINESS | 26 |
| 1.6.1. HOW BUSINESSES LEVERAGE SOCIAL MEDIA FOR MARKETING AND ADVERTISING | 26 |
| 1.6.2. THE IMPACT OF SOCIAL MEDIA ON E-COMMERCE AND CUSTOMER ENGAGEMENT | 27 |
| 1.7. FUTURE TRENDS IN SOCIAL MEDIA | 28 |
| 1.7.1. EMERGING TRENDS IN SOCIAL MEDIA DEVELOPMENT | 28 |
| 2. DISINFORMATION AND FAKE NEWS | 31 |
| 2.1. INTRODUCTION TO DISINFORMATION AND FAKE NEWS | 31 |
| 2.1.1. DEFINITIONS AND DISTINCTIONS BETWEEN DISINFORMATION, MISINFORMATION, AND FAKE NEWS..... | 31 |
| 2.1.2. THE SIGNIFICANCE OF STUDYING DISINFORMATION | 32 |
| 2.2. DETECTION OF DISINFORMATION | 33 |
| 2.2.1. THE DANGER OF DISINFORMATION | 34 |
| 2.2.2. HISTORICAL EXAMPLES OF DISINFORMATION CAMPAIGNS | 35 |
| 2.3. THE ROLE OF SOCIAL MEDIA IN DISINFORMATION | 36 |
| 2.3.1. HOW SOCIAL MEDIA PLATFORMS COULD CONTRIBUTE TO THE SPREAD OF DISINFORMATION | 37 |
| 2.3.2. THE IMPACT OF ALGORITHMS, ECHO CHAMBERS, AND FILTER BUBBLES | 38 |

| | |
|--|-----------|
| 2.4. CONSEQUENCES AND IMPACT OF DISINFORMATION | 39 |
| 2.4.1. EFFECTS ON PUBLIC PERCEPTION, DECISION-MAKING, AND SOCIETY..... | 39 |
| 2.4.2. REAL-WORLD CONSEQUENCES AND NOTABLE CASE STUDIES | 40 |
| 2.5. DETECTION AND MITIGATION STRATEGIES | 41 |
| 2.5.1. FACT-CHECKING, VERIFICATION METHODS, AND TOOLS | 42 |
| 2.5.2. LEGAL, REGULATORY, AND POLICY RESPONSES TO DISINFORMATION | 42 |
| 2.6. ETHICAL CONSIDERATIONS AND CHALLENGES | 43 |
| 2.6.1. ETHICAL DILEMMAS IN ADDRESSING DISINFORMATION | 44 |
| 2.6.2. BALANCING FREE SPEECH WITH DISINFORMATION CONTROL..... | 44 |
| 2.7. FUTURE DIRECTIONS IN COUNTERING DISINFORMATION..... | 45 |
| 2.7.1. EMERGING TRENDS IN DISINFORMATION AND RESPONSE STRATEGIES | 46 |

CHAPTER 3: ARTIFICIAL INTELLIGENCE **47**

3. AI (ARTIFICIAL INTELLIGENCE) **48**

| | |
|--|-----------|
| 3.1 INTRODUCTION TO AI..... | 48 |
| 3.1.1. DEFINITION AND OVERVIEW OF AI | 48 |
| 3.1.2. HISTORICAL DEVELOPMENT AND MILESTONES IN AI | 50 |
| 3.2. TYPES AND APPLICATIONS OF AI..... | 51 |
| 3.2.1. DIFFERENT CATEGORIES OF AI (E.G., NARROW AI, GENERAL AI) | 51 |
| 3.2.2. PRACTICAL APPLICATIONS OF AI ACROSS VARIOUS INDUSTRIES | 52 |
| 3.3 AI AND SOCIAL MEDIA..... | 53 |
| 3.3.1. THE ROLE OF AI IN SOCIAL MEDIA PLATFORMS | 53 |
| 3.3.2. AI-DRIVEN CONTENT RECOMMENDATION AND USER INTERACTION | 54 |
| 3.4. ETHICAL AND SOCIAL IMPLICATIONS OF AI | 55 |
| 3.4.1. THE IMPACT OF AI ON EMPLOYMENT, PRIVACY, AND SOCIETY | 57 |
| 3.5 CHALLENGES IN AI DEVELOPMENT..... | 59 |
| 3.5.1. TECHNICAL AND ETHICAL CHALLENGES IN AI RESEARCH AND IMPLEMENTATION..... | 59 |
| 3.5.2. REGULATION, POLICY, AND ETHICAL CONSIDERATIONS | 61 |
| 3.6 FUTURE TRENDS AND RESEARCH DIRECTIONS IN AI | 62 |
| 3.6.1. EMERGING TRENDS IN AI TECHNOLOGY | 62 |

CHAPTER 4: METHODOLOGY **63**

4. METHODOLOGY PART **64**

| | |
|--|-----------|
| 4.1 OUTLINE OF THE OBJECTIVES..... | 64 |
| 4.1.1 MIXED-METHOD APPROACH..... | 65 |
| 4.1.2. RESEARCH DESIGN | 66 |
| 4.1.3. DESCRIPTION OF THE RESEARCH DESIGN | 66 |
| 4.1.4. JUSTIFICATION OF THE RESEARCH DESIGN | 66 |
| 4.1.5. EXPLANATION OF THE RESEARCH DESIGN..... | 67 |
| 4.2 CHOOSING THE TARGET GROUP | 68 |
| 4.2.1. THE CHOSEN SAMPLING FRAME..... | 69 |
| 4.2.2. THE CHOSEN SAMPLING METHOD FOR THE SURVEY | 70 |
| 4.2.3. SAMPLING SIZE DETERMINATION | 70 |

| | |
|---|----------------|
| 4.3 SURVEY STRUCTURE AND EXPLANATION | 71 |
| 4.3.1 DATA COLLECTION METHOD | 73 |
| 4.3.2 JUSTIFICATION OF CHOOSING A SURVEY | 75 |
| 4.3. INTERVIEW SECTION..... | 76 |
| 4.3.1 IMPORTANCE OF INTERVIEWS IN RESEARCH | 76 |
| 4.3.2 DEVELOPMENT OF INTERVIEW PROTOCOL | 77 |
| 4.3.3 DESIGNING INTERVIEW QUESTIONS..... | 78 |
| 4.3.4 SELECTION OF PARTICIPANTS..... | 78 |
| 4.3.5 CRITERIA FOR PARTICIPANT SELECTION..... | 80 |
| 4.3.6 INFORMED CONSENT | 82 |
| 4.4 DATA COLLECTION METHOD | 82 |
| 4.4.1 SETTING FOR INTERVIEWS | 83 |
| 4.4.2 CONDUCTING THE INTERVIEWS | 84 |
| 4.4.3 RECORDING AND TRANSCRIPTION..... | 84 |
| 4.5 METHODOLOGY RESULTS | 85 |
| 4.5.1 DATA ANALYSIS SURVEY | 85 |
| 4.5.2 RESULTS OF INTERVIEWEE 1 | 91 |
| 4.5.3 RESULTS OF INTERVIEWEE 2 | 93 |
| 4.5.4 COMPARISON OF BOTH INTERVIEWS | 96 |
| 4.5.5 FINAL ANALYSIS AND REFLECTION | 98 |
| 4.5.6 THEORETICAL ASSESSMENT OF FINDINGS | 100 |
| CONCLUSION | 105 |
| LIMITATION..... | 106 |
| OUTLOOK..... | 107 |
| BIBLIOGRAPHY:..... | 109 |
| APPENDIX..... | 120 |

Table of Figures

| | |
|--|----|
| Figure 1: How to choose the best social media platform for your business. Source: Greene (2019)..... | 13 |
| Figure 2: AP Twitter hacked; no attack at White House, (Today, 2013) | 25 |
| Figure 3: Elections 2016: Trumps fere media helped keep cost down, but fewer doners provided more of the cash, (Sultan, 2017)..... | 32 |
| Figure 4: Dr. Oz criticized for saying panic about coronavirus is worse than the virus itself, (Crump, 2020)..... | 41 |
| Figure 5: Who owns AI ChatGPT, (DeepBrain AI, n.d.)..... | 49 |
| Figure 6: Screenshot of Google Forms showing the introduction to the survey, Source: https://docs.google.com/forms/d/1_n2cG1smWDpqGqfXctiLFNX0oMb7TC5JO1GyRZv2xG4/edit | 75 |
| Figure 7: Picture of Christoph Schumacher, Source: https://www.linkedin.com/feed/ | 80 |
| Figure 8: Picture of Jonny Preuss, Source: https://www.linkedin.com/feed/ . | 80 |

Introduction

Donald Trump, coronavirus, war in Ukraine. We all know these topics and headlines from standard media such as television, radio and social media. Imagine, news can be distributed across the globe in a matter of seconds, anyone can edit this news and play journalist for a moment or even get involved as an activist. This privilege, which in the past was only available to people who have been trained in the field for years and know how to do proper research and find the right information, is now also available to laypeople through social media platforms who indiscriminately put information out into the world. And this is exactly what makes the internet and social media so dangerous. How is it still possible for normal social media users to distinguish between true and false on platforms where anyone can freely express their opinion?

In the last 20 years, communication has changed rapidly, both interpersonally and politically. Thanks to platforms such as Facebook, Twitter and Instagram, people can now be located anywhere in the world and still communicate with each other and stay connected in a matter of seconds (Bala, 2014). This development can be defined as a technological revolution. However, it must be clear that such revolutions, especially those that happen so quickly and on such a huge scale, always bring challenges with them. As already mentioned, the biggest change is that producers and consumers become one and the same person and accordingly, as most people have no training in how to search for information, it can quickly happen that information is communicated incorrectly or even information is communicated that is simply not true. The latest technological revolution, which has been available to the public for a few years now, is artificial intelligence. Artificial intelligence is basically a technology that has been created and trained by humans and therefore machines and computer programs are being developed to help automate human tasks and, in the best-case scenario, replace them as much as possible and, above all, do the work of humans faster than would even be possible for humans. To do this, AI uses external data to make the work as precise and human-like as possible. This technology is now so well developed that it is impossible to imagine everyday life without it. Due to its fast and smart functioning and its ability to learn quickly, artificial intelligence has had an extreme impact on industries such as healthcare, supply chains, but also on the media industry (PK, F.A., 1984). The topic of fake news is also more topical than ever, as journalists are no longer needed to get news out into the world. What's more, the news that is spread by normal social media users on social media is not checked by anyone,

as everyone simply needs access to social media and a profile through which the news can be posted. So, if this person reads facts that might be false, does not question them and posts them anyway, this false news could reach many thousands of people and misinform them about issues. Now voices are being raised that AI could be another danger in connection with the spread of fake news. Since artificial intelligence is so new but is also just a machine and therefore mistakes can happen, many assume that AI could lead to the spread of fake news instead of recognizing and curbing it. These statements also result in the research question of this thesis, which is as follows: “The role of AI in battling disinformation and fake news on social media: How can AI be used to identify fake news on social media?”

Through the research and the various research methods used in this thesis, the aim is to find out whether artificial intelligence is more likely to curb fake news and make life on social media easier, or whether AI has the opposite effect and only makes the situation worse. In addition, the results of this thesis will be used to find out in which areas there is still room for further research.

Chapter 1: Social Media

1. Social Media

Social media is the tool that rules the day. You would think that everyone would at least know what social media is and how or what to use it for. 90% of online users have at least one account on at least one social media platform (Taprial/Kanwar, 2012). But what exactly is social media? Social media is a tool with which people can network online and interact with each other. The best-known platforms are X, formerly Twitter, Instagram and YouTube (Social Media Overview – Communications, 2024)

1.1. Introduction to social media

Social media has dramatically changed the lives of everyone who has access to it over the last 15 years and will continue to do so (Bik/Goldstein, 2013). In a matter of seconds, it is now possible to check messages, contact friends and family within a second who live far away and even place food orders product orders from without having to leave the house (Miller, et. Al. 2016). What used to be unthinkable is now made possible by platforms such as Instagram, Facebook, YouTube and many more (Bik/Goldstein, 2013). Bik & Goldstein (2013) define it as any web presence where users can write their own thought up in form of a blog or share their lives in forms of posts.

1.1.1. Definition and overview of social media

Social media is a group of internet-based applications that are based on Web 2.0, which does not only let the user be the consumer of the content but also allows the user to become the creator himself. These various platforms, which were mentioned above, are used to create content that enables different users to exchange their thoughts with each other and where brands or private people can create user-generated content (Gundecha & Liu, 2012). There are different types of social media, such as the traditional media which are radio, television and newspapers, or the more modern platforms that fall under the term 'Social Media' like Facebook, Twitter, Instagram, and other online platforms. Social media has not only changed topics such as holiday planning, music tastes and research for school projects, but also how people can interact with each other daily in a digital way (Dewing, 2010). It helps people to

share photos and videos, send messages to each other and express their thoughts and opinions in blog posts, allowing them to network further with others. This is not only something that only private individuals can do. With the help of social media even companies gain a better access to their customers (Dewing, 2010). These relatively new platforms offer people the opportunity to be in constant dialogue with each other at a speed that no medium has yet managed to surpass (Gundecha & Liu, 2012). Over the last ten years, social media has become an integral part of society, and it is now hard to imagine life without it (José, 2013). The various platforms are so firmly anchored in both private and professional life that some businesses don't make the effort of creating a website. They sell their products on these platforms. Social media also help people to keep in touch with others from all over the world, be it family members who live abroad or in another city, or people who have met through these platforms (José, 2013). The fact that these social media platforms have become so ubiquitous is shown by the fact that there are now many different tools on Facebook, Twitter, Instagram, and the like and comment button that are designed to enable users to enjoy the platforms even more and express their feeling toward a post (Murthy, 2018). However, social media is not only about being used for private or professional purposes, as mentioned above, but political situations and discussions are also increasingly becoming a topic on the platforms. Social media has already changed people's lives and will continue to do so. These platforms are more popular than ever, especially in the advertising industry. Both for companies and private individuals. Nowadays, people don't even need an education to make money from advertising on social media, they can learn it themselves with the help of YouTube and just start (Murthy, 2018). This rapid development of it is therefore forcing almost everyone to get used to this new way of exchanging, communicating and to adapt (José, 2013).



Figure 1: How to choose the best social media platform for your business. Source: Greene (2019)

1.1.2. Historical development and evolution of social media platforms

In recent years, social media has become increasingly popular and successful, helping people to share information and their opinions with others. The popularity of these platforms is growing, and more and more new ones are being added that can displace older platforms (Kong, et. Al. 2020). But how did it all start?

The innovation of the World Wide Web in 1991, when Tim Beyers-Lee managed to combine hypertext technology with the Internet, laid the foundation for social media. Initially, social media was used to allow people to join a channel or form small groups with each other to help them interact with like-minded people (Van Djick, 2013). In the last two decades, however, social media has become more and more integrated into people's everyday lives and has become indispensable. Since the late 1990s, more and more social media websites have been launched. The first one called "SixDegrees.com" appeared in 1997. There were already several different platforms and websites, many of which did not manage to establish themselves (Dewing, 2010). It all began in 1994 when Justin Hall started personal blogging in text form and realized that other internet users were interested in what he was writing (Sajithra/Patil, 2013). After a slow start, blogging grew in popularity and became more widespread in 1999. In the years that followed, blog posts developed more into the direction of user-based content, which was more likely to be in forms of pictures and later videos (Sajithra/Patil, 2013). Large platforms such as Facebook and Twitter can keep themselves alive through constant new developments and even newer software tools and create a better user experience for the people using these platforms. In 2012, for example, Facebook reported that it had 1 billion users worldwide (Dewing, 2010). By the end of 2012, two-thirds of American adults were using any type of social network, rising further to 83% among 18–29-year-olds (McIntyre, 2014). These changes are so large that some historians consider this a break-through-moment in the history of mass communication. Even though these numbers look good for the platforms, there are always ups and downs when it comes to users. Many realize that social media is not always doing them any good and would agree to take longer-term breaks from them (McIntyre, 2014).

1.2. Impact of social media on Communication

The new way of communicating and interacting through social media is clearly noticeable. If one looks around in a restaurant these days, one sees people sitting at a table with other people, but they're still on their smartphones checking their social media or even posting something instead of talking to the other person. Experts claim their worries by saying that there will be no turning back when it comes to this issue (Subramanian, 2017).

1.2.1. Changes in communication patterns and styles

Communication from person to person and, above all, mass communication has changed frequently and rapidly in recent decades. The urge to share something with others, be it a birthday party, a great new opportunity that came up, or a wedding, all these events are shared and publicized on social media nowadays to show everyone else what a great life they have (Bala, 2014).

If one compares communication in general and especially interpersonal communication with the communication style from 20 years ago, it would have been unthinkable to exchange information without at least having to be in front of one another. People had to actively arrange meetings to keep up to date with new events happening in the lives of others (Ahmar, 2023). Nowadays, the various platforms offer the opportunity to build relationships with people who are in geographically different locations and have never seen each other in person, have grown up in different cultures, but still have the same interests. Ahmar (2023) claims that through social media came a revolutionary change that is mainly evident in the speed and accessibility of communication. Even if one looks at the rapid spread of news from all over the world, it is noticeable that people used to have to wait until the evening to watch the daily news on TV, whereas nowadays it appears on everyone's smartphone via direct message only minutes after the news event occurred. Even if the classic means of mass communication still exist and have been the leading ones on the market for several centuries, social networks have managed to overtake them in popularity, so that the former market leaders nowadays orientate themselves towards Instagram, Facebook, and Twitter to find out the latest trends and news (Bala, 2014). This ensures that these platforms are strong competitors for traditional media. All of this because they are so fast-moving and everyone who has access to the internet has the possibility to access them (Bala, 2014).

But it's not just communication between friends or internet friends that has changed or become possible thanks to social media. Communication from politicians to their Follower or from celebrities to their fans has also changed. Nowadays, they no longer just give interviews or television speeches, but address their target group on an almost daily basis, which makes communication more personal and therefore public people more approachable (Syarifuddin, et. Al. 2021). Today Celebrities film themselves in their kitchen, make room tours of their houses or show what they bought recently which makes them seem very approachable and real. The fact that Internet users can also comment, like or share the posts and story contributions means that they spread much faster, and users get the feeling that they are getting very close to these unapproachable people. Since politicians noticed a drastic change and a possible benefit in social media, they are increasingly using it as an election campaign tool rather than traditional media, since it can help them gain popularity and especially trust which is much needed for elections and being able to beat the opponent (Syarifuddin, et. Al. 2021).

1.2.2. The role of social media in interpersonal relationships

The generations growing up in the 21st century are experiencing a major shift in the way people communicate with each other (Moorman/Bowker, 2011). Nowadays, as already mentioned above, it is no longer necessary that people must meet face to face to deliver a message or news, but it is enough for them to type it on their smartphone and transmit it to the desired addressee in a matter of seconds (Moorman/Bowker, 2011). The question that arises is whether this type of communication has positive or negative effects on people. Only recently, however, has this type of communication proven to be very helpful, especially during the Covid-19 pandemic (Syarifuddin, et. Al. 2021). When schools, restaurants, cultural institutions, and others were closed due to the lockdown and people were not allowed to leave their homes, they had almost no choice but to communicate with their friends and relatives via social media platforms to stay in touch and up to date on how their loved ones manage the situation of a worldwide pandemic. Platforms such as Zoom made it possible for students to attend classes from home without having to go outside and having to risk of infection, and the same was the case for employees (Syarifuddin, et. Al. 2021). Regular person-to-person interactions are important for well-being and the psyche

(Pennington, 2021). The lockdowns have forced citizens to stay at home, leaving them with only social media to keep in touch and maintain some form of interpersonal interaction (Pennington, 2021). Research has shown that Facebook has been the most used network during the pandemic and surveys on the platform has shown that people have found the network to have a particularly positive impact on their well-being during that time (Salman, 2021). Researchers even found out that writing a message on Facebook was enough to make people feel less lonely (Pennington, 2021). It was also found that people who felt very lonely were able to improve their mental health because they got the feeling of being connected to other people through social media (Pennington, 2021).

Although social media had many advantages in times like the Covid-19 pandemic, as already mentioned communication with relatives or friends, these platforms, as Salman (2021) interestingly claims, have brought negative aspects at the same time. Since everyone was at home and constantly scrolling through social media, everyone was getting new insights about the pandemic in a short amount of time, but also some insights that were not true or were fake. This fake news spread quickly as users sent it back and forth, sometimes causing unnecessary panic (Salman, 2021). People also didn't know whether they could still trust the local news channels which made radical groups like the Corona Deniers form quickly (Jaspal/Nerlich).

1.3. Social Media and Information Dissemination

Nowadays, not only private internet users use social media to communicate with their friends, but also companies and even the government and news services use it to spread news around the world as quickly as possible. Tools such as blogs, wikis and others make this happen. What used to be broadcasted on TV only a few times a day is now spread in seconds via social media (Osatuyi, 2013).

1.3.1. How social media has transformed information sharing

The dissemination of information has never been more important in the 21st century than when the corona pandemic broke out. People constantly wanted to know what the new infection numbers were, what the new regulations were or what the general situation looks

like or would be. Health centers would therefore provide the media with information as quickly as possible (Kudchadkar/Carroll, 2020).

The fact that news can spread quickly on social media is particularly helpful for the population in crisis situations such as the coronavirus pandemic that arose in 2020. However, Zhu, et. Al. (2018) say that this can also lead to misunderstandings or false information, which can cause a stir. As anyone with internet access can access social media, anyone can spread anything on these platforms, i.e., people without specialist knowledge on certain topics can share their information there, even if it is incorrect or has been falsified (Zhu, et. Al. 2018).

As social media is a real-time and instant communication platform, it was used at the time especially for creating information and networking between organizations and patients. This is because social media worked with an algorithm that recognizes what interests its internet users have and uses hashtags that show users similar content matching those hashtags (Kudchadkar/Carroll, 2020). However, the rapid dissemination of information does not only take place from government to user, but also among friends, which strengthens the credibility of their information, as people believe their friends even more than the news services. For this reason, groups such as the lateral thinkers formed during the coronavirus pandemic, as they allowed themselves to be influenced by their friends' news (Zhu, et. Al. 2018). The fact that users nowadays also have several platforms to choose from, which they can access and get information from, allows them to build up different groups or communities that deal with the same views and opinions of theirs, which can lead to users being fed more false information and thus the dissemination of information can happen across platforms (Zhu, et. Al. 2018).

Sharing news on social media not only has an influence on the spread of news in society but also influences the reward centers in people's minds (Lee, et. Al. 2011). This is also one of the reasons why people like to share their lives or achievements on social networks, as it gives them the feeling that they are doing something right when people like it or comment. This does not only work with political news or global news, but also private news that they share with their friends (Lee, et. Al. 2011).

1.3.2. Influence of social media on news distribution and consumption

Social media has become indispensable when it comes to the dissemination of news. Users' social media feeds always display posts from their friends, but also from news services (Bergström/Belfrage, 2018). In 2007, a study found that social media had little economic impact on the news industry, but this study was conducted before the collapse of newspaper advertising revenue. However, a study conducted in 2016 showed that the power of social media is much stronger than assumed. Smaller newspapers such as local newspapers are particularly affected and could be displaced by the new media platforms (Zhao, 2019). Many people are not entirely comfortable with the topic of advertising in news services, which is why they are losing credibility compared to social media. As the news market on social media continues to grow, so does the interest in understanding how the consumption of news content is changing (Zhao, 2019).

The internet has managed to greatly reduce the cost of distributing news and advertising. For example, online magazines and news services, or the ability to read news on a digital device instead of physically, means that the printing of newspapers and magazines can be reduced, if not eliminated, which is saving publishers money (Flaxman, et. Al. 2013). But social media does not only simplify the process of sharing for publishers. Private individuals can also use various platforms to share stories in the form of text, video, or photos with their friends. In addition, social media, which works with an algorithm, changes what is displayed to internet users and can therefore offer personalized content (Flaxman, et. Al. 2013). The content that users receive depends entirely on how they behave on those platforms, but also on how their friends behave and what interests they show (Bergström/Belfrage, 2018). Nowadays it is known that social media is responsible for discussions of all kinds, such as political issues, climate change or pandemics and so many other topics (Etta, et. Al., 2022). This can initially be seen as a good thing. However, scientists' express concerns that it is difficult to use such algorithms for political content (Flaxman, et. Al. 2013). With such an algorithm, users are only being shown content that relates to content that they have previously been interacting with, which can lead to them being denied other political views and perspectives and thus start to think in a one-sided way (Flaxman, et. Al. 2013). News, such as the Covid-19 pandemic, has spread rapidly through social media and there have been many rumors that have been received as news. This phenomenon is known as infodemic,

and it has been proven that it can have a negative impact on people's behavior (Etta et. Al., 2022).

1.4. User Behavior and Psychology on social media

Social media helps young adults to meet their basic needs. Recognition in the form of likes as soon as a post is published or being in constant contact with others. People need human contact, regardless of whether this takes place online or offline (Bölen, 2022).

1.4.1. User motivations for using social media

Most people use all forms of social platforms on an almost daily basis. Access to the internet is no longer only possible via PC as it used to be, but also via mobile devices, which means that users have constant access to social networks. Users in the United States now spend more than a quarter of their online time on these platforms (Hoffman/Novak, 2012). One of the reasons for this is that nowadays it has become easy to communicate with friends and family by simply creating a post on Instagram, Facebook or other platforms showing what is happening in the person's life (Luo/Hancock, 2020). This willingness to share has almost become a trend, happens daily and it is hard to imagine life without it. This means they can always be in touch with their loved ones (Luo/Hancock, 2020). This is a form of self-presentation that would not even be possible in real life (Hoffman/Novak, 2012). However, the question arises as to why users disclose so much information without hesitation, even though it is known that the internet brings various dangers, especially when it comes to data (Luo/Hancock, 2020). The need to communicate relates to the so-called "uses and gratification" phenomenon (Al-Menayes, 2015). This refers to the fact that certain behaviors by people leads to a need or desire of being satisfied. Real-time contact via social media is also part of this phenomenon and thus satisfies people's need to have interpersonal contact, even if it is only online. Furthermore, social media satisfies needs such as affection, control, relaxation, inclusion, and others (Al-Menayes, 2015). Social media enables people to interact with each other, regardless of their time zone or continent. With this, users often interact with people they have never met, nor will they ever meet them, because their friendship will only be online (Hoffman/Novak, 2012). Numerous positive tools are offered on the

platforms to make people's lives easier, such as the ability to access different literature or it makes it easy for people to educate themselves independently through videos and blog posts (Hoffman/Novak, 2012). Whiting and Williams (2013) found that people like to use social media because it makes it easier to find people with the same interests and to network and connect with them. There is also the point of always being informed about the latest news (Whiting/Williams, 2013). As mentioned above, social media also enables users to educate themselves and access a lot of information throughout the many platforms. For many people, however, it is also a place to relax and unwind from everyday life through entertainment. Using the platforms is therefore enjoyable, fun, and entertaining. But social media is also very useful for internet users, as they offer many convenient ways to make people's everyday lives easier. One example of this is online shopping, which is a convenient way of shopping since the consumer only needs to make a purchase via their smart device and get the product sent home (Whiting/Williams, 2013).

1.4.2. Psychological effects of social media use

As described in the chapter above, social media offers a range of positive aspects, which is why users like to access these platforms. Ostic et al. (2021) found that the use of Facebook or WhatsApp can have a positive effect on users and their psychological well-being, as these platforms make it easy for people to stay in contact with others and thus release a sense of belonging in people. This is mainly because social media offers the opportunity to network with people who share the same interests. This has a positive effect on interactions such as acceptance or trust (Ostic, et. Al., 2021). Research relating to social relationships in social media has shown that seniors who spend a lot of time on social networks and thus network with like-minded people show that their cognitive and motor functions improve positively in old age instead of deteriorating (Hoffman/Novak, 2012). Interestingly, social contacts also influence pain sensitivity. The study found that seniors with a larger network are not as sensitive to pain as those with a small network. This means that it does not matter whether it is a high-quality network or one that is maintained offline, it still has a positive effect on people's mental and physical well-being (Hoffman/Novak, 2012).

However, there are also negative effects of social media on people. Especially when it comes to excessive and problematic use of it (Schivinski, et. Al., 2020). More and more people,

especially young people, are using social media. In the USA, around 90% of young people spend time on social media every day. Today, it is not yet entirely clear whether these platforms are either the cause of mental illness or can help prevent such illnesses (Primack, et. Al., 2017). Research on precisely this topic has shown that people who consume social media in excess tend to notice a deterioration in their psychological well-being and negative changes to their general health (Schivinski, et. Al., 2020). One example of the negative effects of social media is the constant social comparison with others, which is often an upward comparison. Comparison is a process in which people seek to improve their self-image by observing other people and their lives (Lee, 2020). For example, the number of likes under a post on Facebook or Instagram puts people under a lot of pressure as they see that other users get more likes on their posts, which makes them more popular and can cause envy. These comparisons make self-presentation more and more extreme, as users hope to create a positive effect on their followers and thus come across better (Lee, 2020). Other worrying findings relating to mental health problems also compel us to investigate the topic further. People who upload more pictures of themselves for the very reason mentioned above may show signs of narcissistic traits (Keleş, et. Al. 2019). However, uploading lots of pictures of yourself is considered normal nowadays because many social network users reveal a lot about themselves. Constant comparison can also impair personal and social development. Studies investigating on the topic of depression with young adults have also shown that there might be a small correlation with social media. Another phenomenon caused by social media are sleep disorders, as social media consumption usually takes place while sitting, which leads to less physical activity. Because the various channels offer so much different content, users often feel overwhelmed (Keleş, et. Al. 2019). A survey by Bevan, Gomez, and Sparks in 2014 found that users between the ages of 18-70 who spend a lot of time on social media are more prone to stress and mental illness than people who do not spend as much time on these platforms (Sheldon, et. Al. 2019). The survey also revealed that people who share a lot, especially negative events, suffer from extreme stress and lose quality of life. Since users can connect with each other in large numbers and thus build up a large network, many of them stated that they are overwhelmed by the many messages they receive each day and the constant exchange even with strangers and often respond to messages with a very long delay (Sheldon/Rauschnabel/Honeycutt, 2019). Depression can also be caused using social media, not necessarily by the amount of time people spend on

social platforms but how they use it. By scrolling through other profiles, seeing their vacations, or seeing them being invited to an event where oneself was not invited to can lead to loneliness and envy (Aalbers, et. Al. 2019). In addition, social media are places where people often present a rather unrealistic portrayal of their everyday life (Schmidt, 2022). For many users, however, it is difficult to distinguish between what is unrealistic and what is not. Constantly observing the lives of others can also lead to increased personal expectations of themselves, resulting in people feeling that they have failed because others have more or do more in their lives (Schmidt, 2022). The same also applies to body ideals. Especially many women present themselves while exercising or their results after going to the gym or what they have eaten. This can create a distorted image of the "perfect" body ideals, especially in young women, and in the worst case can lead to an eating disorder (Schmidt, 2022).

1.5. Social Media and Society

Social media is playing an increasingly important role in people's social lives. Nowadays, it is possible for anyone with access to the internet to educate themselves in any subject area and almost always for free. Through videos on YouTube, students can simply open their laptops at home and have subjects such as math or biology explained by Youtubers within a few minutes instead of taking expensive private tutoring lessons. Furthermore, students can nowadays study or do group works with fellow students without having to leave their homes thanks to the video call platforms like Zoom or Teams (Amedie, 2015).

1.5.1. Influence of social media on politics, activism, and public discourse

Due to the accessibility for everyone, people with no good intentions can make themselves heard on social media, because these platforms can offer bad people opportunities that would be more difficult, unpleasant, or hardly possible offline. Above all, users can hide behind their screen and protect their identity with a fake account (Amedie, 2015). This is precisely why cyber bullying is so dangerous, but at the same time an easy thing for those who want to bully others. Thanks to platforms such as Facebook, Instagram, Twitter and others, bullies no longer even must look their victim in the face. They can expose the person to everyone

else without anyone knowing who is behind the account. This puts much more pressure on young people, as they are afraid that someone will be willing to expose them again. It can also lead to a distorted image and make something like bullying more acceptable as it seems more normal (Craig et. Al., 2020). The National Children's home study in Britain found that one in four children today report being bullied online (Amedie, 2015). This is because the accounts that bully often also receive praise or likes for their posts and hate comments from others who join the bully (Craig et. Al., 2020). This effect of bullying becoming normalized for adolescents is called the "disinhibition effect" (Craig et. Al., 2020). Another negative influence of social media on our society is that it can become an addiction. As mentioned above, social media can trigger a sense of belonging in people and stimulate the reward center in the brain. As a result, people may spend a lot of time on social media and neglect their work or life offline (Akram/Kumar, 2017). Fraudsters also like to hang around on social platforms, as they have an easy game due to the large amount of advertising and can sell or trick users into buying something that doesn't exist. Another danger and a highly controversial issue is data protection. It is well known that every user's data is collected on the internet and personalized user profiles are created to display advertisements or posts that the user might like based on their previous behavior. However, there are often data leaks that could expose private data such as addresses (Akram/Kumar, 2017). In addition to that social media is also helpful for terrorists. These groups have quickly spread on these platforms because it helps them to recruit new members, spread their propaganda or fundraise so that they have money for new operations (Amedie, 2015). Professor Weimann from Haifa University in Israel has been studying the connection between terrorists and social media for some time. In 2013, a terrorist group hacked a news account on Twitter and spread the news that a bomb had exploded in the White House and that Obama had been injured. This news caused the stock market to plummet. Even though this news were not true, this news had a major effect on the economy (Amedie, 2015).

But they can also have a positive influence on our society. Through various articles or YouTube videos in which something is explained, students can now educate themselves outside of school without the need for a teacher. The speed at which news spreads on the internet also ensures that our society is always informed and knows everything that is happening around the world (Akram/Kumar, 2017).



Figure 2: AP Twitter hacked; no attack at White House, (Today, 2013)

1.5.2. Ethical and privacy concerns related to social media

Nowadays, people are increasingly willing to disclose their data on the internet to make better use of online services (Jozani, et. Al. 2020). Since Internet use is no longer limited to a computer but can be done at any time via a smartphone, comfort with online services is in high demand and the willingness to give away data to these online services by the user is much higher. By accepting cookies, which is done with just one click, websites collect data about the user to display personalized content such as advertising (Lavin, 2006). A cookie is a small file that allows the website to analyze and store the user behavior of the person that is looking at the website (Lavin, 2006). However, the disclosure of data can lead to an increased risk of data breaches (Jozani, et. Al. 2020). This process gives both individual and private users cause for concern when it comes to data protection (Ullah, et. Al. 2020). However, it is not only cookies that can analyze behavior; portable, internet-enabled devices, such as smartphones, also have built-in sensors that store user behavior and can pass it on to companies which they can use for their advertising. For example, the country in which the user lives can be quickly determined and the user's approximate income bracket can be determined based on the neighborhood the user lives in (Jozani, et. Al. 2020). With more awareness of social media, people are increasingly concerned about where their data ends up and where it is processed. There is also often talk in the media about data leaks that lead to private data, such as home addresses or recent purchases, being published by mistake (Gruzd, et. Al. 2018).

1.6. Social Media Marketing and Business

The new way of communicating through social media has not stopped at companies. With the shift from traditional media to social media, companies also need to rethink how they communicate with their customers when it comes to marketing. This is because it is now possible for companies to build a huge community via these media platforms, both globally and internationally, which was virtually impossible before. Just as social media can be seen as a blessing for companies, these platforms can quickly become a curse and destroy a company's reputation within seconds or within few days. Because customers or users of social media always can express their opinion under a post as a comment, they can quickly bring down a company if they do not like the way it communicates or interacts with customers, or they simply don't like the product (Saravanakumar/Sugantha/Lakshmi, 2012).

1.6.1. How businesses leverage social media for marketing and advertising

The development of the internet has made great leaps in recent years. In addition to the conventional search function, which enables internet users to obtain any information, social media can be used to convey people's feelings and thoughts. In less than five years, social media has established itself as a medium that is here to stay (Tsimonis/Dimitriadis, 2014). Not only have people's dynamics changed due to social media, but also those of companies, as they have realized that they can achieve their goals with the help of these platforms. Topics such as targeted advertising are particularly helpful for companies as they can better avoid wastage. Companies that focus on advertising and data analytics make a profit by collecting data and reselling it (Ullah, et. Al. 2020). Even if many companies are not yet too comfortable with social media and the new way of producing and publishing advertising, it is essential for companies to have a presence on the platforms (Tsimonis/Dimitriadis, 2014). Social media also offer companies the opportunity to show a certain personality through a constant exchange with customers, whether it is a comment under a picture or a private message. It is therefore no longer just the name of the company, but also a certain character that the company shows on social media (Tsimonis/Dimitriadis, 2014). Television, radio, and others are almost ignored by many companies these days, as they have an existing customer base on social media that they know is interested in the company and its products

(Singh/Singh, 2018). In addition, these platforms not only offer companies the opportunity to advertise more cost-effectively, but also bring faces and thus a certain closeness to the customer (Singh/Singh, 2018). Even small companies are increasingly turning to social media. Many no longer even create their own proper website, as they advertise the link to the products via their social channels. In 2012, around 73% of small businesses also used social media to promote their products and many companies were in the process of implementing this (Jones, et. Al. 2015). Precisely because it is becoming more and more normal and easier for customers to find out about a product on social media using images and video material, it is becoming increasingly important for companies to position themselves here and have a good marketing strategy. A Companies' work with influencers is also very helpful, as social media users trust influencers and their opinions and are often more likely to buy a product through the influencer than through the company itself (Jones, et. Al. 2015).

1.6.2. The impact of social media on e-commerce and customer engagement

Nowadays, the size of a company no longer makes a difference, as most companies are now on social media and even could grow as well (Singh/Singh, 2018). Especially when it comes to advertising products, social media is particularly suitable due to its algorithm and because companies can reach exactly the audience they want to reach. What makes these platforms particularly interesting for companies is that companies can engage in direct dialogue with their customers and thus receive direct feedback (Singh/Singh, 2018). The online world provides the infrastructure that companies need to carry out online communication (Valerio, et. Al. 2019). These platforms help companies to increase the popularity of their online shops, raise brand awareness among customers and ensure the support of online customers. Today, it is no longer just important to sell, but to create a certain level of trust with customers so that they keep coming back to the brand. Platforms such as Instagram and others can help with this, as the brand is no longer just a logo, but also a voice or a face (Valerio, et. Al. 2019). The description by Farook and Abeysekara (2016) is interesting. They say that social media has changed the power relationship between brands and customers. In the past, it was the company that attracted the customer through advertising and sold products. Today it is still the same, but the customers have many more opportunities to give their opinion or feedback on the product or the entire handling process, which can be

good for companies on the one hand, as they can always improve, but can also be bad if too many customers complain publicly. Through extreme customer loyalty, companies can succeed in ensuring that customers will buy again and again (Farook/Abeysekara, 2016).

1.7. Future Trends in Social Media

It used to be the case that new technological trends first took place in the business sector and were only then adopted in the private sector. An example for that is the computer. It was initially only available in the workplace until it made its way into people's homes. Things are different today. Social media, for example, was first established for private use until companies finally realized that they also needed to be active on social media for advertising reasons (Urbach/Ahlemann, 2016). Artificial intelligence has also been on everyone's lips for several years and will bring about enormous change when it comes to society, science, politics, and technology (Paaß/Hecker, 2020).

1.7.1. Emerging trends in social media development

Social media is here to stay and will continue to evolve for a long time to come. It doesn't matter whether it's for private individuals or companies, everyone needs it. Companies will increasingly ensure that they can guarantee good communication between customers and the brand, because social media is essential for companies to establish trust, which is particularly important to customers nowadays and will increase the chance of people buying products. In addition, more and more money is being invested in platforms such as Instagram and Facebook as social media is fast-moving and a strategy often needs to be changed or customers have new demands on a brand more quickly (Linke/Zerfass, 2012).

Chapter 2: Fake News

2. Disinformation and Fake News

The term fake news has been used more and more frequently in recent years and describes false news or information that circulates and is published on the Internet in particular (Research Subject Guides: Fake News/Misinformation/Disinformation: What Is Fake News? n.d.).

2.1. Introduction to Disinformation and Fake News

The term fake news was named word of the year in 2016, but there is still no exact definition for it. On one hand the dictionary defines it as the creation of misleading content where the consumer can decide for themselves what they want to believe or not. On the other hand, news services or magazines that actively publish news on social media or mainstream platforms that are not official news services but are not intended to be satire are described as fake news (Machete/Turpin, 2020).

2.1.1 Definitions and distinctions between disinformation, misinformation, and fake news

Only 15 years ago, the term "fake news" was mainly used for satirical purposes. The term was often used in magazines such as the German "Postillion" or in satire shows such as the "Daily Show" in the USA (Zimmermann/Kohring, 2018). At the latest since the unexpected election of Donald Trump as US President, the term "fake news" has become known around the globe. This term describes the distribution of information, especially via social media, where it is not clear whether this information is correct, was checked before publication or is even false (Geifert, 2018). The issue of fake news has already gone so far that it can influence elections, as any false news about a candidate is spread and voters could be lost as a result. This spread of fake news ensures that the free formation of opinion of the individual and the knowledge that people acquire on topics can be wrong and thus their freedom of opinion is taken away. Furthermore, the definition of "fake news" is that it is knowingly misrepresented to cause confusion or to lure people onto a certain track in terms of their thoughts and opinions (Zimmermann/Kohring, 2018). Disinformation describes information

campaigns designed to enable states to intervene in the political decisions of other states (Bildung, 2023). For example, in the US election campaign between Donald Trump and Hillary Clinton, it was proven that Russia ensured that favorable information and news about Trump appeared in the media and harmful information about Clinton. Disinformation is therefore false information that is spread with the intention of persuasion. There is another term called Misinformation which is basically the same as disinformation, except that this information was not spread with a certain intention (Bildung, 2023).

Politics as well as the economy and society suffer from this false information (Vosoughi/Roy, 2018). Not only important elections are influenced by it, but also the stock market. Rumors spread by the media caused stocks worth 130 billion US dollars to fall when it was reported that Barack Obama had been injured in an explosion in the White House. The problem here is that nowadays anyone with access to social media can spread such false information within seconds. Artificial intelligence, which is now so well developed that it can make images and videos look deceptively real, is also responsible for the fact that internet users can hardly distinguish between true and false (Vosoughi/Roy, 2018).



Figure 3: Elections 2016: Trumps fere media helped keep cost down, but fewer doners provided more of the cash, (Sultan, 2017)

2.1.2. The significance of studying disinformation

Since there is a vast amount of information on the Internet that can be uploaded by any Internet user, it is difficult or even impossible for many to find the most relevant, accessible,

and high-quality information of them all (Hernon, 1995). Another problem here is that not all information available on the Internet is automatically true information. Internet users who obtain their information here would therefore always have to check whether the sources from which they obtain information are true and not lies or propaganda. As this is an effort for everyone, very few people do it and run the risk of being misinformed about events (Hernon, 1995). Young adults, who are constantly on social media or on the internet in general and are in contact with other users daily who often do not even know each other personally, can quickly become victims of false information, as they are less scrutinizing and naive at a young age (de Vicente Domínguez, et. Al., 2021). Interestingly, Spanish, and Portuguese students are the ones who are very aware of fake news but are most likely to fall for it. De Vicente Domínguez, et. Al. (2021) conducted a study to investigate how young people deal with fake news. It depended on the gender and on the course of study of the students. The study also found that it is very important to young adults between the ages of 18 and 25 with a higher level of education that fake news can be better recognized. As many as 76.5% of respondents have already taken part in a course to help educate them on this topic. Many of those who knew about the course but had not taken part stated that the reasons for this were lack of interest, lack of time or the belief that they knew how to recognize fake news. In addition, the respondents believe that the main reasons for creating fake news are more clicks and followers or to manipulate society (de Vicente Domínguez, et. Al. 2021).

When recognizing fake news, it is important to detach oneself from one's own ideologies and to be open to opinions and different topics (Batailler, et. Al. 2022). People tend to accept information more readily if it agrees with their ideologies. Someone who is a supporter of a political party and therefore hears fake news about a politician they like and who represents their ideologies is very likely to accept the fake news without reconsidering it and proofreading it (Batailler, et. Al. 2022).

2.2. Detection of Disinformation

Fake news is becoming more and more widespread. This is because anyone with access to social media platforms can act and spread fake news (Ahmed, et. Al. 2021). In the past, the perpetrators of fake news wanted to get people to click on their links and land on their websites to get attention and, above all, they did this for monetary purposes, as they could

earn money by people who click on the links. Another reason for fake news in the past was that hackers were behind such links and could find out information about companies by clicking on those certain links, as they could attach malicious codes to them and launch a cyberattack. This is precisely why researchers are working on developing tools to help recognize them more quickly (Ahmed, et. Al. 2021). They often use machine learning for this purpose, as it is hardly possible to recognize fake news with the eye. As the number of these false news items continues to grow, they have started to train the machine learning algorithm to recognize fake news and everything that makes them up. Studies have found that people can hardly recognize lies in texts, which is why it is even more essential to rely on machine learning (Burkhardt, 2017). Since fake news are mostly spread via social media and by fake accounts called "bots", it is difficult for users to recognize what is real and what is false. The so called "bots" are accounts that pretend to be a real person but are not actually real and provide other internet users with false information (Burkhardt, 2017). The probability that a normal internet user will recognize the information from the bot and identify it as false is around 50 percent (Burkhardt, 2017).

2.2.1. The danger of disinformation

Nowadays, disinformation is so widespread and harmful to people that the Secretary-General of the World Health Organization (WHO) has declared a so-called "infodemic" (Simion, 2023). An infodemic is false information about current events spread by private actors. For example, when the corona pandemic broke out in 2020, there were many different theories and information about how it broke out, how to behave and what the future will look like (Blasig, 2020). This disinformation is mostly spread online and on social media and is mostly about political situations, conspiracy theories and propaganda (Simion, 2023). Another problem is that people are so often confronted with disinformation that they hardly know what exactly to believe in anymore. Even if they have true information in front of them and it can even be verified, there is still a high level of distrust towards information from the internet. In 2018, the Edelman Trust Barometer found that trust in online media among the British population is less than 25% and trust in the government is as low as 36% (Simion, 2023). The danger of disinformation is that it shapes the opinions of internet users. While this disinformation is increasingly present and widespread through social media, there have

been many events in the past that have influenced people's opinions and actions (Michaelis, et. Al. 2022). The Jews were portrayed as lesser people in Germany during World War II to justify the genocide. The city of Pripyat, where the biggest nuclear catastrophe happened, was not evacuated at first because the people received false information from the government to cover up this catastrophe (Michaelis, et. Al. 2022). Disinformation can come with such force that it can even help to influence election results. In a closer look at the 2016 US election, government investigations have revealed that the Russian government ran a campaign of disinformation in the United States right before the election of President Trump (McKay/Tenove, 2021). While it cannot be proven exactly whether these campaigns ultimately helped Donald Trump become president, the Russian government made increased use of social media for its campaign. The strongest social media activity was carried out by the St. Petersburg-based Internet Research Agency (IRA). They used fake usernames to spread their messages to citizens and were able to reach at least 126 million people. Situations like this show firstly that disinformation can be well hidden by the government by using fake usernames and thus not coming across as the government but as normal citizens passing on news, secondly it shows how much impact such false information can have on important and historical events like the elections and thirdly it shows how difficult it is for citizens to distinguish this false information from true news and also how few citizens bother to scrutinize such news (McKay/Tenove, 2021).

2.2.2. Historical examples of disinformation campaigns

Social media has a completely new structure when it comes to content distribution. Where journalists used to write and publish articles that were proofread beforehand, content can now be published by any user without filtering or fact-checking. Depending on the algorithm and how people react to the post, normal internet users can even build up a reach as large as that of big news channels like Fox News and CNN (Allcott/Gentzkow, 2017).

Fake news existed before, but at the latest since the US elections in 2016, it has been on everyone's lips (Bovet/Makse, 2019). Especially on the social media platform Twitter, which is now called X, confusing posts that shared false information were repeatedly published. As described in the previous chapter, this disinformation not only originated from private internet users, but was also systematically published on social media by the Russian

government. They not only used fake accounts to spread the news, but also bots, which are automated fake accounts that can spread the news in a matter of seconds (Bovet/Makse, 2019). The problem here is that research has found that 62% of US adults get their news from social media. Furthermore, fake news were often shared more than actual news because people either believed it or simply found it funny, which was also why Donald Trump brought more to the forefront in terms of his media presence during the election (Allcott/Gentzkow, 2017). One study found that around 27.4% of Americans eligible to vote had read an article that was pro-Trump or pro-Clinton, but which were fake (Guess/Nyhan/Reifler, 2018). Most of these articles, whether pro-Trump or pro-Clinton, were heavily biased towards Donald Trump. It is also interesting to note that older Americans were more likely to view fake news, which may be because older people are not as familiar with social media as younger people. But there is hope that internet users are now taking a closer look and scrutinizing news. Nevertheless, there is still room for improvement. One in four US citizens now carry out a fact check before believing information from the internet. However, only 62% state that they know something about fact checks (Guess/Nyhan/Reifler, 2018).

2.3. The Role of social media in Disinformation

Fake news has a huge impact on our society today. In 2013, for example, the Twitter account of a major news agency in the USA was hacked and the news was spread that there had been several explosions in the White House and that Obama was reportedly injured (BBC News, 2013). Within hours of this tweet, several stocks plummeted, and people were shocked. Shortly after the fake Twitter post, the White House press secretary addressed the media and said that this news was not true and that Obama is alive and healthy, causing the stock markets to quickly recover (BBC News, 2013). Social media platforms are very fast-moving. When the coronavirus pandemic hit Europe in 2020 and everyone had to go into lockdown, there was a 20-76% increase in the use of social media. This increased the spread of fake news even more. At the time, more than 46,000 new Twitter posts were associated with fake news (Naeem/Bhatti/Khan, 2021).

2.3.1. How social media platforms could contribute to the spread of disinformation

Almost everyone remembers the day of the election in 2016. People woke up, looked at their smartphones and saw that Donald Trump had been elected as the new President of the United States. The reaction of many people was that they wanted to share the news with their friends. But why is that? The reason for this is that social media has a psychological effect on people and how they think.

Social media platforms are there so that people can share all kinds of things, be it pictures, videos, or blogs with their network (Choi/Toma, 2014). This sharing evokes a basic need in humans, because it helps people to relate to each other (Choi/Toma, 2014). So as soon as something happens in the world, people quickly feel the urge to share and communicate this news with friends and family who live all over the world and thus satisfy their need (Choi/Toma, 2014). Other positive effects of social media on people are that they can expand their knowledge (Akram/Kumar, 2017). It is very easy to learn new things by simply turning on their smartphones (Akram/Kumar, 2017). Additionally, information retrieval is made much easier through social media. Unlike standard media such as television or radio, viewers and listeners only receive new information on certain topics hourly or perhaps every half hour. On social media, it's different. Here, news are being spread every minute and they are always available to the users with access to social media platforms (Akram/Kumar, 2017). Social media help in spreading fake news, as theoretically, anyone can create an account and post whatever they want, if it does not violate the guidelines of the respective social network (Taprial/Kanwar, 2012). Furthermore, these platforms offer the possibility for posts to go viral within hours, reaching an enormous number of people worldwide (Al-Rawi, 2019). Many people intentionally spread false news to manipulate the opinions of others or at least to make them reconsider their own opinion and their beliefs (Bovet/Makse, 2019). While this information may sound distressing, there are now studies indicating that Americans and Europeans are considering whether to post something they are not sure about because if fake news comes out, it can damage their reputation (Altay, et. Al. 2022). 89% of surveyed Americans stated that it is important to them that media organizations often fact-check and correct misinformation if they have made mistakes. 63% of surveyed Americans also indicated that they avoid media organizations that have previously spread fake news, and 50% of Americans have already avoided people in their social circle whom they believed were spreading fake news (Altay, et. Al. 2022).

2.3.2. The impact of algorithms, echo chambers, and filter bubbles

Relevance is one of the basic concepts of information science in social media. The user expects the algorithm to have accumulated relevant knowledge about them and then show them suitable results on search engines or social media platforms (Zimmer, et. Al. 2019). For several years, there has been the so-called algorithm on social media. This algorithm works as follows: If a user looks at another user's post for a longer time which can be around five seconds or longer, interacts with it in the form of comments, a like or by sharing the content with someone else, this post is given a higher weighting for the algorithm in connection with the user. However, this algorithm is constantly being adapted and changed (Zimmer, et. Al. 2019). It collects data from users, such as which posts they have liked, how long the user watches a video, or which links they have clicked on. Thus, user behavior is analyzed to display user-generated content to the user that is expected to appeal to them and be liked by them (Al-Garadi et. Al., 2019).

However, on social media, there are also the so-called "Echo Chambers." Here, the user is shown posts from people who think the same or similarly as themselves. This can lead to the user feeling reinforced in their opinion and not realizing that there are other ways of thinking. Another negative effect of this is that it can lead to the formation of groups that push each other further in a certain direction of thought, eventually adopting an extreme position (Cinelli et. Al., 2021).

Another phenomenon are the so-called "Filter Bubbles." These analyze the online behavior of people, i.e., they look at which websites the user has visited and with which content they have been interacting to create a profile of the user and categorize them into different groups (TED, 2011). As a result, these filter bubbles learn about the user's current interests. Eli Pariser says in his TED Talk that he wanted to test these filter bubbles and started liking and commenting on politically right-oriented posts and interacting with people who support these contents and opinions (TED, 2011). This led to him no longer seeing comments that contradicted the whole narrative. He started to only see comments that supported the opinions of the right-wing party. However, a friend of his, who is not politically right oriented, could see numerous comments that opposed the narrative and provided counterarguments. The problem here is that filter bubbles take away the user's opportunity

to be possibly encouraged to rethink and thus restrict their freedom of expression (TED, 2011).

2.4. Consequences and Impact of Disinformation

The rapid spread of fake news harbors many dangers. One of the biggest dangers of fake news is that people can increasingly lose faith in serious news (MARADUN, n.d.).

Precisely because it is so difficult to distinguish real information from false information, social media users do not even know what to believe (MARADUN, n.d.). This became particularly clear during the coronavirus pandemic, when rumors arose that the virus came from a bat or that it could be transmitted through mosquito bites. This news appeared credible and was then quickly spread with the help of social media and word of mouth and believed by many people (van Der Linde, et. Al. 2020).

2.4.1. Effects on public perception, decision-making, and society

The biggest problem of the impact of fake news on our society is that people not only lose trust in politics but also in their fellow human beings (ENDRE, n.d.). According to a survey by the Frankfurter Allgemeine Zeitung, every second German stated that they perceive fake news as a very serious threat (von Boehn, H. V, n.d.). There are various reasons why they see it this way. In the past, news were only broadcasted on a few television channels or briefly read on the radio. However, today there is an exorbitant pool of channels, like social media or many other television news stations, through which one can obtain news (von Boehn, H. V, n.d.). Another reason is that anyone can publish claims or express their own opinion on any topic without anyone fact-checking the information (ENDRE, n.d.). Thus, the user is no longer just a consumer but can also simultaneously become an author and a source of information. This already gives an idea of how quickly and how many rumors can arise as a result. Social bots also play a role in this. They adapt to the consumer's needs and often comment in masses on such posts, leading the algorithm of such websites to believe that this certain post is particularly popular among people and ensures that it is played more frequently (von Boehn, H. V, n.d.).

2.4.2. Real-world consequences and notable case studies

The coronavirus outbreak in Europe at the beginning of 2020 brought a significant shift to the work, education, and social life of almost everyone (Rocha, et. Al., 2021). The pandemic was suddenly the most important thing reported in the news, every channel broadcasted a large amount of information throughout the day, which brought even more confusion to people. To understand how to cope with this situation, people consumed more media (Bok, et. Al. 2021). In the United States, fake news began to spread faster and faster through social media, causing many people to cling to the idea that it was not a very serious disease (Bok, et. Al. 2021). Due to the rapidly increasing numbers of deaths caused by the virus, there was uncertainty about this disease, and people needed more information about it (Rocha, et. Al., 2021). Because of the lockdown, everyone spent extra time on social media to stay connected to their family and friends. Therefore, people were looking for answers, and others tried to provide answers, leading to the spread of many fake news and conspiracy theories about this virus (Rocha, et. Al., 2021). From biological weapons produced in China to coconut water that was supposed to defeat the virus, every kind of information was represented. Even President Trump suggested that injecting disinfectants could cleanse the body from the inside. All these examples are concerning, especially as fake news, like the suggestion by Donald Trump, poses an enormous threat to the health of the population. The situation of the abundance of various information was barely controllable, prompting the government to begin warning against fake news and promoting websites capable of detecting fake news (Balakrishnan, et. Al. 2022). As mentioned earlier, in such situations, especially through filter bubbles and algorithms, extreme opinions can form. That's why during the pandemic, the topic of the German word "Querdenker" (conspiracy theorists) became more prominent (Koos, 2021). More and more people had lost trust in the government, didn't believe what the government was saying and suggesting and took to the streets in large numbers to protest the COVID rules and restrictions (Koos, 2021). What made people particularly doubtful at the time was that many categorized Covid 19 as a normal flu. Although there were repeated reports of serious cases of illness or even deaths, not everyone who felt ill had a severe course of the disease (Bok/Martin/Lee, 2021). Some people only had a mild cold (Bok/Martin/Lee, 2021).

Another reason why people protested on the streets is that they have personally experienced social grievances (Koos, 2021). A survey conducted by the University of Konstanz at a

“Querdenker” demonstration revealed that only one in five believes what a health expert says, and 93% state that they consider the pandemic measures to be exaggerated. Interestingly, the survey also found that people are less upset about the health-related measures; instead, they find the measures exaggerated because they caused financial difficulties for many (Koos, 2021). It didn't just stop at demonstrations. People started going to secret events, meeting up with friends in larger groups and doing other things where they could get infected or infect others (Bok, et. Al. 2021). Afterwards they went home, taking the risk of infecting the people they live with together with the virus and that they could get a severe course of the disease (Bok, et. Al. 2021). Due to the constant fake news, the term infodemic was introduced, through which it could be observed that new types of cybercrime emerged, which mainly capitalized on the panic of health consequences. However, this cybercrime no longer only originated from private individuals or small businesses, but also from large companies or even governments (Gradoń, et. Al., 2021). Countries such as China and Russia have targeted Western European countries (Gradoń, et. Al., 2021).



Figure 4: Dr. Oz criticized for saying panic about coronavirus is worse than the virus itself. (Crump, 2020)

2.5. Detection and Mitigation Strategies

It is now becoming increasingly important to recognize fake news as quickly as possible, as it is shared very frequently on social media and therefore spreads quickly. Facebook, for example, accounts for 50% of fake news traffic, whereas it only accounts for 20% of traffic for legitimate websites. With more than 60% of American adults getting their news from social media, it is more urgent that fake news is recognized better and faster nowadays (Pérez-Rosas, et. Al., 2017).

2.5.1. Fact-checking, verification methods, and tools

The urgency of stopping the effects of fake news is becoming increasingly high, as it can not only jeopardize people's lives, but also pose a threat to national security. The problem with social media is that users can spread false information and change it afterwards (Zhu, et. Al., 2018). Whenever the topic of fake news is mentioned anywhere, the term "fact-checking" is not far away (Nieminen/Rapeli, 2019). Fact-checking involves analyzing and evaluating important political speeches or news items that appear on social media. In the USA, this evaluation is usually carried out by journalists who are trained to analyze news and stay informed. Since the US elections in 2016 at the latest, fact-checking has become an integral part of political journalism and journalism in general (Nieminen/Rapeli, 2019). However, it has not only arrived in the USA, but also in other countries. In Eastern Europe, various non-profit organizations are even working to stop fake news (Nieminen/Rapeli, 2019). The problem with such fact checks is that it takes a lot of time to analyze speeches by Trump, Obama, and other well-known politicians (Guo, et. Al. 2022). This can take hours or even days. However, since fake news spread quickly, professional fact-checkers are unable to check all articles and speeches on their own and studies have shown that less than half of the published articles are being checked.

As mentioned above, this work is very time-consuming, which is why an automated fact check is already being considered, which will take place with the help of artificial intelligence (Guo, et. Al. 2022). Interestingly, however, DeVerna, et. Al. (2023) argue that artificial intelligence should not be helpful in the fight against fake news. In a study they conducted, they found that although ChatGPT works well when it comes to false headlines, it quickly indicates that it is not entirely sure whether this is true. Although this is faster than a human checking these facts, there is no exact certainty about whether it is true or false, whereas a human could say exactly if it is or not (DeVerna, et. Al. 2023).

2.5.2. Legal, regulatory, and policy responses to disinformation

The French President Emmanuel Macron warned in 2018: „If we do not regulate the Internet, there is the risk that the foundation of democracy will be shaken (Tenove, 2020). “

A study found that 83% of Europeans believe that fake news is a major threat to democracy. In Taiwan, 67.5% believe this and in the US, 68% of those who were asked, believe that fake news will extremely shake trust in politics (Tenove, 2020). In Germany, publishers, journalists, and academics are calling for higher penalties for fake news so that it can be better combated (tagesschau.de, 2020). There are calls for the various social media platforms to adhere to stricter rules. The EU Commission called on platforms such as Twitter and Facebook to do more reporting on fake news and to work more closely with Fact-Checking. There have been calls for clearer sanctions to serve as a deterrent (tagesschau.de, 2020). There are no clear guidelines for states to respond to disinformation taking place abroad, even if it takes place during US elections (Tenove, 2020). As mentioned above, Russia launched a major disinformation campaign during the US election campaign to help make Donald Trump president for their own benefits (Tenove, 2020). Nevertheless, attempts are being made to counter foreign disinformation. NATO, for example, is coordinating its military and intelligence capabilities to intervene if necessary. The G7 countries have set up a rapid response mechanism that monitors intelligence services to be able to recognize election interference or disinformation more quickly and act against it (Tenove, 2020).

2.6. Ethical Considerations and Challenges

Newspapers are a form of information transmission in which the facts on which they are based are scrutinized to ensure that they are correct. Ethical journalism is about uncovering mistakes and correcting them. If you look at the ethics of the media, every step is scrutinized and selected. From the story it is supposed to be about, to news gathering, production, presentation, and publication. All steps are closely monitored (Hadžialić/Phuong, 2020). The so-called Code of Ethics does not specify exactly how journalists should behave and what they should write down, but it does provide guidance and resources to help journalists make better ethical decisions. This is particularly important in the fight against fake news because journalists are ultimately the ones who can recognize and improve these stories more quickly. Journalists have a very important role in terms of fake news because they must realize that they are the ones who should uphold values such as truth and accuracy in their research (Hadžialić/Phuong, 2020).

2.6.1. Ethical dilemmas in addressing disinformation

News and information about the situation in the world should not cause harm but rather inform people or make them think. However, this is not easy in times of social media. Fake news can even cause physical harm, as we have seen during the coronavirus pandemic (Hennell, et. Al. 2020).

During the pandemic, Donald Trump suggested that people infected with the virus should inject themselves with disinfectant to rid themselves of the virus (Trump Suggests "injection" of Disinfectant to Beat Coronavirus and "clean" the Lungs, 2020). People have followed this advice, with health consequences. Doctors then said that injecting disinfectant or other means to keep clean is very dangerous and that people who want to end their lives often use this method to kill themselves (Trump Suggests "injection" of Disinfectant to Beat Coronavirus and "clean" the Lungs, 2020).

Journalists need to be careful about what exactly they publish. Because published content can cause a great deal of damage, especially if it is not checked beforehand (Hennell, et. Al. 2020).

2.6.2. Balancing free speech with disinformation control

Everyone has the right to say what they want to say. However, due to fake news, it is becoming increasingly difficult to believe what people say and there are plans to regulate this more. However, there is a dichotomy here. As most countries live in a democracy, this is not so easy, especially as democratic societies are obliged to protect people's personal freedom of expression (Santuraki, 2019). On the other hand, society is obliged to ensure that the self-esteem of its fellow human beings and their safety is guaranteed, which cannot be provided by fake news or something like hate speech (Loewy, 1992). Since all people are equal in front of the law in a democracy, freedom of expression is unavoidable (Loewy, 1992). Freedom of expression is therefore also the perfect proof that people have different views, which is a good sign (Loewy, 1992).

It is precisely because fake news are becoming a danger to society that it is necessary to regulate it to prevent chaos and possible violence (Santuraki, 2019). In England, a newspaper article was published that was considered hateful towards migrants, so that even the United

Nations High Commissioner for Human Rights intervened and demanded that action should be taken against it. In general, people have spoken out and demanded that governments ban fake news immediately, as it jeopardizes security and peaceful coexistence (Santuraki, 2019). Like mentioned above, Koos (2021) said that because of fake news conspiracy theorists went and demonstrated on the streets during the Covid 19 pandemic because they believed fake news. Nowadays, people are exposed to extreme dangers and overload because of fake news. This is because they no longer only exist in text form but are now also published on various social media platforms in the form of images or video content that is not real but fake (Chauhan, et. al., 2022). In March 2023, a picture of Pope Francis wearing a white puffer jacket from the brand called Balenciaga was published. This post spread so quickly within seconds that suddenly everyone was talking about it (Di Placido, 2023). After a short time, it turned out that this picture was not real but generated by artificial intelligence. Although there is no danger to humanity behind such an image, considering how real artificial intelligence can make images look, this could quickly become a problem if more serious topics are falsified (Di Placido, 2023).

2.7. Future Directions in Countering Disinformation

In a survey conducted by Rossing (2023), respondents were asked what they see as future trends when it comes to combating fake news and disinformation. The most frequently mentioned answer is that people are becoming increasingly skeptical about news, as the issue of fake news has become increasingly well-known, especially since the coronavirus pandemic. Respondents pay particular attention to social media. They say that they think social media has amplified the effect of fake news. They also say that we currently live in a world where an opinion was not just an opinion, but nowadays everyone has their own version of the truth (Rossing, 2023).

It is also expected that there will be more and more specialization in topics such as media literacy. Courses will be offered or taught at school that explain how to recognize fake news. The survey also revealed that it can help to recognize fake news by looking at manipulative tactics rather than the content of the news. This allows people to quickly realize whether it is real news or news that is fake (Rossing, 2023).

2.7.1. Emerging trends in disinformation and response strategies

The question that arises is what impact fake news and false information will continue to have on people, especially as they are already having fatal effects (Dornan, 2017). Social media platforms and search engines such as Google have already announced that they want to reprogram their algorithms so that they recognize fake news or even ensure that it cannot be published at all. "Trending topics" on social media are to be verified by news services before they are published so that the circulation of false information can be prevented. The problem with this, however, is that the bots that spread this fake news will also adapt their algorithm as a result or find other ways to continue exploiting social media platforms (Dornan, 2017). Work is also already underway to introduce stricter rules for fact-checking. Websites such as Snopes.com, FactCheck.org and PolitiFact.com are there to check such facts and find misinformation (Dornan, 2017). Another trend that is suitable for the future, but is already being used today, is machine learning and artificial intelligence to recognize fake news (Berrondo-Otermin, et. Al. 2023). Machine learning is a sub-area of AI and deals with the development of algorithms. These algorithms then enable computer systems to study user behavior and learn from it (Gradoń, et. al., 2021). There are two types of machine learning: supervised learning and unsupervised learning (Berrondo-Otermin, et. Al. 2023). In supervised learning, the system is actively trained by receiving predetermined data. In unsupervised learning, the system learns from unlabeled data. It is important that the system recognizes patterns and relationships in the data (Berrondo-Otermin, et. Al. 2023). Combating disinformation, especially when it comes to acute issues such as the coronavirus pandemic, is a problem for society. These problems for society require the involvement of all sectors of the population, such as the government, the economy, and citizens (Gradoń. et. Al. 2021).

Chapter 3: Artificial Intelligence

3. AI (Artificial Intelligence)

Artificial intelligence helps computers and machines to act like humans, adopt human behavior and learn human decision-making. These technologies are so advanced that they can understand human language and interact with people. One example of artificial intelligence is Siri from the Apple brand, which people can talk to, ask it questions or give it commands (What Is Artificial Intelligence (AI)? | IBM, n.d.).

3.1 Introduction to AI

Almost everyone has been familiar with the term artificial intelligence since the introduction of ChatGPT on November 15, 2022 (Krettek, 2023). This is a chatbot that an internet user can ask any question they want and this tool answers and enables follow-up questions and corrections. This enables even students to get their questions answered more quickly and especially more targeted than for example on Google. The accuracy of the data that ChatGPT gives out is debatable since it is a tool that is trained by the people who ask the tool questions (Krettek, 2023). As one can see, artificial intelligence has arrived in people's everyday lives, whether young or old. But it has been around for a long time now, especially in the field of smart communication with Apple Siri or Alexa from Amazon (Gherheş, 2018).

3.1.1. Definition and overview of AI

Artificial intelligence is a machine that can learn everything that humans know and do and is being trained by humans itself (Lenzen, 2023). The areas that AI can learn include languages, knowledge representation, perception, and image analyses and much more. The machines are trained by computer science and use it to analyze the behavior of people on the internet, thereby storing data (Lenzen, 2023).

The aim of AI is therefore to develop a system that can solve or process problems independently (Wittpahl, 2019). Some of them deal with solutions for computer systems and others are concerned with comparisons to human intelligence (Wittpahl, 2019). It is also the name for a technology that machines can develop. They can create anything by artificial means and display human behavior without necessarily having to be a living being. They can do all this because they have learnt and been trained to do so (Mijwel, 2015). If one

looks at all the technologies that exist today, be it computers or smartphones, they were all developed based on humans. That's what the concept of machine learning tools (Mijwel, 2015).

The ChatGPT tool recently came to public attention. GPT stands for Generative Pre-trained Transformer and is one of the best-known language models used in artificial intelligence (Matt, 2024). ChatGPT is a chat bot developed by OpenAI. It is trained by humans submitting a set of text files to the tool. The way it works is that people can ask the tool questions, and the tool gives the user answers to the questions. If the answer is wrong, the one who asked can tell the ChatGPT tool that the answer was incorrect, and it can continue to learn from the feedback. However, the fact that anyone can train this chatbot also raises ethical concerns, as misinformation can often be spread via the tool (Matt, 2024). Nowadays, with the latest developments in AI, the programming language is becoming increasingly important (Mijwel, 2015). Just as humans think with their language and keep themselves alive by performing various functions with it, AI does the same. Nowadays, there are even robots that can speak and behave like humans (Mijwel, 2015). In 2018, former German Chancellor Angela Merkel even spoke on stage with an AI-controlled robot (Kreye, 2018). This robot doll is called Sophia and is programmed to imitate a human. It even masters 60 different facial expressions. They talked about topics such as artificial intelligence and cyberfeminism. Sophia is already so well known that she was even invited to appear on talk show host Jimmy Fallon's show. However, there are some complaints about the robot doll Sophia, as there are fears that using her as an example could lead the development of such machines in the wrong direction (Kreye, 2018). There are also voices saying that engineers should not humanize machines. These machines should become the subjects of humans and not humans themselves (Kreye, 2018).



Figure 5: Who owns AI ChatGPT, (DeepBrain AI, n.d.)

3.1.2. Historical development and milestones in AI

Since the end of the 20th century, AI has not only been concerned with games and mathematical representation systems but has also developed into machine learning and deep learning (Wittpahl, 2019). Machine learning focuses on trying to get a computer to learn from its experience and then use this experience to solve tasks (Wittpahl, 2019). Deep learning is a sub-area of machine learning where the computer system uses large data sets of neuronal networks to analyze ((Mijwel, 2015). As with humans, the data is analyzed in the brain and a forecast is then made (Wittpahl, 2019). Artificial intelligence first officially emerged in 1956, when the first conference on AI was held at Dartmouth College (Mijwel, 2015). LISP was then developed in 1957. LISP stands for List Processing Language and is a programming language for artificial intelligence and is now one of the oldest programming languages of them all (Mijwel, 2015). From the 1970's onwards, there was an upswing in the field of artificial intelligence. At that time, its systems were still used to recognize disease diagnoses, but towards the end of the 1970s it became clear that these systems could also be used in other areas. Over the years, these were then adapted to help solve real-life problems (Mijwel, 2015). In the early 2000s, there was an upswing in the field of neuronal networks. Once it was understood what was hindering the problem of training multilayer networks, various solutions were found, and research could continue (Buttazzo, 2023). After the first successes were achieved with neural networks, large companies such as Google, Microsoft and Facebook became aware of them and saw a great opportunity for the future in neuronal networks and invested in them (Buttazzo, 2023). Due to the enormous amount of data collected by these networks, researchers were able to analyze it and then train and test the new machine learning models again and again. In 2012, a neuronal network called AlexNet overtook its competitors for the first time, which is why large companies once again took notice (Buttazzo, 2023). Artificial intelligence has developed to such an extent that it is now used by many companies to gain an innovative advantage over their competitors (Nath/Pandey/Somu/Amalraj, 2018). Another advantage of using AI in companies is that many processes in these companies can be mechanized by the systems and employees can focus on tasks that AI still finds difficult, such as building relationships with customers (Nath/Pandey/Somu/Amalraj, 2018).

3.2. Types and Applications of AI

ChatGPT is the latest form of AI and the one that most people are familiar with (Krettek, 2023). The main purpose of this application is to make people's everyday lives easier in many ways. For example, it can write emails, create outlines for presentations or summarize texts within seconds. This tool can even create websites. Not only the text that the website should contain, but also the HTML code (Hypertext Markup Language), which is the most common programming language (Krettek, 2023).

3.2.1. Different categories of AI (e.g., narrow AI, general AI)

There are different types of artificial intelligence. Narrow artificial intelligence is a system that focuses on a single task to be performed by a machine. One example of this is the chatbots that customers talk to when they need customer service or want to return something. Other examples include the product suggestions you get when you reopen Amazon or self-driving cars (Bodell, 2023). However, this type of artificial intelligence also harbors risks. For example, there may be a malfunction if the software fails or the hardware exhibits errors (Nath, et. Al. 2018). Such an error can have particularly serious consequences for self-driving cars. For this reason, engineers build the system in such a way that there are several systems that can intervene if one should fail. Another problem that exists not only with Narrow AI, but with everything that is programmed is a malicious cyberattack. This can be carried out by the government or by specialized agencies that launch hacker attacks (Nath/Pandey/Somu/Amalraj, 2018). Another category of artificial intelligence is general artificial intelligence. This is used in an area of AI in which the aim is for this software to build up human-like intelligence and be trained (Arcas, 2023). The aim of this software is that it will eventually be able to solve or take over tasks for which it was not initially intended, but which it has learnt (Arcas, 2023). These robots have been outperforming human capabilities in industry for some time now. They work faster, more accurately and more cost-effectively, which is not even and will never be possible for humans (Buttazzo, 2023). As a result, more and more jobs that used to be carried out by real employees are being replaced by robots or machines which on one hand can be good for the industry but not particularly good for the economy. Although this already happened during the industrial revolution, AI is not only replacing the physical part of people's work but can also take over

mental tasks and information-based tasks, thus jeopardizing many human jobs. One example of this is that self-driving taxis have been permitted in California since June 2022, which is the beginning of self-driving road transport. In Paris, too, many of the trams already drive themselves. Although it brings a lot of advantages, such as lower costs or fuel consumption, it will cost many people their jobs (Buttazzo, 2023).

3.2.2. Practical applications of AI across various industries

The ever-advancing developments in artificial intelligence are leading to more and more uncertainties regarding technological capabilities and security. As AI continues to grow, it has also piqued the interest of governments, resulting in more investment in research (Mutawa/Rashid, 2020). As a result, AI systems are increasingly being integrated into everyday processes, allowing researchers to collect and improve data (Mutawa/Rashid, 2020).

AI is already being used in many areas of the public sector, as AI technologies and cognitive computing can be used to collect large amounts of data and thus make better decisions in the public sector. It can be used to carry out public health checks, for example in restaurants. Especially in cases such as the Covid 19 pandemic, such technologies can be used to ensure the safety of citizens. AI systems can also be used in education (Mutawa/Rashid, 2020).

The learning behavior of pupils and students has changed fundamentally because of social media. Hardly anyone looks at books anymore; instead, information is increasingly being taken from platforms such as YouTube and the other social media platforms (Müssig, 2021). Search engines such as Google and YouTube are based on AI by looking at user behavior and analyzing what users are increasingly searching for to present them with the best results. The use of AI in search engines saves data such as learning time, dwell time and clicks. Although experts see a future for AI in learning, there is currently not too much demand for it, as there is not enough investment in this area (Müssig, 2021). As already mentioned, the topic of artificial intelligence is currently developing rapidly, which is why governments are investing more money than usual in those technologies (Mutawa/Rashid, 2020). The US government has increased its annual growth rate in technology by 54.3% from 2018 to 2021 (Mutawa/Rashid, 2020). This makes it more important for governments to apply these technologies in the public sector (Mutawa/Rashid, 2020). One area of application could be

AI in smart cities. Researchers have the opinion that it would not be a mistake to use AI for a better cityscape. Since most cities today are equipped with sensors, telecommunication systems and more, AI should also be used here (Allam/Dhunni, 2019). By collecting data from different neighborhoods, it is possible to find out what citizens want from their city. This data is then helpful for politicians who must make decisions about the cityscape. As a result, they can make their decisions regarding open spaces, buildings, and streets, for example, dependent on what citizens want and thus enable them to feel more comfortable in their city (Allam/Dhunni, 2019).

3.3 AI and Social Media

Social media has become a big part of people's lives. Not only privately, but now also professionally (Sadiku, et. Al. 2021). Every day, several million people interact with posts on Facebook, YouTube, Instagram, etc. and discuss various topics with acquaintances, friends and even strangers. Through opinion leaders such as influencers, the use of social media has increased once again and companies have realized that their own products can be marketed well on the platforms. Artificial intelligence makes it even easier for companies to target their products and advertising videos and posts to the right audience, as AI can provide data on user activity on the various platforms (Sadiku, et. Al. 2021).

3.3.1. The role of AI in social media platforms

Artificial intelligence is appearing more and more in social networks. AI helps internet users to have an optimized customer journey. If you take Facebook as an example, you can see that the company uses AI to carry out image recognition. This then helps users to find their friends on Facebook more easily and thus stay connected with them (Devang, et. Al. 2019). Artificial intelligence also helps a company like Facebook to collect and filter customer data and then display the content that the user wants to see. AI is also very helpful in the field of marketing, as it can analyze the purchasing behavior of customers and thus estimate what the customer is currently looking for or even in which area of life the customer is currently in (Devang, et. Al. 2019). As one can see, AI is playing a major role on social media. It is increasingly being used in social media marketing. There are now many AI tools that write

social media adverts for companies (Sadiku, et. Al. 2021). This allows companies to stand out from their competitors who do not yet use AI. Such systems also help to find the right target group for the company and to play out the right adverts on the right platforms (Sadiku, et. Al. 2021). Companies often use machine learning technologies to increase brand awareness and thus attract more customers (Micu, et. Al. 2018). These tools can also recognize the preferences and interests of internet users and then target the company's adverts to the right target group which helps to prevent wastage. AI-based analysis tools have also been developed to analyze customers' social experiences, making it easier for companies to improve (Micu, et. Al. 2018). AI tools therefore help companies to monitor and analyze their social media presence. These tools can analyze which online content has a particularly high level of engagement from internet users, which can help optimize audience, image, and sentiment analysis (Capatina, et. Al., 2020). However, researchers fear that the new AI technologies cannot yet be properly used by marketers of social media platforms to discover visual concepts that are essential for the interpretation of images on the networks (Capatina, et. Al., 2020). Therefore, in the future, AI software will be developed that integrates these deep learning algorithms to better apply them to social media and thus determine user-defined categories. Initial attempts to implement this have already been made (Capatina, et. Al., 2020).

3.3.2. AI-driven content recommendation and user interaction

Artificial intelligence is becoming increasingly well-known and is being used more and more frequently in human interaction in the form of chatbots. These bots are often found in customer service when a customer has a question or a concern regarding a product (Aoki, 2020). The most important characteristics of chatbots are, firstly, simplicity. It should be sufficient if there is only one input field and it should be avoided that the process is too complex for the customer, otherwise they may drop out (Müssig, 2021). In addition, they should follow a very simple, human-like communication, which means that they should respond in short and concise sentences. They should also be able to respond to feedback and overcome uncertainties (Müssig, 2021). If the chatbot is unable to resolve the customer's problem, the customer is then forwarded to a real person who then tries to solve the problem (Aoki, 2020). A chatbot is therefore a computer-aided program that establishes

communication between the human and the bot. Since this technology has continued to improve in recent years and has become more and more accurate, it has already become a popular technology and is being used more frequently. Conversations with chatbots have become faster over the years and they can now respond better and better to people's concerns (Aoki, 2020).

There are three types of chatbots. Firstly, there are the menu-based chatbots. These chatbots work by holding conversations with the user by giving them various options. So, if a customer has a problem with their goods, this chatbot asks what exactly the problem is and offers various scenarios to choose from. The customer can then select a scenario, which the chatbot collects and analyzes and therefore offers more specific scenarios to choose from until a solution is found (Gupta/Hathwar/Vijayakumar, 2020). The second type of chatbots are keyword recognition chatbots. These chatbots are programmed to recognize certain keywords and listen to what the customer types to give the right answer (Aoki, 2020). The third type of chatbots are contextual chatbots which use technologies such as speech recognition and machine learning algorithms, making them the most advanced bots on the market. The purpose of these bots is to recognize the user's intentions and provide a thoughtful response (Gupta, et. Al. 2020).

3.4. Ethical and Social Implications of AI

Some researchers predict that it will be hard to imagine our lives and everyday routines in the future without robots doing the jobs that humans used to do (Gherheş, 2018). However, many critics are expressing concern. If artificial intelligence gets out of human hands and out of control, it could be dangerous for society. Even people like Stephen Hawking, Bill Gates, Apple co-founder Steve Wozniak and Tesla co-founder Elon Musk warn of a world where artificial intelligence surpasses that of humans, even though they all use artificial intelligence themselves (Siri at Apple and self-driving car systems at Tesla) for their technologies (Gherheş, 2018). Artificial intelligence is now used in robots and even self-driving cars. It can also be used to recognize hazards in cars driven by humans. The cooperation between humans and technology is already well developed, but the question arises as to what could happen if such technology fails and people could even be harmed because of the failure (Buttazzo, 2023). Looking at self-driving cars as an example, there are

pros and cons. For example, if a pedestrian crosses the road, the AI-tools within the car can recognize this and, in the best-case scenario it will make the car stop. However, if the AI decides to take evasive action by changing lanes, it may crash into another car, which may have more occupants, although the law does not allow the number of human lives to be outweighed. All in all, this situation would be a danger for people. For this very reason, it is important to train and program the systems in such a way that they are not only equipped with necessary information, but also with ethically justifiable information (Buttazzo, 2023). Isaac Asimov defined three laws for such cases in 1942.

First Law: *A robot may not injure a human being or, through inaction, allow a human being to come to harm.*

Second Law: *A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.*

Third Law: *A robot must protect its own existence as long as such protection does not conflict with the First or Second Law. (Der Mensch und die Dinge – „Zeroth Law – Das nullte Gesetz“ von Gamut Inc in der Tischlerei der Deutschen Oper Berlin, n.d.).*

However, as these laws are very unclear, another law called Zeroth Law was introduced, which states that a robot may not harm humanity or allow humans to be harmed by doing nothing (Buttazzo, 2023). So, if someone threatens to destroy humanity, a robot has permission to eliminate them. However, the programming of such problems and situations could cause confusion. Since the robot cannot accurately assess whether it is better to save only one person instead of five, or vice versa, to save five instead of one person, it could, for example, be programmed to want happiness for people, which doesn't sound bad at first, but it could get happiness by removing organs from a healthy person to save five others (Buttazzo, 2023). The topic of machine learning raises several ethical concerns. As these systems learn through big data analyses, it becomes particularly problematic when they receive personal or non-representative data. In addition, the exact functioning of AI is quite difficult to determine (Stahl, et. Al., 2022). One issue that raises major concerns here is data control. This raises the question of how confidentially the data is processed by the system, especially if it is personal data. It is considered critical because the AI has access to large data sets that could contain personal data and thus the link could lead to the identification of

individuals. Another concern is how accurate the data is. Especially when it comes to important data, such as healthcare data, accuracy is essential (Stahl, et. Al., 2022). In addition, these systems are currently being trained for men and women. So, if a diagnosis is to be made for a man, but the algorithm makes a mistake and makes a diagnosis for a woman, this could have fatal consequences. Another issue is transparency. Because the systems are so new and in the middle of their development, some of their developers do not even understand the inner workings of these systems. This lack of transparency is already a concern but could also lead to distortions in data sets. In addition, the lack of transparency also raises concerns about responsibility and liability when using AI systems (Stahl, et. Al., 2022). There are several points in society where the use of AI is viewed critically. These are often higher-skilled jobs but also, for example, how AI can support economic power, where concerns arise as it is strongly associated with intellectual property (Stahl, et. Al., 2022). Another area where the use of AI is viewed critically is when AI is used for law enforcement, as vulnerable groups could be jeopardized if mistakes are made (Bothmann, n.d.). As already mentioned, AI is increasingly being used in the healthcare sector. There are concerns that human contact in care, for example, could become increasingly rare if AI were to take over this role. What people are most worried about, however, is the uncertainty. AI is currently still very new and not particularly mature, which means that criminals can turn this to their advantage (Stahl, et. Al., 2022).

3.4.1. The impact of AI on employment, privacy, and society

The use of artificial intelligence in everyday life is a subject of much debate. Although this technology brings enormous benefits, it also has its downsides. The progressive development of these systems means that human assistance is no longer needed in many areas (Tai, 2020). It is assumed that AI will bring a major social change. Currently, people must work hard to get good grades to earn a good salary and live a good life of it. AI will change this, because then it will only be important to be able to train and program such systems so that it can do the work for people. This means that people will become increasingly lazy (Tai, 2020). It is also feared that human contact will decrease because of this technology, as AI will no longer make face-to-face dialogue necessary. It is also assumed that many jobs will be replaced by artificial intelligence. Since it is already becoming increasingly common to see work that

was previously performed by humans now being carried out by machines (Rigner, 2019). One example is the assembly lines in the automotive industry or the self-service checkouts in some shops (Rigner, 2019). Another negative aspect concerns AI itself. As it is a system that needs to be trained and is constantly learning new things, it is possible that humans will lose control of these systems at some point and they will start to train themselves, which can lead to unforeseen problems (Tai, 2020).

However, artificial intelligence also has positive aspects. For example, scientists could develop AI that targets medical diagnoses and treatments. Robots can also be used for interventions, making the work very precise and thus increasing a patient's chances of recovery or minimizing the risk of errors during interventions (Ahuja, 2019). Computer systems have already been used to test how cooperation with AI works. Here, the patient's data was given to the computer, and the diagnosis came back within a very short time. Treatment options for the diagnosis are also being suggested by it (Ahuja, 2019). Another advantage of AI is that it is a machine and therefore has no signs of fatigue (Tai, 2020). The often-costly mistakes that humans make because they are unfocused or emotionally distracted by personal circumstances do not occur with AI systems (Tai, 2020). Artificial intelligence is used nowadays particularly in magnetic resonance imaging (MRI) research. Although these are already very good and advanced, the search for better algorithms to recognize specific diseases continues (Ahuja, 2019). AI can also be used to make better remote diagnoses. This means that patients do not have to leave their homes, but doctors can examine the patient with the help of a robot and then make a diagnosis. This can relieve the burden on medical staff and in cases such as the coronavirus pandemic, as these systems can be helpful since patients can be examined without infecting staff (Tai, 2020). Most of the literature focuses on the negative effects of artificial intelligence on people's lives, but it is often forgotten that it also has positive aspects (Stahl, et. Al., 2022). One of the advantages is that AI can collect and process large amounts of data very quickly, allowing decisions of all kinds to be made more quickly. It performs these processes at a very high speed, which a human would never be able to do in the same amount of time. Artificial intelligence also makes it possible to optimize processes, which increases productivity in the economy and makes it faster, leading to prosperity in society (Stahl, et. Al., 2022). Normally, such processes are carried out by humans, which means that they must work longer on them and spend resources on them, which AI does not have to do, which is why the use of AI reduces

the environmental damage caused by production (Stahl, et. Al., 2022). The use of artificial intelligence in business is often associated with three factors. Automated business processes, autonomous driving, and higher product quality (Gherheş, 2018). Automated processes using AI mean that more can be produced or processed more quickly, thus increasing productivity gains. Autonomous driving can save costs as drivers no longer need to be paid and it is more environmentally friendly as the vehicles maintain a constant speed. In terms of product quality, AI can also work more accurately and quickly, which can increase sales. In addition, there is no need to pay staff (Gherheş, 2018).

3.5 Challenges in AI Development

Artificial intelligence can be a support in many areas of life, but it can also bring challenges. Caution is required when programming AI. For example, if an AI system is developed for use in medicine, the system should be programmed exclusively for this purpose and not, for example, for another purpose like writing good articles (Cheng, et. Al., 2021). In the medical field in particular, the AI's decision can make the difference between life and death, which is why smaller AI companies should not be blinded by the current hype surrounding AI and should be aware of the importance of their work (Cheng, et. Al., 2021).

3.5.1. Technical and ethical challenges in AI research and implementation

In 2018, at least a dozen countries presented their strategy for artificial intelligence. The industry is also adapting to this change and developing its own AI principles (Cath, 2018). However, care must be taken here to ensure that the industry does not intervene too much in the topic and try to determine it, as the use of AI in various sectors, such as finance, education, or criminal justice, is not without danger (Cath, 2018). However, initial concerns are emerging. Most AI companies are based in the United States. Accordingly, these technologies are trained according to US preferences, lifestyles and laws. There are also concerns about data protection. Data protection regulations in the EU differ from those in the USA. However, if an AI system is trained according to American guidelines, it could cause problems when used in the EU (Cath, 2018). In addition, AI is often described as a black box that is so complex that it is difficult for humans to understand. This also raises

concerns, as poor results could be achieved, or incorrect behavior of the AI could occur. The large online marketing groups, such as Google, Facebook, and Amazon, can collect a lot of data through their large network of customers and make this available for new AI-based services (Cath, 2018). On the one hand, this helps the development of AI, but also the companies, as they are involved in the regulatory processes. Cath (2018) sees this as a problem, as these companies will have more influence on AI regulation. For this reason, the former German councilor Angela Merkel and President of the European Commission Ursula von der Leyen introduced the General Data Protection Regulation (GDPR) in 2019 (Siegmann/Anderljung, 2022). In 2020, the first proposal for laws for the digital market was presented (Orszag, 2023). The aim of these laws is to regulate the safety of people's personal data so that the large corporations like already mentioned above are not too integrated into the decision-making processes of AI regulations and cannot exploit it in their own favor. This keeps the digital markets competitive (Orszag, 2023). These laws were passed in the summer of 2022. In particular, "gatekeeping", i.e. the marketing of companies in search engines, is to be prohibited and it is expected that all large technology companies will be categorized as such (Siegmann/Anderljung, 2022). Companies and policy makers interested in regulating AI should be mindful of the GDPR as the system may spread globally. In addition, they and companies not based in the EU could seek greater involvement in the EU regulatory process. As a result of the EU directives, the AI industry should therefore ensure that its systems also include topics such as explain ability, transparency, and fairness. If these systems fail to do so, the European market for these technologies will disappear (Siegmann/Anderljung, 2022). Another reason to consider is cybercrime. Precisely because the liability framework has not yet been fully clarified, it is currently still difficult to identify people who commit cybercrime and then punish them. There are various trends in cybercrime where AI can help. Most relate to the Internet of Things, with ransomware attacks being launched. These are cyberattacking where data is encrypted or stolen, and people or companies are then threatened with the release of the data to start a ransomware attack (Velasco, 2022). This type of attack is expected to continue, as it is estimated that in the next five years there will be more than 2.5 million devices online that are used not only for personal use, but also in industry or infrastructure (Velasco, 2022). Fake news is also becoming increasingly popular and spreading faster. This is mainly due to AI-controlled bots that spread such news quickly, which is intended to mislead society (Velasco, 2022). In

addition, these bots and fake news are causing a loss of trust in the media (Velasco, 2022). The challenges that the world will face in connection with AI are clear. In addition to the legal challenges, there is also the aspect of technical challenges, which is causing concern among scientists (Dhar, 2016). A major challenge will be to combine machine learning with symbolic AI. Symbolic AI does not refer to data in the form of numbers, but rather to the representation of knowledge. Furthermore, it is currently still too dangerous to let the technologies learn unsupervised, as there is a risk that they will develop in the wrong direction. It is still a long way from being at the level of humans, although the question arises as to whether this should be the goal at all (Dhar, 2016).

3.5.2. Regulation, Policy, and ethical considerations

The greatest fear of scientists and politicians is that artificial intelligence will develop to such an extent that it will begin to surpass human intelligence. This often raises the question of how these systems should be regulated. These systems are now able to formulate hypotheses on their own and can also develop further themselves (Dhar, 2016). The problem here is that if these machines take over everyday life and are widely used, which they already are, and then start to program themselves in a way that humans cannot understand, it will probably no longer be possible to switch them off. If machines start programming themselves in a direction that humans do not like, this could have significant consequences for humans. For example, these machines could harm a party because the machine's principles do not match those of the party. In 2016, the Google car had an accident for the first time. Even though the car was programmed to drive autonomously, it was also programmed to recognize traffic behavior (Dhar, 2016). Although nothing else happened in the accident and no one was hurt in a serious way, this shows that machines can also fail or malfunction and possibly harm humans. The cyberneticist Norbert Wiener warned 50 years ago that there would be considerable damage if people lost control of their machines (Dhar, 2016). The biggest fear, however, is that the data from these technologies will be misused by being made public or by violating data policies. In European law, for example, privacy is a fundamental human right (Forcier/Gallois/Mullan/Joly, 2019). This is precisely why the General Data Protection Act (GDPR) was introduced in 2018 and is being followed strictly. Every company must comply with this regulation, even if many still have problems with it

or data is often published inadvertently (Kosinski/Kosinski, 2024). Even large companies such as Meta or Amazon are not excluded from this. This law specifies how companies based within Europe, but also outside Europe, should manage the handling of data (Kosinski/Kosinski, 2024).

3.6 Future Trends and Research Directions in AI

AI is becoming an increasingly important topic due to its diverse applications. However, research has shown that more research is needed, especially to understand AI even better in terms of the strategies of the systems. Research areas include the training of potential workers to further train these technologies, as well as legal and ethical issues (Loureiro, et. Al. 2021).

3.6.1. Emerging trends in AI technology

Artificial intelligence is already widely present in people's everyday lives. This will not change in the future. Researchers say that this technology will offer great benefits to the public, especially when it comes to areas such as medicine, transportation, the environment, or criminal justice (Bundy, 2017). Machine learning, for example, still has room for improvement. It already works well, but Zhou (2016) says that if, for example, circumstances in the environment or surroundings were to change, the system would not be able to adapt and would have to be reprogrammed accordingly. A machine that can learn its functions should be reusable. It should be able to adapt quickly with the help of new data with which the system is trained. This topic is becoming increasingly important in the field of AI, as the world is developing very quickly, and new events or regulations are constantly occurring (Zhou, 2016).

Chapter 4: Methodology

4. Methodology Part

In the following section, the methodological procedure required for the evaluation and analysis is explained in more detail for a better understanding. To better explore the relationships, the focus is first placed on the appropriate research design chosen to achieve the best possible results in relation to the research question and the literature review. This is followed by a detailed description of the sample and the population used to answer the research question. After that an explanation follows of how the methodology was approached and what results were obtained from the research. Finally, the data cleansing and the limitations are explained in more detail.

4.1 Outline of the objectives

This section explains why the methods were chosen. Methods in general bring clarity to the research and transparency is ensured. This makes it easier for the reader to understand the process that was carried out. In addition, a clear methodology ensures, among other things, that bias and errors can be avoided. In the methodology section, the connection between the literature section and the research results is explained in more detail, but differences between literature and research are also suggested. The research that is set up can also guide other authors in the future who deal with the same or a similar topic or expand on results.

Coming to the appropriate research methods, three objectives were defined in advance, which were to be researched by the thesis and on which the methods section was based. The first objective is to research the attitudes and mindsets of social media users towards AI-supported mechanisms that are used to detect fake news. It also aims to explore what social media users believe about how effective AI algorithms are at detecting fake news. And the third objective is to explore the ethical and social significance of AI-supported fake news detection mechanisms. The first two objectives are to be researched through a survey in which normal social media users are questioned. Since the third objective can hardly be answered by people without much knowledge on the topic of AI, it there were interviews conducted with experts. This means that the form of the methodology is a mixed-methods approach.

4.1.1 Mixed-method approach

To gain a comprehensive insight into whether artificial intelligence could have an influence on how fake news is perceived on social media, a quantitative online survey was conducted in this research project. This approach was chosen because the research question of this project focuses particularly on the normal social media user. Furthermore, this method offers advantages such as the collection of data from many people, which leads to a broad perspective on the research topic. A survey also offers quantitatively relevant information. As already mentioned above, the research question relates to the general social media user, who can be reached by conducting a survey and thus a conclusion can be drawn for the overall popularity. In addition, the anonymity of such a survey offers participants a certain degree of security, which is why it is expected that they answer truthfully. The standardization of the survey also makes it possible to largely exclude confounding variables. All in all, this method proves to be the quickest and most efficient method suitable for the period of writing the thesis. A survey makes it possible to receive answers from as many people from the population as possible with a low budget and little effort. The Google Forms tool was used to program and implement the questionnaire. It is a free tool that is accessible to everyone having a Google account.

An expert interview was chosen as the second part of the methodology. It was initially unclear whether an interview was necessary for this research, as the research question relates to general social media users. Since the general social media user is unable to provide ethical and social assessments on the topic of AI, it was decided to do an interview with two experts to gain more information about the ethical and societal part which is supposed to be discovered with the third objective mentioned above. The advantage of consulting experts is that they are well and deeply familiar with the topic of artificial intelligence and can therefore offer deep insight information on that subject matter. Experts can also name possible long-term solutions and strategies that will help to get the topic of AI under control which can provide information for further research on that topic. Such interviews support and validate the credibility of the research results.

It is then interesting to compare the results of the two methodological approaches and analyze whether the opinions of experts and general social media users differ with what the results show and how it differs with the literature that was found.

4.1.2. Research design

In this section the research design is being explained and examined more for both, the survey and the interview. This helps the reader to understand what the thought of the researcher was in creating this design and how it helps to answer the research question in the end.

4.1.3. Description of the research design

The research question "The role of AI in battling disinformation and fake news on social media: How can AI be used to identify fake news on social media?" chosen for this research project examines the role of artificial intelligence in identifying and combating fake news on social media. Specifically, the question will examine how these AI-supported systems can help to combat fake news on the internet. It has not only recently become known that fake news can be spread quickly on social media. This research question is intended to highlight the urgency of this issue. In addition, the topic is intended to draw attention to the potential of AI, be it positive potential through the fight against fake news or support in the spread of fake news. The question is also intended to find out what challenges and limitations AI systems pose and will pose in the future and what can be done about them. As already mentioned above, a mixed-methods approach was chosen for this work, which means that not only quantitative research in the form of a survey, but also qualitative research in the form of two expert interviews were carried out at the same time. This decision was made because three objectives were set at the beginning, two of which can be researched using a survey to gain answers of the standard social media user. The third objective, however, goes beyond the knowledge of the normal social media user, which is why it was decided to consult an expert through an interview. In this way, the researcher can view the complex topic from different perspectives. In addition, the topic is highly topical, which is why it is even more important to include the viewpoint of an expert.

4.1.4. Justification of the research design

The topic of artificial intelligence is relatively new and has been on everyone's lips, especially since the invention of ChatGPT. The topic is also complex and multi-layered, as there are many different opinions, and the technology is developing rapidly. This is one of

the reasons why a mixed-methods approach was chosen, as both a survey and an interview can be used to analyze this complexity comprehensively, both through numbers and results, as well as through words and thoughts. It is also interesting to take up the opinion of an expert with the help of an interview. Looking at the research question, it becomes clear that a certain scope must be analyzed and researched. This is only possible if an in-depth qualitative insight is ensured but also if broader data is revealed through quantitative research. As already mentioned, the topic is very complex, which is why it is important to have a holistic understanding of the topic. This can be achieved well through a mixed methods approach, as the combination of interview and survey allows you to look at different perspectives from different groups of people. In this way, trends and experiences can be examined, particularly through the qualitative research method. The research question mainly relates to general social media users, who can be easily reached and interviewed through a survey. Taking a sample is particularly helpful for complex and highly topical issues such as artificial intelligence, as data can be collected and analyzed from a larger number of people. In the chosen research method, between 100 and 200 people are asked the same questions, to which they can respond anonymously. The results of this survey can then be generalized to the entire population. The high number of participants and their different answers creates a certain diversity and ensures the credibility of the results. As already mentioned, the answers to the survey help to project them onto the rest of the population, which contributes to a better understanding of people's attitudes towards AI technologies and the topic of fake news.

However, AI is used in practice in relation to fake news, which is why it makes sense to gather qualitative insights, especially in connection with the ethical and social impact.

4.1.5. Explanation of the research design

By choosing an interview (qualitative) and a survey (quantitative) as the research method, the results can be viewed from multiple perspectives. Through the interview, in-depth, qualitative perspectives on the topic of AI in the fight against fake news can be explored. The advantage of a qualitative interview is that the participants can express in their own words what their opinions and concerns are on the topic, which can lead to a deeper understanding. Interviews can also lead to a better understanding of why the person

formulated their answer the way they did. It also provides a lot of information in the form of statements that can influence the participant's perception of the topic. This is different with a survey, as participants can choose their answers, but with a pre-defined choice of words and answers. However, surveys can better capture topics such as trends and patterns, as a survey allows a larger sample to be taken. This means that the answers to the topic are presented in a more generalized way. Due to its large number of participants, a survey offers the opportunity to measure the opinions of people from different groups. This can for example relate to demographic origin or age.

In combination, the two research methods provide a comprehensive understanding of the role of AI in the fight against fake news on social media from multiple perspectives. On the one hand, more general data is collected through a survey, and on the other hand, more in-depth data is collected through an interview with an expert. This makes it possible to explore general trends on the topic, but also to include individual experiences and opinions in the analysis, which contributes to a more comprehensive understanding of the research topic.

4.2 Choosing the target group

A sample of around 100-200 participants was taken for the survey. Participants were required to be between the ages of 20 and 80, use social media and live in either Europe or the United States. There are several important and valuable reasons for this demographic. One of these reasons is that Europeans and Americans have different social values and different attitudes towards technology, which is why it is interesting to investigate the behavioral differences of these two continents. Another reason, which is particularly relevant for this research, is that US citizens were included here because the US elections will take place in 2024, the year this research paper will be published. Since the topic of fake news was a particularly big issue in the US elections in 2016 and the literature section of this thesis addresses this, it was decided to investigate how American citizens view the topic of artificial intelligence in connection with fake news. The American population is particularly interesting to analyze, as they are especially susceptible to fake news. As mentioned, in 2016 the topic of fake news became particularly popular, which is why it is important to find out how Americans feel about it in a year where elections are taking place again and to find possible ways on what to do better. In terms of age, it is interesting to analyze the different

experiences and perceptions of social media and AI between the different generations. It is interesting to find out whether the older generation, which is not as familiar with the dangers of social media and artificial intelligence as the younger generation, is more likely to fall for fake news in the form of text, video, or images. It is also essential to explore whether it is the younger or older generation that tends to use AI-based fake news detection tools. By comparing the responses from the different generations, it should become clear whether AI-based solutions to defeat fake news are understandable and usable. By choosing such a large age group, a wide range of people who use social media is captured, who have already had different experiences with fake news on social media.

Europe and the USA are both markets that are relevant for artificial intelligence companies. The answers that emerge from the survey regarding the user-friendliness of AI systems that can recognize fake news or the attitudes of respondents towards the topic offer an enormous pool of topics for further research in the future. AI companies can also use these results to initiate further or new development steps and adjust or adapt to user preferences. Another reason to take a closer look at these two regions is that by comparing the responses, differences, and similarities in relation to the topic or perceptions can be better analyzed.

4.2.1. The chosen Sampling Frame

Regarding the research question which is “The role of AI in battling disinformation and fake news on social media: How can AI be used to identify fake news on social media?”, it is primarily important that people who use social media at least regularly (at least once a day) take part in the survey. The answers of these people provide an insight into whether they often encounter fake news on social media or whether they have already considered using AI tools to check for fake news. In addition, user behavior and perception in relation to fake news and artificial intelligence can be researched. In addition, many people use social media to stay up to date and obtain news. For the interview, two people who work with artificial intelligence will be interviewed. This is important to gain an insight into the topic of AI on a deeper level. Not only can technical information be gathered, but also information about what ethical consequences experts fear.

4.2.2. The chosen sampling method for the survey

The age group of 20 to over 80-year-olds was chosen for this research project as the different generations have a different understanding of social media. Younger generations generally use social media platforms more often and integrate them more often into their everyday lives than older generations. Accordingly, the younger generations are more enlightened about what trends there are, what topics are discussed on the platforms and what dangers social media entails. Another criteria for selecting the target group is that they must use social media at least once a day and consume its content. This means that the participants in the survey open these apps, share and/or comment on posts once a day or more. These platforms include Facebook, Instagram, and Twitter. As already mentioned, participants who live either in Europe or in the USA were explicitly surveyed. This is partly because the research project also relates to European literature, but also because the presidential elections will take place in the USA in 2024. These were the starting signal for the term "fake news" in 2016, which is why it is interesting to ask Americans how they deal with it eight years later and if they will expect same situations as in 2016. This sampling is intended to find out how social media users respond to artificial intelligence, whether and how they use artificial intelligence tools and how they deal with the topic of fake news. By choosing the target group described above, it can be ensured that relevant results can be found out about how the different generations deal with artificial intelligence in relation to fake news. The participants for the used survey have been drawn from Instagram and private connections.

4.2.3. Sampling size determination

136 participants were interviewed for the survey. This size was chosen because a larger number of participants guarantees quantitative relevance. The large number of participants means that the evaluated data can be analyzed and interpreted accurately. In addition, multiple participants minimize the risk of sampling errors, and the results can represent the population referred to in the research project. Furthermore, a survey of 136 participants can be realized for the given time of the research project. Another reason for the size of the sample is that this allows the different results and behaviors of the generations with AI and social media to be categorized, as it can be assumed that 136 participants will give several different answers. This helps to gain different insights into the role of AI in relation to the

general social media user. With 136 participants, valuable statistical analyses can be carried out and thus conclusions can be drawn about how the effectiveness of AI on fake news in social media affects the public. These results can then be compared with or added to the existing literature. Since the topic of artificial intelligence is highly complex and social media is used differently by many people and has a different meaning for many, it is important for the research project to ensure diversity among the test subjects in terms of age groups and the location of the respondents. The strategy to recruit participants is that participants were contacted via both Instagram and LinkedIn. LinkedIn was used to recruit for the American network, while recruitment via Instagram focused on European participants.

4.3 Survey structure and explanation

The next part explains how the survey was set up, how it was structured and why it is relevant to ask those certain questions that were decided to be asked in the survey. The structure of the survey was introduced with an introductory text that describes the topic and explains the purpose of the survey, which helps the participant to decide whether to take part in the survey or not. It was also explicitly mentioned that the survey is anonymous, since this gives the participant more trust in answering truthfully. The participant was then asked whether they would like to take part in the survey. The first three questions are exclusively general questions about the demographics of the respondents, for example, the question about age, origin and profession. Questions five and six belong to the category "Social Media Usage" in which the participants are asked about how often they use social media during the day and what they use it for. The first six questions were therefore used to obtain more general information. The main reason for this was to show the respondent that the survey can be completed quickly and is not too strenuous. This prevents them from dropping out after the first few questions. Questions five and six are essential to find out what role social media plays in the respondent's life. Questions seven, eight and nine serve to gain a better understanding of how respondents perceive fake news. By asking what the participants understand by fake news, a better insight can be gained into how the respondents' own interpretation of the topic is. Question eight can be used to find out how often the respondents think they come across fake news from social media in their everyday life. This allows an

assessment to be made of how important the topic is to the person and how they might be attuned to it. Question nine asks respondents to give examples from real life. This information serves to increase credibility and the result can ultimately be better transferred to the overall population. Questions ten to 13 belong to the category "Efforts and Perception on overcoming Fake News". Question ten is intended to find out how effective they find current methods for battling fake news, and it is possible to find out how effective the respondents find these methods. The answers can also be used to filter out what can be improved in general about the methods. Question eleven aims to find out how respondents normally carry out a so-called fact check to explore how skeptical they are in general and to find out which tools they use for this, if they use any. Regarding AI in particular, these answers can be used to explore the frequency with which AI fact checks are used and to identify whether there are suggestions for improvement or other solutions for the AI. This leads on to question 12, which aims to find out whether normal social media users are already aware of AI-supported tools that can detect fake news and thus whether the social media are communicating this well enough. Question 13, which asks whether the participants in the survey have already encounter such tools, also helps to get a more precise overview of how social media users really accept these tools. The next five questions belong to the category "Attitude towards Fake News and Artificial Intelligence". Question 14 is intended to find out how serious the respondents perceive the topic of fake news and artificial intelligence to be. This information can be transferred to the general population. The following question is intended to find out what the challenges could be that could be posed by AI tools that are used to identify fake news. This allows us to explore what the obstacles to AI are in the eyes of the respondents and encourages future suggestions for improvement. Question 16 is interesting, as it aims to find out whether normal social media users would give more credibility to information if it had been checked by an AI tool. The answers suggest how much people trust AI, but also what fears they still have about it and what could be improved in this respect. Question 17 is particularly important as it aims to find out whether respondents see AI as something that suppresses the spread of fake news. This question shows whether people trust that AI will reduce the spread of fake news or whether they believe that it will have the opposite effect and that AI is harmful rather than useful. Question 17 is an introduction to question 18, in which respondents are asked to explain what possible concerns they have about the use of AI. The next two questions relate to the category

"Mitigation Strategy". When asked about whether respondents think education is necessary, the answers can provide information about how important they think it is to provide more education and the answers also show whether there is still a need. By asking about trust in news that can be seen on social media, it is possible to refresh how much the trust of social media users has already been abused and how the credibility of news on social media is perceived. Questions 21 to 23 belong to the category "Transparency and Effectiveness of AI". Question 21 is intended to explore how important transparency is to the respondents when it comes to how the fact-checking tools are and how they work. This can be used to find out how much trust there already is in AI and whether transparency could strengthen trust. This can be used to find out what could lead to greater acceptance of AI. All in all, the survey aims to provide a better understanding of how social media users react to fake news, how aware they are of the issue and what tools they are already using to fight it. The survey also serves to find out how and whether artificial intelligence is already implicated in this process.

4.3.1 Data Collection Method

The survey contains 20 questions with different answer options. The aim was to find between 100 and 150 participants for the survey, each from the USA or Europe, aged between 20 and 80 and using social media. The first five questions are general questions about the demographics of the participants. They ask how old the participants are, where they come from, what they work as and whether and why they use social media. These questions are essential for the survey and serve to assess who the participants in the survey are and whether criteria such as age, origin and profession play a role in the results. The other 15 questions delve deeper into the subject matter. In the sixth question, participants are asked to describe in their own words what fake news means to them. This is important to recognize the participants' definitions and to see if they differ and vary. The following question is about whether the participants believe that they often come across fake news on social media. The answers show how aware the participants are of fake news and whether they are careful with it. Question 8 is a voluntary question where the participant is asked to name a case they remember where they were unsure whether it was fake news or not. These answers help to identify how often or rarely people stumble across such fake news on social media. The next

question is about how effective the respondents perceive methods for combating fake news to be. The answers provide information about people's perception of such methods. The tenth question is about what steps respondents take when they come across alleged fake news. This makes it possible to determine how respondents deal with fake news and which tools they use. The next question is about whether the participants are aware of the efforts of social media platforms in the fight against fake news. The answers show whether the social media platforms communicate this well or whether there is still room for improvement. Question 12 asks whether the participants have ever used AI tools for fact checking. This question serves to find out whether people are generally aware that such fact checks are possible and that such tools exist. In addition, the answers also show how trustworthy these tools appear to people. Question 13 is a voluntary question that can only be answered if the previous question has been answered. If the previous question was not answered, this question can be let out by the participant. The next question is about the participants' concerns on the topic. The answers provide information about how urgent this topic is and how important it is to find solutions as quickly as possible. In the next question, the participant should say what challenges they find in the use of AI to identify fake news. The answers should clarify public concerns about AI. Question 16 is about whether the participant would be more likely to trust news if it had been checked by an AI beforehand. This question is designed to find out how trustworthy the AI appears to people and whether the AI has an influence on user behavior. The next question asks whether the participant is of the opinion that AI tools are helpful or rather cause even more damage in terms of fake news. The participant's assessment can be used to explore how they perceive the potential but also the risk of AI and whether they are aware of this. Question 18 asks about possible concerns about AI. Here it is also possible to find out how aware the public is of the risks of AI. In question 19, the respondent is asked for an assessment regarding the further training of students or people in general regarding AI. The answers can be used to identify how important media literacy is perceived to be for the public. The last question is about how the presence of fake news changes trust in information. This makes it possible to find out how trust has already changed and whether solutions are needed here.

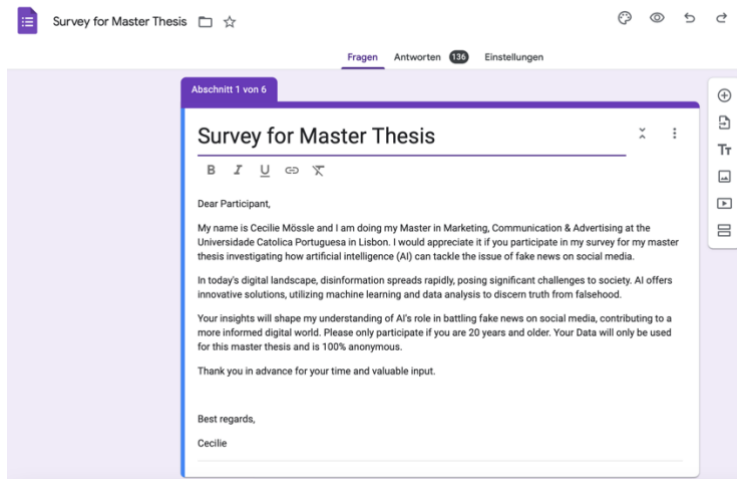


Figure 6: Screenshot of Google Forms showing the introduction to the survey, Source: https://docs.google.com/forms/d/1_n2cG1smWDpqGqfXctiLFNX0oMb7TC5JO1GyRZv2xG4/edit

4.3.2 Justification of choosing a Survey

As can be seen, a survey is important for a topic like this and generally for a work like this, as there are various reasons that show that with the type of methodology, extraordinary results can be achieved that help with the research. A survey helps the researcher to directly obtain primary data at first hand, which applies exactly to the research question. Since the researcher steers the respondent in a certain direction by asking questions that fit the research question, the respondent cannot go too far afield with their answers. In addition, the advantage of such primary data is that it is up-to-date, and the researcher can therefore be sure that it is not outdated, unlike secondary sources, such as literature, which is older. In addition, a survey offers target group-specific findings. The researcher only sends out the survey to suitable participants or the participant becomes clear in the introductory text of the survey at the latest what it is about and whether they themselves are qualified to take part in the survey and answer the questions. Through a survey, mainly quantitative data is collected, which helps to recognize trends and patterns of the topic and accordingly a deeper analysis of the topics can be carried out. In addition, the quantitative results allow the researcher to make better suggestions for the future. As a survey always involves quantitative results, it is also always easy to compare these results with other results from other research that has already been carried out, as these are figures. The answer options can also be easily compared within the survey, because the survey gives each participant the same questions, but the scope for interpretation is large and everyone understands the question in their own

way. This makes it easier for the researcher to proceed. The survey also allows the researcher to make firm recommendations for action, as they can easily recognize what is still missing on the topic or what could be worked out. It can also make recommendations as to which areas should be researched further. This increases the practical benefit for other researchers and enhances the value of the thesis. In addition, a thesis should have an empirical basis. The fact that a survey can be created quickly and free of charge makes this type of research particularly suitable for students. They also provide a good basis as they are based on the actual opinions and experiences of the respondents. Surveys are therefore recommended if you want to collect quantitative data, but also want to leave the option of collecting qualitative data open. They offer valuable insights into the target group that the survey deals with, which is relevant for the work and research.

4.3. Interview section

In this part it is explained how the interview was structured and why the participants were chosen. This will help the reader to understand the importance of the data that is provided from the experts.

4.3.1 Importance of Interviews in Research

For the second part of the methodology, the research method of an interview was chosen. There are three objectives for this research project. Two of these objectives can be achieved and answered by the survey described above. However, the third objective, which reads as follows: "Explore ethical and societal meaning of AI-powered fake news detection from the perspective of the internet user.", can only be answered by someone who is more familiar with the field than just the general social media user. Therefore, experts were chosen to do an interview.

However, interviews also have other advantages. An interview can provide much deeper insights into the investigation of the research question that cannot be achieved by other methods, as interviews reflect the exact views and thoughts of the interviewee. The interviewee shares his or her opinion and insights, which leads to many different findings. Especially with a more complex topic such as AI, which is also very current and new, it is

important to get insight information from someone who is familiar with it to better understand the complexity of the topic. In addition, although there is an interview structure that includes a certain number of questions, the conversation may lead to additional unplanned questions that can also be asked of the interviewee, providing even more insight into the topic. This type of interview is called a semi-structured interview. Another advantage of interviews is that they can produce results that were not expected. Although one or two results can already be assumed, further insights or ideas can emerge through the exchange with the expert and his experience that would never have emerged from respondents to a survey. Another advantage, which is particularly helpful with a mixed-methods approach, is that the answers from the survey can be validated and underlined by the respondent's answers in the interview, which increases trustworthiness and value. As an interview is a qualitative type of research, as mentioned above, the spoken word of the interviewed expert and therefore their thoughts and experiences are captured. These insights can be particularly revealing and therefore an interview is a good complement to a survey. In addition, interviews can be a good basis that can be mentioned later in the outlook and can be a stimulus for further research. In summary, it can be said that the use of an interview as a research method can generate extensive insights that can make a major contribution to the state of knowledge in research, since not everyone can have access to high performing experts.

4.3.2 Development of Interview Protocol

While preparing the interview protocol, the focus was on the third objective of this research project. This objective reads as follows: “Explore ethical and societal meaning of AI-powered fake news detection from the perspective of the internet user”. The interview questions asked in the interview focus solely on this specific objective. To develop the right questions for the interview, literature research was conducted beforehand, which is written down in the literature section of this thesis. A semi-structured interview was chosen for the interview in this research paper. It is an interview in which questions that will be asked are written down in advance, but there are also questions allowed that occur to the interviewer during the interview and could also be relevant to the research. The reason for this is that the interviewees are very well versed in the field of AI and it can therefore be assumed that

further questions will arise during the interview, which can also be clarified by this type of interview. Since AI is a very complex topic, a semi-structured interview is a good option because it can provide an even deeper insight into the subject matter than a structured interview, since the person interviewing the expert can ask questions about certain experiences and other relevant questions that occur. Since the interviewee has a lot of knowledge, an interview in which unplanned questions are also asked maintains a certain flow of speech and can lead to more information than expected in advance. In addition, this type of interview allows the interviewee to continue and deepen their thoughts, which means that the data collected is of a higher quality. It also has advantages for the interviewer, as complex topics are better addressed and there is the opportunity to ask further questions if necessary.

4.3.3 Designing Interview Questions

The questions asked in the interview are open questions that allow the interviewee to be precise with their answer, but also allow for extensive answers. In addition, the interview should contain semi-structured questions, so that on the one hand the flow of speech is maintained, but on the other hand also allows the interviewer to go into more detail on certain points with the help of unplanned further questions. The aim of the interview is to find out what ethical and social significance the implementation of AI-controlled tools that can identify fake news has. The questions of the interview are all based on the objective and as well on the literature from this thesis. The interview contains eleven questions that are written down but as it has been said before, this is a semi-structured interview and the possibility lays with the interviewer to ask further questions if any occur.

4.3.4 Selection of Participants

For the section of the Interviews, it was decided to interview two people that work with AI, have a big knowledge about it and can therefore give qualitative answers. The first person is Christoph Schumacher, who is a professor at the University of Karlsruhe in Germany and a professor at Massey University in Auckland, New Zealand, in the field of innovation and

economics and is the director of the company Knowledge Exchange Hub. Schumacher was chosen because he programs AI technologies for the New Zealand government and is familiar with the subject. Precisely because he works for the government, he also must observe the development of AI in other countries and can therefore provide useful input on what experiences he has already had or what developments are still to come for social media users and if he can give any suggestions. He can also provide an assessment of possible dangers as well as positive developments. Schumacher was chosen because it was assumed that, due to his work as a professor at the University of Karlsruhe and the University of Auckland, he would be able to make an academic and student-friendly contribution and, due to his work with artificial intelligence, would be able to provide qualitative and high-quality data and provide deeper insights. Due to his work with AI, he can also provide technical insights if necessary and, above all, give an insight into how AI will or can develop further. His expertise also allows him to respond to additional questions that might be asked in a semi-structured interview. Precisely because Schumacher is a professor and works with AI, he can provide high-quality insights into both practical and theoretical topics. The survey, which was conducted as part of the mixed-methods approach, involved respondents from both Europe and the USA. Schumacher comes from Europe and lives partly in New Zealand. For this reason, he can also make a good comparison between the different nations. Another reason why it is interesting to interview an expert who works for a government is that they can explain real case studies, which makes it more understandable and insightful for the reader. These examples are helpful to compare with the literature section. It is also assumed that Christoph Schumacher is aware of the advantages and disadvantages of artificial intelligence and addresses them accordingly. Here, too, it is assumed that he will provide examples from real life. The second person who was interviewed is Jonny Simon Preuss. He works at the consulting company Horvath as a Senior Data Scientist and Consultant. Jonny Simon Preuss also has a strong background in both business and technical fields and is well educated when it comes to Artificial Intelligence. He was chosen because even though he has a strong background and works with artificial intelligence, he is almost 20 years younger than Christoph Schumacher which makes it interesting to see if the age difference has an influence on the opinion on artificial intelligence. Both participants followed the same Interview protocol and structure.

The choice for only two interviewees was made because two experts are sufficient to answer the objective. With these informative answers given by both experts, it is possible to easily compare their answers to each other and see if there are differences even between experts. In addition, the interview is conducted in connection with a survey, which is why it is sufficient to conduct only two interviews for this research project.



Figure 7: Picture of Christoph Schumacher, Source: <https://www.linkedin.com/feed/>



Figure 8: Picture of Jonny Preuss, Source: <https://www.linkedin.com/feed/>

4.3.5 Criteria for Participant Selection

The decision for the two participants in the Expert interviews had several reasons. Jonny Simon Preuss for example has a high expertise in Data Science and AI. Christoph Schumacher for example is a well-known expert in the AI world. Not only does he teach a course in Big Data Analytics at MIT as a professor, but he also has experience in working with AI. He developed the AI-driven tool called GDPLive, which processes large data sets for the New Zealand government to gain insights in real time. To emphasize GDPLive again: This is the world's first AI-powered real-time GDP measurement, with which Schumacher was able to prove his excellent expertise, that even the New Zealand government wants to work with the tool. It can therefore be assumed that he can provide good input regarding the

issue of battling fake news with the help of AI and give important insights that will help answer the research question. His academic background can also help with this work, as it lies in the fields of economics, engineering and international business. He can therefore demonstrate a differentiated understanding of how AI can be used sensibly to detect fake news on social media using AI. His work on the Shared Prosperity Index, which is an index driven by big data, only goes to show what a wealth of knowledge he has in the field of AI and big data. This index shows how society deals with AI, which can be traced back to the research question. Fake news is not just a technical problem, but also one that affects society in many ways. The fact that he works with AI on precisely this topic means that he can provide a good insight into the matter. In addition, Schumacher not only works with AI, but has already published several documents on AI, which he has spent a long time working on and researching. He has not only conducted research in the field of AI, but also in areas such as mathematical economics, machine learning and game theory. This means he can give very good input in an academic sense, as he is also a professor himself and has researched topics himself and knows what the expectations are and how best to answer interview questions. His advice to the health organization he is assisting with the government shows that he understands how the fight against fake news on social media should be tackled through the help of AI. Christoph Schumacher's expertise in areas such as AI, big data and other areas are very important for the research of this work and will be able to provide particularly good input.

Jonny Simoon Preuss has an exceptionally strong background when it comes to data science and business analytics, as he studied this both at bachelor's level and in greater depth at master's level. He has also been working in the field for several years. Due to his extensive knowledge in the field, he has above all the technical knowledge of how AI algorithms work and can provide good input in this regard. His knowledge can help to work out how artificial intelligence can be designed and optimized so that fake news on social media can be stopped more efficiently. He also has very good practical knowledge from creating data analytics practices and understands the complexity behind such tools. Preuss can therefore give practical examples of how AI could be implemented in the real world. Through his many years of working with AI, he can also provide input and real examples that can provide concrete evidence of practical insights that can support the arguments in this paper. Jonny Simon Preuss thus combines his technical and practical experience and his deep

understanding of how the AI tools work. Neither of the interviewees had to be acquired or researched, as both participants are people from the author's personal network.

4.3.6 Informed Consent

For both interviews, both interviewees signed an Informed Consent in advance. Informed consent is a document that is primarily used in research, as it ensures that the people interviewed for interviews, for example, are aware in advance that the statements may be used for the work and may even be published. However, a signature also ensures that the researcher can use the statements without any problems (Metro South Health, 2024).

4.4 Data Collection Method

The interviews that were conducted were all held via the video platform Zoom. During the interview, the interviewer recorded the interview, both in video form and in audio form. The reason for that is that the interviewer can better transcribe the interview and the interviewee's spoken words afterwards. A total of eleven questions were asked. Since it is a semi-structured interview and therefore further questions may arise during the interview, the number of eleven questions may vary. It was assumed that the interview would last around 30 minutes but could potentially take longer since the questions could extend in number.

The first question of the interview is about finding out from the interviewee what role they think AI plays in detecting fake news on social media. This question should help to understand the interviewee's perspective and attitude towards the topic. The next question is about the ethical aspects that should be considered when introducing such AI systems. Preserving ethical aspects is important when it comes to technologies such as AI, which is why it is informative to hear from an expert about their views on this and whether they are considering expansion options. The third question is about how transparent AI systems should be towards internet users. This question can be used to find out how important transparency is to the interviewee as an expert, but also how this transparency could be used to strengthen the trust of internet users. The next question relates to the advantages and disadvantages of using AI systems on society. This question is intended to provide the expert's assessment of how AI systems could affect society. The next question refers to

possible abuses that could arise from AI on social media. As experts are good at assessing such risks, the question is intended to provide information on which risks can be minimized. The sixth question is about how the privacy of internet users can best be protected when they use such AI systems. The expert's answer should provide information on what technical solutions are available for this and also show whether or not internet users can build trust in the AI tools. The seventh question equates the judgment of AI with human judgment and asks whether AI could replace human judgment. The interviewed expert should provide an insight into the extent to which AI can work independently and could replace something like human judgment. The next question is about an assessment of the extent to which people would access such systems at all. This question is intended to explore how user-friendly the systems are and which aspects could be improved. In line with this question, question 9 deals with how these AI tools could be made more accessible to the public. This allows the expert to provide suggestions for improvement and ideas for making the systems more receptive. The tenth question asks whether the trust of internet users can be influenced by AI. Here, the expert can give their assessment of the topic and, if necessary, make suggestions as to how trust could be strengthened. The last question is about whether users can also influence AI. Here, the advantages but also the dangers can be emphasized. As it was said before it is possible that new improvised questions arise from the already given questions.

4.4.1 Setting for Interviews

The first interview with Christoph Schumacher was conducted on 17.06.2024. The second interview was with Jonny Simon Preuss and was conducted on 16.07.2024. Since Schumacher and Preuss are acquaintance of the author of this thesis, they were both contacted via personal contact and therefore no canvassing was necessary. Before the interview was conducted, the interviewees were informed about the various conditions of the interview and had to sign a form (see appendix) in which they agreed that they had read and accepted all the conditions. These conditions include, among others, that the participant agrees to be part of the study and that they are conducting the study voluntarily. They are also free to decide whether they wish to be named or conduct the interview anonymously and are informed that the data will be used exclusively for the study. In addition, the participants can refuse to answer questions at any time during the interview and can also

insist after the interview that answers should not be used or even that the entire interview should not be used. The interview is conducted via the online video platform Microsoft Teams. Since Christoph Schumacher is in New Zealand it was made sure that the interview was held at an appropriate time due to the time difference. The interview and video are recorded via Microsoft Teams and transcribed during and after the interview using the tool Microsoft Teams Transcriber.

4.4.2 Conducting the Interviews

Both interviews were prepared in advance. This means that the researcher thought about questions in advance. The interviews were both conducted with the same questions to contrast the different and common statements of the two experts in the analysis. The interviews were not conducted in person. This was simply not possible since Christoph Schumacher lives in New Zealand and Jonny Simon Preuss lives in Paris. Therefore, both interviews were conducted via the online platform Teams Microsoft, in form of a video call. Both interviews lasted a little more than 30 minutes and both the moderator and the interview participant had their cameras on during the interview. Questions were discussed question by question, but since it is a semi-structured interview, it happened from time to time those further questions arose based on the answers, which were then also discussed.

4.4.3 Recording and Transcription

Both interviews were recorded. This was clearly communicated in advance by the moderator and took place with the prior consent of the interview participants. They were recorded both on video and with sound. On the one hand, these recordings were made using the Microsoft platform called Microsoft Team, which allows video calls to be recorded. On the other hand, the audio of both interviews was also recorded with a smartphone to ensure that there is at least one audio recording in case the platform Microsoft Teams fails. The Microsoft Teams platform also offers the option of transcribing the spoken word in the respective calls directly into a document. However, small transcription errors often occur, which is why it is important to also record the audio via the smartphone of the interviewer to be able to revise the document with the transcription.

4.5 Methodology Results

In the following part the results of both methodologies, the interviews and the survey are being evaluated for a better understanding of the results.

4.5.1 Data Analysis Survey

This chapter summarizes and explains the results of the study conducted for this thesis.

The survey was open to participants for a total of six weeks and 136 people took part in the survey. The first question, which asks about age, shows that 62% of respondents are between the age of 20 and 34. This means that this age group is predominant in this study. This shows that young adults are more familiar with social media and are more active on these platforms, which is why they are better able to participate in the survey than older generations. The other age groups show a lower percentage. Thus, 8% of the respondents are between 35 and 44 years old, 9% of the respondents are between 45 and 54 years old, another 14%, and thus the second largest age group in the survey, are people between 55 and 65 years old and the remaining 7% are between 65 and 80 years old. Even though most respondents are from the younger generation, the even distribution of the older generation shows that the topic is not only of considerable relevance to young people but also older generations show interest.

Another important aspect of the survey was to find out whether the participants came from Europe or the USA. Most respondents, 70%, were from Europe, with the remaining 30% coming from the US. These different locations could contribute to different results as the countries have different regulations and laws that regulate fake news. People coming from the US might have a different attitude towards fake news and artificial intelligence than people coming from Europe. The number of 70% Europeans and 30% Americans can be explained because the experts network is bigger in Europe which is why the survey reached more people in Europe.

There are also differences in the professional status of the participants. Most respondents (40%) stated that they work full-time. The second largest group (30%) stated that they were students. This shows that many respondents are currently still in education of some kind but does not mean that they are less familiar with the topic than respondents who are already working, as they constantly must deal with critical thinking and literature, including

literature on AI or fake news or social media in general, in their everyday learning life. 26% of respondents said they were self-employed and 12% said they worked part-time. This shows that people in different work positions nevertheless all have a certain level of concern about the topic of fake news. The smallest grouping at 4% are pensioners, suggesting that due to their older age they are not as familiar with social media as younger people, let alone have social media accounts at all. When asked how often respondents use social media, 85% said they use social media several times a day. This clear result shows the great relevance of social media in people's lives and how much it accompanies them in their everyday lives. 7% of respondents stated that they only use social media once a day and 8% stated that they only use social media every other day or less. Among other things, participants were asked to indicate what they mainly use social media for. There are several possible answers to this question. Interestingly, 85% of people said that they use social media to keep up to date with the latest news and information. According to the survey, 78% said they use social media for entertainment and 65% said they use social media to network. 49% of participants said they use social media to keep up with trends and 47% said they use social media to educate themselves. This shows that social media also has a more profound purpose for some participants and that they see it as a tool to constantly educate themselves. Only 27%, and therefore the smallest group, stated that they use social media to keep up to date politically. As politics is often discussed and discussed on social media, the other answer options were considered more often. This could be since, as mentioned above, the respondents in the entire study are mainly from the younger generation and use social media to switch off. This question allowed the participant to use more than just one options to answer.

In the sixth question, participants were asked to write down their definition of fake news. All 136 participants answered this question, but as many answers were repeated, here are the five most frequently mentioned answers. Most participants simply described the term fake news as "false information". However, others also defined the term as "political propaganda". Another definition that often came up was "mass media without credible sources" and "manipulative content". In addition, "panic-inducing false information". These various definitions show that people are aware of the negative impact fake news has on humanity. It can also be seen that the respondents have already dealt with the topic more often, as they were aware of how they define the term. The answers also show that there is great concern that fake news could have a strong influence on people's opinions and that

fake news even often has political influence or is being published because of politics. In the next question, participants were asked to indicate how often they think they come across fake news when they are on social media. There were five possible answers to this question, with the first option being "very frequent" and the fifth option being "not so much". The result was that 47% of participants believe that they encounter fake news frequently. Of these 47%, 15% stated that they even come across fake news very frequently and 32% stated that they come across fake news frequently. 31% of participants stated that they encounter fake news on social media moderately often, 17% stated that they encounter fake news less often and only 5% of respondents stated that they hardly ever see fake news on social media. However, the large figure of 47% shows that many people are aware of what fake news looks like and that they question the news they consume or might know how to identify fake news. The results also show that social media still offers too much scope for fake news and does not curb it often enough. The next question asked participants to say whether they could remember an instance in which they had come across fake news. 136 participants responded to this question, with many answers being repeated. Here are the five most common answers. Most participants remember seeing fake news in connection with Donald Trump during his election campaigns whether it has been in 2016 or in 2020. Another case that many participants mentioned was the corona pandemic, where it was often unclear what was true and false. Another common response was that many people cannot believe the news about wars, such as the current war between Russia and Ukraine or the war between Israel and Palestine, because too much false information is being spread. Many also criticized the amount of false information that is spread about sports, such as soccer, but also about sports influencers who spread false information to sell items like protein powders or other things that support ones sport journey. These responses show that in many areas, people are already finding it difficult to trust the information they see.

In the next question, participants were asked to indicate how effective they found the current methods to help stop fake news. There were five possible answers here, with 1 being "very effective" and 5 "not very effective". Most participants, 38%, said that they found the current methods less effective and 27% said that they did not find them effective at all. 28% of participants stated that they find the methods mediocre and a total of 7% stated that they find the methods effective to very effective. The results clearly show that the methods, be it the methods offered by the individual social media platforms or websites that offer fact checks,

need to be expanded. In the next question, participants were asked to indicate how they proceed if they suspect fake news. There were several possible answers to this question. Most respondents (86%) stated that they check the source that is spreading the news. This shows that credibility for the participants depends on how trustworthy the original source is. 53% of respondents stated that they often discuss suspicious news with friends to find out whether it is fake news or whether friends have other information about it. 51% of participants stated that they enter supposed fake news into search engines to see if other sources have also reported it. A slightly smaller group of 20% stated that they use fact-checking websites to check such news. This shows that there is still a need for education on these websites and that people could be made more aware of them. Only 3% stated that they don't act at all when they come across fake news. For this question it was again possible to choose more than one answer option. In question eleven, respondents were asked to give examples of efforts to identify fake news from social media platforms. 136 respondents answered this question. As many answers were duplicated, here are the five most frequently mentioned answers. The most common answer was "Twitter Community Notes". This is a Twitter function that allows users to exchange information with each other. Others said that they noticed that during the Covid pandemic, Instagram stories had a disclaimer when a story was about the Covid pandemic or the Covid vaccine. Many stated that they had already received warnings on platforms about content that was labeled as "potential fake news". Many also stated that they had already seen fake news being removed from social media platforms. All of this shows that transparent measures are in place, although there is still room for improvement. In question 12, participants were asked whether they had ever considered using AI systems to check fake news. This question was answered by 83% with "Yes" and 17% with "No". In the next question, participants were asked to state in their own words how effective they find the AI tools. For this question, participants were able to formulate their own answers but participants were also able to skip the question if they did not want or have an answer. That is why only 25 participants answered this question. Here are the five most common answers. The most common response was that these tools could still be improved and should be easier to use. Another common statement is that many do not find it effective because they see the AI tools as a machine and therefore not trustworthy, as many mistakes can still happen. Many also find these tools confusing rather than useful, which also shows that these systems still need to be improved. However, some have also

said that they believe these tools are effective, as they have often seen certain questionable content being removed from platforms. Others say that AI can only be as good as the data it has. This means that if AI has bad data or even false data, it can only give out false information, which is why it still has room for improvement. Question 14 asks how concerned respondents are about the presentation of fake news on social media. There were five possible answers here, where 1 means "very concerned" and 5 means "not concerned". Here, 42% of respondents stated that they are very concerned about the current situation with fake news. 33% stated that they are concerned. This means that a total of 75% of respondents are concerned or very concerned that there is so much false information online and easy to access. Furthermore, 13% stated that they are moderately concerned. Only 12% stated that they are hardly or not at all concerned about the fake news situation on social media. The next question asks what challenges the participants see with AI. For this question, respondents were asked to answer in their own words what 136 participants had done. Here are the five most common responses. One challenge that was frequently mentioned was that the AI itself can create fake news that another AI can no longer classify as fake news. In addition, the challenge was also expressed that the AI could also start labeling true information as false. It was also frequently mentioned that it must be borne in mind that AI is a technology and therefore technical errors can always occur. Another challenge that was frequently raised by respondents is that AI is created by people with their own opinions and therefore already has a certain bias towards certain topics or political directions. Some respondents also questioned whether the AI is already developed enough to recognize images that have been edited using Photoshop. The next question was designed to find out whether people would be more likely to believe information if it had been checked by an AI beforehand. There were five possible answers here, with 1 meaning "I would trust AI" and 5 meaning "I don't trust AI". It emerged that people are very divided on this. 39% of respondents stated that they would neither trust nor not trust AI. Surprisingly, only 8% said that they would tend to trust such information and 12% said that they would not trust the information despite fact-checking with AI. 27% stated that they would find the information slightly more trustworthy if they knew that the AI had looked over the information and 15% stated that they would trust the AI less. Question 17 is designed to find out whether respondents think that AI is responsible for the spread of fake news, among other things, or not. 60% of participants say that AI would curb fake news on social media and 40% of

participants say that they think AI is responsible for the spread of fake news. The next question asked respondents to indicate whether they have concerns about AI systems. This was a voluntary question in which 84 out of 136 participants responded. As the answers were often duplicated, here are the five most common responses. Many participants are concerned that AI will gain political influence in advance and thus adopt a certain position. Many also said that they were concerned that AI could get out of control if humans did not monitor its work. Other participants mentioned their concern for personal data. They are afraid that the AI could make a mistake and publish personal data. Participants also expressed concern that they did not know how reliable the tools really were in the end and that they would still have to carry out double checks. Another point that was mentioned more often is that the AI could also be tricked by the AI. Question 19 asks whether the participants think that more education about artificial intelligence and fake news would be useful. The answer to the question is clear. 88% of respondents agreed that it would be useful to consider this. Only 12% of respondents thought that it would make no difference, let alone not even be necessary. The next question asked respondents whether they thought that the presence of artificial intelligence was changing trust in information and in which direction trust was changing. For this question, the participants were able to formulate their own answer.

As the 136 answers were often repeated, here are the most frequently mentioned answers from the participants. Many said that their level of trust in information is lower due to AI and that they need to do more fact checking. In addition, it was often said that trust is decreasing because nowadays anyone with internet access can publish information or contribute to its dissemination. Many also responded that they no longer believe any source except sources that they already know and that are trustworthy. Many participants also say that they find the situation very worrying. Only a few said that they felt that it did not affect their trust at all. The last question was about how important the participants' felt transparency was. This question had a scale where 1 was "very important" and 5 was "not important". Here, 58% voted that it is very important to them that the AI tools show transparency. 24% of respondents stated that it is important to them. 14% of respondents said they were indifferent and only 5% said it was not important or not important at all to them whether the AI is transparent.

All in all, the survey provides a great understanding of the demographics of the participants, how they use social media and what they think of fake news. The data that was collected

makes it very clear that there is a lot of fake news on the internet and especially on social media and that a social media user simply must deal with it. It is also clear that the situation urgently needs to improve and that AI technologies can be expanded, and more education should be provided.

4.5.2 Results of Interviewee 1

The interview conducted with Christoph Schumacher is about whether artificial intelligence can help to recognize and curb fake news on social media and what challenges there are in doing so and the ethical actions of AI. The questions have already been described in detail above. Christoph Schumacher's statements offer a profound insight into the topic of AI in connection with fake news from an expert's perspective. In the following section, the most important topics and statements are written down in detail to provide an understanding of the topic and possible solutions. The discussion can be divided into central points and main topics.

The first main topic is the role and of AI in recognizing fake news. Schumacher says that he believes that in the case of AI, it depends on what data is being fed to the systems so it can recognize fake news. Generative AI systems such as ChatGPT, for example, can also make mistakes with such information and mark content as correct even though it is incorrect. Schumacher says that a central problem of AI is that it currently still has difficulty distinguishing between true and false, especially when a lot of false information already exists. He therefore recommends that the solution is for these AI systems to work more closely with search engine machines or tools to be able to check certain information and thus interpret it correctly as true or false by looking at other sources that contain the same kind of news. Many large AI tools are already doing this, but many still need to catch up. The next main issue raised in the expert interview are the ethical considerations and accountability. Schumacher says that with AI, it can never be 100% certain whether it is right or not. Since AI is a machine that learns and is trained by humans, it may be trained incorrectly or with bad intentions. One hopes that this is not the case, but you will always find people who have bad intentions and uses something powerful like AI for their own advantage. He compares the situation with that of an atom. On the one hand, atoms in nuclear power plants generate electricity and can thus supply entire cities. However, if these atoms

fall into the wrong hands, they can become a nuclear bomb that ultimately destroys everything good that has been achieved by it so far. In Schumacher's opinion, AI is no different. He therefore believes that users should always question information in advance. He also says that when people hear artificial intelligence, they automatically trust that it is an excellent and error-free tool because it has the word "intelligence" in it. He compares it to a doctor. A doctor often makes a diagnosis that is not necessarily correct, but people believe it because the diagnosis was made by an educated doctor. It's no different with AI. That is why the expert says that it is the responsibility of the internet user to double-check the information themselves and to not believe everything that the AI says. The next main topic that the interview can be divided into is education and awareness. Schumacher says that it is extremely important to educate people who use social media on how to recognize fake news. The big problem is that people have too much trust in AI and assume that it is telling the truth because it is seen as a reliable source. However, it is essential to scrutinize AI, as it is only a machine and can therefore make mistakes. Christoph Schumacher also believes that transparency is important and that it is important to offer this to the user. Even though he says that transparency about how AI works would not help much, as it is incomprehensible to someone who has nothing to do with AI professionally. However, he says that social media users would feel more comfortable if they were at least given the option to learn how the AI algorithms work and act. Another key issue is the potential misuse and dangers of AI. Christoph Schumacher says: "Anything that can do good can also do evil." By this he means that AI was created with the ulterior motive of helping humans and curbing the mistakes that humans make. The only problem with this is that AI is trained by humans, which means that mistakes can happen, or humans can train AI maliciously. Another topic is trust in AI and what influence there is from AI. Schumacher says that trust in AI is greater than assumed. Although people tend to be skeptical when asked, when it comes to verifying information, they place more trust in AI than previously assumed. Schumacher, who is a professor at universities in Karlsruhe and Auckland, among others, has already noticed that his students often use tools such as ChatGPT when writing an essay to make writing an essay easier. He notices that some students don't even check what data the tool gives them, but simply trust that the information is correct. He then notices this when correcting, as he is familiar with this area. However, caution is advised here. AI is a tool that is trained by humans. AI only works because it is fed data and information by humans.

However, this data can be incorrect, whether maliciously or not, and the AI can be trained incorrectly. The AI can then turn the data into stories that don't exist and present them as truth. The last topic addressed in the interview is the prospects and development of AI. Christoph Schumacher says that AI is something that will remain with us for a lifetime. It is here to stay. He also says that AI will help to continue to recognize fake news and that AI is getting better every time. However, this also means that the creation of fake news is getting better and better and the stories that generative AI can create are becoming more and more credible.

In conclusion, the statements of the interviewed expert Christoph Schumacher provide deep insights and his own opinion on what the role of AI is in identifying fake news. Here, it is essential that users and AI work together to avoid incorrect information and minimize risks.

4.5.3 Results of Interviewee 2

The interview conducted with Jonny Simon Preuss is about whether artificial intelligence can help to recognize and curb fake news on social media and what challenges there are in doing so and the ethical actions of AI. The questions have already been described in detail above. Jonny Simon Preuss' statements offer a profound insight into the topic of AI in connection with fake news from an expert's perspective. In the following section, the most important topics and statements are written down in detail to provide an understanding of the topic and possible solutions. The discussion can be divided into central points and main topics.

The first topic addressed in the interview was the role of AI in detecting fake news. Jonny Simon Preuss says that AI plays a crucial role in identifying and filtering fake news and that it is hard to imagine life without it. The advantage of AI is that it can quickly analyze large amounts of data and recognize patterns that indicate whether a post on social media is fake news or not. The fact that masses of fake news are published on social media these days means that the filtering process is accelerated by AI. Humans would not be able to do this at the same speed. However, Preuss mentions that humans and AI also work together. In certain cases, the AI marks the content that may contain false information and then passes it on to a human for final verification. He cited posts by famous people such as Donald Trump as an example of this. With people like him, it is better if the AI does not make the final decision as to whether the post is real or fake. In this case, it is better if a human makes this decision.

However, if it is a post by a private individual, and therefore does not have such great significance for the public sphere, it is often the AI that makes the final decision as to whether it is fake news or not. It is also important to remember that AI is constantly and rapidly evolving. Tools are constantly being developed to help improve the accuracy and increase the effectiveness of AI. Another topic that came up in the interview is the ethical consideration and the risk of AI bias. AI must be trained in such a way that it respects users' freedom of opinion. Particularly in Europe, care must be taken to ensure that this is respected. Preuss therefore says that it is important to ensure ethical implementation so as not to jeopardize the rights of internet users. One problem that Jonny Simon Preuss sees is that AI can be biased. This is because AI is trained by humans. He cites the example of the social media platform X and its owner Elon Musk. Theoretically, Elon Musk has the power to train the AI systems that are supposed to filter fake news on platform X to remove posts that do not correspond to his own opinion and thus the AI would be biased and would prevent the user from hearing other opinions. In short, the AI is only as good as the human who trains it and can therefore be biased. This is precisely why many internet users are still skeptical about AI, which is why it is important for people who train AI to work as transparently as possible. Another topic that was addressed in the interview is the transparency and the so-called "black box" that has already been briefly mentioned. For a better understanding of the AI for the social media user, it is important that the AI communicates transparently why it decides that something is fake news or not. So, if a social media user's post is blocked or even removed by the platform, it is helpful if the user is informed about the decision-making process and understands why the AI acted the way it did. One problem that could arise, however, is the black box that is imaginative. This is an algorithm that many AIs use, but which provides no explanation for its decision-making processes. This means that AI's that work with a so-called black box are somewhat less accurate than those that work without a black box. Although the tools without a black box may lose some accuracy, they also offer better transparency. Another topic that was addressed in the interview was the advantages and disadvantages of using AI. Jonny Simon Preuss has the opinion that the speed of AI is a great advantage for all areas in which it is used. The fact that social media is very fast-moving means that thousands of false pieces of information are spread around the world by many thousands of accounts every day, every minute. Without AI, humans would not be able to contain this anywhere nearly as quick as they can with AI. This is a huge advantage,

especially when it comes to topics such as election campaigns or highly topical issues. It can ensure that people are provided with the right information in the best possible way. Nevertheless, Preuss also points out that there are also disadvantages to AI. He says that AI is not perfect. As he has already mentioned, it may be biased and there is a risk that it will be trained “maliciously” or for the benefit of the trainer. It is also heavily dependent on its technology and technologies are known to have failures. Above all, the reason for the bias is a cause for concern, as it deprives people of their right to form their own opinions and also their right to freedom of expression. When asked about the misuse and dangers of AI, Jonny Simon Preuss replied that platforms could train their AI to ban unpopular opinions from the network under the pretext of labeling them as fake news. However, this would violate laws such as the GDPR (General Data Protection Regulation), which applies in Europe and is intended to ensure transparency, among other things. In addition, there is always the problem with AI that it cannot give a complete guarantee as to whether the decision was right or wrong. This can lead to the AI classifying a post that is rather true than false and vice versa. However, Preuss is not overly concerned about data protection. He says that if all framework conditions are adhered to during training and ideally, they are not trained on user data, there should be no concerns. These tools are trained to search through posts, whereby they pay attention to conspicuous sentences or keywords and less to who wrote the post. Various laws, such as the AI Act, ensure that people's privacy is protected, at least in Europe. There is also a talk of humans in the loop, which means that an AI will never work alone and there will always be a human in the background who looks over the AI's decisions. According to Jonny Simon Preuss, this is particularly important as AI will never be able to completely replace human judgment. Especially in critical and important cases, a human will always be consulted. Another question dealt with the acceptance of AI tools. Jonny Simon Preuss is of the opinion that the acceptance of such tools depends on the individual. People who deal a lot with political issues are aware that fake news exists and want to use AI tools to ensure that they are consuming truthful information. People who are aware of the problem of fake news are also receptive to such tools. According to Preuss, people who do not really engage with such tools could be picked up by educational campaigns. For example, platforms that use AI tools should educate their users about the benefits they bring. Tutorials and user-friendly design can also increase user acceptance. The problem that Jonny Simon Preuss currently sees with regard to usage is that people do not yet trust AI. If communication about

the decision-making processes is improved somewhat here, user trust can be increased and it is also important that users can always see that the AI is mainly making the right decisions. To summarize, the emphasis of the interview is that AI will never replace human assessment and human perspective on something and will therefore always work closely with humans but never without humans. Human acceptance of these tools can also be improved through more accurate results and more education in this regard.

4.5.4 Comparison of both Interviews

The interview was held with two experts who, due to their jobs, have deep insights into the topic of artificial intelligence. Nevertheless, among the many similar answers, there are also many differences, which are listed here. Both interview guests agree on the role of AI in detecting fake news. They both have the opinion that AI plays a major role. Jonny Simon Preuss emphasizes that AI uses large amounts of data to identify patterns, which speeds up the process of detecting fake news. Christoph Schumacher says that it is important for AI to work closely with search engine algorithms to quickly verify information or classify it as fake. The two also agree on ethical considerations. The topic of ethics still has room for improvement for both. Jonny Simon Preuss emphasizes that it is important that AI tools should be trained in such a way that they respect people's freedom of opinion and, above all, should not receive biased training data from their programmers. Christoph Schumacher also says that it is the responsibility of social media users to check the accuracy of the AI's decisions. Even an AI cannot always be right, which is why users should check information themselves using search engines. Like Preuss, he also warns that AI tools could lead to misuse if they are trained in a biased way. Both also agree that it is essential for the tools to focus on transparency. Preuss says that greater transparency would give users more trust in AI. Schumacher takes a similar view but says that it would not help much if the decision-making process and the technology and working methods of the AI were disclosed, as someone who is not familiar with technology or AI is unlikely to be able to understand these technical processes. However, he says that it wouldn't be wrong if users were at least given the option to see how it works so that they feel integrated. Both agree on the topic of misuse of AI. Both say that there is a danger with AI that people who train it or people who know how to use it can use it for abusive purposes. Using the example of Elon Musk and Platform

X, Jonny Simon Preuss points out that platform owners could ensure that the AI is trained in such a way that unpopular opinions are suppressed from the outset on their platform, which would result in people being denied their freedom of expression and the ability to form their own opinions. Christoph Schumacher points out that you must be careful, as anything that can do good can also do evil and that you have to expect that there will always be someone in humanity who misuses these tools or trains them for their own benefit. Both also point out that AI can never be 100% correct and will never work completely on its own and could never replace humans. Jonny Simon Preuss says that it is always important to let a human make the final decision, especially in important cases or when public figures make statements. Christoph Schumacher, on the other hand, makes a good point and claims that while AI is a tool that can certainly make an enormous contribution to the detection of fake news, it can also have the opposite effect and boost the spread of fake news by making the wrong choice.

Even though they gave many similar answers, there were also different answers to some questions. Preuss, for example, said that in the development of AI tools, attempts are now being made to train the AI tools so that they can recognize fake news without having to access databases. Schumacher, on the other hand, says that more attention should be paid to ensuring that the AI tools work more closely with search engines such as Google and others to use their databases and the various articles to check whether a critical article is true or false. There are also differing opinions when it comes to reviewing AI decisions. Jonny Simon Preuss says that there is the so-called “human in the loop” system, where an AI makes the decision, but a human always looks over this decision again and it is the human who then makes the final decision. However, Christoph Schumacher says that the end consumer always has their own responsibility. Social media users should always check for themselves whether the post they are reading is genuine or fake. The two also give different answers when it comes to the influence of users on AI. Jonny Simon Preuss sees the influence on the AI more from the perspective of feedback. He says that users could give feedback to tell the AI trainers where there is still room for improvement but does not go into whether every human could theoretically train the AI. Christoph Schumacher takes a different view, as he carried out an experiment in which he faked several articles about a fictitious Colombian jumping spider and then published them. When he did this quite often and very credibly, he asked ChatGPT if there really was a Colombian jumping spider and if ChatGPT could tell

him something about the spider, whereupon ChatGPT told him such a credible story that it would have been impossible to recognize it as fake. So, Schumacher says that it is perfectly possible for any human to train the AI and lure it onto a false trail.

4.5.5 Final Analysis and Reflection

The problem of fake news on social media exists and will not disappear in the near future. To address the original question of whether artificial intelligence has an impact on the fight against fake news on social media, a mixed methods approach was used, firstly by asking 136 normal social media users what they know and think about the topic and secondly by conducting two interviews with experts who are familiar with the topic of AI at expert level. The results of both studies are being compared below to get a better impression of whether the opinions of experts and those who are not experts differ greatly or even show similarities. There are clear opinions on the role of AI in detecting fake news. Experts are of the opinion that AI can become an advantage, as it has an overview of all the data that has ever been collected on the topic within seconds. Accordingly, AI can quickly recognize patterns and work faster than humans would. However, it is also emphasized that the AI must work together with search engines to verify the information as quickly as possible. The results of the survey on the topic of detecting fake news showed that 83% of respondents have already used or seen such AI tools to detect fake news, although opinions on the effectiveness of these tools are divided, as some still see potential for improvement. It is interesting that the opinions of the two AI experts do not differ, but normal social media users are still skeptical about AI tools and question their effectiveness. This leads to the conclusion that there should be more education for normal social media users and deeper insights should be provided about how AI works. The experts and social media users agree on the topic of ethical considerations and transparency. The experts emphasize how important ethical considerations are. Above all, however, care must be taken to ensure that freedom of opinion is respected on social media and that the AI tools are not biased or already trained with a particular opinion on certain topics. It is also emphasized that AI is a machine that can be flawed, which is why the experts stress that it is also the responsibility of the user to double-check facts despite AI tools. Although both experts also agree that there should be much greater transparency for users regarding how AI tools work, they do not agree on whether

this functionality is useful and understandable for every normal social media user, as it is often very complex. However, 58% of respondents in the survey stated that the transparency of AI systems and algorithms is very important to them to understand why certain decisions were made and to see how the AI works. In addition, 88% of respondents stated that they believe that more education about such AI tools or AI in general would positively change the attitude of users towards AI and that more people would start using these tools. Both the experts and the respondents in the survey agree on the challenges and dangers of AI, such as the misuse of AI. The experts warn of the dangers of AI, especially that the platforms that use AI tools can use them to suppress unpopular opinions and thereby make it appear that the opinions held by the platform owners are the right ones. The example given here is that Elon Musk, who bought the social media platform X, could use AI tools to present his own opinion as the correct one and thus influence thousands of users. So, the experts say that something that can do something good can also be misused by humans. Furthermore, both experts agree that AI can never be 100% correct and can also make mistakes, which is why it is important to never let it work without human supervision. In the survey, 42% of respondents also said that they were concerned that AI could make incorrect decisions or be manipulated by malicious humans and thus lose control. 40% of respondents also stated that they believe that AI could do more harm than good when detecting fake news on social media. This result shows that caution is required with AI and, above all, in its development to ensure that no serious mistakes are made. When it comes to the effectiveness and user behavior towards AI, both the experts and the respondents to the survey are not always in agreement. The survey found that 85% of respondents use social media several times a day and use it mainly for entertainment, news and information. The respondents say that they check the information they see either by using search engines or by checking the source of the article. On the one hand, the experts say that the final decision of the AI should never be made by the AI alone, but that a human should always look over it again, especially for important articles and topics. On the other hand, they emphasize that the user should assume that the AI could always be wrong and that it is therefore the user's responsibility to check the facts for themselves. Both experts also discuss the issue of whether social media users themselves can help train the AI. On the one hand, Jonny Simon Preuss believes that user feedback could be an important resource, while Christoph Schumacher says that every user could theoretically retrain the AI by putting in a great deal of effort. It can therefore be seen

from this that even the experts are not in complete agreement when it comes to the development of AI. When it comes to trust in AI and prospects, the statements are quite similar. Both experts agree that AI will never be able to completely replace human judgment. Jonny Simon Preuss sees the future of humans and AI working together, whereas Christoph Schumacher is rather skeptical that AI can ever be truly trustworthy. Even though Christoph Schumacher is skeptical, he still says that AI is here to stay and there needs to be a way to deal with AI without it getting out of control or being misused. The respondents to the survey also stated that only 8% would fully trust a message that had been verified by AI, whereas 27% of respondents said that they would be rather skeptical about verification by AI. This shows that 35% in total of the participants are not sure about whether they can trust artificial intelligence or not. It would be helpful here if the platforms provided more information or ran campaigns to show users which AI tools are available, which are trustworthy and how they work.

All in all, although social media users are not trained in the topic of AI, they often have similar answers to the experts. However, it is also clear to see that the experts discuss and argue on a completely different level, as they have more insights. The comparison between the both chosen methodologies show, that the topic of AI and whether it can stop fake news from being spread is such a new topic that even experts that have studied artificial intelligence and work with it don't have a clear answer on the topic.

4.5.6 Theoretical Assessment of Findings

In this chapter, the literature part and the literature that was used for this work is compared in detail with the methodology part and at the same time similar statements as well as different statements are highlighted and compared with each other. The comparison of the two parts of the work is important because it allows the validity of the statements and the relevance of the topic to be emphasized. By comparing one's own research results with the results that are already available, it is possible to recognize what should be researched further in the future and which topics are particularly relevant. For the comparison, the most important statements of the literature part and the methodology part are compared and equated with each other.

Both the literature results and the methodology results address the topic of the evolution of communication media. The literature points out that social media has become very important in the media world today. It has even reached the point where it is now the most widely used medium especially amongst the younger generation and it is assumed that it will overtake old media such as newspapers, television and radio, as social media is already used today to exchange information with friends and to receive news. Exactly this point is also taken up as a statement in the methodology part and it is emphasized that social media is the platform on which a lot of news is spread and, among other things, has also the largest platforms for facilitating the spread of fake news. The topic of social interaction is also emphasized in both parts. The literature part describes that social media is the reason why people can connect on a new level and thus build a network, which would have been unimaginable before the time of social media. This has its advantages as nowadays you can be friends with people you have never met, get to know cultures without visiting the country and much more. On the other hand, it also has its disadvantages, as people can access an enormous pool of information through social media, some of which is not verified and is therefore simply false. This topic is also addressed in the methodology part. Especially regarding the verification of information on social media. It discusses how social media can influence user opinions and what options there are for verifying information that you see on social media. It also discusses the fact that social media has changed people's communication behavior, as platforms such as Facebook, Instagram and others make it easy to hide behind a fake profile and then spread information, which is particularly easy since it is important to respect freedom of expression on social media platforms. Another topic that is addressed in both parts is the technologies for detecting fake news. The literature section mentions that something like machine learning and algorithms are already being developed to help filter out and recognize false information as quickly as possible. This includes not only fake news in text form, but also so-called deepfakes, such as fake videos and images. More and more efforts are being made to make these tools user-friendly so that users can quickly check for themselves whether the news they see is fake or real. The interviews in the methodology section also discuss how accessible these AI tools are for normal social media users. Here, the experts also talk about the fact that these tools are being developed or are even already in use, but that users do not know too much about fake news and AI and therefore find the tools too complicated or do not even know where to find them. This is also underlined by

the result of the survey, which states that although AI tools have already been used more frequently by users, users who have already used such tools agree that there is still room for improvement when it comes to its efficiency and how trustworthy they are for social media users. The psychological and social impact on people is also addressed in both parts. The examples of the 2016 US-elections and the Brexit are mentioned in the literature section. In the 2016-US elections in particular, wrong news emerged ongoing, and it was even revealed that Russia partially controlled election programs and election propaganda in the USA to help Donald Trump win. This election campaign was also the first time that Trump publicly used the term fake news in connection with misrepresented news. However, it was also through AI that a lot of fake news was transmitted at that time. The experts from the interviews in the methodology part also say that although AI can also be used to detect fake news and is already helping with this, they emphasize that it is particularly important that humans always look over the decisions made by AI. Especially when it comes to statements made by important public figures, such as Donald Trump. Normal social media users also like it when the decision-making processes of these AI tools are disclosed to everyone so that they can understand why a certain decision was made. Both sections also address concerns regarding ethics and transparency. The literature section addresses the fact that the use of algorithms and artificial intelligence is viewed critically, as they are always trained with a bias. This means that systems could be created that successively delete information that does not correspond to the opinion of the biased trained AI. There is a similar discussion in the methodology part. The experts say that an AI is still trained by humans, which means that humans can also train it in a malicious way or for their own benefit. For example, Elon Musk could ensure that only his opinion is allowed to be expressed on the social media platform X and all other opinions that are published are deleted by the systems and marked as false. In addition, normal social media users want more transparency to be able to understand the decision-making processes, although the experts say that it would be difficult for a “normal” person to understand these decision-making processes, as the explanations would be too technical and at the level of understanding would be too high for a normal social media user. The topic of transparency and user control is also mentioned in both parts. The literature section mainly talks about the need for more transparency in the processing of data. It is also mentioned that it is important to comply with laws such as the GDPR (General Data Protection Regulation), which primarily applies in Europe. Data protection is viewed

more strictly in Europe than in the USA, which is why the law was introduced in the first place. This law is also mentioned by the experts, who point out that it is particularly important to protect user data and to train the AI tools accordingly so that they do not exploit data and, in the worst case, pass it on to the general public or companies. The survey revealed that normal social media users want more transparency when it comes to how AI works. However, this raises the question of how comprehensible these processes are for normal consumers. They are often very technical and not understandable for someone who is not familiar with AI.

While the common statements that cover both the literature and the methods used in this thesis are interesting for the research of this thesis, even more interesting are the different statements that highlight which areas of the topic can be further researched. One topic where there are differences is trust in artificial intelligence. In the literature section, the risks to users from fake news on social media are identified. The literature states that there is a risk of AI being biased and influencing the user's opinion. This is also what the experts say, and both the literature and the experts advise that AI should never work without the critical eye of a human and that humans should always review the AI's decisions. Although this is currently the case and AI does not work alone, normal social media users are skeptical about AI and do not trust it. This statement contradicts both the literature and the experts, as they say that AI is a tool that is an enrichment with human supervision. There are also discrepancies when it comes to transparency. The literature states that the lack of transparency of AI is the reason why people do not trust AI tools. This statement is also underpinned by the survey that was conducted, which found that normal social media users would also like to see more transparency. However, the experts say that although they agree with this, and say it should be more transparent, they claim that there should be more education for normal social media users so they can really understand the decision of the AI properly, because if there is no education, users would not understand it, if not be more confused than they were before. Interestingly, opinions differ on the effectiveness of AI in detecting fake news. While the literature states that AI is indeed capable of detecting and curbing fake news on social media, the literature also states that consumer trust in AI technologies is rather low, even though these technologies are more likely to detect fake news than the normal social media user. The difficulty here is seen in the fact that many users do not even know what tools are available and how they can use them, and their basic

attitude towards AI tends to be skeptical. This result was also found in the survey conducted in the methodology section. Here, respondents stated that they were rather skeptical about the effectiveness of AI tools. The respondents stated that they believe that these AI tools still have potential for improvement. However, the experts' statements differ here from the statements in the literature and the survey. This is because the experts are convinced that AI will be an indispensable tool for detecting fake news on social media. The experts' only condition is that they must work closely with search engines to be able to check certain information quickly and, in the best case, a human employee will look over the results of the AI tools again. The topic of education and transparency is treated differently in the two sections. The literature section states that it is important to educate social media users at a young age about the existence of fake news and how to recognize it, and that young adults should not simply accept information because someone they know has posted it or because it is on the internet but should learn to question this information critically. When it comes to transparency, the literature mentions that so-called filter bubbles created by the algorithm can restrict freedom of expression. Accordingly, social media users should be warned and repeatedly that such filter bubbles exist so that users are aware that they should question information or google those information for fact checking. The methodology section revealed that both experts and social media users are of the opinion that more education and transparency regarding the functioning of AI are essential. However, there is no mention here of the fact that education on the topic should already take place at school, but rather that it should be seen more generally and that users should simply be made aware of it through campaigns and be trained at the same time. They also believe that if users were to learn how the AI tools work, they would have more confidence in the AI's decision-making process, because it would be easier for them to understand why something was flagged by the AI. When it comes to ethics, both the literature and the experts and normal social media users that have been surveyed agree. They all have the opinion that strict rules and laws should be introduced here, as otherwise freedom of expression would be violated. The dangers here are the filter bubbles, which can be used primarily by social media platform owners. These have control over the platforms and their algorithms and artificial intelligence and can therefore ensure that the opinions of normal social media users are influenced or even distorted. When it comes to the prospects, there are more common statements than different ones. Both the literature section and the methodology section state that AI is here to stay and

will be included in the future. However, both the literature and the methodology advise that the exchange and cooperation between AI and humans should always remain close so that AI does not get out of control or end up being responsible for spreading fake news itself. The literature as well as the experts and social media users surveyed also agree on the topic of data protection. It is important to ensure that users' data is protected. The experts say that they do not assume that the AI would give out data if a social media user uses an AI tool to find out whether messages are real or fake. However, the literature section says that there has already been one major scandal, namely the Cambridge Analytica scandal in 2018, where data was unlawfully collected from millions of Facebook users to be used for advertising campaigns. A scandal like this must not happen again and, accordingly, both the literature and experts are calling for data protection laws to be complied with and tightened.

In conclusion, it can be said that although different groups with different levels of knowledge regarding artificial intelligence and fake news on social media were interviewed in this work, the statements are very similar to those in the literature that was used and researched for this work. Although the statements in the literature tend to involve a detailed analysis of the historical development, as the literature section is intended to introduce the reader to the topic, and the methodology section is more concerned with the ethical challenges of AI on social media, the statements always boil down to the same thing, namely that AI is our future, but that caution is required, as it is still new and there will always be people who have bad intentions with tools that were actually created to prevent bad things from happening. The juxtaposition and the resulting findings illustrate how complex and new the whole topic of AI in social media is. Even if AI still has a long way to go, experts believe that it will be an enrichment in many parts of the lives of normal social media users.

Conclusion

At a time when news, both good and bad, can spread incredibly quickly through platforms such as Facebook, Instagram and the like, fake news has an extreme impact on our society, because this fake news can not only misinform people, but can also influence their opinions, such as opinions on politics or certain events such as the corona crisis. The various misinformation that society is constantly confronted with means that people no longer know

what is real and what is fake news, which is why trust in the media, in politics and, above all, in technology is being lost. In view of these challenges, the question arises as to how artificial intelligence can be used to prevent this problem in the best-case scenario and curb the spread of false information. The research carried out shows that AI can play an essential role in curbing fake news on social media. The AI tools can process large amounts of data quickly and recognize patterns that help to sort out fake news. These data patterns are not only applicable to texts, but the artificial intelligence can also be trained to recognize images, videos and sound. Of course, AI is also just a technology, and technologies can always make mistakes, be set up incorrectly or be trained incorrectly. There may also always be people who train these technologies incorrectly so that they can use them for their own purposes. Even if many people are still rather skeptical about the use of AI tools, it is a technology that is here to stay. If AI remains under human observation and collects a lot of data on how fake news can best be recognized, it is a helpful method for quickly curbing fake news on social media. While it won't eliminate fake news, it can control the amount of fake news and eliminate it faster than if humans were to view the fake news.

Limitation

Although extensive research was carried out in the thesis and results were achieved, there are also points that were not considered.

First and foremost, artificial intelligence is a new technology that is constantly evolving and being researched daily. Although progress is already being made with the technology and this work can only record the status of this technology, it must be borne in mind that the results of this study could look different in six months' time than they do at present. Furthermore, it is always possible for AI to make mistakes and unexpected situations could arise. Another point that was not considered in this paper relates to the survey that was conducted to answer the research question. Here, European and US citizens were asked about fake news and artificial intelligence to check whether there are different opinions based on origin. However, the survey was conducted without separating the responses of American and European citizens. Although participants were asked which country they came from, they were not directed to specific questions relating to their country of origin. This would

have provided a better overview of which answers Americans give on the topic and which answers Europeans give. Another limitation also relates to the survey conducted. In this survey, 136 people were interviewed. Since this thesis is based on the data from the survey, the results may not be decisive for the broader masses. 136 participants are already a lot and sufficient for the time the author worked on the thesis, but it is recommended to continue researching the same question, but to interview more people.

Outlook

In this work, it became clear that social media play a major role in how fake news is spread. They are the tool that makes it easy for anyone to spread fake news simply and often without consequences. When it comes to artificial intelligence, it turns out that it can do both good and bad things. On the one hand, AI is a system that can help to quickly identify fake news and thus mark it as such, or even remove it from the internet altogether. On the other hand, AI is still a very new system that can and will be used for malicious purposes. Precisely because AI is so new, several questions arise for the future. The most important question that arises is how AI can be further developed so that it becomes a reliable tool for identifying such false messages. Now, AI is still often flawed. In addition, the topic of AI is very fast-moving and will continue to be very innovative and dynamic. In the future, it will certainly be necessary to focus all research efforts on improving the algorithms that are able to recognize false news with a high probability and to know in advance whether it is true news. Another important point that should be considered in future research is education on social media. This helps the population and normal social media users to better recognize fake news even without AI. Platforms such as Instagram and Twitter could offer short, free workshops that quickly explain to users which features they should look out for when viewing news on social media. Because if the user already has a slight education, they will start to question more and not believe everything they read. In addition, an AI tool should also be programmed in such a way that it recognizes when a person is trying to maliciously train the tool for their own purposes. In future research, AI should also be trained in such a way that social media users can also contribute to better training the AI to recognize fake news.

All in all, the topic of artificial intelligence is very fast-moving and highly complex. It is a topic that will determine our future, which is why it is more important to conduct thorough research to improve every single area of AI.

Bibliography:

Aalbers, G., McNally, R. J., Heeren, A., De Wit, S., & Fried, E. I. (2019). Social media and depression symptoms: A network perspective. *Journal of Experimental Psychology: General*, 148(8), 1454.

Ahmar, A. S. (2023). View of changes in communication patterns in the digital age. <https://journal.arrus.id/index.php/soshum/article/view/1992/1297>

Ahmed, A. A. A., Aljabouh, A., Donepudi, P. K., & Choi, M. S. (2021). Detecting fake news using machine learning: A systematic literature review. arXiv preprint arXiv:2102.04458.

Ahuja, A. S. (2019). The impact of artificial intelligence in medicine on the future role of the physician. *PeerJ*, 7, e7702.

Akram, W., & Kumar, R. (2017). A study on positive and negative effects of social media on society. *International journal of computer sciences and engineering*, 5(10), 351-354.

Al-Garadi, M. A., Hussain, M. R., Khan, N., Murtaza, G., Nweke, H. F., Ali, I., ... & Gani, A. (2019). Predicting cyberbullying on social media in the big data era using machine learning algorithms: review of literature and open challenges. *IEEE Access*, 7, 70701-70718.

Allam, Z., & Dhunny, Z. A. (2019). On big data, artificial intelligence, and smart cities. *Cities*, 89, 80-91.

Al-Menayes, J. J. (2015). Motivations for using social media: An exploratory factor analysis. *International Journal of Psychological Studies*, 7(1), 43.

Aldwairi, M., & Alwahedi, A. (2018). Detecting fake news in social media networks. *Procedia Computer Science*, 141, 215-222.

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of economic perspectives*, 31(2), 211-236.

Altay, S., Hacquin, A. S., & Mercier, H. (2022). Why do so few people share fake news? It hurts their reputation. *new media & society*, 24(6), 1303-1324.

Al-Rawi, A. (2019). Viral news on social media. *Digital journalism*, 7(1), 63-79.

Amedie, J. (2015). The impact of social media on society.

Aoki, N. (2020). An experimental study of public trust in AI chatbots in the public sector. *Government Information Quarterly*, 37(4), 101490.

Arcas, B. a. Y. (2023, October 10). Artificial general intelligence is already here. NOEMA. <https://www.noemamag.com/artificial-general-intelligence-is-already-here/>

Bala, K. (2014). Social media and changing communication patterns. *Global Media Journal: Indian Edition*, 5(1).

Balakrishnan, V., Ng, W. Z., Soo, M. C., Han, G. J., & Lee, C. J. (2022). Infodemic and fake news—A comprehensive overview of its global magnitude during the COVID-19 pandemic in 2021: A scoping review. *International Journal of Disaster Risk Reduction*, 78, 103144.

Batailler, C., Brannon, S. M., Teas, P. E., & Gawronski, B. (2022). A Signal Detection Approach to Understanding the Identification of Fake News. *Perspectives on Psychological Science*, 17(1), 78-98.

BBC News. (2013). AP Twitter account hacked in fake “White House blasts” post. BBC News. <https://www.bbc.com/news/world-us-canada-21508660>

Bergström, A. & Belfrage, M. J. (2018). News in social media. *Digital Journalism*, 6(5), 583–598. <https://doi.org/10.1080/21670811.2018.1423625>

Berrondo-Otermin, M., & Sarasa-Cabezuelo, A. (2023). Application of Artificial Intelligence Techniques to Detect Fake News: A Review. *Electronics*, 12(24), 5041.

Bik, H. M., & Goldstein, M. C. (2013). An introduction to social media for scientists. *PLoS biology*, 11(4), e1001535.

Bildung, B. F. P. (2023, 26. Oktober). Fake news, misinformation, desinformation. bpb.de. <https://www.bpb.de/shop/zeitschriften/izpb/medienkompetenz-355/539986/fake-news-misinformation-desinformation/>

Blaßnig, M. (2020). „In Zeiten der Infodemie. Staatliche Desinformation in der Corona-Krise“. *juridikum*, (4), 433-442.

Bodell, L. (2023, December 26). How narrow AI will disrupt every industry. *Forbes*. <https://www.forbes.com/sites/lisabodell/2023/12/26/how-narrow-ai-will-disrupt-every-industry/?sh=1b45941d7ec1>

Bok, S., Martin, D. E., & Lee, M. (2021). Validation of the COVID-19 Disbelief Scale: Conditional indirect effects of religiosity and COVID-19 fear on intent to vaccinate. *Acta Psychologica*, 219, 103382.

BOTHMANN, L. KÜNSTLICHE INTELLI-GENZ IN DER STRAF-VERFOLGUNG. CYBERKRIMINALITÄT, 55.

Bovet, A., & Makse, H. A. (2019). Influence of fake news in Twitter during the 2016 US presidential election. *Nature communications*, 10(1), 7.

Burkhardt, J. M. (2017). Combating fake news in the digital age (Vol. 53, No. 8, pp. 5-9). Chicago, IL, USA: American Library Association.

Bundy, A. (2017). Preparing for the future of artificial intelligence.

Buttazzo, G. (2023). Rise of artificial general intelligence: risks and opportunities. *Frontiers in Artificial Intelligence*, 6, 1226990.

Bölen, H. (2022). Einfluss von sozialen Medien auf junge Erwachsene und ihr Konsumverhalten: eine Mixed-Methods Studie.

Capatina, A., Kachour, M., Lichy, J., Micu, A., Micu, A. E., & Codignola, F. (2020). Matching the future capabilities of an artificial intelligence-based software for social media marketing with potential users' expectations. *Technological Forecasting and Social Change*, 151, 119794.

Cath, C. (2018). Governing artificial intelligence: ethical, legal and technical opportunities and challenges. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 376(2133), 20180080.

Chauhan, R. S., Connelly, S., Howe, D. C., Soderberg, A. T., & Crisostomo, M. (2022). The danger of "fake news": How using social media for information dissemination can inhibit the ethical decision-making process. *Ethics & Behavior*, 32(4), 287-306.

Cheng, J. Y., Abel, J. T., Balis, U. G., McClintock, D. S., & Pantanowitz, L. (2021). Challenges in the development, deployment, and regulation of artificial intelligence in anatomic pathology. *The American Journal of Pathology*, 191(10), 1684-1692.

Choi, M., & Toma, C. L. (2014). Social sharing through interpersonal media: Patterns and effects on emotional well-being. *Computers in Human Behavior*, 36, 530-541.

Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrociocchi, W., & Starnini, M. (2021). The echo chamber effect on social media. *Proceedings of the National Academy of Sciences*, 118(9), e2023301118.

Craig, W., Boniel-Nissim, M., King, N., Walsh, S. D., Boer, M., Donnelly, P. D., ... & Pickett, W. (2020). Social media use and cyber-bullying: A cross-national analysis of young people in 42 countries. *Journal of Adolescent Health*, 66(6), S100-S108.

Der Mensch und die Dinge – „Zeroth Law – Das nullte Gesetz“ von Gamut Inc in der Tischlerei der Deutschen Oper Berlin. (n.d.). *Nmz - Neue Musikzeitung*. <https://www.nmz.de/kritik/oper-konzert/der-mensch-und-die-dinge-zeroth-law-das-nullte-gesetz-von-gamut-inc-der>.

Dewing, M. (2010). *Social media: An introduction* (Vol. 1). Ottawa: Library of Parliament.

DeVerna, M. R., Yan, H. Y., Yang, K. C., & Menczer, F. (2023). Artificial intelligence is ineffective and potentially harmful for fact checking. *arXiv preprint arXiv:2308.10800*.

Devang, V., Chintan, S., Gunjan, T., & Krupa, R. (2019). Applications of artificial intelligence in marketing. *Annals of Dunarea de Jos University of Galati. Fascicle I. Economics and Applied Informatics*, 25(1), 28-36.

De Vicente Domínguez, A. M., Beriain Bañares, A., & Sierra Sánchez, J. (2021). Young Spanish Adults and Disinformation: Do they identify and spread fake news and are they literate in it?. *Publications*, 9(1), 2

Dhar, V. (2016). The future of artificial intelligence. *Big Data*, 4(1), 5-9.

Di Placido, D. (2023, March 27). Why did 'Balenciaga Pope' go viral? *Forbes*. <https://www.forbes.com/sites/danidiplacido/2023/03/27/why-did-balenciaga-pope-go-viral/?sh=708846e84972>

Dornan, C. (2017). *Dezinformatsiya: The past, present and future of fake news. Series of reflection papers, Canadian Commission for UNESCO.*

ENDRE, M. Fake-News, Verschwörungsgedanken und po-litischer Extremismus–Vertrauen in einer sich fraktionierenden Gesellschaft. *ertrauen–ein risk*, 31.

Etta, G., Cinelli, M., Galeazzi, A., Valensise, C. M., Quattrociocchi, W., & Conti, M. (2022). Comparing the impact of social media regulations on news consumption. *IEEE Transactions on Computational Social Systems*.

Farook, F. S., & Abeysekara, N. (2016). Influence of social media marketing on customer engagement. *International Journal of Business and Management Invention*, 5(12), 115-125.

Flaxman, S., Goel, S., & Rao, J. M. (2013). Ideological segregation and the effects of social media on news consumption. Available at SSRN, 2363701.

Gruzd, A., & Hernández-García, Á. (2018). Privacy concerns and self-disclosure in private and public uses of social media. *Cyberpsychology, Behavior, and Social Networking*, 21(7), 418-428.

Forcier, M. B., Gallois, H., Mullan, S., & Joly, Y. (2019). Integrating artificial intelligence into health care through data access: can the GDPR act as a beacon for policymakers? *Journal of Law and the Biosciences*, 6(1), 317-335.

Gherheș, V. (2018). Artificial Intelligence: Perception, expectations, hopes and benefits. *Romanian Journal of Human-Computer Interaction*, 11(3), 219-230.

Gelfert, A. (2018). Fake news: A definition. *Informal logic*, 38(1), 84-117.

Gradoń, K. T., Hołyst, J. A., Moy, W. R., Sienkiewicz, J., & Suchecki, K. (2021). Countering misinformation: A multidisciplinary approach. *Big Data & Society*, 8(1), 20539517211013848.

Guess, A., Nyhan, B., & Reifler, J. (2018). Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 US presidential campaign.

Gundecha, P. & Liu, H. (2012). Mining social media: A Brief Introduction. In *New Directions in Informatics, Optimization, Logistics, and Production* (S. 1–17). <https://doi.org/10.1287/educ.1120.0105>

Guo, Z., Schlichtkrull, M., & Vlachos, A. (2022). A survey on automated fact-checking. *Transactions of the Association for Computational Linguistics*, 10, 178-206.

Gupta, A., Hathwar, D., & Vijayakumar, A. (2020). Introduction to AI chatbots. *International Journal of Engineering Research and Technology*, 9(7), 255-258.

Hadžialić, S., & Phuong, V. T. (2020). Media ethics within the fake news challenges during the Covid-19 pandemic. *Studia i Analizy Nauk o Polityce*, (2), 33-46.

Hennell, K., Limmer, M., & Piacentini, M. (2020). Ethical dilemmas using social media in qualitative social research: A case study of online participant observation. *Sociological research online*, 25(3), 473-489.

Hernon, P. (1995). Disinformation and misinformation through the internet: Findings of an exploratory study. *Government information quarterly*, 12(2), 133-139.

Hoffman, D. L., & Novak, T. (2012). Why do people use social media? Empirical findings and a new theoretical framework for social media goal pursuit. *Empirical Findings and a New Theoretical Framework for Social Media Goal Pursuit* (January 17, 2012).

Jaspal, R. & Nerlich, B. (2022). Social representations of COVID-19 Skeptics: denigration, demonization, and disenfranchisement. *Politics, Groups, and Identities*, 11(4), 750–770. <https://doi.org/10.1080/21565503.2022.2041443>

José, V. D. (2013, 12. August). Understanding social media logic. <https://deliverypdf.ssrn.com/delivery.php?ID=854006004112116116071028021076111069053032069011093069025100071025001116086107065006030110036003123013013080100092008118114108028051002032011066082086025064011081068040012032016069070114102067098115114093024094118003084097113006103082019007085079084067&EXT=pdf&INDEX=TRUE>

Jones, N., Borgman, R., & Ulusoy, E. (2015). Impact of social media on small businesses. *Journal of small business and enterprise development*, 22(4), 611-632.

Jozani, M., Ayaburi, E., Ko, M., & Choo, K. K. R. (2020). Privacy concerns and benefits of engagement with social media-enabled apps: A privacy calculus perspective. *Computers in Human Behavior*, 107, 106260.

Keleş, B. Y., McCrae, N. & Grealish, A. (2019). A Systematic review: The Influence of social media on Depression, anxiety, and Psychological Distress in Adolescents. *International journal of adolescence and youth*, 25(1), 79–93. <https://doi.org/10.1080/02673843.2019.1590851>

Kong, Q / Mao, W. / Chen, G. / Zeng, D. (2020) Exploring trends and patterns of popularity stage evolution in social media. (2020, 1. Oktober). *IEEE-Journals & Magazine | IEEE Xplore*. <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8428538>

Koos, S. (2021). Die „Querdenker“. Wer nimmt an Corona-Protesten teil und warum?: Ergebnisse einer Befragung während der „Corona-Proteste“ am 4.10. 2020 in Konstanz.

Machete, P., & Turpin, M. (2020). The use of critical thinking to identify fake news: A systematic literature review. In *Responsible Design, Implementation and Use of Information and Communication Technology: 19th IFIP WG 6.11 Conference on e-Business, e-Services, and e-Society, I3E 2020, Skukuza, South Africa, April 6–8, 2020, Proceedings, Part II* 19 (pp. 235-246). Springer International Publishing.

Kudchadkar, S. R. & Carroll, C. L. (2020). Using social media for rapid information dissemination in a pandemic: #PEDSICU and Coronavirus Disease 2019. *Pediatric Critical Care Medicine*, 21(8), e538–e546. <https://doi.org/10.1097/pcc.0000000000002474>

Kosinski, M., & Kosinski, M. (2024, February 23). How to implement the General Data Protection Regulation (GDPR). IBM Blog. <https://www.ibm.com/blog/general-data-protection-regulation-implementation/>

Krettek, C. (2023). ChatGPT. *Die Unfallchirurgie*, 126(3), 252-254.

Kreye, A. (2018, July 18). Angela Merkel plaudert mit der Roboter-Puppe Sophia. *Süddeutsche.de*. <https://www.sueddeutsche.de/digital/kuenstliche-intelligenz-merkel-plaudert-mit-der-roboterpuppe-und-das-ist-ein-problem-1.4033933>

Lavin, M. (2006). Cookies: What do consumers know and what can they learn?. *Journal of Targeting, Measurement and Analysis for Marketing*, 14, 279-288.

Lee, C. S., Ma, L., & Goh, D. H. L. (2011). Why do people share news in social media? In *Active Media Technology: 7th International Conference, AMT 2011, Lanzhou, China, September 7-9, 2011. Proceedings 7* (pp. 129-140). Springer Berlin Heidelberg.

Lee, J. K. (2020). The effects of social comparison orientation on psychological well-being in social networking sites: Serial mediation of perceived social support and self-esteem. *Current Psychology*, 1-13.

Lenzen, M. (2023). *Künstliche Intelligenz: Was sie kann & was uns erwartet* (Vol. 6302). CH Beck.

Linke, A., & Zerfass, A. (2012). Future trends in social media use for strategic organisation communication: Results of a Delphi study. *Public Communication Review*, 2(2).

Loewy, 1992

Loewy, A. H. (1992). Freedom of Speech as a Product of Democracy. *U. Rich. L. Rev.*, 27, 427.

Luo, M., & Hancock, J. T. (2020). Self-disclosure and social media: motivations, mechanisms, and psychological well-being. *Current opinion in psychology*, 31, 110-115.

Loureiro, S. M. C., Guerreiro, J., & Tussyadiah, I. (2021). Artificial intelligence in business: State of the art and future research agenda. *Journal of business research*, 129, 911-926.

MARADUN, L. U. Digital media: The danger and consequence of fake news in democratic societies.

Matt. (2024, February 6). What GPT stands for and what is ChatGPT? AutoGPT Official. <https://autogpt.net/what-gpt-stands-for-and-what-is-chatgpt/>

McKay, S., & Tenove, C. (2021). Disinformation as a threat to deliberative democracy. *Political Research Quarterly*, 74(3), 703-717.

McIntyre K. (2014), View of the evolution of social media from 1969 to 2013: a change in competition and a trend toward complementary, niche sites. (o. D.). <https://thejsms.org/index.php/JSMS/article/view/89/43>

Metro South Health. (2024, July 2). Informed consent. <https://www.metrosouth.health.qld.gov.au/patients-and-visitors/going-to-hospital/informed-consent>

Michaelis, M., Jafarian, J. H., & Biswas, A. (2022, June). The Dangers of Money and Corporate Power Relating to Online Disinformation. In 2022 23rd IEEE International Conference on Mobile Data Management (MDM) (pp. 470-475). IEEE.

Micu, A., Capatina, A., & Micu, A. E. (2018). Exploring artificial intelligence techniques' applicability in social media marketing. *Journal of Emerging Trends in Marketing and Management*, 1(1), 156-165.

Mijwel, M. M. (2015). History of Artificial Intelligence Yapay Zekânın T arihi. no. April, 2018.

Miller, D., Sinanan, J., Wang, X., McDonald, T., Haynes, N., Costa, E., ... & Nicolescu, R. (2016). How the world changed social media (p. 286). UCL press.

Moorman, J., & Bowker, A. (2011). The university Facebook experience: The role of social networking on the quality of interpersonal relationships. *The American Association of Behavioral and Social Sciences Journal*, 15(1), 1-23.

Murthy, D. (2018). Introduction to social media, activism, and organizations. *Social media and society*, 4(1), 205630511775071. <https://doi.org/10.1177/2056305117750716>

Mutawa, M. A., & Rashid, H. (2020, August). Comprehensive review on the challenges that impact artificial intelligence applications in the public sector. In *Proceedings of the 5th NA International Conference on Industrial Engineering and Operations Management (Vol. 10)*.

Müssig, C. (2021). Künstliche Intelligenz für Lernende und Lehrende. *Pflegezeitschrift*, 74(6), 36-40.

Naeem, S. B., Bhatti, R., & Khan, A. (2021). An exploration of how fake news is taking over social media and putting public health at risk. *Health Information & Libraries Journal*, 38(2), 143-149.

Nath, M. P., Pandey, P., Somu, K., & Amalraj, P. (2018). Artificial intelligence & machine learning: the emerging milestones in software development. *IJRSI*, 5(IX), 36-44.

Nieminen, S., & Rapeli, L. (2019). Fighting misperceptions and doubting journalists' objectivity: A review of fact-checking literature. *Political studies review*, 17(3), 296-309.

Orszag, P. (2023, October 16th). Do not underestimate the ‘mega-Brussels effect’ of EU-US co-ordination. Financial Times. <https://www.ft.com/content/9173e128-b702-43cc-8c63-552a1dad4ac3>

Osatuyi, B. (2013). Information sharing on social media sites. *Computers in Human Behavior*, 29(6), 2622-2631.

Ostic, D., Qalati, S. A., Barbosa, B., Shah, S. M. M., Galvan Vela, E., Herzallah, A. M., & Liu, F. (2021). Effects of social media use on psychological well-being: a mediated model. *Frontiers in Psychology*, 12, 678766.

Paaß, G., & Hecker, D. (2020). *Künstliche Intelligenz: was steckt hinter der Technologie der Zukunft?* Springer Vieweg.

Pennington, N. (2021). Communication outside of the home through social media during COVID-19. *Computers in human behavior reports*, 4, 100118.

Pérez-Rosas, V., Kleinberg, B., Lefevre, A., & Mihalcea, R. (2017). Automatic detection of fake news. arXiv preprint arXiv:1708.07104.

Primack, B. A., Shensa, A., Escobar-Viera, C. G., Barrett, E. L., Sidani, J. E., Colditz, J. B., & James, A. E. (2017). Use of multiple social media platforms and symptoms of depression and anxiety: A nationally representative study among US young adults. *Computers in human behavior*, 69, 1-9.

Research Subject Guides: Fake News/Misinformation/Disinformation: What is Fake News? (n.d.). <https://subjectguides.lib.neu.edu/fakenews>

Rigner, A. (2019). AI-based machine vision for retail self-checkout system. Master's Theses in Mathematical Sciences.

Rocha, Y. M., de Moura, G. A., Desidério, G. A., de Oliveira, C. H., Lourenço, F. D., & de Figueiredo Nicolete, L. D. (2021). The impact of fake news on social media and its influence on health during the COVID-19 pandemic: A systematic review. *Journal of Public Health*, 1-10.

Rossing, J. (2023). The disinformation landscape of the future: a qualitative perspective (Master's thesis, University of Twente).

Sadiku, M. N., Ashaolu, T. J., Ajayi-Majebi, A., & Musa, S. M. (2021). Artificial intelligence in social media. *International Journal of Scientific Advances*, 2(1), 15-20.

Sajithra, K. / Patil, R. (2013). Social media–history and components. *Journal of Business and Management*, 7(1), 69-74.

Salman, A. (2021). Media dependency, interpersonal communication, and panic during the COVID-19 Movement Control Order. *SEARCH Journal of Media and Communication Research*, 13(1), 79-91.

Santuraki, S. U. (2019). Trends in the Regulation of Hate Speech and Fake News: A Threat to Free Speech?. *Hasanuddin Law Review*, 5(2), 140-158.

Saravanakumar, M., & SuganthaLakshmi, T. (2012). Social media marketing. *Life science journal*, 9(4), 4444-4451.

Schivinski, B., Brzozowska-Woś, M., Stansbury, E., Satel, J., Montag, C., & Pontes, H. M. (2020). Exploring the role of social media use motives, psychological well-being, self-esteem, and affect in problematic social media use. *Frontiers in Psychology*, 11, 617140.

Schmidt, A. (2022). *Schönheitsideale in sozialen Medien-Eine empirische Untersuchung zum Einfluss von Körperidealen und der Body Positivity Bewegung auf Studierende der Hochschule Meißen* (Doctoral dissertation).

Sheldon, P., Rauschnabel, P., & Honeycutt, J. M. (2019). *The dark side of social media: psychological, managerial, and societal perspectives*. Academic Press.

Siegmann, C., & Anderljung, M. (2022). The Brussels effect and artificial intelligence: How EU regulation will impact the global AI market. *arXiv preprint arXiv:2208.12645*.

Simion, M. (2023). Knowledge and disinformation. *Episteme*, 1-12.

Singh, M., & Singh, G. (2018). Impact of social media on e-commerce. *International Journal of Engineering & Technology*, 7(2.30), 21-26.

Social Media Overview - Communications. (2024, June 24). Communications. <https://communications.tufts.edu/marketing-and-branding/social-media-overview/>

Stahl, B. C., Antoniou, J., Ryan, M., Macnish, K., & Jiya, T. (2022). Organisational responses to the ethical issues of artificial intelligence. *AI & SOCIETY*, 37(1), 23-37.

Subramanian, K. R. (2017). Influence of social media in interpersonal communication. *International Journal of Scientific Progress and Research*, 38(2), 70-75.

Syarifuddin, T. I., Budiman & Purwaningsih, T. (2021). Regional Head Communication patterns on social media in handling the COVID-19 pandemic. *IOP conference series*, 717(1), 012020. <https://doi.org/10.1088/1755-1315/717/1/012020>

Tai, M. C. T. (2020). The impact of artificial intelligence on human society and bioethics. *Tzu-Chi Medical Journal*, 32(4), 339.

Taprial, V., & Kanwar, P. (2012). *Understanding social media*. Bookboon.

TED (2011). Beware online “filter bubbles” | Eli Pariser. On Youtube. <https://www.youtube.com/watch?v=B8ofWFx525s>.

Tenove, C. (2020). Protecting democracy from disinformation: Normative threats and policy responses. *The International Journal of Press/Politics*, 25(3), 517-537.

Trump suggests “injection” of disinfectant to beat coronavirus and “clean” the lungs. (2020, April 24). NBC News. <https://www.nbcnews.com/politics/donald-trump/trump-suggests-injection-disinfectant-beat-coronavirus-clean-lungs-n1191216>

Tsimonis, G., & Dimitriadis, S. (2014). Brand strategies in social media. *Marketing Intelligence & Planning*, 32(3), 328-344.

Ullah, I., Boreli, R., & Kanhere, S. S. (2020). Privacy in targeted advertising: A survey. arXiv preprint arXiv:2009.06861.

Urbach, N., & Ahlemann, F. (2016). Der Wissensarbeitsplatz der Zukunft: Trends, Herausforderungen und Implikationen für das strategische IT-Management. *HMD-Praxis der Wirtschaftsinformatik*, 53, 16-28.

Valerio, C., William, L., & Noémier, Q. (2019). The impact of social media on E-Commerce decision making process. *International Journal of Technology for Business (IJTB)*, 1(1), 1-9.

Van Djick, J. (2013) *The culture of connectivity*. (o. D.). Google Books. https://books.google.fr/books?hl=de&lr=&id=t5RpAgAAQBAJ&oi=fnd&pg=PPI&dq=the+beginning+of+social+media+&ots=pn5AYTsqJI&sig=Mhwkz138jznZ_-VWnhCWhzCDDyw&redir_esc=y#v=onepage&q=the%20beginning%20of%20social%20media&f=false

van Der Linden, S., Roozenbeek, J., & Compton, J. (2020). Inoculating against fake news about COVID-19. *Frontiers in psychology*, 2928.

Velasco, C. (2022, May). Cybercrime and Artificial Intelligence. An overview of the work of international organizations on criminal justice and the international applicable instruments. In *ERA-Forum* (Vol. 23, No. 1, pp. 109-126). Berlin/Heidelberg: Springer Berlin Heidelberg.

von Boehn, H. V, n. d. Fake statt Fakt–wie Fake News uns beeinflussen.

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *science*, 359(6380), 1146-1151.

What is artificial intelligence (AI)? | IBM. (n.d.). <https://www.ibm.com/topics/artificial-intelligence>

Whiting, A., & Williams, D. (2013). Why people use social media: a uses and gratifications approach. *Qualitative market research: an international journal*, 16(4), 362-369.

Wittpahl, V. (2019). *Künstliche Intelligenz: Technologien| Anwendung| Gesellschaft* (p. 270). Springer Nature.

Zhao, M. (2019). Social media sharing and online news consumption. Social Science Research Network. <https://doi.org/10.2139/ssrn.3328864>

Zhu, H., Wu, H., Cao, J., Fu, G., & Li, H. (2018). Information dissemination model for social media with constant updates. *Physica A: Statistical Mechanics and its Applications*, 502, 469-482.

Zhou, Z. H. (2016). Learnware: on the future of machine learning. *Frontiers Comput. Sci.*, 10(4), 589-590.

Zimmer, F., Scheibe, K., Stock, M., & Stock, W. G. (2019). Fake news in social media: Bad algorithms or biased users?. *Journal of Information Science Theory and Practice*, 7(2), 40-53.

Zimmermann, F., & Kohring, M. (2018). „Fake News“ als aktuelle Desinformation. Systematische Bestimmung eines heterogenen Begriffs. *M&K Medien & Kommunikationswissenschaft*, 66(4), 526-541.

Appendix

| |
|------------------------------|
| INFORMED CONSENT FORM |
|------------------------------|

Researcher: Cecilie Moessle

Advisor: Fernando Ilharco

Purpose of the Study

This study intends to provide a better understanding of the way of how artificial intelligence can be used to identify fake news on social media. Your participation in the study will involve an online interview with an estimated length of 30 minutes. This interview will be recorded and transcribed for later analysis. You may pass on any question that makes you feel uncomfortable. At any time, you may notify the researcher that you would like to stop the interview. A summary of the results will be available to participants upon request.

Subject's Understanding:

- I agree to participate in this study.
- I understand that my participation is voluntary.
- I understand that I will not be identified by name in the final product if I don't agree.
- I am aware that all the recordings will be kept confidential.
- I understand that I may withdraw from the study at any time and request that any data collected not be used in the study.

Signature: _____

Full Name: _____

Date: _____

INFORMED CONSENT FORM

Researcher: Cecilie Moessle

Advisor: Fernando Ilharco

Purpose of the Study

This study intends to provide a better understanding of the way of how artificial intelligence can be used to identify fake news on social media. Your participation in the study will involve an online interview with an estimated length of 30 minutes. This interview will be recorded and transcribed for later analysis. You may pass on any question that makes you feel uncomfortable. At any time, you may notify the researcher that you would like to stop the interview. A summary of the results will be available to participants upon request.

Subject's Understanding:

- I agree to participate in this study.
- I understand that my participation is voluntary.
- I understand that I will not be identified by name in the final product if I don't agree.
- I am aware that all the recordings will be kept confidential.
- I understand that I may withdraw from the study at any time and request that any data collected not be used in the study.

Signature: _____

Full Name: Jonny Simon Preuss

Date: 16.07.2024

INFORMED CONSENT FORM

Researcher: Cecilie Moessle

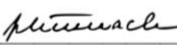
Advisor: Fernando Ilharco

Purpose of the Study

This study intends to provide a better understanding of the way of how artificial intelligence can be used to identify fake news on social media. Your participation in the study will involve an online interview with an estimated length of 30 minutes. This interview will be recorded and transcribed for later analysis. You may pass on any question that makes you feel uncomfortable. At any time, you may notify the researcher that you would like to stop the interview. A summary of the results will be available to participants upon request.

Subject's Understanding:

- I agree to participate in this study.
- I understand that my participation is voluntary.
- I understand that I will not be identified by name in the final product if I don't agree.
- I am aware that all the recordings will be kept confidential.
- I understand that I may withdraw from the study at any time and request that any data collected not be used in the study.

Signature: 

Full Name: Christoph Schumacher

Date: 17/06/2024

Transcription of Interview with Christoph Schumacher

Cecilie Katharina Mössle

What role do you think AI plays and can potentially play in detecting fake news on social media?

Christoph Schumacher

Tricky question. Because even our generative AI's are trained with data and they can't detect whether that news is actually correct or not. For example, use ChatGPT.

If somebody were to feed a certain amount of information into the net and it appears often enough ChatGPT would consider this to be accurate information and would use it. So generative AI certainly can't solve that issue. It would then require some form of search engine that would simply try to accumulate data to back up any information and you could expect this to happen in the future, just as you would run a background check on certain things on anything that the information is out there, you would use a very sophisticated search algorithm to check information, but the problem always is once enough information is out there, whether it's correct or incorrect, it becomes more and more difficult to differentiate truth from fiction.

Cecilie Katharina Mössle

So you would suggest that an AI tool works together with a search engine tool or platform?

Christoph Schumacher

Absolutely it would have to. But it always depends. AI is a huge field. So, it would always be the question what type of AI you would want to use, and we see now the more modern ones are also linked to uh search engine. I don't know if you use Bing or Microsoft products, they're linking search engines to generative AI still extract certain information and then models it and then and presents it. So yes, absolutely. There would have to be some form of search algorithm be involved here absolutely.

Cecilie Katharina Mössle

Do you think that there are any ethical considerations that should be made when introducing an AI supported system to detect fake news?

Christoph Schumacher

In AI at least where we are at this stage can't, because every AI is trained on certain information. And even the training itself creates biases in the knowledge base of that AI. So an AI cannot as such determine the correctness. If you're thinking well, we should get to a stage where it should, if ChatGPT was to tell you something, it should have checked whether that is correct. Should this happen? Yes, I think so. It's the usual kind of due diligence process you'd expect. Can it do so? I believe that would be very, very difficult to do. Therefore, I think the responsibility would sit with the user rather than with the AI. If an AI presents you with information, it would be up to you to check that this information is correct before using it. I think the due diligence should be at the part of the user and not part of the software presented.

Cecilie Katharina Mössle

Do you think in the future it might be possible that the user doesn't have to do anything and AI could be almost 100% right?

Christoph Schumacher

No, because if AI gets better, then AI will also be used to create fake news and therefore fake news will become so good that it's almost impossible to detect that it is infected fake news, and we are already almost there. When you hear about celebrities suddenly presenting a good or advertising good, which they've never done because somebody used AI to create an image that is simply not correct. But as AI improves in detecting fake information, this is the same process as to improve the quality of fake news, so I don't think we'll ever get to a point where AI can 100%, take it because these things go hand in hand. The better our ability is to detect fake news, the better also, the ability to create fake news and therefore umm, it will always be difficult to figure out who's a step ahead.

Cecilie Katharina Mössle

So you would just suggest to just educate people on how to detect fake news or how to handle social media?

Christoph Schumacher

It absolute absolutely. I believe the important part is education on that part. People have great trust in these AI's because it's this buzzword "intelligence" and so it must be something very clever and people do believe, and I've seen this when students submit essays where they clearly used an AI and I just look at certain values and numbers in there and instinctively I go that can't be right. And Students didn't even bother to check, simply because they assumed if the AI tells you this, it must be right. It's almost like with a doctor. If somebody is a doctor and tells you something you believe that they must be right, because this person is a doctor and people almost have a similar strange view about AI. If ChatGPT tells you this must be wrong, when I still maintain generative AI's are simply the most skilled bullshitters that they are, they can create content and tell it in a way that is so good that you almost feel it must be true, but in fact, they're just extremely skilled in packaging news, whether that news is correct or incorrect.

Cecilie Katharina Mössle

What level of transparency do you think an AI tool that detects fake news should offer users regarding their function? And their decision-making process.

Christoph Schumacher

Tricky question, because there's a fine balance in providing information and user being able to understand what is going on. So, what you're wondering is if I understand that correctly, is if an AI follows a protocol to determine whether news is correct, or whether it believes it to be fake, should it tell the end user what the decision process is. I'm sure it could be a help function where that information can be acquired, but I would doubt that the average user would have any idea on how the decision process works.

Cecilie Katharina Mössle

So it wouldn't make sense for the normal user that don't work with AI?

Christoph Schumacher

Yeah, I think they would find it very difficult if suddenly your tool tells you that it is using a cake cluster algorithm where it creates these subgroups and then finds changes between these clusters, then it reduces it to create an $n - 1$ cluster and views how information of that one cluster that is dissolved is distributed now, and so on and so forth. An end user would probably don't understand that. So these things are not always intuitive so would I like to know? Sure. Would many people be able to derive anything from that information?

Possibly not.

Cecilie Katharina Mössle

I think people will feel safer if they know they could look it up, but the question would be if they would understand it.

Christoph Schumacher

Yes, I'm sure. These things should be available, as in any analytics tool. If you use for example, SPSS, which is one of the most used data analytics tools and you just press a button and it gives you an outcome, but then you can go into the handbook and see for example, what algorithm did it use, what confidence intervals did it imply and so on. So you can look this up but most people won't, but people who are interested could yes, so I could easily imagine that any platform or any tool would have some kind of repository where you can retrieve the decision process.

Cecilie Katharina Mössle

Do you think there are or could be cases of misuse or possible dangers of using AI tools in social media?

Christoph Schumacher

Look, everything that can do good can do bad, and every time such tools are made available, some people will use it for the right reasons and some people will use it for the wrong reasons. This could be very tricky, because if you now rely on AI to detect fake

from non-fake news and you trust in in the AI then you could easily have false positives and in anything we do we have these. There's for example they I will consider something to be fake news when it's in fake not and even the opposite is terrible when it is correct, and you believe in it when in fact it was fake so somebody could systematically create fake news supported so that AI's won't detect it and then get away with it. So, absolutely, if you rely on an automated process to tell truth from fiction, it can be manipulated.

Cecilie Katharina Mössle

Do you think in the future it will do more good or more bad?

Christoph Schumacher

That I don't know. The usual thing is with developments, once something is out there, it's out there. You can't take it back, so AI is here to stay and it will do good and it will do bad. Only time will tell. You'd always hope that the good dominates the bad, but you can never exclude it. The problem could also be that bad parts might have far bigger impact than the good has done before. I mean, you know, discussions always come up with splitting the atom of the very good thing to power our nations but of course it could also be used to create a terrible bomb, and it only takes one bad event with that bad bomb to destroy all the good it has done before. So, we don't know, we hope, that it's used correctly and it's the responsibility now of the legislators to put safeguards in place to try as much as they can and to ensure that these tools are used in the right way and stop the misuse of these tools. But yes, every time you have something that opens the door to misuse, you'll find somebody who will do it. I mean think about cybersecurity, what a huge industry this now is and think about it it's just there to prevent people from misusing the Internet and what a big industry that has become. So, AI will be no different.

Cecilie Katharina Mössle

Is it always ensured that AI powered systems that detect fake news keep the users private information while they identify the fake news, for example, if I put out fake news and AI is trying to detect if my news are right or wrong, will my personal information be kept private or is there a possibility that it can come out?

Christoph Schumacher

Of course, there's always a possibility can come out as soon as you are on the Internet, nothing is private anymore. I'm sure there are rules and measures in place to prevent this, but once it's there, it's there and I don't think AI can ensure any of those things. If an algorithm can find the information, so can someone else. I believe that risk is implicit when using social media that somebody finds out who you are. When you send a message through the Internet, you can't stop somebody capturing this information somewhere. When you use your Email, do you know that your Email account is safe and wherever your data is being stored and which database it is and that somebody's not accessing it?

Cecilie Katharina Mössle

We assume, I thought we don't know.

Christoph Schumacher

We hope, and we still know it doesn't happen because governments do control certain things. They look for words in messages you receive and take you accordingly. So we don't want to think about it, but now nothing that happens in the Internet is safe.

Cecilie Katharina Mössle

Umm do you think AI can ever replace human judgment in the fight against fake news or do you think humans should always have a look on the AI and work together with it?

Christoph Schumacher

Well, I would be greatly worried if I ever was to become autonomous. That is the scary part. The next question of course, would be what fake news is as such and if even humans can figure out what is in and what isn't, but I think ultimately the responsibility lies with the user and not with your platform or AI. There needs to be safeguards in place and that is the responsibility of any company who offers such a tool. But at the end it is the human who uses that information to make the final judgment.

Cecilie Katharina Mössle

How receptive do you think social media users are to AI tools that are fighting fake news, and do you think they use them frequently or less frequently?

Christoph Schumacher

I'm guessing here because I have no empirical evidence of that. I'd be surprised if social media users spent time on figuring out if news are correct or incorrect. Also, the way our brain works with the different heuristics we use if something happens or if we are told something often enough, we tend to believe it's true. And if it's some famous person that tells the story, we also tend believe even more that it's true. The whole part about social media to a large extent is entertainment, it's pleasure, it's to be fun, which means we don't want to engage that analytical part of our brain because that means we use up energy, thoughts, time and these things we don't want to when it comes to social media. So I'd be surprised if too many social media users check facts before they distribute information they've got from social media or believe in something I think it's almost counterintuitive to the whole idea of social media.

Cecilie Katharina Mössle

So do you think if platforms even would offer AI tools or make it very easy to fact check people are basically using social media with a turned off brain and wouldn't even consider using it?

Christoph Schumacher

Well, the question would be if such a tool existed, why doesn't the platform use it as a filter before? So don't put it on the end user. Platforms should put that filter in front before it reaches the user and social media platforms do such things because they screen content for certain things such as profanities, nudity, offensiveness, and so on, and remove these things if they are observed. So, you would expect this to already happen to the best of their knowledge or use such filter and indirectly such tools exist. If you read something, you can go onto one of the search engines and look a bit closer whether they just are true or not. I

would expect the platform providers to check that if such a filter exists before it reaches the end user, rather than putting everything out there and then having the end user apply.

Cecilie Katharina Mössle

How can those AI tools be more accessible and understandable to the public.

Christoph Schumacher

They have become very, very accessible, right? Because there are some state-of-the-art free AI's out there, all the big system providers have their own AI. Google has one, Microsoft has its own version of them, so certain ones are free. Some you must subscribe to, but there's plenty out there that you can already use and it accessible and I think we've seen just in the last year a huge search in people using these AI's. I think if you go back two years, for example, few people would have known what ChatGPT is, and it existed with open AI before that. Now that everybody speaks about it. So, I think a lot has happened in the last year when people are far more aware of these tools that are out there and big companies provide AI for example Microsoft 365, you already now can get copilot in form of AI that helps you with things and Google as well. So, I think these bigger companies are all heading that way anyway.

Cecilie Katharina Mössle

Do you think that the trust of the Internet user can be influenced by AI tools? And if so, how can it be influenced?

Christoph Schumacher

I'm quite sure that that can be achieved. Simply because people believe AI is this magic tool. It all over the news that AI predicts certain things and knows everything so people think, wow, if it comes from there, it must be better simply because AI has powers that we don't have, it has access to so much information and then there's a lot of good in this. So yes, I believe if there were such tools, people would be more trusting in this information if they knew that it was checked by the AI and considered it OK.

Cecilie Katharina Mössle

AI tools can kind of influence the user, but can the user also influence the AI?

Christoph Schumacher

In a way, yes, because AI's are trained and feed information that ultimately comes from the user, I mean this is how a lot of generative AI is built up they are out there for free and whenever you interact with them, they learn from it and from this information, so would I call it learning? Maybe not, but certainly it can be influenced, as I've said, if I were now deliberately to, for example, put out there that $1 + 1 = 3$ and I write papers, generate papers, put it out there, and generative AI uses probability distributions to collect information, and if something is the dominant view, it will grab this information to put something together. So if I was to make a huge effort to convince the world that $1 + 1 = 3$, I probably could, and eventually generative AI, when I asked what does $1 + 1$ will say 3.

Cecilie Katharina Mössle

OK, didn't you try something like this? I think my father told me something about it.

Christoph Schumacher

Yes, I did. I did exactly this and I actually I communicated this to open AI and I created a fake spider name and wanted to know and assumed that it can use its magnetic receptors to guidance where are completely made up and got a lovely story written that was so convincing that you needed to be an expert to build its correctness. But this is exactly the things I I was referring to. AI's are built on information that's presented to them. Papers that have been fed and if incorrect information has been put into the system A I will eventually believe it's correct. So, yes AI can be influenced by end users, but it would take quite a little liberate action because it's using big date. That's why we call it big data. It needs to be big. It needs to happen often enough.

Cecilie Katharina Mössle

So with your example, you didn't just type it in once you worked on it for a longer period?

Christoph Schumacher

No, it was a first question and I simply wanted to check what comes out so I didn't feed this, but it was simply the way these generative AI's work that I thought, OK, they can be tricked, and this is exactly what I did. I combined different things and implied that they belong together and yes, AI believed this and went all the way with it and suddenly told me that the Colombian jumping spider can use its receptors and its eyes to which... the Colombians jump Spider doesn't even exist. I knew that there are jumping spiders. I knew Colombia as a country. I put these two together. I had heard of birds using the earth magnetism to guide their way and then the AI figured out or this is correct. Colombia is a country. There's a spider. AI combined this information, and it made up a whole essay wrote. Then I asked it to write a big student essay about this, explaining it, and it went into detail. It was a fantastic read, so much soul that when I read it, I thought maybe this really exist? Maybe by chance made something up that does exist. It was that good. So once you understand how these AI's work, you can use that to create information or mislead them in a way.

Cecilie Katharina Mössle

And then so the information now is out there. Is it possible to quickly tell AI that it is not true?

Christoph Schumacher

Actually I've did the same thing or three months ago. I asked the same question and this time ChatGPT answered that there is no such thing as a Colombian jumping spider. So, it just means they're getting better and better, and are being more precise. And what things can be done and checking the links exist or don't exist? I mean, I'm quite sure I can find a correlation between sunshine hours in China in the performance of a stock just because I can find a correlation between doesn't mean one causes the other. This is where the check needs to happen, so generative AI's are getting better and figuring out just because two things move in the same way doesn't mean one causes the other, so there's a big question of correlation and causality which are different things, and I think AI's are getting better in detecting what is pure coincidence and what is actually fact.

Cecilie Katharina Mössle

Yeah, that's already it. Thank you so much for your insights!



Transcription of Interview with Jonny Simon Preuss

Cecilie Mössle

What role do you think AI plays, and can potentially play, in detecting fake news on social media?

Jonny Simon Preuß

Of course AI plays an important role in this kind of thing. I believe that it will definitely help to filter out misinformation that is simply very, very obviously false, as AI is able to quickly analyze large amounts of data and recognize patterns that may indicate such misinformation. This supposed misinformation is then forwarded to human moderators, who can then check the facts. However, it is sometimes a little more difficult with social media in particular, as these are platforms where anyone can express their opinion publicly. This is why tools are currently being developed that can detect fake news without having to rely on databases. These tools are being trained to become increasingly accurate and effective in the fight against fake news. Of course, it depends on whether the social media platforms want to work with these tools.

For people who want to spread fake news, this means that they have to make small changes to factually correct information. This is the only way to deceive the AI and change public opinion bit by bit.

Cecilie Mössle

Do you think that there are there any ethical considerations that should be made when introducing AI-supported systems to detect fake news on social media?

Jonny Simon Preuß

Yes, ethical considerations are essential when introducing AI-supported systems for detecting fake news. Of course, any AI that works with content shared on social media must somehow operate within the lines of freedom of expression. In Europe, it has to respect user privacy more than in America and the big issue is of course always whether the AI that tries to recognize fake news has been trained correctly, i.e. whether there is some kind of bias within the training data with which it is trained. If there is such a bias, then it may be that news that was classified more strongly as fake within the training data is later also classified as fake, although it may simply have been disproportionately included in the training corpus of the training data. Care must therefore be taken to ensure that information is not distorted by the programming of the AI. Transparency in terms of how the AI works and makes decisions is also crucial, but this could either be towards the regulators (e.g. so they are aware of the AI models being used) or, in a more drastic step, towards the customers so they know why a message is not being displayed to them or their own post has been deleted. Then, of course, it's about ensuring that decisions made by this AI are kept transparent, so this can either be in the direction of the regulatory organizations, for example, in the EU, compliance with guidelines such as the GDPR or the AI Act or, for example, in the UK there is also a kind of authority for algorithm transparency and they then of course make sure that a large social media operator labels it when it uses such an AI. The operator must then disclose how this AI works and how certain news is filtered out. Of course, it can now go in the direction of customers or social media users, who then have to find out why, for example, a post was blocked by them or labeled as fake news.

Cecilie Mössle

What level of transparency do you think such AI tools should offer users regarding their functioning and decision-making processes?

Jonny Simon Preuss

The problem with such an AI is of course always the so-called black box problem. Such algorithms usually have to sacrifice some accuracy if they are to be more explainable. AI deep learning model, which learns by itself which things are fake news or not, can spit out a very good result in the end, but it may not necessarily be able to tell you how it came to this result and such better explainable algorithms, which can show how they came to this certain result, are usually somewhat worse and lower in accuracy than such black box algorithms might be. If you then say that we want to be transparent and be able to justify why we have come to a result, why something is fake news, then you might at least have to make the compromise of giving up some accuracy in return. Of course, AI tools should offer users a high level of transparency, including clear explanations of how they recognize fake news and what criteria they use in their algorithms. Users should be able to understand why certain content is flagged and have the ability to challenge decisions, especially if it is user-generated content and not just the republication of an article.

Cecilie Mössle

Is such a black box a standard for every AI tool?

Jonny Simon Preuss

So black box ultimately just means that you run the AI model and at the end it comes out whether a text is fake news or not. But there is no explanation behind the decision. You have to imagine this black box figuratively or fictitiously, if you want to remove this black box, then you have a model that says, for example: The content piece is fake news and I have made this decision with, for example, 5 features in my model, the most important of which is the feature, which is based on which link behind it is the other important feature, for example keywords.

So a black box can explain to you which specific properties of an AI model played a role in detecting fake news. Self-learning deep learning models don't have that, but they create their output themselves and therefore can't tell you in the end what's behind their decision.

Cecilie Mössle

In general, what do you think are the advantages and disadvantages of using these AI tools from a societal perspective?

Jonny Simon Preuss

The benefits include the ability to quickly and efficiently detect and curb the spread of fake news, thereby protecting public discourse and democracy.

The disadvantages, however, include potential biases in AI systems, the risk of over-reliance on the technology and concerns about privacy and freedom of speech.

Cecilie Mössle

Do you think there are or could be cases of misuse or possible dangers of using AI tools in social media?

Jonny Simon Preuss

Of course, in extreme cases, a platform could use this case of fake news detection as an excuse to block any unpopular opinions.

If you think about how Twitter could theoretically develop with Elon Musk, for example, you could say that at some point he could twist Twitter's fake news detection so that unpopular political opinions or opinions that he personally doesn't like are blocked and thus denied to the public. Ultimately, it's about who decides which facts are correct. After all, it is ultimately a human decision as to what goes into the training data for such an AI model.

Cecilie Mössle

This means that the AI models can theoretically also be biased.

Jonny Simon Preuss

Correct. Every AI model is only as good as the training data that is used, and that is ultimately always a human decision. One disadvantage of this AI is, of course, that no model achieves 100% accuracy, which is why it can happen that the AI recognizes something as fake news even though it is actually true news. But it can also happen the

other way around. Because there is always a small risk of something being misidentified, people may lose trust in the AI.

Cecilie Mössle

Is it always ensured that AI-powered systems that detect fake news keep the users privacy while identifying false information?

Jonny Simon Preuss

As I said, no AI is perfect, but now with regard to user data, these AI models are actually trained so that they don't need to access private user data. I don't believe that an AI is trained to know which user is posting this fake news, it's purely about the text or the metadata of the article that is being distributed.

Cecilie Mössle

So you think that if I write a text on Twitter now, but don't necessarily publish it under my name but anonymously, the AI can't find out that it was me?

Jonny Simon Preuss

Yes, so from my point of view, the AI that looks for fake news should be able to recognize whether the post is fake news or not by only looking at the headline, the words used, the phrases, the link behind it, etc. So the models that we tested during our studies really focused purely on the content and metadata of the article.

Cecilie Mössle

Do you know of any laws that support the regulation of AI?

Jonny Simon Preuss

Yes there are laws such as the GDPR in Europe that aim to protect user privacy, but the implementation and enforcement of these laws in relation to AI tools can be complex and inconsistent. The most recent law that has been enacted at EU level is the AI law, although I don't know how it applies to fake news.

Cecilie Mössle

Do you think AI can ever replace human judgment in the fight against fake news, or should humans continue to work alongside AI?

Jonny Simon Preuss

Of course, AI can never completely replace human decision-making, which is why we often talk about “human in the loop” scenarios for things like this. For example, if Trump posts something, I don't know if you would completely rely on an AI to check this scenario. In such important cases, you would rather have a human look over it and make the final decision as to whether the post is fake or real. However, if it is a normal social media user who publishes something, it is more likely to be an AI that checks it. As I said, an AI is not perfect, which is why a human decision is better in some scenarios.

Cecilie Mössle

How receptive do you think social media users are to AI-tools to fight fake news on social media and do you think users tend to use them frequently or not frequently?

Jonny Simon Preuss

I would say that it differs which type of user that is. The willingness of social media users to use AI tools to combat fake news simply depends on various factors, such as political opinion, technological openness, does the person know that the problem of fake news exists, how is the topic communicated by the platforms, political perspective in the sense of how skeptical my counterpart is of the authorities. If people get the feeling that their freedom is being taken away when their own posted content is looked over again and there is a possibility that it will be blocked, they may be less receptive to such tools than someone who has no problems with them and whose posts are not flagged as fake. I believe that some users appreciate the efforts to curb misinformation and use these tools frequently, while others are skeptical or unaware of their existence. For better acceptance, raising awareness and proving the effectiveness of these tools can help increase their use.

Cecilie Mössle

How can AI tools for detecting fake news be made more accessible and understandable to the public?

Jonny Simon Preuss

Classically, you can launch educational campaigns. If Instagram were to run small campaigns and point out to people that they can post within Instagram's guidelines, but make them aware that their posts will be looked over, users would at least not be able to claim that they had never heard of it. Among other things, tutorials could be added to the platforms and simply make sure that it is user-friendly.

Cecilie Mössle

Do you think that the trust of internet users can be influenced by AI tools? If so, how?

Jonny Simon Preuss

Internet users can definitely be influenced, especially of course when they see that the right decisions are being made by AI. Ultimately, this is probably also about communication. For example, if a platform like Instagram has been running this for a while, you might at some point read a figure somewhere that informs the user how much fake news is actually filtered out and that the platform says that it has become safer thanks to the AI. This could also increase trust. The platforms should therefore focus on transparency here. If, on the other hand, AI tools are perceived as opaque or biased, trust can be undermined.

Cecilie Mössle

What role do you see for internet users in the design of AI-supported tools to overcome fake news? Can users also influence the AI?

Jonny Simon Preuss

Internet users can play an important role in the development of AI-supported tools by providing feedback, reporting inaccuracies and suggesting improvements. Of course, you can always do this with a short survey asking the user whether they like it or not, or you

can provide a link to a feedback questionnaire. Some AI tools allow interaction with users to improve their algorithms, e.g. by upgrading or downgrading content, allowing the AI to learn from human input.

Cecilie Mössle

And do you think it's possible to train AI as a private individual or as a private internet user, or can only the so-called trainers do that?

Jonny Simon Preuss

I do think that user feedback can play a role as another block of training data. The AI makes the decision as to whether it is fake news or not and ultimately has no feedback on it. A simple example in the data science setting is the classic example of programming an AI that predicts which person survived the Titanic accident. In the case of the Titanic, it is known which of the passengers survived and which did not. So you know, after you have run your AI and the AI tells you that passenger 1 survived, for passenger 2 the AI has said that he did not survive, etc., you can check whether this prediction is correct, because you already know what the results are and you can then say how accurate your AI is. With fake news, of course, you try to create a certain set of fake news as training data, so you give it a bunch of New York Times articles, for example, and say that's not fake news and then you give it a bunch of articles from a made-up newspaper like, say, Breitbart News or something and then you look at what's fake and what's not fake. But you probably don't have tons of articles available to train the AI and this user feedback can be used for that. Of course, you shouldn't put it on the gold scale, because you can't let it go back in unfiltered without at least checking it. But it could be that this user feedback is then converted into a larger corpus and training data, in which you then take the content piece that was posted and the user feedback and then perhaps have it checked again by humans and then use it to retrain the AI. Let's say it was last trained on January 1, 2024 and since then new articles have appeared and so you can then continue to train the AI and have it trained on new emerging pages on the Internet, have new keywords trained and so on. So this training dataset should be constantly expanded, that's true, and this hints that the user could be a starting point for the people you have in the background who provide this training dataset or review the articles that flow into the training of the AI, perhaps by making the decision a

little easier for them or providing a cross-check. So in short: Yes, user input can be used, but it should not be allowed to flow in unfiltered.

Cecilie Mössle

If I were to upload such articles to Google now and put a lot of effort into the articles and somehow write papers that are actually fake and, for example, report that there are blue elephants, could it happen that a platform like ChatGPT would then at some point think that there really are blue elephants?

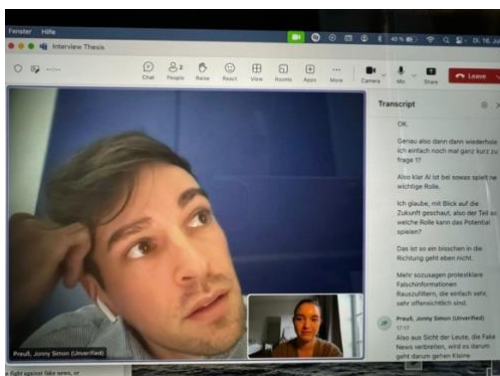
Jonny Simon Preuss

It is certainly a problem at the moment that many people are putting certain information on websites that they have created quickly in order to be able to link backlinks to certain pages. With the help of ChatGPT, you can program a website within 5 minutes. You then simply buy any URL and do this 100 or 1000 times and always write the same information in it in certain variations, so that your actual target page can get to the first page of the search results on Google through this jumble of extra information that is similar to what you want to promote and it is pushed through all the articles on Google. If this is also fake news, I hope that A Google itself has such control mechanisms that run in the background and B that such an AI model then also has a filtering step in the training corpus to somehow filter out such pages. But yes, it is a problem that has not yet been finally solved.

Cecilie Mössle

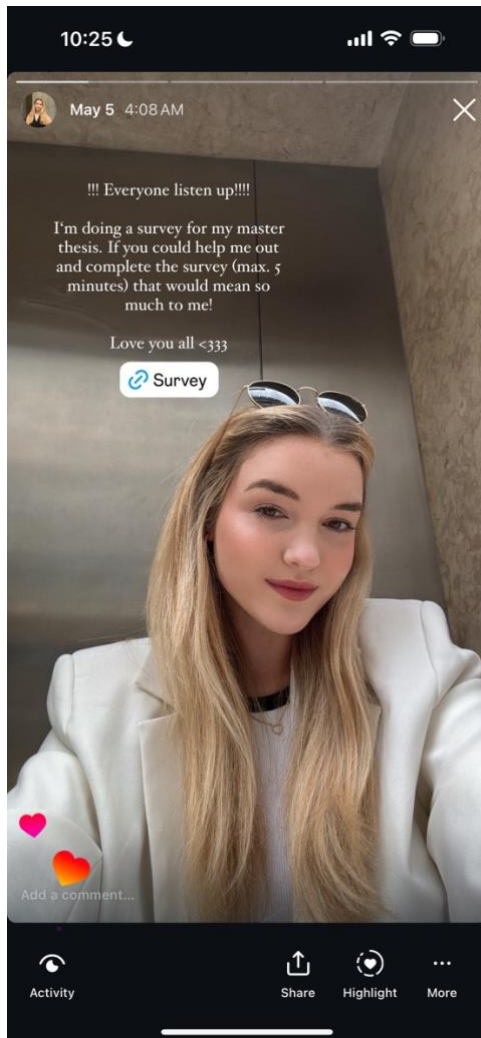
OK.

Thank you for the Interview!



Call to action to participate in survey

To reach out as many participants as possible the author used her own Instagram account @cecimoessle where she has around 50.000 Follower.



Survey on Google Forms

Survey for Master Thesis

B *I* U [↗](#) ~~X~~

Dear Participant,

My name is Cecilie Mössle and I am doing my Master in Marketing, Communication & Advertising at the Universidade Catolica Portuguesa in Lisbon. I would appreciate it if you participate in my survey for my master thesis investigating how artificial intelligence (AI) can tackle the issue of fake news on social media.

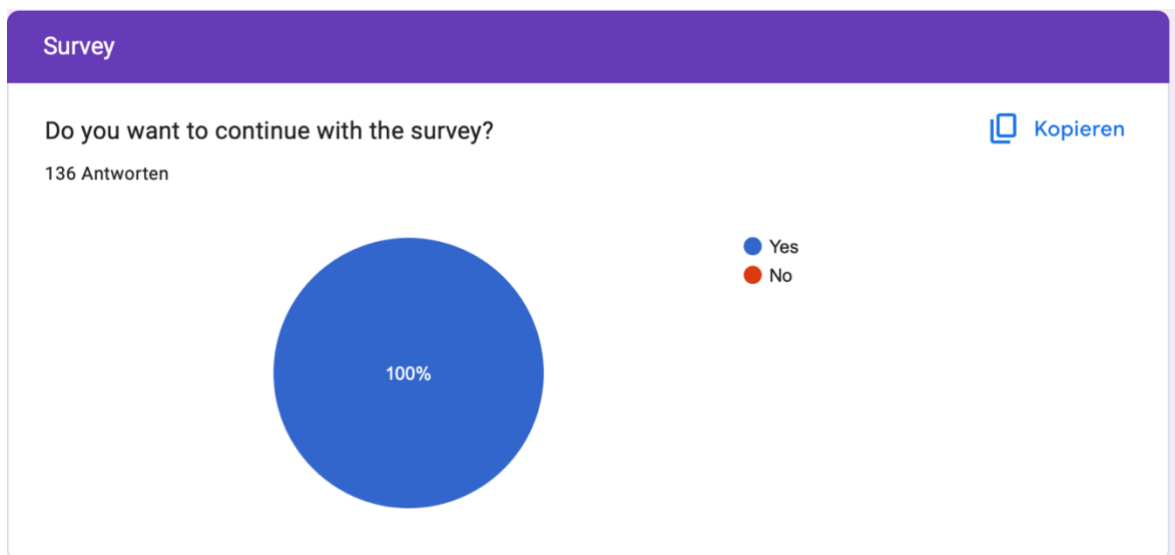
In today's digital landscape, disinformation spreads rapidly, posing significant challenges to society. AI offers innovative solutions, utilizing machine learning and data analysis to discern truth from falsehood.

Your insights will shape my understanding of AI's role in battling fake news on social media, contributing to a more informed digital world. Please only participate if you are 20 years and older. Your Data will only be used for this master thesis and is 100% anonymous.

Thank you in advance for your time and valuable input.

Best regards,

Cecilie

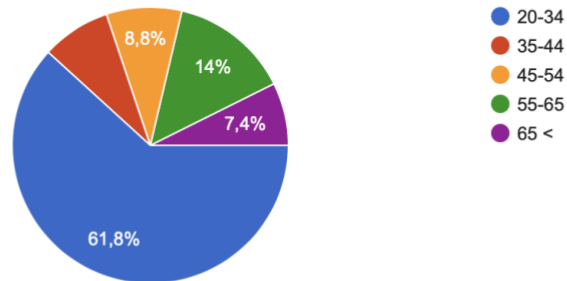


Survey

1. How old are you?

[Kopieren](#)

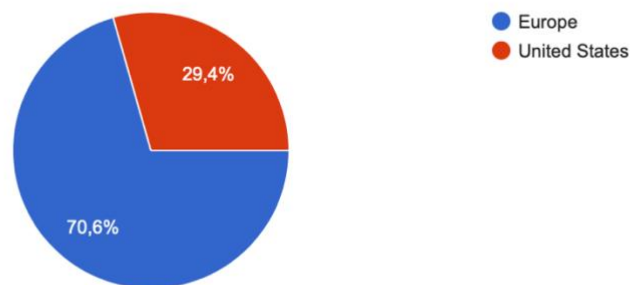
136 Antworten



2. Where are you from?

[Kopieren](#)

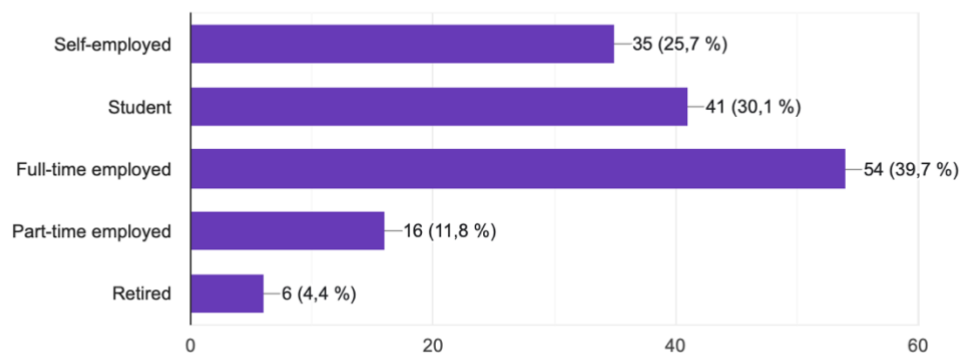
136 Antworten



3. What is your profession?

[Kopieren](#)

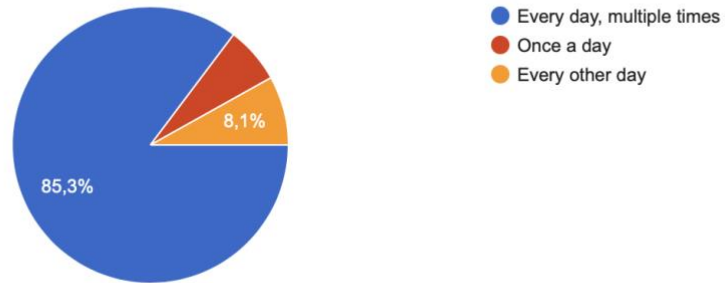
136 Antworten



4. How often do you use social media?

 Kopieren

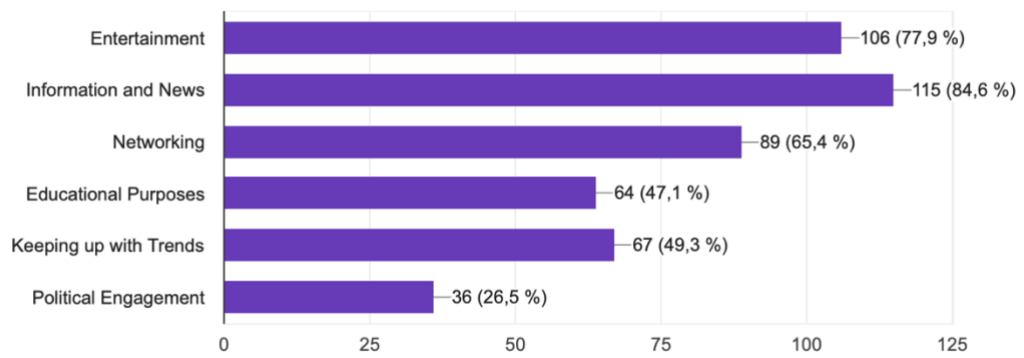
136 Antworten



5. What do you use social media for (more than one option possible)?

 Kopieren

136 Antworten



6. Please describe in some words what the term "fake news" means to you?

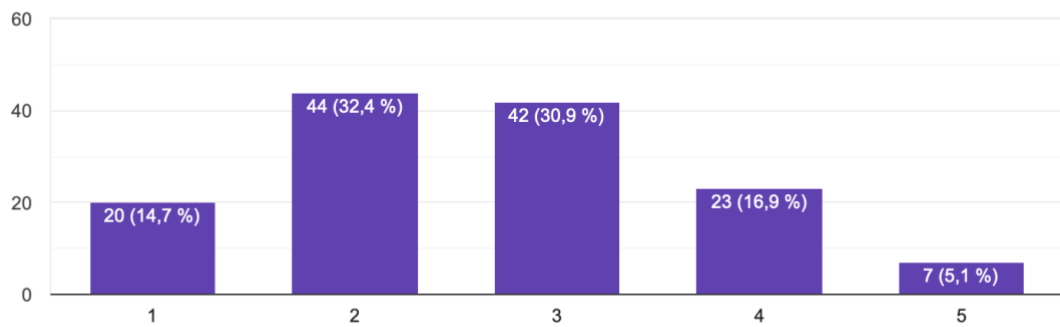
136 Antworten

- wrong information
- Wrong portrayal of what actually happened
- Propaganda
- Wrong information that is spread
- Mass Media, because they are lying usually
- Wrong informations
- Propaganda promoted by parties with biased agendas.
- misleading information,Fake news has become a significant concern in today's digital age.
- False Information

7. How frequently do you think you come across fake news on social media?

 Kopieren

136 Antworten



8. Can you recall an instance where you were unsure about the credibility of the information you saw on social media? If not just answer with "no".

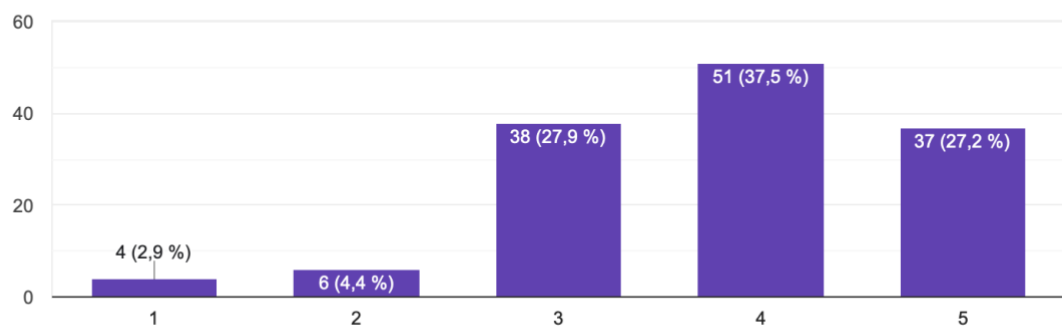
136 Antworten

| |
|--|
| No |
| no |
| No |
| Yes |
| Many times - mostly political content |
| Zelensky purchasing a luxury yacht |
| ORR / mass media |
| Yes there is a lot of false information about the actions of the Israel army |
| About the war |

9. In your opinion, how effective are current methods in battling fake news on social media?

 Kopieren

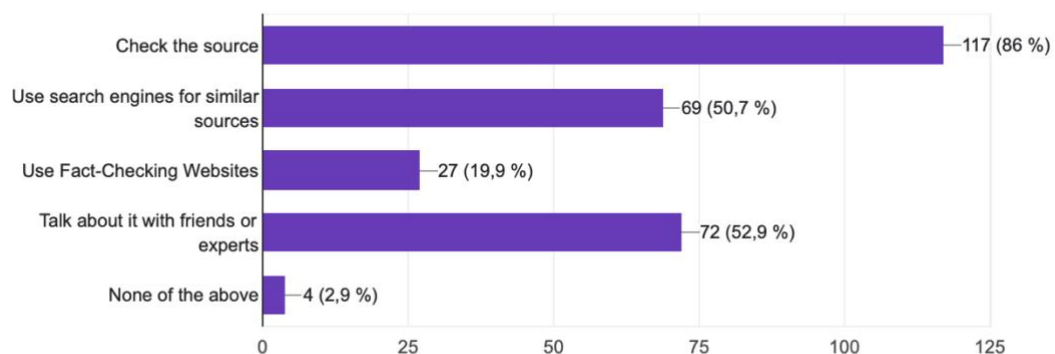
136 Antworten



10. What steps do you usually take (if any) to verify the accuracy of news or information shared on social media?

 Kopieren

136 Antworten



11. Are you aware of any efforts by social media platforms to overcome disinformation and fake news?

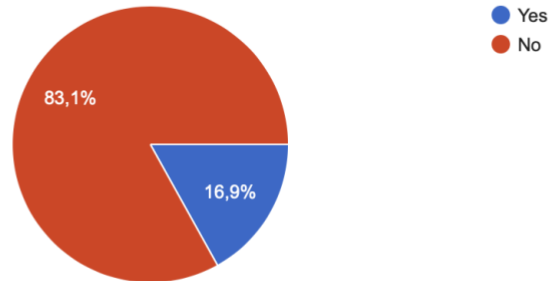
136 Antworten

- No
- Yes
- no
- No
- Not really
- yes
- no
- TikTok
- I don't know any

12. Have you ever encountered Artificial Intelligence (AI) tools designed to identify fake news on social media?

 Kopieren

136 Antworten



13. If so, how effective do you think have these AI-powered features been in identifying fake news in your opinion? If not, skip the question.

25 Antworten

Relatively effective, at least on X

Unfortunately, artificial intelligence cannot solve this problem—at least not with something as simple and comprehensive as an algorithmic “fake news detector.”

They can also be wrong i've noticed

Not mature, yet

Pretty effective

Considerably effective

Not very effective

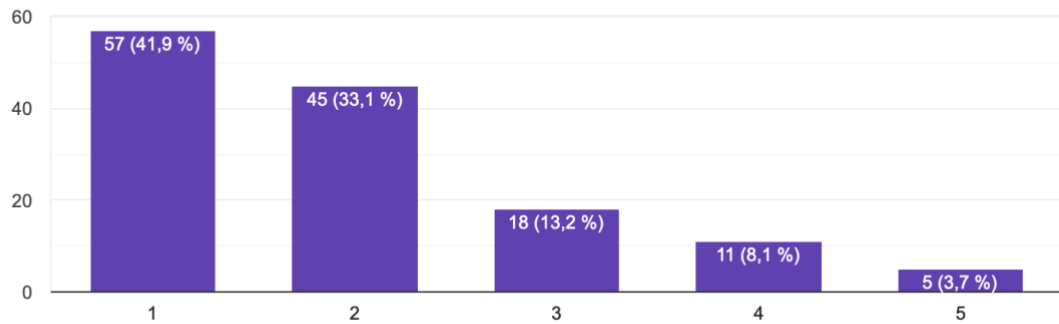
N/a

Some efforts to eliminate bots

14. How concerned are you about the presence of fake news and disinformation on social media?

 Kopieren

136 Antworten



15. What do you think could be challenges of using AI for identifying fake news?

136 Antworten

people questioning if AI is giving them the neutral stance

No

Double check

What if AI gets hacked and trained with bad intentions

AI are programmed left orientated and useless

The mafia

Maybe sometimes AI could not pick up fake news

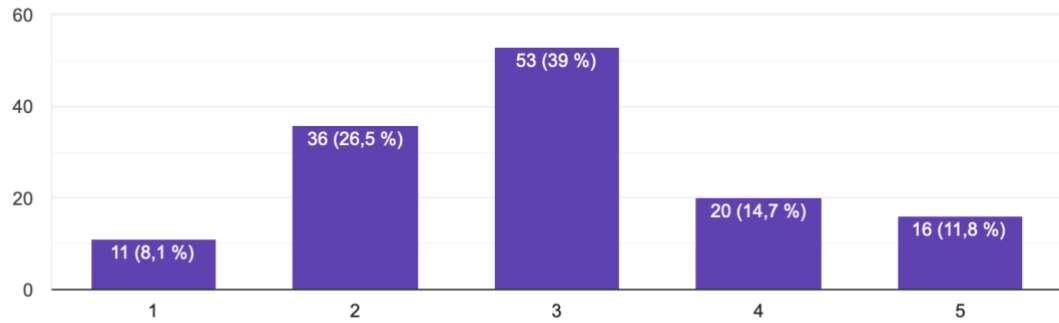
not clear

It's just a computer learning from the internet and the internet has bad info.

16. Would you be more likely to trust a news story or information shared on social media if it was verified by an AI-powered system?

[Kopieren](#)

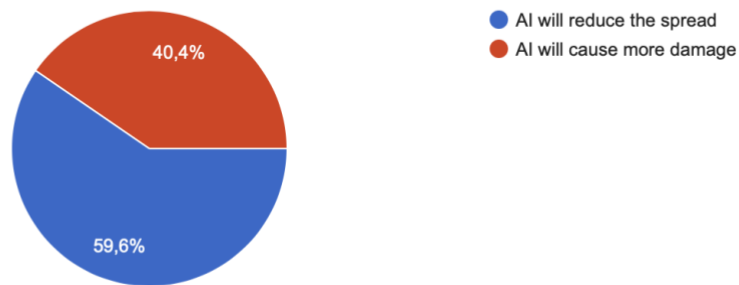
136 Antworten



17. Do you think AI technology could play a significant role in reducing the spread of fake news on social media or do you think it will cause more damage?

[Kopieren](#)

136 Antworten



18. What concerns, if any, do you have about the use of AI to identify fake news on social media?

84 Antworten

It's a good question

no concern

It's going to get it wrong sometimes

No concerns. Would be helpful

Menschenmassen in die falsche Richtung zu bewegen, zu manipulieren

I think that AI might serve as another tool for fact-checking but not more

AI source and trading data has to be fully transparent as otherwise it could even amplify the problem

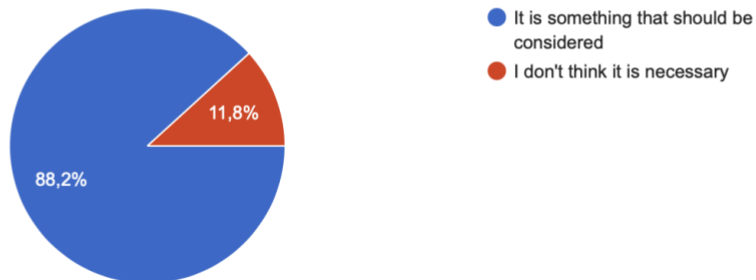
Because there is a possibility of a false verification of fake news will cause more difficulties and confusion

political motivation

19. Do you believe that educating users about how to spot and report fake news could help mitigate its spread?

 Kopieren

136 Antworten



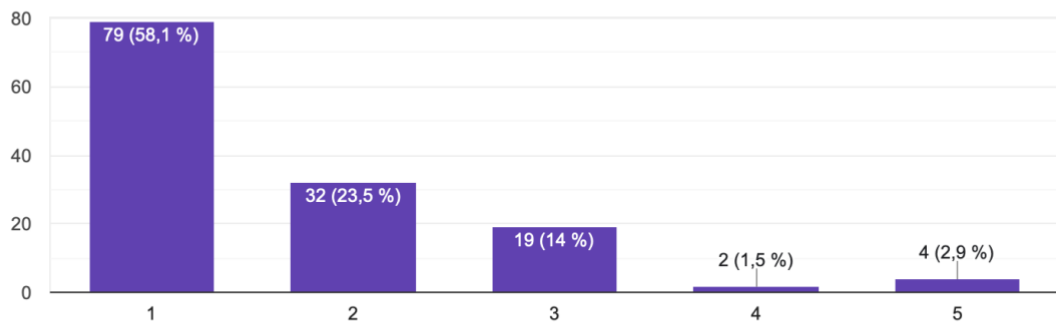
20. How do you think the presence of fake news on social media platforms affects your trust on the information you see there?

136 Antworten

- No
- yes
- I tend to double check sources I trust
- Yes
- I don't trust the media in general
- I think it makes me question everything a lot more
- None
- I use social media for mostly entertainment, so a lot of the time I am not concerned whether the info is true.
- does not affect it much

21. How important is transparency in the algorithms used by AI systems to identify fake news to you? [Kopieren](#)

136 Antworten



Survey

Thank you for your participation!