

Bachelor Thesis

University of Twente, Enschede &
Westfälische Wilhelms-Universität, Münster

DEEPPAKES AS A THREAT TO DEMOCRACY **Perceptions, Challenges, And Implications Of Deepfake Discourses** **In Democracies**

Paula Tabea Berenzen

s2815818

Programme: Public Governance across Borders (MST)

Supervisors: Dr. Ringo Ossewaarde, Dr. Azadeh Akbari

Date of Presentation: 29th June 2023

Word Count: 11946

ABSTRACT

The aim of this bachelor thesis is to examine *in what ways the relationship between democracy and deepfakes is discussed by the governments of Austria and UK*. To achieve this, concepts of critical theory, specifically those of the Frankfurt School, with a focus on scholars such as Adorno, Horkheimer, Anders, and Marcuse, will be employed. A qualitative content analysis will be conducted on text documents, including policy papers from both countries. Additionally, newspaper articles will be analysed to gain an overview of the discourse development in both countries. Austria and the UK are considered two highly developed European countries, and given the impact of Brexit, it is important to understand how the UK is developing without EU regulations. A coding scheme is provided for operationalization of the core theoretical concepts, and the tool ATLAS.ti will be used to analyse the collected documents. As discussion within society is crucial for shaping opinions in a democracy, a content analysis will provide insights into both the stated and unstated meanings. Key findings indicate that while both the UK and Austria perceive deepfakes as a threat to democracy, they adopt different approaches in terms of regulation and identifying the specific dangers posed by deepfakes.

TABLE OF CONTENTS

1. Introduction.....	1
1.1. Background And State Of The Art.....	1
1.2. Knowledge Gap And Research Questions.....	2
1.3. Research Approach.....	3
2. Theory.....	5
2.1. Introduction To Core Theoretical Concepts.....	5
2.2. The Frankfurt School And Its Approach To Technology.....	5
2.2.1. The Dialectic Of Enlightenment: Adorno And The “Myth Of Progress”.....	6
2.2.2. The One-Dimensional Reality: Marcuse's Critique Of Technological Domination.....	6
2.2.3. Problematic Technology And Its Impact On Democracy.....	7
2.3. Günther Anders And The Abuse Of Technology.....	8
2.3.1. Promethean Gap And Shame.....	8
2.3.2. Günther Anders Perspective On Problematic Technology.....	9
2.4. Theoretical Approach In Relation To Deepfake Technology.....	10
2.4.1. Frankfurt School And Its Approach To Deepfakes.....	10
2.4.2. Günther Anders And His Approach To Deepfakes.....	10
2.5. Conclusion.....	11
3. Methods.....	12
3.1. Case Description.....	12
3.2. Method Of Data Collection.....	13
3.3. Method Of Data Analysis.....	14
3.4. Conclusion.....	17
4. Analysis.....	19
4.1. Introduction.....	19
4.2. Deepfake Discourse In Newspapers: An Increasing Fear?.....	19
4.2.1. Discourse In Austria: From Fear To Regulation.....	19
4.2.2. Discourse In The UK: From Fear To Action.....	20
4.3. Deepfake Discourse In Governments: Regulations Are Necessary.....	22
4.3.1. Austrian’s Government: “Deepfakes As A Major Threat”.....	22
4.3.2. UK’s Government: “Concerning But Not Threatening”.....	23
4.4. Comparing Deepfake Discourses In Newspapers: Austria Vs. UK.....	24
4.5. Comparing Deepfake Discourses In Policy Papers: Austria Vs. UK.....	26
4.6. Do Deepfakes Raise A Threat To Democracy?.....	28
5. Conclusion.....	30
5.1. The Discussion Of Democracy And Deepfakes.....	30
5.2. Suggestions For Future Research.....	31
5.3. Practical Implication For Policy And Governance.....	32
6. List Of References.....	33
7. Appendix.....	36

1. INTRODUCTION

1.1. BACKGROUND AND STATE OF THE ART

The widespread use of technology in our daily lives has a profound influence on society. Although technology has made life more convenient for most people, it also poses significant dangers, particularly in terms of its impact on democracy. Especially new AI technology is seen critical among politicians as well as within society. The Cambridge Analytica Scandal, for example, revealed how a political consulting firm harvested Facebook user data without authorization and subsequently used it to target political advertising during the 2016 US Presidential Election. The scandal highlighted the potential impact of social media on democratic processes. This research is not particularly interested in how technology influences lives but rather on the critical discussion of recent AI technologies and its relationship towards democratic systems. A recent example of technology that poses a threat to democracy are deepfakes, which are defined as "*[...] an audiovisual record created or altered in a manner that the record would falsely appear to a reasonable observer to be an authentic record of the actual speech or conduct of an individual.*" (Malicious Deep Fake Prohibition Act 2018). Although deepfakes have mainly been used to create pornographic content and stigmatize women as a target group (Diakopoulos and Johnson 2019, p. 2089), political actors are becoming increasingly vulnerable to deepfake technology.

The current literature states that deepfakes can be seen as a threat to democracy, so it is critically discussed within science (Chesney and Citron 2019; Diakopoulos and Johnson 2019; Dobber et al. 2020; Masood et al. 2021; Ray 2021; Vaccari and Chadwick 2020). This means, on the one hand, that politicians' integrity can be harmed through the spreading of deepfakes about them. On the other hand, deepfakes can harm democratic elements such as elections and national security (Diakopoulos and Johnson 2019). Dobber et al. (2020, p. 86) stress that citizens' ability to detect deepfakes is limited, which can negatively influence their attitude towards politicians. Further, it is of an urgent matter that the believability of videos and information might also be questioned in the future, leading to difficulties in forming public debates on the internet (Vaccari and Chadwick 2020). In contrast to these findings, Hameleers et al. (2022) state that the power of deepfakes in destabilizing democracy should not be overstated and that it remains an under-researched topic, making it difficult to draw definitive conclusions. On an institutional stage, researchers point out the

urgency that deepfakes could influence the integrity of democratic elections since they might be used to manipulate information voters are using for their decision-making (Diakopoulos and Johnson 2019, p. 2073; Ray 2021, p. 983). Additionally, Ray (2021, p. 987) highlights the essentiality of deepfakes posing a threat to elections by altering vote preferences and undermining trust in democratic institutions. Also, Masood et al. (2021, p. 3974) emphasize the significance of deepfakes in terms of their potential to disrupt elections, manipulate facts, and defame public figures, which has led to growing concerns in the media and politics. The lack of legal and policy regulations to address these threats adds further complications (Chesney and Citron 2019, p. 1819).

1.2. KNOWLEDGE GAP AND RESEARCH QUESTIONS

Current research has already investigated the extreme consequences that deepfakes have on individuals and shown that they are a threat to democracy. This was primarily done by examining legal regulations and referring to the ethical stage that deepfakes may harm. What has not been researched is the discourse held about deepfakes in the public sector and newspaper coverage, and how this threat to democracy is critically discussed within different countries. Researchers also state the necessity of conducting more research on the usability of deepfakes (Diakopoulos and Johnson, 2019, p. 2084). By focusing on current discourses, this research aims to add a new perspective to the existing research that has not focused on the use of language, even though it is quite an important part of shaping opinions in society. This is of great importance because if society and politics are not discussing the significant impact deepfakes have on society, especially the believability of online information, the threat of deepfakes can easily spread without addressing the horrific consequences they might pose to democracy. To fill this knowledge gap, the aim of this interpretative research is to examine *in what ways the relationship between democracy and deepfakes is discussed in the countries of Austria and the UK*.

This is an important research question because research has stated that deepfakes can pose a threat to democracy, and if countries do not take this danger seriously, it can harm their democracies in the future. Since public discourse has a massive impact on individuals' opinions and perceptions of potential dangers, it is necessary that the discourse aligns with the current threats identified in scientific research. Additionally, this research contributes to the current state of the art by providing empirical evidence of the existing problem. Comparing two countries is useful for orientation, especially since they are both European countries. Due to the Brexit, there is a significant difference in how independent Austria and the UK are in terms

of regulation. The media, in this research, namely newspaper reports, plays a significant role in the societal context of democracies as it shapes opinions and builds the framework for politics. Furthermore, political discussions are held within newspapers, which mirrors the development of discussions due to the steady reporting of new findings (Schneider and Toyka-Seid 2023). Therefore, this can influence and frame the discourse within society, making it crucial to answer **sub-question one**: *How has a public discourse on deepfakes developed within newspapers in Austria and the UK?* The governmental perspective is of great importance since politicians need to ensure that democracy is protected, which means being aware of new dangers. Additionally, the government can regulate and change laws to protect its country. It is crucial to know how the discussion has developed and if governments take deepfakes as an increasing danger for democracy. Therefore, **sub-question two** is: *How has the policy discussion towards deepfakes developed within the national governments in Austria and the UK?*

To put the results in a broader context, it is interesting to compare the UK with Austria. Comparing two politically similar countries offers more insights, whether deepfakes are seen differently in newspaper discourse or political discussions. Furthermore, the EU/non-EU membership might show that the EU has an influence on how topics are discussed. **Sub-question three** and **sub-question four** are, therefore: *What are the similarities and differences in public discourse in terms of the meanings attributed to deepfakes within the national newspaper coverage in Austria and the UK? What are the similarities and differences in the policy discussion in terms of the meanings attributed to deepfakes within both national governments in Austria and the UK?*

1.3. RESEARCH APPROACH

This bachelor's thesis aims to gain new perceptions to determine if the latest research, which states that deepfakes can pose a threat to democracy, is also a realistic threat for governments and in newspaper coverage in discussions about deepfakes. This insight is important since it is of great interest to consider deepfakes as a potential danger to democracy and to actively work towards preventing or decreasing possible threats. The Frankfurt School, which is used as a theoretical framework for this research, already had a critical perspective on the impact of problematic technology on society. Furthermore, a discussion is the first step towards acting and preventing a threat. Therefore, this thesis will examine the discussions about deepfakes and their relationship towards democracy through a qualitative content analysis. Newspaper articles from popular newspapers and available

policy documents will be analysed using a coding scheme with ATLAS.ti to explore the relationship between deepfakes and democracy in discussions. In the end, Austria and the UK will be compared to gain stronger insights into the deepfake discussion. A content analysis can answer the research question because it is important to focus on the discussions that are held about deepfakes. This research approach is justified because discourses take place through language. Analysis works best for written texts, and since this research is interested in both political discourse and newspaper discourse, which represent societal discourse, it is necessary to analyse policy and newspaper documents through textual analysis. To achieve this, the paper is structured in the following way: The research questions mentioned above will be theorized and core concepts discussed in section two. Critical theory, more implicit the Frankfurt School, which has a critical approach towards the influence of technology on democracy, will be used to provide a theoretical framework. In section three, the reasons for choosing Austria and the UK as cases will be explained, the method of qualitative content analysis will be described, and the procedure for data collection will be explained. In section four, the results of the analysis will be described and set in context by answering the sub-questions. The final section is the overall conclusion.

2. THEORY

2.1. INTRODUCTION TO CORE THEORETICAL CONCEPTS

Given that a critical content analysis forms the core of this research, and in order to provide a theoretical background to support the research question, critical theory has been chosen as the theoretical framework. Specifically, core concepts of the Frankfurt School scholars Adorno, Horkheimer, and Marcuse, along with the perspectives of Günther Anders, a scholar also associated with the Frankfurt School, will be introduced. The idea that problematic technology can influence society in a negative way and impact core democratic values will be used to examine the relationship between deepfakes and democracy. By generally discussing problematic technology and presenting core concepts of the Frankfurt School and their perspective on the relationship between democracy and technology, the point will be made clear how deepfakes are considered as problematic technology. As the Frankfurt School has consistently adopted a critical perspective on technology, the theoretical concept can support the analysis. Additionally, Günther Anders, who extensively analysed the abuse of technology in his work "Die Antiquiertheit des Menschen", will have his core principles highlighted and placed in the context of the relationship between problematic technology and democracy. Lastly, the theoretical concepts will be applied to the context of deepfake technology.

2.2. THE FRANKFURT SCHOOL AND ITS APPROACH TO TECHNOLOGY

The Frankfurt School, founded including scientists as Horkheimer, Adorno, and Marcuse, is a school of social theory that emphasizes the role of culture and ideology in shaping social reality and calls for radical social transformation. While contemporary critical theory partly ignored technology, the first generations of the Frankfurt School took technology into their ideas (Delanty and Harris 2021). It was concerned about the impact of technology on human life and its role in modern societies (Wiggershaus 1995). Adorno and Horkheimer (1944) saw technology as closely connected with the question of nature itself and linked it to capitalism, where a leading society could evolve, increasingly controlled and manipulated by powerful institutions such as governments and corporations.

2.2.1. THE DIALECTIC OF ENLIGHTENMENT: ADORNO AND THE “MYTH OF PROGRESS”

In the dialectic of enlightenment, which can be seen as a core theoretical concept, Adorno and Horkheimer (1944) argue that the emphasis on reason and progress during the enlightenment has resulted in an oppressive culture where individuals are subordinated to collective interests and technological forces. This concept critiques the "myth of progress" that forms the foundation of modernity and suggests that this myth has led to the creation of a society that is increasingly irrational, destructive, and alienating (Hindrichs 2017). This decelerates democratic processes. The original project of the enlightenment, which aimed to free humanity from ignorance and superstition through reason and science, has transformed into its opposite. Regarding the relationship with technology, they describe it as the knowledge that generates methods to exploit the labour of others. The remaining unknown knowledge serves as tools to "learn from nature how to use it to dominate both nature and human beings completely" (Adorno and Horkheimer 1944, p. 2). Furthermore, technology gains power through those who hold economic positions of leadership in society. However, they did not reduce technology solely to capitalism but identified it as having its own dynamics (Delanty and Harris 2021). Adorno and Horkheimer (1944) expressed concerns about different developments within modernity and for the future. They argued that the negative effects of technology were attributed to the capitalist system, characterized by the dominance of corporations over individuals. Currently, the technology employed in the culture industry is limited to standardization and mass production, disregarding the distinctive elements that once differentiated the logic of work from that of society. Examples include mass media such as radio, television, and entertainment like movies, which consequently discourage critical thinking and individuality (Adorno and Horkheimer 1944, p. 95; Arato and Gebhardt 1977). Furthermore, technology erodes the individuality of humans, leaving them solely focused on pursuing their own purposes.

2.2.2. THE ONE-DIMENSIONAL REALITY: MARCUSE'S CRITIQUE OF TECHNOLOGICAL DOMINATION

In his novel "One-Dimensional Man," Marcuse describes a one-dimensional reality where individuals are subjected to a manufactured consensus (Marcuse 1964). In this reality, their desires and aspirations are shaped and channelled within the existing system. The dominant ideology perpetuates a false consciousness that suppresses radical and transformative thought.

As a result, individuals' critical potential is neutralized, and they become passive consumers and conformists. According to Marcuse, this one-dimensional reality hinders the realization of true liberation and authentic individuality. He advocates for the development of a new form of radical consciousness and a liberating praxis that can challenge and transcend the oppressive nature of advanced industrial societies. Marcuse states, "*[t]oday, domination perpetuates and extends itself not only through technology but as technology, and the latter provides the great legitimation of the expanding political power, which absorbs all spheres of culture*" (Marcuse 1964, p. 158). Moreover, Marcuse describes technology as a tool for creating more effective forms of social control and cohesion, which leads to the understanding that the technological society is "*a system of domination*" (Marcuse 1964, p. 8).

Given the totalitarian characteristics of this society, the conventional belief in the "neutrality" of technology can no longer be maintained. Technology itself cannot be separated from its intended purpose; it is ingrained in the very conception and implementation of techniques. Within the realm of technology, culture, politics, and the economy converge into an all-encompassing system that either absorbs or rejects alternatives. The productivity and capacity for expansion within this system serve to stabilize society and confine technological advancements within the boundaries of control. Technological rationality has become synonymous with political rationality (Marcuse 1964, p. 10; Luke 2000). In this realm, technology serves as a significant justification for the lack of freedom in humanity and demonstrates the "technical" impossibility of achieving autonomy and self-determination (Paddison 2017). This state of unfreedom is not perceived as irrational or political but rather as a surrender to the technological apparatus that enhances comfort and boosts labour productivity. Consequently, technological rationality defends and reinforces the legitimacy of dominance, while the instrumentalist perspective of reason leads to a society characterized by rational totalitarianism (Marcuse 1964, p. 92).

2.2.3. PROBLEMATIC TECHNOLOGY AND ITS IMPACT ON DEMOCRACY

Overall, the Frankfurt School believed that technology was a double-edged sword. While it had the potential to liberate and empower people, it was also a tool of domination and control. They argued that it was essential to create a society in which technology was used for the benefit of all, rather than just the few who held power. In the context of democracy, the Frankfurt School mainly defines mass media (television, radio, etc.) as problematic, but also machines that prioritize efficiency and profit. They view these technologies as

dehumanizing and reifying individuals. The relationship between what they call *problematic* technology and democracy can be described as follows: through technology, humans are subjected to domination, which supports a totalitarian system. Therefore, technology poses a threat to democracy because humans should be free individuals and not trapped in a totalitarian system. Furthermore, technology diminishes critical thinking and individuality, which undermines a core principle of democracy, namely pluralism. In this regard, technology is used to reject and suppress alternatives to the totalitarian system, further harming pluralism. By utilizing technology as a tool for social control and cohesion, the relationship with democracy becomes problematic because freedom is a core principle of democracy and should only be limited for justified reasons, not for social control. As Marcuse describes, technology leads to a system of rational totalitarianism, which stands in stark contrast to democracy. In summary, the relationship between what the Frankfurt School identifies as problematic technology and democracy is highly challenging, as technology undermines core democratic values. However, the first generation of the Frankfurt School had its limitations. Therefore, it is crucial to examine the relationship between technology and democracy from a different perspective.

2.3. GÜNTHER ANDERS AND THE ABUSE OF TECHNOLOGY

Günther Anders was a German philosopher who wrote extensively on the relationship between technology and society. One of his key concepts was the "abuse of technology", which he argues was a pervasive and dangerous phenomenon in modern societies. In his book "Die Antiquiertheit des Menschen" (Anders 1956a/b), he describes how technology already interferes in human lives, independently influencing human behaviour, which he sees as a significant danger.

2.3.1. PROMETHEAN GAP AND SHAME

By introducing the principle of the "promethean gap", he explains what he sees as problematic circumstances of industrialization. The "promethean gap", which means that there is a gap between *"what humans can produce with the help of technologies and the capacity to imagine the negative effects these technologies can have"* (Fuchs 2017, p. 582), is one of Anders' core concepts. Anders argues that this could lead to a sense of disorientation and alienation as people become increasingly dependent on machines and lose touch with their physical and emotional experiences. In addition, "promethean shame" means that humans seek to become identical with machines and, therefore, lose their humanity (Fuchs 2017, p. 589), and is strongly connected to the "promethean gap". The apocalyptic potentials of technologies

and society have become so large that the “promethean gap” creates apocalyptic blindness (Anders 1956, p. 233-308). In that regard, Anders argues that machines have become pseudo-persons because humans are fallible and cannot be trusted (Anders 1962, p. 504). He continues that as machines become increasingly capable, powerful, and pervasive, the sense of being a mere human, particularly in the realms of work and social engagement, becomes increasingly demeaning and degrading. Additionally, he argues that the power of technology overshadows human thinking and acting, posing a threat to the democratic values of freedom. Consequently, he argues that humans lose their connection to reality because they mainly deal with technology instead of interpersonal relationships (Kluge 2014).

2.3.2. GÜNTHER ANDERS PERSPECTIVE ON PROBLEMATIC TECHNOLOGY

Anders' perspective on technology is specifically related to nuclear technology. By applying his perspective to 21st-century technology, specifically AI, the following can be argued: Anders highlights a disparity between the creations we produce and our capacity to envision their impact. Consequently, there is a failure to assume responsibility for the repercussions generated by our artifacts. This concept of a "promethean gap" signifies an increasing divide between the technologically mediated effects we have on the world - socially, politically, and economically - and our capability to empathize, contemplate, assess, and emotionally engage with the consequences of our creations (Müller 2016, p. 12). The "promethean gap" encompasses various aspects: it represents a disparity between our capability to produce and our capacity to envision the implications of our creations. It also signifies a disconnect between our proficiency in utilizing technological products and our ability to emotionally connect with actions mediated by technology. Furthermore, it reflects an expanding divide between knowledge and genuine understanding. As Anders suggests, we may possess the ability to invent and manufacture immensely powerful nuclear weapons, yet we struggle to fully grasp the catastrophic consequences that their utilization would entail (Schwarz 2019, p. 104). Converting his perspective to the technological development of 21st century, one can see that the human world becomes fully technologized and datafied. As Anders sees the dehumanization of technology, one can argue that unless the social, economic, and environmental impact of technology is taken into account, it will become a bigger challenge in the future (Schwarz 2019, p. 105 f.). Therefore, the capitalist logic of efficiency and profit reigns in technological development, which not necessarily harms democracy but accepts a certain dehumanization and loss of pluralism.

2.4. THEORETICAL APPROACH IN RELATION TO DEEPPFAKE TECHNOLOGY

The deepfake technology has developed rapidly in the 21st century. In this section, the perspectives of the Frankfurt School and Günther Anders will relate to the concept of deepfakes and how the critique that technology undermines democracy by these scholars is applicable to deepfakes.

2.4.1. FRANKFURT SCHOOL AND ITS APPROACH TO DEEPPFAKES

It can be argued that deepfake technology exploits the labour of others, not in the sense that it eases and therefore devalues the work of humans, but rather that it exploits the integrity of popular people by faking their arguments or statements. Furthermore, one can argue that through mass media and therefore also deepfake videos on (social) media, critical thinking is decreasing. Humans quickly believe videos without questioning their content. Due to easy access to information, deepfakes can undermine democracy even faster because false information is shared, and critical thinking is marginalized. Additionally, deepfakes might contribute to the decrease in the pluralism of opinions, as sharing content without critical thinking might lead to the unification of opinions on certain topics. This aligns with Marcuse's argument of manufactured consensus that can be achieved through deepfakes, which specifically aim to spread a certain opinion and image of the person in the video. Technology impedes the realization of authentic individuality, and this also applies to deepfakes. Individuals are unable to differentiate between real and fake, resulting in the loss of authenticity. Through social control, deepfakes could blur opinions, and the technology itself is not neutral but biased towards a particular political position which can harm democracy. Furthermore, by imitating humans, deepfakes contribute to dehumanization because they demonstrate that humans are replaceable in the digital space, dehumanizing individuals and objectifying human beings.

2.4.2. GÜNTHER ANDERS AND HIS APPROACH TO DEEPPFAKES

Deepfakes can be viewed as an "abuse of technology," as stated by Günther Anders. Undoubtedly, deepfakes are affecting human behaviour. His claim that humans are losing their connection to reality can be applied to deepfakes. Not only is recognizing deepfakes challenging, but the credibility of what some people say or how they behave is no longer questioned. This implies that reality is lost for some people who consume deepfakes. Anders refers to the problem that human beings do not stop and see boundaries and that not everything that can be produced should be produced ("promethean gap"). This can be connected to

deepfakes. Although humans *could* produce deepfakes, it does not mean that it *is* necessary to produce them. The question arises as to what deepfakes contribute to society and whether they cause excessive harm. His argument of machines becoming human is also applicable to deepfakes because deepfakes can display and imitate real human beings, which is degrading to individuals. Through deepfakes, humanity becomes replaceable in the digital world, and thus the “promethean shame” can also be applied to deepfakes as a critical technology. Once again, the argument of pluralism can be extended to deepfakes and their ability to steer content in one political direction. In conclusion, Günther Anders' theoretical assumptions can be applied to the concept of deepfakes, illustrating his arguments with real-life examples.

2.5. CONCLUSION

This section has presented a theoretical framework for examining the relationship between democracy and deepfake technology. As it can be seen, the Frankfurt School and Günther Anders argue that technology is replacing human tasks and influencing human thinking and behaviour, resulting in a loss of power and capabilities. This poses a threat to democracy because it limits individual freedom and could restrict freedom of speech if critical thinking is suppressed. Critical theorists view technology as a danger resulting from capitalism, which oppresses many people within the system. Günther Anders sees technology as an abuse that exerts a dominant power over human behaviour and can overshadow human thinking and acting. Consequently, humans lose their connection to reality and to emotional and physical experiences. In this theoretical context, deepfakes can be seen as potential technological abuse, as their influence blurs the line between real and fake news, leading to people believing a lot of what they read online. Therefore, concerning the research question, a theoretical answer suggests that the relationship between democracy and deepfakes should be viewed critically. Deepfakes impact human freedom, and particularly regarding fake news and disinformation, they can manipulate people's perceptions. Hence, deepfake technology poses a threat to democracy, and the relationship is more negatively influenced than positively. In the analysis, this theoretical answer will be evaluated to determine whether there is empirical evidence supporting these assumptions. In the next section the methodology will be introduced for operationalizing the given theoretical assumptions.

3. METHODS

This bachelor thesis aims to examine the relationship between deepfakes and democracy as discussed in Austria and the UK. The deepfake discussion in political spheres began 2017, when the Obama deepfake occurred. Therefore, this analysis will use 2017 as a starting point and end with March 2023 as ending point for the analysis. In this section, it will be justified why a qualitative content analysis is the appropriate method to answer the research question. The selected cases of the UK and Austria and the reasons for choosing these countries will be given. The method of data collection will then be described, explaining how the data is collected and why it is appropriate for this thesis. Lastly, the data analysis will be explained, describing the methods, and further explaining the process of the analysis. A coding scheme will be provided to operationalize the theoretical concepts for the analysis.

3.1. CASE DESCRIPTION

In this study, the discussion of deepfakes in newspaper as well as within governments of UK and Austria will be examined to see if the threat of deepfakes is visible. Why is it crucial to compare UK and Austria as two countries? Deepfakes have become a prominent topic in the United States and Canada, and examining their development within Europe, particularly in well-developed countries, is of great interest because deepfakes can interfere in democratic processes. With the European Union (EU) serving as a legally binding institution, the discussions surrounding deepfakes might vary between EU and non-EU member countries. Given the Brexit and UK's reputation as a highly developed yet increasingly liberal society in terms of societal topics, understanding how AI is discussed in the UK is particularly fascinating. Specifically, it is intriguing to compare their AI discussions with those of an EU member country to identify any differences. The UK, on the one hand, is generally receptive to AI technology while also linking it to ethical values. Austria, on the other hand, is also a well-developed country with a conservative government. The country's interest in AI is increasing, and it is actively working on regulations to govern its use. Given that Austria is somewhat reliant on future EU regulations, it is worth exploring whether it is only following the minimum requirements or taking more stringent actions. Comparing these two countries provides an intriguing case study due to their EU/non-EU membership. Additionally, both countries are democracies with conservative governments that are becoming more progressive and liberal on societal topics such as AI technology, making them compelling to compare in terms of their views on AI opportunities and risks. Regarding ongoing research indicating that

deepfakes pose a potential threat to democracy, it is interesting to see whether these countries government's view deepfakes as a danger and whether such concerns are being discussed.

Since the interest relies on these countries, the analysis includes only documents that are about these specific countries. Because the focus is on the discussion of deepfakes, possible identification of patterns, their relationship towards democracy and the meaning deepfakes are given, a content analysis is an appropriate method to answer these research questions. In discussions, language is a key factor to analyse, which implies that a textual analysis fits the research questions. Because the interest relies on the governments' perspective, policy documents from UK and Austria will be analysed by looking for similarities and differences (sub-question two & four). To receive a general overview of the discussion within the countries and how news frame the discourse, newspaper perspective will be included to see how deepfakes are discussed and what similarities and differences occur in both countries (sub-question one & three). To demarcate, only newspaper articles of UK and Austrian newspapers will be used.

3.2. METHOD OF DATA COLLECTION

In a first step, secondary data will be used, meaning that no data will be collected by me. The data sources will include policy documents and newspaper articles. Keywords as "deepfake(s)", "AI", "deepfake policy documents", and "deepfake newspaper articles" will be used for data collection from media and government documents. To find suitable newspaper articles and answer sub-question one, the second step contains an advanced Google search, focusing on established newspapers that have published articles on deepfakes between 2017 and 2023. Due to time constraints, only newspaper articles from one publisher will be used. An advantage of this approach is that it enables an in-depth examination of the development of the deepfake discussion. However, a disadvantage is that focusing on only one newspaper publisher may lead to a lack of variety in the data. Since major developments are often shared among big newspapers, there could be a lot of repetition that might influence the perception of the discussion about deepfakes. However, the results of the analysis of deepfake discussion in newspaper articles will only be applicable to the selected newspaper company. For the UK, *the Guardian* will be used for the analysis, while for Austria, the newspaper *Der Standard* will be used. The Guardian is one of the most reputable newspapers in the UK, funded by its readers and highly independent from political influence (The Guardian, About Us). Similarly, Der Standard is

also considered one of the most trustworthy newspapers and has won several awards for its reporting (DER STANDARD 2023). Therefore, both newspaper publishers can be seen as reliable and trustworthy sources for this analysis.

The search function on each website will be used to filter articles from 2017 to March 2023 that discuss deepfakes. To answer sub-question two about policy documents, both UK and Austria have a snapshot paper/Aktionsplan Deepfakes which will be the source of analysis. This will give an insight on how deepfakes are discussed in a governmental environment. These documents reflect the theoretical concepts of critical theory because they are critical of deepfakes as an AI technology. Furthermore, it needs to be considered that newspaper articles often polarise, and a misrepresented picture of the reality can be shown. However, if there is a societal discussion in the selected countries on how deepfakes poses a threat to democracy, it is most likely to find them in newspaper articles. Policy documents might see deepfakes as a potential threat to democracy, reflecting the democracy/technology relationship explained in the theory section. Therefore, it is crucial to see the political perspective on deepfakes.

3.3. METHOD OF DATA ANALYSIS

A content analysis is defined as *"the intellectual process of categorizing qualitative textual data into clusters of similar entities, or conceptual categories, to identify consistent patterns and relationships between variables or themes"* (Julien 2008, p. 120). It is independent of theoretical perspective and can be used in both quantitative and qualitative analyses. In qualitative analysis, that is used here, content analysis aims to answer the "why" question and therefore involves interpretation. This means that *"text is open to subjective interpretation, reflects multiple meanings, and is context dependent"* (Julien 2008, p. 120). Content analysis can also identify conscious and unconscious messages within a text, and through coding, researchers can interpret not only the concepts that are present in the text but also the absence of certain words. For a robust content analysis, validity and reliability are core principles (Krippendorff 2004; Mayring and Fenzl 2014). For qualitative analysis, researchers seek trustworthiness and credibility. This can be achieved by involving more than one researcher and having them independently analyse the data to ensure reliability. Textual analyses are often conducted using research tools. In this thesis, ATLAS.ti will be used. ATLAS.ti is a comprehensive tool for analysing data, serving as a repository for various components, such as primary documents or their links, quotes, code words, notes, memos, links, and more (Friese 2012). It replaces the traditional analogue way of using pencil and paper to facilitate textual analysis. With content analysis, the above-mentioned data will be analysed. This is done with

the help of a coding scheme. Coding in the context of data analysis is understood as “*representing the operations by which data are broken down, conceptualized and put back together in new ways*” (Flick 2009, p. 307).

By using this methodology, it becomes feasible to categorize thematically related items and break down extensive units of analysis, which results in structured and significant outcomes. For a clearer overview, the coding scheme is divided into German and English terminology, as documents from both languages are used. However, it is not a different coding scheme. The key concepts in the thesis are “democracy” (Figure 1/2) and “deepfake technology” (Figure 3/4). To make these concepts measurable, three key features are identified which will then be operationalized to two keywords each (see figures). The collected data will be analysed in ATLAS.ti using the provided coding scheme. The coding scheme will filter out significant parts of newspaper articles and government documents where the use of deepfakes and their relationship with democracy can be identified. The relevant features of democracy include *harming freedom, threat for democratic processes, and tools to provide safety*. The chosen keywords aim to observe harm, threats, or protection of these features, with *freedom, democratic processes, and safety* declared as essential features of democracy. In terms of deepfakes, the features of *general use, AI technology, and information provider* will be examined. In alignment with Anders (1956a/b) and Adorno and Horkheimer (1944), the coding scheme is based on the concepts of deepfakes as a problematic technology. These features do not refer to deepfakes as a tool with a specific purpose, but rather as a technology that can be used in various ways. Nevertheless, specific keywords are tailored to identify patterns that may indicate how deepfakes could harm the democratic features mentioned earlier. The next step will involve further analysis of these parts to determine how key concepts are discussed. Identifying frequently used keywords may reveal patterns that can be addressed later in the analysis. Given that the main research question examines the relationship between deepfakes and democracy, it is essential to investigate this relationship in detail. To achieve this, the context will be examined in-depth to identify instances that suggest a relationship between deepfakes and democracy.

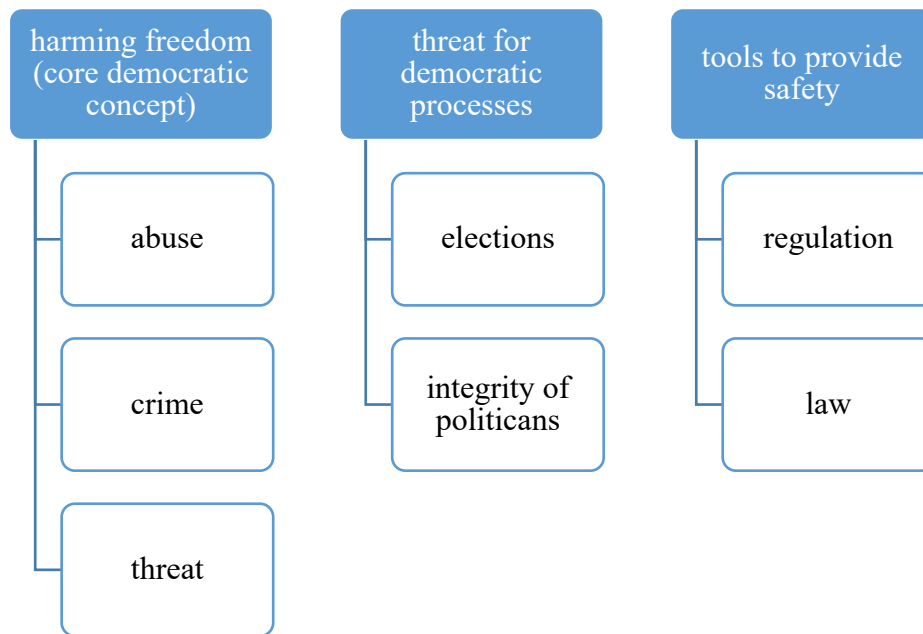


Figure 1: Coding scheme for “democracy” concept (English)

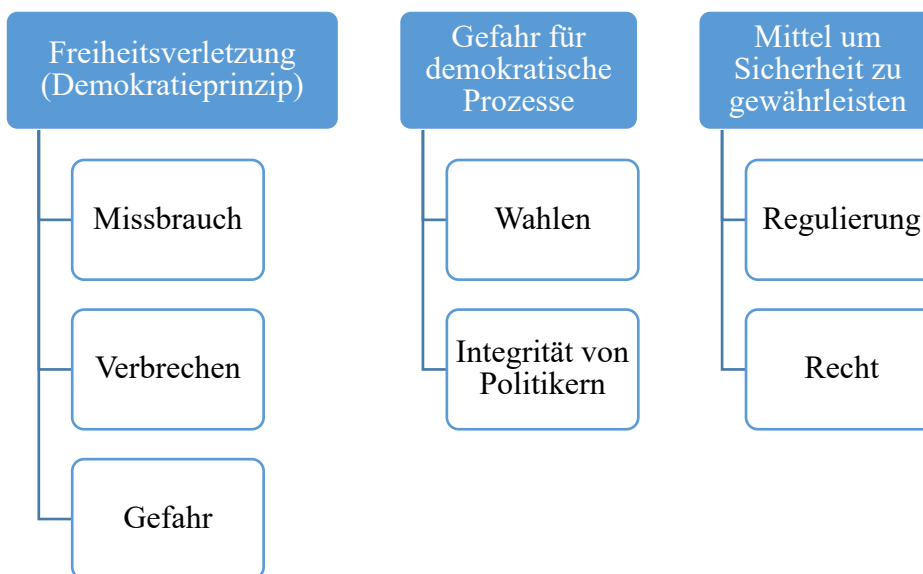


Figure 2: Coding scheme for “democracy” concept (German)

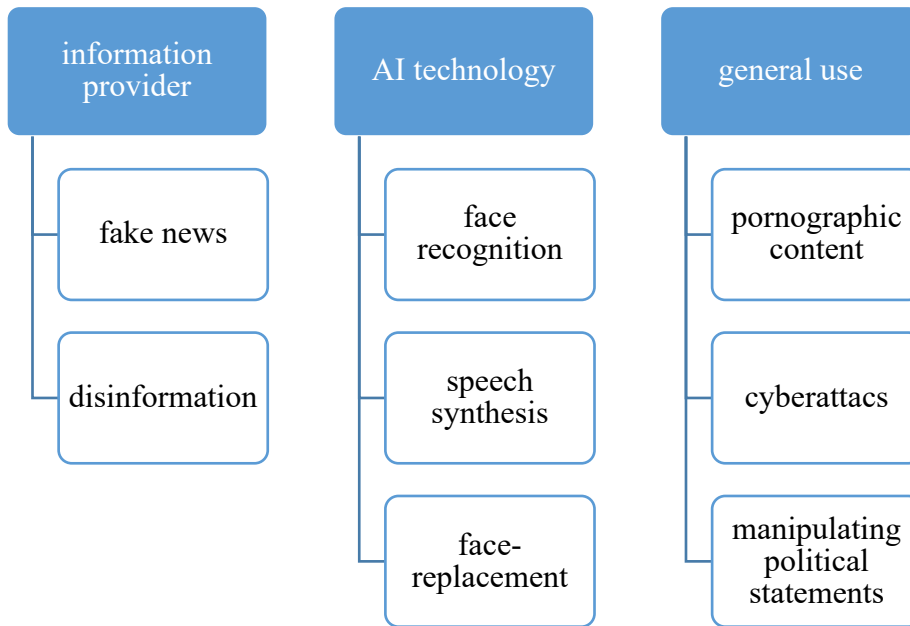


Figure 3: Coding scheme for “deepfakes technology” concept (English)

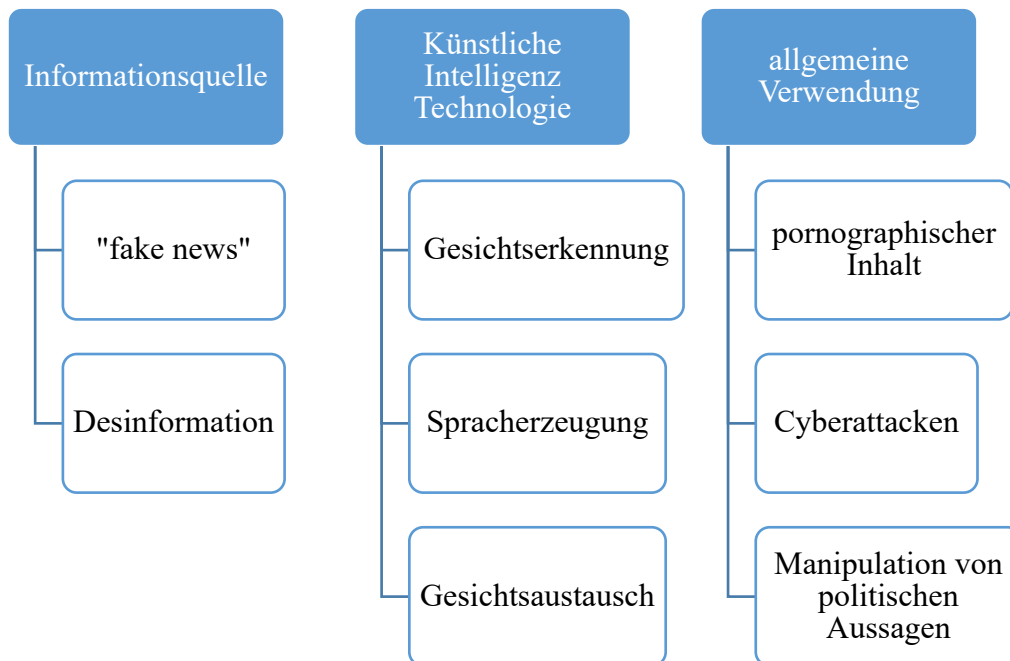


Figure 4: Coding scheme for “deepfakes technology” concept (German)

3.4. CONCLUSION

To answer the research questions, the newspaper articles will be analysed using the coding scheme in ATLAS.ti to determine how media is framing the discussion of deepfakes (answer to sub-question one). A similar approach will be taken with the political documents, using the coding scheme to figure out how the relationship between democracy and deepfakes is described, as well as to identify any upcoming future strategies (answer sub-

question two). Once the analysis is complete and the two sub-questions are answered, a comparison of the cases in Austria and the UK can be made to determine if they have different discussions about the relationship between deepfakes and democracy (answer sub-question three/four). In the end, the main research question can be answered by summarizing the results of the four sub-questions. To conduct the analysis, the coding schemes for democracy and deepfakes will be used to filter out the relevant parts regarding their relationship. Next, the parts will be analysed in-depth, and the results will be collected to offer a complete picture in the end. In the next section, the results of the analysis will be explained, and the sub-questions will be answered.

4. ANALYSIS

4.1. INTRODUCTION

In the following chapter, the results of the textual analysis will be published, separated by each sub-question. The analysis wants to show that, in accordance with the core concepts of the Frankfurt School, the discourse of deepfakes is in both, policy documents and newspaper coverage, negatively connotated. The point that should be made here is that deepfakes poses threats to democracy. In other words, deepfakes are an empiric danger and the consciousness of this threat is developing rapidly. By comparing Austria and UK with each other, new perspectives are given on how two European countries see deepfakes as a general similar danger, but the discourse and the impact on strategies is differing significantly. Also, surprisingly, both countries come to the same outcome that deepfakes pose as threat to democracy, however, the reasons for this conclusion are partly similar, partly different ones. By first examining the discourse development of each country separately, an in-depth insight will be given on the key results of the analysis. By comparing the discourses in a next step with each other, more generalized observations can be presented which entails to answer the research question properly.

4.2. DEEPFAKE DISCOURSE IN NEWSPAPERS: AN INCREASING FEAR?

4.2.1. DISCOURSE IN AUSTRIA: FROM FEAR TO REGULATION

Surprisingly, it can be said that the discourse in Austria has developed into two directions. On the one hand, in the last few months, recent events have brought deepfakes more into focus, highlighting their danger in the political sphere. On the other hand, numerous articles discuss new technology/software that aims to prevent people from believing deepfakes. It is interesting to note that initially, when the discourse on deepfakes began, there were no available technologies to detect them. However, now several software programs have been developed with a recognition rate between 94% and 97.5% for detecting deepfakes (Der Standard 2022c). However, shocking results indicate that human recognition of deepfakes has not improved over the past years. *"Participants were correct only 48% of the time, which is slightly worse than what would be expected from random guessing (50:50)"* (Der Standard 2022h)*. This suggests that individuals are not prepared or aware of deepfakes and how to

*The original citations are in German and can be found in the appendix.

deal with fake news. Furthermore, it aligns with the concept of the "myth of progress," which leads to irrational human thinking. It also resonates with Anders' perspective on technology overshadowing human thinking. This leads to the main point that politicians are also not aware of the threat by deepfakes. An unpleasant incident involving the mayor of Vienna has raised questions about the competence of politicians and the level of preparedness in Austria to protect itself from deepfake attacks. The mayor accidentally communicated with a deepfake of Witali Klitschko. It is unclear whether crucial information was leaked, but it shows that the government is ill-prepared. This reflects the developments of the last two to three years, where deepfakes have become a major concern for the political sphere and are not only used for pornographic content but also for spreading political fake news.

While in 2020 the consensus sounded as a danger that is still far away: *"It becomes dangerous and offensive as soon as it is used to fake political speeches or pornographic material"* (Sommavilla and Stajić 2020), by May 2022 it had shifted to: *"Deepfakes are used to manipulate democratic processes. Involving key figures in politics or the economy poses a significant security risk because the identification of artificial manipulation is difficult to prove or trace"* (Der Standard 2022f). It is comical that deepfakes are now viewed as a democratic threat, whereas a few years ago, they were only considered dangerous. Horrifically, while in 2018 there was concern about using deepfakes in war: *"Imagine what would happen if these videos ended up online and spread rapidly. [...] In this scenario, a 'negative actor' falsifies geopolitical events to achieve certain goals"* (Schmid and Al-Youssef 2018), now they are being used in the Ukrainian war: *"Through 'digital warfare,' it seems that the aim is to shake trust in politics and discredit Ukraine and its supporters"* (Scherndl 2022b). For instance, this aligns with Günther Anders' concept of technology abuse. While he originally referred to nuclear weapons, now an online weapon in the form of deepfakes has emerged. Generally, the language used to be fearful and alarmist, now it is still alarming but in an objective and calm manner. Furthermore, there has been no regulation from the political side, which changed in 2022 and explains the recent different approach. Surprisingly, even though the threat is increasing, the discourse has shifted from alarmingly presenting deepfakes as a great fear to discussing how to regulate, prevent, and minimize the harm they pose to democracy.

4.2.2. DISCOURSE IN THE UK: FROM FEAR TO ACTION

The discourse development in The Guardian started in late 2018/2019, where deepfakes were seen as a big danger to democracy. An expert says, *"[deepfakes] could be weaponized in ways that weaken the fabric of democratic society itself"* (Schwartz 2018), which confirms the

assumption that deepfakes are a threat to democracy. However, this sentiment is changing over time. One year later, it is mentioned that deepfakes have a "*subversive potential*" (Hunt 2019) according to an expert. At the same time, he says, "*deepfakes pose less of a threat than fake news articles, which are harder for the platforms to detect*" (Hunt 2019). This is a surprising development since deepfakes are often seen as a better version of fake news. In 2022, the consensus is that deepfakes are dangerous, but the language used is settled and calmer: "*Deepfake political videos are also a fast-growing issue. There are worries about how they might be used to spread misinformation and influence elections*" (Mahdawi 2022). Therefore, there are worries about the impact of deepfakes on democracy, but since the government is also dealing with deepfakes, it has become less of a dramatic debate. Moreover, some experts see deepfakes not as a threat to democracy but as a main issue of non-consensual use for pornographic content, even in 2022: "*[...] we should generally be concerned about this tech, but [...] the main problem with deepfakes today is their use in non-consensual deepfake pornography, rather than information*" (Blackall 2020).

Also, a remarkable point is how the impact of deepfakes is discussed. While some experts fear that the videos and content themselves are harming democracy, a new phenomenon is described: "*The problem may not be so much the faked reality as the fact that real reality becomes plausibly deniable*" (Sample 2020). If there is the possibility of everything being fake information, individuals will become sceptical about online content, which might lead to the development of people "choosing" what they see as real and false information. This could raise serious issues, also in a political manner. For instance, this poses the opposite of Anders' concept of technology making humans irrational and instead strengthens human scepticism. However, the outcome of destabilization still aligns with his theoretical concept. Another development is that deepfake fails in other countries are mainly discussed, whereas the focus in the UK lies on the fact that government dealt with deepfakes early (Chivers 2019). This is also used to distance itself from the EU and justify the outcome of Brexit, but also to take the role of a nation that is in control of the situation. Interestingly, in one article, the critique on the extent to which the newest technological possibilities should be followed is mentioned. "*As well as asking how we stop deepfakes, we need to ask why someone thought they'd be a good idea to begin with*" (Beard 2019). This not only shows how the fear of deepfakes is developing: it conforms with Anders concept of "promethean gap" and for instance the necessity of using every technological potential. All in all, the development of deepfake discourse is calming down from more dramatic phrasing. Nevertheless, there is a disagreement among experts on how dangerous deepfakes are, especially concerning the

political sphere. Being able to establish regulations as one of the first countries is something that constantly stands out in the debate about deepfakes.

4.3. DEEPFAKE DISCOURSE IN GOVERNMENTS: REGULATIONS ARE NECESSARY

In the following section, the analysis of the leading policy documents of Austria and the UK will be presented. It shows the discourse development within the governments. The development will be presented by introducing the topics developed in the coding scheme.

4.3.1. AUSTRIAN'S GOVERNMENT: "DEEPFAKES AS A MAJOR THREAT"

The most remarking aspect of the policy document is that the Austrian government perceives deepfakes as a much greater threat than something beneficial for society. *"In Austria, the threat posed by deepfakes is real"* (Aktionsplan Deepfake 2022, p. 23), indicating that Austria views deepfakes as a significant threat. This perception is further emphasized by the statement that a *"fight against deepfakes"* (Aktionsplan Deepfake 2022, p. 23) is taking place. This not only employs a strong metaphor to depict deepfakes as a threat but also underscores the seriousness of the issue. In the policy paper, certain goals are presented, one of which is the protection of the integrity of democracy: *"Consistent protection of the integrity of our democracy and the democratic formation of will against external influence"* (Aktionsplan Deepfake 2022, p. 24) is crucial. This statement implies that if the government does not act, deepfakes pose a more realistic threat to democracy. It agrees with the Frankfurt School that problematic technology undermines democracy. However, on the contrary side, there is a warning against excessive regulation, particularly in the context of satire or the art scene. *"The labelling of deepfakes in the field of artistic expression and freedom of speech should ensure that there are no limitations on this form of art or fundamental freedom. Satire is fundamentally not about deceiving, unlike disinformation"* (Aktionsplan Deepfake 2022, p. 10). If deepfakes are used for artistic expression and not for disseminating disinformation, there should be no limitations. The Frankfurt School, for instance, would disagree since especially entertainment technology attacks individuality (Adorno and Horkheimer 1944). This implies that deepfakes are not inherently dangerous, but the threats perceived by the Austrian government are nonetheless significant.

It is also remarkable to note that the policy document strongly emphasizes the protection of individual rights (*Persönlichkeitsschutz*) because democracy is also about protecting individuals. Furthermore, it highlights the impact of deepfakes on the relationship between citizens' trust and democratic processes that can be influenced by deepfakes: *"When these*

elements spread on the internet, they disrupt democratic processes and the trust of citizens" (Aktionsplan Deepfake 2022, p. 12). This clearly indicates that there is a fear of deepfakes undermining democracy in the minds of current politicians and aligns with Frankfurt School thoughts of diminishing democracy. In conclusion, it can be summarized that while Austrian policy documents acknowledge the use of deepfakes in culture and art, they also recognize a much greater threat and danger to democracy. Therefore, regulation, legislation, and a comprehensive strategy are of paramount importance in addressing deepfakes in the coming years.

4.3.2. UK'S GOVERNMENT: "CONCERNING BUT NOT THREATENING "

Striking results show that the extent to which deepfakes are seen as a threat is likely smaller than expected. For example, *"[s]keptics argue that the fears surrounding deepfakes mostly relate to what might come to pass rather than what already has. The reality is that it remains difficult to create convincing forgeries, even for seasoned users of the underlying technology"* (Snapshot paper 2019, p. 12). This implies that more is made of deepfakes than it actually is, and people should not overestimate their abilities. Furthermore, the government states that the ability of deepfakes is too low to make a significant impact: *"We are yet to see a convincing deepfake of a politician that could distort public discourse"* (Snapshot paper 2019, p. 1). Nevertheless, it is clearly stated that existing deepfake videos are used to harm the integrity of politicians (Snapshot paper 2019, p. 13), which aligns with Frankfurt Schools' theoretical approach to problematic technology. In contrast, it is surprising that legislation is not seen as the ideal way of dealing with deepfakes: *"[...] [U]sing legislation to contain deepfakes would be both ineffective and counterproductive: ineffective because it is difficult to identify the makers of deepfakes, with many residing on foreign soil, and counterproductive because new legislation could have the unintended consequence of curbing the use of visual and audio manipulation techniques for socially beneficial uses"* (Snapshot paper 2019, p. 14). Anders would disagree on that and question the benefit of the technology ("promethean gap") Therefore, the fear of regulating too much and misbalancing the relationship between freedom and safety is indeed difficult. Thus, *"detection tools should at least be accurate and accessible to those who need them"* (Snapshot paper 2019, p. 16). This would raise awareness for employees who work extensively with digital content. On the other hand, it would decrease the spread of fake news and disinformation and make media more trustworthy. All in all, the UK government

sees the responsibility to react. However, democratic freedom is a strong ideological part and needs to be protected as well.

4.4. COMPARING DEEPPFAKE DISCOURSES IN NEWSPAPERS: AUSTRIA VS. UK

Firstly, it can be said that the differences outweigh the similarities. In both newspapers, the main topics referring to deepfakes have been the "general use of deepfakes" as a means of *manipulating (political) statements*. References to democracy, "harmed democracy" has been found the most used topic and deepfakes being seen as a *threat* was most found in the analysed articles. This means that both newspapers have a similar focus. Secondly, it is surprising that both newspapers are reporting on the positive ways in which deepfakes can be used. Not only in the art and entertainment sector but also, for example, how they can help in police investigations: *"Dutch police have received dozens of leads after using deepfake technology to virtually bring to life a teenager almost two decades after his murder"* (AFP 2022). To illustrate, deepfakes help in upholding democracy and the rights of individuals, which is an unexpected finding concerning Frankfurt School's theory as seeing in technology only a threat. However, it must be stated that these reports, where deepfakes are seen positively, are in the minority. Thirdly, they both report on the dangers that deepfakes can pose to the political sphere. The articles clearly highlight the initial recognition of the need for regulation when addressing deepfakes. As the discussion progresses, the articles delve into the implementation of regulations and how they are expected to mitigate the threat posed by deepfakes to the political sphere and democracy. This also aligns with the thoughts of the Frankfurt School, which sees technology as a threat to society: *"As well as considering the threat to privacy and national security, both scholars became increasingly concerned that the proliferation of deepfakes could catastrophically erode trust between different factions of society in an already polarized political climate"* (Schwartz 2018). The Austrian newspaper also supports this viewpoint: *"Riparbelli called on governments to enact comprehensive laws to regulate 'synthetic media' and the deepfake industry"* (Der Standard 2023), because *"it is urgently necessary for politics to also engage with deepfakes and develop strategies to protect the affected individuals"* (Der Standard 2020f). Additionally, in both newspapers, there is a general tendency to highlight the negative impact of deepfakes on democracy. However, depending on who is asked in the articles, some experts are less worried than others. This indicates that different opinions are presented, but the extent to which deepfakes harm democracy remains uncertain.

Concerning differences, the results can be separated into two parts: on the one hand, *how* newspapers are reporting about deepfakes in terms of language, and on the other hand, *what* they report. First, the language used in both newspapers is slightly different. Whereas the language in the Austrian newspaper is alarming but neutral, stating, "*deepfakes - such as those of politicians - pose, according to the government, 'a significant security policy risk' and 'constitute a threat to our democracy and societal cohesion'*" (Der Standard 2022d), the UK newspaper uses a lot of metaphorical language that is emotionally charged, such as, "*[t]hey could be weaponized in ways that weaken the fabric of democratic society itself*" (Schwartz 2018). Using dramatic language often has a deeper effect on humans as it appeals to their emotions rather than their rationality. Assuming this, it would mean that the UK's newspaper sees the threat as greater than the Austrian's newspaper. Second, for *what* is reported about, the Austrian newspaper provides a lot of information about the technological process, how to fight and detect deepfakes, which adds to the neutral language used. Moreover, it frequently describes how deepfakes are used, emphasizing that, apart from the few initial political events where deepfakes were used, the main danger lies in non-consensual pornographic content. "*For instance, the faces of famous women are retouched into pornographic material by the thousands. Furthermore, video manipulations often lead to attempts of extortion*" (Der Standard 2021c). In contrast, the focus of the UK newspaper regarding deepfakes is how they can/will or have already influenced elections. "*Critics say the technology can be used to create bogus videos to manipulate elections, defame someone, or potentially cause unrest by spreading misinformation on a massive scale*" (AgenceFrance 2019) or "*[w]e have to inoculate the public before deepfakes affect elections*" (Parkin 2019).

A controversial side note is that Austria primarily reports about Austria itself, whereas the UK newspaper describes numerous scenarios in other countries, as if it wants to distract from its own situation. Nevertheless, Austria mentions other countries mainly in the context of collaboration and the creation of international or European strategies or regulations against deepfakes, while the UK aims to address deepfakes on its own and frequently mentions that it was the first country to publish regulations against deepfakes. This behaviour in news reporting indicates that the UK wants to distance itself from the EU and strengthen the decision of Brexit, which is a logical behaviour but surprising in manners of how to deal with problems that concern the whole world. Overall, it is evident that both newspapers have a generally similar stance, but how and what they report about deepfakes differs. Both see deepfakes as a threat and danger, but also acknowledge their potential as opportunities and entertainment. This partly reflects the concepts of the Frankfurt School,

which sees technology as a unique threat to society and its engagement. These concerns can be found in both newspapers.

4.5. COMPARING DEEPAKE DISCOURSES IN POLICY PAPERS: AUSTRIA VS. UK

Previously, the development of the discussion on deepfakes within the national governments of Austria and the UK was presented. In this chapter, the most prominent similarities and differences will be evaluated by providing examples. According to the coding scheme, there are relatively equal findings within the policy documents regarding the mention of deepfakes as an AI technology, its general use, and its role as an information provider. The same applies to the connection between deepfakes and democracy, which is almost equally addressed in both documents. In the next step, the content similarities will be presented. Firstly, a significant similarity is that both policy documents highlight the danger of deepfakes to democracy. However, in the UK document, it is still more perceived as a *potential* danger, which illustrates the hesitance to call deepfakes a threat to democracy. The Austrian perspective, on the other hand, sees it as a more recent and imminent threat, preparing for potential upcoming scenarios and considering it a greater danger than the UK perspective. For example, the Austrian policy document mentions: *"A false video created using AI to deceive or manipulate with political intent can pose a significant threat to the integrity of a democracy. As an example, consider a video showing a head of state or a government official saying things that subsequently lead to mass demonstrations and a government and state crisis. When these elements spread on the internet, they disrupt democratic processes and the trust of citizens"* (Aktionsplan Deepfake 2022, p. 12).

Similarly, the UK document expresses concerns: *"Deepfakes are viewed by many as a critical threat to individuals and society. 'You thought fake news was bad? Deepfakes are where truth goes to die,' reported The Guardian in November 2018. 'Deepfake videos threaten the world order' ran a headline in The Times in February 2019. Whether or not these fears are justified will depend on the quality of deepfakes, how they are used in practice, and the ability of the public to tell fact from fiction"* (Papershot 2019, p. 10). This is controversial and illustrates how technology can influence rational thinking, which aligns with Adorno and Horkheimer. Additionally, it is interesting that both countries emphasize that the use of deepfakes in satire or art contexts should not be limited, which contradicts the stance of the Frankfurt School scholars who explicitly state that this entertainment sector poses a significant threat to democracy (Adorno and Horkheimer 1944). Both countries acknowledge the importance of technological development as a valuable addition to the entertainment sector.

Furthermore, they recognize the need to act against the use of deepfakes as providers of disinformation and fake news. They are proactive in preventing potential future incidents rather than waiting and reacting to potential concerns.

Secondly, they also share the fear that excessive regulation might be counterproductive or ineffective, with the UK particularly emphasizing the importance of minimizing restrictions on freedom. Moreover, both countries acknowledge the difficulty in finding regulatory solutions that are effective and practical. It is challenging to identify the creators of deepfakes and establish a causal relationship between deepfakes and specific outcomes, such as election results. There, Anders' perspective on the necessity of using all technological resources without questioning the consequences serves as an argument against further establishing deepfakes in the societal context. Nevertheless, there are also differences between the UK and Austria. In the UK policy document, deepfakes are defined as *"visual and audio content that has been manipulated using advanced software to change how a person, object, or environment is presented."* (Snapshot paper 2019, p. 3) In the Austrian document, deepfakes are defined as *"an umbrella term for various forms of audiovisual manipulation, including video, audio, or both. Typically, AI-based technology is used to create deepfakes. Deepfakes are perfectly faked videos, images, or audio in which individuals are made to say things or appear to engage in actions that never actually took place."* (Aktionsplan Deepfake 2022, p 8). Here, the definitions differ significantly because the Austrian definition assigns a more dangerous value to deepfakes, while the UK's definition is objective and neutral. This indicates that there is still no clear vision of the potential threat that deepfakes pose, which aligns with current research (Hameleers et al. 2022).

Additionally, the Austrian definition implies that deepfakes are only used for "manipulation," although later it is mentioned that they can also be used e.g., in the art sector. In the Austrian document, it is also referred to "perfectly faked videos," suggesting that there is no chance of identifying them or that the technology has advanced to the point where it poses a significant threat to democracy. Moreover, both countries have different approaches to regulations. Austria aims to expand research and education, establish task forces, and implement legal regulations for future incidents. In contrast, the UK is more focused on detecting deepfakes and educating the public. The UK also considers not only the government but also the media and research community as responsible actors in the process. Another crucial difference is that Austria's language is much more alarming than the UK's. The UK expresses a cautious approach, avoiding excessive regulation and waiting

to assess the harm caused by deepfakes. *"However, there is still much that is unknown about the effectiveness of different interventions, including their potential for unintended consequences. The government, tech companies, and media forensics specialists must continue exploring and piloting new containment measures, while being mindful not to squeeze out beneficial uses of audio and visual manipulation"* (Snapshot paper 2019, p. 17). Austria, conversely, expresses a high level of fear and urgency. It says *"In Austria, the threat posed by deepfakes is real"* (Aktionsplan Deepfake 2022, p. 23), emphasizing the need for solutions as soon as possible to protect democracy. Austria also emphasizes *"the protection of democracy, the protection of the individual, national security, and technological developments"* (Aktionsplan Deepfake 2022, p. 25) as priority.

This underscores the urgency of acting and aligns with the assumption that deepfakes pose a threat to democracy. Furthermore, Austria actively seeks international and European solutions and cooperation, whereas the UK's snapshot paper does not mention international cooperation at all. Overall, the UK and Austria share many similarities in recognizing how deepfakes can harm democracy and the dilemma of balancing citizen protection with the freedom to use this technology. The significant difference is that the UK has a more relaxed approach, while Austria sees deepfakes as a major threat, particularly to its democracy. Therefore, Austria's inclination towards regulation is stronger than that of the UK. However, both countries view deepfakes as potential threats to democracy, and the policy papers represent a first step in regulating and protecting democracy.

4.6. DO DEEPFAKES RAISE A THREAT TO DEMOCRACY?

To answer sub-question one regarding the *development of public discourse on deepfakes within newspapers in Austria and the UK*, the analysis reveals that the discourse in the Austrian newspaper has transitioned from an alarming state to a more objective and calm tone, focusing on exploring technological solutions. With the gradual implementation of regulations in Austria, the urgency for government intervention has diminished. The newspaper now reports on how existing regulations may prevent or address deepfake incidents, such as the deepfake Klitschko talk with Vienna's mayor. Similar observations have been made for the UK. The discourse has shifted from a dramatic language emphasizing the threat deepfakes pose to democracy to a more composed tone. However, an ongoing discussion remains regarding the extent of the danger deepfakes pose to democracy, particularly in relation to government regulations. In other words, the discourse in both countries has become less fearful and more focused on understanding and managing the danger. Nevertheless, deepfakes have gained

greater prominence due to the increasing threat they pose. Differences lie in their approach and coverage of deepfakes. The UK predominantly addresses deepfakes as a threat to elections and society, while Austria still perceives the primary danger to be the impact on women and pornographic content. Both countries acknowledge deepfakes as a potential threat and hazard, but they also recognize their possibilities and entertainment value. Therefore, sub-question three, which investigates *the similarities and differences in the meaning attributed to deepfakes within national newspaper coverage in Austria and the UK*, is answered.

Regarding sub-question two on the *development of policy discussions on deepfakes within the national governments of Austria and the UK*, the analysis suggests that Austrian policy documents not only acknowledge the cultural and artistic applications of deepfakes but also express heightened concerns about their impact on democracy. Consequently, prioritizing regulation, enacting legislation, and implementing comprehensive strategies are deemed crucial for effectively addressing the challenges posed by deepfakes in the future. In the UK, the government recognizes the need for a proactive response. Deepfakes are perceived as a present danger to democracy, and the government emphasizes the importance of preventive strategies, including education and enhanced detection capabilities. However, the preservation of democratic freedom, which holds significant ideological value, is also emphasized, particularly regarding the positive use of deepfakes in the entertainment sector. To address sub-question four regarding *the similarities and differences in the policy discussions and the meanings attributed to deepfakes within the national governments of Austria and the UK*, it can be said that both countries acknowledge the detrimental impact deepfakes can have on democracy and the challenge of balancing citizen protection with the freedom to utilize this technology. However, a notable distinction arises in their attitudes. The UK adopts a relatively more lenient approach, while Austria perceives deepfakes as a significant and specific threat, particularly to its democratic processes. As a result, Austria demonstrates a stronger inclination toward implementing regulations compared to the UK. Nevertheless, both countries recognize deepfakes as potential threats to democracy, and their policy papers serve as initial steps toward regulating and protecting democratic systems.

5. CONCLUSION

5.1. THE DISCUSSION OF DEMOCRACY AND DEEPFAKES

So, what is the answer to the main research question of *how the relationship between democracy and deepfakes is discussed in Austria and the UK?* In a policy context, deepfakes are recognized as having a significant impact on democracy. They not only harm the trustworthiness of political statements but also pose threats to national security, elections, and the integrity of politicians. These issues are being discussed within the governments and covered in newspaper reports in both Austria and the UK. The abuse of non-consensual pornography created with deepfakes also raises concerns about individual rights and the potential distortion of society by deepfake technology. Additionally, the newspapers highlight instances where deepfakes have harmed political discussions, deceived politicians, and even been used as weapons in conflicts like the Ukrainian war to spread fake news and disinformation. While there are minor discussions on how deepfakes can strengthen democracy, such as their use in crime investigation or identifying fake news and disinformation, it is important to remember that deepfake technology can also be used in the entertainment sector, adding value to pluralism, and supporting the "free development of personality," which is an important individual right.

However, overall, the threats posed by deepfakes outweigh the positive aspects, as evident from the focus of political and newspaper discussions. The current developments call for continued vigilance regarding the rapid advancement of this technology. In terms of the discussions on the relationship, it can be argued that the initial warnings and pessimism have transitioned into a more serious recognition of the threat, without overreacting or instilling fear in society. The emphasis is on redirecting attention towards the technology and educating society to be aware and adopt a critical view of online content, particularly in the political sphere. In summary, the approaches taken by Austria and the UK differ. While they both acknowledge key dangers and threats to democracy, their reactions and language used vary. While Austria's government expresses alarm and perceives deepfakes as a serious threat, the newspaper attempts to report in a calmer manner. In contrast, the UK shows a reversal of this dynamic. This difference is also reflected in their future strategies, with the UK showing more hesitancy towards regulation compared to Austria. In conclusion, the relationship between democracy and deepfakes in Austria and the UK is acknowledged to have a significant impact on various aspects of democracy. Although discussions on potential positive uses exist, the

overall focus is on recognizing and addressing the threats posed by deepfakes, albeit with varying approaches and language in Austria and the UK.

5.2. SUGGESTIONS FOR FUTURE RESEARCH

How did this research fill the knowledge gap? This research conducted a comprehensive analysis of how two different yet comparable countries discuss deepfakes. While previous research mainly focused on the *consequences of deepfakes* and the *formulation of regulatory measures*, this thesis delved into the *discussions* taking place within each country, including the perspectives of various actors such as government policy papers and newspaper reporting. Additionally, it provided insights into *whether* and *to what extent* deepfakes are recognized as a threat to democracy, rather than simply assuming they *are* a threat. With regards to the Frankfurt School, it can be clearly stated that core concepts such as the "promethean gap," "myth of progress," and the overall perspective of problematic technology and Anders' view of abusing technology are evident in the results of the analysis. The most significant difference is highlighted by Adorno and Horkheimer (1944), who argue that technology, especially in the entertainment sector, exploits the fear of losing individuality and thereby undermines a core concept of pluralism. However, the analysis reveals that deepfakes, particularly for entertainment purposes, should be allowed. This may also challenge the limits of a theory whose core concepts were established 70 years ago.

Nevertheless, this research contributes to the existing literature by shedding light on the discussion within specific European countries, offering new perspectives. The findings align with those of Diakopoulos and Johnson (2019), highlighting the potential harm deepfakes pose to democratic elements such as elections and national security. "*The surprising low number of participants who recognized the deepfake as being manipulated is a clear sign that public awareness and knowledge of deepfakes should improve*" (Dobber et al. 2020), which can be confirmed by this research. It emphasizes the urgent need for education and awareness campaigns to address the deepfake problem. The newspaper analysis, particularly in the UK, reveals significant concerns about deepfakes eroding trust in information, aligning with Vaccari and Chadwick (2020) argument regarding the challenges of public debates in online discourse. While Ray's (2021) research suggests that deepfakes have already influenced elections, this fear is predominantly expressed in UK discourses without concrete evidence of actual impact on voting outcomes. While Chesney and Citron (2019) suggest that deepfakes can defame public figures, this research indicates that deepfakes can harm politicians' integrity, albeit not to the extent of destroying their political reputation.

Most deepfake attacks could be detected and disproven. Therefore, this research partially aligns with these findings. For future research, it would be interesting to conduct a quantitative comparison, including a broader range of EU countries and comparing them with other countries (USA, Canada, or Asian countries). Additionally, examining collaboration and international/European strategies would be valuable, particularly in light of the EU's recent publication of tackling deepfakes in European policy.

5.3. PRACTICAL IMPLICATION FOR POLICY AND GOVERNANCE

So, what needs to be done by policymakers? Given the developed insights, it is crucial to involve tech companies in policy strategies, providing subsidies for the development of deepfake detection technology because we as humans are not capable of detecting deepfakes alone. Investing in education and awareness campaigns is also essential because the danger lies in an uneducated and unaware society where deepfakes can thrive and destroy discussion and debates. Society and individuals play a significant role in determining the effectiveness of deepfakes as a threat to democracy. Therefore, raising awareness in schools but also for example make advertisements about this technology will demonstrate their presence and encourage a critical view towards online content. As the analysis shows, implementing measures such as watermarking deepfake videos could also help raising awareness and transparency, particularly when deepfakes are used in campaigns or advertisements. It was demonstrated that governments, given their powerful position, would be wise to compel tech companies, especially social media platforms like Meta and Google, to develop strategies and technologies for detecting deepfakes. Collaborating with media companies and equipping them with deepfake detection tools is another crucial step in preventing the easy spread of fake news, as can be concluded with the analysis results. The government could provide financial support for such initiatives without interfering in the work of tech companies. Lastly, governments must treat the threat of deepfakes seriously by actively working, revising, and addressing how deepfakes interfere with political and democratic processes. All in all, if deepfakes are recognized as a threat to democracy and governments take appropriate measures, it is still possible to prevent their proliferation on the internet and mitigate the spread of fake news and disinformation. Taking action and fostering collaboration among society, industry, and global partners are key components of these efforts.

6. LIST OF REFERENCES

- Adorno, T. W., & Horkheimer, M. (1944). *Dialectic of Enlightenment*. Stanford, CA: Stanford University Press.
- Anders, G. (1956a). *Die Antiquiertheit des Menschen (Band 1). Über die Seele der zweiten industriellen Revolution*. München.
- Anders, G. (1956b). *Die Antiquiertheit des Menschen (Band 2). Über die Zerstörung des Lebens im Zeitalter der dritten industriellen Revolution*. München.
- Anders G. (1962). Theses for the Atomic Age. *The Massachusetts Review* 3(3), 493–505.
- Arato, A., & Gebhardt, E. (1977). *The Essential Frankfurt School Reader*. New York : Urizen Books.
- Chesney, B. K., & Citron, D. K. (2019). Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security. *California Law Review*, 107(6), 1753–1819. <https://doi.org/10.15779/z38rv0d15j>.
- Delanty, G., & Harris, N. (2021). Critical theory and the question of technology: The Frankfurt School revisited. *Thesis Eleven*, 166(1), 88–108. <https://doi.org/10.1177/07255136211002055>.
- DER STANDARD. (2023). DER STANDARD. *Unternehmen*. Last retrieved: 26.06.2023. <https://about.derstandard.at/unternehmen/>.
- Diakopoulos, N., & Johnson, D. G. (2019). Anticipating and Addressing the Ethical Implications of Deepfakes in the Context of Elections. *Social Science Research Network*. <https://doi.org/10.2139/ssrn.3474183>.
- Dobber, T., Metoui, N., Trilling, D., Helberger, N., & de Vreese, C. (2020). Do (Microtargeted) Deepfakes Have Real Effects on Political Attitudes? *The International Journal of Press/Politics*, 26(1), 69–91. <https://doi.org/10.1177/1940161220944364>.
- Flick, U. (2009). *An Introduction to Qualitative Research*. SAGE.
- Friese, S. (2012). *Getting to know ATLAS.ti*. SAGE Publications Ltd, <https://doi.org/10.4135/9781529799590>.
- Fuchs, C. (2017). Günther Anders' Undiscovered Critical Theory of Technology in the Age of Big Data Capitalism. *TripleC*, 15(2), 582–611. <https://doi.org/10.31269/triplec.v15i2.898>.
- Hameleers, M., Van Der Meer, T. G., & Dobber, T. (2022). You Won't Believe What They Just Said! The Effects of Political Deepfakes Embedded as Vox Populi on Social

- Media. *Social media and society*, 8(3), 205630512211163. <https://doi.org/10.1177/20563051221116346>.
- Hindrichs, G. (2017). Max Horkheimer/Theodor W. Adorno: Dialektik der Aufklärung. Berlin, Boston: De Gruyter. <https://doi.org/10.1515/9783110448764>.
- Julien, H. (2008). . Content analysis. *SAGE Publications*, Inc. Given, L. M. (Ed). <https://doi.org/10.4135/9781412963909>.
- Kluge, S. (2014). Menschenverbesserung in einer Welt ohne Mensch? Zur Aktualität und Kritikwürdigkeit von Günther Anders' Diagnose einer Antiquiertheit des Menschen. *Jahrbuch Für Pädagogik*, 2014(1), 83–104. https://doi.org/10.3726/265764_83.
- Krippendorff, K. (2004). Content analysis: An introduction to its methodology (2nd ed.). *Thousand Oaks*, CA: Sage.
- Luke, T. W. (2000). One-Dimensional Man. *Organization & Environment*, 13(1), 95–101. <https://doi.org/10.1177/1086026600131006>.
- Malicious Deep Fake Prohibition Act. (2018). <https://www.congress.gov/bill/115th-congress/senate-bill/3805>.
- Marcuse, H. (1964). One-Dimensional Man: Studies in the Ideology of Advanced Industrial Society. Boston: Beacon Press.
- Masood, M., Nawaz, M., Malik, K. M., Javed, A., & Irtaza, A. (2021). Deepfakes generation and detection: state-of-the-art, open challenges, countermeasures, and way forward. *Applied Intelligence*. <https://doi.org/10.1007/s10489-022-03766-z>.
- Mayring, P., & Fenzl, T. (2014). Qualitative Inhaltsanalyse. In J. Blasius & N. Baur (Eds.), *Handbuch Methoden der empirischen Sozialforschung* (pp. 543–556). Springer VS. https://doi.org/10.1007/978-3-531-18939-0_38.
- Müller, C. (2016) Prometheanism: Technology, Digital Culture and Human Obsolescence. London: Rowman & Littlefield.
- Paddison M. (2017). Art and the Concept of Autonomy in Adorno's Kant Critique. In M. J. Thompson (Eds.), *The Palgrave Handbook of Critical Theory* (pp. 291-308). Springer.
- Ray, A. M. (2021). Disinformation, Deepfakes and Democracies: The Need for Legislative Reform. *University of New South Wales law journal*, 44(3). <https://doi.org/10.53637/dels2700>.
- Schneider, G., & Toyka-Seid, C. (2023). Vierte Gewalt. *BPB. junges Lexikon*. <https://www.bpb.de/kurz-knapp/lexika/das-junge-politik-lexikon/321342/vierte-gewalt/>.

- Schwarz, E. (2019). Günther Anders in Silicon Valley: Artificial intelligence and moral atrophy. *Thesis Eleven*, 153(1), 94–112. <https://doi.org/10.1177/0725513619863854>.
- The Guardian. (n.d.). The Guardian. *About us*. Last retrieved: 26.06.2023. <https://www.theguardian.com/about>.
- Vaccari, C., & Chadwick, A. (2020). Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News. *Social media and society*, 6(1), 205630512090340. <https://doi.org/10.1177/2056305120903408>.
- Wiggershaus., R. (1995). The Frankfurt School: Its History, Theories and Political Significance. *Constellations*, 2(2), 268–291. <https://doi.org/10.1111/j.1467-8675.1995.tb00031.x>.

7. APPENDIX

7.1. OVERVIEW OF DATA

POLICY DOCUMENTS

GOV.UK. (2019). *Snapshot Paper - Deepfakes and Audiovisual Disinformation*. Last retrieved 26.06.23. <https://www.gov.uk/government/publications/cdei-publishes-its-first-series-of-three-snapshot-papers-ethical-issues-in-ai/snapshot-paper-deepfakes-and-audiovisual-disinformation>.

Parlament Österreich. (2020). Erschließungsantrag der Abgeordneten Douglas Hoyos-Trauttmansdorff, Kolleginnen und Kollegen betreffend Strafrechtliche Verfolgung Deepfakes. Last retrieved: 26.06.23. https://www.parlament.gv.at/dokument/XXVII/A/372/imfname_784533.pdf.

MEDIA DOCUMENTS

UK (The Guardian Articles from 2017-2023)

AgenceFrance – Presse in Shanghai (2019). Chinese deepfake app Zao sparks privacy row after going viral. *The Guardian*. Last retrieved: 26.06.23. <https://www.theguardian.com/technology/2019/sep/02/chinese-face-swap-app-zao-triggers-privacy-fears-viral>.

AFP (2022). Dutch police create deepfake video of murdered boy, 13, in hope of new leads. *The Guardian*. Last retrieved: 26.06.23. <https://www.theguardian.com/world/2022/may/23/dutch-police-create-deepfake-video-of-murdered-boy-13-in-hope-of-new-leads>.

Bartlett, J. (2019). You can't believe a word any of these people is saying – that's the "deep fake" era for you. *The Guardian*. Last retrieved: 26.06.23. <https://www.theguardian.com/commentisfree/2019/jun/16/you-cant-believe-a-word-any-of-these-people-is-saying-thats-the-dep-fake-era-for-you>.

Beard, M. (2019). To fix the problem of deepfakes we must treat the cause, not the symptoms. *The Guardian*. Last retrieved: 26.06.23. <https://www.theguardian.com/commentisfree/2019/jul/23/to-fix-the-problem-of-deepfakes-we-must-treat-the-cause-not-the-symptoms>.

Blackall, M. (2020). Channel 4 under fire for deepfake Queen's Christmas message. *The Guardian*. Last retrieved: 26.06.23.

<https://www.theguardian.com/technology/2020/dec/24/channel-4-under-fire-for-deepfake-queen-christmas-message>.

Busby, M. (2019). Facebook needs regulation to combat fake news, say MPs. *The Guardian*.

Last retrieved: 26.06.23.

<https://www.theguardian.com/technology/2019/feb/18/facebook-regulation-fake-news-mps-deepfake>.

Chivers, T. (2019). What do we do about deepfake video? Deepfake – the ability of AI to fabricate apparently real footage of people – is a growing problem with implications for us all. *The Guardian*. Last retrieved: 26.06.23.

<https://www.theguardian.com/technology/2019/jun/23/what-do-we-do-about-deepfake-video-ai-facebook>.

Cohen, N. (2020). Don't bank on Britain's foppish, lazy elites to save us from deep fakery. *The Guardian*. Last retrieved: 26.06.23.

<https://www.theguardian.com/commentisfree/2020/aug/01/in-an-era-of-deep-fakery-who-can-we-trust-to-tell-us-the-truth>.

Hern, A. (2019). I thought online fakes would cause an infocalypse. *The Guardian*. Last retrieved: 26.06.23. <https://www.theguardian.com/commentisfree/2019/jan/02/online-fakes-hoaxes-internet-deepfakes>.

Hern, A. (2020). Facebook bans “deepfake” videos in run-up to US election. *The Guardian*.

Last retrieved: 26.06.23.

<https://www.theguardian.com/technology/2020/jan/07/facebook-bans-deepfake-videos-in-run-up-to-us-election>.

Hern, A. (2022). Online safety bill will criminalise “downblousing” and “deepfake” porn. *The Guardian*. Last retrieved: 26.06.23.

<https://www.theguardian.com/technology/2022/nov/24/online-safety-bill-to-return-to-parliament-next-month>.

Hern, A., & Milmo, D. (2023). Everything you wanted to know about AI – but were afraid to ask. *The Guardian*. Last retrieved: 26.06.23.

<https://www.theguardian.com/technology/2023/feb/24/ai-artificial-intelligence-chatbots-to-deepfakes>.

Hill, A. (2022). Criminal reforms target “deepfake” and nonconsensual pornographic imagery.

The Guardian. Last retrieved: 26.06.23.

<https://www.theguardian.com/law/2022/jul/07/criminal-reforms-target-deepfake-and-nonconsensual-pornographic-imagery>.

- Hunt, E. (2019). Deepfake danger: what a viral clip of Bill Hader morphing into Tom Cruise tells us: Are deepfakes a threat to democracy? The creator of a series of viral clips says he is raising awareness of their subversive potential. *The Guardian*. Last retrieved: 26.06.23. <https://www.theguardian.com/news/shortcuts/2019/aug/13/danger-deepfakes-viral-video-bill-hader-tom-cruise>.
- Marks, G. (2022). Are these guys for real? How to keep your business safe from deepfakes. *The Guardian*. Last retrieved: 26.06.23. <https://www.theguardian.com/business/2022/aug/21/deepfake-video-audio-fraud-business-protection>.
- Mahdawi, A. (2022). A deepfake video of Joe Biden singing the children’s song Baby Shark isn’t funny – it is deeply worrying. *The Guardian*. Last retrieved: 26.06.23. <https://www.theguardian.com/commentisfree/2022/oct/26/a-deepfake-video-of-joe-biden-singing-the-childrens-song-baby-shark-isnt-funny-it-is-deeply-worrying>.
- Milmo, D. (2023). Age checks, trolls and deepfakes: what’s in the online safety bill?. *The Guardian*. Last retrieved: 26.06.23. <https://www.theguardian.com/technology/2023/jan/17/age-checks-trolls-and-deepfakes-whats-in-the-online-safety-bill>.
- Milmo, D., & Sauer, P. (2022). Deepfakes v pre-bunking: is Russia losing the inforwar?. *The Guardian*. Last retrieved: 26.06.23. <https://www.theguardian.com/world/2022/mar/19/russia-ukraine-infowar-deepfakes>.
- Oltermann, P. (2022). European politicians duped into deepfakes video calls with mayor of Kyiv. *The Guardian*. Last retrieved: 26.06.23. <https://www.theguardian.com/world/2022/jun/25/european-leaders-deepfake-video-calls-mayor-of-kyiv-vitali-klitschko>.
- Parkin, S. (2019). Politician fear this like fire: The rise of the deepfake and the threat to democracy. *The Guardian*. Last retrieved 26.06.23. <https://www.theguardian.com/technology/ng-interactive/2019/jun/22/the-rise-of-the-deepfake-and-the-threat-to-democracy>.
- Paul, K. (2019). California makes “deepfake” videos illegal, but law may be hard to enforce. *The Guardian*. Last retrieved: 26.06.23. <https://www.theguardian.com/us-news/2019/oct/07/california-makes-deepfake-videos-illegal-but-law-may-be-hard-to-enforce>.
- Pomerantsev, P. (2021). Trust no one: Inside the World of Deepfakes by Michael Grothaus review – disinformation’s superweapon. *The Guardian*. Last retrieved: 26.06.23.

<https://www.theguardian.com/books/2021/dec/16/trust-no-one-inside-the-world-of-deepfakes-by-michael-grothaus-review-disinformations-superweapon>.

Roth, A. (2021). European MPs targeted by deepfake video calls imitating Russian opposition. *The Guardian*. Last retrieved: 26.06.23.

<https://www.theguardian.com/world/2021/apr/22/european-mps-targeted-by-deepfake-video-calls-imitating-russian-opposition>.

Sample, I. (2020). What are deepfakes – and how can you spot them?. *The Guardian*. Last retrieved: 26.06.23. <https://www.theguardian.com/technology/2020/jan/13/what-are-deepfakes-and-how-can-you-spot-them>.

Schwartz, O. (2018). You thought fake news was bad? Deep fakes are where truth goes to die. *The Guardian*. Last retrieved: 26.06.23.

<https://www.theguardian.com/technology/2018/nov/12/deep-fakes-fake-news-truth>.

Sullivan, H. (2021). Move over, Deep Nostalgia, this AI app can make Kim Jong-un sing I will Survive. *The Guardian*. Last retrieved: 26.06.23.

<https://www.theguardian.com/technology/2021/mar/12/wombo-ai-app-kim-jong-un-sing-i-will-survive-deepfakes>.

Taylor, J. (2022). From Trump Nevermind babies to deep fakes: DALL-E and the ethics of AI art. *The Guardian*. Last retrieved: 26.06.23.

<https://www.theguardian.com/technology/2022/jun/19/from-trump-nevermind-babies-to-deep-fakes-dall-e-and-the-ethics-of-ai-art>.

Austria (Der Standard Articles from 2017-2023)

Al-Youssef, M. (2018). Welche Möglichkeiten Fake News in Zukunft haben werden. *Der Standard*. Last retrieved: 26.06.23.

<https://www.derstandard.at/story/2000089573757/welche-moeglichkeiten-fake-news-in-zukunft-haben-werden>.

Al-Youssef, M. (2020). Die Zukunft der Fake-News ist deep. *Der Standard*. Last retrieved: 26.06.23. <https://www.derstandard.at/story/2000120315626/die-zukunft-der-fake-news-ist-deep>.

Der Standard. (2018). Deepfake: EU-Kommission warnt vor gefälschten Politikervideos. *Der Standard*. Last retrieved: 26.06.23.

<https://www.derstandard.at/story/2000094525676/deepfakes-eu-kommission-warnt-vor-gefaelschten-politikervideos>.

- Der Standard. (2019a). "Rise of Skywalker": Neuer "Star Wars"-Film lässt Deepfake-Welle erwarten. *Der Standard*. Last retrieved: 26.06.23. <https://www.derstandard.at/story/2000112470977/rise-of-skywalker-neuer-star-wars-film-laesst-deepfake-welle>.
- Der Standard. (2019b). Stimme wie Obama: Deepfake-KI lässt Nutzer klingen, wie sie wollen. *Der Standard*. Last retrieved: 26.06.23. <https://www.derstandard.at/story/2000098708642/stimme-wie-obama-deepfake-ki-laesst-nutzer-klingen-wie-sie>.
- Der Standard. (2020a). Britische Queen klagt in Deepfake-Video über Meghan und Harry. *Der Standard*. Last retrieved: 26.06.23. <https://www.derstandard.at/story/2000122773387/britische-queen-klagt-in-deepfake-video-ueber-meghan-und-harry>.
- Der Standard. (2020b). Kim und Putin: Deepfake-Autokraten warnen USA vor Untergang ihrer Demokratie. *Der Standard*. Last retrieved: 26.06.23. <https://www.derstandard.at/story/2000120364619/kim-putin-deepfake-autokraten-warnen-usa-vor-untergang-ihrer-demokratie>.
- Der Standard. (2020c). Deepfake lässt Hitler und Stalin im Duett singen. *Der Standard*. Last retrieved: 26.06.23. <https://www.derstandard.at/story/2000119794156/deepfake-laesst-hitler-und-stalin-im-duett-singen>.
- Der Standard. (2020d). Sebastian Kurz beim Song Contest: Deepfake-Scherz auf Youtube. *Der Standard*. Last retrieved: 26.06.23. <https://www.derstandard.at/story/2000114847718/sebastian-kurz-beim-song-contest-deepfake-scherz-auf-youtube>.
- Der Standard. (2020e). Twitter will böartige Deepfake-Videos löschen. *Der Standard*. Last retrieved: 26.06.23. <https://www.derstandard.at/story/2000114194685/twitter-will-boesartige-deepfake-videos-loeschen>.
- Der Standard. (2020f). Deepfakes werden erstmals Thema im Nationalrat. *der Standard*. Last retrieved: 26.06.23. <https://www.derstandard.at/story/2000120442561/deepfakes-werden-erstmals-thema-im-nationalrat>.
- Der Standard. (2020g). Assembler: Tool von Google-Schwester soll Deepfakes enttarnen. *Der Standard*. Last retrieved: 26.06.23. <https://www.derstandard.at/story/2000114706895/assembler-tool-von-google-schwester-soll-deepfakes-enttarnen>.

- Der Standard. (2021a). Niederländische Parlamentarier fielen auf Deepfake-Streich herein. *Der Standard*. Last retrieved: 26.06.23. <https://www.derstandard.at/story/2000126164587/niederlaendische-parlamentarier-fielen-auf-deepfake-streich-herein>.
- Der Standard. (2021b). Forscher entwickeln Software, die fast jedes Deepfake erkennt. *Der Standard*. Last retrieved: 26.06.23. <https://www.derstandard.at/story/2000125029252/forscher-entwickeln-software-die-fast-jedes-deepfake-erkennt>.
- Der Standard. (2021c). STANDARD-User nicht viel besser als der Zufall. *Der Standard*. Last retrieved: 26.06.23. <https://www.derstandard.at/story/2000124957110/erkennung-von-deepfakes-standard-user-nicht-viel-besser-als-der>.
- Der Standard. (2022a). Falscher Klitschko: Sprach Michael Ludwig wirklich mit einem Deepfake?. *Der Standard*. Last retrieved: 26.06.23. <https://www.derstandard.at/story/2000136929726/falscher-klitschko-sprach-michael-ludwig-wirklich-mit-einem-deepfake>.
- Der Standard. (2022b). „My Face My Choice“: Deepfakes sollen helfen, unsere Privatsphäre zu schützen. *Der Standard*. Last retrieved: 26.06.23. <https://www.derstandard.at/story/2000141045458/my-face-my-choice-deepfakes-sollen-helfen-unsere-privatsphaere-zu>.
- Der Standard. (2022c). Deepfakes: Eine neue Ära der Täuschung steht bevor. *Der Standard*. Last retrieved: 26.06.23. <https://www.derstandard.at/story/2000139484409/deepfakes-eine-neue-aera-der-taeschung-steht-bevor>.
- Der Standard. (2022d). Kampf gegen Deepfakes: Innenminister Karner spricht von „Steigerungsstufe zu Fake News“. *Der Standard*. Last retrieved: 26.06.23. <https://www.derstandard.at/story/2000136043079/kampf-gegen-deepfakes-innenminister-karner-spricht-von-steigerungsstufe-zu-fake>.
- Der Standard. (2022e). Fakecatcher: Intel kann Deepfakes in Echtzeit entlarven. *Der Standard*. Last retrieved: 26.06.23. <https://www.derstandard.at/story/2000140984734/fakecatcher-intel-kann-deepfakes-in-echtzeit-entlarven>.
- Der Standard. (2022f). Regierung will Verbreitung von Deepfakes verhindern. *Der Standard*. Last retrieved: 26.06.23. <https://www.derstandard.at/story/2000136041438/regierung-will-verbreitung-von-deepfakes-verhindern>.
- Der Standard. (2022g). Forscher entwickeln KI, die Deepfakes mit 98,5 Prozent Wahrscheinlichkeit erkennt. *Der Standard*. Last retrieved: 26.06.23.

<https://www.derstandard.at/story/2000135336666/forscher-entwickelt-ki-die-deepfakes-mit-98-5-prozent-wahrscheinlichkeit>.

Der Standard. (2022h). Menschen sind wirklich schlecht darin, echte Gesichter von Deepfakes zu unterscheiden. *Der Standard*. Last retrieved: 26.06.23.

<https://www.derstandard.at/story/2000133611271/menschen-sind-wirklich-schlecht-darin-echte-gesichter-von-deepfakes-zu>.

Der Standard. (2023). Deepfakes verbreiten Propaganda für Kommunistische Partei Chinas. *Der Standard*. Last retrieved: 26.06.23.

<https://www.derstandard.at/story/2000143330503/deepfakes-verbreiten-propaganda-fuer-kommunistische-partei-chinas>.

Fellner, S. (2019). Perfektes Deepfake-Video: Boris Johnson lobt Jeremy Corbyn. *Der Standard*. Last retrieved: 26.06.23.

<https://www.derstandard.at/story/2000110975264/perfekte-deepfake-videos-von-boris-johnson-und-jeremy-corbyn>.

Kropshofer, K. (2019). Zao: Aufregung um neue Deepfake-App, die erschreckend gute Resultate erzeugt. *Der Standard*. Last retrieved: 26.06.23.

<https://www.derstandard.at/story/2000108167817/zao-aufregung-um-neue-deepfake-app-die-erschreckend-gute-resultate>.

Palinger, J. (2021). Wie KI Falschinformationen produziert – und uns damit täuscht. *Der Standard*. Last retrieved: 26.06.23.

<https://www.derstandard.at/story/2000127500558/wie-ki-falschinformationen-produziert-und-uns-damit-taeuscht>.

Pichler, G. (2022). Ludwig und der falsche Klitschko: Die Deepfake-Gefahr droht real zu werden. *Der Standard*. Last retrieved: 26.06.23.

<https://www.derstandard.at/story/2000127500558/wie-ki-falschinformationen-produziert-und-uns-damit-taeuscht>.

Pramer, P. (2020). Wie sich mit Deepfakes Geschichte manipulieren lässt. *Der Standard*. Last retrieved: 26.06.23.

<https://www.derstandard.at/story/2000120975968/wie-sich-mit-deepfakes-geschichte-manipulieren-laesst>.

Red. (2018). US-Verteidigungsministerium entwickelt Tools gegen Deepfakes. *Der Standard*. Last retrieved: 26.06.23.

<https://www.derstandard.at/story/2000084973088/us-verteidigungsministerium-entwickelt-tools-gegen-deepfakes>.

Scherndl, G. (2022a). Ludwig und der falsche Klitschko: Wie konnte es so weit kommen?. *Der Standard*. Last retrieved: 26.06.23.

<https://www.derstandard.at/story/2000136908295/nach-klitschko-fake-botschaften-sollen-gespraechе-koordinieren>.

- Scherndl, G. (2022b). Wiens Bürgermeister telefonierte mit Fake-Klitschko. *Der Standard*. Last retrieved: 26.06.23. <https://www.derstandard.at/story/2000136884227/wiens-buergermeister-ludwig-soll-anruf-von-fake-klitschko-bekommen-haben>.
- Schmid, F., & Al-Youssef, M. (2018). Deepfake-News täuschen alle Sinne – und könnten Kriege auslösen. *Der Standard*. Last retrieved: 26.06.23. <https://www.derstandard.at/story/2000074430944/deepfake-news-taeuschen-alle-sinne-und-koennten-kriege-ausloesen>.
- Sommavilla F., & Stajić, O. (2020). Ohne automatische Deepfake-Erkennung sind wir bald aufgeschmissen. *Der Standard*. Last retrieved: 26.06.23. <https://www.derstandard.at/story/2000118585546/ohne-automatische-deepfake-erkennung-sind-wir-bald-aufgeschmissen>.
- Stepanek, M. (2022). Dem falschen Klitschko auf der Spur. *Der Standard*. Last retrieved: 26.06.23. <https://www.derstandard.at/story/2000138077510/dem-falschen-klitschko-auf-der-spur>.
- Weiss, S. (2022). Deepfake-Expertin Nina Schick: “Jede Identität ist heute gefährdet“. *Der Standard*. Last retrieved 26.06.23.. <https://www.derstandard.at/story/2000137059050/deepfake-expertin-nina-schick-jede-identitaet-ist-heute-gefaehrdet>.
- Wilkens, A (2022). Österreich bringt internationales Verbot für Deepfake-Technik ins Spiel. *Heise online*. Last retrieved: 26.06.23.. <https://www.heise.de/news/Oesterreich-bringt-Verbot-fuer-Deepfake-Technik-ins-Spiel-7123179.html>.

7.2. TRANSLATION GERMAN CITATIONS

Citation From: 4.2.1. Discourse In Austria: From Fear To Regulation

1. "Participants were correct only 48% of the time, which is slightly worse than what would be expected from random guessing (50:50)".
Original: „Die Teilnehmer lagen dabei nur in 48 Prozent der Fälle richtig und damit sogar etwas schlechter, als es bei zufälligem Raten (50:50) zu erwarten gewesen wäre“ (Der Standard 2022h).
2. "It becomes dangerous and offensive as soon as it is used to fake political speeches or pornographic material".

Original: „Gefährlich und verletzend wird, es sobald damit politische Ansprachen oder pornografisches Material gefaket werden“ (Sommavilla and Stajić 2020).

3. "Deepfakes are used to manipulate democratic processes. Involving key figures in politics or the economy poses a significant security risk because the identification of artificial manipulation is difficult to prove or trace".

Original: „Deepfakes werden verwendet, um demokratische Prozesse zu manipulieren". Schlüsselpersonen in der Politik oder in der Wirtschaft bergen ein erhebliches sicherheitspolitisches Risiko, weil die Identifizierung einer künstlichen Beeinflussung schwer nachweisbar oder nachvollziehbar ist“ (Der Standard 2022f).

4. "Imagine what would happen if these videos ended up online and spread rapidly. [...] In this scenario, a 'negative actor' falsifies geopolitical events to achieve certain goals".

Original: „Man stelle sich vor, was passiert, wenn diese Videos im Netz landen und sich dort in Windeseile verbreiten. [...] Dabei will ein ‚negativer Akteur‘ geopolitische Ereignisse fälschen, um bestimmte Ziele zu erreichen“ (Schmid and Al-Youssef 2018).

5. "Through 'digital warfare,' it seems that the aim is to shake trust in politics and discredit Ukraine and its supporters“.

Original: „Durch ‚digitale Kriegsführung‘ solle wohl das Vertrauen in die Politik erschüttert und die Ukraine und ihre Unterstützer diskreditiert werden“ (Scherndl 2022b).

Citations: 4.3.1. Austrian's Government: "Deepfakes As A Major Threat"

1. "In Austria, the threat posed by deepfakes is real".

Original: „In Österreich ist die Bedrohung durch Deepfakes real“ (Parlamentsbeschluss 2022, p. 23).

2. "Fight against deepfakes"

Original: „Kampf gegen Deepfakes“ (Parlamentsbeschluss 2022, p. 23).

3. "Consistent protection of the integrity of our democracy and the democratic formation of will against external influence".

Original: „Konsequenter Schutz der Integrität unserer Demokratie und der demokratischen Willensbildung vor Einflussnahme von außen“ (Parlamentsbeschluss 2022, p. 24).

4. "The labelling of deepfakes in the field of artistic expression and freedom of speech should ensure that there are no limitations on this form of art or fundamental freedom. Satire is fundamentally not about deceiving, unlike disinformation".

Original: „Bei der Regulierung von Deepfake-Videos sind die relevanten Grund- und Persönlichkeitsrechte zu berücksichtigen und ist insbesondere auf den besonderen Schutz

der Meinungsäußerungsfreiheit und der Kunstfreiheit zu achten“ (Parlamentsbeschluss 2022, p. 10).

5. "When these elements spread on the internet, they disrupt democratic processes and the trust of citizens".

Original: „Wenn sich diese Elemente im Internet verbreiten, stören sie demokratische Prozesse und das Vertrauen der Bürgerinnen und Bürger“ (Parlamentsbeschluss 2022, p. 12).

Citations From: 4.4. Comparing Deepfake Discourses In Newspapers: Austria Vs. UK

1. "Riparbelli called on governments to enact comprehensive laws to regulate 'synthetic media' and the deepfake industry".

Original: „forderte Riparbelli die Regierungen auf, umfassende Gesetze zur Regulierung der ‚synthetischen Medien‘ und der Deepfake-Industrie zu erlassen“ (Der Standard 2023).

2. "It is urgently necessary for politics to also engage with deepfakes and develop strategies to protect the affected individuals".

Original: „Daher ist es dringend notwendig, dass sich auch die Politik damit beschäftigt und Strategien zum Schutz von Betroffenen entwickelt" (Der Standard 2020f).

3. "Deepfakes - such as those of politicians - pose, according to the government, 'a significant security policy risk' and 'constitute a threat to our democracy and societal cohesion’".

Original: “Deepfakes – etwa von Politikerinnen oder Politikern – bergen laut Regierung ‚ein erhebliches sicherheitspolitisches Risiko‘ und ‚stellen eine Bedrohung für unsere Demokratie und das soziale Gefüge dar““ (Der Standard 2022d).

4. "For instance, the faces of famous women are retouched into pornographic material by the thousands. Furthermore, video manipulations often lead to attempts of extortion“.

Original: „So werden etwa zu Tausenden die Gesichter berühmter Frauen in pornografisches Material hineinretuschiert. Und immer wieder kommt es mit den Videomanipulationen auch zu Erpressungsversuchen“ (Der Standard 2021c).

Citations From: 4.5. Comparing Deepfake Discourses in Policy Papers: Austria Vs. UK

1. "A false video created using AI to deceive or manipulate with political intent can pose a significant threat to the integrity of a democracy. As an example, consider a video showing a head of state or a government official saying things that subsequently lead to mass demonstrations and a government and state crisis. When these elements spread on the internet, they disrupt democratic processes and the trust of citizens".

Original: „Ein falsches Video mittels KI hergestellt, um mit politischer Absicht zu täuschen oder zu manipulieren, kann eine erhebliche Gefahr für die Integrität einer Demokratie darstellen. Als Beispiel sei ein Video genannt, das ein Staatsoberhaupt oder ein Regierungsmitglied zeigt, das Dinge sagt, die dann in Folge zu Massendemonstrationen sowie Regierungs- und Staatskrise führen. Wenn sich diese Elemente im Internet verbreiten, stören sie demokratische Prozesse und das Vertrauen der Bürgerinnen und Bürger“ (Parlamentsbeschluss 2022, p. 12).

2. "an umbrella term for various forms of audiovisual manipulation, including video, audio, or both. Typically, AI-based technology is used to create deepfakes. Deepfakes are perfectly faked videos, images, or audio in which individuals are made to say things or appear to engage in actions that never actually took place".

Original: „Der Begriff ‚Deepfake‘ wird als Überbegriff für verschiedene Formen der audiovisuellen Manipulation einschließlich Video, Audio oder beides verwendet. Typischerweise wird zur Erstellung von Deepfakes eine auf Künstlicher Intelligenz (KI) basierte Technologie verwendet. Deepfakes sind perfekt gefälschte Videos, Bilder oder Audio in denen Personen Aussagen in den Mund gelegt werden oder in denen sie scheinbar Handlungen begehen, die in Wirklichkeit nie stattgefunden haben“ (Parlamentsbeschluss 2022, p 9).

3. "the protection of democracy, the protection of the individual, national security, and technological developments"

Original: Schutz der Demokratie, Schutz des Individuums, nationale Sicherheit und technologische Entwicklungen (Parlamentsbeschluss 2022, p. 25).