

The Digital Condition:

Conceptualizing Arendtian Political Action in Digital Worlds

Master's Thesis submitted for the
Research Master in Philosophy
Utrecht University
December 4, 2023

Jasper Verheul
Student number: 6083803

Supervisor: Dominik Klein
Second reader: Maurits de Jongh



**Utrecht
University**

Contents

- Introduction..... 4
 - [1] The digitalization of society 4
 - [2] Philosophies of digital worlds 8
 - [3] Understanding political action in digital worlds 10
 - [4] An Arendtian framework for political action 10
 - [5] Outline of the chapters 12
- Chapter 1: Digital Worlds 14
 - [1] Interactive, social, complete 14
 - Interactive 15
 - Social..... 15
 - Complete 16
 - [2] Simulated, virtual, digital 17
 - Simulated..... 18
 - Virtual 19
 - Digital..... 20
 - [3] Digital realism 21
 - [4] Phenomenology of digital worlds 24
 - [5] Conclusion: Conceptualizing digital worlds 27
- Chapter 2: Political Action 30
 - [1] Labor, work, and political action..... 30
 - [2] The nature and importance of political action 32
 - [3] Plurality 34
 - Distinction..... 35
 - Equality 35
 - [4] Freedom 36
 - Unpredictability 37
 - Irreversibility..... 38
 - [5] Totalitarianism 38
 - [6] Conclusion: Conceptualizing political action..... 41
- Chapter 3: Plurality in digital worlds 42
 - [1] Distinction in digital worlds..... 42
 - Taking distinction to digital worlds 42
 - Two perspectives on digital distinction 44
 - Distinction in the Matrix..... 46
 - [2] Equality in digital worlds..... 48
 - Connection between actors 49

Political moods	49
Creating political moods in digital worlds	50
[3] Conclusion: The theoretical obstacles for plurality in digital worlds.....	52
Chapter 4: Freedom in digital worlds.....	53
[1] Unpredictability in digital worlds.....	53
Digital dice and deeds	55
[2] Irreversibility in digital worlds.....	57
Acting with a backspace button	58
[3] Conclusion: Theoretical obstacles to freedom in in digital worlds.....	59
Chapter 5: Building digital societies	61
[1] Life in a constructed digital world.....	62
[2] The threat of digital totalitarianism.....	65
The Netherlands and indistinguishability	66
The United States and inequality	67
China and predictability	69
Russia and reversibility.....	71
[3] Conclusion: Enabling political action in digital worlds.....	73
Conclusion	76
Literature.....	81

Introduction

[1] The digitalization of society

We live an age of “digital transformation” (Guillot, 2023). No matter how ordinary the role of digital devices might appear in our daily lives, these are epochal times. Although the apparent everydayness of digital technology manages to conceal how truly impactful it is on the human experience, the shift in this experience should not be ignored. The modern way of living is being radically transformed, not between but within generations. With the introduction of new digital tools, the possibilities for their users grow, but so do the possible dangers. This rapid change has outdated our systems of governance for matters like privacy, security, responsibility, and commerce (Flyverbom et al., 2019). More fundamentally, however, it has left us with a void in our philosophical understanding of the human condition: being in relation to digital technology (*homo digitalis*). Filling this void is one of the most pressing tasks of contemporary philosophy.

The impact of digital technology on the human experience appears to be growing exponentially. It changes this experience not only by giving us new options, like constant access to information and global communication, but also by transforming old ways of doing things. Increasingly many human activities can be performed and consequently are being performed through digital mediums. These activities range from political acts like e-voting to personal activities like socializing with friends to ceremonies like attending a wedding or a funeral through a digital medium, something that would have been frowned upon a decade ago.

This trend of digitalization of society has attracted many optimists. Projects like the Metaverse try to take it to its next logical step (Floridi, 2022). They promise to develop a digital medium that is all-encompassing, meaning that every aspect of life can be experienced through it. The goal is no longer to digitalize separate aspects of the human experience, but instead to create an integrated *digital world* that many users can inhabit. The idea here is that this digital world would offer a substitution for the non-digital world. It would be a virtual space in which all aspects of the human experience that the physical world has to offer could be experienced. To create a complete substitution that renders the physical world unnecessary is still too ambitious, even for projects like the Metaverse, but this idea of an all-encompassing digital world speaks to the imagination. It could be seen as the ideal version of digital environment projects. Of course, the concept of digital worlds predates the Metaverse. They have featured in science fiction stories for decades already. Notably, the concept was used in the movie *The Matrix (1999)*, in which our world in fact turns out to be a digital world,

designed by malicious robots to keep us docile. It is likely, however, that humans rather than malicious robots will be the ones to design the first digital worlds.

These worlds will probably not be complete replicas of the physical world like in *The Matrix*. Instead, its creators will likely be tempted to leave out things that we do not want. Pain and illness, for example, are not experiences that are likely to be coded into designed digital worlds. Other inconveniences like traveling long distances to visit places or having to wait for certain things to happen might also be left out because they are not deemed necessary, although some might find value in these experiences. On the other hand, people will have the option to add many new experiences to digital worlds. If so desired, one could fly like a bird across physically impossible landscapes, appear to others in any way one wants, and play elaborate games like the ones depicted in the movie *Ready Player One* (2018). Conversations between any inhabitants of this world could be had at any time. Large amounts of computing power and information would be ready at one's fingertips. The possibilities would seem limitless.

However, an important aspect of digital worlds is sometimes forgotten in these eager predictions of new possibilities. When someone fully lives out their life in an environment where every action is mediated by digital technology, then this digital environment will shape that life in a fundamental way. Just like a life in the real world is bound by natural laws, a digital life will be governed by its own set of digital laws. While these laws might seem unintuitive to those who did not grow up in a digital environment, and might be so intuitive to those who grew up in it that they fail to recognize them, they nonetheless will play an important role in shaping a digital life. The difference between natural laws and digital laws does not have to be a metaphysical one, which is an issue that we will return to later. Just like 'an apple falls to the ground' can be a natural law that does not accurately describe the event on a metaphysical level, 'a digital apple can be deleted' could be a digital law that helps us understand the event without metaphysically describing it. Here, we do find differences, since a real apple cannot be deleted in the same way a digital apple can. Understanding these differences will be an important part of understanding the experience of life in a digital world.

Naturally, the different experiences people will have and the activities they will perform in digital worlds will be unique, but they can be roughly divided into three categories. First, people will have a private life. This includes everything a person needs for their personal well-being and satisfaction. In

one form or another, people would require food and sleep in a complete digital world.¹ On top of that, people will have amusement like sports and games. They could meet other inhabitants of the digital world and talk to them, create friendships and even engage in romantic relationships. Second, people will work in the digital world. This includes every kind of creation, from writing an essay to creating art to constructing the digital environment itself. This work will often involve collaboration between people on bigger projects, meaning that large specialized information networks will emerge, similar to professions in the physical world. Lastly, people will engage in political activity. Because the digital world will be an environment that is shared by many people with different wants and needs, this will require a system to decide how the digital world will be run. This political system could have many forms, but if it is similar to modern democracies, then there will be speeches, debates, campaigning, and periodic elections all happening within the digital world.

Moving these three categories of activities from the physical world to a digital world will probably not go without its fair share of problems. These problems might come in the form of bugs and errors, manageable malfunctions that might annoy users but will have no considerable impact on their lives. The true problems might arise when some human activities turn out to be fundamentally reliant on some non-digital aspect that cannot be translated into the digital world. These challenges could arise in any of the three categories. There might, for instance, be a need in people's personal lives for physical connection. This might cause difficulties with engaging in true valuable friendships or romantic relationships that imbue a life with meaning. In work, there might be a need for a physical context that goes beyond an information network. In addition, it might be impossible to produce something lasting and unique in a digital world. Finally, the performance of political action in digital worlds might face enormous challenges. The perfect multipliable, ephemeral and non-temporal nature of digital objects might obscure the unique wants and needs of people in a digital world. Trustworthy information might be nigh impossible to come by in such an environment. In the worst-case scenario, it might result in a constant paranoid battle with Descartes' evil digital demon. This might result in a populace of isolated lonely people that is extremely susceptible to totalitarian government forms. This would mean that the individuality and freedom of people living in a digital world could not be guaranteed, a scenario that makes Orwellian dystopias look like the nice alternative.

¹ The activities of eating and sleeping might be simulated in the digital world while they are taking place in real life, or they might require someone to temporarily disconnect.

Although a digital world dystopia might still sound like a science fiction scenario, the technological advancements of the last decades are slowly moving it into the realm of realistic possibilities. We have reached the point where such ambitious goals are realistic enough for big investors to at least give it a shot. Even if digital worlds are still far away, we will most likely see a continuation of the digitalization of society. This will mean the introduction of newer and more capable digital environments. A better understanding of the human condition in digital worlds will also give us crucial insights into the human condition in these smaller digital environments.

Today, more than ever, our common sense of reality is in jeopardy. So far, the digital age has not ushered in a time of endless possibility and shared knowledge, but rather one of extreme volatility and post-truth. At the same time we are faced with large-scale challenges that require unprecedented amounts of agreement and cooperation to be solved. The COVID-19 pandemic forced whole nations to swiftly adapt to social distancing measures. This meant that many activities that were previously done on location, and where being physically together was considered an essential part of the activity, suddenly had to be performed through digital mediums. The technology to facilitate this was quickly developed, but it became apparent that moving a whole society to digital platforms is, in practice, still extremely problematic. However, the digital infrastructure that was developed remains available and is still being used.

Although these digital technologies helped to keep a large part of society running without being physically together, digital platforms also turned out to be an obstacle for collective action. The widespread circulation of factually incorrect information about the virus on these platforms hindered governments in their measures against it. Likewise, the critical urgency of climate change requires a common ground rather than an increasingly divided digital experience.

Before we turn digital worlds into an ideal to strive for, we should make sure that these worlds could and will be good worlds for people to live in. Technological advancement is not a deterministic force that is beyond our control. Its course can be altered and steered by those who develop it, those who use it, and the legislative instruments that bind it. Ultimately, we are capable of making choices that harness the potential of technology to improve people's lives. Making the correct choices, however, asks for a good understanding of the problem. For this goal, understanding the human condition in a digital environment is essential. Investigating the human condition in digital worlds will help us to better grasp the problems at hand. This will allow us to address current problems with virtual

technology and anticipate future developments, like that of artificial intelligence, to guide us on the path of digitalization.

[2] Philosophies of digital worlds

The topics of digital environments and the relation between humans and technology have been studied by many philosophers in many different contexts. Concerns about the quality of life in and alongside digital environments have mainly been raised through a critique of social networking platforms. These platforms include traditional networking sites (Facebook, LinkedIn), media sharing sites (Youtube, Reddit, Instagram, TikTok), platforms for microblogging (Tumblr, Twitter), messaging apps (WhatsApp, Messenger, WeChat), and social gaming platforms (Steam, Twitch, Discord) (Vallor, 2022). Social networks are currently the closest contenders for digital worlds in popular use, although attempts to incorporate virtual reality technology in these networks are starting to emerge, like the Metaverse and VRChat. Discussions of such networks usually focus on a specific issue that they introduce. These are topics like privacy (Nissenbaum, 2009), identity (Ess, 2010), friendships (Elder, 2014; Vallor, 2012), cybercrime (Douglas, 2016), and alienation (Turkle, 2011).

Recently, much attention has been given to the question if political action and discourse are threatened by social networks. For instance, many have raised epistemological concerns, addressing problems like information bubbles (Nguyen, 2020) and the rise of conspiracy theories (Uscinski & Enders, 2023) as possible threats. Others have identified the increased potential to influence elections as a possible challenge (Snyder, 2018, pp. 227–232). A different way of discussing this issue is by asking the question if political action and discourse are possible through the mediation of digital platforms (Boyd, 2008). A common approach that has been taken is to use a theoretical concept that is foundational for political action, like Habermas' concept of the public sphere, and investigate if this concept could be realized in a digital environment (Bohman, 2008; Dahlberg, 2001; Ess, 1996). Another approach is to develop a negative concept as a threat to politics like digital totalitarianism (Capurro, 2011). Much of this discourse is based on a certain underlying theory about the nature of digital environments and what is theoretically possible in them.

These theories can, of course, be very different from each other, sometimes resulting in contrasting evaluations of digital environments. Two philosophical traditions that have dealt with questions in this area are that of phenomenology and philosophy of mind. Within the phenomenological tradition itself, many different approaches have been explored to understanding digital environments. One of

these approaches originates from the conceptualization of technology by Martin Heidegger (Heidegger, 1954/1977). One of the key insights that is shared by these approaches is that technology and society co-constitute each other and that they continually draw upon each other for their meaning (Itrona, 2017). This means that if someone was to communicate with another person through a digital platform, the presence of that digital platform would alter the way they became disclosed to one another, which is to say they phenomenologically appear different. They would now show up in each other's world of meaning as 'being contactable' through the platform for instance.

This theory informed some of the early critiques of the possibilities of digital environments, spearheaded by thinkers like Albert Borgmann and Hubert Dreyfus. Their worry is that because people in digital environments are not embodied, not only in the physical sense but also the phenomenological sense, something is missing from the experience in these environments. Borgmann emphasizes the shallow nature of digital beings and their disconnectedness from the physical world. This disconnection means that users of digital platforms are placed outside the network of meaning of this physical world, which prevents them from being seen and judged (Borgmann, 1993). Dreyfus continues this line of thinking and claims that digital environments lack a sense of risk. Because we do not present ourselves in our physical fragility, there is no risk in any decision that we make or action we perform. This leads to a lack of commitment to these decisions and actions, ultimately resulting in a loss of proper meaning (Dreyfus, 1999, 2001).

Very recently, a positive stance on digital environments was taken by David Chalmers (2022), coming from the philosophy of mind tradition. More specifically he has been a defender of a computational theory of mind, claiming that the mind is fundamentally computable. In his recent work, Chalmers turns his attention to the question of reality, which creates a different perspective that is difficult to directly compare to the phenomenological tradition. Informed by his background in computational theory of mind Chalmers defends the strong position of digital realism. This position claims that there is no fundamental difference in reality between physical and digital objects. If one would find themselves inside a digital environment, then the digital objects that they would encounter in that environment are real to them in exactly the same way that physical objects in a physical environment are real to me. From this equivalence, Chalmers goes on to make further claims about the possibilities of leading a good life in digital worlds. Since there is no fundamental difference, the idea is that there is exactly as much of a possibility to live a full and good life in a digital world as in the physical world. A digital world could have meaning or it could lack it, but if it lacks meaning, this is not because of the fact that the world is virtual (Chalmers, 2022, pp. 303–304).

[3] Understanding political action in digital worlds

While much academic attention has been given to the question if political action can function in digital environments like social networks, this question has not been thoroughly explored in the context of full digital worlds. The fact that much attention has been given to this question in other contexts, demonstrates the underlying assumption that the possibility for effective political action is vital for people to be able to lead a full and good life. Without it, a dystopian society lies waiting in the background. It is important that questions about political action as such should be included in the discussion about digital worlds, because the conclusions of these questions will heavily influence our verdict of digital worlds and our orientation in their development.

While both the early phenomenological critiques of digital worlds and the later positive position on digital reality by Chalmers have raised many important points about the relation between humans and digital worlds, these discussions have so far failed to properly address political action. Therefore, this thesis will discuss the question if digital worlds are a suitable platform for political action. Taking an analysis of the nature of digital worlds, digital objects and the human experience in digital environments by the phenomenological and philosophy of mind traditions as a foundation, I will investigate both the possibility of political action and the practical realization of such action in digital worlds. The primary research question that this thesis explores is **whether and under which circumstances digital worlds provide a conducive environment for political action.**

Since there do not exist any theories of political action that are tailor-made for an investigation of digital worlds, the secondary objective of this thesis will be to restate an existing theory of political action in order to make it fit for this purpose. The theory for political action that this thesis will repurpose for the investigation of digital worlds is the one that has been developed by Hannah Arendt. The secondary research question is therefore **how Arendt's political action theory can be applied to effectively analyze and understand political action within digital environments.**

[4] An Arendtian framework for political action

Arendt develops a theory of political action in her book *The Human Condition* (1958/2018) that appears promising for the purpose of analyzing digital worlds for a couple of reasons. First, she conceptualizes the category of political action by contrasting it with two other important categories of action she calls labor and work. This tripartition of human action can be employed to make a distinction between the kinds of action we perform in digital worlds. This will allow us to isolate the

category of political action in digital worlds and investigate the extent to which it is possible in such environments.²

The way Arendt conceptualizes these categories is also useful because it has roots in the phenomenological tradition but does not continue in the same direction as many later phenomenologists. This means that while her theory is an excellent vantage point from which to address the worries that phenomenologists might have about digital worlds, it is also a good bridge towards contemporary problems that have been raised in other corners of philosophy.

Even though Arendt did not live long enough to witness the profound impact that digital technology has had on the world to the extent that we are familiar with today, her theory proves its relevance even in this digital age. Especially on questions about totalitarianism, freedom, and collective action, topics that have become very relevant in the last few years, her ideas are of great value. Especially the concept of totalitarianism has reared its head in the context of digital societies.

The concept of political action itself, as it is developed in *The Human Condition*, can be seen as an antidote to totalitarianism, a concept she already had explored more fully in *The Origins of Totalitarianism* (1951/2017). In her work, Arendt discusses several constituents and necessary conditions for political action. An important concept here is the public sphere, an environment that allows for political action to occur, but is dynamically kept open by the ongoing performance of political action. The question if a public sphere can be maintained in a digital environment is important for understanding the possibility of politics in digital worlds.

Two concepts that are central to Arendt's political action are plurality and freedom. In essence, plurality is the condition that every participant in political action is uniquely distinct, but in a way that lets them perform political action together. This is related to two other concepts, that of equality and distinction, both of which might be under pressure in a digital environment. Freedom is also a necessary condition for political action. According to Arendt, political action is necessarily a free act. This freedom, in turn, rests on the concepts of unpredictability and irreversibility, both of which might also be jeopardized in digital environments. Through these concepts that Arendt developed, the possibility of political action in digital worlds can be investigated.

² Labor and work are interesting concepts in their own right, that might provide useful for an analysis of other categories of action in digital worlds, but we will only concern ourselves with political action here.

[5] Outline of the chapters

The concept of a digital world is still largely undefined, as tends to happen with predictions of future technology. It is important to get clear on the difference between digital worlds that completely facilitate someone's experience of life and prevalent instances of partial digital environments and online social networks. We also need to break down some of the properties of digital objects and digital worlds in order to effectively analyze them. Therefore, chapter one will start with a conceptualization of digital worlds. The first section of chapter one will set the bounds for the kind of digital worlds that are under investigation. Those worlds are interactive, social and complete. Then, section two will distinguish between the regular world and the digital world along the dimensions of simulation, virtuality and digitality. The final sections of chapter one will introduce the two theories of digital realism and phenomenology as a theoretical foundation for our analysis of digital worlds.

Next, the Arendtian theory of political action and the different concepts that constitute this theory must be explained more carefully. This will be the objective of chapter two. The first section of chapter two will characterize political action as a mode of action that is contrasted by labor and work. Then, section two will elaborate on political action as a dynamic mode of action that is performed in a space of appearance. The third and fourth sections of chapter two will introduce the conditions of plurality and freedom as necessary conditions for political action. Finally, the fifth section of chapter two will establish totalitarianism as the antithesis of political action.

Afterwards, chapters three, four, and five will apply the Arendtian theory of political action to the context of digital worlds. The guiding question of these chapters will be the possibility of Arendtian political action in digital worlds. The question of this possibility can be divided into two parts. First, we will ask whether political action is theoretically possible in any digital world. This question will be the focus of chapters three and four. They will discuss the conditions of plurality and freedom respectively in a context of theoretical digital worlds. As such, these chapters will be an argument for the theoretical application of Arendt's framework in the context of digital worlds.

Next, we will ask if, given the theoretical possibility of political action in digital worlds, such action is likely to emerge in realistically conceived digital worlds. This will be the focus of chapter five. The first section of chapter five will give an outline of a realistically conceived digital world. Then, section two will discuss the threat of totalitarianism along four vignettes of different political situations. Finally, the third section of chapter five will propose some principles that might enable political action in

digital worlds instead. Additionally, this chapter will illustrate the value of Arendt's framework for an applied understanding of digital worlds.

Chapter 1: Digital Worlds

In order to address the question to what degree political action is possible in digital worlds we first need a conceptualization of digital worlds. This chapter will introduce three core features that a digital world must have to properly investigate the possibility of political action, in the sense that without them we could not even start this investigation. Additionally, three different dimensions that differentiate digital worlds from the real, actual, and non-digital world will be clarified. Next, this chapter will examine two influential philosophical approaches to conceptualizing digital worlds, that of virtual realism and the phenomenological approach. By considering these different perspectives, this chapter aims to provide a comprehensive understanding of the ontological nature of digital worlds that bridges the gap between the analytical and continental philosophical traditions. Grounding our conceptualization of digital worlds in these distinct philosophical perspectives will allow us to develop a nuanced understanding of the potential for political action in digital worlds. This chapter will conclude by positioning ourselves between digital realism and the phenomenological approach along the three essential requirements and differentiating dimensions discussed in the first two sections. This will deliver a conceptual framework that will allow us to effectively evaluate the possibility of political action in digital worlds.

[1] Interactive, social, complete

There are many different ways to conceptualize and define digital worlds. Not all of these conceptualizations and definitions are equally useful to examine the possibility of political action or effectively evaluate the potential for political agency in digital spaces. Therefore, I will first give a characterization of the kind of digital worlds that we will investigate on the basis of three essential requirements. Defining these limits of the kind of digital world that is under investigation will allow us to effectively discuss questions about political action in later chapters.

Firstly, digital worlds should be *interactive*, meaning that action and engagement is facilitated by them. This aspect is fundamental because it forms the basis of all complex forms of action, which certainly includes political action. Secondly, they should be *social*, meaning that they are inhabited by multiple people who are able to interact with each other.³ This aspect is crucial as it allows for communication, cooperation, and coexistence, all of which enable the political dimension of action.

³ One could argue that the social requirement could also be fulfilled by having artificial intelligences to communicate with. Although the philosophical questions about personhood that this raises are important, we will only concern ourselves here with political action performed by human beings amongst each other. However, artificial intelligence could still have an impact on this political process, as we will see later.

Thirdly, they should be *complete*, meaning that no interaction is possible outside the digital world. This ensures that all interactions and consequences are within the boundaries of the digital world, which allows us to effectively evaluate the possibility of political action in digital worlds by isolating the effect that such worlds might have on political action. Establishing these requirements will deliver a basic conceptual framework that enables us to understand the potential of political action in digital worlds. This section will introduce each requirement by considering first how its antithesis, a digital world that does not meet the requirement, is not suitable for analyzing political action, and then examining how introducing the requirement will benefit this objective.

Interactive

One could consider a digital world without interactivity. This would resemble something like a very realistic movie or an experience machine (Chalmers, 2022, pp. 31–33). In order to emulate a credible world, this cinematic experience could be incredibly immersive. We can even imagine that someone in such an environment could look around and perhaps even move their perspective, while all kinds of events are unfolding in this digital world. Nevertheless, such a world would essentially be a passive experience. Someone or something other than the one who experiences the world determines what will happen every step of the way. In such a scenario, the possibility for any kind of action is eliminated. A world without interactivity would not be useful for analyzing political action, because it lacks the possibility for action in the first place.

In order for people to act in a digital world, therefore, it should be *interactive*. This means that someone inside a digital world can alter some aspects of the world itself. The digital world should then, in turn, react to a user's actions or give feedback to those actions. This constant interactive exchange between a person and their environment forms the basis of every kind of human action. Without it, the experience in a digital world would be like a ghost trapped in an environment that it cannot act in, but merely observe. The condition of being interactive, on the other hand, does not necessarily mean that everyone in a digital world should be able to alter the digital architecture that supports the world itself. This can be left to a group of specialized engineers. It simply means that people should be able to interact with the digital world from within this world.

Social

Secondly, one could consider a digital world without any social dimension. This would look more like a modern single-player video game. A player in such a game can probably interact with their

environment. In many single-player games, interacting is an essential component to advance through levels, whether it be a fighting game where one has to evade attacks and knock out opponents, a platforming game where one must maneuver through obstacles, or a puzzle game where interaction with the environment is required to solve the puzzle. Although playing a single player game might be an interactive experience, in contrast to observing a non-interactive digital world, it is also essentially a solitary experience. It lacks the presence of other people and the opportunity of social interaction, both of which are fundamental elements of politics. Therefore, such a world cannot serve as an effective tool for analyzing political action.

Instead, a digital world that allows for political action should be *social*. For a digital world to be social, there should be multiple individuals inhabiting the same digital environment. This requirement is partly captured by terms like MUD (Multi-user Domain) (Borgmann, 2000, p. 188). However, it is not enough to simply change the single-player experience into a multiplayer one by introducing other individuals into the environment. For a digital world to be truly social in a political sense, each individual should be able to communicate. Furthermore, they should not only be able to interact with the environment, but also with each other. In a political environment the words and deeds of one individual affects the lives of others. Thus, this should be the case in a digital world as well.

Complete

Lastly, the most common interpretation of digital worlds is one that is not complete but partial. Examples of these are social digital environments on the internet, which are typically social media platforms (Boyd, 2008; Papacharissi, 2004). These are digital spaces that accompany the physical world and facilitate social interaction in addition to it. These spaces are interactive in the sense that you can post content on them and evaluate other content through likes and dislikes. They are definitely social since social interaction is the main purpose of these platforms. Whether this means that political action would be possible on social media platforms is an important question that will not be pursued here however. Since the aim of this thesis is to analyze the possibility of political action in digital worlds considering the digital nature of such worlds, social media platforms make for a difficult subject. They are digital environments that are complexly interwoven with the physical world. Therefore, it is difficult to find out if political action is possible because or despite the digital nature of these environments.

In order to isolate the effect of digitality itself on political action, this thesis will discuss complete digital worlds. In contrast to partial digital worlds like social media platforms, inhabitants of a complete digital world do not perform actions and have no experiences outside of that world. This means that every action and experience is mediated by digital technology. Also in contrast to partial digital worlds in which we generally talk of users, in the case of a complete digital world we can properly talk of inhabitants because individuals fully live in it and have no existence outside of it. An assumption here is that immersion in a digital world to the point that one inhabits it, drives the need for political action. If someone is merely a user that has the choice between different worlds, then political involvement in a digital world might be nice, but does not carry the existential weight that characterizes political action. This does not mean that this existential immersion is automatically present in complete digital worlds, as we will discuss later, but when one has the option to step outside their digital life, the immersion is likely to disappear.

The second argument for incorporating the complete requirement is that it will prove useful for our analysis. If we assume that political action is possible in a totally non-digital world, then any possible impediment to political action in a complete digital world will be a consequence of the digital nature of such a world and not because of some other factor. This allows us to isolate the effect of digital worlds on political action.

[2] Simulated, virtual, digital

In addition to defining the properties of digital worlds that are required for political action we need to examine three dimensions in which digital worlds differentiate themselves from the real, actual, and non-digital world. Firstly, it is essential to examine the extent to which digital worlds are *simulated*, meaning that they are an imitation or a model of the real world. Secondly, we should consider the way in which such a world is *virtual*, as the extent to which meaning and interaction are being mediated by some (technological) medium. And thirdly, we should take the *digital* foundation of such a world into account, referring to the underlying method of its construction. This section will discuss these dimensions by contrasting them to the way the world is real, actual, and non-digital respectively.

These terms have been used in many ways and received different interpretations both in popular language as well as in academic literature. They are, however, not always useful as antonyms and synonyms for the physical world. There are ways in which this world itself, which is sometimes also

referred to as the brick-and-mortar world, is simulated, virtual and digital. Therefore, in the rest of this thesis, the physical brick-and-mortar world will be referred to as the *regular world* when it is contrasted with *digital worlds*, and the terms *real*, *actual*, and *non-digital* will be employed as contrasts to the ways in which digital worlds are *simulated*, *virtual*, and *digital*.

It is important to comprehend what the dimensions of simulation, virtuality, and digitality mean for the purpose of analyzing the possibility of political action in digital worlds because they are what differentiate them from the regular world in which such action is possible. This will allow us to position ourselves between two different philosophical approaches that have been applied in such an analysis, namely the digital realist and the phenomenological approach.

Simulated

A common way of describing a digital world is by calling it a *simulation* (Chalmers, 2022, pp. 9–10). Although this term applies particularly well when a digital world is supposed to be a re-creation of or even a substitution for the physical world, not all digital worlds are simulations. And even if they are, they can be simulations to higher or lower degrees, so using the term simulation synonymously with digital world can be misleading.

Simulations are always imitating something. Unlike static imitations like pictures or sculptures, however, simulations are replications of the behavior of some environment. If we take a certain environment that behaves a specific way under certain circumstances, then a simulation of that environment will behave in exactly the same way under the same circumstances. But, it is questionable if all circumstances of the regular world could ever be considered. Every simulation that currently exists is a partial simulation, and is therefore inaccurate at a certain degree.

Take the example of simulating a tree. One could simulate a tree so that it behaves as it would in the real world under the circumstances of observing the tree or wind blowing through it. One could walk around this simulated tree and observe it from multiple angles just like you would with the real tree. The simulated leaves could rustle in the wind just like the real leaves. If this simulation did not take the physical resistance of the real tree into account in its set of circumstances, however, one could move through the simulated tree, which would quickly break the illusion and make the difference between the real tree and the simulated tree apparent. A simulation could theoretically be a total replica of something real, but it is practically always the case that certain elements of a real environment are simulated, while others are left out.

A digital world can be a simulation of the real world to the degree that it takes elements of the real world into account. In addition to solely being a simulation, however, a digital world could also introduce new elements that do not exist in the real world. In a digital world, it is conceivable that trees could be elastic like rubber bands for example. This particular property would not be considered a part of the simulation of real trees, nor would it represent some element of those real trees that has been excluded from the circumstances of the simulation. Digital worlds cannot be defined as simulations of the real world. Instead, the degree of simulation should be understood as a dimension that differentiates the digital world from the real world, both in the sense that it does not simulate aspects of the regular world and that it simulates aspects that do not exist in the real world.

Virtual

The term *virtual*, like simulation, is subject to a range of many different interpretations in both popular and academic discourse. There are two common ways of using the term. First, the virtual could be opposed to the actual, where it acts like a layer of meaningful virtual appearance on top of the meaningless actual state of the world (Norton, 1972). Second, it could be used to refer to the realization of a virtual environment. In this sense it is commonly used as part of the term virtual reality, especially since the advent of VR-technology. In this way of using the term, the digital foundation of virtual reality is usually implied (Chalmers, 2022, Chapter 10). In order to separate the virtual and digital dimensions of digital worlds, virtuality will here be understood more generally as any way in which a virtual environment is realized.

The contrast between the virtual and the actual is useful because it allows us to avoid opposing the virtual with the real. Making this opposition would imply that virtual things and worlds are not real. As such, the meaning of virtual would be akin to 'artificial' or 'simulated'. Opposing the virtual to the actual, however, allows us to understand the ways in which the virtual is real. We can think of the virtual as a layer of meaning on top of the actual. Consider a wooden chair, for example. The virtual dimension of this chair is the way it appears to the observer as a chair. The actual dimension consists of the things that physically underlie this appearance. When one looks at a digital wooden chair through a screen, the actual physical foundation of this appearance has changed from wood to the, still very real, light-emitting-diodes of my display, but the virtual appearance remains that of a wooden chair.

Thinking of virtuality as a way of mediating interaction works very well with this real-but-not-actual interpretation. However, it focuses more on the way the virtual dimension is realized. In this sense, virtuality refers to the mediation of interaction that the medium or technology makes possible, e.g. one can see a chair or other individuals on the screen and interact with these objects and people through a computer. Because a specifically digital realization of virtuality is typically implied in this way of defining the virtual, especially when talking about virtual reality, the term might become redundant as a synonym for digital mediation; virtual reality would simply mean digital reality. Here, however, I will understand virtuality more generally as a dimension of digital worlds that differentiates it from the physical world through a different actualization of meaning and interaction where these are mediated by some medium.

Digital

The final dimension of *digitality* refers to the specific method in which both simulation and virtuality in digital worlds are realized. The core of this method consists of encoding information into a binary string of numbers which can be manipulated, quickly communicated, and decoded into information. This method has its own theoretical and practical history and it is the foundation of the information and communication revolution of the last decades. Additionally, it turned out that digitalization is incredibly useful for simulation and virtualization purposes. Therefore, we will concern ourselves here with a digital realization of such simulated and virtual worlds. Further questions whether simulation or virtuality can be effectively realized in some non-digital way and if political action might be realized in some artificial but non-digital world are interesting but go beyond the scope of this thesis.

There are many ways in which the regular world could also be considered partly digital. The dimension of digitality contrasts digital worlds to the regular world through the extent to which virtuality in digital worlds is realized through a digital framework. By understanding digitality in this way, we avoid synonymizing virtual worlds with digital worlds. We also avoid thinking of every digital world as a simulation of the regular world. The dimension of digitality both grants possibilities and introduces limitations. These possibilities and limitations of digital worlds can be different from either the regular world, simulations of the regular world and other realizations of virtuality.

In other words, just like the regular world is governed by certain natural laws, so will the digital world be governed by digital laws. These laws are determined by what is either possible or impossible for that digital technology to realize. An object that one encounters in the regular world, like an apple,

will always fall to the ground. Likewise, a digital apple has its own laws. Digital objects are, for instance, *perfectly multiplicable*, meaning that a perfect duplicate of any digital object can be made at any time, they are *ephemeral*, meaning that they can be completely erased without leaving a trace in the world, and they *lack temporality*, which means that they endure through time without deteriorating (Kim, 2001). There does not necessarily have to be a metaphysical difference between objects in the regular world and digital objects, as we will see in the following section. These differences will still have an impact on one's experience of a digital world however. When I talk about a *digital world*, I refer to a world that is governed by these digital laws.

Now that we have defined the three dimensions of simulation, virtuality and digitality of digital worlds that contrast them to the real, actual, and non-digital world respectively, we are ready to discuss two different philosophical approaches to analyzing digital worlds. These approaches will result in different perspectives on the interactive, social, and complete requirements of digital worlds and how the simulated, virtual, and digital dimensions of such worlds affect these requirements. We will continue in the last section of this chapter to position a conception of digital worlds between these approaches.

[3] Digital realism

A first philosophical approach to characterizing digital worlds is a position that I will call digital realism. One of the core tenets of this position is the claim that digital worlds are physical in exactly the same way that the regular world is. According to Chalmers, this leads to the proposition that a digital world could behave in exactly the same way as the regular world does, which is backed up by a thought experiment of perfect simulations (Chalmers, 2022, Chapter 2). This makes digital realism a good candidate for giving a positive evaluation of the possibilities of digital worlds.

Chalmers defends digital realism by making a structuralist argument (Chalmers, 2017, 2022, Chapter 22). According to structural realism, a mathematical or computational structure is either all there is to the world or it is all we can know about it (Chalmers, 2022, p. 389; Ladyman, 2020). This means that the reality of objects is defined by their underlying structure: how they stand in relation to other objects and how they behave in relation to each other.

For example, consider the game of chess. According to structural realism, the reality of the knight on the chessboard is defined by the structure of the game itself. The piece has a position on the board, it can move in a certain way, and it will interact with other pieces in a manner determined by the game's structure. The physical actualization of the piece, such as a carved wooden figure, or the name given to it, such as 'knight', are not relevant to its reality.

Most importantly, for our purposes, structural realism asserts that if one thing or environment has a particular structure and is assumed to be real, then anything that shares that structure is equally real. In the case of digital worlds, if the structure of the digital environment is indistinguishable from that of the regular world, then it should be considered equally real. Therefore, digital worlds can be just as real as the regular world if their underlying structures are identical.

If we take a perfect digital simulation as a thought experiment, we can see how a structuralist view can take us to digital realism. For a perfect digital simulation we disregard practical limitations such as memory space and power supply. In such a perfect simulation it is possible to simulate the causal structure of reality as it is described by the laws of physics, given that these laws can be described by digital models. The perfect digital simulation could have atoms and gravitational laws, all represented by digital objects. Since this simulation is not bound by practical constraints, it would be possible to model the entire physical universe in this digital environment. That means that the structure of a perfect digital simulation could be indistinguishable from that of physical reality, even though this structure is realized in a different way. If, like the structural realist claims, reality is defined by structure, that would mean that a perfect digital simulation is real in exactly the same way as the regular world is real.

The consequence of digital realism is that the property of digitality of a digital world does not differentiate it from the regular world. If a digital world would lack something, it would not be because it is digital, but because of some other factor. In this sense, digital realism could be captured in the following proposition: if the physical world exists, then digital worlds exist. You could, for instance, disagree that physical laws actually describe reality. Perhaps physical laws only approach reality or perhaps an epistemic barrier prevents us from grasping reality as it is. Although these may be legitimate concerns, they challenge the antecedent of the digital realist proposition, not the proposition itself. According to digital realism, the digital nature of a digital world does not fundamentally prevent it from attaining the status of reality.

An interesting facet to the digital realist position that cannot be fully discussed here is the digital requirement of this position. One could, for instance, challenge this digital requirement, while remaining a virtual realist. This would mean that you adopt the position that virtual environments can be as real as physical reality, but argue that this virtual environment must be realized in some other non-digital way. The question if any other non-digital ways of realizing a virtual environment allow for more possibilities might be to pursue further, but since this thesis is concerned with digital worlds we will restrict ourselves to digital realism, rather than exploring other variants of virtual realism.

On the question of the possibilities of digital worlds, digital realism nudges towards a positive evaluation of such possibilities. It offers, for instance, convincing rebuttals to arguments that introduce some sense of naturalness as a requirement for proper action or communication. A digital realist could point out that action and communication rarely happen in a totally natural setting. Some level of artificiality in our environment is nothing out of the ordinary. Someone who lives in an urban area, for example, is likely to almost exclusively encounter artificially constructed spaces. Some people might value natural environments, but this is nothing more than an “optional preference” (Chalmers, 2022, p. 307). If one accepts that the artificiality of these spaces does not affect the possibility of human action and communication within them, it seems like a small step to substitute them with totally digital environments.

An important assumption that digital realism makes, however, is that people and the environments they inhabit can be understood separately. This does not suggest, however, that we should focus our attention to a digital realist interpretation of minds. It is a far stronger claim and more difficult, though not impossible, to defend that minds can be realized digitally than to argue that the environment between those minds can be realized digitally. The question of action, what people can and cannot do in a digital environment, remains central to this thesis however. This means that, without delving into the intricacies of the mind, we do need to understand the relation between people and their environment. This being the case, however, it is still not immediately obvious that we can conceptually separate people and their environment, as we will see in the next section.

In conclusion then, digital realism is a philosophical position that is able to accentuate the way in which digital worlds are real. By doing so, it allows us to positively define the possibilities of digital worlds by drawing on what is possible in the real regular world. Although many further questions could be asked about the practical realization of digital worlds, digital realism offers a considerably

optimistic outlook on the theoretical possibilities of such worlds. If no aspects of reality exist that cannot be modeled in a perfect simulation, such a simulation would most likely allow for the same kind of action as the regular world. There are, however, other ways of interpreting digital worlds that do not take the reality of a digital environment, but the relation between such an environment and the people living in it as the point of focus.

[4] Phenomenology of digital worlds

A second way of philosophically characterizing digital worlds is by adopting a phenomenological perspective. Phenomenology can be seen as a distinct discipline within philosophy that concerns itself with phenomena, how things appear in our experience in a meaningful way. It can also be regarded as a historical philosophical movement that originated in the early twentieth century (Smith, 2018). Either way, it is a complex tradition with numerous different interpretations. We will concern ourselves here with the phenomenological approach that Albert Borgmann and Hubert Dreyfus have developed around the turn of the millennium to launch a critique of the upcoming phenomenon of digital worlds (Borgmann, 1987, 2000; Dreyfus, 1999, 2001). Both of these thinkers were informed by the theoretical foundation of phenomenology as it was laid out by Martin Heidegger (1954/1977, 1927/2008).⁴

This section will start by introducing some of the central concepts of this phenomenological approach. Then, two insights that are relevant to this thesis will be introduced. The first is the *device paradigm*, introduced by Borgmann (1987), that illustrates the way in which digital objects appear to us as easily available while concealing the complexity of how they relate to the world. The second are the notions of *situated risk* and *shared moods*, which are essential for meaningful experience, but might be absent in digital worlds according to Dreyfus (2001).

According to the phenomenological approach, thinking of people as living in a world of external objects does not properly explain how these objects acquire their meaning or how human action functions in relation to those objects. A common assumption in Western philosophy is that individuals encounter specific objects, like a hammer or a nail, and from these instances extract a general concept of hammers and nails. One of the central claims that Heidegger puts forward is that, while this theoretical framework of meaning is useful for understanding theoretical concepts, in order to understand how the world becomes meaningful for human beings, it needs to be turned

⁴ Additionally, Dreyfus also draws heavily on the work of Merleau-Ponty and Kierkegaard

around (Heidegger, 1927/2008, pt. 1). Instead of moving from objects to general concepts, he introduces the idea of a background world of meaning that is always already there, and in which we are always already situated. From our situatedness in this background world we project a range of possible actions, some of which are actualized, thereby in turn creating a new situatedness. From this background world, things show up as meaningful to us through the kind of actions they afford. A hammer shows up as something we can hammer with and a nail shows up as something that can be hammered (Heidegger, 1927/2008, p. 98). It is only later that we assign explicit definitions like hammer and nail to these already meaningful things. Showing up will be used here as a technical term denoting the way things appear as meaningful in this phenomenological sense.

Two important ideas can be derived from this conceptual framework. First, people cannot be understood separately from their environment, especially not from the tools and technology that furnish this environment. For example, consider when someone has a friend living far away. That friend will be a different person depending on whether they own a phone or not. When they do own a phone, they will show up as contactable, whereas in the latter case they will not. According to phenomenologists, showing up as contactable or not is not just a trivial matter of how people appear to each other, but a non-trivial difference in the way their meaning is constituted. In other words, people and the phenomena they encounter, including the tools and technology that shape these phenomena, have a co-constitutive relation (Heidegger, 1954/1977, p. 12).

Second, Heidegger introduces the idea that the way phenomena show up depends on the mood we are in (Heidegger, 1927/2008, pp. 172–179). This idea can be intuitively explained by taking examples like an optimistic joyful mood and a pessimistic somber mood. When in a positive mood, objects and possibilities for action will appear more joyful and optimistic, while in a negative mood the opposite is the case. Furthermore, Heidegger postulates that moods are important to the way people live together. So called shared moods, like the ones that crowds experience during an exciting show or football game, allow people to feel like they belong together and consequently allow them to act together. In this case the word ‘attunement’ better captures the meaning of the original term ‘*Stimmung*’ for mood.

Borgmann employs the first of these ideas to develop a critique of the increasingly prevalent role that digital technology plays in societies. He does so by presenting a conception of technological devices that illustrates how these devices show up to us as easily accessible commodities while at the same time concealing the complexity of how they relate to the world. Borgmann calls this way of

understanding technology the *device paradigm* (Borgmann, 1987, Chapter 9). Take a heater that replaces the hearth of a house for example. We can warm up a room by simply turning on the heater, which makes warmth an easily accessible commodity. The heater then shows up as something that can warm the room by the turn of a knob and the house shows up as something that can easily be heated. In the case of a hearth, warming up the room shows up as a complex activity. One has to collect the firewood, place it in the hearth, ignite it and maintain the fire. In fact, however, the heater has a way more complex relation to the world. Gas must be extracted from the earth and moved to a central boiler where it heats up water that runs through the heater. However, this complexity is hidden from the user of the device .

The *device paradigm* can be applied to digital worlds as well. For the inhabitants of such a world, the digital environment will be easily accessible. They could move their digital bodies around with similar ease as in the regular world, but the complexity of how the world is realized will be hidden. In this way, these inhabitants of digital worlds are detached from the reality of their world. To a certain extent this veiling of complexity is necessary to effectively interact with digital objects (Borgmann, 1987, p. 44), and it is part of the appeal of digital worlds in the first place (Borgmann, 2000, p. 189). If we consider this from the digital realist position, we can see that these approaches result in contradicting conclusions. Where a digital realist would argue that because the structure of a digital world is the same as the real world, there is no difference between the two, a phenomenologist could reply that how that structure is realized does matter for the inhabitants of digital worlds because they have a constitutive relation to them. In other words, even though the complexity of a digital world might be hidden from its inhabitants, they are in part defined by it.

From this conclusion arises a twofold concern. First, we might worry that our detachment from the complexity of the devices that define our world might have negative consequences. Second, we might fear that by mediating interaction between people through the easily accessible front end of our devices we would risk reducing the complexity of those people to mere available commodities.

Dreyfus employs the second concept of *shared moods* to describe what might be missing from interaction in digital worlds (Dreyfus, 2001). He imagines a digital classroom where both the teacher and students are present via screens. What is different from such a teaching environment compared to a physical classroom, Dreyfus argues, is that a teacher in a physical classroom can feel the mood of the attending students and can discern their interest, skepticism or loss of attention. In a digital classroom, such a shared mood is difficult to achieve. Even when students are able to digitally raise

their hands or share their emotional status, the physical subtleties that are involved in being in a room together appear difficult to simulate in a digital environment.

One counterargument to these concerns is that a solution might be offered by transferring more information of people's body language into the digital environment and removing the barrier of users having to actively input this information. However, Dreyfus maintains that even a very sophisticated simulation of people's bodies might still not be enough to create the right conditions for being together digitally (Dreyfus, 2001, pp. 56–57). He points out that in an ideal classroom scenario a certain sense of risk is involved (Dreyfus, 1999, 2001, p. 54,57,87,88). This risk is essential to making the class a meaningful experience. Through the possibility of asking questions, both the teacher and the student risk not knowing the answer for example. But this risk is not something like social embarrassment, though that might definitely be involved.

What Dreyfus aims at is a more fundamental *situated risk*. By being situated as a teacher in a classroom, someone inherits the responsibility that comes with being a teacher. When a teacher blunders, they are not at liberty to change their role, like someone in a digital world might do by changing avatars, but have to correct the mistake in order to be considered a good teacher. Likewise the students experience their own variants of situated risk. Only this kind of risk might allow all participants of a classroom to experience a shared mood of *being in a teaching/learning environment*.

We have now explored two ways in which the phenomenological approach can be used to voice concerns about the possibilities of digital worlds. The device paradigm tells us that even though a digital world might simulate a structure that is similar to the real world, the underlying phenomenological difference between those structures might still be significant. Situated risk as a requirement for shared moods tells us that even though we might interact socially through digital mediums, we need to be properly situated in a digital world to make that interaction meaningful. This phenomenological approach creates a contrast with the digital realist approach that mainly focuses on the ways in which digital worlds are similar to the real world.

[5] Conclusion: Conceptualizing digital worlds

The first section of this chapter developed a fundamental understanding of digital worlds. This conceptualization introduced the properties of being interactive, social, and complete as necessary

properties for effectively investigating the possibility of political action in digital worlds. The second section of this chapter introduced three dimensions according to which digital worlds can be differentiated from the real, actual and non-digital world. These dimensions can be used to highlight the multiple ways in which digital worlds and the real worlds can be different as well as the ways in which they are similar.

As we saw in sections three and four of this chapter, however, different philosophical positions might lead to contrasting interpretations of the possibilities and drawbacks of our fundamental understanding of digital worlds. Since these contrasts will affect the possibility of political action in digital worlds, we need to consider them in the terms of our initial conceptualization. Therefore, we will now position ourselves between the discussed approaches along the requirements of interactive, social, and digital. This will allow us to incorporate arguments from both approaches in our further discussion of political action.

First, we will examine the philosophical perspectives on the requirement of *interactivity* in digital worlds. The way that one defines the nature of this interactivity, will impact what kind of action is possible in a digital world and how it will compare to action in the regular world. Digital realism proposes that digital environments and the digital objects in those environments are real just like the non-digital world. When an actor interacts with a digital world, this causes real changes that will in turn influence the actor. Digital realism posits that, since the causal structure of the regular world could be simulated in a digital world, the possibilities of actions will be the same. According to phenomenology, however, interaction with a digital world that is differently realized than the non-digital world will affect the actor because they are constituted by the digital environment they interact with. This means that even if the causal structure of the regular world can be simulated, the digital world might still phenomenologically show up in a different way.

Second, consider the philosophical divergence towards the requirement of digital worlds to be *social* in order to engage in political action. The way one defines interactivity and communication between people and what one takes to be required for people with different wants and needs to live together in a digital world will impact what kind of social life can be realized in a digital world and how it will compare to the regular world. From the digital realist perspective, the digital world in between inhabitants could be seen as an artificial construction that replaces other artificial constructions like cities that are common-place in the regular world. In the digital world then, social life could be mediated in a way that is similar to the regular world. The phenomenological perspective posits,

however, that there is a difference between the virtuality of the regular world and a digital world. A virtuality that reduces people to the availability of devices might be detrimental to meaningful social interaction. In addition, an absence of situated risk could mean that people live next to each other in a digital world, but not together.

Third, the requirement to live in a *complete* digital world also leads to different positions based on the philosophical perspective one takes. One could argue that living in a complete digital world is a matter of mediating every action through digital technology. In that case one could perhaps still choose to enter and exit a digital world. According to the phenomenological perspective, however, to be in a complete digital world might mean that one is completely situated in such a world. To be in such a state might mean that one is born in a digital world without an alternative.

How the two different philosophical perspectives of digital realism and phenomenology further differ in their interpretation of the ontology of digital worlds and the way individuals are positioned in relation to such worlds are complex questions. Although these questions merit more academic attention in their own right, this thesis is primarily concerned with the possibility of political action in digital worlds. As such, the differences between these positions are only relevant to the extent that they help shed light on the question of political action. How these differences will affect this question will be discussed in greater depth later. Before we turn our attention to this, we need to establish a clear definition of political action, which will be the focus of the next chapter.

Chapter 2: Political Action

In her book *The Human Condition* Hannah Arendt develops, among other things, a conceptual framework of three fundamental modes of human activity (Arendt, 1958/2018). These modes are that of *labor*, those actions that are necessary for the sustaining of life; that of *work*, those actions that produce something durable in the external world; and that of *political action*, those actions whereby people actively reveal “their unique personal identities [...] in the human world” (Arendt, 1958/2018, p. 179). One of the underlying motivations for Arendt’s project is to reveal the way that totalitarian societies, especially those that are grounded in Marxist ideology, grant the highest value to the activity of labor. This drives her to “rethink the hierarchy of [those] modes of activity” (Kohn, 2000, p. 123), in a way that emphasizes the central importance of political action to human life.

If it turns out that political action, as Arendt conceptualizes it, cannot be realized in digital worlds, or is impeded in them, then Arendt’s critique of totalitarian societies will be a helpful analytical tool in any analysis of digital worlds. In order to properly investigate political action in digital worlds, we first need to understand the way that Arendt conceptualizes it. That will be the aim of this chapter. The first section of this chapter will expand on the distinction between labor, work, and political action. Section two will illustrate both the importance that Arendt allots to the political mode of action and how she understands it as something that cannot be taken for granted. Then, sections three and four will discuss two concepts that are essential to the realization of political action, namely those of plurality and freedom. Finally, the last section will explore the concept of totalitarianism as the antithesis of political action.

[1] Labor, work, and political action

Arendt differentiates between three fundamental modes of human action. She names these modes labor, work, and action (Arendt, 1958/2018, p. 7). The last of them, action, is unique to humans and essential for human beings to exist *as people*. A life without action, she emphasizes, “has ceased to be a human life because it is no longer lived among men” (Arendt, 1958/2018, p. 176). In this sense, Arendt’s term ‘action’ is both used to denote something more specific than general human action and something more exceptional than the “much larger field of what is conventionally called politics” (Kateb, 2000, p. 131). This thesis will use the term *political action* for what Arendt calls action. This will avoid confusion when talking about labor and work as being different modes of human action and highlight how political action is fundamentally dependent on a political context, as well as differentiate it from conventional politics. In order to understand Arendt’s concept of political action,

we first need to discuss her theoretical framework of modes of human action. This will allow us to define political action in contrast to labor and work.

The first mode of action that Arendt introduces is that of *labor*. These are all activities that are required to sustain life. Eating, drinking, and procreation are typical examples of labor activities. Importantly, labor is always a necessity. Nobody ever chooses to labor. It is imposed upon people by the necessity of having to stay alive. There might still be other dimensions to an activity. Eating a meal might be inherently enjoyable for example. It does not have to be labor exclusively. The activity of eating a meal is only labor to the extent that it contributes to sustaining life. Therefore, labor only has one product, which is life itself.

Although a certain degree of labor is per definition a necessity for everyone and everything that is alive, there are different ways in which the amount of labor that is necessary can be reduced. One such mechanism is the distribution of labor among individuals. The harvesting of grain, for instance, is an example of labor to the extent that it is necessary for the production of bread, but not every consumer of bread has to labor in the grain field. Importantly, labor can also be reduced through the introduction of external tools.

External tools that can reduce labor can be the product of the second mode of action, which is that of *work*. These are activities that produce something lasting and external to the individual who is performing that activity. Examples of these activities are the production of tools like a hammer or a combine harvester, or the building of a house. Work is not limited to the production of concrete and pragmatic objects, however. A policy maker creating a political plan and an artist creating art are examples of work as well. Through the process of work, humans produce an external world that is filled with artifacts. These artifacts can serve a myriad of functions and possess cultural meanings. Just like an activity is labor to the extent that it produces and sustains life, an activity is work to the extent that it produces these external artifacts.

The final mode of action is that of *political action*. Unlike labor, political action is not something that is driven by necessity, and unlike work, political action does not produce anything external to the performer. An activity is political action, according to Arendt, to the extent that it reveals the actor as an individual in a world that is shared with other individuals. In this sense, the product of political action is people. The term that Arendt herself uses is 'men' (Arendt, 1958/2018, p. 7), but here we will denote the product of political action as 'people' or 'personhood'. This makes political action, for

Arendt, something integral to being human. Someone that does not perform political action is not a person, but “is literally dead to the world” (Arendt, 1958/2018, p. 176). Consequently, political action is a process that requires perpetual engagement. In order for humans to exist as people they need to continually reveal themselves to others through political action. This is a strong requirement. An interesting question that could be pursued elsewhere is to what extent beings are still human without being people. In any case, political action is a significant aspect of human existence and indispensable for living a meaningful life.

In addition, Arendt emphasizes that the realization of political action is fundamentally dependent on labor and work. Without labor, human beings would not be able to survive. In a life that is constantly on the brink of survival, however, there remains no space for political action. It is only through the distribution of labor among individuals or the alleviating of labor through tools that space for political action can be created. Likewise, political action requires a stable external world, which can only be created through work. The totality of external artifacts that are the product of work create something that Arendt calls the *human artifice*. Without such an artifice, the world “could never be a reliable home for [people]” (Arendt, 1958/2018, p. 167). Although labor and work provide the essential infrastructure for political action, they are not yet sufficient conditions.

[2] The nature and importance of political action

The previous section has introduced the concept of political action as a mode of human action. It has been contrasted to two other modes of human action, that of labor and work. This has revealed political action as a concept that is essential for human beings since they would not be people without it. In addition, political action was defined as an activity that continually needs to be performed, and fundamentally depends on work and labor. This section shall expand on these notions by discussing how political action and personhood are related, how political action is able to emerge, and how it is inherently a dynamic process. This discussion will introduce two essential concepts to political action, that of plurality and freedom, which will be the focus of the following sections.

Political action and personhood are deeply linked for Arendt. The previous section discussed how Arendt characterized labor as being driven by necessity. It is not necessary for every person to labor, however. Even though a bare minimum of effort is required to sustain one’s life, theoretically almost

any aspect of labor can be alleviated or distributed to others.⁵ Similarly, not every person has to work. Because the human artifact is made up of durable objects, one could make use of the objects that others have left behind without ever adding something of their own. No one, in stark contrast, can refrain from political action and still remain a person (Arendt, 1958/2018, p. 176). This does not mean that a being that does not perform political action “would no longer be human” (Arendt, 1958/2018, p. 10), but such a human would lose their personhood.

How, then, should we understand this personhood that is the product of political action?

Importantly, the link between these concepts reveals that, for Arendt, appearing in the world constitutes the existence of a person, not the other way around.⁶ In other words, being a person is not a static property that human beings are born with and retain for the rest of their lives.

Personhood has to be created by appearing in the world through political action, which has certain requirements. Arendt describes this world, in which human beings appear as people, through her concept of a space of appearance.

A space of appearance is a virtual space that is created wherever people act together. That is to say “where people are *with* others and neither for nor against them” (Arendt, 1958/2018, p. 180). So, it is different from acting in concert. Rather, it is acting with a shared sense of each other’s personhood, creating a “space where I appear to others as others appear to me” (Arendt, 1958/2018, p. 198). Through acting politically, an actor’s personhood is revealed to everyone in the space of appearance except to the actor themselves. This is an important point for Arendt, because if political actors became revealed to themselves, this would make their appearance an external object that can be created and manipulated, which is characteristic of the work mode of action. Instead, there is a great unpredictability and risk involved with appearing to others through political action.

What, then, are the conditions that allow human beings to act together and reveal themselves to others? Two important concepts that enable this process, according to Arendt, are plurality and freedom. She defines plurality as “the basic condition” (Arendt, 1958/2018, p. 175) of political action. Acting together, for Arendt, means that people appear to each other as possessing the property of plurality. This means that they appear as uniquely different, yet with a common ground that is their

⁵ One could even imagine technology alleviating labor at the most basic level, where feeding tubes and sperm banks replace nourishment and procreation.

⁶ Arendt further claims that “the reality of the world is guaranteed by the presence of others, by its appearing to all” (Arendt, 1958/2018, p. 199).

shared personhood. What these properties entail will be discussed in further detail in section three of this chapter. Furthermore, in order to reveal themselves as people in the space of appearance, the acts of human beings must be governed by freedom. This means that through their action, they bring something unpredictably new into the world that has irreversible consequences. What acting in this fashion entails will be discussed in section four of this chapter.

Whenever and wherever people act together on the basis of plurality and according to the principle of freedom, they open a space of appearance. The flipside of this coin, however, is that whenever people stop acting together in this way, the space of appearance closes. Like personhood, the space of appearance is not a static environment, but a dynamic process. It cannot be created as a durable external object. This underlines the importance of the concepts of plurality and freedom. Whenever these concepts are lost, the space of appearance closes, which in turn means that personhood itself is lost. In such an environment, human beings might still be able to survive through labor and create an external world through work, but they will not inhabit this world as people.

In conclusion, the significance of political action in digital worlds becomes evident when viewed through the conceptual framework offered by Arendt. This conceptual framework of political action urges us to pose the question of the possibility of such action in digital worlds in a specific way. The question central to this thesis cannot be addressed by assuming the personhood of human beings and subsequently asking if digital worlds enable them to act politically in a way that communicates this personhood. Instead, the crucial question of inquiry is whether digital worlds facilitate the proper conditions for human beings to reveal themselves to each other in such a way that establishes their personhood in the first place. Only when these conditions are met, are human beings able to act together, which would mean that political action is possible in digital worlds. In order to answer this question in the context of digital worlds we must understand the concepts of plurality and freedom as they are developed by Arendt. Furthermore, we must discuss the political implications that might arise from their absence, which is the possible emergence of totalitarianism.

[3] Plurality

For Arendt, political action rests on the fundamental condition of plurality (Arendt, 1958/2018, p. 175). This means that the space for political action only opens when humans appear to each other as satisfying the condition of plurality. Arendt further defines this condition of plurality as consisting of both *distinction* and *equality*. This requires plural beings to be both different from each other, while

sharing the quality of personhood. This section will discuss the condition of plurality by considering the nature and requirements of both of these two aspects. This will help us understand the condition of plurality in a way that allows us to critically analyze the possibility of its realization in digital worlds.

Distinction

Distinction is the aspect of plurality that requires plural beings to be different from each other. The kind of difference that Arendt has in mind here is unlike the basic way in which all things are different from other things. To explain what she means, she contrasts distinction with otherness. All objects that we can differentiate between in the world possess the quality of otherness, simply by virtue of being one thing rather than another. The quality of distinction, on the other hand, is only possessed by things that are alive and have a distinct set of wants and needs. Everything that is alive is distinct in this way. A plant might be thirsty in the sense that it requires water, and two plants from the same species could be distinct from one another when one, for example, has just received water and is not thirsty anymore.

The kind of distinction that Arendt has in mind as an aspect of plurality is still different from this basic idea of distinction, however. It could be characterized as *unique distinction* instead. This unique distinction, that shows a being's plurality, is both different from "otherness, which [it] shares with everything that is, and distinctness, which [it] shares with everything alive" (Arendt, 1958/2018, p. 176). In order to possess unique distinctness, beings must be able to express themselves as a person with unique wants and needs. A dog, for instance, might be able to communicate hunger or thirst, but, according to Arendt, it is not able to communicate *that* it is hungry.⁷ This is the ability that allows people to reveal themselves in their unique distinction. They are able to express themselves, not merely as something that has hunger, but as someone who is hungry.

Equality

In addition to unique distinction, Arendt argues that *equality* is one of the defining aspects of plurality. While beings must appear uniquely distinct from each other in order to satisfy the condition of plurality, there is also a sense in which they must appear as being alike. What, then, do beings share when they appear as equals in their plurality? In this context Arendt does not refer to a kind of

⁷ For Arendt, this is one of the important dividing lines between humans and animals. How sharp this line truly is, might be interesting to explore further.

social or economic equality, but to a way of understanding others. For Arendt, equality allows beings to “understand each other and those who came before them [...] and foresee the needs of those who will come after them” (Arendt, 1958/2018, p. 175). In order for two humans to be equal they both need to understand each other as uniquely distinct beings. Furthermore, they need to understand each other as being part of the same space of appearance, a space that some have inhabited before and others will inhabit in the future. Although each person in a space of appearance is uniquely distinct, they are all equal in their distinctness and in their relation to this space. Only if people appear to each other equally as such, can they satisfy the condition of plurality.

Notably, Arendt develops the concept of equality in such a way that it fundamentally allows for unique individual expression. For her, equality in political action does not mean that people should be equalized in any common dimension of difference like race, social class or gender. Any equalization of such dimensions would imply a certain way that political action should be performed or what its actors should be like, which is exactly what Arendt tries to avoid. However, there might exist a danger here of ignoring structural differences in society that prevent the performance of political action. A theory of political action that looks to overcome these differences might have to take them into account more thoroughly.

The concepts of distinction and equality are crucial to the condition of plurality, which in turn is a prerequisite for political action. In order to act with each other as plural beings, both of these aspects must be realized. That implies that for the inhabitants of digital worlds to engage in political action, they must be able to reveal their unique distinction, and understand each other as equal within their shared space of appearance. However, before delving into this requirement in the context of digital worlds, we must first turn our attention to a second concept that is necessary for political action: freedom.

[4] Freedom

The condition of plurality describes the context that is required for political action. It is only in a context where beings appear to each other in their plurality, that the possibility of a political act emerges. Arendt calls those political acts “deeds” (Arendt, 1958/2018, p. 178). Through performing deeds, an actor reveals who they are as a person in the space of appearance. However, there are some requirements in addition to the context of plural actors, for the performance of those deeds themselves. To perform a deed, for Arendt, is to act under the principle of *freedom*.

Deeds are the only type of action that can be considered truly free. In this sense freedom is not strictly a condition of political action, but a consequence of it. When we analyze the aspects of freedom, however, we can also conceive of freedom as a precondition, because any action that would lack these aspects would not be political action. To qualify as free, an act needs to exhibit the aspects of *unpredictability* and *irreversibility*. First, a deed is unpredictable if the actor cannot anticipate the consequences of the action that they perform. Second, a deed is irreversible in the sense that once it has been carried out, it cannot be undone. This section will explore these two aspects of deeds by considering how their absence would result in a loss of freedom. This will allow us to better understand freedom as the second condition of political action alongside plurality.

Unpredictability

The first aspect of free deeds is their *unpredictability*. When an actor performs a deed they can never predict its outcome. Earlier we already touched on the notion that in the space of appearance an actor reveals their personhood to everyone except themselves. Both these aspects of political action emanate from the same idea. The unpredictability of deeds arises from both the inability of people to predict “who they will be tomorrow, and [to predict] the consequences of an act in a community of equals” (Arendt, 1958/2018, p. 244). If an actor were able to predict the outcome of an action, the action as such would not be free according to Arendt. Likewise, in order to appear to others as a free person, an actor can never anticipate who they will become as a result of their actions.

The connection between freedom and the unpredictability of deeds becomes evident when we examine the scenario in which an actor is able to predict the outcome of their deeds. In such a case the outcome of a deed becomes something outside of the actor that can be worked towards. One could adjust their behavior in such a way that the desired consequences of the act are reached. This would mean that the personhood of an actor becomes an object that can be created, turning the actor into an object rather than a person. As a result, the action would be fixed by the desired outcome, instead of being the outcome of freedom. This would make it more likely to be a product of the work mode of action. In any case, it would no longer be considered political action.

Irreversibility

The second aspect of free deeds is their *irreversibility*. This is the notion that once a deed has been performed, it cannot be retracted by the actor. Rather, its consequences are now permanently integrated in the space of appearance wherein it was performed. This space is opened and held open by the continuous performance of political action itself. Being able to retract a deed would affect this very process and therefore affect how humans appear as people.

Again, we can see how this process would be affected by comparing it to the working mode of action. In this mode the actor is able to destroy what they have created. If this were possible for political action, then the actor would be able “to *undo* what he has done as he undoes an unsuccessful object, by means of destruction” (Arendt, 1958/2018, p. 238). This would mean that the outcome of political action, which is personhood, would be an object like everything that makes up the human artifact. These objects could be removed, by the actor or by others. As such, an act that is reversible can never be free because it is always bound by this contingency.

Both unpredictability and irreversibility present crucial aspects of free acts. In their absence, an act would not align with the principle of freedom, and therefore would not be an example of political action. This turns freedom into a precondition of such action. So far, we have discussed the two conditions of plurality and freedom as necessary for political action. This has revealed four requirements of political action, namely distinction, equality, unpredictability, and irreversibility. Before we can explore how these requirements could be met in digital worlds, we first need to consider their absence in the regular world. This will give us insight into what a world without political action would look like.

[5] Totalitarianism

It is largely due to circumstantial factors that political action turned out to be one of Arendt’s main academic interests. Being from both Jewish and German descent, the dangerous political situation created by the Nazi Party forced her to leave her country of birth. In her eyes, it was exactly the failing of political action that had allowed such a calamity. The absence of political action had allowed for the rise of totalitarianism. As such, the majority of Arendt’s work can be considered to be an attempt to find an antidote to totalitarianism, which has not vanished with the end of the Second World War, but continues to be a threat to modern societies. Totalitarianism itself is a term that has been widely used in both academic and everyday discourse. Commonly, it is understood as a political system that aims to exert total control over individuals and public opinion. In Arendt’s

conceptualization, however, it can also be understood as the antithesis of political action. This definition is useful for our purposes because it allows us to understand totalitarianism as a political system that influences the possibilities of human action. In this sense, totalitarianism aims to control and align the actions of its subjects.

In this section, I will examine the concept of totalitarianism by considering the absence of the requirements for political action. This will define a space that is vulnerable to the emergence of totalitarianism. Such a space would be characterized by the *lack of distinction and equality*, where actions become *predictable and reversible*. In addition, I will delve into the concept of loneliness, which Arendt introduces as a phenomenon that is both a product and a catalyst of totalitarianism. Here, I will discuss how loneliness and negative conditions of political action are related. Understanding these concepts will enable us to investigate them in the context of digital worlds. This will both allow us to critically evaluate the likelihood of the emergence of totalitarianism in digital worlds, and establish possible remedies to this emergence, which will be the focus of chapter five.

What would action *without distinction* look like? The relevant distinction to consider here is narrowly defined as unique distinction, which pertains to the ability of humans to communicate themselves as being a people with unique wants and needs. This means actors without distinction can still differ from one another in their otherness, simply by being different actors. It also means that they can have unique sets of wants and needs. A lack of unique distinction, however, would mean that actors would not be able to express their wants and needs as being uniquely their own and constituting their personhood. This would make actors essentially replaceable, because if an actor is only defined by a set of properties without the ability to express that these properties are possessed by a person, then every actor sharing those same properties would be deemed equivalent.

Action *without equality* would not mean that the actors have different social or economic standings. Rather, it would mean that they do not understand each other as sharing a space together as people. Instead of having a common ground of equality that enables political action, such actors would understand each other as being fundamentally different from one another. From the perspective of one actor, others might still inhabit the same space, but they do not inhabit it together as people, and they would lack a sense of shared humanity.

What would it mean for action to be *predictable*? As we saw earlier, this would make the outcome of action a product that can be worked towards. Furthermore, this would eliminate the freedom of

action because every act would be bound by its results. In turn, this would mean that in an environment where every action is predictable people would lack the ability to act freely. It might be difficult to imagine a situation where every act can be predicted. After all, humans are highly unpredictable and their ability to act in original ways can look like miracle working when “seen from the viewpoint of the automatic processes which seem to determine the course of the world” (Arendt, 1958/2018, p. 246). However, instead of thinking about the question in absolutes, it might be helpful to recognize predictability as a sliding scale. With enough data, human action can be predicted to a certain extent. The degree to which this is possible might be the degree to which action loses its freedom.

In a similar fashion, action loses its freedom when it becomes *reversible*. To the extent that the outcome action can be reversed, it is also possible to manipulate this action. Importantly, the consequences of political action become part of the shared space of appearance. When someone reveals themselves through political action, this action becomes part of the story of who someone is. Of course, consequent actions can change public opinion about a person, but political action can never undo what has been done; the threads of the story cannot be unwoven. Because political action and personhood are so closely related in Arendt’s conceptual framework, the ability to reverse the outcome of action would mean that it is possible to manipulate someone’s personhood. Again, it might be difficult to envision a scenario where the outcome of political action could be totally reversed. However, the mediation of political action through digital technology and the embedding of the consequences of action in digital artifacts might expand the possibilities for such manipulation.

In environments where political action is inhibited, the threat of totalitarianism increases. Such an environment would lack the presence of distinct and equal actors, and action would be predictable and reversible. Arendt introduces the concept of loneliness to illustrate the way in which such an environment elicits totalitarianism. She contrasts loneliness with solitude (Arendt, 1951/2017, p. 625). People are in solitude when they are not in close proximity to others. Loneliness, however, emerges when the space between people is destroyed, regardless of physical proximity. This space that is destroyed refers to the “space of appearance” that enables political action and allows humans to live together as people. Since the space of appearance is dynamic rather than static, this destruction occurs whenever the conditions for political action are absent. Therefore, lonely people are undermined in their ability to engage in political action, which could enable the emergence of totalitarian politics.

When people share a space without being able to act politically, they risk being alone together (Turkle, 2011), and according to Arendt, the threat of totalitarian domination is most pressing when “loneliness [...] has become an everyday experience” (Arendt, 1951/2017, p. 627). This further increases the urgency of investigating the phenomenon of loneliness, because while the potential space for political action has a fleeting nature, the consequences of mass loneliness might be more enduring. The task of chapter five, then, becomes to identify the ways in which the conditions for political action might be at risk in digital worlds. This will allow us to evaluate the risk of loneliness in such worlds, which could create fertile ground for the emergence of totalitarianism.

[6] Conclusion: Conceptualizing political action

So far, this chapter has extensively explored the concept of political action as it is developed by Arendt. This discussion has brought to light the intimate connection between political action and personhood, as well as the dynamic nature of the space in which political action is possible. In addition, this chapter examined the conditions of freedom and plurality, which introduced four requirements for political action: that of distinction, equality, unpredictability, and irreversibility. Now that we have acquired a nuanced understanding of the concept of digital worlds and a conceptual framework of political action, the objective of the following chapters becomes to apply this framework to the context of digital worlds. Through this investigation, we will achieve a better understanding of the question if political action is possible in digital worlds, as well as to evaluate the potential threat of totalitarianism in digital worlds. The next two chapters will explore the theoretical possibility of political action in digital worlds while addressing these concerns.

Chapter 3: Plurality in digital worlds

The previous chapter introduced plurality as one of the conditions of political action. This chapter will take the condition of plurality and analyze the theoretical possibility of its emergence within the context of digital worlds. In order to analyze this theoretical possibility, the two sections of this chapter will focus on the two requirements for plurality: *distinction* and *equality*. Each section will consider the theoretical possibility of meeting these requirements in digital worlds respectively.

Section one will discuss two different perspectives on the requirement of digital *distinction* that are provided by the digital realist and the phenomenological approach. Then, this section will analyze some features that digital worlds must include in order to meet the requirement of distinction. Section two will consider the requirement of political *equality* and relate it to Dreyfus' concept of shared moods that was introduced in chapter one. This will establish the concept of *political moods* as a requirement for political action in digital worlds. The final part of section two will consider the theoretical possibility of creating these political moods in digital worlds. By placing the requirements of distinction and equality in the context of a digital world, this discussion will further our understanding of the condition of plurality in such worlds, and thus for the possibility of political action.

[1] Distinction in digital worlds

The preceding chapter introduced distinction as an essential requirement for supporting plurality, which is one of the conditions of political action. So far, our understanding of distinction has only considered the regular world. This section will focus on the possibility of having distinction in the realm of digital worlds. Therefore, we must theoretically bridge the regular and digital environment. This section will first consider what the dimensions of simulation, virtuality, and digitality can tell us about the difference between these environments regarding distinction. Subsequently, this section will examine both the digital realist and the phenomenological perspective on digital distinction. This will result in a set of requirements for distinction in digital worlds. The final part of this section will consider the theoretical feasibility of these requirements.

Taking distinction to digital worlds

The definition of distinction as a requirement for political action that was introduced in the previous chapter resulted in two aspects of distinction. First, for someone to be distinct, they need to be able to occupy a unique position in the world. This position includes a specific set of wants and needs that

a person possesses. Second, because distinction was narrowly defined as *unique distinction*, in order to be distinct, a person must not only possess a distinct position in the world, in the sense that they have a set of wants and needs that is unique to them, but also be able to communicate their wants and needs as belonging to a unique individual. Both of these aspects must be attainable in a digital world to enable distinction.

The regular world affords the possibility of distinction. This follows from the way Arendt conceptualized distinction, namely as a part of the human condition in the regular world (Arendt, 1958/2018, pp. 175–176). That is not to say that distinction is an innate trait that is automatically possessed by every human being in the regular world. Rather, it represents the theoretical possibility for every human being to achieve distinction in the regular world. This means that both aspects of distinction are theoretically attainable in the regular world. Consequently, the main question of this section is: to what extent are these aspects of distinction also attainable in a digital world?

When moving conceptually from the regular world to a digital world, we can first consider the dimension of *simulation*. To what extent can the features of the regular world be simulated in a digital world to allow for distinction? This raises the further question whether distinction itself is something that needs to be simulated, or that distinction will emerge naturally once a simulation of the regular world is complete enough. In addition, one could distinguish between the two aspects of distinction regarding their need to be simulated. It could be the case that having a unique position in the digital world is something that needs to be simulated by assigning a set of digital objects to belong to a specific inhabitant, by giving them a digital avatar for example, while the ability to communicate themselves as being people, rather than just a set of digital objects, follows automatically from having human beings as the inhabitants of digital worlds.

When we consider the second dimension of difference, that of *virtuality*, however, it becomes evident that the possibility of distinction might not be automatically present in digital worlds. Otherness might be retained in the actualization of digital worlds, because fundamentally the construction of such a world is based on binary states of physical transistors. This means that even when two parts of code that describe two different people are identical, these two parts of code are actualized by two physically different building blocks, i.e. different transistors. Distinction, however, appears to be something that occurs in the domain of the virtual. We defined virtuality as a dimension of meaning and interaction that is mediated by some medium. Distinction, in turn, was defined as a specific set of wants and needs, but unique distinction, which is what we are interested

in as a requirement for plurality, requires that someone is able to communicate those wants and needs as uniquely belonging to a person. Although having a specific set of wants and needs could be realized without meaning and communication, these virtual concepts are required for the ability to express oneself to others as a distinct person. That means that the virtual dimension of digital worlds becomes an important consideration for distinction. Theoretically, the virtual dimension of digital worlds could facilitate meaningful interaction without enabling unique distinction. In such a scenario humans in digital worlds would be able to talk meaningfully about all kinds of things, but might not appear to each other as distinct people.

Because the virtual dimension is mediated, it becomes important for understanding the possibility of distinction to take this mediation itself into account, which in our case is the *digital* foundation of digital worlds. This digital foundation might introduce new ways for distinction to emerge, or it might inhibit or eliminate the possibility for distinction. A feature of digital objects that is particularly relevant to the question of distinction is that of *perfect multiplicability* (Kim, 2001, pp. 99–100). This property of digital objects makes it possible to create a perfect copy of any digital object at any time. In addition to determining the way digital objects operate, perfect multiplicability might also affect the ability of an actor to occupy a distinct position in a digital world, since an actor's wants and needs might be shaped by the environment one finds themselves in. When every aspect of that environment is perfectly multiplicable, as is the case in a complete digital world, then it might be impossible to occupy a unique position, because every position would be replicable. Furthermore, even when occupying a unique position is possible, perfect multiplicability might still inhibit the possibility of communicating this position to other actors as belonging to a person, because their uniqueness, which is an essential element of personhood, would be lost.

Two perspectives on digital distinction

The two philosophical perspectives that were introduced in chapter one might result in contrasting analyses of distinction in digital worlds. On the one hand, digital realism might emphasize the realness of digital environments (Chalmers, 2022, pp. 198–202). This could develop into an argument for the possibility of digital distinction that makes use of the perfect simulation thought experiment. On the other hand, the phenomenological approach might emphasize the need for people to show up as being distinct. This means that even though otherness and plain distinction could be simulated, unique distinction could still be obscured by pitfalls in the virtual dimension, such as the device paradigm. I will now consider these two perspectives.

According to digital realism, digital worlds are as real as the regular world. According to our definitions, that boils down to saying that every actual aspect of the regular world can be simulated. In the context of distinction that means that when distinction is not possible in digital worlds, this is not because of the digital actualization of such worlds. But would that mean that distinction is automatically present in digital worlds?

The ability to communicate themselves as uniquely distinct is something every human possesses, but that does not mean that everyone appears as distinct, even in the regular world. Unfortunately, in fact, it is quite common for people to lose their unique distinctness in the regular world. In these cases, people become mere statistics, and lose their personhood. That this can happen on a large scale in political movements was exactly what urged Arendt to develop the concept of distinction in the first place. A digital realist might say that this could be equally the case in a digital world. Pointing out that someone could lack distinction in a digital world, however, would not entail that distinction is theoretically impossible in digital worlds. It might still be the case that distinction is not possible in digital worlds, but it would be because of the digital nature of such worlds (Chalmers, 2022, pp. 303–304).

Instead, a digital realist could employ the perfect simulation thought experiment to argue for the theoretical possibility of digital distinction. In the case of a perfect simulation, the laws of physics that are present in the regular world, to the extent that they are known, would be simulated in a digital world. In such a perfect simulation, there would not be any noticeable difference for inhabitants of the digital world. Everything in the digital world would behave in the same way that it would in the regular world. That means in order to communicate themselves as unique people in a perfect situation, all they would have to do is behave just like they would in the regular world when they communicate themselves as unique people. The requirements that digital realism poses are that the inhabitants of digital worlds are people that possess the ability to communicate themselves as uniquely distinct and that there exists no noticeable difference between their actions in the regular world and their equivalents in the perfectly simulated digital world.

Following the phenomenological approach, however, one could argue against this that the abilities of the inhabitants of digital worlds and nature of the digital world itself cannot be considered as separate requirements because they cannot be understood separately (Heidegger, 1954/1977, p. 12). In addition, because phenomenology emphasizes the interplay between the apparent world of

usable devices and the concealed background world, it sheds a different light on the requirement of having no noticeable difference between the regular and digital world (Borgmann, 1987, p. 44).

For distinction, this means that the digital foundation of a digital world will have an impact on the way that people are able to distinguish themselves and how they distinguish others. Just like possessing a phone causes someone to show up to others as being available to call, as we saw in the chapter one, being in a digital world causes someone to show up as being mediated through a digital medium. Consequently, both the limitations of this digital medium as well as the possibilities that it affords become part of someone in a sense that is relevant to their distinction. When someone communicates their position in the world, it is relevant to the audience, for instance, whether this position is perfectly multiplicable. This is the case, even when this perfect multiplicability is not always apparent because it is part of the background world from which objects and people phenomenologically show up and therefore determines the way in which they show up.

If digital worlds impose a way of showing up that makes distinction theoretically impossible, then this voids the digital realist requirement of having no noticeable differences. This could be the case if the concern of the device paradigm applies to digital worlds. In that case, people in digital worlds could be defined by their device-like interchangeability. Like the knob of a heater conceals the complexity of the processes that make the process of heating available, the digital veil between people in digital worlds could hide the complexity of their unique distinction. Per definition, a distinct individual is not interchangeable. If the phenomenological appearance of people in digital worlds is one that theoretically allows for two exact copies of a person to exist, the possibility of distinction is in peril.

Distinction in the Matrix

So far, I have discussed some questions related to the theoretical possibility of distinction in digital worlds. Distinction was conceptualized as a unique position in the world that captures someone's personhood. To be distinct, someone must not only occupy this position, but also be able to communicate it as such. Taking this possibility from the regular to the digital world introduces lines of inquiry related to the dimensions of difference. Which elements of distinction must be *simulated* and to what extent? Since the possibility of distinction is determined by the *virtual* dimension of the digital world, it becomes relevant to ask in what way *digital* properties like perfect multiplicability affect it theoretically.

Depending on the philosophical perspective one takes, these questions can be interpreted differently. Digital realism urges us to consider the similarities between the regular and the digital world, positing that a digital world without noticeable differences from the regular world will result in the same theoretical possibilities for distinction for both worlds. Phenomenology, on the other hand, invites us to consider the way the unapparent parts of digital worlds still affect the possibility of their inhabitants to distinguish themselves.

In the popular movie *The Matrix* (1999), the audience is presented with a digital world that is a perfect replica of earth. This situation resembles the perfect simulation that a digital realist describes. The protagonist of the movie, Neo, is plugged into the digital world the moment he is born in the regular world. To him, there exists no noticeable difference between the two. Furthermore, he is a human being that would have the ability to be distinct outside of the digital world, in the regular world. Next to humans like Neo, the movie introduces another kind of inhabitant of the digital world: agents. These are computer programmed entities that cannot exist outside of the digital world. They are able to manipulate the digital world to their desire, an ability that Neo ends up acquiring after he awakens to the truth of his situation.

Now, we can ask the question, are either humans like Neo or agents in *The Matrix* able to be distinct, meaning first, are they able to occupy a unique position, and second, are they able to communicate this position as belonging to a person? First, since agents are purely programmed, they are perfectly replicable. This would appear to exempt them from occupying a unique position in the digital world, since a perfect copy of any agent could exist. However, is this the kind of uniqueness that we require for distinction? After all, even when perfect copies of agents exist, they still occupy their own position in the digital world, they are the same fragment of code that functions in different places of the program so to say. Posing the question in these terms forces us to make a decision. Do we link distinctness to the code or its function? If we say that uniqueness of function is required for distinction, then agents cannot be distinct, because their function can be perfectly replicated. If we say uniqueness of the code is required for distinction, then agents could be distinct, but so could all other digital objects.

Conversely, humans do not have a function in the digital world that can be replicated, because, according to Arendt, they are not the product of work, and they are capable of dynamically determining their function through political action (Arendt, 1958/2018, p. 7). This means that they

necessarily occupy a unique position in the digital world. There might still be a theoretical hurdle for them to acquire distinction in a digital world, however, because of our second question. Are they able to communicate this position as uniquely belonging to a person? Arendt's definition of personhood, that we are adhering to here, fundamentally refuses that a person could possess a predetermined function. This means that in order to communicate personhood, others must be able to tell the difference between humans and functional code that is part of the digital world, like agents.⁸ The problem that the device paradigm introduces, is that exactly this ability might get lost in digital worlds. According to Borgmann, devices act like a veil in front of the complexity of the world (Borgmann, 1987, p. 44). If, in a digital world, this veil is applied to humans, then their distinction might be funneled through a device filter. As such, even though a distinct human might be the originator of actions in a digital world, this information is fundamentally inhibited from being shared with others, because one cannot differentiate people from non-people through the devices that are supposed to communicate it.

In conclusion, while humans occupy unique positions in the digital world due to the absence of a function that can be replicated, the device paradigm might impede the effective communication of their distinction in a digital world. This would mean that a simulation of the regular world can only guarantee distinction when it allows people to discern distinct others from functional parts of the digital world.

[2] Equality in digital worlds

A second essential requirement for political action that was introduced in the previous chapter is that of equality. In contrast to distinction, that requires political actors to be unique and diverse, equality requires them to have something in common. What actors should have in common is a shared sense of being in a political space. The question that this section will discuss is if this sense of being in a shared political space can theoretically emerge in a digital world. And if this is possible, what is required to facilitate this equality in digital environments?

First, we will take a look at the concept of shared moods as defined by Dreyfus, which was introduced in chapter one. This concept can be used to develop a way of understanding connection between actors, offering an alternative perspective to the digital realist approach. In the following

⁸ One could make the case that agents are, in fact, people, but this gets into a debate about the extent to which humans are functional/computable, which goes beyond the scope of this thesis.

discussion I will discuss this difference and highlight the significance. Next, I will argue that the concept of shared moods can be applied to Arendt's concept of equality, thereby presenting political moods as an essential requirement for political action. Finally, this section will consider if political moods can theoretically emerge in digital worlds. This will help us to assess whether meaningful political action can take place in digital worlds.

Connection between actors

Equality, as it is developed by Arendt, requires actors to be similar in some regard, while remaining fundamentally unique (Arendt, 1958/2018, p. 175). This similarity can be understood as a shared sense of being together in a political environment. What is meant with this 'shared sense' can differ based on the philosophical approach one takes. Similar to the arguments for distinction, a digital realist will likely posit that equality is a phenomenon that emerges in the regular world, and that it therefore translates into digital worlds once the elements of the regular world that facilitate this shared sense of being are simulated. In order to acquire a better understanding of equality in digital worlds, one could ask which specific elements of the regular world need to be simulated. Likely candidates are epistemological or cognitive requirements, suggesting that actors must either know of or be aware of each other's presence in the digital realm.

Although these requirements might perform well in describing a sense of being together, they may fall short of capturing a deeper sense of connection between actors. A phenomenologist would likely argue that describing this deeper sense of connection is only possible by considering similarities in the phenomenological experiences of actors. By taking this approach, however, we have switched from considering elements of the external world, to aspects of the actors themselves, although a phenomenologist will likely argue that these things are inseparable. This makes it difficult, however, to agree upon a list of elements that need to be simulated in a digital world in order to facilitate equality. Instead, a phenomenologist might look at the differences and similarities between the regular and digital world within their virtual dimensions. The question for the phenomenologist, then, becomes if the funneling of information about others through digital technology in some way affects the way that one is able to connect phenomenologically with them.

Political moods

Dreyfus takes up this problem of connection in digital worlds from the phenomenological angle. He starts from the observation that in digital environments there is either a different sense of being

among people, or this sense is lacking altogether. His idea is that this sense does not emerge from the sum of sensory information inputs, which would imply that we could enable equality in digital worlds if the technology is advanced enough. Instead, Dreyfus introduces the concept of "shared moods" (Dreyfus, 2001, pp. 108–110). A shared mood emerges wherever people have a similar way of phenomenologically interpreting the world. Clear examples of shared moods can be found during football games. Here, a crowd of thousands can experience a shared sense of support for their team. This does not only mean that they experience a similar emotion when their team scores, like elation or excitement, but that their experiences are also connected in a meaningful way; the goal is a shared victory that plays a similar role in each of the supporters' virtual world of meaning.

This concept of shared moods can be applied to our question of understanding equality as a requirement for political action in digital worlds. Shared moods provide a framework for understanding the shared sense of being in a political space that creates equality. This shared sense, then, transforms into a shared mood. It would be a specific kind of shared political mood, in which individuals perceive the world as political. Shared moods function like a phenomenological pair of glasses, presenting the world and its possibilities to the subject in a specific way. Consequently, the world would have a different meaning for people who are in a shared political mood, compared to those who do not share this mood.

What would have to be part of the shared sense of meaning of people in a political mood is the idea that each of them is a distinct individual that is part of the same environment as a political actor. This creates the idea of a political community. This community is not bound by a common objective, as is the case in the football match, but by a common commitment to opening and keeping open a virtual space of appearance. This space, then, allows those who are in a shared political mood to act politically.

Creating political moods in digital worlds

This brings us to the question if shared political moods could emerge in digital worlds. Would the creation of such a mood necessarily follow from simulating elements of the regular world in a digital world? A digital world, or a part of this world, would have to be meaningful enough to a group of people to open a space of appearance and continue performing political action within it together.

This, however, might be where a problem with digital worlds might arise when we consider one of Dreyfus' other concepts, that of "situated risk" (Dreyfus, 1999, 2001, p. 54,57,87,88). The idea is that

in order for something to be meaningful in a way that enables the emergence of a shared mood, there needs to be a certain type of risk involved. This risk is created by being in a phenomenological situation. What makes taking up such a position risky, is that as long as you have the position, you face the consequences of negative actions within that meaningful frame of reference. A football fan and supporter of a specific team faces the risk of their team losing, but also experiences a meaningful reward when the team wins.

So, what happens when a crowd of supporters support their team in a digital world? Can the same sense of shared excitement be felt in a digital stadium?⁹ Dreyfus argues that the required level of existential commitment to a risky position in digital worlds might be harder to attain because it is easier to switch your position (Dreyfus, 1999, pp. 108–109). By being digitally present in a stadium, instead of in person, one does not expose themselves to the same level of risk because it is easier to abandon the situation. This might be because there is a fundamental risk of harm or embarrassment associated with putting one's body in the fray (Dreyfus, 2001, pp. 57–59); someone might think twice about saying something about the opposing team when their supporters are standing in front of them in the regular world. By comparison, someone could easily ridicule an opponent in a digital space and leave to avoid the consequences, which is a common occurrence in online gaming lobbies.

But there is also a sense of non-commitment in digital worlds that goes beyond this type of anonymity. This might be caused by the abundance of options that are available in such worlds and the ease with which one switches between them. When one team disappoints, there might be so many other teams to choose from that the type of existential commitment required for situated risk is undermined. Additionally, in the regular world many people support their local football team exactly because it is local, whether it be from their town or their country. There exists a natural tendency to existentially commit to things that are local in this way, to one's physical neighbors. In a digital world, this effect might be entirely absent.

When we return to equality, we can see some similarities. If we understand equality in digital worlds as a political mood that is shared by its inhabitants, then we can ask the question if the sheer abundance of possibilities in a digital world affects the emergence of such a mood. We might envision a myriad of digital political communities that one could join or leave at a moment's notice.

⁹ This situation is very similar to modern eSports events. Although in big eSports tournaments, it is common to have competitors play the game while the audience is physically present, suggesting that this presence has some unreplaceable value.

And because there might not be any cities or countries that are local in the same way as they are in the regular world, it might be difficult to existentially commit to a single one. In order to meet the requirement of equality in digital worlds, it might be necessary to simulate certain physical boundaries that prevent people from abandoning a situation, and artificially create a sense of locality. This could allow people to share a political mood in a digital world, which could afford a shared sense of equality.

[3] Conclusion: The theoretical obstacles for plurality in digital worlds

In conclusion, the examination of the requirements of distinction and equality in the context of digital worlds has revealed some prerequisites for plurality in such worlds. Individuals must be able to occupy a unique position in the digital world and communicate themselves as *distinct* beings. This communication might be hindered by the functional nature of digital worlds that prevents making distinctions between non-functional human beings and functional parts of the digital environment. In order to be *equal* in a digital world, individuals should share it as a political environment. This could be accomplished by being in a political mood together. This requires a level of commitment, however, that is in conflict with the many opportunities that a digital world presents. This means that both the device paradigm and the impossibility of situated risk might form theoretical obstacles for plurality in digital worlds. Before delving into the practical limitations of plurality in digital worlds, I will discuss the requirements for the second condition of political action, that of freedom, in the context of digital worlds.

Chapter 4: Freedom in digital worlds

Now that we have considered the condition of plurality in a context of digital worlds, it is time to turn our attention to the second condition of political action, that of freedom. Chapter two established a conceptual difference between plurality and freedom. While the former can be understood as a necessary condition of the actors themselves for political action, the latter represents the mode of action that is required for political action. Specifically, political action must be performed under the principle of freedom. Arendt's framework suggests that such action must be both unpredictable and irreversible.

Taking these concepts to the context of digital worlds allows us to raise questions about the theoretical possibilities of digital worlds regarding freedom. Our goal is to employ Arendt's framework and interpretation of freedom to better understand the possible theoretical challenges for political action in digital worlds. First, we can ask if action can be unpredictable in digital worlds. To understand what unpredictability in digital worlds would look like and how it could affect political action, the first section of this chapter will consider the concept of unpredictability in the regular world. Then, I will discuss the role of algorithms in relation to this question. The second section of this chapter will focus on the theoretical possibility of irreversibility in digital worlds. This discussion will start with an exploration of the ways that digital objects can be erased in digital environments. Then, I will consider how the possibility of erasing digital objects might affect action in digital worlds. The discussions in this chapter will give us a better theoretical understanding of Arendt's concept of freedom in digital worlds, and give us a foundation for an investigation of totalitarianism as the absence of freedom in the next chapter.

[1] Unpredictability in digital worlds

The concept of unpredictability in the regular world can be understood in at least two different ways, from an individual or a universal perspective. It can either mean that someone who performs an action cannot predict the outcome of that action at the moment of performing it, or that the action itself is fundamentally unpredictable, meaning that neither the actor nor anyone else with any amount of knowledge or advanced technology could theoretically be able to predict the outcome of the action. When someone rolls a die, it would typically make sense to claim that the outcome of that action is unpredictable; any side of the die has the equal chance of one in six to end up facing upwards. However, perhaps someone has weighted the die beforehand, unbeknownst to the actor throwing it. In this case, the action would be unpredictable in the first sense, while being predictable

in the second sense. But there is no need for deceptive intentions to create this scenario. With enough visual data, a computer might calculate and predict which face will end up facing upwards while the die is falling. This means that even though the dice thrower performs an action that is unpredictable to them, the computer introduces predictability in the second sense.

This kind of predicting outcomes in the regular world on the basis of data is very common. In order to make such calculations, one has to make a model of the regular world that is calculable. Of course, such models never predict the regular world with absolute certainty. Physics-models try to reach a high level of predictability, and arguably come the closest to certainty, but even they break down when pushed to extreme precision. Models become even more unpredictable when they try to account for human action, like behavioral or economic models. This means that an argument can be made that human action is, even in the second sense, fundamentally unpredictable in the regular world.

However, the question that will be discussed here is: what happens to this unpredictability in models that account for human action, when human action is performed exclusively in a calculable environment? This is in line with the limitations for our investigation that we set in chapter one. We are interested, here, in social and digital worlds. That is to say, worlds that are inhabited by multiple people, where their interactions are mediated by digital technology. The question that this section tries to clarify is how the digital constitution of this interaction affects the concept of unpredictability in that interaction.

Arendt considers unpredictability to be a necessary requirement for free deeds, which are actions that are performed in the public space of appearance and whose effects and consequences are beyond the control of the actor (Arendt, 1958/2018, p. 244). It is not immediately clear in her argument if this unpredictability should be considered in the first sense, from the perspective of the actor, or in the second more general sense.¹⁰ What is more important for Arendt is that unpredictability allows her to differentiate action, which lacks a presupposed purpose, from work, which does possess one. One could imagine a person living in a digital world whose actions are totally predictable by the engineers of the world, but who is ignorant about this, much like the dice roller who rolls a weighted dice and one could engage in a philosophical debate about the

¹⁰ Arendt does differentiate between the “basic unreliability of men” and “the impossibility of foretelling the consequences of an act within a community” (Arendt, 1958/2018, p. 244), but both of these branches of unpredictability could be understood from the individual or the universal perspective.

implications of such a scenario, questioning if unpredictability from the actor's perspective does indeed lead to free action.

Here, I will not engage in this debate, however, because our focus lies on the question if digital worlds can theoretically provide a conducive environment for political action according to Arendt's framework. This means that we will adopt her notion of freedom, which requires unpredictability. If there exists a digital world in which this unpredictability can occur, then digital worlds can theoretically meet this requirement. Therefore, I will consider a digital world with inhabitants that understand the predictive power of their environment. This means taking the universal perspective on unpredictability. When unpredictable action turns out to be impossible for these people, it will not be because of some lack of knowledge on the part of the individual, but on account of the inherent conditions of the digital world. Hereby, I attempt to isolate the effect that digital technology, through which all communication and interaction is filtered in a digital world, has on unpredictability. Political action that is performed in the regular world might be unpredictable because models cannot fully capture this world. But how unpredictable are actions that are fully performed in a digital world, which is a model in and of itself?

Digital dice and deeds

When we move the example of a person throwing a die from the regular world to a digital world, we can see that there might exist fundamental differences between these scenarios in the way unpredictability functions. For instance, consider how a dice throw in a digital world would operate. Say someone wants a random result out of six possible outcomes. However sophisticated the immersive experience of such a dice roll might be, there could be a simulation of a falling three-dimensional die in the digital world, effectively, it comes down to asking a computer to generate a random number between one and six, which cannot be inherently random. Because a digital world is computable, this entails that there are no random elements inherent to the world itself. "The values of random processes cannot, by definition, be the output of a function. Random processes can be *exploited* by a computation, [...] but no computational technique can amount to a mere sequence of random choices" (Piccinini & Maley, 2021). The definition of computability prevents randomness from appearing inside the code of a digital world itself. In order to create randomness, programs

usually exploit external sources of randomness, such as radioactive decay or the unpredictable movement of subatomic particles.¹¹

However, when a computer relies on an external unpredictable process in order to generate randomness and unpredictability, this puts the possibility of creating randomness at tension with the philosophical position of digital realism. According to this view, there exists no element of the regular world that cannot be simulated in a digital environment. But if we want to generate a random number using some external source X, then in order for this number to be truly random, X should not be calculable. Therefore, X would also be a part of the regular world that cannot be simulated in a digital world, which would contradict digital realism. The conclusion of this is that digital realism excludes the possibility of computer generated unpredictability entirely.

In the predictable digital worlds of digital realists, then, it might not be possible to create an unpredictable situation by throwing a die. But perhaps humans might still be able to act unpredictable. A human being might, after all, still make unpredictable choices when the options are determined. Even a game of chess, where the options a player has are strictly defined, can play out in wildly unpredictable ways when played by humans. But does that allow for enough unpredictability, and the right kind of unpredictability to perform free deeds?

One of the concerns that followed from Borgmann's device paradigm was that by reducing complexity through the introduction of devices, we are commodifying the world (Borgmann, 1987, Chapter 9). The world we experience is, then, filled with things that are always available, while we lose track of the complexity that goes into making those things available. We can reframe this worry to the problem of unpredictability. If we do not adhere to the digital realist position, we might argue that a person in the regular world has an incalculable array of options for action. And even if we say that the number of different actions in the regular world is calculable, this number might be so vast that it is not possible to simulate them digitally. In both cases, moving to a digital world reduces options that someone has for performing action.

By introducing the idea that there exists a limited number of options for action, we increase predictability. This could be because the metaphysical reality of a digital world is inherently more

¹¹ This is analogous to the difference between pseudo-randomness and true-randomness. For the former, programs rely on a predictable algorithm that appears unpredictable to a human user, while for the latter, external physical processes are used as the basis for randomness.

predictable than the regular world. Possibly, the very idea of there being a limited number of options itself could also cause predictability, by changing the psychology of action. If one believes that they can only perform a certain set of actions, they might be less likely to perform actions that are outside that set, even if they are theoretically possible. If, through any of these effects, predictability is increased in digital worlds compared to the regular world, this fundamentally undermines the potential to perform free deeds.

[2] Irreversibility in digital worlds

The second requirement for free action is that of irreversibility. I will first consider Arendt's conception of irreversibility in both the regular world and the digital world. Subsequently I will explore different ways in which irreversibility might be compromised in digital worlds. The next section will consider how reversibility might lead to an absence of situated risk in digital worlds and how this might be fundamentally at odds with the condition of freedom.

Central to irreversibility is that the consequences of political action cannot be undone. Once someone has performed an action, it will leave a lasting impact on the world and, more importantly, on others who inhabit that world. What has happened cannot be undone. This is a natural to the way many things occur in the regular world. When a physical object gets destroyed, it might be restored. But it will never be exactly like it was before.

Take the example of the regime of a country that takes down the statue of a political figure with whom they do not agree. In the aftermath of that regime's downfall, the citizens of the country might want to reinstate the statue. They might succeed in restoring the original or creating a very precise replica, but realistically the damage will still be part of the statue in some way. On the other hand, the original removal of the statue by the regime was likely to have left traces. Furthermore, the symbolic effect of the act cannot be undone. After all, the taking down of the statue was not some attack on the physical structure, the actual dimension, of the statue, but rather on the idea it represented, its symbolic – or virtual – dimension. The history of the act of taking down the statue will irreversibly be part of the public space in which it was performed.

Now we can pose the question if the same is true for digital worlds. There are, of course, new possibilities in digital worlds. Digital objects behave differently from most physical objects, even though a digital realist would argue that they are fundamentally the same. Unlike most physical

objects, digital objects can be deleted, restored, and perfectly duplicated. Admittedly, with enough effort, parts of the regular world can be deleted, restored, or duplicated as well. Nevertheless, in a digital world, these changes can happen more abruptly and absolutely. This raises the question whether these differences affect the irreversibility of action in digital worlds.

Acting with a backspace button

There is an interesting paradox to the contingency of digital objects in digital worlds that can also be experienced on the modern internet: while everything on the internet is fleeting on the one hand because it can be deleted with the click of a button, nothing seems to ever really disappear. Because every part of a digital world is based on a piece of digitized code that describes it, the way to destroy parts of a digital world is by erasing that code from the memory system it is stored on. When that is done thoroughly, the destruction is absolute and it will be like that part of the world never existed. When only one copy of the code remains, however, on a small memory device like an USB stick for example, the destroyed part of the world can be restored to its former state like nothing ever happened.

This paradox of digital worlds affects the concepts of creation and destruction. Every action in such worlds, whether it involves destruction of some part of the world or the creation of something new, is always done under the possibility of it being undone, and this undoing can be swift and absolute. Furthermore, whereas different parts of the regular world are more difficult to create or destroy, everything in a digital world has the same foundation and can be created and destroyed in the same way. A large building in the regular world, for example, is difficult to construct, but will also be difficult to tear down, creating something durable. Spoken words are also difficult to silence in the regular world, even though strict governmental policies sometimes achieve this. In a digital world that simulates someone voicing their opinion by speaking, on the other hand, sound waves do not travel through physical matter. Rather, simulated sound travels through simulated matter. And because the process is constructed in such a manner, it can also be undone this way.

Does this reversibility of things in the digital world affect the possibility of political action? Intuitively, the process that is affected is that of work. This is the mode of action that relates to durable external objects, like buildings and statues. Although it is a worthwhile question to ask how work will be affected in digital worlds, I will focus here solely on political action. One of the ways that the reversibility of digital worlds might prohibit political action is by causing an absence of situated risk.

Situated risk was introduced in the first chapter as a fundamental risk that endows action with meaning. In the third chapter, it reared its head when discussing shared moods in digital worlds. Situated risk arises where people choose to play a role in the world and accept full responsibility for that role, which means that they are not at liberty to switch to any other role. This comes both with the risk of having to accept the consequences of the role, which might be especially hard when things go wrong, and the risk of not being able to lead any other life. One of the compelling aspects of digital worlds lies exactly in the absence of this kind of risk. Dreyfus uses avatars and the ability to change them as an example of how someone could avoid risk in digital worlds (Dreyfus, 2001, p. 98,115). This could be refuted, however, by pointing out that there are usually whole social systems behind digital communities that will not always allow casual risk avoidance.

Instead, I argue that, if anything, the absence of situated risk in digital worlds is caused by the radical reversibility of such worlds. Because all action in a digital world is done under the constant threat of the undoing of that action, it might lose its risk, but it also loses the promise of having a lasting impact and being meaningful. Of course, no action in the regular world is ever guaranteed to be meaningful, but the promise of the regular world is that it at least has the potential to be. In digital worlds the uncertainty of this promise might be pushed to such an extreme that situated risk becomes impossible.

So, although the regular world is also to some extent reversible, digital worlds might exhibit a degree of reversibility that diminishes the meaning of political action. If that is indeed the case, action in digital worlds would not be irreversible in the way that Arendt deems essential for freedom of action. Therefore, political action would be impossible in digital worlds.

[3] Conclusion: Theoretical obstacles to freedom in in digital worlds

This chapter discussed the ways that digital worlds might be predictable and reversible. These properties would be at odds with the requirements that Arendt gives for freedom, the second condition of political action. In the previous chapter, we saw that digital worlds might prohibit people from communicating their unique distinction and that without shared moods it might be impossible to be equal in such worlds. Both of these deficiencies might undermine the first condition of political action, that of plurality. None of these discussions ended with definitive conclusions about the possibility of political action in digital worlds. The main point that these chapters tried to convey is

that living in a digital world changes the human condition. This does not necessarily mean that political action will be impossible in this new 'digital' condition, but simply it is not a given. Therefore, the theoretical possibility of political action in digital worlds should remain an area of further philosophical investigation.

In the next and last chapter, the conditions and requirements of the previous will be worked out further, but the focus will shift towards their applicability to realistically conceived digital worlds, as opposed to theoretical ones. In other words, even when political action is theoretically possible in digital worlds, how likely is it that political action can actually be performed in such worlds? Therefore, the next chapter will examine possible challenges and strategies to enable political action within digital worlds.

Chapter 5: Building digital societies

The aim of the previous chapters was to establish a theoretical foundation for a discussion of political action in digital worlds on the basis of Arendt's framework. These chapters focused on the question of the *theoretical* possibility of political action in such worlds. Simply put, the question was: can we theorize a digital world in which political action is possible? In this chapter, the focus will shift to a different set of questions. Assuming that the answer to the previous question is positive, and political action in digital worlds is possible, we will explore whether political action is *enabled* to emerge in digital worlds. When we construct a digital world, what are the challenges that might inhibit political action and how can digital worlds be designed to facilitate such action? Additionally, the aim of this chapter is to develop the theory of political action by Arendt as a workable theory for understanding digital worlds as they emerge in our societies. Where the aim of the previous two chapters was to make this theory of political action workable for a theoretical analysis of digital worlds, the emphasis here will be on an applied analysis of such worlds.

This chapter will attempt to shed light on these questions through the following structure. The first section will consider some examples of complete digital worlds that we find in the present-day to assess how the challenges for political action in realistically conceived digital worlds relate to the challenges that are present in the regular world. The second section discusses the threat of totalitarianism in digital worlds. In *The Origins of Totalitarianism* (1951/2017), Arendt explores the concept of totalitarianism in her analysis of the recent regimes of national-socialism and Stalinism. She defines totalitarianism as a new ideology that sought total control over all aspects of public and private life. I will mainly explore totalitarianism as a political organization of a society that negates distinction and equality while promoting predictability and reversibility. As such, totalitarianism will be understood as the antithesis of the concept of political action that Arendt later develops in *The Human Condition* (1958/2018). The urgent question that this section explores is how digital worlds might play into the hands of totalitarianism. The third and last section of this chapter will offer some suggestions for addressing these challenges in digital worlds. This will hopefully further our understanding of digital worlds and provide direction for constructing worlds that enable political action.

[1] Life in a constructed digital world

So far, we have analyzed theoretical complete digital worlds without considering practical limitations. It is important to note that these worlds do not currently exist. Even if we could create such worlds at some point, transforming whole societies into digital worlds will be a gradual process. This means that, in order to get a thorough understanding of political action in digital worlds, we have to consider realistically conceived digital worlds while acknowledging these constraints. We should also take into account the political and historical contexts in which such digital worlds will likely be created. First we should ask ourselves what such realistic digital worlds look like, and what they might look like in the future. Then, we should consider to what extent the concepts of plurality and freedom, as we have discussed them so far, apply to these digital worlds. In other words, where do realistically conceived digital worlds position themselves between the regular world and theoretical digital worlds?

As a start, let us consider what a digital world in the present-day might look like. There are many examples of *partial* digital worlds, such as online forums and social media platforms where a considerable amount of interaction is facilitated through digital technology. Even if the members of these digital communities might spend a lot of time on the internet, that interaction essentially is just an addition to their life in the regular world. For the sake of clarity, however, I will examine examples of *complete* digital worlds in present-day society. At the very core, a complete digital world requires individuals to experience all facets of life through a digital lens.

Although they are rare, cases of people who live such a lifestyle do exist. One of the most famous examples of people who choose to move their lives entirely to a digital setting are Hikikomori. This term was originally used around the 1970s to describe a phenomenon in Japanese society where people withdrew from social interactions and lived their lives exclusively from the confinement of their home (Kato et al., 2020). Initially, this was believed to be specifically linked with Japan's social culture and work pressure. However, this hermit-like behavior has since been recognized as a worldwide phenomenon that might even have increased during the Covid pandemic. The Hikikomori lifestyle is now recognized in many different countries (Rooksby et al., 2020). Hikikomori are typically defined by a lack of social interaction, but this only relates to regular face-to-face social interaction. Many Hikikomori spend a considerable amount of their time in social communities on the internet, and there might even be a link between internet usage and their radical confinement (Kato et al., 2020). This would suggest that Hikikomori are living in the closest equivalent of a present-day

complete digital world. This means that, even though Hikikomori are still a marginal part of society, they are a useful example to ground our analysis of realistically conceived digital worlds.

How, then, would the life of a Hikikomori be categorized according to the modes of action that Arendt proposes, that of work, labor, and political action? Labor entails those processes that keep someone alive in their closed-off room. It is not difficult to imagine that someone could acquire food while staying at home, most obviously by ordering it online. This would only be possible in an urban setting, however, and that is exactly where the phenomenon of Hikikomori is rooted. This lifestyle might still result in the occasional face-to-face encounter, for instance when someone delivers the food or collects the trash. When these interactions are swift and kept to the bare minimum however, as fits with the general behavior of a Hikikomori, they are essentially superfluous. From home, Hikikomori could do exercises to keep their body healthy. It would even be possible to have a romantic relationship online, and with the help of in vitro fertilization two people could even have children without ever meeting each other in the regular world, although according to the current standards of romantic relationships this would probably be considered quite dystopian. Leaving aside the normative considerations of such a setup, we can see how it would be possible for a present-day Hikikomori to “labor” in an urban setting.

Working from home is also possible, both in the colloquial and the technical sense that Arendt defined. Work is the mode of action that produces external artifacts. These artifacts could, however, just as well be digital artifacts. There are many jobs available that require nothing but digital work, which could be done without ever leaving the house, especially in knowledge economies that are increasingly focused on information industries. In addition, one could create digital art. This has been an increasingly popular way of creating art in general, and it has even become feasible to earn a living through freelance service networks. This means that it is also quite possible for a present-day Hikikomori to “work.”

Performing political action as a Hikikomori, on the other hand, will be less straightforward. If one draws the conclusion that political action is theoretically possible in digital worlds, then Hikikomori could also theoretically perform such action from the isolation of their home. However, from a psychological perspective, it is a concern that Hikikomori have been defined by a “desire to become invisible from society” (Rooksby et al., 2020, p. 399). Of course, this does not necessarily imply that everyone living a complete digital lifestyle desires to withdraw from society. But significantly, it has been a trend among Hikikomori to avoid political participation. Additionally, there might be practical

challenges regarding plurality and freedom. These challenges will be discussed in the following section.

When someone confines themselves to their home and experiences the world through their digital devices, they are always dependent on an outside world of people who are living non-digitally. These people make sure that those in the digital world have access to necessities such as food, supplies, and other resources. They build the physical houses that the inhabitants of digital worlds live in and they produce and maintain the technology that supports the digital world. This does not mean that living and performing political action in a complete digital world is strictly impossible. On the contrary, this situation seems to be in line with Arendt's view of how political action first came into existence in ancient Greek society (Arendt, 1958/2018, p. 194,197-199). In that specific historical setting, it was far from common that someone could perform political action in the first place. This was an activity that was reserved for the most elite members of society and only made possible by an incredible amount of labor and work that was performed by those who did not partake in the political scene. It is well-known that in ancient Greece, the leisure required for political action was often facilitated by the institution of slavery.

A situation where some are not able to partake in politics could be mirrored and exaggerated by digital worlds. If political action would indeed be possible in such worlds, then those who make it possible might not even appear in these digital worlds. While the laborers and workers in ancient Greek society did not enter the space of appearance and thus moved to the background of that society and its history, those who work and labor in modern societies might, likewise, be moved to the background because they are unable to enter the digital world. This might result in a political community in digital worlds that is oblivious to the factors that make their world possible.

And even though the background of a digital world might be invisible from within, it plays an important role in the way such a world will be constructed. Digital worlds will not be created instantly in neutral conditions. They will gradually emerge in political environments that govern their development. This is another aspect that differentiates realistically conceptualized digital worlds from theoretical ones. For example, a digital world might emerge in a society that is already totalitarian. Then, this digital world will most likely be a continuation of that political situation. In any case, digital worlds will, to some extent, reflect the values and social constructs of the society they emerge in.

In conclusion, complete digital worlds are feasible in modern societies. They will be akin to the lifestyle of Hikikomori. This lifestyle currently carries negative connotations, but that might change as digital worlds are normalized. It is an open question if we should want to normalize it, however. This depends on many factors. In this thesis I argue that one of these factors is the possibility of performing political action in digital worlds. Although it is relatively clear that labor and work can be performed from a present-day digital world, this is not the case for political action. In the last chapters, we discussed the theoretical possibility of such action. Here, I want to draw attention to the fact that the background from which a digital world emerges is also an important consideration for this question. Even if political action would be theoretically possible in digital worlds, the political environment that they emerge in and the way they are subsequently constructed, affect the possibility of political action in digital worlds as well. In the next section, I will consider the possible challenges for political action in the construction of digital worlds.

[2] The threat of digital totalitarianism

Political action is possible in the regular world, but it is not a given. The space of appearance, in which political action can take place is a virtual space that must constantly and consciously be maintained. This process can be disrupted or even made impossible in many ways, such as through political systems that inhibit political action. Such political systems can and will also emerge in digital worlds. In this section I want to consider the ways that digital worlds interact with these systems. This is a two-way relationship. The previous section addressed the fact that the political environment in which a digital world is created, influences the development of such a world. On the other hand, the properties of digital worlds might also influence how political systems develop in such worlds. The question that is central to this chapter is, then: do the properties of digital worlds enhance or inhibit the emergence of political action?

Totalitarianism is the political system that poses the largest threat for political action, according to Arendt. Her use of the term totalitarianism has a historical context. It became a commonly used umbrella term in the English-speaking world after the Second World War to define the new threat that emerged in the form of authoritarian and communist governments. The idea was that the catastrophe that had occurred in Germany under the Nazis was not a localized incident, but something caused by a harmful totalitarian political model that could emerge anywhere (Kuklick, 2022, pp. 124–131).

I will not use the term totalitarianism in this way, nor will it acquire the complex meaning that the term had across much of Arendt's writing. Instead, I will restrict myself to the specific meaning it has in relation to political action. In this way, totalitarianism can be understood as the antithesis of such action. Totalitarianism as a political system, then, is defined by its tendency to prevent the majority of people living under it to perform political action. As such, totalitarianism is a political system that renders people indistinguishable and unequal and their actions predictable and reversible. The question under discussion in this section is whether digital worlds play into the hands of these totalitarian tendencies. In the next section I will consider how, on the flipside, digital worlds might encourage political action and what design choices might nudge towards this encouragement.

In order to create a structured discussion of the four totalitarian aspects of indistinguishability, inequality, predictability, and reversibility in relation to digital worlds, I will sketch four vignettes of political systems around the world where the digitalization of society has influenced at least one of these aspects. These vignettes will use the case of the Netherlands to clarify indistinguishability, that of the United States for inequality, that of China for predictability, and that of Russia for reversibility. Here, it is important to note that none of these aspects determine these societies. They can be, and usually are, present across all societies. After illustrating the connection between digitalization and these totalitarian aspects, I will consider how digital worlds might exacerbate them, and therefore inhibit political action.

The Netherlands and indistinguishability

The Netherlands is a bureaucratic country. This is understandable, since it is a very densely populated country that requires a lot of organization to function the way it does. In some instances, however, an over-reliance on bureaucracy and the lack of proper checks and balances for these systems can cause harm to the citizens of a country. This became abundantly clear during the Dutch childcare benefits scandal (Omtzigt & Vlieger, 2021, pt. 4). In the early 2000s the Dutch government implemented a system that allowed parents to apply for financial aid for childcare. This system turned out to be sensitive to fraud as one could quite easily apply for the benefits without actually making use of childcare. In response to this fraudulent activity, the Dutch government toughened their anti-fraud policy and demanded suspected wrong-doers to immediately pay back all the benefits they received (Omtzigt & Vlieger, 2021, pp. 120–121).

As part of their strategy to find and punish fraudulent activity, the Dutch government made use of machine learning algorithms to analyze the digital profiles of childcare benefits applicants. These

algorithms turned out to be biased (Omtzigt & Vlieger, 2021, p. 141). They exceedingly pointed out people with foreign backgrounds as targets for closer investigation. As a result, not only the offender, but also many innocent users of childcare benefits suffered greatly, since the amounts they had to pay back were often substantial.

On top of the ethical discussions about explainable A.I. that emerged from the childcare benefits scandal, it also revealed a deeper problem concerning the relationship between the government and its citizens. Within the bureaucratic system that the government used to distribute financial benefits fairly, the ones who received those benefits no longer appeared as distinct people. Once someone had received the tag of offender, their distinct identity became invisible and they were treated as if they were identical to proven offenders. In practice, this also meant that innocent people were put on a digital blacklist that kept track of them and prevented them from using other government benefits or sometimes cost them their job (Omtzigt & Vlieger, 2021, pp. 141–142). Thus, a well-intended program by the government to give benefits to those who deserved it was transformed into a bureaucratic monster that eliminated people's distinction, because the inability to consider every case on an individual level led to an over-reliance on systems and algorithms to make decisions.

In a digital world, these situations might occur more often and impact larger groups of people. In chapter three, the device paradigm was discussed as one of the theoretical hurdles for distinction in digital worlds. When people disappear behind a veil of computability, their distinction gets reduced to generalizable statistics. In the case of child benefits in the Netherlands, where the task of individually considering each applicant was considered to be infeasible, it turned out to be tempting to trade distinction for efficiency. In a digital world, where there is more digital data available than in the regular world, and distinct people are more likely to be seen through a device paradigm, political decisions that damage distinction could be more common.

The United States and inequality

In the last decade, the United States has witnessed substantial political polarization. This polarization is also reflected in American society, where groups of people find it difficult to co-exist because of opposing belief systems. The most prominent example of this belief-based separation is the rise of conspiracy thinking (Uscinski & Enders, 2023). So-named conspiracy theorists often reject unifying institutions like the government, the electoral system, or academic institutions.

On top of creating an epistemically divided society, this polarization has amplified inequality. Although social and economic inequality are relevant topics that the United States have struggled with for centuries, and which also might have been influenced by the recent increase in polarization, I will leave those out of the discussion for now. The kind of inequality I want to consider here is the counterpart of equality as I developed it according to Arendt's framework in chapter three. Equality in that understanding means to have a sense of connection that allows people to open a shared space of appearance in which they can perform political action. Collective institutions help this process because they establish and maintain an equal common ground. The fact that not the way that collective institutions function in practice, but rather the very idea of those institutions is being renounced by a growing group of people in the United States is harmful to the sense of equality among citizens.

Many have already pointed to the increasingly important role that social media plays in public debates and evidence-gathering as a main factor in fueling this inequality (Snyder, 2018, pp. 227–232; Uscinski & Enders, 2023). Content algorithms are rewarded for feeding users controversial topics because the economic model behind social platforms depends on engagement and watch time. Echo chambers cause people to get stuck on information islands that are specifically tailored to their existing opinions (Nguyen, 2020). All of these elements contribute to an environment that encourages the emergence of groups, like conspiracy thinkers in the United States, that separate themselves from the public debate. And, for a large part, this is only possible because our communication is increasingly mediated by digital technology.

It is not a long shot to think that someone in a contemporary digital world, like a Hikikomori, might be especially susceptible to this kind of influence. They predominantly communicate with others through digital media and have no other ways of checking their sources than through what they find online. If it turns out that digital media are in fact damaging to equality, then this effect will transfer to digital worlds as well. Instead of a connected digital world that is filled with communities that exchange ideas between them, these communities will be more like isolated islands of in-groups that define their identity by opposing them to the out-group.

This is not necessarily harmful for equality, however. In fact, the feeling of equality within groups might be very strong. And as we discussed earlier, every inhabitant of a digital world might themselves already be considered to be part of an isolated group when they are compared to those

who live outside of the digital world. Still, there might be arguments against this new, strong kind of equality that is formed in small, isolated groups.

First, one would ideally hope that many members could be part of the same equal group and that it does not take a large out-group to realize this, especially since those individuals outside of the group tend to be “othered” and de-humanized. Second, the process of fragmentation into many in-groups is harmful to existing equal communities. We see this happening in American society, where all members of that society are supposed to be equal according to the political institutions that they have built. When this community of equals is split into ever smaller groups, this process destabilizes the whole political landscape, making it harder for everybody to perform political action. In theory, new institutions that stimulate equality could emerge within those smaller groups, but in practice this rarely happens. Those groups tend to produce communities of people who are the same in the sense that they act and think the same, but who are not equal according to our understanding of the concept. According to Arendt, equality means to be distinct, but equal in access to the space of appearance. However, the sameness that often arises in small in-groups tends to compromise this distinction, thereby undermining their equality.

Lastly, the isolation might increase to such an extent that we are not left with small splintered groups but with isolated individuals. Following our insights from Dreyfus’ concept of shared moods, a shared political mood is essential for a sense of equality. However, such a shared mood requires all who are involved to take situated risks. Even if situated risk would be theoretically possible in digital worlds, currently these environments seem to be discouraging individuals from taking such risks. Both Hikikomori and conspiracy theorists, who share a tendency to spend much time on digital platforms, are often characterized by an avoidance of responsibility rather than by the risky and full commitment that is associated with situated risk.

China and predictability

China takes a fundamentally different approach to their political organization than the United States. It has a one-party system with a government that exerts considerable, if not total, control over its citizens. Like all examples in this section, a digital world that emerges in China’s political conditions will most likely have its own idiosyncrasies. Even though challenges for all requirements of political action might arise to a greater or lesser degree, China’s recent development into a surveillance state and its attempts to implement a digital social credit system (SCS) indicates that it is set on increasing predictability (Buckley & Mozur, 2019; Cabestan, 2020; Mac Síthigh & Siems, 2019, pp. 12–19). This

predictability will be an inevitable challenge for political action in digital worlds that are developed under this system.

How can surveillance and an SCS introduce predictability? In chapter four we discussed the theoretical challenges for predictability in digital worlds. There, the metaphysical idea of having a certain number of options for action played an important role. However, predictability can also be created by someone or some institution that is able to exert force. The idea of having an incalculable number of options for every action might sound appealing in theory, but in a reality in which someone is held at gunpoint, there is not much of this freedom left.

If it turns out that certain actions have negative consequences for those performing them, the likelihood of those actions being performed decreases. If the Chinese government tracks one's actions and records them in a SCS, these consequences could be very impactful and long-lasting. Since the Chinese government controls many aspects of life, individuals face the risk of losing access to vital facilities like transportation, accommodation, and employment. This means that many options for action could be eliminated. An additional challenge arises when substantial amounts of audio and visual data are collected. On top of the information that this could provide for an SCS, predicting algorithms could be applied to these data sets.

This could increase predictability in two ways. First, it will support the aforementioned repression of options. When previous behavioral patterns would indicate that someone is likely to perform a certain action, such as committing a crime, then this can be preemptively steered or prevented. Second, as I discussed in chapter four, there could be a psychological element to the existence and application of predicting algorithms. Even if no control is exerted based on these predictions, people could start behaving more predictably because of the hypothetical downsides of unpredictable behavior.

So far, the digitalization of Chinese society has contributed to its capability of both surveillance and SCS management. Therefore, it is likely that the circumstances in a digital world will further increase this capability. Additionally, there is another fundamental difference between digital worlds and the regular world. In a digital world, every action is traceable. While these traces could be secured or hidden, similar to the traces that someone might leave in the regular world, in the example of a realistically conceived digital world many of these options are under government control or regulation. One needs a house, a computer with internet connection, access to food and water, and

sufficient protection from outside threats in order to live undisturbed in a stable digital world. When a government aims to increase predictability, a digital world could prove to be an environment where this is possible to a very high degree and with relative ease.

Russia and reversibility

To many in the western world, Russia has become known as one of the most fervent and effective producers of disinformation. They have actively spread alternative information as a way of exerting influence in foreign countries. This has, for instance, allowed them to influence elections, though the extent of the influence is difficult to determine. Above all, disinformation has proven to be a useful weapon in warfare. Although the idea itself is very old, spreading confusion amongst your enemies on the battlefield has been recognized as a vital strategy for thousands of years, Russia has excelled in applying digital technology to achieve this goal. One of the most prominent examples of Russian information warfare is and continues to be their effort in the Ukraine war (Snyder, 2018, Chapter 5).

In 2014, Russia employed an unusual tactic in Ukraine. They invaded the Crimean Peninsula while denying that an invasion was taking place (Snyder, 2018, pp. 162–163). When armed Russian soldiers in unmarked green uniforms without proper insignia, referred to as “little green men,” were crossing the border and occupying government buildings, they kept up the pretense that nothing was happening (Shevchenko, 2014). Evidence of disturbances was twisted to make it seem like local protests against the Ukrainian government. Such an outrageous transformation of the truth, to the extent that it worked, was enabled, or at least supported, by the inherent uncertainty of digital evidence. When acts of war are recorded on digital media, it is often difficult to determine the authenticity of this digital footage.¹² This is especially the case in unpredictable situations like battlefields. This difficulty is often enhanced by the strategic motivations of parties involved to create obscurity in a so-called fog of war. Digital media have transformed modern battlefields since developments can be followed by the international community on a minute by minute basis. Yet, the apparent clarity that digital media provide can be misleading.

There is a double uncertainty in digital footage because one can almost never be sure of either the time of the recording or the authenticity of the content itself. During conflicts it is common for old footage from different contexts to be shared on social media platforms as if it was currently taking

¹² Strictly speaking, digital media do not have ‘footage’ as film or audio records have. The fact that we still talk about digital footage exemplifies the problem at stake.

place. This is possible because digital evidence usually has no timestamp baked into it that links it to a specific moment. In this sense, digital descriptions of reality are atemporal. In addition, digital footage that is presented as evidence is susceptible to editing or complete fabrication. Especially since the advent of artificial intelligence, the boundaries of what is possible in this regard have shifted significantly. As a result, one can never be entirely sure that digital evidence accurately describes an event in the regular world.

For irreversibility, it is important that an action, once performed, can never be retracted without leaving a trace. Actions should have consequences. In the political situation of Russia, however, these consequences are undermined. The ambiguity of digital evidence is actively exploited for such purposes. As a result, a Russian invasion of the Crimean Peninsula may one day appear nonexistent, and on the other day be presented as local protests against the Ukrainian government. In this scenario, actions that were performed in reality can be transformed or erased, and actions that never happened can have political impact. As a result of the efforts of the Russian government, all action is performed within a reversible context, which severely impedes the possibility of performing political action.

When we move from the situation of present-day Russia, to one where a substantial part of the population lives in a digital world, we can see how reversibility might become an even more acute problem. In the regular world, the consequences of actions are typically aligned with the truth by those who participate in the space of appearance or those who document it, like diarists, journalists, or historians. We assign great value to reports that are made by spectators that are physically present, even though their impressions of unfolding events are limited to an individual perspective and often colored by interpretation. This value is illustrated by the fact that even in dangerous conflict areas, the most reliable sources are reporters who put their lives at risk to report on unfolding events from the ground. Such checks on actions could become almost impossible from the confinement of a room that is connected to a digital world. This is also the case when one argues that the real irreversibility of action takes place in the minds of others. In such a case, the external world might be totally reversible, but those who have experienced previous iterations will be irreversibly influenced. When all action is mediated through digital technology, however, there is never a guarantee that any action will reach other people without intervention. An action could already be totally changed or negated before it even reaches another human living in their own digitally encased bubble.

[3] Conclusion: Enabling political action in digital worlds

Digital worlds will not necessarily steer us in the direction of totalitarianism. Although the previous section highlighted the possible threats for political action when digital worlds carry over and amplify certain totalitarian tendencies from the regular world, digital worlds could also be employed to compensate for those tendencies in order to enable or amplify political action instead. In order to accomplish this, digital worlds should be constructed with certain principles in mind. These principles should be aligned with the requirements for political action. That means that digital worlds should be constructed in such a way that they facilitate distinction, equality, unpredictability, and reversibility. This section will go over these principles and give some examples of how they might be incorporated into digital worlds.

In the case of distinction, this principle requires that people in a digital world should be valued as unique individuals and that they should be enabled to communicate their position in the world. In light of the increased potential for automation and statistical analysis, the importance of case-by-case approaches should be emphasized in digital worlds. This would require policies that limit the amount of automated decision-making that can take place when those decisions concern individuals. Active safe-guarding of the communication channels between individuals and bodies of governance becomes very important to ensure that their relation stays balanced.

Equality should be promoted as well to create the conditions for political action. In present-day digital worlds, it is already very common to form online communities. In addition, as it turns out, people will unite themselves as easily behind popular eSports teams as behind football teams, even if online spaces do not afford the possibility of situated risk. When constructing a digital world, these social platforms and activities should be kept in mind as possible avenues for promoting equality. Digital spaces where people can create a sense of belonging but that do not require an in-group versus out-group dynamic are not directly financially profitable. But when constructing a digital world, investments in such spaces could be made with the aim of fostering political action. Just like governments invest in playgrounds and social spaces in the regular world in order to provide places to meet others and increase social cohesion, such investments should also be done in digital worlds.

Where the promotion of equality might require more regulation, digital worlds need the opposite to ensure unpredictability. In order to create the conditions for political action, there should always be space for action that is not pre-calculated. This is only possible when certain areas of a digital world

are not under surveillance and are not subject to predictive algorithms.¹³ Adhering to this principle in digital worlds effectively comes down to creating and protecting digital private spheres. These spheres would be connected to the public digital world, but not accessible to anyone but the owners. Such a setup requires a more radical delimitation of personal boundaries in digital spaces than we currently experience. One could argue that adopting the Hikikomori lifestyle, centered around living at home, could afford some privacy, since action within that physical space cannot be monitored. Digital privacy, however, should also include features that allow others to be invited into a digital private sphere and the ability to move around in the digital world without constant observation. Just as initiating political action alone from one's room proves challenging in the regular world, a similar difficulty arises in a digital world. This strongly suggests that the boundaries of unpredictability should be expanded in such a digital realm.

Lastly, in order to enable political action in digital worlds, these worlds should be constructed according to the principle of irreversibility. This means that the consequences of an action should be preserved to a certain extent. While it is difficult to evaluate the authenticity of digital artifacts, this process could perhaps be supported by the very technology that enables the creation of fake digital artifacts. Artificial intelligence could be employed to trace the creators of digital content in the public domain.¹⁴ These tools could be used by organizations of journalists in the digital world to distinguish between political action that is performed by individuals and digital noise.

Using the requirements of distinction, equality, unpredictability, and irreversibility as principles when constructing digital worlds will guide their development in a direction that enables political action rather than totalitarianism. Although the previous section illustrated some examples of how these principles might be applied in digital worlds, their application will depend on how and in which social-political context a digital world evolves.

Effectively incorporating the principles for political action into the construction of digital worlds is a topic that goes beyond this thesis. By defining these principles in a context of realistically conceived digital worlds, however, we have acquired a better understanding of the possible challenges for political action in such worlds. Moreover, this endeavor has developed a theory of political action, provided to us through the work of Arendt, and made it a workable theory for an investigation of

¹³ Blockchain technology might prove to be a useful tool for creating such private areas.

¹⁴ This technology should only be used in the public domain and not in the private domain so as to maintain this boundary that is required for unpredictability.

digital worlds. In addition, this has shed light on broader contemporary issues of digital technology through the lens of political action. This will aid us in both our understanding of those issues and in our complex navigation through the digital landscape.

Conclusion

The introduction of this thesis pointed at a void in our understanding of the human condition. This void has been created by the recent emergence of digital technology. Especially in developed societies, the role of such digital technology in daily life is increasingly pervasive. So much so that neither the classical understanding of technology as a tool, nor the realization that technology and society co-constitute each other is sufficient to describe the impact of digital technology. Recent philosophical explorations of digital worlds, like the one by Chalmers, underpin the need for a new understanding of the digital human condition. The very idea of a digital world, pushes the analysis of technology to a further stage. Not only can our world be shaped by digital technology, it could be constructed out of it.

One of the reasons why the impact of such a pervasive relationship with technology is not clearly understood is because we lack a philosophical framework to effectively discuss digital worlds. In order to improve this understanding, the aim of this thesis has been twofold. The primary research question that we were concerned with was whether and under which circumstances digital worlds provide a conducive environment for political action. The secondary research question of our investigation was how Arendt's political action theory can be applied to effectively analyze and understand political action within digital environments.

Chapter one initiated this investigation by elaborating on the concept of digital worlds. Firstly, this chapter established a concept of digital worlds that determined the limits of our investigation. Such a digital world should be *interactive* and *social*, in order to be able to raise questions about action between people, which is an essential feature of political action. It should also be *complete*, in order to isolate the effect of the digital world on political action. Next, the first chapter contrasted digital worlds against the regular, non-digital world along three dimensions: *simulation*, *virtuality*, and *digitality*. Lastly, this chapter discussed two philosophical positions with contrasting views on the capabilities and challenges of digital worlds, that of digital realism and the phenomenological position. By incorporating these perspectives into our analysis, we formulated a concept of digital worlds that acknowledged the reality of such worlds while also considering their phenomenological experience.

Having laid a conceptual foundation of digital worlds in chapter one, the second chapter turned to the concept of political action. This chapter developed the ideas of Arendt into a framework for political action that could be used for an analysis of digital worlds. This was done by first identifying political action as a mode of action that can be contrasted with labor and work. Additionally, political action was established as a dynamic process that can be performed in a space of appearance. In order to provide us with some footing for analyzing the possibility of political action in digital worlds, chapter two then outlined four requirements for such action: distinction, equality, unpredictability, and irreversibility. Chapter two, then, developed *distinction* and *equality* as requirements for those who perform political action, and *unpredictability* and *irreversibility* as requirements for the way that such action is performed. Lastly, totalitarianism was introduced as a threat for worlds that do not meet these requirements, which gave us a basis for understanding digital worlds without political action.

Then, chapters three and four took the four requirements for political action and considered the theoretical possibility of such action in digital worlds. In the case of distinction, we discovered that digital worlds might prevent people from communicating themselves to others as unique individuals in digital worlds. Equality might be hindered in such worlds by the inability to create collective moods between individuals. Unpredictability could be theoretically limited in digital worlds because the computability of such worlds imposes a calculable amount of options for action. In turn, irreversibility could be hindered in digital worlds because the processes of creation and destruction differ from the regular world. Thus, based on the four requirements for political action, chapters three and four identified a set of theoretical limitations for such action in digital worlds.

Finally, chapter five applied the requirements for political action to realistically conceived digital worlds. Here, we moved from theoretical considerations to practical application. This allowed us to identify the practical challenges for political action that potentially could lead to the emergence of totalitarianism. These challenges were illustrated by four vignettes of different societal contexts we see in our world. Assuming that the society in which digital worlds emerge will affect the particularities of such worlds, chapter five identified crucial elements of these different societies that might prove to be especially challenging to political action when they are carried into digital worlds: model-thinking in the Netherlands, splintered communities in the United States, surveillance in China and the creation of alternate realities in Russia. These cases highlighted four tendencies in our world that can serve as warning signals, cautioning against pathways leading to digital realms that foster

totalitarianism. Based on the identified challenges, chapter five also proposed some measures that could be taken to enable political action instead of totalitarianism.

By developing the concept of digital worlds in such a way that the question of political action can be discussed, this thesis has fundamentally progressed our understanding of the impact of digital technology. Although complete digital worlds have been the subject of discussion before, especially in the writings of Chalmers and Dreyfus, it often remained unclear what the concept entailed exactly. This thesis has rendered the concept of digital worlds operational by framing it in such a way that it can be analyzed as a platform for political action. This has allowed for a comparison with the regular world along the dimensions of simulation, virtuality and digitality, and facilitated an evaluation according to insights from both digital realism and phenomenology. This framework both acknowledges the metaphysical correspondence between the regular and digital world, and the phenomenological differences between them.

Crucially, by adopting the political action theory of Arendt and applying it to our concept of digital worlds, we have acquired a better understanding of this theory. Most importantly, this framework has proven highly valuable in understanding the impact of digital technology on the human condition and in elucidating how political action functions in this *digital condition*. This theory has demonstrated its relevance not only in a theoretical analysis of digital worlds, but also in practical, applied investigation.

Finally, the analysis of digital worlds yields valuable insights into pressing contemporary issues surrounding digitality. While the primary focus has centered on political action, the identified challenges provide a crucial vantage point for addressing urgent issues like privacy, identity, friendships, cybercrime, and alienation in digital worlds.

Our approach of analyzing political action in digital worlds has inherent limitations as well. By defining a specific concept of digital worlds, we implicitly neglect other interpretations of such worlds. What would a digital world look like without the social requirement of being inhabited by multiple people, for instance? One could imagine a digital world where people can only interact with artificially intelligent networks instead of other people. Or how about a digital world where both people *and* artificial intelligences co-exist? Since the rise of A.I. technology, these questions have become increasingly relevant and they could potentially reshape our understanding of the digital

condition. Consequently, future research of digital worlds should pay attention to the intersection with topics in artificial intelligence.

The requirement of digitality itself could also be limiting to our investigation. It might very well be that digitality is not a sufficient concept to understand the nature of digital worlds because they might, in fact, be constructed out of elements that are non-digital. In addition, breakthroughs in the field of quantum computing could cause our digital paradigm to be outdated. In any case, the nature of digital worlds and their relation with the regular world should remain a topic of investigation.

The investigation of political action in digital worlds could also be extended beyond an Arendtian interpretation of action. Perhaps other theories would lead to different evaluations of the possibility of political action in digital worlds. It would be interesting to compare these differences because it could lead to a more profound understanding of both the concept of digital worlds and that of political action.

There also remains something to be said about the application of the framework that this thesis developed. Although the fifth chapter identified some of the challenges for political action, these examples certainly do not exhaust the discussion of totalitarianism in digital worlds. Likewise, the principles that were suggested to enable political action can be elaborated on in further detail. As mentioned before, there are many links to be found between the conclusions of chapter five and contemporary issues surrounding both digital worlds and political action. By further incorporating the Arendtian framework in the discussion of the subjects, further research has the potential to articulate guiding principles for political action in digital worlds more clearly.

Although we might still be far removed from a future where complete digital worlds are a common environment in which to live one's life, digital technology has already started to fundamentally change the human condition and the societies we inhabit. This is a process that cannot be undone. What can be achieved, however, is to determine the direction that digital technology will take us. If this road does indeed lead to digital worlds, then it is of the utmost importance that these worlds are fit for humans to live in, allowing them to live a full and good life. This means that we should construct them as such. We can only achieve this by developing a deep understanding of humans, machines, and the complex interactions between them. To deepen our understanding of the digital

condition has been the focus of this thesis, and remains one of the most pressing tasks of future philosophical research.

Literature

Arendt, H. (2017). *The Origins of Totalitarianism*. Penguin Classics. (Original work published 1951)

Arendt, H. (2018). *The Human Condition* (Second edition). University of Chicago press. (Original work published 1958)

Bohman, J. (2008). The Transformation of the Public Sphere: Political Authority, Communicative Freedom, and Internet Publics. *Information Technology and Moral Philosophy*, 66–92.

Borgmann, A. (1987). *Technology and the Character of Contemporary Life: A Philosophical Inquiry*. University of Chicago Press.

Borgmann, A. (1993). *Crossing the Postmodern Divide*. University of Chicago Press.

Borgmann, A. (2000). *Holding On to Reality: The Nature of Information at the Turn of the Millennium*. University of Chicago Press.

Boyd, D. (2008). Can Social Network Sites Enable Political Action? *International Journal of Media & Cultural Politics*, 4(2), 241–244.

Buckley, C., & Mozur, P. (2019, May 25). How China Uses High-Tech Surveillance to Subdue Minorities. *International New York Times*.

Cabestan, J.-P. (2020). The State and Digital Society in China: Big Brother Xi is Watching You! *Issues & Studies*, 56(01), 2040003. <https://doi.org/10.1142/S1013251120400032>

Capurro, R. (2011). Never Enter Your Real Data. *International Review of Information Ethics*, 16, 74–78.

Chalmers, D. J. (2017). The Virtual and the Real. *Disputatio*, 9(46), 309–352.

Chalmers, D. J. (2022). *Reality+: Virtual worlds and the problems of philosophy* (First edition). W. W. Norton & Company.

Dahlberg, L. (2001). The Internet and Democratic Discourse: Exploring The Prospects of Online Deliberative Forums Extending the Public Sphere. *Information, Communication & Society*, 4(4), 615–633.

- Douglas, D. M. (2016). Doxing: A Conceptual Analysis. *Ethics and Information Technology*, 18(3), 199–210.
- Dreyfus, H. L. (1999). Kierkegaard on the Internet: Anonymity vs. Commitment in the Present Age. *Kierkegaard Studies Yearbook*, 96–109.
- Dreyfus, H. L. (2001). *On the Internet*. Routledge.
- Elder, A. (2014). Excellent Online Friendships: An Aristotelian Defense of Social Media. *Ethics and Information Technology*, 16(4), 287–297.
- Ess, C. (1996). The Political Computer: Democracy, CMC, and Habermas. *Philosophical Perspectives on Computer-Mediated Communication*, 197–230.
- Ess, C. (2010). The Embodied Self in a Digital Age: Possibilities, Risks, and Prospects for a Pluralistic (Democratic/Liberal) Future? *Nordicom Review*, 31, 105–118.
- Floridi, L. (2022). Metaverse: A Matter of Experience. *Philosophy & Technology*, 35(3), 73.
- Flyverbom, M., Deibert, R., & Matten, D. (2019). The Governance of Digital Technology, Big Data, and the Internet: New Roles and Responsibilities for Business. *Business & Society*, 58(1), 3–19.
- Guillot, J. D. (2023, November 22). Shaping the Digital Transformation: EU Strategy Explained. *European Parliament*.
- Heidegger, M. (1977). *The Question Concerning Technology, and Other Essays* (W. Lovitt, Trans.). HarperCollins Publishers. (Original work published 1954)
- Heidegger, M. (2008). *Being and Time* (J. Macquarrie & E. Robinson, Trans.). Harper Perennial/Modern Thought. (Original work published 1927)
- Itrona, L. (2017). Phenomenological Approaches to Ethics and Information Technology. In *The Stanford Encyclopedia of Philosophy* (Fall 2017).
- Kateb, G. (2000). Political Action: Its Nature and Advantages. In D. Villa (Ed.), *The Cambridge Companion to Hannah Arendt* (pp. 130–148). Cambridge University Press.

- Kato, T. A., Shinfuku, N., & Tateno, M. (2020). Internet Society, Internet Addiction, and Pathological Social Withdrawal: The Chicken and Egg Dilemma for Internet Addiction and Hikikomori. *Current Opinion in Psychiatry*, 33(3), 264–270.
- Kim, J. (2001). Phenomenology of Digital-Being. *Human Studies*, 24(1–2), 87–111.
- Kohn, J. (2000). Freedom: The Priority of the Political. In D. Villa (Ed.), *The Cambridge Companion to Hannah Arendt* (pp. 113–129). Cambridge University Press.
- Kuklick, B. (2022). *Fascism comes to America: A Century of Obsession in Politics and Culture*. The University of Chicago Press.
- Ladyman, J. (2020). Structural Realism. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2020). Metaphysics Research Lab, Stanford University.
- Mac Síthigh, D., & Siems, M. (2019). The Chinese Social Credit System: A Model for Other Countries? *The Modern Law Review*, 82(6), 1034–1071.
- Nguyen, C. T. (2020). Echo Chambers and Epistemic Bubbles. *Episteme*, 17(2), 141–161.
- Nissenbaum, H. (2009). *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford University Press.
- Norton, R. (1972). What Is Virtuality? *The Journal of Aesthetics and Art Criticism*, 30(4), 499–505.
- Omtzigt, P., & Vlieger, W. (2021). *Een Nieuw Sociaal Contract*. Prometheus.
- Papacharissi, Z. (2004). Democracy online: Civility, Politeness, and the Democratic Potential of Online Political Discussion Groups. *New Media & Society*, 6(2), 259–283.
- Piccinini, G., & Maley, C. (2021). Computation in Physical Systems. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2021). Metaphysics Research Lab, Stanford University.
- Rooksby, M., Furuhashi, T., & McLeod, H. J. (2020). Hikikomori: A Hidden Mental Health Need following the COVID -19 Pandemic. *World Psychiatry*, 19(3), 399–400.
- Shevchenko, V. (2014, March 11). ‘Little green men’ or ‘Russian invaders’? *BBC News*.

- Smith, D. W. (2018). Phenomenology. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2018). Metaphysics Research Lab, Stanford University.
- Snyder, T. (2018). *The Road to Unfreedom: Russia, Europe, America*. The Bodley Head.
- Turkle, S. (2011). *Alone Together: Why we Expect More from Technology and Less from Each Other*. Basic Books.
- Uscinski, J. E., & Enders, A. M. (2023). What Is a Conspiracy Theory and Why Does It Matter? *Critical Review*, 35(1–2), 148–169.
- Vallor, S. (2012). Flourishing on Facebook: Virtue Friendship & New Social Media. *Ethics and Information Technology*, 14(3), 185–199.
- Vallor, S. (2022). Social Networking and Ethics. In *The Stanford Encyclopedia of Philosophy* (Fall 2022). Metaphysics Research Lab, Stanford University.