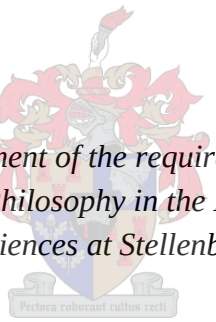
AI & Bioterrorism: An Overview of the Ethical Risks Involved

by
Tristesse Erasmus



Thesis presented in fulfilment of the requirements for the degree of
Master of Philosophy in the Faculty of
Arts and Social Sciences at Stellenbosch University.

Supervisor: Professor JP Smit

December 2021

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

December 2021

Abstract

In modern day society, we are becoming all the more dependent on technology and its continued advancements. Technological advancement, however, is not a wholly beneficial issue. The *dual nature* of technology is highlighted with research and developments that can be both beneficial and detrimental to society. The dual nature thus refers to technological advancements and developments or discoveries in research that have the potential or are likely to harm society just as much as a society can benefit from this innovation. This thesis will focus specifically on the dual nature of Artificial Intelligence (AI) and its related technologies to highlight the potential for AI-enabled bioterrorism. A focus on our ethical obligations as members of society, scientists, doctors, academics, theorists, and the like will be present throughout.

As the dual nature of AI is not something that can be overcome fully, the issue at hand pertains to the prevention and mitigation of bioterror incidents. What ethical measures should be put in place to prevent, detect, mitigate, and respond to AI-enabled bioterror incidents? AI has a long history of doom and gloom attributed to it, where the imaginings of filmmakers and scientists came together to determine that AI could be both our saving grace and our undoing. The notion of the AI-overlord that will become conscious and destroy humanity will not be the focus of this thesis.

This thesis will focus on the notion that bioterrorism planning and perpetuation is eased with the use of biological data and AI. While it is entirely true that bioterror can be perpetuated without access to biological data, the use of biological data enables the attack to be far more accurate, efficient, and effective. Further, by using the currently available AI systems to perform computational tasks on this data, the process of planning and perpetuating a bioterror attack is sped up exponentially. I focus specifically on AI-enabled bioterror, aided by biological data, as bioterror incidents have seen a rapid increase since the 1990s. This forecasts that bioterror will either continue increasing at the same rate or increase over time and reflects the concern of this thesis directly. As technology advances, biological data and AI both contribute to the exponential threat of bioterrorism.

This thesis is an overview of four distinct but wholly interrelated topics, namely Data Mining, Biological Data, Bioinformatics, and AI, and the ethical considerations involved in each level of AI evolution due to the prevalence of the dual nature of technology. This will serve as the reasoning behind the several recommendations that will be made in the final chapter, each focusing on a distinct aspects of bioterror prevention, mitigation, and response.

AI is a dual nature technology, as are the technologies that lead to its development. However, the focus on evil, conscious, super AI is misguided and ignores very real and current issues that affect us today. This kind of AI simply is needed for the perpetuation of biological attacks. No matter how much time and effort we put into the ethical development of AI and its related technologies, the dual nature will remain. Therefore, we must do our best even when we know that we will never fully succeed. If we apply ethical considerations to each level of AI development, we will be in the best possible position to respond to bioterror ethically and effectively.

Abstrak

In die moderne samelewing raak ons al hoe meer afhanklik van tegnologie en die vooruitgang wat dit bring. Tegnologiese ontwikkeling is tog, in geheel beskou, nie slegs voordelig nie. Die *dualistiese aard* van tegnologie is duidelik in navorsing en ontwikkeling wat beide voordelig en nadelig vir die samelewing kan wees. Die dualistiese aard verwys dus na tegnologiese vooruitgang en ontwikkeling of ontdekking in navorsing wat die potensiaal het om die samelewing te kan benadeel, net soveel soos wat die samelewing uit hierdie vernuwing kan voordeel trek. Hierdie tesis fokus spesifiek op die dualistiese aard van Kunsmatige Intelligensie en verwante tegnologie om sodoende die potensiaal vir KI-bemagtige bioterreur uit te lig. ‘n Fokus op ons etiese verpligtinge as lede van die samelewing, wetenskaplikes, doktors, akademici, teoretici en diesulkes sal deurgaan beskou word.

Omdat die dualistiese aard van KI nie in die geheel oorkom kan word nie, het die kwessie op hande betrekking op die voorkoming en mitigasie van bioterreur insidente. Watter etiese beginsels behoort toegepas te word om KI-bemagtigde bioterreur insidente te voorkom, bespeur, te mitigeer and daarop te reageer? KI word gekoppel aan ‘n lang geskiedenis van somberheid en veroordeling daarvan, waar die verbeelding van filmmakers en wetenskaptelikes ontmoet om te bepaal dat KI beide ons redding en ons ondergand sal wees. Die idee van n KI-oorheerser wat ‘n werklikheid sal raak en die mensdom sal vernietig, sal egter nie die fokus van hierdie tesis wees nie.

Hierdie tesis sal fokus op die feit dat bioterreur se beplanning en voortbestaan vergemaklik word deur die gebruik van biologiese data en KI. Alhoewel bioterreur wel kan gebeur sonder toegang tot biologiese data, kan die gebruik van biologiese data sulke aanvalle meer akkuraat en effektief maak. Verder, is dit ook die geval dat die gebruik van huidig beskikbare KI om berekeninge op sulke data uit te voer die beplanning en bestaan van bioterreur eksponensieel bespoedig. Ek fokus spesifiek op KI-bemiddelde bioterreur, d.m.v die gebruik van biologiese data, aangesien insidente van bioterreur vinning toegeneem het sedert 1990. Dit voorspel dat bioterreur of sal anhou groei of dat die toename self sal versnel en is 'n basiese motivering vir die kwesses bespreek in hierdie tesis. Soos tegnologie verbeter sal die gebruik van biologiese data en KI beide tot die eksponensiële gevaar van bioterreur bydra.

Hierdie tesis is n oorsig van vier onderskeie maar in die geheel verbonde onderwerpe, naamlik Data Ontginning, Biologiese Data, Bioinformatika, asook die etiese oorwegings betrokke op else vlak van KI evolusie as gevolg van die voorkoms van die dualistiese aard van tegnologie. Hierdie sal die redenasie agter die verskeie aanbeveelings wees, soos in fie finale hoostuk, waar ek fokus op die aspekte can bioterreur voorkoming, mitigasie, and reaksie.

KI is tegnologie van ‘n dualistiese aard, net soos die tegnologie wat daaraan ontstaan gegee het. Die fokus op bose, bewuste, super KI is egter misleidend en ignoreer vele werklike en huidige kwessies wat ons beïnvloed. Hierdie tipe KI is eenvoudig nie nodig vir die voortbestaan van biologiese aanvalle nie. Ongeag van hoeveel tyd en moeite mens in die etiese ontwikkeling van KI en verwante tegnologie insit, bly die dualistiese aard daarvan ‘n gegewe. Ons moet dus om daardie rede soveel as moontlik doen selfs al besef ons dat ons nooit volledig daarin sal slaag nie. Indien ons etiese oorwegings op elke vlak van KI ontwikkeling toepas, sal dit ons in staat stel om eties en effektief op bioterreur te kan reageer.

Acknowledgements

Without the emotional, mental, psychological, humorous, and financial support of several people in my life, this thesis would have never been possible. I thank you all.

Firstly, I would like to thank my supervisor, Prof. J.P. Smit, who has guided me through two theses in the past three years. I appreciate your continued mentorship and guidance.

Secondly, I would like to thank my partner, Koogan Archary, for believing in me when I struggled to believe in myself and supporting me through years upon years of student life.

Lastly, I would like to thank my parents and little brother. Thea and Sybrandt Erasmus instilled in me the love of academia and education that has led me to not only becoming the person I am today but the person I always dreamed of becoming. And to Liam, your soul is a shining light in this dark world. Even as a young adult, you give me hope for future generations with your love for and mindset about the world. I hope to make you proud as you make me proud to be your sister.

Table of Contents

Introduction

In this thesis, I will demonstrate the clear and fundamentally dual nature of technology and provide an overview of ethical concerns, considerations, and challenges that we face due to the exponential advancement of technology and all its related fields. These considerations do not attempt to judge the situation but to guide and stipulate how to benefit from the dual nature fully while maintaining a healthy concern for the potential negative consequences. *Dual nature* will be defined as the potential of something to be used for supporting and damaging society, something that can be used for just as many unfortunate occurrences as positive outcomes. With regards to technology, the positive side supports and provides for the wellbeing of humankind while the negative side could lead to the harm or destruction of society.

This thesis is considered an overview of the different topics and ethical considerations that are present in the discussion of the dual nature of technology as I do not intend to argue for a specific approach that will solve this problem. This thesis is an overview of these ethical considerations and will not focus on specific ethical discussions or attempt to adjudicate the various standpoints involved. The focus of this thesis will be general ethical considerations that pertain specifically to dual nature technology due to the potential for harm. This is largely because we cannot fully solve this problem and must consider viewing this issue as a tradeoff instead, where we offer up something, like a small portion of our private information, so that we may benefit from technology and all that it is able to provide for us, now and in the future. This thesis will not offer a solution to the problem but mention each concept and field to highlight their dual nature and the reason for concern. This will lead to several recommendations for dealing with this dual nature to show us, quite literally, how we can do our best to protect society.

This dual nature, referred to as the dual dilemma as well, is the focal point of this thesis and serves as the foundation for the ethical issues that will be mentioned. By investigating the dual nature of technology in the fields of data mining, biological data, and bioinformatics, this thesis will emphasize how the dual nature of technology is ever present and of great concern when considering advancements in Artificial Intelligence (AI). The question that arises here focuses on how we can benefit from the dual nature of technology while mitigating the risks involved.

To answer this pivotal question, this thesis will highlight the relationship between the process of data mining, how this contributes to the collection of big data, biological data, and further, big biological data. Data mining can be considered the beginning of this process which inevitably, with its advancements, leads to more and more collected data. As this data continues, it eventually reaches such a vast amount that we consider it to be big data. Data mining is an ever-present process and while it is important and influential in a variety of other industries and fields, this thesis will focus on the kind of data mining that contributes to the collection of biological data and big biological data. Within the field of biological data, data mining is necessary so that this data can be collected at a sufficient, accurate, and timely rate. As with data mining and big data,

as more and more biological data is collected, it eventually reaches the category of big biological data.

The thought may well occur to the reader that while the link and potential harm of the use of data mining in bioterror is obvious, the role of AI is less clear. I think, however, it is a legitimate concern for the following reasons. It is important to note that while biological data is not essential to plan and perpetuate a biological attack, it certainly provides advantages in terms of the time it takes to plan the attack and the overall success of the attack. With the aid of biological data, bioterror attacks can specifically target certain communities, areas, or cities, on small or large scale, with the most effective kinds of biological agents, as can be deduced by geographical disease data. Certain biological agents will be more or less effective on certain groups of people. With detailed data about the general diseases that each of these communities or areas suffer from, specific and localized attacks with the most effective biological agents can be planned with ease.

Bioinformatics had to advance to keep up with this vast amount of data that was being collected in the realm of biological data, increasing its technological abilities in terms of speed, computational capacity, and more. This means that as more and more biological data was collected, and as it continues to be collected daily, technology had to evolve to keep up with the demand for improved computational technology. Furthermore, with the advancements in data mining and biological data, and the required advancements in bioinformatics, the road is being paved for AI that is far more advanced than we currently have access to. AI requires these technologies to develop to a certain extent so that its computational capacity is maximized. To do this, incredibly vast amounts of data is required, which is provided by big data and big biological data. Machine learning, deep learning, and neural networks, developed within the field of bioinformatics, all contribute and are required to advance AI. These are the very elements AI requires to exist, and they must continue to advance and to allow for faster, smarter, and more capable AI to develop.

It is important to note that this thesis will not focus on the concept of a fully autonomous and conscious AI system as the concern of this thesis is relevant to our current AI. The concern here is not predicated on the idea of an AI overlord who seeks to destroy or enslave humanity, but on the currently available AI technologies that can be used to aid the planning and perpetuation of biological terrorism. The simple truth is that we don't need an evil AI superbeing to destroy humanity. This is entirely accomplishable through applying the current advancements in AI technologies and systems to the perpetuation of bioterror. It is important to note that we do not apply an agential understanding of AI, in the sense that it is not responsible for its actions the way a human is responsible for their actions. However, it can be used to aid bioterrorists in their attempts to harm society due to its advanced computation capacity that far outweighs that of the average human. This thesis does not claim that AI is a necessity in the perpetuation of bioterror, simply an incredibly useful and efficient tool with which to plan and execute more complex attacks that are more difficult to detect and treat than what society is prepared for. The risk here is existential as we are faced with the possibility of future pandemics and biological attacks that aren't just potentially synthetic, based on some conspiracy theories.

In Chapter 1 of this thesis, I will define and investigate the topic of data mining as the first step in the technological evolution of the dual dilemma, with specific reference to the potential for malicious intervention and intention that relates to the perpetuation of bioterror. Briefly, the data mining process will be considered to convey how data is mined and how easily and effortlessly data is mined. This will clarify how the field of biological data had the opportunity to expand to big biological data, or the mass collection of medical or scientific data relating to the human body and our existence.

Following on this, this thesis will turn to arguments for and against data mining, highlighting both its advantages and its disadvantages. Ultimately, while I will admit that data mining is highly effective and effortlessly efficient in fulfilling its purpose, it is my view that the ethical risks so far outweigh these advantages if we do not have protective measures in place. This is not to say, and this thesis will not argue, that data mining should cease. This exercise will emphasize that we simply cannot ignore the dangers that we are faced with.

Finally, this chapter will end with privacy and security as a foundation for an explanation about Privacy Preserving Data Mining (PPDM) techniques. This will lead into the second chapter, which focuses wholly on biological data and the ethical concerns, challenges, and risks that arise due to the vast and increasing collection of biological data.

In Chapter 2 of this thesis, the focus will be on the topic of biological data and the ethical concerns that arise at every level of this process. These concerns contribute to the larger concern of this thesis, namely the perpetuation of bioterror through the vast collection of biological data that is available and is becoming available every day. These ethical responsibilities address the issue that some research is of a dual nature, whether we are knowingly conducting potentially dangerous research or not. This will highlight the requirement of the governing principles of ethics we expect of researchers in the name of society's wellbeing.

To follow, this chapter will mention the guidelines related to the ethical governance of databases, as databases focus on the storage of biological data. The concern here is clear when we consider how influential AI has been during the current COVID-19 pandemic and understand that the very same tools that have been used to aid people in this pandemic can be used to perpetuate the very same pandemic. The trend of the dual dilemma and the necessity of trading off advantages and disadvantages will be present throughout this chapter.

Lastly, this chapter will focus on not only bioterror preparation and the ethical concerns that we need to address and keep in mind when preparing for a potential bioterror event or occurrence but the aspect of biodefense and the ethical response to bioterrorism. As has been mentioned previously, biological data is helpful to localizing and specifying the most effective biological attacks. The same is true for the scientists and government bodies as they can use the same data to determine where the most likely attacks are to occur and which agents could be used.

This highlights the dual nature of biological data and its absolute usefulness in either planning a bioterror attack or planning to prevent one. It is important to stipulate that biological data is, just as with AI, not necessarily required for the perpetuation of a bioterror attack. However, it is helpful and effective information to use when determining which harmful biological agents will be most effective.

In Chapter 3, attention will be focused on bioinformatics and its related technologies as the necessary evolution of technology in the development of AI. Bioinformatics is a fundamentally crucial step in the development of artificially intelligent systems that are capable of contributing to biological attacks. Ethical concerns regarding the development of bioinformatic technologies rest on several issues that will be considered here. This leads to ethical challenges that are present in the storage of biological data in data warehouses.

Deep learning is mentioned as a precursor to the description of Artificial Neural Networks (ANNs), where the focus is on several concerns that have evolved with the development of bioinformatics technologies as the need for these advancements grew. Lastly, this chapter will mention two aspects regarding their applications and the concerns that are involved in their use. ANNs are a vital development in the field of deep learning and contribute immensely to the development and advancement of AI technologies.

In Chapter 4, this thesis will turn its focus to AI. By examining AI in terms of the healthcare industry, the prevailing focus on the dual dilemma we face becomes clear in the field of AI as well. Several theorists, including Searle, argue against the possibility of strong AI which is required for the theorized “evil overlord AI” which is supposed to destroy humanity and take over the world. This, however, solidifies my point that we don’t need strong AI to destroy humanity.

The link between bioterrorism and biological data, as mentioned above, might very well be obvious. In the case of AI, its link to bioterror may not be seen as obvious. Note, however, that artificially intelligent systems, as understood in this thesis, offer an immense advantage over and above the skills and capabilities of human beings. The AI simply makes it a lot easier or more convenient to do this. Weak AI is our currently available AI technology, whereby the system can run automated processes to reduce the burden of data analysis. These systems surpass human-level computational ability, exponentially decreasing the time requirements of mass data analysis. The kinds of algorithms required to do this have greater computational ability than our average computers and laptops, but do not come close to the notion of Strong AI, to be discussed later.

This chapter will focus on 5 ethical considerations that pertain to the use of AI in the healthcare field. These considerations serve as a foundation for the following section which highlights how unprepared the world was for the current pandemic to highlight the very apparent and obvious need to improve our preparation for and responses to bioterror and biological occurrences. It should be emphasized that while this thesis does not entertain conspiracy theories

about the origin of the COVID-19 virus, it is a useful example to illustrate what could occur if bioterrorists had access to AI and enough data to develop the perfect, custom biological agent.

The focus of this chapter and of this thesis ends with the recommendations that follow. Each of these recommendations highlight an important aspect of improving our preparedness for future pandemics and biological attacks. These recommendations include changes and improvements to policies, training, resources, the monitoring of potentially harmful research, risk and threat assessments, and biodefence in terms of prediction, detection, and prevention.

At the end of this thesis, the dual nature of technology and the risks involved when considering the influence of AI and its continued and exponential advancements should be apparent. The need for intervention, improvement, and advancements in our preparedness for biological occurrences is abundantly clear when looking at the global response to COVID-19. In finality, this thesis will have provided a grand overview of data mining, biological data, bioinformatics, and AI, the risks of biological occurrences, and the dual nature of technological advancement that may render us in a worse off position than the academically and theatrically asserted notion of the evil, autonomous AI overlord. The goal of this overview is to clarify the ethical considerations that should dictate how AI is developed and used, through applying these considerations to all of the building blocks involved in the development of AI.

Chapter 1
The Data Mining process, ethical considerations, and privacy preservation

1. Introduction

There is much controversy to be discussed concerning the topic of data mining. I will define *data mining* as the process of extracting explicit and implicit data from any form or format, from audio files to photocopies and uploaded textbooks from a century ago. This controversy, on the one hand, is that those in support of data mining argue that the risks and costs of the process are rare when considering the immense benefit of mining data in any sector of the economy. The benefits so far outweigh the risks that it seems quite ridiculous to some of these supporters that data mining could be considered harmful. On the other hand, those who are opposed to data mining due to the privacy and ethical concerns it poses do not intend to claim that we should not be data mining *at all.* The purpose is not to strip the world of the clear benefits that data mining provides, only to prevent misuse, abuse, theft, and privacy violations.

In this chapter, I will take the stance of the latter. The data mining process is a rich, wonderful world where detailed information about billions of different scenarios, people, businesses, and more, is readily available to us. However, as a society, we cannot ignore the potential risks. We cannot claim that the benefits are far greater than the potential cost that may be incurred when people stand to lose their private medical, social identification, and financial information. The argument that I put forward, at this level, is that while data mining is useful and provides great benefits, we need measures in place to protect people's information from malicious actors. Privacy protection must begin at the most fundamental level of technological advancements: the data mining process.

This chapter will serve as a foundation for the closing chapter in this thesis, AI and Bioterrorism, to highlight the most fundamental tool in the creation of AI. Without data mining, AI would be impossible as AI cannot exist without *big data. Big data* can be defined as the mass of collected data that has been collected over the years and through a world-wide platform as well as the related technologies that make this possible (Taylor-Sakyi, 2016; Riahi & Riahi, 2018:524). AI cannot exist without enough data to learn from and therefore requires big data, and big data cannot exist without data mining. Moreover, big data cannot exist without the intense and mass collection of data from all fields, sectors, and countries; this is where it gets its name in the first place (Taylor-Sakyi, 2016; Riahi & Riahi, 2018:524).

2. The Data Mining Process

The focus, here, is the data mining process as a 5-step process. These steps include data cleaning, data integration, data transformation, data mining, and knowledge representation. Each

of these steps is a fundamentally important part of this process, and without the completion of even one, the process falls short and fails (Colonna, 2013:315; Imberman, 2014).

There are a multitude of reasons why data mining is requested or completed, and the applications go far beyond what I could include in this thesis. It is necessary, however, to deal with one common point of confusion with the "knowledge discovery" or data mining process. "The phrase "knowledge discovery in databases" ("KDD"), coined by Gregory Piatetsky-Shapiro in 1989, has a broader meaning than data mining. KDD denotes the entire process of using unprocessed data to generate information that is easy to use in a decision-making context" (Colonna, 2013:315; Imberman, 2014). Further, another point of clarification is necessary regarding the differentiation between the term KDD and data mining.

To clarify the important terminology here, a description of this process by Sarker et.al. (2000:1):

> Knowledge discovery in databases (KDD) is the process of extracting models and patterns from large databases. The term data mining (DM) is often used as a synonym for the KDD process although strictly speaking it is just a step within KDD. DM refers to the process of applying the discovery algorithm to the data.

The point to note here is the statement that, while data mining is often used as a synonym, its technical definition is just one part of the Knowledge Discovery in Databases Process (KDD) (Sakar et.al., 2000:1; Colonna, 2013:315; Imberman, 2014). For this thesis, I will take an informal stance on this and refer to the process as "the data mining process."

For the purposes of what follows, both Van Wel & Royakkers (2004:130) and Madria, Bhowmick, Ng & Lim (1999) distinguish between three kinds of data mining. The first is *usage mining*, which determines how much we use the internet. This will measure the number of times you visit a site or search for the same thing on the internet and the duration of your stay on these sites. The second is *content mining*, which refers to the discovery of what information sources, documents, videos, and the like, contain (van Wel & Royakkers, 2004:130; Madria et.al., 1999). This process analyzes what exactly is contained in each source on the internet.

The third is *structure mining*, which reveals the connections between documents and other sources on the web (van Wel & Royakkers, 2004:130; Madria et.al., 1999). *Sources* in this is defined as any document, video, audio file, poem, article, journal, novel, and anything in between. This has an impact on search engines as well, as related content to the topic of the search can be provided as a result.

The search engines we use, such as Google, Yahoo, Bing, and others, rely on content mining specifically for the purposes of accurate search results (Van Wel & Royakkers, 2004:130;

Khorsheed, 2015).[1] Having analyzed the content of these documents, audio files, and videos, search engines are able to produce results that are the most relevant to the search criteria. The most common keywords that are used in each source aid in the refinement of search results. Those who benefit most from search engines are, of course, *researchers*. Here, I will refer to anyone looking for information on the internet as a *researcher* to broaden the scope of the term.

As mentioned previously, people prefer immediate and accurate search engine results (van Wel & Royakkers, 2004:130; Khorsheed et.al., 2015). When key words or phrases are entered, your expectation is that the search results will display relevant sites and sources. If your keyword search includes "social media" "privacy preservation" and "ethical responsibility", you will certainly be unimpressed with search results about puppies, genetically modified food, or information about Pearl Harbor.

The first step in the data mining process, namely data cleaning, will be mentioned below, as explained in both Sahu, Shrma & Gondhalakar (2011) and Rani & Rao (2018).

2.1. Data Cleaning

The first step of the data mining process is data cleaning. Data cleaning refers to an in-depth *scrubbing* of the data that has been collected from various sources. *Scrubbing* refers to the removal of noise from data. *Noise* can be described as any part of the data set that has been altered for a purpose or is simply irrelevant to the purpose of the mining (Sahu, Shmra & Gondhalakar, 2011:114; Rani & Rao, 2018:5841). Thus, I use the term *noise* to describe either an intentional effort to dilute the success of a potential data miner by changing some of the values present in the data, or it can simply be described as the irrelevant information present in the original data set that has no value for the results of the mining process (Sahu, Shmra & Gondhalakar, 2011:114; Rani & Rao, 2018:5841).

Above, noise was conceptualized in terms of either the removal of unnecessary data or the intentional addition of data to obscure potentially harmful information (Sahu, Shmra & Gondhalakar, 2011:114; Rani & Rao, 2018:5841). On the one hand, *noise* can simply refer to unnecessary information that the data miner does not need for their process. This can be in the form of other kinds of information or information that simply isn't specific to the purpose of the mining process. Consider, for example, the intentions of a data miner who is trying to hack into your banking app to gain access to your accounts. Even if this person can grab all your information all he really needs is a couple of numbers to access your money.

On the other hand, *noise* can be added to a data set to prevent someone from mining and misusing the collected data. This is one available method we can use to protect the sensitive

[1] Accuracy in search engine results is an important factor when considering the value of a search engine. If a search engine cannot provide accuracy, its results are likely to be inconsistent, unrelated, or irrelevant (Van Wel & Royakkers, 2004:130; Khorsheed, 2015).

information that we have stored. In essence, certain parts of the data will be altered, switched out, or is missing entirely (Sahu, Shmra & Gondhalakar, 2011:114; Rani et.al., 2018:5841; Gupta & Gupta, 2019). Picture, for a moment, the information on your banking app. There are details that you can freely share with someone and some details that you need to keep to yourself to protect your finances.

Now, data mining is common on any app, even just keeping track of how often you use it, and this does not immediately pose a clear threat. However, a threat emerges when other kinds of private information, such as account numbers, are accessed by unauthorized parties. You want to believe that your bank has systems in place to keep your information safe. As the data has been thoroughly cleaned by the end of this step, it is ready to be integrated into a common source. The following sections, 2.3-2.5, will focus on the work by Sahu, Shmra & Gondhalakar (2011).

2.2. Data Integration

The second step in the data mining process is data integration. Here, I will mention the integration of data from various sources into one common source. This conveys how effective the data mining process is with analyzing large volumes of information, as stated in Sahu, Shmra & Gondhalakar (2011). This step of the process is used to compile all the relevant data, note the use of the term relevant, that has been cleaned in the previous step. There is no longer any noise or irrelevant information which does not suit the purpose or intent of the process. The sources from which this data has been collected and cleaned can take various forms (Sahu, Shmra & Gondhalakar, 2011:114).

Imagine a highly publicized event with many different journalists capturing information for various forms of media. There are cameras recording videos and taking pictures, people writing down information on notepads, laptops, and tablets; journalists talking to people and recording their words for later use. During the data integration stage of the data mining process, we can imagine capturing all this information from all these sources and putting it all together in a single location in preparation for the next step of the process. This step leaves the data miner with data that has been cleaned and collected in a common source.

2.3. Data Transformation

The third step in the data mining process is data transformation (Sahu, Shmra & Gondhalakar, 2011:114). This step in the data mining process occurs after the various forms of the collected data have been stored together in the same location. For the data mining process to continue after the second step, the data must be transformed from its initial, varying forms to a singular or “collective” form (Sahu, Shmra & Gondhalakar, 2011:114). There is a shift away from articles, pictures, vlogs, news reports, and recordings, to a single location where all this information can be stored.

To put this into a visual image, think about the media. Every day, we can go onto any news website and read news from all over the world. In one country, war is breaking out and people are fleeing their homes. In the next country, children are going to school and working on brand new tablets and laptops. And all around us, for just over the last year, the news has been overflowing with different developments in countries around the world regarding the Covid-19 pandemic. No matter in what form this information was originally collected, when it is read about, all of it is in one single form, available to anyone no matter where they are from.

An interesting aspect of data transformation is that, while it may seem harmless, it can pose real dangers. Our lack of knowledge about such can, in fact, easily begin to pose a danger as well. "The Boiling Frog" analogy specifically applies to data transformation and broadly to data mining as a whole. To be brief, there are two ways to boil a frog with differing levels of success. One option is to toss the frog into the water when it is already boiling. However, the outcome of this is that the frog will immediately jump out of the pot to save its own life. The second option is to put the frog in the water while it is cold and heat the pot thereafter. The temperature will increase slowly, alarming the frog only once it is too hot and the damage has been done.

The problem that arises due to the seemingly harmless data mining that occurs behind the scenes every day is similar in nature to the story of the frog and the pot. If a data miner immediately attacks our banks or our medical records, the danger and potential harm is abundantly clear, and we feel forced to act in self-defense. However, if we become comfortable with seemingly minor data mining, we are less likely to be concerned about more dangerous data mining. Like the frog, we will not notice until it is too late to escape or save ourselves. This topic will be given more attention in a later section of this chapter, with a focus on privacy policies.

2.4. Data Mining

Data mining is a fundamental element of the KDD process.[2] As Sahu, Shmra & Gondhalakar (2011:114) explain, there are three techniques which are applied to transformed data to extract useful information and patterns. These three data mining techniques will be briefly mentioned below, namely clustering, association rule learning, and regression techniques. Each of the following data mining techniques can be used to accomplish interesting and far-reaching goals.

Firstly, clustering is a simple technique that takes data items that are related in some way and groups them together. Within the data set, there may be a couple different kinds of data available about a person or group of people (Sahu, Shmra & Gondhalakar, 2011:114). Now, to visualize this, imagine that each person's data is captured in a table. The table contains data about each important aspect of your life, such as medical data or financial history. If this information

[2] It is important to keep in mind that this thesis uses the informal term "data mining" to refer to the process that is more formally known as KDD, as is described at the start of this chapter. In this description of the DM process, the fourth step is known as "data mining", not to be confused with the umbrella term "data mining" that refers to this entire process informally.

from 10 people would be clustered together, one section of the table would contain everyone's medical data, another would contain everyone's financial status, and so forth.

Secondly, association rule mining places its focus on the relationships between variables (Sahu, Shmra & Gondhalakar, 2011:114). This means that each variable can either be influenced by or can influence another variable in a small and meaningless, or large and profound way. In keeping with the example above, these variables could be anything from financial status to work history. These variables have an intricate and complex relationship as each has an effect on the other (Sahu, Shmra & Gondhalakar, 2011:114). For example, our income determines the level of healthcare we can afford, while the level of healthcare we can afford may be the difference between early and late cancer diagnoses, or our educational history may influence the kind of career we take one, while the career we take on influences our income.

Third, regression techniques are used for a special and important purpose, namely the prediction of future trends and behaviors. This technique, interestingly, is one of the oldest techniques used among data miners and it has survived the tests of time (Sahu, Shmra & Gondhalakar, 2011:114). This is essential for a multitude of industries in a variety of fields, and the first to come to mind is, of course, consumer habits and behaviors. However, it certainly depends on a deeper understanding of the variables that association rule mining focuses on. As an example, consider a pharmaceutical company. They might want to predict the rising trends in medication to keep up with demand. If far more people are contracting the flu each year, they simply must produce increasingly of these medications as the need increases.

2.5. Knowledge Representation

Knowledge representation is the last step in the data mining process. The form the knowledge representation takes is not essential to discuss as it is either up to the individual or company which has requested the mining or the data miners themselves (Sahu, Shmra & Gondhalakar, 2011:114). The data, at this point, has been cleaned, integrated, transformed, and mined. It is ready to be represented to highlight the results or outcome of the process.

A good example of this would be the final copy of a written piece of work. Data is collected through research from various sources and is cleaned through extracting only the useful information which will be used in the final product. The data is integrated into a single draft and transformed into your own voice. The rewrite can be considered the fourth step as the information is clustered together into a few different chapters, each dealing with its own core topic. When the final product is ready, it can be seen as the representation of the mined knowledge.

As data mining has been described, the following section of this chapter will focus on the ethical considerations that we need to bear in mind when discussing the potential hazards of data mining. I will briefly expand on a couple of arguments related to data mining, its benefits and hazards, in an effort to refute these arguments as unjustifiable, invalid, unreasonable, or down-right ignorant.

3. Ethical Considerations in Data Mining

As has been made explicit in the introduction, this thesis does not intend to argue for or against any specific views regarding the ethical management of data mining. However, it is my intention to broadly express the foundational ethical concepts that will be explained more fully in the later chapters of this thesis. More specific attention will be paid to ethical concerns and why we need to concern ourselves with these issues that pertain to all levels of data mining. This relates to situations as minor as transaction mining or as major as the development of advanced AI.

3.1. The Responsibility Argument

There are two sides to the argument about the ethical responsibility we bear with regards to data mining. I will call this the Responsibility Argument. On the one hand, this argument states that people bear responsibility for what they choose to do and where they choose to do it. It is their choice to use the internet and by sharing their information on it, it becomes part of the public data (van Wel & Royakkers, 2004:134).

I will concede to this point to a limited extent. Personal responsibility plays a pivotal role in our everyday lives, interactions, and experiences. We all bear several responsibilities on any given day and take ownership of our choices. “As the Internet evolves, its users and uses grow and diversify globally” (Howard et.al., 2002). However, and this is where my disagreement comes in, the problem isn’t this simple. This is not an issue that we can blindly accept and hide behind the claim that personal responsibility is an overarching measure when considering the ethics of this situation.

Our dependence on technology is abundantly evident as we rely on it for every aspect of our day. In the morning, an alarm rings, signaling the start of a new day. The news, social media, and work sites are all accessible through phones, computers, laptops, tablets, and the like. To do our jobs, the internet is likely the core necessity to ensure the smooth operation of a company or business. Students all around the world log onto their university or college sites to access correspondence, assignments, information, and more. Labs, hospitals, and medical centers run more smoothly by using improved technology with every passing decade. The world is fundamentally dependent on technology for traveling, banking, researching, working, gaming, calling, texting, driving, navigating, production, education, and more. As I have established a foundation for our need for technology, I will now focus on why the argument that we are personally responsible for our use of the internet is limited and ultimately fails.

I argue, on the other hand, that it is almost impossible for modern students and workers to go without the use of the internet. The systems we have in place in this world require the use of some form of technology. It is not my intention to say that we couldn’t live a life less dependent on the internet and technology. There are a variety of situations where we don’t necessarily need

to use the internet.[3] However, as students and as workers, we are at a severe disadvantage in various aspects of our lives if we refuse to use the internet. Our teachers, lecturers, managers, coworkers, and whoever else will not be able to contact us with any new assignments, results, schedule changes, or new tasks that need to be completed.

Without the use of technology, practices slow down to an unacceptable speed. While this may have been the case years ago before technology was as big a part of society, the world does not function at the same pace anymore. Students, workers, and others cannot negate the use of the internet without causing grave detriment to their lives. Therefore, while personal responsibility remains important, people should not be subjected to privacy violations simply because they performed normal functions within society.[4]

3.2. The Consent Argument

I have entitled the next argument, the Consent Argument, and it is much like the previous argument about personal responsibility in that people consent to their data being mined when they use the internet. So, like bearing personal responsibility for your own use of the internet, you inexplicably consent to the potential consequences of using the internet (van Wel & Royakkers, 2004:134). The difference between this argument and the previous rests on the main claim of each. While the responsibility argument blames the user for their actions, whether knowingly or not, the consent argument tries to dictate that the user agreed to this kind of exploitation and focuses on the idea that the user knew what would happen and did it anyway. So, while both arguments pertain to our responsibility towards our use of the internet and blame the victim in this scenario, responsibility towards something and consenting to something are different.

As stated in both Van Wel & Royakkers (2004) and Norval and Henderson (2017), as you consent to the potential consequences of using the internet, you cannot be upset when your data is mined by a malicious third-party. “Some have argued that social data posted online are freely available for research” (Van Wel & Royakkers, 2004; Norval and Henderson, 2017).

This, however, isn’t much of an argument if we consider the real-world barriers that prevent people from being “aware” in the first place. If you’re not aware of something, and there are barriers in place that keep you misinformed, ill-informed, and unaware, it simply cannot be claimed that you consented. In other words, it is necessary to acknowledge these barriers and find solutions to address them. Until we have done that, the argument that a person simply consents to these potential privacy violations has no real value. There are some factors need to be considered to determine the extent to which the Responsibility Argument and Consent Argument fail. Below, these considerations are highlighted by Fabian et.al. (2018).

---

[3] Some simple examples may include using a recipe book to cook a meal or dessert, jobs that require manual labor like construction, or watching DVDs or reading books to avoid online entertainment.

[4] Examples may include having to get a bank account to avoid tax violations, having to own a computer for school, or being required to give out personal details to a potential employer.

Readability can be defined as "the ease of understanding or comprehension due to the style of writing" (Fabian et.al., 2018:19). The readability of privacy policies impacts their credibility. It is simple enough to counter that privacy policies and popups about collecting cookies warn people about the possibility of data collection on sites and apps, but this argument is limited.[5]

Against the background of rising privacy concerns, as highlighted by Fabian et.al. (2017:18), "privacy policies seem to represent an influential instrument for increasing customer trust and loyalty." This tells us that people are more likely to trust a website if they have even a mention of a privacy policy. It is important to keep in mind that it cannot be assumed that every privacy policy is written clearly and is understandable to everyone. If people are given information they cannot reasonably interpret correctly, it cannot be argued that they have understood what they are agreeing to, and thus, it cannot be argued that they consented to having their data collected (van Wel & Royakkers, 2004:134; Fabian et.al., 2017:18).

Additionally, privacy policies can be misused or be false. "According to US Legal Dictionary, a *privacy policy* in the online context is related to as "a statement that declares a firm's or website's policy on collecting and releasing information about a visitor" (Fabian et.al., 2017:19). Like ambiguous language in legal contracts that may cause any of the involved parties to be harmed in some way, these privacy policies may convince the reader that their information is safe. To do this, they include ambiguous language and hidden clauses that allow them to mine your data without informing you and without being liable for the collection of data.

A good example of this can be found in the 2016 Facebook Data Scandal where 87 million people had their personal information and private messages stolen through a third-party app (The Great Hack, 2019; Confessore, 2018). By innocently clicking on and completing one of the online quizzes that people have grown used to over the years, the information from their accounts and their entire friend networks' accounts was mined, (The Great Hack, 2019; Confessore, 2018).

Further, it should be stressed that privacy policies don't necessarily hold much value. The user is made to believe that their information will not be mined (van Wel & Royakkers, 2004:135). So, privacy policies may be too difficult to read, they may be written in such a way that the language can easily be misinterpreted and giving a false sense of privacy, and the policies may be a blatant lie. On these three fronts, I claim that the value of a privacy policy, as a tool for protection and informed consent, falls short. Thus, the argument cannot serve to vindicate the idea that people consent to their data being used in this way.

And lastly, the final nail on this coffin for this argument rests on the simple premise that it is utterly unreasonable to expect each user to read and re-read every update to every privacy policy

[5] This part of the argument ignores the reality that people are not necessarily as knowledgeable about these topics, or the terminology used in these privacy policies. This benefits the websites in that they have the freedom to mine your data, under the guise that they do not.

on every site or app that they use. Such vast documents often contain unclear language which is difficult to understand and can easily be used to hide vital information from readers (van Wel & Royakkers, 2004:135). The argument simply cannot be that every person must obtain some advanced degree before using the internet.

I have stated that this argument fails wholeheartedly, and my overarching point stands. We need to protect people, their information, and their overall privacy because the systems we have in place are inefficient. People are exploited daily due to this problem and a blind eye cannot be turned to this pervasive and prevalent issue. Our ethical responsibility towards these people is remarkably simple: protection. This can be understood by looking at a dangerous argument in favor of data mining that poses grave ethical problems.

3.3. The Personalization Argument

I have entitled the next argument the "Personalization Argument", which holds that web-data mining will facilitate and foster personalization. This means that the basic mining conducted on sites really shouldn't bother us because it is to our own benefit. Everybody loves personalized services, and you cannot have personalized service without mining some of your data, such as names, addresses, likes, dislikes, or products you are interested in (van Wel & Royakkers, 2004:136; Yu, 1999). If we look only at the benefits of personalization, however, we ignore the very real risks involved in this process.

A store clerk who knows your name, perhaps a hairdresser who asks about your kid's college choice, or a butcher who remembers that your dogs enjoy marrow bones is far more innocent than web-data mining. To be known by your first name at a store you frequent poses no violation of privacy; after all, to engage with society is to reveal at least something about ourselves, be it only our name. Building customer profiles is often the objective of mining data, where the mining process wants to determine who the customer is and what they do (Cook, 2005; Adomavicius & Tuzhilin, 2001).

To address this argument, I will focus on one overarching point. This argument fails because personalization is not worth the potential violation. In other words, the risks of personalization far outweigh the potential benefits. As stated above, there are innocent situations where revealing our basic information does not pose any risk. However, it has and will continue to impact people disproportionately. As is the case with data mining, "our information is stored, analyzed, collected, and compared to other data" (Kanan & Babu, 2018:163). This data, as initially non-violating or non-threatening as it may seem, can, in fact, be used by malicious actors to reconstruct potentially sensitive data. This information may include names, addresses, ethnicity, race, and the like. To highlight the danger that could be posed by this, I will mention an example of a bank:

> Suppose a particular group of people, Group x, fall into a particular racial category. Group X members have all opened accounts at the same bank in their area. The bank has all their information stored in its systems, such as the place of employment, dependents, residential address, race, gender… This sensitive information could be extracted, which in turn, could allow another group, say Group z, to determine where to coordinate a racially motivated attack.

Without privacy measures in place, this is very well possible. Humanity has proved, repeatedly, that malicious actors will go out of their way to exploit, undermine, and destroy others for their own selfish interests. We see this on the news daily, with hate crimes and racially motivated attacks occurring and reoccurring for the same old reasons as people had 400 years ago and further.

It is the year 2021 and the world is becoming all the more dangerous. The necessity of protection against this type of mining and against these kinds of malicious actors is becoming increasingly essential as incidents of motivated attacks will continue to occur and could increase in frequency, as can be predicted by looking at history, or even by simply reviewing the past few years.

3.4. The Advertisement Argument

Next, I have entitled the argument the "Advertisement Argument" which is in favor of data mining. This is similar to the personalization argument but focuses specifically on the personalization of advertisements. The more data we mine, the better and more accurate our collection of data and extrapolated information (van Wel & Royakkers, 2004:136; Bhatia & Hasija, 2016). As stated by Bhatia and Hasija (2016), this data can be exploited to serve the customers better and offer them advertisements, or so the argument claims.

This is particularly true with targeted advertisements. We don't like ads that pop up simply because we have discussed the topic around our technology or accidentally clicked on a Facebook advertisement. Proponents of this argument will claim that customers feel special and valued when targeted advertisements are presented accurately. To add to this, this argument claims that privacy protectors and data miners should work together to ensure fewer unsolicited advertisements (van Wel & Royakkers, 2004:136). This is odd because the very advertisements we are talking about are unsolicited.

To counter this point, I argue that profiling customers in stores and on websites creates far too many concerns to ignore. Data is mined and customer profiles are created, this is repeated, and added to the growing collection of data as frequently as possible. This is the type of relentless cycle that is never satisfied that it has enough information. Knowing extensive details about each and every customer poses far greater privacy risks than initially considered, as is described above with the example about the bank. As explained by Van Wel & Royakkers (2004:136), while we

may not have a lot to be concerned about initially, the very same data could pose potential privacy violations to the people involved.

There is a variant of this argument that looks at a specific kind of data mining. This argument claims that a specific kind of mining, namely non-personal data mining, does not pose violations and is therefore an irrelevant concern. Non-personal data collection of many people simply assesses crowd behavior and does not target any particular person individually (van Wel & Royakkers, 2004:136). Evaluating purchasing habits of grocery store customers simply gives the store a better idea of which products are most popular. In turn, this allows them to ensure that their customers are always satisfied with the products that are available at that store (van Wel & Royakkers, 2004:136). However, this data can be used in a discriminatory fashion which violates the privacy and potentially the safety of the customers. The personalization argument has been challenged by Van Wel & Royakkers (2004) on the basis of the value of individualism. In support of this, I will describe individualism as a fundamental value of Western culture.

One of the most fundamental values of Western culture is individualism (van Wel & Royakkers, 2004:136). Profiling through web data mining can lead to *de-individualization*. We can define *de-individualization* as the judgment of people and treating them based on their group characteristics. As is expressed by Van Wel & Royakkers (2004:136), using group profiling to make decisions about marketing or creating policies threatens the individualism of each of these people. This is an even greater concern when these profiles somehow become public.

We cannot deny that there are individuals and groups who seek to damage others based on various characteristics such as skin color, gender, age...the list goes on. Think about some of the greatest tragedies the world has seen, such as slavery, Apartheid, the Holocaust, the ethnic cleansing during the Balkan war. Today, while these tragedies are still ongoing in similar forms, a new threat begins to emerge.[6] Too much of our data is available and we know how easily it can be exploited and has been exploited in recent years.

To end, the argument that non-personal data collection does not infringe on a person's right to privacy only holds true in the smallest of scales. Given any circumstance where more than a first name is collected, the information may become personally identifiable long before we notice that anything has happened. The protection of people's data is an ethical responsibility as it centers itself around the value of privacy and safety.

Prior to a description of privacy, I will clarify the importance of, and risks associated with, predictive analysis. To support this claim, the recurring data breaches at Facebook in 2016 and 2021 are appropriate examples.

[6] The tragedies referred to here include, as a brief example, the current and ongoing wars in certain countries in the Middle East and Africa.

4. Predictive Analysis

As mentioned above, Facebook has seen recurring data breaches, the most noteworthy of which occurred before the 2016 Presidential Election in America. In 2016, 87 million Facebook users had their personal account information and private messages mined. This could very well be an influential factor regarding undecided voters during the 2016 Presidential election in America. It was revealed, through the Congressional Hearing that took place two years later in 2018, that Mr. Zuckerberg was aware of this mass violation of privacy and chose to stay silent, keeping those affected in the dark (The Great Hack, 2019).

4.1. Predicting Trends

Unfortunately, the argument that non-personal data, mentioned previously, is irrelevant does not hold up as even general knowledge about a person can give a malicious actor insight into how they can manipulate these people. This is exactly what happened with the Facebook data scandal of 2016. People's thoughts and posts were used to identify and influence the undecided voters (The Great Hack 2019). While profiling crowd behavior in a grocery store may seem inconsequential, profiling in larger industries can easily pose a threat to not only individualism but the safety of the public. On top of this, at this very moment in 2021, Facebook is in the spotlight again for another breach of privacy related to user profiles, affecting more than half a million people (O'Sullivan, 2021).

Predicting trends may pose a threat to individuals, governments, and society as a whole. "Predictive analytics involves several steps through which a data analyst can predict the future based on the current and historical data" (Kumar and Garg, 2018:31). However, as stated in Edwards (2019), it cannot be claimed that predicting trends should be ceased as predicting the use of products and services is a fundamentally important part of society for two specific reasons.

Firstly, Edwards (2019) states that predicting trends with the demand of goods ensures that companies and manufacturers know which goods are in higher demand (Edwards, 2019). In turn, this allows them to produce enough of these goods to meet the demand. Secondly, Edwards (2019) states that it reduces the risk of unnecessary expenditure as the products and goods that are not in much of a demand will be produced to a lesser degree (Edwards, 2019). This is true in every sector, as stores need to know what clothes and food to stock more or less of and pharmaceutical companies need to be able to produce enough medicine when there is an uptick in its use. Edwards (2019) emphasizes:

> Retailers often use predictive models to forecast inventory requirements, manage shipping schedules and configure store layouts to maximize sales. Airlines frequently use predictive analytics to set ticket prices reflecting past travel trends. Hotels, restaurants, and other hospitality industry players can use the technology to forecast the number of guests on any given night in order to maximize occupancy and revenue.

4.2. Predictive Analytics

Predictive analytics is, as explained in both Edwards (2019) and Kumar & Garg (2018), one aspect of analyzing data that specifically deals with predicting trends through the collection and analysis of data. This type of prediction can produce insights about the future in a variety of sectors and industries. Organizations can use both current and past data to predict trends days or years into the future (Edwards, 2019). Edwards (2019) states that the usefulness and effectiveness of predictive analysis is obvious as its market growth is expected to reach $10.95 billion by next year, 2022.

Predictive analysis is useful in various sectors. These include the maintenance of aircrafts, improving vehicle manufacturing plans, determining the impact of the weather on the country's power supply, predicting trends in financial markets, improving the use of raw materials, developing statistics from trends in crime, and more (Edwards, 2019). As highlighted by Kumar and Garg (2018:33), predictive analytics provides opportunities for the growth in the types and volume of data, the development of more user-friendly computers, and an advantage in both business and the economy.

For example, picture a pharmaceutical company attempting to discover trends in medication use. Edwards (2019) states that predictive analysis is a far more reliable and accurate form of prediction than other forms of prediction. During flu season, for example, pharmaceutical companies have the responsibility to produce enough flu shots and enough medication that treats the symptoms of flus and colds. If they are behind on production or if they underprepared the production of medications and shots, the people they are supplying will be fighting it alone. In turn, this causes a loss of trust in these companies and a loss of income for the company itself as people will turn to other producers in their time of need.

Predictive analysis, while ultimately important, can be exploited by malicious actors to perpetuate their desires. "The developments in the field of artificial intelligence and machine learning have changed the world of computation where intelligent computation techniques and algorithms are introduced" (Kumar & Garg, 2018:37). The collection of medical data poses a risk should the data be mined and used to perpetrate an attack on a particular area or group of people. Biological warfare is something that can be achieved even on a small scale, even when there is no clear war going on around us. With the advent of the 1980s, genetic engineering technology and techniques have increased national and international concern for safety and security (Vogel, 2019).

Biodefense refers to, as described by Vogel (2019) defensive strategies and methods that can be taken against the threat of a biological attack, whether it be weapons or biological agents. This connection between Big Data and biodefence is a new topic compared to data mining itself, with the earliest publication about this connection occurring only in 2014, in "National and Trans-

National Security: Implications of Big Data in the Life Sciences". This source will be paid more attention in the next chapter of this thesis on biological data.

Predictive analysis, just as is the case with data mining, has clear benefits while posing grave risks and potential consequences. As has been stated, while we enjoy the benefits of mining data, big data, and predictive analysis, the potential implications for AI are far-reaching and pose enough of a risk to warrant further investigation. The concern here lies in the preservation of privacy within big data and data mining, so that it is possible to reap the rewards of data collection without compromising the safety of private or personally identifiable information.

The following section will focus on Privacy Preserving Data Mining (PPDM) as a means to mitigate privacy and security concerns in the data mining process.

5. Privacy Preservation

5.1. Privacy and Security

Privacy preservation techniques can protect against data and databases from malicious actors. To distort data is to protect its content from being used as a source for personally identifiable information (Sahu, Shmra & Gondhalakar, 2011:114). One way we can distort data has already been described above, namely the addition of noise. Here, the focus is the purpose and fundamentals of privacy and security, privacy concerns, and PPDM. This supports my argument that data mining requires methods of privacy protection to prevent abuse and misuse of private or personally identifiable information.

Further, as described by Qi and Zong (2011:1343), the purpose of privacy preservation is three-fold "In order to protect privacy in released databases, people have proposed a lot of effective data mining technology to hide sensitive information. The purpose of privacy protection is as follows (1) Hide sensitive information contained in the original data; (2) data between hidden and original have the same characteristics (3) get the same data accuracy as original data set" (Qi & Zong, 2011:1341).

This has become necessary as the uptick in data collection renders traditional methods ineffective and therefore requires improved privacy preservation. Qi & Zong (2011:1341) claim that there are two important aspects to consider here, namely protecting personally identifiable information and using the data advantageously. The first, as stated by Qi & Zong (2011:1341), is concerned with information such as home address or ID card numbers and attempts to devise methods that do not allow sensitive or personally identifiable information to be revealed during the process of mining and applying the data. Secondly, Qi & Zong (2011:1341), considers the advantages of mining and applying the data in a way that does not reveal sensitive information (Qi & Zong, 2011:1341). A question that may arise here considers the interchangeable use of the terms 'privacy' and 'security.'

Therefore, it is necessary to determine the difference between "privacy" and "security" as these two terms are often used interchangeably. By looking at Shah and Gulati (2016:40), it is clear to see that both are entirely separate and wholly related issues. This refers to how the concepts exist individually but work together when the protection of peoples and data is concerned. PPDM finds its purpose in ensuring that sensitive data is protected through "vigorous methods." Shah and Gulati (2016:40) claim:

> The three fundamentals of security are Confidentiality, Integrity and Availability [28]. In context of Census data, security can be termed as the facility for controlling person-specific access to information, protecting it from unauthorized disclosure, modification, loss, or destruction of this information. Security can be accomplished through controls based on operational and technical know-how.
>
> In contrast, privacy is specific. It can be termed as a right of an individual to keep his/her personal information from being disclosed. Privacy can be accomplished through policies and procedures. A person's personal information which may lead to his identification may not be disclosed under ethical grounds.

This tells us that security is a broader concern that includes the problems of unauthorized access and disclosure, alteration to the data, or the loss of the information as a whole, while privacy is more specific to the ethical concern of the basic human right to privacy and deals with personally identifiable information. In terms of accomplishing security and privacy, security requires a measure of control over the data itself, while privacy is the ethical question of the duty to keep data safe. As this differentiation has been established, privacy concerns will be the next focus point.

Privacy, as understood above, is a fundamental issue in the development of data mining. It is a part of the fundamental right we have as human beings to keep some details about ourselves private. However, with privacy preservation, we often face the loss of information. This means that, after the data mining process has been completed, some of the PPDM methods will have resulted in some loss of information. "Many methods like attribute removal, anonymization, randomization, aggregation on numeric values is applied on data sets to provide privacy. These methods incur information loss in some situations too" (Shah & Gulati, 2016:41).

This suggests that privacy may result in the loss of some essential information after the data mining process has been completed. It is, in fact, one of the costs of privacy preservation methods, that may very well lead to inaccurate mining results. This does not mean that continuous innovation in this field is a waste of time, on the contrary, it explicitly expresses the ever-increasing need for better methods of privacy preservation.

5.2. Privacy Preserving Data Mining

PPDM has become one of the most prominent issues in data mining (Chahal & Gulia, 2018:3484-85). This has pushed researchers to develop approaches that would protect against information leakage while allowing data mining to occur. The following PPDM techniques and evaluative methods are detailed in both Chalal & Gulia (2018) and Shah & Gulati (2016).

PPDM techniques, as explained by Chalal & Gulia (2018:3485), rely heavily on *Secure Multi Party Communication* (SMPC). *SMPC* is defined as "a computation protocol at the end of which no party involved knows anything else except its own inputs the results, i.e., the view of each party during the execution can be effectively simulated by the input and output of the party." A chain of grocery stores would be an appropriate example (Chalal & Gulia, 2018:3485).

Imagine there are three popular grocery chains who have collected non-personally identifiable consumer habits based only on the purchases in the store, over the period of three months. These three stores can use SMPC to share the results of their data collection processes, sharing only the number of purchases for every product or every product category. For example, the sale of low-fat milk versus the sale of milk as a group, where every brand and type are included in the category. These stores may want to do this to predict trends in purchases in their broader area. SMPC allows these stores to share the results of the data they collected safely, without allowing any potential personally identifiable information to be shared about their customers. Shah & Gulati state that there are a number of techniques that can be used to achieve PPDM.

Anonymization Based Techniques, as mentioned in Shah & Gulati (2016:42), refer to situations where data must be published publicly, in its original form. Encryption and perturbation are not possible in this situation. However, something still needs to be done to protect against the potential disclosure of identity. Anonymization is possible through some methods like suppression, removal of data, permutation, swapping, and more (Shah & Gulati, 2016:42).

Perturbation Based Techniques refer, as stated in Shah & Gulati (2016:42), to the distortion of data prior to the process of data mining. One concept, addressed earlier in this thesis, is relevant here, namely noise. It is possible to use noise to create a perturbation-based solution to the problem of privacy preservation. Noise is added to data sets to hide or disturb some information so that it cannot be discovered.

Fuzzy Algorithms refer, as mentioned by Shah & Gulati (2016:42), to techniques that prevent a great loss of information while achieving anonymization. "The algorithms merge similar records into clusters. Each cluster formed is distinct from other clusters and the records of each cluster are not distinguishable from those of other clusters" (Shah & Gulati, 2016:42). This means that the process of clustering, as mentioned briefly above, is used to perpetuate anonymization, and prevent significant loss of information.

Neural Network Based Techniques refer, as described by Shah & Gulati (2016:42), to the use of a biological basis for the design and production of a computational or mathematical model, forming a 'neural network' of computational ability and power. Shah and Gulati (2016) state that computation occurs in the framework of the structure of neurons in a brain. Neural networks form a critical point in the discussion of AI and this topic will be revisited in length in the following chapters. What remains of PPDM rests on its effectiveness, success, and accuracy (Shah & Gulati, 2016:42).

The evaluation of PPDM algorithms can be mentioned briefly to convey what parameters and criteria are sought after in this process, as stated in Chalal and Gulia (2018:3486). Firstly, Chalal & Gulia (2018:3486) state that the *performance* of the algorithm is an important aspect of its effectiveness and accuracy. The way in which this is measured is by determining the time required by the algorithm to complete the process and preserve the necessary level of privacy. Secondly, Chalal & Gulia (2018:3486) state that the *utility* of the data refers to the amount of information that is lost and its influence on the functionality of the data. This means that the data should measure according to how accurate and effective it is after the PPDM process.

Third, Chalal & Gulia (2018:3486) state that the *uncertainty level* of the algorithm refers to "a measure of uncertainty with which the sensitive information that has been hidden can still be predicted" (Chalal & Gulia, 2018:3486). This suggests that PPDM is measured on its ability to hide data well enough that it cannot reasonably be predicted. Lastly, Chalal & Gulia (2018:3486) state that *resistance* refers to the ability of the algorithm to tolerate other data mining models and algorithms. This means that PPDM is evaluated based on its ability to defend, in simple terms, against other forms of mining that may not be as intent on preserving sensitive information as PPDM (Chalal & Gulia, 2018:3486).

6. Conclusion

This chapter has focused on data mining and the process involved to highlight this fundamental ingredient in improving AI efficacy and why we must be concerned about the quantity and quality of its related materials. Further, ethical considerations that pertain to data mining were mentioned to clarify how quickly or easily ethical compromise can become a reality. Finally, this chapter mentioned a few methods for privacy preservation as examples of how we can manage the protection of data and data privacy.

Data mining acts as the first level of interest in this thesis. Developments in data mining are vital to data collection and protection. One form of data that is regularly mined is that of biological data. This thesis places an emphasis on biological data as it can be exploited to improve the efficiency and effectivity of bioterror attacks. The kind of data included in biological data, as will be mentioned below, can be used to determine localized attacks and the most harmful agent for that area or community. As has been mentioned previously, while this kind of data is not a

requirement for the perpetuation of bioterror, its use can speed up the planning process exponentially.

The following chapter focuses on biological data as one specific kind of data that is mined, occurring in the medical and scientific communities. This data holds the same dual nature as data mining, whereby it can be used to plan and prevent biological attacks. This chapter will connect the topics of biological data and AI as a means to highlight their compatibility both in terror prevention and perpetuation.

Chapter 2
Biological Data, Bioterrorism, & Biodefense

1. Introduction

The previous chapter of this thesis focused on data mining in detail as the foundation for this chapter, where biological data is seen as one kind of the rich field of data. The ethics that pertain to data mining, such as considerations of personal responsibility, privacy, and profiling and discrimination, pertain to biological data as well.

Data miners and scientists alike have an ethical responsibility towards the everyday citizen to ensure that their desire and/or need to mine data does not violate anyone's privacy or right to privacy. Biological data, as one kind of data, requires the same ethical consideration as data mining, with specific focus on medical or clinical data. Without data mining, data mining algorithms, techniques, methodologies, and the like, biological data would be stuck, growing too slowly to facilitate innovation and discovery efficiently.

In this chapter, biological data will be viewed as a necessary medium AI requires to contribute to the perpetuation of biological attacks. It is necessary to stress that the AI itself is not the perpetuator of the attacks, it is simply the mindless minion, tasked with something it must carry out. The same holds true for biological data. These tools are used to plan and execute attacks and simply speed up or evolve the process. The developments and advancements in data mining and storage technologies have allowed the collection of biological data to increase to such an exponential degree that it has led to big biological data. Big data, in the context of biological data, can be defined as the result of incredibly vast and fast data mining technologies and algorithms in the fields of medicine, science, and biology.

The vast and increasing collection of biological data poses ethical issues and considerations due to the concern that this information faces the same dual dilemma as data mining, namely that it can be used for malicious purposes. The real-world implications of using biological data to perpetuate biological attacks is a foremost concern for the scientific and medical communities. Biological data is incredibly important to theoretical, practical, and experimental knowledge from these fields as advancements and developments are made possible at an ever-increasing rate. Biological data, however, also provides the necessary theoretical and practical data AI requires to contribute to bioterrorism. It should be stressed that we consider AI to be a contributor as opposed to the perpetuator of these attacks as it simply does what is required of it through the data that is available to it.

Terror is one of the most fundamental concerns societies and communities around the world face, but we are moving into an era where we face bioterror threats from more than just the average extremist group. As there has been a stark increase in bioterror events since the 1990s, we must concern ourselves with the nature of these attacks. It is vital to consider the dual nature of

biological data for two specific reasons. On the one hand, biological data analysis can aid malicious actors in the perpetuation of a biological attack. On the other hand, biological data analysis can be used by societies and governmental bodies to predict the most likely attack and the type of agent that is likely to be used. In our modern-day society, AI poses incredible safety and security concerns specifically because of its ability to harness data and extrapolate knowledge at an alarming speed. AI does not need to be conscious to aid in bioterrorism, it simply needs the directive to do so from a human being. At its core, AI relies on data mining and its developments just as biological data does.

In this chapter, the first section will focus on biological data and the ethical responsibility of the medical and scientific communities to abide by privacy and confidentiality, informed consent, and other fundamental considerations. The second section will use the nature of biological data to highlight the importance of privacy and ethical decision making in every level of biological data. After this description, the third section will turn to a fundamentally important aspect of this, namely biological databases.

The third section focuses on biological databases to highlight the vast nature of biological data storage requirements. This emphasizes the volume of biological data, and in turn, emphasizes how much of it is easily accessible to AI and malicious actors. The fourth section is a case study on COVID-19, whereby the involvement of AI in the management of pandemics, epidemics, and localized outbreaks, is mentioned.

The fifth section will investigate bioterrorism in detail, attempting to emphasize the real-world threat of biological attacks through AI. As has been mentioned above, it is important to keep in mind that while biological data itself is not a necessity in the perpetuation of terror, it provides benefits and advantages to the bioterrorist. This emphasizes the potential threat as biological data aids in the accuracy, efficiency, and effectivity of bioterror attacks. With enough information, a bioterrorist can determine the kinds of agents that will be most useful in localized areas or cities. There are several considerations pertaining to ethics and privacy that will emphasize the need for responsible research and innovation, or in other words, preventing a problem from occurring by using privacy preservation technologies. This relates to biological data as the protection of biological data from malicious actors is essentially the first round of defense against data hacking or theft.

The last section of this chapter will connect big biological data and biodefence. Should governments or societies be faced with bioterror attacks perpetuated through the use of AI, ethical decision making becomes paramount in terms of the expected response or possible countermeasures. This means that biological data and big biological data, due to their dual nature, are vital to bioterror prediction and prevention. The most likely attacks, areas, and agents can be determined by using the biological data that is available in our databases.

2. Defining Biological Data & The Ethical Responsibility of Researchers

Biological data is a broad term and includes anything from patient data to disease data, genomics to proteomics, and experimental biology. Due to new advancements in this field, such as high throughput sequencing techniques and micro array analysis, copious amounts of biological data are being collected (Achan et al., 2012). *High throughput sequencing technology i*s a technique that "sequences multiple DNA molecules in parallel, enabling hundreds of millions of DNA molecules to be sequenced at once" (Churko et al., 2014). This advancement is a reach from the previously used technology, namely low throughput sequencing or "Sanger sequencing", which was capable of a lot less.

This is a major advantage, as the faster sequencing is accomplished the more data can be collected (Churko et al, 2014). *Micro array analysis* is a modern advancement in the field of cancer research and treatment. This technique allows for the analysis of vast quantities of samples, old or new, and hastens the collection of biological data (Govindarajan et al., 2012). [7]

Achan et.al (2012) states that the collection of biological data does not exist without its difficulties, similar to the difficulties faced in the data mining process. Firstly, Achan et.al. (2012) highlights that the amount of data is an issue in this field because this causes difficulties in warehousing and computation. This tells us that there is so much data that it is a challenge to analyze, store, and send. Secondly, Achan et.al. (2012) states that this kind of data takes various forms and is collected from a variety of sources. This poses problems in terms of the integration of the data, as was mentioned in the data integration portion of the data mining process.

Thirdly, Achan et.al. (2012) explains that the nature of the data is rather dynamic, and this poses a unique challenge as new data and new relationships between previously discovered data are being discovered with great frequency. An example provided by Achan et.al., (2012) is clear and efficient: "For example, a new protein can be discovered. Likewise, a previously unknown interaction can be detected between a pair of proteins present in the database. Thus, the data is dynamic in nature." This description clarifies that the nature of the data has a direct impact on the ability to process and collect it for further use.

Due to the vast amount of biological data that is being collected at an ever-increasing pace, it becomes all the more important to address ethical responsibility of researchers in the scientific and medical fields. The principles below have been distinguished by Weinbaum et.al. (2019) and will be used as necessary background for the ethical concerns that are applicable when these principles are violated.

[7] These speed up advancements in fields such as cancer research, providing researchers, biologists, and doctors around the world with improved understanding of the disease. This improves their ability to develop treatments and medications.

2.1. Duty to Society

Weinbaum et.al. (2019:5-6) state that the first principle is Duty to Society. This can be defined as the responsibility of the researcher and research to contribute to or facilitate the society or community's well-being. Throughout history, differing communities and cultures have a collective desire to protect the well-being of people. "The primary premise of duty to society is that research must not be undertaken if it produces no benefit to society" (Weinbaum, et.al., 2019:5-6). This means that researchers have a responsibility not to their research necessarily, but to their participants.

2.2. Beneficence

The second principle, as described by Weinbaum et.al. (2019:11), is Beneficence, referring to the responsibility of the researcher to consider the participant's welfare as the main aim. This includes ensuring that the risks never overshadow the benefits. This is a cornerstone of research in the scientific and medical communities. Beneficence can be considered to reconcile the difficulty between the need to experiment and perform clinical trials, and providing the highest standard of care (Weinbaum, et.al., 2019:11).

2.3. Conflicts of Interest

Weinbaum et.al. (2019:15) states that the third principle is conflicts of interest. The core premise of this principle is that research simply should not let outside influences, such as financial incentives, influence or bias their research. While there is a difference between a biased researcher and a biased participant, both must be considered as conflicts of interest that alter or impede the true results of the study (Weinbaum, et.al., 2019:15). Just as researchers may act under bias, participants could be biased towards the outcome of the trial. These could include intending to circumvent the trial or keeping any negative outcomes or symptoms to themselves, in fear of being cast from the trial.

2.4. Informed consent

The fourth principle, as Weinbaum et.al. (2019:19) explains, is informed consent. "All research participants must voluntarily agree to participate in research, without pressure from financial gain or other coercion, and their agreement must include an understanding of the research and its risks" (Weinbaum et.al., 2019:19). This means that, should someone become involved in a trial, they must do so fully by their own volition. They should not be approached and incentivized to join a trial, and if they do choose to participate, they must have all the information before starting the trial. Informed consent will be described in more detail in a later section of this chapter.

2.5. Integrity

Weinbaum et.al. (2019:23) states that the fifth principle is integrity, or the adherence of scientific and medical researchers to the values of truthfulness and honesty. This has a rather substantial impact on the quality and accuracy of the results gained from the study or experiment. Researchers may face pressures to alter or boost the results of their studies, which directly violates the principle of integrity (Weinbaum, et.al., 2019:23). Further, doing so alters the results of the trial, potentially placing people's lives in danger.[8]

2.6. Non-discrimination

The sixth principle, as mentioned in Weinbaum et.al. (2019:19), deals with non-discrimination. This relates to an explanation in the first chapter of this thesis, detailing the possibility of discrimination due to mining personal or personally identifiable information. Researchers should not, under any circumstance, reduce or deny the benefits of studies or research to any group for any reason (Weinbaum et.al., 2019:19). This principle intends "to guarantee that human rights are exercised without discrimination of any kind", including gender identity or sexual orientation, family or marital status, age, or disability, social or economic status, politics, language, race, or any other discriminatory reason (Weinbaum et.al., 2019:19).

2.7. Non-exploitation

Weinbaum et.al. (2019:26) states that the seventh principle is similar to the sixth but deals with the other side of the issue. Non-exploitation requires researchers to refrain from taking advantage of or perpetuating the exploitation of participants. This principle intends to prevent researchers from singling out specific communities within to conduct research, where this community will bear the worst of the burdens by participating in the study (Weinbaum, et.al., 2019:26).

3. The Nature of Biological Data & Ethical Considerations

3.1. The Nature of Biological Data

Biological data has been broadly considered above in terms of its wide scope, the advancements that led it to become such a vast collection of data, and the ethical principles on which scientific research must be conducted. Here, I will focus on the nature of biological data in terms of data heterogeneity, data in high volume, data accuracy and consistency, data organization, data sharing, data integration, and data curation and provenance, as stated in Wooley & Lin (2006).

[8] If, for example, researchers hide the fact that their drug causes certain harmful or even fatal side-effects, the participants could face severe consequences, even death. This, in turn, impedes the scientific discovery that could be possible in this field or through this specific trial.

I have selected these characteristics from Wooley & Lin (2006) as they pertain directly to the ethical considerations that will follow. The nature of biological data, as will be seen below, supports its potential use for both the benefit of society and potential bioterror.

To begin with, Wooley & Lin (2006) describe the first important characteristic I will mention here, *data heterogeneity.* This refers to the diversity of biological data. Wooley and Lin (2006) state that the complexity and variety of data forms, formats, types, and the like, pose quite a challenge for modern day biology (Wooley & Lin, 2006). These include, among others, as explained in Wooley & Lin (2006), firstly, sequence data from human and non-human DNA. Secondly, genetic map and pathway data. Thirdly, geometric information or the shape/structure of cells and other biological entities, and fourthly, imagery related to both artificial and natural forms.

The second characteristic, as stated in Wooley & Lin (2006), worth noting is the large volume of biological data as an essential characteristic. I have described a few diverse kinds of biological data above to highlight the stark variety in biological data types. Biological data demands high volumes of data as it details not only one, but multiple levels, aspects, elements, parts, and more, of biological organisms (Wooley & Lin, 2006).

The third relevant characteristic, as described by Wooley & Lin (2006), focuses on data accuracy and consistency as two key factors in biological data. Inconsistency is possible even in protocol-, or instrument-dependent processes. The likelihood and frequency of these inconsistencies increase the more parties are involved in the experimentation (Wooley & Lin, 2006). The fourth characteristic, noted by Wooley and Lin (2006), is the organization of data. This is a pivotal factor in data acquisition, use, and storage as incomprehensible, uninterpretable, or inaccessible data is, truthfully, useless. The organization of data requires databases which serve as research tools and are available to the community (Wooley & Lin, 2006). The classification of databases and examples will be the focus of the next section of this chapter, *biological databases.*

The fifth characteristic, as mentioned by Wooley & Lin (2006), is *data sharing*, the result of a shared consensus in communities of scientists that findings necessarily have to be reproduceable. This means that data must be shared, or it won't be possible for other researchers to test and validate the results of an experiment or trial (Wooley & Lin, 2006). However, the practice of sharing research and findings is not met world-wide. An old joke in the community of life science research: "the data is mine, mine, mine..." This is due, in part, to the profitability of this kind of data for biotechnology, pharmaceuticals, and bioinformatics. Non-profit organizations or universities are not immune to external pressure, as these institutions are often funded by the for-profit industries (Wooley & Lin, 2006).

The sixth noteworthy characteristic, as described by Wooley & Lin (2006), the aspect of data integration form, will be briefly mentioned here. Wooley and Lin (2006) state that as the sharing of data has been established, the focus can shift to integrating the mass collection of data. Digitally, biological data is constituted as "data bits", as mentioned in the previous chapter, which

form part of "data sets", which in turn, form part of databases. Data integration is required at the data bit, data set, and database level (Wooley & Lin, 2006).

The last relevant characteristic, mentioned by Wooley & Lin (2006), is data curation and provenance, or the process of organizing or selecting items and the origin of data. Just like animals and man, data adapts and evolves as time goes by. The process of data curation is essential as both data as a result of analysis, study, or experimentation and raw data are contained within these ever-expanding databases (Wooley & Lin, 2006). There is a likelihood that errors in one database may be propagated to other databases, causing the same and unfamiliar problems. Similarly, updates to data may not occur between databases in a timely manner, resulting in related problems. This clarifies that the curation of data is simply essential to its validity, integrity, and usefulness (Wooley & Lin, 2006). Curation allows data to have a value that endures time.

3.2. Ethical Considerations

Based on the characteristics mentioned above, several ethical considerations pertain to biological data and the use and storage thereof. A few of these have been selected in the literature by Vayena & Madoff (2019), which will be used to briefly describe the ethical concerns and considerations that pertain to not only the collection of biological data but its purpose as well.

3.2.1. Privacy

The first issue, as is stated in Vayena and Madoff (2019), considers privacy to be one of the most challenging ethical issues when discussing big biological data. On the one hand, data sharing needs to occur to increase the utility or usefulness of the data, but it is a concern that access to the data must increase to allow sharing to occur in the first place. This paves the way for malicious actors to gain access to data.

On the other hand, Vayena & Madoff (2019) state that the measures and protocols that have been in place to address traditional accessibility issues cannot necessarily provide protection against modern-day problems. Modern-day actors and technologies can therefore easily gain access to any kind of data they wish to see. Additionally, consent is a massive concern. "The consent model is stretched to the breaking point when what is in prospect is consenting to unknown uses and linkages of data by various unknown users" (Vayena & Madoff, 2019).

3.2.2. Control of Data & Data Sharing

The next ethical concern, as described by Vayena & Madoff (2019), rests on the control of data and data sharing. Time can be taken to address data sharing parameters between institutions and between countries on everyday matters. However, during unprecedented public health emergencies, such as the one we are currently facing with COVID-19, any problems or issues data

sharing perpetuates are brought onto the forefront (Vayena & Madoff, 2019). Some of the barriers faced with regards to data sharing include economic barriers, logistical barriers, ethical factors, legal barriers, and political barriers. These are major influences in the realm of data sharing and can alter or impact data in a magnitude of diverse ways, both positively and negatively (Vayena & Madoff, 2019).

3.2.3. Non-state Actors

The third issue Vayena & Madoff (2019) focus on is a vital consideration pertaining to the ethics of biological big data. This rests on the non-state actors in the public health sphere. To determine an accurate account of the ethics that should be involved in this topic, consideration must be paid to the capabilities and nature of biological agents that could be used (Vayena & Madoff, 2019). Non-state actors include corporations while excluding any government involvement. A wonderful example is described by Vayena & Madoff (2019), "Consider, for example, that a private company running analytics on large data is able to discover—perhaps without even seeking to do so—a pattern of disease related to a behavior, a region threatened by an outbreak, or even a particular individual as a source of an infectious disease outbreak".

3.2.4. Mitigation of Harm

In terms of the fourth issue, Vayena & Madoff (2019) state that the mitigation of harm must be at the forefront of ethical considerations in the field of big biological data. The harms that are included here are discrimination, social exclusion, the infringement of individual or basic rights, stigmatization, damaged social and family relationships, and more. As claimed by Vayena and Madoff (2019), "The risk of these harms is in numerous ways exacerbated because big data approaches to public health are in an early stage of technical development". This means that technology has not advanced enough to protect people from the potential harm that can be perpetuated by biological big data (Vayena & Madoff, 2019).

With a brief mention of the nature of biological data, its complexity and diversity are clearly understood. Due to its complexity, diversity, and its high-volume nature, biological data has led to biological big data, where so much data is collected that it essentially exceeds fathomability. The accuracy, consistency, and organization of big biological data poses complications and difficulties due to its large and complex nature. Data sharing and integration need to be discussed thoroughly to deal with issues such as accessibility in both positive and negative lights. However, this thesis will not go into detail about this matter.

Ethical considerations that are relevant to the nature of biological data range from the simplicity of accessibility to the complexity of ethical data sharing, the prevention of malicious actors, and the mitigation of harm. These ethical issues are just as interrelated as the nature of biological data, each just as important as the next.

The following section will take a look at biological databases as a necessary requirement for the vast volume of biological big data and the ethical governance of data in databases.

4. Biological Databases & The Ethical Governance of HBGRDs

Biological databases are the collection of massive amounts of data and data types. Zou et.al. (2015) details that these databases are developed for a range of different purposes and are curated according to different methods and levels. “As biological data accumulate at larger scales and increase at exponential paces, thanks principally to higher-throughput and lower-cost DNA sequencing technologies, the number of biological databases that have been developed to manage such data deluge is growing at ever-faster rates” (Zou et.al., 2015).

The main objectives of biological data are tenfold, including the storage, sharing, and organization of data in a way that is easily searchable and structured understandably. Further, web application interfaces (APIs) which stem from this increased collection and improved organization. This will provide computers with the ability to integrate and exchange data between a larger variety of databases autonomously (Zou et.al., 2015). Here, the focus is on the scope of data coverage, biocuration, and examples of human-related databases to highlight the vast storage requirements of big biological data. This emphasizes how important the ethical governance of biological databases is, as is stated in Zou et.al. (2015).

4.1. Databases

The scope of biological data is complex, comprehensive, and specialized. Typical examples of biological databases that collect and store a comprehensive variety of data include DNA Data Bank Japan (DDBJ), European Molecular Biology Laboratory (EMBL), and GenBank. Established in 1988, Zou et.al. (2015) state that these databases became known as the International Nucleotide Sequence Database Collaboration with the aim of collecting and analyzing both RNA and DNA sequences.

Based on the data curation level, Zou et.al. (2015) states that biological databases can be categorized into three groups, namely derivative databases, primary databases, or secondary databases. “Primary databases contain raw data as archival repository such as the NCBI Sequence Read Archive (SRA), whereas secondary or derivative databases contain curated information as added value,” (Zou et.al., 2015). In terms of the method of biocuration, the velocity, volume, and variety of biological data continuously requires advancements or improvements to collaborate, integrate, and annotate data correctly, accurately, and efficiently. Biological databases can be categorized into two classes, namely “expert-curated databases” and “community-curated databases” (Zou et.al., 2015).

Next, human-related databases, as described in Zou et.al. (2015), will be mentioned here. “Decoding the human genome bears great significance in, from a theoretical view, unveiling human evolutionary history, and from an application view, exploring personalized medicine against diverse diseases”. For the purposes of this thesis, DNA databases, RNA databases, Protein databases, and disease databases will be touched on lightly to emphasize the scope and far-reaching nature of biological data in databases (Zou et.al., 2015).

Zou et.al. (2015) highlights that DNA databases primary function to profile genetic variations, form associations between genotypes and phenotypes, establish the reference genome, and more. RNA databases encompass a range of RNA data.

Protein databases, as stated by Zou et.al. (2015), “collects universal proteins, identifies families and domains, reconstructs phylogenic trees, and profiles structures of proteins”. Disease data are vital databases in the fight against disease. This is just as true now, as the COVID-19 pandemic continues into its second year. As an example, 14.6% of deaths worldwide were caused by over 200 forms of cancer, as of 2015. This emphasizes that disease data is essential to understanding and developing not only treatments, but preventative measures, for all kinds of disease (Zou et.al., 2015).

Biological databases are incredibly vast and complex, containing all our biological information from our metabolic or brain function to our protein or molecular interactions. This section has, thus far, expressed how vitally important the collection of biological data is. However, just as is the case with data mining, the potential for misuse should be a major concern for scientists, doctors, and the general population worldwide. This concern, of course, stems from the fear of bioterror and the application of AI to perpetuate bioterror on small and big scales. While it is possible to perpetuate bioterror without biological data and without super-AI, access to biological data allows terror organizations to perpetuate far more accurate and efficient attacks. This means that they can use the biological data of, say, a city, to determine which biological agent would be most effective on each community or area.

It is essentially important that databases are protected from malicious actors and therefore we have an ethical obligation to safeguard the information contained within these databases. The focus will now turn to the following guidelines related to the ethical governance of biological databases, as prescribed by the Organization for Economic Co-Operation and Development (OECD) with respect to Human Biobanks and Genetic Research Databases (HBGRDs). The list of principles and best practices mentioned below are prescribed by the OECD.

4.2. Principles & Best Practices: Use of HBGRDs

Firstly, the OECD (n.d.:4) describes some general elements related to the use of HBGRDs as some background information on the subject. These principles and practices will serve as a

foundation for the following section about the governance of databases. The OECD (n.d.) distinguishes the following principles and best practices with regards to the use of HBGRDs.

The OECD (n.d.:4) provides a few examples of these principles are provided: a) fostering research should be the primary objective of HBGRDs; b) ethical principles and legal frameworks should be applied to the establishment, government, management, operation, access, use, and discontinuation of these databases; c) human freedoms and rights should be respected at all times; and d) procedures and policies should be in place "for the procurement, collection, labelling, registration, processing, storage, tracking, retrieval, transfer, use, and destruction of human biological materials, data and/or information" (OECD, n.d.:4).

Next, the OECD (n.d.:4-5) offers a few examples of best practices: a) ensuring the availability of information regarding the databases in terms of uncertainties and risks, as well as the underlying rationale for the existence of the database; b) independent research committees should be responsible for the approval or rejection of "the establishment, governance, management, operation, access to, and use of the HBGRDs and its protocols and processes for research activities" (OECD, n.d.:5); and c) stigmatization or discrimination towards a group, family, or individual, regardless of their contribution to these databases, should be respected and protected through reasonable measures at all times (OECD, n.d.:4-5)

4.3. Principles & Best Practices: Governance of HBGRDs

Secondly, the OECD (n.d.:6) focuses on the principles and best practices related to the governance of biological databases. These principles and practices are inherently related to our ethical considerations when discussing biological data and biological big data. The OECD provides these principles in relation to the governance of HBGRDs.

As claimed by the OECD (n.d.:6), the principles involved in the governance of biological databases are as follows: a) transparency and accountability are of paramount importance; b) the governing structure of HBGRDs should be accessible to the public; c) the well-being and rights of participants in these databases demand the design of a governing structure to prevent an overemphasis on research interests; and d) mechanisms for oversight should be in place at every level of these databases.

The OCED (n.d.:7) states that the best practices involved in the governance of biological data are: a) during the process of reviewing databases, attention must be paid to instances where data was used in a way that differs from what was expected by the participant when giving consent; b) ethics and legality must be of paramount importance when researchers, personnel, and partners carry out their tasks; c) during the oversight process, the individuals should be "drawn from diverse areas of expertise of relevance to the nature and purpose" of these databases; and d) the need for continuous review and modification of policy and procedure as time passes is a fundamental requirement for these kinds of databases (OECD, n.d.:7).

These guidelines, principles, and best practices, in my view, are of vital importance to the ethical government of biological data in databases. The emphasis on the facilitation of research could be quite difficult to maintain when research interests move beyond reasonable respect and care for the participant. Measures such as transparency, accountability, and public accessibility are required to ensure the sanctity of this data is upheld by its users. There is a large emphasis on the rights and freedoms of the participants as well, and I believe this extends to the impact this may have on the general public. The public, for any reason, could easily be impacted by the research that is conducted and the people that participate in this research, in both positive and negative ways.

As with data mining, big data, and biological data, biological databases face the dual dilemma as well. Just as we use them for the advancement of science and in the effort to cure disease, malicious actors could use them against us as well. There has been a controversial and conspiracy-minded debate regarding the COVID-19 virus and its potential origin in a lab as a manufactured disease. While this thesis will not speculate on conspiracy, the principle stands. Currently, society faces an ever-increasing threat from AI and the vast increase of the collection, integration, and storage of biological data in databases, as this is the kind of information necessary to perpetuate small-scale biological attacks.

The following section of this chapter will focus on the COVID-19 virus and the impact or influence of AI on the diagnosis, tracking, treatment, and other aids AI has provided in the fight against COVID-19. This will emphasize how easily AI can be integrated with biological data and stand as the necessary foundation for my argument that the combination of AI and biological data can lead to devastating bioterror events in the near future.

5. COVID-19 & The Use of Artificial Intelligence in a Pandemic

Virologists have been warning us for a long time and we have not listened. I quote Nobel Laureate Joshue Lederberg, “The single biggest threat to man’s continued dominance on the planet is the virus” (Henig, 2020). This has been up for discussion for a long time. In the 1980s, Edward Kilbourne, an expert in infectious disease, spoke at a conference on Long Island about “Genetically altered viruses and the environment” (Henig, 2020).

His premise was that the most contagious, lethal, and most impossible to control virus would rise in the near future. He called this virus the Maximally Malignant Mutant Virus, or MMMV. “As Kilbourne described it, MMMV would have the environmental stability of poliovirus, the high mutation rate of influenza virus, the unrestricted host range of rabies virus, and the long latency of herpes virus. It would be transmitted through the air and replicate in the lower respiratory tract, like influenza, and it would insert its own genes directly into the host’s molecules, like HIV (Human Immunodeficiency Virus)” (Henig, 2020).

It is true that COVID-19 does not exactly bear all the qualities of MMMV, but the similarities are striking. In December of 2019, the news emerged that a new virus had been discovered in the Hubei province, Wuhan, China (Henig, 2020; Ayukekbong et.al., 2020:2). We have all heard enough speculation about which animal may be responsible for the virus, and we have certainly heard more than enough conspiracy theories about government involvement in the creation of the virus itself. At the time, the potential of destruction should have been more evident to everyone, but it is well over a year later and millions of people around the world show little to no care for the harm that has been caused and will continue. As mentioned previously, where the virus came from is not the concern here. The COVID-19 pandemic is a good example to show what the world could experience if the use of AI and its related fields is perfected to determine specific and localized attacks or even develop specific agents or diseases for use based on our data. Instead of studying this research to determine how to help and heal people, AI could easily assess the same data to determine the best way to harm society.

This has, partly, forced the use of AI as a decent portion of the worldwide population has contributed to a potentially massive increase in the spread of the virus.[9] Arora et.al. (2020) provides two ways in which AI can be used with regards to COVID-19. AI can be used to develop early warning systems for disease outbreaks and predicting the virus' spread (Arora et.al., 2020).

AI has played and continues to play a crucial role in the prediction, tracking, diagnosis, and reducing the burden on healthcare workers. Further, AI plays a role in protein-structure prediction and the development of vaccines (Arora et.al., 2020). With the application of biological data in AI, AI is all the more capable of being effective in the roles mentioned. Without this access to biological data, AI would not be nearly as useful as it would not have the knowledge it needs to be efficient and effective. These uses of AI will be described below in brief alongside a brief description of the ethical issue that pertains to each use.

5.1. Prediction, Tracking, & Privacy:

AI has the ability to predict and track the spread of viruses. To do this, the artificially intelligent system is able to figure out where clusters of disease begin to form and how they travel. In the same way, AI can locate areas in cities or cities themselves where certain diseases are more prominent. This will help determine the kinds of biological agents that will be most effective in these areas. Information or data about the spread of disease is essential to the effective and efficient dispatching of treatment, aid allocation, and more (Arora et.al., 2020; Chang, 2020). A few different COVID-19 tracking apps, hereafter referred to as "CTAs", have been developed in an effort to curb the spread of the virus (Klar & Lanzerath, 2020:1).

While this immediately seems like an intensely positive innovation, and is, without a doubt, abundantly necessary in a time like this, there are ethical privacy concerns. This kind of technology

[9] This refers to people's ignorance of preventative measures to curb the spread of COVID-19. While many people were staying safe, social distancing, wearing masks, others refused or ignored these measures.

is cause for concern (Klar & Lanzerath, 2020:1). Further, as Klar and Lanzerath (2020:3) state, it is quite possible to hack CTAs and exploit the data contained within. "Moving beyond the usage for COVID-19 containment via 'surveillance creep', governments might exploit the crisis to establish and retain tracking data from citizens which could, in theory, be used in other contexts". And of course, this invasion of privacy could easily extend to malicious actors outside the government.

5.2. Contact Tracing & Surveillance:

Through AI, governments and healthcare facilities can conduct contact tracing. The AI system can "augment mobile health applications where smart devices like watches, mobile phones, cameras, and an entire range of wearable devices can be employed" (Arora et.al., 2020; Chang, 2020). Abuhammad et.al. (2020) states that there are a multitude of ethical concerns regarding contact tracing and the surveillance that it involves, namely a) privacy, in that people are being tracked, which leads us to the next issue, b) voluntariness, where the voluntariness of the participant must be taken into account, c) beneficence, the concern for who benefits from the data, and d) incentives, referring to the influence of incentives on citizens (Abuhammad et.al., 2020).

5.3. AI in Healthcare: Informed Consent, Safety, & Privacy

AI can aid in the early diagnosis of COVID-19, the development of vaccines, and reduce the burden on healthcare workers. This includes speeding up the diagnostic process for faster results on tests, contributing to the study of protein structure-prediction for the development of vaccines, and aiding medical workers in taking care of their patients (Arora et.al., 2020; Chang, 2020). It is important to keep in mind that AI can only do this effectively and efficiently if it has large-scale access to biological data. Gerke et.al. mentions several ethical concerns related to the use of AI in healthcare. These include, firstly, informed consent applied to the use of AI in "imaging, diagnostics, and surgery will transform the patient-client relationship". Secondly, safety, which relates to the concern that AI learns from synthetic cases and can therefore recommend incorrect treatments. Thirdly, transparency, serves as the foundation for patient confidence. Lastly, biases relate to the potential for bias to be inferred by AI for a variety of different reasons (Gerke et.al., 2020).

While it is true that AI still has a way to go, it has proven itself to be of use during this global pandemic. I argue that the opposite is also true, as with data mining, as with biological data. AI can and will likely be used in the near future to perpetuate small-scale and eventually large-scale biological attacks, aimed at harming society. It must be kept in mind that access to biological data and the continued advancements and developments in AI are not crucial requirements, without which these attacks cannot occur. However, these technologies and advancements continue to make it easier for malicious actors to plan effective and efficient biological attacks. Now, it is not to say that this will be the doing of a conscious, fear-inspiring, cruel AI superbeing. Humans have

always had a hand in the greatest tragedies to befall humanity, and the potential for bioterror through AI is a massive concern.

The following section of this chapter will turn to bioterrorism and historical acts of bioterror to emphasize the historically proven nature of some human beings to perpetuate violent and horrific acts of bioterror. This will emphasize that it is simply a matter of time before the use of AI is co-opted by those with a modern taste for bioterrorism.

6. Bioterrorism & Ethical Considerations in Preparation

The definition of bioterrorism worldwide will differ from person to person. For the purposes of this thesis, I will define *bioterrorism* as "the threat or use of biological agents by individuals or groups motivated by political, religious, ecological, other ideological perspectives". In this definition, I will include the intention to simply create terror and chaos (Carus, 1998:3). While some may include the intent to threaten or intimidate governments or societies as a whole, this seems to ignore other significant potential motivations for acts of terror, such as the nature of biological attacks to affect masses or achieving a specific goal unrelated to swaying, intimidating, or threatening governments (Carus, 1998:3; Grundman, 2014).

Carus (1998:10) provides the trends in bioterror from the start of the 1900s to the end of the 1990s to illustrate how stark the increase in terror attacks had become. This clearly describes the increase of concern for these attacks. They are becoming more prominent as technology develops and advances. The automated AI processes currently available to society are sufficient for the purposes of planning a bioterror attack as this automation reduces the burden on the terrorists. This automated AI is much faster and more efficient at determining which population or city to target and what agents or variety of agents are most appropriate for these specific areas. "We are increasingly seeing terrorist groups looking into the feasibility and effectiveness of chemical, biological, and radiological weapons" (Carus, 1998:10). Between the years of 1900-1960, a scope of 60 years, no more than three and often only one terror attack was perpetuated in a ten-year period. However, during the 1990s, a dramatic and terrifying increase in the number of biological attacks occurred, totaling 153 terror attacks in a ten-year period (Carus, 1998:11; Grundman, 2014).

Famous examples of bioterror attacks include a few interesting and terrifying occurrences. One of the most prominent examples of the severe potential aftermath of biological attacks occurred in the 1300s (Riedel, 2019). During the Siege of Caffa, Tartar forces began to fall ill as an epidemic spread among them. These forces catapulted the bodies of the plague victims over the wall to infect the citizens. This was the start of the Black Death, which killed 70-200 million people (Riedel, 2019).

Interesting examples include the use of leprosy blood in wine in 1495, whereby the Spanish sold infected wine to their French enemies (Riedel, 2019). In 1710, Russian troops used the same

method as during the Siege of Caffa, where they catapulted infected bodies over the walls of a Swedish city to infect the inhabitants. In 1863, during the Civil War in America, clothing from smallpox and yellow fever patients were sold by the Confederates to the U.S. troops (Riedel, 2019).

On the other side of this discussion, one prominent example stands out as we have not considered the testing of these biological agents. It is claimed that Japan tested a variety of biological agents on an estimated 10 000 prisoners of war during WWII. These included anthrax, plague, cholera, meningococcal infection, and more (Riedel, 2019).

When we consider biological attacks over the past few decades, a few come to mind. In 1984, the Rajneesh cult infected salad bars in Dallas, Oregon with salmonella, causing over 700 people to become ill. In 1995, the cult Aum Shinrikyo, launched the Sarin gas attack on the metro in Tokyo. One of the most well-known biological attacks is the Anthrax attacks that occurred in 2001, in the U.S. only a short time after 9/11. This affected 22 people in total (Barras & Greub, 2014).

This section will briefly mention a few examples of biological agents and their use in a potential bioterror attack. Among these are Anthrax and the plague, classified as "category A agents" by the CDC (Central Disciplinary Committee), which require considerable risk management and planning. There are three categories of biological agents, according to the CDC, namely A, B, and C, as described by Grundman (2014):

- Category A: Anthrax, botulism, smallpox, the plague, and viral hemorrhagic fevers like Ebola, etc. These are easily dispersed and highly deadly.
- Category B: Brucellosis, Epsilon Toxin, Glanders, Melioidosis, Ricin Toxin, Typhus fever, and more. These agents "result in moderate morbidity rates and low mortality rates."
- Category C: Emerging infectious diseases, Hantavirus, and Nipah virus. These are easily acquired and dispersed and can be fatal.

Let's imagine for a moment that a small-scale bioterror organization is planning an attack on one particular community in a city. They retrieve patient data from the hospitals and clinics in the area and find that the population in this area suffers from respiratory issues. They are able to do this quickly and efficiently because they are not relying on themselves to analyze the medical data that has been collected in even a small area or community. They have automated AI systems to do this for them, saving time and producing far more accurate and efficient results. S As Anthrax shows flu-like symptoms, and as this community has respiratory issues already, it could be a viable choice of biological agent on behalf of the terrorist. The goal will be to hurt as many people as possible, and the longer the anthrax inhalation goes unnoticed, the more people will fall into the lethal trap of anthrax.

Now, Evans & Inglesby (2019) claim that a distinction can be made between responses to biological attacks and the responses to pandemics or accidents, in the sense of ethics. While both

situations are governed norms instructing the use of state power in defense of such events and the emphasis on public health, bioterror is distinguished as an act of war (Evans & Inglesby, 2019).

There are two ethical frameworks, as is described by Evans & Inglesby (2019), from which to prepare for the potential of a biological attack, namely monistic, or the promotion of one value alone (like liberty or utility), and pluralistic, namely "conceiving of the ethics of public health as involving a trade-off among a set of plausible moral values and principles", meaning that a range of values will be considered in lieu of just one (Evans & Inglesby, 2019). One important consideration for any of these views is that of security. Evans and Inglesby (2019) state that security encompasses fairness, liberty, utility, and protection against situations which may threaten these values. Security can be seen as both a feeling of safety and the practice of safety.

Here, I will focus on the preparation for the potential of a biological attack with attention paid to the work of both Evans & Inglesby (2019) and IMF (2007). This was touched on briefly in earlier sections of this chapter. Essentially, the concern here is that while some research may be conducted in the name of disease prevention, the very same research could result in the release of a deadly pathogen.

Dual-use research, or research that poses a dual-dilemma, rests on the conflict mentioned earlier, namely the liberty of scientists versus the well-being of the public. In preparation for the possibility that research, first intended as a saving grace, could contribute to a biological attack, ethical decision making should be based on a variety of ethical principles. These include justice and fairness, accountability and transparency, and respect for individual liberties (Institute of Medicine (US) Forum on Microbial Threats [IMF], 2007).

To begin with, the IMF (2007) states that ethical decision-making in preparation for a pandemic rest on a few important traits. If these traits can be perpetuated, preparedness and planning will be improved (IMF, 2007). The section below will focus on the IMF (2007) perspective on the relevant traits involved in the process of ethical decision-making during a pandemic, namely justice and fairness, accountability and transparency, and individual liberties.

6.1. Influencing Ethical Decision-Making in a Pandemic: The Relevant Traits

The IMF (2007) provides their view that collaborative and considerate discussion is required in the effort to perpetuate ethical decision-making. Responsibility is placed on decision makers to not only be accountable for any ethical issues, but to be aware of these potential issues before the fact. This refers to a requirement of planning and awareness. These traits rest on a few ideas and ideals that need to be considered in these emergent situations (IMF, 2007).

First, in these cases, the IMF (2007) states that justice and fairness are of concern. Justice and fairness, here, refers to the identification of potential burdens resulting from responses to bioterror and who they may impact. If a response does necessitate a burden on society, that it is as

fair as possible among these people. For example, if one potential response to an attack is to lock down an area or another is to return the attack as an act of war, consideration needs to be given to the potential impact of these actions. In times of bioterror and war, consideration needs to be given to how the population will be affected and if there will be a disproportionate impact (IMF, 2007).

Secondly, the IMF (2007) explains that in terms of accountability and transparency, emergency planning and responses to attacks need to be guided by clear ethical frameworks. This avoids the abuse of power or decisions that impact the well-being of the nation severely. While the people in power and the decision-makers are held accountable for their actions through evaluations and assessments, transparency is required to keep the citizens informed, even in times of uncertainty (IMF, 2007).

Third, as described by the IMF (2007), respect must be given to individual liberties, even in times of emergency planning. Therefore, "Liberty-limiting and social-distancing interventions should be based on the best available evidence." However, the IMF (2007) states that any such measures will have an influence on the personal freedoms people tend to hold most dear, such as the a) quarantining of exposed peoples, b) isolation of infected peoples, c) restricting the use of or closing of public spaces, public events, and schools, and d) limitation of travel in and around the country (IMF, 2007).

As a focus has been placed on the emergence of bioterrorism, its exponential increase since the early 1990s is concerning and evident. This served as the basis for a brief explanation of biological agents and an example of how easily they could be used to perpetuate an attack in modern-day society. Next, ethical considerations and principles in emergency planning highlights the values and guidelines that need to be considered and respected during the preparation for bioterror events. The following section closes off the chapter with an explanation of biodefence and the ethics pertaining to the response to such events.

7. Biodefence & Big Biological Data

This section will focus on big data and biodefence to follow from the previous section detailing the reason why biodefence should be a priority (Vogel, 2019). Modern technologies are one of the driving forces behind the increase in bioterror attacks, such as cellphones, social media platforms, geospatial technologies, advanced analytics, sensors, and more. These tools make data collection easier and more accessible (Vogel, 2019).

Vogel (2019) describes that, while there have been obvious advancements on the side of bioterrorists, these same technological advancements can be used for biodefence. "Harnessing digital technologies more effectively to better identify and assess emerging biosecurity threats could lead to improved detection, response, and preparedness measures to eliminate or at least mitigate a bioweapons attack" (Vogel, 2019). Big data is the result of a mass collection of biological data and this data can be used to understand potential threats and develop better

biodefence protocols and measures. It is worthwhile to note that big data will include a range of different data types, including community, public, and confidential data FN (Vogel, 2014). Big data has four main characteristics, as described by Vogel (2019) commonly known as the "4 Vs". These include volume, variety, velocity, and veracity.

Next, I will define *biodefence* as any measure taken to prevent or protect against any kind of biological attack that has the potential to occur (Vogel, 2019). In terms of the U.S., there are two pillars for biodefence, namely awareness of the threat and the surveillance and detection of threats. As stated by Vogel (2019), awareness of threats includes accurate and timely information relevant to bioterror prevention and includes any assessment required to determine the current or potential future patterns that may emerge. To do this, big data is required for its volume, its variety, its velocity, and its veracity.

With updated techniques and technologies, big data can now offer even more in terms of tallying, organizing, and reorganizing, and storing data that would be most useful and most relevant to the fights against bioterror (Vogel, 2019). While it is true that big data is an important and efficient means to benefit society in a positive and effective way, it can be used to exploit, manipulate, or cause harm to politics, society, the economy, or the health of the public. This issue of potential advantages and liabilities can be referred to as the "Dual Use Dilemma" (AAAS-FBI-UNICRI:35)[10].

As is true with any Dual Use Dilemma, big data offers solutions to these scenarios as well. I will address them briefly to emphasize the aspect of dual use and the need for prevention, protection, and preservation of biological data and Big Data. The following section will briefly detail technical, institutional, and individual solutions, as prescribed by the joint publication developed by the American Association for the Advancement of Science (AAAS), the Federal Bureau of Investigation (FBI), and United Nations Interregional Crime and Justice Research Institute (UNICRI).

7.1. Technical, Institutional, & Individual Solutions

Technical Solutions, as described by AAAS-FBI-UNICRI (n.d.:54), prevent access through unauthorized methods, where data encryption and access controls are the primary methods. There are two types of access controls, one in which the user is identified through passwords or access codes, and one which prevents this access by blocking specific people or IP addresses (AAAS-FBI-UNICRI, n.d.:54).

Institutional solutions, as explained by AAAS-FBI-UNICRI (n.d.:57), attempt to prevent or mitigate data, biological security, or cyber risks. These institutions may refer to research

[10] American Association for the Advancement of Science (AAAS), the Federal Bureau of Investigation (FBI), and United Nations Interregional Crime and Justice Research Institute (UNICRI).

institutes, healthcare facilities, and private industry. The measures provided to mitigate these risks include training, education, partnerships with law enforcement, threat assessment task forces, physical security, and oversight and review of activities pertaining to science (AAAS-FBI-UNICRI, n.d.:57).

Individual solutions, as stated in AAAS-FBI-UNICRI (n.d.:58), involve the risks facing the individual, within the context of big data, designing harmful biological weapons and access to the data by malicious actors. While we can address the risk of access as is described above, the design of harmful biological agents is a far more pressing concern to attend to. In order to do this, risk assessment needs to be conducted by individual scientists to determine how harmful their research could be if it ends up in the wrong hands (AAAS-FBI-UNICRI, n.d.:58).

7.2. The Ethical Challenges in Biodefense

Now, biodefence measures which make use of big biological data face several ethical challenges that must be addressed. “There is a great need, however, to address ethical challenges in biodefence,” (Loike & Fischbach, 2013:1). Here, I will consider ethical challenges, as described by Loike & Fischbach (2013), related to a) resource allocation, b) the clinical testing of new treatments, c) the prevention of unauthorized access to labs, and d) the aspect of dual use.

In the first place, Loike and Fischbach (2013:2) explain that the allocation of resources includes the allocation of personnel. This ethical challenge rests on the risk healthcare workers working during a biological attack face. On top of the risk to their lives, these people face additional challenges such as a lack of education in the response to such a situation, the lack of long-term childcare for workers with children, or placing the health of their family first. The allocation of resources such as medicine or equipment poses a great challenge during the peak of a biological attack’s aftermath. The question arises as to whom should receive the last of the available equipment or the priority with medication dispersion when there are 50 critical condition patients who all need those things to survive (Loike & Fischbach, 2013:3).

In the second place, Loike and Fischbach (2013:2) state that the clinical testing of new treatments poses an ethical challenge in terms of biodefence. One of the prominent points in this relates to clinical trials performed in adults. Should these very same trials be conducted in children and infants when the side effects could be potentially harmful, in preparation for the chance of a biological attack? To expand on this, we could face two choices. The first refers to preliminarily conducting potentially harmful clinical trials on young children. The second refers to finding ourselves in the midst of a terror attack with no known treatment against the biological agent for our children. Here, we are faced with an ethical dilemma regarding the protection of our kids. There is no clear answer here, as stated by Loike & Fischbach (2013:3).

In the third place, Loike and Fischbach (2013:2) emphasize that preventing unauthorized access is a great ethical challenge. Measures can be put in place to prevent the hacking of data or

to prevent physical access to the building. However, the concern remains that people could get around these measures and access data they could very well use to harm small or large parts of society. This is not the only concern. “Even more frightening is the possibility that individuals employed by a biosafety laboratory could use their access to pathogens to set up bioterrorist activities” (Loike & Fischbach, 2013:3). As is clear, our ethical responsibility to protect this kind of dual-use data could be subverted by malicious actors. Due to this, ethical biodefence requires consideration of both these factors so that necessary and sufficient countermeasures are already in place at the advent of an attack.

In the fourth place, Loike and Fischbach (2013:2) state that the dual-use nature of this kind of data poses a clear ethical challenge when we compare the benefits and harms of potential or existing research (Loike & Fischbach, 2013:1). While we can benefit from this greatly, will we benefit from it enough to justify the potential misuse of the same data to perpetuate biological attacks? Here, the nature of biodefence and its necessity in the world we live in today has been highlighted. Before defending in any way against a bioterror attack, the ethical considerations at play in such decisions must be discussed thoroughly by the authorities involved.

8. Conclusion

This chapter has considered biological big data and the ethical considerations in doing research and the ethical questions that are raised due to the nature of biological data. The ethical use and governance relating to the storage of this data in databases are of importance. This is the case due to the possibility of theft and misuse of the data to develop bioterror plans. These databases should be closely governed and monitored to ensure there is no unauthorized access and subsequent misuse of the data to perpetuate harm. Further, the ethical concerns raised when using AI to predict, track, contact trace, and assist medical workers during a pandemic are pivotal. At every level, in each aspect, and in every question about biological data, questions of ethics arise.

As the connection between biological data, big biological data, and bioterrorism has been clarified, its foundational impact on AI-enabled bioterror becomes clear as well. Just as the standard development of AI depends on data mining, biological data, and bioinformatics, AI-enabled bioterror is dependent on these fields as well as these fields provide the data and advancements it requires to function appropriately and complete required tasks. To complete this picture, it must be tied to AI. This will be focused on in the final chapter of this thesis. For now, my attention will be turned to the following chapter. Chapter 3, Bioinformatics, will focus on the methods through which this data is gained and their use in the creation of AI, while paying attention to the ethical considerations surrounding machine learning, deep learning, and neural networks.

Chapter 3
Bioinformatics: Ethical Concerns and Challenges

1. Introduction

This Chapter will focus on *bioinformatics* as the use of computers to form a broad understanding of biological data, as the necessary precursor to the following chapter, AI. Bioinformatics provides a unique and required opportunity for the development of bioterror schemes using AI. Thus far, this thesis has focused on data mining as the central to the usefulness of biological data, and this chapter will provide the same meaningful connection between biological data and bioinformatics.

Without bioinformatics, AI could not make use of biological knowledge to perpetuate bioterror attacks. Unfortunately, as with any scientific or technological advancement, data mining increased to include all different branches of human life, including biological data. Our drive to study the human body and cure it from disease and disorder has led to biological big data, which not only fostered the need for but facilitated the development of technologies capable of dealing with this amount of data. The existence of bioinformatics is wholly dependent on the mass influx of biological data and the developments in its field can be, at least partially, attributed to the need for more efficient techniques, algorithms, and systems. AI will require all of this to function as a means to perpetuate bioterror.

Therefore, the first section will clarify the relationship between biological data and bioinformatics as a base for understanding of some core challenges of bioinformatics. Next, the second section will focus on data warehousing and four common errors involved in the warehousing of data. Alongside this, there are some trends in data warehousing that lead to ethical concern. The third section will turn to machine learning to understand how it functions as a precursor to the development of AI. This will highlight both ethical machine learning, in terms of data protection and a commitment to fairness, and ethical considerations in machine learning.

The fourth section of this chapter will focus on deep learning as an advancement of machine learning and a necessary part of AI development. Further, I will mention the ethical concerns involved in the development of ethical algorithms for deep learning. The last section of this chapter will mention neural networks as a manifestation of our development in computing technologies. Included here is the use of neural networks to foster the development of artificially intelligent systems. In conjunction with this, attention will be paid to ethical considerations and challenges of Artificial Neural Networks (ANNs).

This chapter serves as the last foundation for the closing chapter that follows AI, where the focus will turn to the ethical consideration in the development of AI and the concern for the use of AI in bioterrorism.

2. Biological Data and Bioinformatics: Ethical Concerns

Here, I will turn to the relationship between biological data and bioinformatics to highlight its aims before turning to some ethical concerns. Previously, this thesis has defined biological data as "a collection of data that pertains to the field of medicine. This includes patient files, hospital records, research, and more. Biological data is the result of data mining from medical institutions around the world." *Bioinformatics* is seen as the necessary computational techniques that are required to analyze, interpret, and organize biological knowledge and data. Here, I include the term *biological knowledge* to denote the knowledge gained through the analysis and interpretation of medical data as well as the direct or explicit information in the data.

Biological data is flooding in from all over the world at an increasing rate, as can be seen by the doubling of data every 15 months at GenBank, SWISS-PROT, and the like (Luscombe et.al. 2001). Keeping in mind that this was, at this moment, 20 years ago, it is valuable to consider the more recent growth of these databases. Since the year 2000, the data contained in these databases has increased with exponential speed (Tibebu, et.al., 2020). Matt Thompson, during an interview for Caltech Magazine, states "To give an analogy, let's say you throw a ball up in the air. If you know parameters like the ball's mass, the acceleration due to gravity, its initial velocity, and so on, you can accurately predict where the ball will land after a certain amount of time. We want to predict how biological systems will evolve, but it's difficult because there are so many parameters" (Dajose, 2017).

To begin with, bioinformatics has several aims, as prescribed by Luscombe, et.al. (2001). These include, firstly, as stated in Luscombe et.al. (2001), the organization of data in a way that provides ease of access to researchers who are either submitting latest information to the databases or retrieving data necessary to continue their research. Secondly, Luscombe et.al. (2001) explains that bioinformatics aims to analyze data and develop new tools and techniques with this to conduct its analysis. The development of the resources that are required to perpetuate the process of analysis requires not only a deep grasp of biology, but that of computational theory as well (Luscombe, et.al., 2001). Third, Luscombe et.al. (2001) states that bioinformatics aims to obtain and provide useful and meaningful information from the data available. It uses tools and techniques that are available and were developed specifically for the purpose of analysis (Luscombe, et.al., 2001).

Bioinformatics raises several issues for research and researchers. Goodman (2001) distinguished between three specific factors that pertain to the ethical and social issues involved in bioinformatics. These include a) accuracy and error, b) the appropriateness of data, and c) privacy and confidentiality. Each of these will be briefly mentioned below.

2.1. Accuracy & Error

Firstly, Goodman (2001) describes accuracy and error and states that the avoidance of these things sets off what we can call an ethical red flag. Errors, in this context, could have real impacts on people, hence, this is not only a concern of accuracy but an ethical concern as well. “If there are emerging or established standards for database management, for instance, then a system that relies on a database will be more or less useful, reliable, and safe, depending upon whether or not the database is appropriately maintained, tested, augmented, and so on” (Goodman, 2001). Related to the accuracy of data warehousing and bioinformatics systems are a few more issues, such as the risk to people. While there is a vast and wonderful benefit to warehousing and systems, the risk to patients due to errors in computer-aided innovation and discovery involves a concern for public health and individual well-being (Goodman, 2001).

2.2. Appropriateness of Data

Secondly, Goodman (2001) states that the appropriateness of the use of data and the users of data raises some interesting ethical concerns. Goodman (2001) states that this largely concerns the informed consent of patients when their genetic data is used beyond their own diagnosis and care. For example, imagine a doctor who uses the genetic data from his past and current patients to form a predictive system that aids his diagnosis of these conditions. The question to ask is this: does the patient know and consent to this use of their genetic data? Further, consider the instance that this kind of data is collected by some third party for the purposes of adjusting the eligibility of medical insurance cover. It would be difficult to claim, in this instance, that there is valid and informed consent on behalf of the patients (Goodman, 2001).

2.3. Privacy & Confidentiality

Lastly, as described by Goodman (2001), privacy and confidentiality always play a vital role in the medical and scientific fields. Privacy and confidentiality can be threatened by certain actions, such as the transition and maintenance of genetic data through computers. “These threats include bias and discrimination, personal stigma, psychological stress, and tensions within families, among other risks” (Goodman, 2001). It is easy to see why this is a concern as personal freedoms, rights, and relationships can all be influenced by computerized biases and discrimination. It is therefore imperative that we honor the value of individual rights by respecting and protecting privacy and confidentiality in not just the medical and scientific fields, but throughout society and across the world (Goodman, 2001).

3. Data Warehousing & Ethical Challenges

Clearly, due to the sheer amount of biological data, storage is a concern. I will detail the topic of data warehousing to touch on the various kinds of biological databases, the problem of storage, the quality of data, and the kinds of errors that might be present in this kind of database

(Koh & Brusic, 2005). *A Data warehouse* can be defined as the point of storage for data, developed out of necessity due to a phenomenon that emerged in the 1990s. Data warehousing was originally used in the business domain during the era of information explosion as data mining and collection increased exponentially. Biological data warehousing, on the other hand, is quite different. This is simply due to its vastly complex and incredibly diverse nature, as well as the dual nature of this data, that can lead to the harm of people.

As stated in Koh and Brusic (2005), there are only a few examples of biological data warehouses, namely warehouses that contain gene expression data, genomic data, and others. Koh & Brusic provide two common features amongst these kinds of warehouses, namely a) the extraction of data and its integration from a variety of sources, and b) data analysis through data mining algorithms and techniques. Further, data warehouses used in bioinformatics require a few components to be able to operate. These include a) obtaining the data, b) cleaning the data, c) dataset manipulation, and d) independent and joint analysis tools (Koh & Brusic, 2005).

The quality of data in databases, as described by Koh & Brusic (2005), is of fundamental importance. When we refer to the quality of the data, we refer to its validity, accuracy, efficiency, and more. Through some processes such as data cleaning, the quality of data can be improved. While this process often occurs manually, there are some proprietary programs to aid in the cleaning of data where human beings cannot contribute any more. Human ability may influence the quality of the data, requiring some concern (Koh & Brusic, 2005).

A question may arise as to how we assess the quality of data. Data is evaluated based on an agreement between what has been learned from the data and how well this relates to the real world (Koh & Brusic, 2005). While it is not quite possible to obtain perfect data, it is necessary for a biological database to attempt the highest quality of data that is possible. It is essential that data reach the highest attainable quality to promote the accuracy of the results. In turn, data that is noisy or contains errors of any kind will have a reduced accuracy within the results or may lead to results that are entirely inaccurate (Koh & Brusic, 2005). Data cleaning was mentioned in the first chapter of this thesis in more detail.

3.1. Four Typical Errors

Koh & Brusic (2005) distinguish between four errors typically involved in bioinformatics, mainly due to the sheer volume of diversity of the data. Firstly, the attribute level errors refer to individual values that contain errors. This could include anything from a typo to a misplaced value, and therefore, occur most often among the four classes of errors. Secondly, Koh & Brusic (2019) state that the next level is the record-level error. “The record-level errors often result from conflicts between or misplacement of fields within a record” (Koh & Brusic, 2005). This could include features such as comments, links that are not functional, or the incorrect version or accession numbers.

Thirdly, Koh & Brusic (2005) mention that single-source level errors in databases encompass duplicate or conflicting data items in one database. “Examples include exact or fragmentary duplicates having identical content across separate records” (Koh & Brusic, 2005). Lastly, Koh & Brusic (2005) state that multi-source errors in data bases “occur because of imperfect data integration and source synchronization problems.” The issue many of the public databases are facing is a lack of structural descriptions and a lack of functionality, which directly impedes the potential knowledge that can be gained and the accuracy of the result (Koh & Brusic, 2005).

Beyond these errors, data warehouses face some security concerns due to the large volume and the diverse nature of data (Saleem et.al., n.d.:15). Due to this, there are a number of security approaches to curb or prevent the malicious use of this data by actors with nefarious intentions for groups of peoples around the world. The issues that we face here, as described by Saleem et.al. (n.d.), include confidentiality and integrity.

3.2. Confidentiality & Integrity

In terms of confidentiality, Saleem et.al. (n.d.:15) state that the emphasis is on protecting information from unauthorized access by malicious users with the use of access-control measures. These measures have the aim of controlling both the administration and invocation of the database and include auditing and authentication mechanisms. The reality is that the data contained in warehouses all over the world includes enough information about individuals so that the information can be exploited. Without legitimate access-control measures in place, the confidentiality of this data becomes moot and malicious actors gain the freedom to access private data that can be used for bioterrorism.

In terms of integrity, Saleem et.al. (n.d.:15) explains that the focus is on preventing the alteration of data either accidentally or maliciously. This act involves the insertion of data that is false or contaminated, and therefore leaves the data in a lesser state of integrity. To deal with this, Saleem et.al. (2005) emphasize that there are a few different approaches, namely a) restriction-based techniques, b) access-control techniques, c) inference-control techniques, and more. When the integrity of data is compromised, whether accidentally or on purpose, the results are damning. False data can lead to false “discoveries”, and malicious actors can exploit this if they have access to genetic data warehouses, creating falsities about disease progression, treatment, encouraging the use of certain medications for conditions that will be negatively affected by it, and so much more (Saleem, et.al., n.d.:15).

The following section will focus on machine learning in terms of ethical machine learning and machine learning ethics.

4. Machine Learning & Biological Data: Ethical Algorithms

This section will pay attention to machine learning in its three main forms to describe its use and usefulness in several domains of biological data. The term *machine learning* was coined by Arthur Samuel in 1959 and defined as a "field of study that gives computers the ability to learn without being explicitly programmed" (Deo, 2015). This tells us that machine learning is one of the foundations of AI, as one of the most important characteristics of AI is the ability to "think" for itself.

Deo (2015) and Chakraborty et.al. (2017:1) claim that the tasks machine learning attempts and accomplished are of three distinct kinds, namely supervised learning, unsupervised learning, and reinforcement learning. During an interview for Caltech Magazine with Lori Dajose, Lior Pachter states, "Machine learning is the process of using computational tools to predict and learn from data. These tools can be used in a variety of ways, from combing through telescope data to find planets outside our solar system to teaching a computer how to recognize moving objects in order to drive a car" (Dajose, 2017). Below, three kinds of learning will be mentioned briefly, as described by Chakraborty et.al. (2017) and Deo (2015).

4.1. Supervised, Unsupervised, & Reinforcement Learning

Supervised learning is as it sounds, as described by Chakraborty et.al. (2017:1), where the computer is given examples of inputs and what is required of the output by its algorithm. It is a closed-loop feedback system which uses the comparison of the desired output upon input and the actual output at the end of the process to adjust the parameters of the network (Chakraborty, et.al., 2017:1). The goal, in this case, as explained by Deo (2015), is identifying "a general rule that maps inputs to outputs". Any errors that occur are determined by the differences between these values, which requires the repetition of the learning process. This repetition will continue until the process either resolves the error or fails (Deo, 2015).

For example, picture a baking competition where the novice bakers are expected to recreate a picture-perfect version of a celebrity chef's cake. They have the input data, namely the recipe, and they have the expected output, namely the physical example of the cake that is provided. They are required to redo the recipe until they have created an accurate duplicate in taste and look. After round one, one baker used too little salt and must correct this. Another baker found that her oven is a little different from the one used by the chef and adjusts the temperature accordingly during the second round. This goes on and on until the cake is replicated exactly.

Unsupervised learning, as described by Deo (2015), gives and expects more freedom on behalf of the computer itself. The algorithm is expected to work without any previously programmed inputs or expected outputs and identify the structure of the input and the expectation of the output. The output is not dictated in this instance of learning, but parameters can be set, where a training program will teach the algorithm what it is supposed to be looking for (Deo,

2015). To give an example, let's consider a university that has to collect, organize and analyze bursary applications from undergraduate students, in order to determine who the most eligible candidates are. The use of an algorithm to sort through the potential thousands of applications can ease the burden on university staff members.

Each department and the university as a whole will have specific requirements or criteria that need to be met for the applicant to be considered eligible. During the first round, the applications are analyzed to determine the ineligible candidates and remove them from consideration. The algorithm, in this round, may focus specifically on excluding the candidates that don't meet the initial criteria. During the second round, which may consider any range of other criteria to determine which candidates should be awarded the bursaries. This may be about grades, consistency in academic performance, family lineage, finances, or other criteria. During the third round, the most eligible candidates have been selected and the final round determines which candidates are the actual recipients of the bursaries.

Reinforcement learning, as described by Deo (2015), refers to the expectation that the computer will interact with its environment in a dynamic way to accomplish a particular aim. It is required to be able to do this without the aid of a specific algorithm telling it how to function or whether it has reached its prescribed goal (Deo, 2015). With this method, the training model is essentially used to identify the rules of classification or the patterns that are generally present in the data. After the training program, the algorithm essentially follows the same process on one data set, analyzing the set by testing the possible hypotheticals to determine the best one. This hypothesis is then applied to other sets of data, without the necessity of running all the hypotheses again on each data set (Chakraborty et.al., 2017:1).

For example, imagine an autonomous racing car. The car can only be programmed with so many hypothetical situations, it can only know how to respond to so many different situations from its input data. However, there are so many variables with varying degrees of effect on the environment, such as the amount of rain, the conditions of the tar, the durability of the tires, and much, much more. This form of learning is not perfect and poses some concerns. Briefly, picture, for a moment, a modern-day trolley problem. If the car is unable to stop in time and will hit either a grandmother or a small baby, we cannot predict its choice.

Daniel Van Valen, in an interview with Caltech Magazine, shares "You can give a computer some example data sets and teach it how to look for insights. Then, once it has "learned," you can give it a totally new data set to analyze. It's a kind of artificial intelligence, and it has broad applications" (Dajose, 2017). As stated by Chicco (2017):

> In recent years, biological data has been made available to scientists all over the world as not only the data sets expanded, but the online services expanded as well. Machine learning has spread to computational biology due to its ability to "make predictions on them through accurate statistical models" and its "ability to handle large datasets." A machine learning

algorithm is a computational method based on statistics, implemented in software, able to discover hidden non-obvious patterns in a dataset, and moreover to make reliable statistical predictions about similar new data.

This ability to identify patterns automatically is greatly important as the amount of biological data is far too large for any human being to handle (Chicco, 2017). Machine learning, just as with data mining, biological data, and bioinformatics, faces challenges and problems that impede its ability to do its job as effectively and efficiently as possible. As evolution moves forward in machine learning, its infrastructure, the research it has inspired, and its algorithms, managing machine learning models becomes increasingly difficult (Schelter et.al., 2018).

However, machine learning does not exist outside of the clear ethical concerns related to data protection and privacy, the development of fair algorithms, and the issue of transparency. Below, Vayena et.al. (2018) distinguished between a few ethical considerations in machine learning.

4.2. Ethical Considerations in Machine Learning

To begin with, as described by Vayena et.al. (2018), machine leaning algorithms in the field of medicine require ethical attention and concern to be paid to each level or stage of analysis or processing the data is put through. These algorithms are subject to privacy and confidentiality requirements as the use and reuse of medical data requires informed and valid consent. Vayena et.al. (2018) state that it is difficult to determine where authorization has been granted by participants. This poses challenges for using this data in training algorithms. This emphasizes that even with the use of data to train algorithms, the issues of informed consent, protection of data, and privacy arise.

Secondly, Vayena et.al. (2018) states that the development of these algorithms should be committed to fairness. As is true for most of the world, biases are a part of life, even when unconscious. It is entirely possible for poor representative data used to train algorithms to teach the algorithms biases. The issue here is twofold, as Vayena et.al. (2018) claim, namely a) when the data sources “do not reflect true epidemiology within a given demographic”, and b) when the algorithm is trained based on data that focuses predominantly on one type of person, say white, middle-aged men. Within the first case, the “epidemiology” refers to the field of medicine that studies the determinants involved in health-related issues, as well as the distribution or pattern of disease, focused on specific populations (Vayena et.al., 2018).

Vayena et.al. (2018) states that a fair algorithm takes into account that some populations have been over-diagnosed with certain diseases. Within the second case, it is easy to see how an algorithm that learns wholly about one group of people cannot make accurate assumptions about an entirely different group of people based on the same data (Vayena et.al., 2018). For example,

should the data be based wholly on African American women, the data is valid, but might not apply as a study on white, elderly men.

Here, I have focused on machine learning to highlight its function through supervised learning, unsupervised learning, and reinforcement learning. I have stated that machine learning has its own ethical concerns that exist based on its development and application. The following section will turn to Deep Learning and the ethical development of algorithms.

5. Deep Learning: The Ethics of Algorithms

This section will turn to deep learning as a type of machine learning that utilizes unsupervised learning, supervised learning, or both. Ching et.al. (2018) emphasize that impressive results have been observed with the use of deep learning algorithms. Due to this, and the vast, complex, and ill-understood nature of biological and medical data, deep learning is perfectly suited to solve the most complicated problems in these fields.

Deep learning uses artificial neural networks, or essentially an artificial network crafted in the image of the human brain, to conduct complex and multi-layer learning. The way a deep learning neural network functions is remarkably similar to the way that human being's function. We learn from experience, from both new and repeated actions and events (Marr, 2018). Deep learning algorithms contribute to a variety of everyday human services. These include virtual assistants (Siri, Alexa, Cortana), translation services (Google Translate), chatbots and service bots, facial recognition, driverless delivery vehicles or autonomous cars, personalized shopping, and entertainment, and most importantly to this thesis, medicine, and pharmaceuticals (Marr, 2018).

However, as with most topics that have been considered in this thesis, the dual dilemma rears its head. Deep learning can just as easily be used to contribute to the harm of individuals and societies. Here, I will focus an ethical perspective on the development of algorithms and the key problems and solutions associated with increasingly intelligent computational learning. To emphasize the need for ethical algorithms in the development of deep learning, and by extension AI, Tsamados et.al. (2021) states that “Algorithms have become a key element underpinning crucial services and infrastructures of information societies”. Together with this, not only do individuals rely on algorithms, but schools, hospitals, governments, courts, and financial institutions rely on algorithms as well. The fabric of society, today, depends on algorithms.

To begin with, there are three epistemic concerns involved in the ethical development of algorithms. Below, I will briefly mention three epistemic concerns and three normative concerns, as have been distinguished by Tsamados et.al. (2021).

5.1. Epistemic Concerns: Inconclusive, Misguided, or Inscrutable Evidence

In terms of epistemic concerns, Tsamados et.al. (2021) states that inconclusive evidence may lead to actions that are unjustified, where the algorithm produces and output "in probabilistic terms". This contributes to some serious ethical concerns, namely the distraction of the algorithm from the actual problem, thereby being unable to stop or solve the problem, and the uncertainty, incompleteness, and time-sensitive nature of algorithmic data insights (Tsamados, et.al., 2021). While it is argued that this can be overcome by simply entering or uploading enough data for the algorithm to learn from, "recent research rejects this view." One explicit example of this pertains to the ethical risk associated with algorithmic systems that perpetuate structural inequalities such as racism. Data that contributes to these systemic and structural inequalities is rarely corrected and continues to have an effect even today (Tsamados, et.al., 2021). Below, three kinds of evidence will be considered, as described by Tsamados et.al. (2021).

Inscrutable evidence, as described by Tsamados et.al. (2021), may lead to opacity, meaning that some evidence may be so complex that it is impossible to understand, and that this will lead to unintelligible results. This, in turn, leads to reduced scrutiny, accountability, and trustworthiness (Tsamados, et.al., 2021). Now, while transparency in itself is not necessarily considered and ethical principle in itself, it is "a pro-ethical condition for enabling or impairing other ethical practices or principles" (Tsamados, et.al., 2021). Therefore, transparency is a concern because it can either support or hinder other major ethical concerns (Tsamados, et.al., 2021).

Misguided evidence, as explained by Tsamados et.al. (2021), may lead to bias, requiring a focus on the underlying bias of the developer of the algorithm and how these biases can be translated from them to the algorithm. While some scholars may argue that "algorithmic formalist" or a complete adherence to expected forms and rules should be the dominant belief in the development of algorithms, others agree that abstractions are limited by the bias that may exist in the developer. Therefore, as stated by Tsamados, et.al. (2021), there are 5 traps that an algorithm may fall into due to a lack of social context in data, namely a) the failure to include social criteria in system modelling, b) the failure to use appropriate solutions for each individual situation, c) the failure of ignoring the inclusion of the full meaning of social concepts (such as fairness), d) the failure to realize that "the insertion of technology into an existing social system" has an immediate and lasting evolutionary effect on the system, and e) the failure to consider a non-technological solution to a problem (Tsamados, et.al., 2021).

5.2. Normative Concerns: Unfair Outcomes & Transformative Effects

In terms of normative concerns, unfair outcomes may lead to discrimination. Tsamados, et.al. (2021) states that algorithmic fairness is a common goal to mitigate discrimination. This speaks to the ethical concern that discrimination, through the algorithms we all generally require to function in modern-day society, can be perpetuated more powerfully than ever before.

Algorithmic decisions have an incredible impact on people's lives, days, families, and more, and therefore, attention needs to be paid to the fact that algorithms are not capable of offering the same unique human skills that we are able to offer (Tsamados, et.al., 2021).

Transformative effects, as described by Tsamados et.al. (2021), may lead to autonomy and privacy related challenges. Due to the significant impact on algorithms on the lives of people across the world, the autonomy of users has come into question. The autonomy of users can be hindered by the fact that "Algorithm-based services are increasingly featured "within an ecosystem of complex, socio-technical issues"" (Tsamados, et.al., 2021). This refers to a) the ability of algorithms to shape the user's decisions, which is a major ethical concern in the development of AI, and b), the difficulty understanding algorithms, or the information provided by algorithms which hinder the ability of users to make decision. Linked directly to this is that of privacy. The more people interact with algorithms, as we do on a daily basis, the less control they have over access to their information and how it is being used (Tsamados, et.al., 2021).

Deep learning has been mentioned in brief as a form of machine learning and a necessary development preceding the evolution of artificially intelligent systems. While it is a valuable tool and a necessary precursor in the development of AI, algorithms themselves, on which computation learning is wholly dependent, pose ethical challenges and risks. The following and concluding section in this chapter will focus on an important part of the function of deep learning, namely artificial neural networks.

6. Neural Networks: Ethical Development & Use of ANNs

Here, I will mention artificial neural networks as the final part of this chapter to highlight the function of deep learning and how it pertains to AI. Artificial Neural Networks, hereafter generally referred to as ANNs, are essentially the modelling of a computational system based in the neuronal structures in the human brain (Zayegh & Al Bassam, 2018:1). Our neurons all work together to produce our thoughts, our learning, our cognition, and more, and these make decisions for our bodies based on internal data and input data. "The similarity between artificial neural networks and the human brain is that both acquire the skills in processing data and finding solutions through training (Zayegh & Al Bassam, 2018:1).

The human brain is an incredibly complex system of billions upon billions of neurons that are working together continuously to allow your body to perform all the required functions to keep you alive and well. Information is exchanged through the synapses and this information determines the actions and functions of your body. Zayegh & Al Bassam (2018:2) state "In the same way, artificial neural networks consist of simple computing units or "artificial neurons," and each unit is connected to the other units via weight connectors; then, these units calculate the weighted sum of the coming inputs...." The artificial neurons respond to and send information in the same way that our neurons do.

6.1. ANNs & Healthcare

In the interest of healthcare, Shahid 2019 states that there are a few critical ethical and ethically related challenges and concerns that pertain to the use of AI and ANNs in healthcare. As healthcare advances and a greater focus is on a patient-centered model of healthcare, new and complex challenges will be faced. This will require advanced and accelerated ANNs, as described by Shahid (2019).

ANNs, as explained by Shahid (2019), offer a vast variety of uses in the healthcare field, namely a) speech recognition, b) clinical diagnosis, such as the diagnosis of myocardial infarction, c) cancer prediction, d) prediction of the length of the hospital stay, e) the development of drugs, f) the management of the facility, g) the prediction of costs or the like, h) time management, i) resource management, j) and other cost-effective solutions (Shaid, 2019). The main applications of ANNs include diagnosis, prediction, and classification. Two primary areas for concern arise at this point, as described by Shahid (2019), namely consent to share patient data with ANNs and both the user's competence and familiarity with the operation and function of the system. These will be briefly mentioned below.

6.2. Sharing Patient Data & Potential Dependence on ANNs

In terms of the first concern, highlighted by Shahid (2019), the sharing of patient information is restricted by legislation, such as the Health Insurance Portability and Accountability Act (HIPPA), without the informed consent of the patient. The use of ANNs by healthcare facilities and providers poses a serious ethical concern in the sharing of patient data that is inevitably required to gain knowledge and new discoveries through these systems. Without the sharing of these kinds of data, ANNs would be left impeded or useless altogether (Shahid, 2019).

In terms of the second concern, Shahid (2019) claims that it isn't difficult to imagine that healthcare staff can become too dependent on ANNs to aid in the diagnostic process. They may rely too heavily on these systems for diagnoses. With the incorrect diagnosis or treatment, due to the lack of personal deliberation of the data as a responsibility of healthcare professionals, patients' conditions may go untreated. They may face harmful side-effects of the incorrect medications or even find themselves in a terminal state, where proper medical intervention could have saved their very lives (Shahid, 2019).

The use of ANNs in the field of healthcare poses a growing concern for the use of AI in the diagnosis, treatment, prevention, and prediction of diseases and health conditions. Some of these concerns rest on the ethical principles of beneficence, non-maleficence, justice, autonomy, privacy, and confidentiality. These concerns will be focused on in the following and last chapter of this thesis: Artificial Intelligence: Ethical Considerations for the prevention of and response to Bioterrorism.

7. Conclusion

This chapter has focused on bioinformatics as a precursor to the development of AI, which will be the topic of importance in the following and last chapter of this thesis. Bioinformatics can be considered as a technology, at least in part, developed out of the necessity to provide appropriate and effective technology and tools to manage the influx of biological data.

Machine learning, deep learning, and artificial neural networks are highlighted as fundamental tools in bioinformatics and as specifically required developments in the evolution of AI technologies. These methods of learning and computational abilities are useful and efficient not only in bioinformatics but are core parts of developing greater computational intelligence. Within bioinformatics, and including machine and deep learning and ANNs, certain ethical concerns were highlighted, such as the concern of unfair algorithms in machine learning, epistemic and normative concerns in deep learning, and concerns related to the sharing of patient data when using ANNs. These concerns emphasize that each level of AI development and its related technologies must be carefully considered to determine areas for ethical concern and develop methods and techniques to subvert these ethical concerns as far as possible.

The next chapter, Artificial Intelligence & The Ethical Considerations in the Prevention of and Response to Bioterrorism, will refer to several ethical concerns specifically present in the use of AI. These ethical concerns will serve as a foundation for the recommendations to follow at the end of the chapter, each pertaining to a specific area for ethical concern in the preparation for and response to bioterror through AI-enabled technologies.

Chapter 4
AI: Ethical Considerations in the Prevention of and Response to Bioterrorism

1. Introduction

This thesis began with a detailed overview of data mining. Data mining is an essential part of the foundation of Artificial Intelligence (AI), and without it, AI would not be possible. Data mining includes the mining of all kinds of data from all over the world, from millions of different people, thousands of companies, and fields of study, and more. The fundamental point is that in protecting us from disastrous AI systems, the protection must begin at the most fundamental level of the process, namely the collection and mining of data.

The second chapter in this thesis focused on an overview of biological data and its most fundamental ethical consideration for the very same reason, to emphasize how the protection of data must begin at lower levels of the process. Following this, this chapter turned to the nature of biological data and the ethical considerations that need to be given due to the volume, complexity, and diverse nature of biological data. The ease with which AI was adapted to aiding a pandemic emphasizes how easily AI could be used to perpetuate a pandemic.

The third chapter of this thesis detailed the necessary steps in the evolution of data technologies as the necessary precursor to the development of intelligent systems, namely bioinformatics. This chapter paid attention to the nature of bioinformatics and focused on ethical considerations that pertain to the data itself. Next, this chapter paid attention to machine learning as a developed and developing field in the creation of AI. Following this, deep learning is mentioned as the next step in the evolution of machine learning. However, the development of ethical algorithms plays an influential role in the ethical nature of the deep learning system, and therefore, ethical concerns arise.

This chapter will focus wholly on AI and the ethical concerns that pertain to not only AI, but AI in healthcare and AI in bioterrorism. Each of these concerns will be mentioned below to highlight what these concerns are, why they are concerning to us, and how they relate to the use of AI in the medical and scientific fields. As with each of the other chapters and various sections throughout this thesis, the “dual dilemma” is present here. Therefore, we require a larger understanding about each ethical consideration and its potential relation to bioterror. In each of the sections to follow, ethical concern will be described and followed up with its relevance to the field of healthcare and the perpetuation of bioterror.

In the first section, AI will be defined and contextualized with respect to its potential involvement in bioterror through its use in the medical industry. In the second section, the focus is on privacy and surveillance through AI and the ethical challenges involved. In the third section,

the manipulation of behavior through AI is considered due to the effects of manipulation. In the fourth section, recognition will be given to the potential opacity of artificial systems, highlighting the intelligible parts of AI. In the fifth section, focus will be on the potentiality of bias in AI systems designed to make decisions and the effects of this on society.

In the sixth section, the focus will turn to the aspects of autonomy and responsibility with regards to AI systems and how these aspects are irrelevant to the topic at hand. In the seventh section, attention will be paid to COVID-19 and its exposure to our lack of preparedness to fight pandemics. In the eighth section, threat assessments of bioterror will be mentioned on the backdrop of AI to highlight the considerations involved in these assessments. In the ninth section, the topic of national security and AI will be the focus, illustrating the need for AI as well as the concern for AI. The last section will turn to the way forward and recommendations as to how we can prepare, predict, and prevent bioterror incidents.

2. Defining AI: AI in Healthcare and Bioterrorism

Here, I will briefly mention AI to clarify its nature and vast capabilities, its relevance and use in the field of healthcare, and the potential contribution of AI to aid and perpetuate bioterrorism. AI is the development of intelligent machines that are capable of learning both from their training algorithms and the external environment, having human-like or greater than human intelligence, and are able to, to a certain degree, work, interact, and react like human beings (Habeeb, 2017; Paschen et.al., 2019; Saleh, 2019:2).

The development and implementation of intelligent machines is necessitated by the modern and technologically driven world that we live in. The only way for intelligent machines to learn as much as they need to about the external world, to yield the greatest benefits to fields and societies alike, is to have enough data about said external world (Habeeb, 2017). Habeeb (2017) states that "artificial intelligence must have access to objects, categories, properties, and relations between all of them to implement knowledge engineering, initiating common sense, reasoning, and problem-solving power...".

For the purposes of this thesis, it is necessary to distinguish between narrow or weak AI and strong AI. Isaksson (2018:12) states "Narrow or weak artificial intelligence is designed to do a specific narrow task. The narrow/weak AI works towards a predefined goal and is limited by its algorithms in performing its task(s)." This is crucial to my argument with regards to weak AI, whereby it is all a bioterrorist needs to plan and perpetuate bioterror attacks. Strong AI that has not been developed yet is not the concern, although will pose incredible concerns with its birth in the future. This, as has been stated, is not the focus here as it is not a reality now. Weak AI is a reality now and all that is needed to plan and perpetuate the most accurate, the most efficient, and

the most effective bioterror attacks. In terms of strong AI, the distinction is essential that this form of AI is almost, if not entirely, indistinguishable from a human (Isaksson, 2018:12).[11]

In clarifying this differentiation, Isaksson (2019:17-18) quotes John Searle in his 1980s publication, "Minds, Brains, and Programs" (Searle, 1980):

> According to weak AI, the principal value of the computer in the study of the mind is that it gives us an immensely powerful tool. For example, it enables us to formulate and test hypotheses in a more rigorous and precise fashion. But according to strong AI, the computer is not merely a tool in the study of the mind; rather, the appropriately programmed computer really is a mind, in the sense that computers given the right programs can be literally said to understand and have other cognitive states. In strong AI, because the programmed computer has cognitive states, the programs are not mere tools that enable us to test psychological explanations; rather, the programs are themselves the explanations.'

Famously, Searle argues against the possibility of strong AI, stating that AI cannot develop "intentionality" and can therefore never be comparable to human intellect. Searle (1980:24) states, "beliefs, desires, and intentions are intentional states; undirected forms of anxiety and depression are not." He created two propositions that imply three consequences with regards to AI and intentionality.

These propositions, as described by Searle (1980:25), include a) intentionality is a causal feature, specifically found in humans and animals, and b) the representation of a computer program does not imply intentionality. He draws three results or consequences from this. Firstly, the mind and a computer do not share the process of producing intentionality. Secondly, if a mechanism were to produce intentionality, it would need to be equal to the brain in terms of its causal qualities. Thirdly, the creation of intentionality in computers is dependent on a lot more than programming.

Isaksson (2018:17-8) states that a variety of other theorists, including Harry Collins, Hubert L. Dreyfus, and Nills J. Nilsson, all argue against the potentiality of strong AI. However, this is not a focal point of this thesis as strong AI is not needed for the perpetuation of bioterrorism. For the purposes of this thesis, as has been stated previously, an AI superbeing isn't the foremost concern at the moment, considering the real and current threat of bioterror and the multitude of barriers that prevent us from developing super intelligent AI (Isaksson, 2018:17-18).

[11] Discussion about strong AI often focus on the possibility of *The Singularity*. The supposed time at which computers will surpass human-level intelligence (See for instance Ray Kurzweil, "The Singularity is Near: When humans transcend biology.") The notion of the Singularity will not be discussed in this thesis as the focus here is on much less advanced AI systems.

2.1. Applications of AI in the Healthcare Industry

AI, specifically weak AI as we have it now, is used in a variety of different fields and offers an abundance of different uses, including speech recognition, facial recognition, detecting fraudulent use of credit cards, the treatment and diagnosis of disease, targeted ads, the list goes on (Habeeb, 2017; Paschen et.al., 2019). This is due to the various important traits that AI is supposed to possess. These include, as described by Saleh (2019:4), a) the ability to adapt to environments and predict patterns using algorithms, b) the ability to continuously learn from both its algorithms and the external environment, c) the characteristic of thinking forward, in other words, generating insights that allow for better decision-making in the future, d) and the ability to make its own decisions. Together, these traits establish a functional artificially intelligent system, which raises the question about the use of such systems in the field of healthcare (Saleh, 2019:4).

Focusing more directly on AI in healthcare, Sharma (2021) distinguished between several AI applications in healthcare, three of which will be mentioned briefly below. Firstly, Sharma (2021) states that the field of neuroscience can be aided by AI as it helps doctors to remain up to date with the latest research and findings (Sharma, 2021). Secondly, Sharma (2021) claims that AI in thoracic surgery has helped doctors' leaps and bounds in terms of developments in the fields of pathology, radiology, and respiratory medicine (Sharma, 2021).

Third, Sharma (2021) states that AI is useful in the management of cardiac patients in terms of diagnosis and treatment of the disease. This kind of system helps doctors keep track of more than one patient at a time, and more (Sharma, 2021). Fourth, AI is influencing nursing as well. AI will learn to provide “ambulation support, vital sign measurement, medication administration, and infectious disease protocols” as it continuous to develop (Sharma, 2021).

These are only a couple examples of how broad and how specific the aid AI provides in the healthcare field is. Alongside this, Sharma (2021) claims that there are a variety of advantages to using AI in healthcare. These include a) improved management of hospital records, b) clinical decisions will be much easier and faster, c) reducing the workload burden, d) the reduction of administration costs, e) healthcare facilities will be more readily available all due to the reduced cost, and f) the use of wearable healthcare devices will aid in diagnosis of disease and the detection of potential problems (Sharma, 2021).

As with most things in life and in this thesis, the “dual-dilemma” rears its head amidst all the fanfare and appreciation for the wonderful advances that AI has contributed to the world, even in its “weak” form. The following sections will pay attention to five ethical dilemmas that we face when using AI in healthcare and how these may contribute to the perpetuation of bioterror.

3. Privacy & Surveillance

The ethical concern here rests on the right to privacy and the impact of large-scale surveillance of a population. This is a relevant concern today as surveillance of people, with respect to tracking, contact tracing, and more, to keep track of the spread of the COVID-19 pandemic, is a reality. This clarifies that while AI surveillance is beneficial and even critical to curbing the spread of COVID-19, and subsequently the epidemics and pandemics that will follow, we need to be conducting this in a way that is ethical and respects the personal rights and freedoms attributed to individuals. A couple examples of these tracking systems that have been implemented include China's use of AI, Israel's use of AI, and Singapore's specific use of "TraceTogether"[12], which contributed to the development of similar programs in Germany (Shachar, et.al., 2020).

The initial focus here will be defining public health surveillance and offering substantiation for its use. There is some differentiation in the definition of surveillance between countries, some defining it very narrowly and others using a broad definition (WHO, 2017). For the purposes of this thesis, I will define the term broadly as well, as the practice of continual observation of society through the collection, analysis, and interpretation of data in the public health sector with the goal of planning, implementing, and evaluating practices in support of public health (WHO, 2017:14; Nsubuga, n.d.). One distinct objective of systems used in surveillance is the acquisition and provision of data that is capable of guiding the required interventions (Nsubuga, n.d.). WHO (2017:10) explains:

> Disease surveillance has been a basic public health activity since the late 19th century. It is the foundation for initiatives to promote human well-being at the population level. Public health surveillance is the bedrock of outbreak and epidemic response, but it reaches beyond infectious disease. It can contribute to reducing inequalities: pockets of suffering that are unfair, unjust, and preventable cannot be addressed if they are not visible. It is central to understanding the increasing global burden of noncommunicable conditions.

The duty of private actors and public health officials during an emergency, such as public health, includes complying with the regulatory frameworks that govern the use of data and considering any other privacy concern that may be involved. Considerations of privacy, its value, the required action and respect towards it, and its protection of privacy have been discussed for many decades.

Schafer (2011:7) explains that in the span of 100 years, definitions of privacy ranged from "the right to be left alone" (Cooley, 1888) to "the right to be let alone to live one's life with the minimum degree of interference" ("In Privacy and the Law", 1970:45). Westin, in Schafer (2011), defines privacy as "the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others." Further, Miller,

[12] "TraceTogether" is a contact tracing system employed by the Singapore Government to track the spread of the current and ongoing COVID-19 pandemic.

in Schafer (2011:8), defines privacy as "the individual's ability to control the circulation of information relating to him".

The Nuffield Council of Bioethics, however, states that medical professionals and governmental bodies don't always need consent from patients or participants in trials. This can occur if the situation is dire enough, and these institutions and governmental bodies will determine whether the situation calls for data use without consent (Shachar, et.al., 2020). This tells us that this personally identifiable information is readily available to institutions that collect medical data. If this is true, personally identifiable and medical data can be extracted without consideration of privacy and consent by these organizations. These organizations can determine, according to their own criteria or requirements, that there is a greater benefit to using your information without consent. To follow, the question of the ethical considerations of surveillance arises as well as the notion of forgoing informed consent.

Klingler et.al. (2017) distinguished between several ethical issues that should be considered in the discussion surrounding the ethical use of surveillance which specifically pertains to the distinct phases in surveillance. These include, amongst others, background condition issues, design and implementation phase issues, data collection and analysis phase issues, and use of data phase issues (Klinger et.al. 2017).

Firstly, Klingler et.al. (2017:3) states that issues related to background conditions suggest a lack in the framework which guides the judgements made about surveillance systems or if some conditions are not fulfilled that pertain to public health surveillance.[13] These systems of surveillance are developed for a specific purpose, for example, to track the spread of COVID-19. If this tracking system also collects private or personally identifiable information about the people it is tracking, this results in the potential for a violation of private or personally identifiable information (Klingler et.al., 2017:3).

Secondly, as described by Klingler et.al. (2017:4), during the phase of design and implementation, "issues relate to conflicts of priority setting with regards to the type of surveillance system to be implemented and the possibility of ill-designed systems" (Klingler, et.al., 2017:4). This means that during the design of and the implementation of the system, errors can occur that lead to false data, false tracking information, or as mentioned previously, the collection of personally identifiable information. Third, as stated in Klingler et.al. (2017:5), during the data collection and analysis phase, the focus is on breaches of privacy and the foregoing of informed consent. The question of forgoing informed consent is a complex issue and will only be lightly touched on in this thesis. Lastly, Klingler et.al. (2017:5) state that during the use of data phase,

[13] These background conditions refer to the possibility that surveillance systems related to public health could, if the framework that guides them is improperly constructed, collect other information that it is not supposed to collect about people. For example, a surveillance system could collect personally identifiable information about people while attempting to track the spread of a virus or outbreak.

there is the possibility of inadvertently causing harm to patients whose medical data is compromised.

I will briefly focus on this foregoing notion of informed consent and what criteria, or conditions, may allow for the use of medical data without explicit permission or even the knowledge of a patient or participant in a research study. Klingler et.al (2015) distinguishes between a few criteria for foregoing informed consent, namely a) due to the principle of harm, whereby protecting the health of the public justifies defying consent, b) and this includes vulnerable populations such as children, c) due to the proportional nature of harm, wherein the benefits of surveillance outweigh the potential harm, and d) due to the minimalization of privacy infringement (Klingler, et.al., 2017:5).

The line between ethical data use without informed consent and unethical use of data without informed consent is thin and difficult to define. However, ethical considerations need to be considered when conducting surveillance for public health. The goal would be to conduct this kind of data collection and analysis without any potential of privacy breaches, but as this is a bit unrealistic, attention must be given to preserving as much of people's privacy as is within our capabilities. The following section will focus on the ethical concern pertaining to the potential of AI to manipulate the behavior of people.

4. AI & Manipulation of Behavior

With the use of AI, the issue of manipulating behavior is compounded. In Muller (2020:6) it is highlighted that it is commonplace in society today that advertisers, online sellers, and marketers use exploitative methods to influence people's behavior in perfectly legal ways. "Of course, efforts to manipulate behavior in a way that undermines the autonomous rational choice are ancient, but they may gain a new quality when they use AI systems" (Muller, 2020:6). Given our intense, and certainly increasing, interaction with systems that collect our data, knowledge about how we can be manipulated is gained by these artificially intelligent systems.

The collection of this data leads to algorithms targeting small groups or even individuals who are "vulnerable to 'nudges,' manipulation, and deception." In fact, contributing to the manipulation of people on the internet is a core value in the business sector. Muller (2020:7) states with support:

> Furthermore, social media is the prime location for political propaganda. This influence can be used to steer voting behavior, as in the Facebook-Cambridge Analytica 'scandal' (Woolley & Howards 2017) and – if successful – it may harm the autonomy of individuals (Susser, Roessler, & Nissenbaum 2019).

To begin with, as can be viewed in Susser et.al. (2019:1-2), it is necessary to provide the scope of and define online manipulation. The scope of online manipulation concerns several

factors, such as knowledge of a person's interests, preferences, employment, education, financial standing, friends, habits, acquaintances, health, and more. "It enables them to better understand what motivated their targets, what their weaknesses and vulnerabilities are, when they are most susceptible to influence and how most effectively to frame pitches and appeals" (Susser, et.al., 2019:2) The use of the term "targets" here is noticeable and appropriately defines the malicious nature with which people's data is being collected.

Defining manipulation is varied just as accounts of privacy are, but it is worth it to attempt a slight narrowing of the broadest definitions. Broadly, manipulation is defined as the act of controlling or steering someone or something, as in a doctor's use of surgical equipment or the use of cockpit controls to steer a plane. This extends to examples of controlling systems and institutions, where the scandal at Cambridge Analytica is an excellent example of attempted voter manipulation.

Susser et.al. (2019:3) state that there are varying definitions here, where some claim that manipulation refers to "non-rational influence", meaning that manipulation refers to influence of people by "circumventing their rational, deliberate decision-making" (Wood 2014). On the other hand, manipulation has been defined by some to mean "a form of pressure", with 'blackmail' as a prominent example, as noted by Kligman & Culver in Susser et.al., (2019:3).[6]

For the purposes of this thesis, the definition provided by Susser et.al. (2019:4) will be appropriated, wherein simply, manipulation refers to "hidden influence." We are not aware of this influence, but it does have an impact on our thoughts and actions (Susser et.al., 2019:4). This is a truly terrifying notion. This leads us to a necessary distinction that must be made to distinguish hidden influence from other influences, namely persuasion and coercion, and deception.

As is stated by Susser et.al. (2019:4), persuasion and coercion are quite distinct from one another. While persuasion rests on offering reasoning or reasons to support the intended influence, coercion is much more concerning as it refers to placing a constraint or a limit on someone's options in such a way that they are inevitably forced to make the only "rational decision" available. What distinguishes these from other forms of influence, however, is that we are more likely to be aware of persuasion or coercion when it is happening to us or the people around us. These are, therefore, explicit attempts at manipulation.

On the other hand, as claimed in Susser et.al. (2019:4), deception is less clear, less explicit, and poses concerns in its own way. These kinds of manipulation are "hidden", they are dangerous, and they happen without our knowledge to a significant extent. We may be faced with the outcome of a situation, being unable to explain why we acted in that way prior to the action itself. I will define deception as a kind of manipulation which covertly, unlike persuasion and coercion, may influence someone by, for example, promoting beliefs that are false. It is important to note that this goes much deeper, and therefore includes influence by enacting guilt, temptation, seduction, and using the emotions or desires of people against them (Susser et.al., 2019:4).

The notion of manipulation may be varied but rests on the idea that people are not entirely free to make their own decisions or choose what they genuinely want. AI is capable of providing a much deeper and far-reaching influence, both explicitly and implicitly, through persuasion or the like, and the power of this influence will only continue to increase as AI becomes smarter, faster, and more advanced. In a later section in this chapter, attention is paid to the notion of autonomy in relation to ethical responsibility. For now, the following section will focus on the opacity of intelligent systems, or the potential for unintelligible systems.

5. The Opacity of AI systems

The concern here rests on the idea that these artificial systems conduct their learning and their outputs in a way that isn't intelligible to humans and is incredibly difficult to understand, even for the creator of the system. Surden and William (2016) explain that technological opacity "applies any time a technological system engages in behaviors that, while appropriate, may be hard to understand or predict, from the perspective of the human users". There are three different forms of opacity, as described by Burrell (2016), namely a) intentional secrecy, b) technical illiteracy, and c) complexity of outputs.

Firstly, Burell (2016) states that whether the AI systems act on their own accord or are programmed to do so by a malicious actor, the implication of intentional secrecy is, of course, that something that shouldn't be happening, is happening (Surden & Williams, 2016; Burrell, 2016). So, if a company is intentionally keeping secrets, we can assume some foul play.

Secondly, Burrell (2016) emphasizes that the technical illiteracy of people in general plays to the advantage of the malicious actor (Surden & Williams, 2016; Burrell, 2016). Not only do we have a malicious actor furiously working away behind the scenes, collecting data for their nefarious plot, the people who are being affected most directly are entirely unaware that this is happening. This provides a much greater advantage to the malicious actor intent on using the data to perpetuate a biological attack (Herzog, 2019).

Third, Burrell (2016) states the inner workings of decision-making in artificial systems, or put simply, how AI makes it decisions, is not possibly interpretable through human understanding and cognition. In other words, we can understand the output that the system provides but we cannot understand how it has come to this decision to a large enough extent; this part of the AI is commonly referred to as the "black box" (Surden & Williams, 2016; Burrell, 2016). However, even though it is difficult to interpret, and while it may be possible to form some understanding of these processes and inner workings, the level of complexity involved in understanding how the AI makes decisions leaves us at a disadvantage. So, now, not only is our data being stolen without our knowledge but the very people who developed those systems *cannot understand how the AI came to its decision* (Surden & Williams, 2016; Burrell, 2016).

The question that arises is thus, clearly, “How do we protect people from something not even the creators are fully capable of understanding?” Some argue, like Danaher and Robbins (2019), “...a harm requirement for explicability could prevent us from reaping all the possible benefits of AI”. This means that, essentially, the benefits outweigh the risks, and is similar to the initial argument in chapter 1 that mentions the benefits of data mining as a whole. If this is the case, we need not worry because whatever violation occurs is nothing compared to the benefits.

Another response, as described by Nyholm & Gordon (2021:11), to the problem presented by the opacity of an artificial system rests on the idea that we should simply develop ways of understanding the decision-making of the artificial system better, so that even if we cannot understand it fully, we have partial comprehension and, therefore, far less concern (Nyholm and Gordon, 2021:11). This argument, however, will fall into the same pitfall as it is trying to mitigate. I reiterate, we’re incapable of developing an artificially intelligent system that we can understand fully, largely due to the “black-box” issue that is described briefly below. Any attempt to mitigate this will still leave us in a position where we cannot comprehend most of what the system is doing. This leaves a lot of room for malicious involvement, and ultimately, does not solve the true problem. As they say, it is like putting a band-aid on a stab wound.

Neither of these responses account for the potential of harm if artificial systems are free to mine medical or personally identifiable information unencumbered. Claiming that the benefits outweigh the risks is a difficult call to make in an environment where the scope of these harms is complex and difficult to understand. However, we aren’t really in a position to be demanding more, unfortunately. This is the best we can do, and further, we hope to improve. Providing transparency in terms of AI requires the release of information about AI, increasing the inevitability of AI being hacked, attacked, or leaking information. This issue is called the “transparency paradox”, as explained by Burt (2019):

> Call it AI’s “transparency paradox” - while generating more information about AI might create real benefits, it may also create new risks. To navigate this paradox, organizations will need to think carefully about how they’re managing the risks of AI, the information they’re generating about these risks, and how that information is shared and protected.

The transparency paradox is just another dual dilemma. We have no way of understanding these systems fully, and our attempts to make AI more explainable have only brought us so far. Even with the potential benefits unencumbered access to medical and patient data in the development of drugs, treatments, improved diagnoses, and more, the potential for a biological attack seems like a great concern. The following section will expand on the potential and consequences of bias in intelligent systems.

6. Bias in Artificial Intelligence

Here, I will focus on the potential for bias in artificial systems and the direct impact on everyday society. “Bias typically surfaces when unfair judgements are made because the individual making the judgement is influenced by a characteristic that is actually irrelevant to the matter at hand, typically a discriminatory preconception about members of a group” (Muller, 2020:8). This applies not only to humans, but to AI as well.

Nyholm and Gordon (2021:10) explain that biases are extensive and include, amongst others, hiring bias, where one gender is preferred over another, sexual bias, where sexual orientation influences decisions, and a range of racial biases. There are at least three reasons for bias in intelligent systems, as described by Nyholm & Gordon (2021:10), namely bias in data, bias in algorithms, and bias in outcomes.

In the first place, as described by Nyholm & Gordon (2021:10), training artificial systems with biased data not only forces the system to learn bias, but it also exacerbated the bias as well. This leads to systematic bias in the system, whereby the system itself cannot recognize the bias as real bias (Nyholm & Gordon, 2021:10). Secondly, Nyholm & Gordon (2021:10) state that the biases of the developer are either intentionally or unintentionally, consciously, or unconsciously put into the algorithm. Lastly, Nyholm & Gordon (2021:10) emphasize that bias related to the outcome of the process may be influenced by historical biases, whereby the historical expectation of, say crime in one area, is used to predict the level of crime years later, when it may not be the case anymore. Using the generated outcomes, organizations such as the police could be informed about where additional security is required, even when the neighborhood in question has much lower crime rates than the historical data might indicate (Nyholm & Gordon, 2021:10).

To substantiate the description above by Nyholm & Gordon regarding the reasons for bias is data and algorithms, Muller (2020:8-9) offers three forms of biases in artificial systems, namely learned biases, cognitive biases, and statistical or historical biases. Firstly, Muller (2020:9) states that learned biases refer to the biases gained through learning and are, in some cases, hidden from the person with that very bias (Muller, 2020:9). Secondly, as described by Muller (2020:9), cognitive biases fall on the behalf of the developer, who instill biases in the algorithms based on their most primary beliefs, which may impede the judgement of an artificial system. Lastly, in relation to outcome bias, historical or statistical bias refers to outdated data is used and likely does not hold the same bearing on the current situation as previously (Muller, 2020:9).

Bias leads to a range of discrimination, and focusing on healthcare, the consequences can be dire when considering the implications for people’s health. Just as physicians operate under their own unconscious or conscious biases, AI in healthcare poses the same issue, and the same concerns are clear in both cases. It is worth clarifying that the suggestion here is not that biases in humans and biases in AI are the same, simply that biases exist within human beings and in AI.

Bias in healthcare is an age-old problem that continues to prevail in this industry (Zestcott et.al., 2016:529).

Now, I will focus on AI and the potential biases mentioned above. The potential for discrimination is exacerbated by the three forms of biases. The artificial system is able to learn to the data that is available. Data bias, in whichever form it takes, influences the artificial system at its foundation. Further, AI is influenced by its training algorithms, in that the biases of the developer become a part of its functionality. The bias, at this point, is rather ingrained in the AI, and only exacerbated by the last form of bias, namely historical or statistical bias, wherein old data is used inappropriately to assess certain conditions and qualities. This leads us to consider the incredible scope of the biases that are perpetuated through AI systems, and the impact of algorithmic bias in healthcare systems.

6.1. The Core Challenges of Algorithmic Bias

There are three core challenges, as distinguished by Panch et.al. (2019), that healthcare faces due to algorithmic bias in the artificially intelligent systems it uses. As described by Panch et.al. (2019), a standard of fairness is required but presents an impossible issue. “Algorithms are trained on data from the world as it is, which creates the need for additional stewardship, which is complicated as there is no broadly recognized quantitative summary metric for fairness and hence evaluation is ultimately qualitative and subject to implicit biases of the evaluators” (Panch et.al. 2019:2). This tells us that this lack of a standard of fairness leads to even more bias as the evaluation of biases is perpetuated by human beings.

Secondly, Panch et.al. (2019:2) turn to the variability of healthcare system design which due to the diversity of the people involved and the common objectives for the system and people. Diverse cultures, distinct environments, differing socio-economical profiles, alternate lifestyles, evolving preferences, and alternative genetics have influenced and continued to influence the healthcare systems that are in place. While the idea is to develop AI with data that represents the incredible diversity of people, the reality is that “data are not uniformly available for all socioeconomic groups” (Panch, et.al., 2019:2). This causes an imbalance and a greater potential for bias.

Lastly, as stated in Panch et.al. (2019:2), the black box is a major concern when considering AI, but specifically pertains to a challenge in the field of healthcare as well (Panch et.al. 2019:2). As mentioned previously, the black box is the part of the artificial system where decisions are made and is not interpretable by humans. The problem is that physicians and healthcare professions need to know how the system came to the decision about diagnoses, treatment, and surgical options (Panch, et.al., 2019:2). Without this, these workers cannot “check” these decisions to make sure the answer that has been provided is correct. In turn, this means that errors cannot be avoided.

Attention was given to the aspect of bias in artificial systems to highlight the real-world implications of the perpetuation of human bias in these systems. These biases may even offer a reason for the perpetuation of a biological attack, as bias is a powerful motivator within the human condition. As it is possible to implicitly, unintentionally, or unconsciously input bias into algorithms and artificial systems, it is just as possible, if not more possible, that malicious actors will exploit this for their ill-intentioned goals. The following section will turn to autonomy and responsibility of systems.

7. AI, Autonomy, & Responsibility

There is a lot of doomsday discussion regarding the development or creation of a fully autonomous AI. Note, however, there is no need for autonomous AI in the perpetuation of a biological attack. AI, as far as it has come today, can already do this, so we don't need a super-intelligent, evil, autonomous system to evoke the end of the world. This image of the larger-than-life, truly gruesome, ill-intentioned super system may very well happen one day, as it is not for us to say at this point where we might be in 50 years, or even 10 years. However, our current problem is not predicated on a super system that does not yet exist. It is predicated on the very real, very current increase in bioterror across the globe, and the increasing ability of AI contribute to the development of biological agents, the planning and coordinating of attacks, as well as the dispersal of the agent. It doesn't take years to plan when a system can sort through data at a much faster speed than humanly possible.

However, the question of the responsibility of an AI system is an interesting notion. We admit that we only have what we can call "weak AI" now, so any discussion about the responsibility, morality, and autonomy of an AI system is, at this point, theoretical and speculative. We cannot yet ascribe any of these intrinsically human qualities to the systems we have today, no matter their level of advancement past the first computers.

Before we go any further, let's pause for a moment to consider what would constitute autonomous, moral, and responsible super systems. Therefore, this idea is not appropriate to the consideration of our current developments in AI, as described by Totschnig (2020). Therefore, we cannot yet attribute true autonomy to our current AI in the same sense we would attribute it to a human. The ability to make autonomous decisions is not currently ascribed to the AI we have developed thus far but certainly applies to the human being who uses or operates the AI system. It is important to note that autonomy implies the capacity to make decisions and act entirely on its own accord, which even our most advanced current AI falls short of. With a discussion of the autonomy of a bioterror attack, the consideration falls on the terrorists, not the tools they used to perpetuate the attacks. If this were the case, we would blame Little Boy and Fat Man for the destruction of Hiroshima and Nagasaki in 1945 instead of the U.S.

Further, Totschnig (2020) provides a brief overview of three potential views of moral autonomy. To begin with, Daniel Dennet (1998) argues that, while robots are not considered to be moral agents now, they may be so in the future. Secondly, in contrast, Selmer Bringsjord (2007) states that a moral autonomous system will never exist. Third, a completely different view by Joseph Emile Nedeau claims that we are the amoral agents and that autonomous systems are, in fact, the moral agents (Totschnig, 2020).

Responsibility is also a difficult quality to attribute to the current AI we have today. AI is, by any means of the term, heavily influenced by its human creators, as was stated in the sections above detailing the issue of bias. And with any other machines, even programmed machines, and by extension any other manufactured product, when they fail, we do not blame them. If someone is coerced into committing an act of terrorism, or a system is used to cause a terrorist act, we place the blame on the person who perpetuated the attack, assisted the attack, but not the system or the victim of coercion.

For the purposes of this thesis, it is not necessary to consider the qualities of autonomy, morality, or responsibility of an AI system as these qualities cannot, with our current developments, be attributed to these systems. They can only be attributed to the human creators. These creators are responsible for the outcomes of their creations or actions. It is not Little Boy or Fat Man that destroyed major cities in Japan and killed hundreds of thousands of people. We blame the manufacturers, the parties who commission such weaponry, the people who profit from such weaponry, and the governments and countries who approve such measures. The potential for AI to become autonomous should always be monitored while maintaining a strict focus on the human participants in the situations we may face today. Anyway, as was claimed earlier in this section, these qualities are not required to simply use our current AI for the perpetuation of bioterror. The mining of data, the developments of biological data, and the advancements of bioinformatics give our current Ai all it needs to be used to perpetuate bioterror, and even when it is, we cannot blame AI or hold it accountable.

The following section returns to the topic of the current and ongoing COVID-19 pandemic, which was previously mentioned, to emphasize the far-reaching use of AI in the management of a pandemic. The focus will turn to the lessons learned from the world's lack of preparedness for the current pandemic as another reason why we need to be better prepared in the future.

8. COVID-19: Exposing the World's Lack of Preparedness

While this thesis has addressed the fact that AI has been ultimately useful in various areas of curbing the spread and treating the result of the COVID-19 virus that, now, almost two years after first emerging in Wuhan, is still ongoing and continues to affect several countries wave after wave, this pandemic has exposed our lack of preparedness to fight not only pandemics but bioterror as well. As stated by NATO (2021):

> The ongoing COVID-19 crisis has exposed global vulnerabilities to biological threats. As of February 2021, the total number of confirmed COVID cases is 120 million globally with 2.6 million persons having died of the disease (European Center for Disease Prevention and Control, 2021). The wide-reaching and disruptive consequences of the pandemic challenge the ability of national governments, public health authorities, medical services, and international organizations to respond effectively.

Advances in the fields of science and biotechnology have been crucial to the effective use of AI in the curbing of the spread of COVID-19 and the treatments required. We cannot ignore the presence of the dual dilemma here as it makes explicit the potential threats, we face from the very same technology we rely on today to save us from the current pandemic.

As has been claimed in this thesis, we don't need an autonomous, super-intelligent AI system to perpetuate bioterror attacks. Our existing technology and the exponential increase in our technological advancements make it possible today. We can see the devastating effects of even small-scale bioterror attacks and epidemics by looking at the 2001 U.S. anthrax attacks and the 2018-2020 Ebola outbreak in the Congo (NATO, 2021:2). Therefore, we clearly need not concern ourselves just yet with an AI overlord as current technology enables AI to aid the planning and execution of a bioterror attack, on both small and large scales, and is dependent only on the presence of a malicious human actor. As has been stated previously in this thesis, terror is not a rare occurrence, and we have no shortage of terrorists.

It should be noted that there is support and belief in the ability of our new and continuously developing technologies to be just as good at preventing bioterror attacks as the malicious actors improve their technology over time as well. NATO (2021) explains:

> The rapid development of Emerging and Disruptive Technologies (EDTs) like Artificial Intelligence (AI), biotechnology, Big Data and Advances Analytics (BDAA), is likely to dramatically improve our ability to prevent, detect, and contain biological threats, whether deliberate attacks or naturally occurring pandemics. AI has a notable application in the rapid identification of pathogens due to its ability to process copious amounts of data for pattern analysis and information extraction.

As true as this is, it is not the purpose of this thesis to say that AI is only a negative entity, but to specify and to emphasize the dual use of AI and all its related technology to foster the notion that we need to be prepared for what's coming. We were not prepared for COVID-19, and regardless of the claims that it is a bioterror attack, it is teaching us an important lesson: wake up and prepare.

Here, this brief description has highlighted the current and ongoing COVID-19 pandemic as it has exposed a few matters of reality. The following section will detail the complexity of

biological threat assessments to stress the importance of pre-response preparedness and continued amendments of responses as the technological landscape changes.

9. AI & Bioterror: Threat Assessments

As has been stressed repeatedly in this thesis, the advancements in biotechnology, AI, and any related technologies are increasing and continue to increase at exponential rates. The difficulty in conducting and completing biological threat assessments rests on the dual dilemma this thesis has emphasized in all of AI and its related technology throughout. Kernchen (2020) states:

> Threat assessment is an information fusion task which consists of assessing projected future scenarios to determine whether adverse incidents are likely to occur. With regard to biological risks, threat assessment involves obtaining timely, accurate, and relevant intelligence related to the malevolent use of biological agents, and the recognition of existing and future trends and patterns in the evolving threat of biological weapons.

Due to not only the most recent COVID-19 pandemic, but the outbreaks involving Zika, Ebola, SARS, dengue, MERS, and good old-fashioned influenza, it is becoming clearer and clearer that we need preventative measures that are effective. These measures must be effective against not only these natural occurrences but the biological attacks that are likely to gain popularity as AI eases the road to pathogen manipulation and more (Kernchen, 2020:2). Kernchen (2020:3) states that there are two things to keep in mind with regards to biological threat assessments, namely a) the relevance of tacit knowledge and b) the ever-changing biotech landscape, considered below.

9.1. Tacit Knowledge & The Ever-Changing Biotech Landscape

Firstly, Kernchen (2020:3) states that when assessing the potential for future bioterror attacks by malicious actors, tacit knowledge is required. Tacit knowledge will be defined here as knowledge that is difficult to extract from the source of the data. This tacit knowledge aids in the assessment of the dual nature of biotechnologies and AI (Kernchen, 2020:3). In other words, the ability of these technologies is to, essentially, look beyond the data and infer patterns or likelihoods of occurrences is a vitally essential element when assessing potential threats.

Secondly, Kernchen (2020:3) states that the ever-changing biological landscape poses both opportunities for improvement, which is favorable, and opportunities for malicious exploitation, once again due to the dual dilemma. There are four notable, primary chances to the biotech landscape to keep in mind, namely a) the increasing convergence of disciplines such as the life-sciences and non-life sciences, b) the expansion of biology to include other disciplines such as engineering, c) the globalization of biotech and the life sciences, and d) the funding that is afforded to these technologies and disciplines (Kernchen, 2020:3).

Both of these aspects will continue to have an influence on biotechnology and its use in the perpetuation of bioterrorism. The value of AI and biotech is incredible, especially in the healthcare and biological fields. We can continue to reap the rewards of this if we are, in conjunction, developing the field of biodefence as well. While we focus on developing preventative and security technologies that will protect us from bioterrorism perpetuated by AI, ethics remains a vital consideration.

The principles and ethics of the dual-use nature of biotech and related technologies give rise to a variety of considerations. Rodrigues (2015:6) states the principles that are relevant to this discussion include, a) respect for the rights and dignity of humans, b) respect for the global environment and humanity's welfare, c) access to science on an equal basis, d) considerations of the future of the earth and future generations, and e) non-maleficence and beneficence (Rodrigues, 2015:6).

Further, Rodrigues (2015:6) ascribes the following principles as well, namely a) protection of the environment, future generations, and human dignity and rights, b) considerations of benefit, harm, privacy, confidentiality, capacity to consent, non-stigmatization, non-discrimination, equality, and justice, and c) respect for culture, diversity, vulnerabilities, and integrity (Rodrigues, 2015:6).

The ethical concerns, as explained by Rodrigues (2015:7), rest on the following, including a) safety principles, b) precaution, c) freedom of research, d) transparency, and e) justice. Situations where ethical concerns in research may arise include, a) the use of classified materials, information, or techniques, b) the use of restricted or dangerous materials, and c) the use of results from research that poses dangers to society or individuals (Rodrigues, 2015:7).

This section has briefly demonstrated the complexity of biological threat assessments and highlighted the principles and ethical concerns that are present within dual-use research. The following section will focus on AI and National Security.

10. AI & National Security

It is necessary to center AI in the realm on National Security as the use of AI to perpetuate bioterrorism poses direct and explicit threats to the security of nations all over the world. In terms of national security, we can frame the definition of AI according to 5 ideas. There are several applications and challenges of AI in national security which highlights the dual-use nature of AI in National Security, and by extension, bioterrorism.

To begin with, while no commonly accepted definition exists, and beyond the definition provided at the beginning of this chapter, there are a few more considerations to take into account when the focus is national security. These state that AI is, firstly, a system that can perform tasks without much oversight or gain knowledge from the provided data and its own experiences.

Secondly, regardless of its construction, it is a system that requires, at least, human-like planning, cognition, perception, communication, or learning. Thirdly, a system that can act or think similarly to human beings and functions with neural networks and other cognitive architecture. (Congressional Research Service (CRS), 2020:1-2). Below, CRS's focus on the applications of AI for National Security will be mentioned briefly as the perspective of AI held by CRS (2020).

10.1. Applications of AI for National Security

According to this definition, an artificially intelligent system is at least human-like and can perform some human-like tasks based on its various abilities as an artificial system. Due to this nature, it has an incredible variety in its application to national security, as described by the CRS (2020) below. Firstly, the CRS (2020) states that the use of AI is expected to benefit the fields of intelligence, reconnaissance, and surveillance. The benefit of AI, here, is that it automates some forms of work, freeing up a lot of time for the human analysts[14]. (CRS, 2020:10). Therefore, AI can be used in this capacity to identify security and bioterror threats far, far more quickly than human beings are able to.

Secondly, the CRS (2020) explains that AI has applications in the field of logistics, especially when considering military logistics. A fitting example of this is the application of AI in the Air Force, where AI is used for the prediction of required aircraft maintenance on an individual level, based on the individual needs of each aircraft. "This approach, currently used by the F-35's Autonomic Logistics Information System, extracts real-time sensor data embedded in the aircraft's engine and other onboard systems and feeds the data into a predictive algorithm to determine when technicians need to inspect the aircraft or replace parts" (CRS, 2020:11).

Third, as described by the CRS (2020), AI has what can only be described as incredible potential for the realm of control and command in the military field. The ultimate goal here is to centralize and link the planning and the execution of space, cyberspace, sea, land, and air-based operations. This will give them the power, in the rather near future, to use AI to "fuse data from sensors in all of these domains to create a single source of information" (CRS, 2020:12-13). This may extend to the determination of operations by the AI system, as has been demonstrated in the medical field, where AI has, in some cases, taken over diagnostics for doctors, simply supplying the best course of action without much consideration of the validity of the diagnosis on behalf of the doctor (CRS, 2020:12-13).

Here, the focus was on a few applications of AI, specifically in relation to national security. These apply heavily to the military sector and the overall government. The Department of Homeland Security (DHS) and The Department of National Intelligence (DNI), in a joint publication, mention a few vulnerabilities and risks involved with AI. First, as described by DHS

[14] This includes sifting through data or hours of drone footage, drastically reducing the time these analysts spend on this kind of work. This frees them to focus on other, more complicated work that can't be left to an artificially intelligent system.

& DNI (2018:13), the concern rests on the inability or the lack of advancement of humans to understand and use technology as it progresses. In other words, if we neglect keeping up with the advancements in technology, we will leave room for malicious actors to exploit vulnerabilities in AI (DHS & DNI, 2018:13). This is particularly concerning when considering bioterrorism and the application of AI in the planning and execution of such attacks.

Secondly, as described by DHS & DNI (2018:13), the integrity of the data has an impact on the performance of algorithms and the capabilities of AI. Therefore, as data can be corrupted by hackers or malicious actors, and this in turn can influence the output of the algorithm, the integrity of the data is essential. "If an object detection system used by warfighters is compromised with this type of attack, it can misidentify certain objects (i.e., labelling an enemy tank as a tree), or the system may fail to detect certain objects altogether" (DHS & DNI, 2018:13). This clarifies that, in the event of a potential bioterror attack, AI that is programmed to detect anomalies or patterns could be influenced by malicious actors so that they do not detect an impending attack.

10.2. Mitigation of Risk

To mitigate these risks and others the DHS & DNI (2018:19) joint publication offers a guide to decision-making regarding the use of AI, "on how to begin to identify, adopt, and implement specific standards for specific needs in the national and homeland security context." To begin with, as stated by DHS & DNI (2018:13), the context of the security issue must be identified. At this point, consideration is paid to the impact of the system on security, whether explicit or implicit, and whether the use of this system is guided by the priorities set out in accordance with national defense strategies. Next, the DHS & DNI (2018:19) explain that the application of the system will pose certain risks and issues that need to be determined by questioning, on the one hand, how a malicious actor may exploit this system, and on the other hand, which aspects of processes such as decision-making could be exploited (DHS & DNI, 2018:19).

Following this, the DHS & DNH (2019:19) mention another key point regards the metrics that determine if a system is working accurately, trustworthily, or appropriately, and the measures that help determine whether the results of the algorithm are worthy of confidence and trust. The user of the AI system must be aware of the internal functioning of the system to prevent unnecessary errors or unnoticed intrusions (DHS & DNI, 2018:19).

Further, as described by the DHS & DNH (2019:19), the user must have confidence in the results of the system, or the entire process ends up being a waste of time with unsatisfactory results. This has an important impact when the incorrect or misguided results are not noticed and used in research or the like (DHS & DNI, 2018:19). For example, hackers or malicious actors may attempt to corrupt data unbeknownst to the researchers that are conducting clinical trials or researching new treatments. This, inadvertently, causes them to use false or misguided data in their trials and research, leading to potentially harmful or even fatal consequences for patients.

Lastly, the DHS & DNH (2019:19) emphasize that the focus must be turned to identifying key assumptions, limitations, and considerations. The following questions pertain to this issue: What do these standards miss or ignore? What qualifies as an effective or ineffective standard? In what ways could a malicious actor exploit these standards or faults in the standards? There are, of course, many more questions we could ask (DHS & DNI, 2018:19). When considering these key elements, the standards of AI use place a focus on continuous improvement, which is what is necessary in the fight against bioterror. We need to be aware, not only of the standards required or expected, but of the potential pitfalls, failures, and harms that could be attributed the AI, in support of the dual dilemma.

The following section will finish off this chapter with an in-depth look at AI and the dangers of its use in the perpetuation of bioterror.

11. AI & Bioterror: The Way Forward

This thesis has placed a large interest in the dual dilemma that is present in data mining, biological data, bioinformatics, and AI. The way forward is one riddled with obstacles, where new and old technologies will continue to pose threats, where our advancements are also the advancements of our enemies. The better protected be become, the harder malicious actors work to advance past our available measures.

Technological and scientific advancement is exponential; thus, society needs to prepare itself for an on-going battle that will last for decades to come. While there is no ultimate protection from AI without the complete destruction thereof and considering that AI is necessary for all walks of human progress, the only solution is to do our best. And our best needs to be far beyond any malicious actor with ill intentions. As stated by the U.S. National Biodefense Strategy (2018:8), “The risks from biological threats cannot be reduced to zero – but they can and must be managed. Wide-ranging threats require a comprehensive approach to minimizing the risks.”

This closing section of this chapter will detail the viable solutions and recommendations for the continued use of AI while protecting society from biological threats malicious actors may perpetuate using AI.

To begin with, Brundage et.al. (2018) explains that there are three security domains where attention must be focused as the evolution of these domains over time must be adapted to if the intention is to prevent malicious use of AI and its related technologies (Brundage, et.al., 2018). These domains include digital security, physical security, and political security.

In terms of digital security, Brundage et.al. (2018) states that the threat of cyberattacks increases over time due to the increased ability of AI to complete automated tasks. Here, the time spent on the perpetuation of the attack is essentially nullified when compared to the time a human

would spend completing the same task. In the same breath, as AI advances, it will gain all the greater control of the digital sphere, having greater and greater access to data that may be used to perpetuate biological attacks.

Secondly, as is stated by Brundage et.al. (2018), physical security refers to the domain that is threatened when AI is used to carry out drone strikes or use other autonomous weapons systems. As with the previous domain, the increasing advancements in AI lead to an increase in these kinds of threats, where people may lose their lives. The potential for drone strikes as AI advances may very well be used to disperse biological agents, rendering it rather simple for a malicious actor to disseminate a crippling agent over a large area of a population.

The last domain Brundage et.al. (2018) refers to is political security, wherein mass-collected data is at an increasing chance of exploitation for the purposes of persuasion, deception, or manipulation. The capacity to analyze human behavior is an advancing quality of modern AI, and thereby its ability to influence human beings will increase as well (Brundage, et.al., 2018).

Due to the ever-changing domains mentioned above, based on Brundage et.al., (2018), the landscape of the potential threats changes as well and requires our adaptation. On the one hand, new threats may emerge, while on the other, we will face an increasing number of new or never-seen-before threats. New threats, such as the perpetuation of bioterror and the use of newly discovered biological agents, may arise due to the actions of malicious agents.

As technology advances and AI becomes increasingly capable, predictable, and unpredictable threats will emerge. In terms of the expansion of current threats, even while protective measures evolve, “A natural effect would be to expand the set of actors who carry out particular attacks, the rate at which they can carry out these attacks, and the set of potential targets” (Brundage, et.al., 2018).

Above and beyond this, the character of threats may change as well, presenting its own range of concerns and issues. As the capabilities and computational capacity of AI increases, the very character or nature of threats will change (Brundage, et.al., 2018). With the ability to finely target populations or communities, the difficulty in attributing blame arises. When the malicious actor hides behind AI and considering the current and future vulnerabilities of our AI that can be exploited by such actors, it is most clear that our advancements simply must keep up with this changing landscape of security domains and threats (Brundage, et.al., 2018).

Considering the security domains and the changes to the landscape of threats, preparedness and ethical responses become paramount. Based on the principles of AI ethics for the Department of Defense (DoD) in the U.S, the use of AI must be guided, even for purposes of national defense and security, by ethical principles such as responsibility, traceability, equitability, governability, and reliability.

Firstly, the DoD (n.d.) states that responsibility as a principle of AI ethics relates to the expectation that human beings should be held accountable for their judgements regarding the use of AI systems, in the development of the system, throughout its use, and in respect to any potential outcome it may lead to. Further, just as we place the blame on the perpetuator of malicious actions, we must hold our own developers accountable for the errors, flaws, losses, is misguided actions of their systems (DoD, n.d.).

Secondly, the DoD (n.d.) emphasizes that the traceability principle focuses on the advancement of technical experts (DoD, n.d.). In other words, we need to be able to discern where problems and issues may arise during ever process, as well as discerning points for exploitation and areas that have already been exploited, either by accident or purposefully (DoD, n.d.).

Third, the DoD (n.d.) states that equitability principle refers to the assurance that bias, whether intentional or unintentional, is not present in the development or execution of AI systems, especially where these biases may harm people. This is true in the case of both combat and non-combat systems. This tells us that the bias in our own systems must be identified and addressed to ensure that adequate preparation for bioterror.

As stated in DoD (n.d.), governability principle rests on the ideal that all participants involved in the development of AI should develop systems that "fulfill their intended function while possessing the ability to detect and avoid unintended harm or disruption" and this must include "human or automated disengagement of deactivation of deployed systems" that are escalating unintentional or unpredictable behavior (DoD, n.d.). In other words, we need a failsafe in place to protect against unwanted behavior or uses of AI systems. This will allow us to prevent many of the potential risks of using AI, including the potential for our AI to be hacked and used to perpetuate bioterror.

Lastly, the DoD (n.d.) explains that the reliability principle is the evaluative measure to determine the security, the safety, and the strength of these systems. This evaluation must be ongoing, continuing with the systems as they themselves and the developers use them in advance (DoD, n.d.). Therefore, issues, problems, areas of vulnerability, or areas that are exploitable must be continuously identified, evaluated, and addressed to avoid any intentional or unintentional harm to the population or community (DoD, n.d.).

These ethical principles emphasize how we can continue to use AI to our benefit while mitigating the risks. The common ground will be shared below to address the core values, principles, or ethical considerations when using AI in the preparation for and response to bioterror. The ethical principles of responsibility, traceability, equitability, governability, and reliability are both general in the sense that they are, mostly, universally shared, and specific, in that they explicitly govern the required actions and responsibilities of the developers and users of AI.

The following section of this chapter, and the closing section of this thesis, will offer seven recommendations regarding AI-enabled bioterror incidents. The first recommendation holds that we must take the dual nature of technology seriously in order to do anything about it in the first place. The second recommendation focuses on the requirement for concrete and effective policies regarding the development and use of AI. The third recommendation focuses on the implementation of greater resources and training to ensure that staff working in these fields are up to date on technology and thus does not lead to destructive errors.

The fourth recommendation focuses on the growth and development of AI engineering. The fifth recommendation relays the necessity of monitoring potentially harmful research as this can result in harmful consequences for society and communities. The sixth recommendation focuses on risk awareness and security issues that pertain to AI-enabled bioterrorism. The last recommendation highlights several aspects of biodefence that are pertinent to consider when responding to bioterrorism.

11.1. Recommendation 1: Understand the threat of the dual dilemma

AI, as with most other technologies mentioned in this thesis, is a dual-use system. All recommendations to follow rest on the dual-use nature of AI, highlighting the need for awareness. Following Brundage et.al. (2018:51), understanding this nature of AI provides the researchers with the awareness that their discoveries need to be evaluated for potential harmful applications. This awareness guides the responsibility of research to prioritize the well-being of humanity.

However, it is very well possible that the same technology we develop to counteract possible bioterror events may be exploited to perpetuate the same attack it was designed to prevent (Brundage et.al., 2018:51). Given the essential eventuality that our necessary protective measures are developed from technology that could be used to counteract such measures by a malicious actor, all organizations, researchers, and government bodies involved in AI development are hard pressed to employ the recommendations to follow.

11.2. Recommendation 2: Create & Implement Concrete Policy

When creating policies on the ethical development and use of AI systems and related technologies, great attention and focus must be given to collaboration with experts and researchers on the subject (Brundage et.al., 2018:51). Policymakers cannot put forward accurate or even relevant policies without the required knowledge that can only be gained from the perspectives and experience of the developers, technical researchers, and subject experts. This avoids irrelevant policies, uninformed policies, interferential policies, and more, that may impede scientific progress while providing no clear benefit (Brundage et.al., 2018:51).

Further, within these policies, focus must be given to several aspects of AI, especially when medical data is involved. These include a) the development of these policies must align with our

ethical values and obligations towards humanity's well-being, b) the development of new or modern ethical guidelines as new issues emerge now and, in the future, c) the development of these policies must account for the potential that AI systems may reveal personally identifiable information, and d) the development of these policies must focus on inclusivity in gender, age, socioeconomic or geographical groupings, and more (Connected Health, n.d.).

11.3. Recommendation 3: Implement Training & Increase Resources

As has been mentioned previously, an issue arises when the workforce is not up to date about the advancements of technology nor knows how to use them efficiently. This leaves room for error and gives freedom to malicious actors to work behind the scenes (DoD & DIB, n.d.). Training and educational programs are a requirement to aid the workforce in keeping up with modern technologies and adaptations to the technologies they are used to. This applies to corporate workers, scientific researchers, technological developers, and more. Without proper training and education on these important technologies and, of course, why this is a concern, errors, misconceptions, and malicious use could see an exponential increase (DoD & DIB, n.d.).

In order to do this, resources and funding must be provided. In Brockmann et.al. (2019:42-44), it is stated that, depending on the sector, the kind of resources they require will differ. On the national government level, one focal point is the increase of outreach programs that connect academic institutions, independent communities, and industries alike. On a smaller level, such as academic institutions, resources such as mandatory courses and seminars focusing on ethics could be provided in the name of biosafety considerations (Brockmann et.al., 2019:42-44). This reaches students as they become a part of the academic community and reinforces ethical ideals throughout their education.

11.4. Recommendation 4: Grow & Improve AI Engineering

AI engineering should be supported by all stakeholders involved as its growth and development is of fundamental importance to not only our technological advancements but to our safety and security as well (DoD & DIB, n.d.). This growth and continued development should be based on sound and ethical practices. To foster this growth, programs can be implemented to encourage the younger generation to get involved in this and related fields, giving hope for the future. The more people that become passionate about the field, the further and the faster it will progress.

However, keeping in mind the dual nature of AI systems, growth and improvement needs to apply to our assessment of these systems (DoD & DIB, n.d.:10). Here are a few questions that can be asked at this point. Does this system include any algorithmic or data source bias? Are there loopholes in its security that can be exploited by malicious actors? Does the system perform its intended function within a certain time frame and to a certain quality? Is the source information accurate or incomplete? The assessment of such a system rests not only on its potential

shortcomings but the inevitable outcome of the exploitation of those flaws or failures (DoD & DIB, n.d.:10).

11.5. Recommendation 5: Monitor Potentially Harmful Research

As has been mentioned in this thesis, our use and development of AI needs to advance as our enemy's technology advances. Therefore, as is explained by Brockmann et.al. (2019:42), much potentially harmful research exists across the world as we attempt to develop our systems to a superior level. This must be monitored not only by the institution where the research is conducted, but requires, especially when the research involves risks to national security, government oversight. When researchers and institutions display a lack of respect for society's well-being, a solution can only be provided by a governing body of higher standing (Brockmann et.al., 2019:41-42).

In the interest of ethical scientific exploration and with the focus on society's well-being, we need some form of oversight in place to ensure that uncaring or unethical researchers do not create something we cannot protect ourselves against. This is an age-old concern. There are varying levels where this applies, whereby the private sector could monitor the use of drones, setting out clear guidelines that may prevent the use thereof for malicious purposes. Brockmann et.al. (2019:43) emphasizes that on the level of the academic institution, assessment, interdisciplinary oversight, and collaboration between institutions worldwide is required.

11.6. Recommendation 6: Risk-awareness for Decision-Making

At the very start, risk management is of essential importance when dealing with dual-nature technologies such as AI. The methodology that is applied will vary slightly from institution to organization to governing body, but the core ethical value remains the same universally: maintaining the benefits we receive bountifully from AI and its related technologies while managing any security risks that may arise due to our ethical obligation to prevent harm to society through our research. Being aware of the potential risks and security issues leads to improved and most applicable or appropriate decision-making.

Decision-making when managing potential future risks is rather different from managing a bioterror attack (NBS, 2018:9). Thereby, the skills we employ to manage everyday breaches and errors are minor level compared to what is used or needed to detect, predict, or prevent an AI bioterror attack. Decision-making in a time of bioterror must occur rapidly and according to the best possible data or intelligence. This can only occur if small-scale risk management practices are implemented and used as a basis for the development of larger strategies. A further benefit is that small-scale risk management is ongoing and will highlight or allow for the detection of larger concerns as well (NBS, 2018:9).

11.7. Recommendation 7: Biodefence: Prediction, Detection, Prevention, Preparation

Lastly, there is the matter of biodefence. As stated by Brockman et.al. (2019:42), biodefense encompasses several aspects. These include the prediction of potential bioterror attacks, the detection of real-time occurring threats, measures to mitigate against or prevent such attacks, and the preparation of the response to the attacks. To do this, of course, scientific research needs to be conducted to assess vulnerabilities, risk areas, the potential for loopholes in security measures, the addition of corrupted data or theft of data my malicious actors, and much more. Based on this evidence, the prediction of problem areas will be conducted to determine additional areas for research. Further, prediction will serve the detection of impending attacks, and the improvement and development of both preventative measures and preparation for responses (Brockmann et.al., 2019:42).

Not only do we require government or nationwide capacity to prevent, detect, and prepare responses to bioterror but local governments must be involved and educated as well. This will allow for faster recognition of potential bioterror attacks and potentially resulting epidemics or pandemics and an increased rate of detection of attempts to acquire biological agents or related materials. Further, institutions working with biological agents that could be exploited for malicious purposes must be overseen to improve and apply biosafety regulations and biosecurity measures appropriately. Based on the gained knowledge and awareness from the endless research into the subjects at hand, responses can be planned and executed to the greatest possible efficiency. To aid this, global cooperation on bioterror surveillance increases the universal awareness and knowledge of potential bioterror threats, leading to more research, which in turn leads to improved and more appropriate responses to such incidents.

The response to bioterror incidents will revolve largely around what kind of attack has been perpetuated. As we cannot predict for every potential event, focus is generally on the recurring agents (i.e., anthrax). When new biological agents emerge, AI will be more necessary than ever as researchers and governments alike will scramble to identify the new biological agent while investigating its composition, the symptoms it causes, and the potential treatment.

12. Conclusion

Even with all our preparation, the intervention into such a matter would inevitably be delayed, just like your immune system may take a few days to develop an immune response to a new virus it has never experienced before (much like COVID-19). As was stated at the start of this section, we cannot prepare for everything, nor can we reduce the impact of such events to nothing. However, “doing our best” is a rather incredible task, with respect to bioterror defense, that upholds the universal ethical values that we collectively believe in, whereby the well-being of people is the utmost concern and duty.

This chapter focused, on the one hand, on ethical considerations that pertain to the development and use of artificially intelligent systems. This included but was not limited to privacy, manipulation, and autonomy. Each of these ethical considerations highlights the dual nature of Ai technologies and highlights the need for intervention. On the other hand, the recommendations that follow serve as foundational and guiding principles, each interested in a specific element of the discussion surrounding AI-enabled bioterrorism. These include, but are not limited to, a focus on policy, training and resources, and ethical response to bioterror. It is pertinent to note that these recommendations are not capable of defending against every possibility of this dual dilemma, but we need these guidelines and principles to prepare for and respond to bioterror most effectively and ethically.

Chapter 5
Conclusion

This thesis has demonstrated the dual nature of technology, present in every level of technological advancement and development of AI. *Dual nature* has been defined as the potential of technology to be used both beneficially and detrimentally. Due to this dual nature, technology and research must be monitored to determine which developments may result in harm to society. Further, this can determine how to mitigate the potential detrimental consequences while continuing to enjoy the benefits that the very same technology offers. As this thesis is an overview of the different topics, fields, and ethical considerations present in the development of AI evolution, attention was paid to three core levels of AI development. The intention here was not to detail specific ethical discussions that make grand claims about ethics but to bring to light the sheer spectrum of potential ethical blunders and their potentially detrimental consequences given the connection to AI and bioterrorism.

The first core level of this development is data mining, where massive amounts of information is mined, analyzed, and stored. AI requires this vast amount of data to be developed and learn enough from its training data to be effective, efficient, and useful in its core task. However, as was highlighted in the first chapter of this thesis, even at this level, ethical consideration must be paid as this data influences the AI and its decision-making process, as is explained in the final chapter more fully.

Biological data, as a specific kind of data that is collected, refers to medical and patient data from medical or research institutions, whether private or public. Biological data is the focus of this thesis due to the larger concern contained here, namely the use of biological data to perpetuate AI-enabled bioterror. Without ethical use or governance of biological data, such as a focus on ethical organization, analysis, storage, and more, of these databases, and without protective or safety measures in place to protect the data during the mining stage, the analysis stage, and the storage stage of data collection, biological data can be exploited to perpetuate biological attacks.

The next step in the evolution of AI-enabling technologies is that of bioinformatics, where the development of machine learning, deep learning, and artificial neural networks (ANNs) is required to manage the vast amount of biological data that is collected. These forms of learning all require a lot of data to function and perform their functions, and their development is necessitated by the influx of data that needs to be managed. This works hand in hand to speed up the development of these technologies. Ethical concerns in bioinformatics pertain largely to the ethical development of algorithms, the ethical storage of data in warehouses, and how this influences the ethical development of AI. This ensures that the least possible bias is achieved, regardless of the origin of this bias, as either encoded by the developer, either consciously or subconsciously, or as learned through its training data.

These three core levels of technological development led to the development of increasingly improving artificially intelligent systems. Just as science progresses exponentially, AI and its related technologies continue to develop exponentially as well. This has led scientists, theorists, academics, and filmmakers to the conclusion that any form of AI that develops consciousness will immediately set out to harm, enslave, or destroy humanity in some way. This thesis has no intention of commenting on this aspect, or this degree, of AI development as it is not relevant to the real-world problem at hand, namely AI-enabled bioterror, which is possible with the AI that we have currently.

The focus of this thesis was based on the notion that access to and use of biological data improved and exponentially decreases the time needed to plan a biological attack. This allows for far greater accuracy in determining the correct areas or groups to attack. These areas may be sought out for a number of reasons, such as religion, political beliefs, or for no reason at all other than to harm or cause chaos. The planning of the attack itself is more effective when combined with biological data as specific or most harmful biological agents can be selected based on their potential impact on the areas that have been chosen for an attack. The time spent on the analysis of the data is significantly reduced by the use of AI systems as described above. Therefore, while biological data and AI are to essential requirements for the perpetuation of bioterror, they certainly provide a number of benefits.

The last chapter of this thesis looked at several ethical considerations in the development and use of AI, such as privacy violation, manipulation of behavior, the opacity of these systems, bias, and the irrelevance of autonomy and responsibility. This emphasized that ethical concern and consideration must not only be paid at every level of AI technological development but in the way we use it as well. These concerns rest on providing security, safety, and protection to the public and society.

The recommendations that have been given cover several vital aspects of the processes of preparation, prevention, detection, mitigation, and response to bioterrorism. Firstly, it is simply impossible to do anything about bioterrorism if we do not take the threat seriously. Terrorism is a reality in our world and as criminals improve their technologies, we must do so as well. The second recommendation relies on the implementation of concrete policies pertaining to the ethical development of AI, whereby policy makers must work together with experts to develop the guiding principles and policies.

However, none of this will be useful if we do not implement and require the proper training of staff and provide them with the resources they need to use these technologies effectively and ethically. This ensures that ethical conduct and principles are upheld in the name of the protection of society and its people. Further, it is recommended that AI development and engineering is grown and improved in a way that is ethical, useful, and efficient. This must apply to our assessment of

these systems to ensure there is no bias, lack of or inability to function, and no involvement or interference from malicious actors.

From this, the next recommendation follows, where the monitoring of potentially harmful research is required. We cannot hide from or ignore the fact that research and technology are always faced with the dual dilemma and must have measures in place to identify research that could cause harm. The fifth recommendation centers on the ideas of security, risk-awareness, and decision-making. Dual nature technologies require this kind of assessment so that ethical and efficient decisions can be made about how to deal with the potential harm of research or technology.

Finally, this thesis offers a recommendation regarding biodefence and details the requirement of ethical consideration even in our response to bioterrorism. The response to a bioterror attack will be individual as there are a vast variety of attacks that could occur. This means that each occurrence of bioterrorism could be completely different from the previous or the next. Stressing the importance of our own development of AI, it is apparent at this point that we need not only fear the use of AI by malicious actors but the failure of our AI to detect and aid the prevention or mitigation and response to bioterror incidents. Our ability to respond to bioterror is influenced by every topic of this thesis, where responsibility and ethics need to be taken seriously at even the most fundamental level of data mining.

In conclusion, this thesis has emphasized the clear dual nature of AI and its related technologies and the objective requirement of ethical concern and consideration at absolutely every level of this field and its related fields. It is a reality that we could never fully escape from this dual nature but every effort we put into mitigating ethical compromise of data and its related technologies puts us in a better position than we were before. With privacy preserving data mining measures, ethical and secure collection, use, and storage of biological data, the ethical development of machine learning and deep learning algorithms, and the conscious ethical development of AI, we put ourselves in the best possible position to prevent bioterror before it occurs.

Bibliography:

1. AAAS-FBI-UNICRI (US) Forum on Microbial Threats. 2007. *Ethical and Legal Considerations in the Life Sciences.* Washington, DC: American Association for the Advancement of Science, 35-58
2. Abuhammad, S., Khabour, O. & Alzoubi, K. 2020. COVID-19 Contact-Tracing Technology: Acceptability and Ethical Issues of Use. *Patient Preference and Adherence*, [Online]14:1639-1647. Available: https://www.dovepress.com/covid-19-contact-tracing-technology-acceptability-and-ethical-issues-o-peer-reviewed-fulltext-article-PPA [2021, 11 August]
3. Achan, P., Warrier, A. & Chitturi, B. 2012. Biological Data Handling Methods. [Electronic], Research Gate. Available: https://www.researchgate.net/publication/266177806_Biological_Data_Handling_Methods [2021, 29 July].
4. Adomavicius, G. & Tuzhilin, A. 2001. Using data mining methods to build customer profiles. *Computer*, 34(2):74-81, doi: 10.1109/2.901170
5. Arora, N., Banerjee, A. & Narasu, M. 2020. The role of artificial intelligence in tackling COVID-19. *Future Virology*, 15(11): 717-724.
6. Ayukekbong, J., Ntemgwa, M., Ayukekbong, S., Ashu, E. & Agbor, T. 2020. COVID-19 compared to other epidemic coronavirus diseases and the flu. *World Journal of Clinical Infectious Diseases*, 10(1):1-13.
7. Barras, V. and Greub, G., 2014. History of biological warfare and bioterrorism. *Clinical Microbiology and Infection*, 20(6):497-502.
8. Bhatia, V. & Hasija, V. 2016. Targeted advertising using behavioral data and special data mining, in *2016 Eighth International Conference on Ubiquitous and Future Networks (ICUFN)*, 937-942
9. Boulanin, V., Brockman, K. & Brauer, S. 2019. *Bio Plus X: Arms Control and the Convergence of Biology and Emerging Technologies.* SIPRI Publications, 41-43.
10. Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., Dafoe, A., Scharre, P., Zeitzoff, T., Filar, B., Anderson, H., Roff, H., Allen, G., Steinhardt, J., Flynn, C., O hEigeartaigh, S., Beard, S., Belfield, H., Farquhar, S., Lyle, C., Crootof, R., Evans, O., Page, M., Bryson, J., Yampolski, R. and Amodei, D., 2018. *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation*. [Electronic] Available: https://www.researchgate.net/publication/323302750_The_Malicious_Use_of_Artificial_Intelligence_Forecasting_Prevention_and_Mitigation.
11. Burt, a. 2019. The AI Transparency Paradox. *Harvard Business Review.* [Online] Available: https://hbr.org/2019/12/the-ai-transparency-paradox [2021, 12 August]
12. Carus, S. 1998. *Bioterrorism and Biocrimes: The Illicit Use of Biological Agents since 1990.* Washington, DC: National Defense University, 3-11.
13. Chakraborty, I. & Choudhury, A. 2017. Artificial Intelligence in Biological data. *Journal of Information Technology & Software Engineering*, 07(04):1
14. Chalal, H. & Gulia, P. 2018. Techniques and algorithms of PPDM. *International Journal for Scientific Research & Development*, 3(4):3484-3485.

15. Chang, A. 2020. Artificial Intelligence and COVID-19: Present state and future vision. *Intelligence-Based Medicine*, 3-4. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2666521220300120 [2021, 11 August].
16. Chicco, D. 2017. Ten quick tips for machine learning in computational biology. *BioData Mining*, 10(1).
17. Ching, T., Himmelstein, D., Beaulieu-Jones, B., Kalinin, A., Do, B., Way, G., Ferrero, E., Agapow, P., Zietz, M., Hoffman, M., Xie, W., Rosen, G., Lengerich, B., Israeli, J., Lanchantin, J., Woloszynek, S., Carpenter, A., Shrikumar, A., Xu, J., Cofer, E., Lavender, C., Turaga, S., Alexandari, A., Lu, Z., Harris, D., DeCaprio, D., Qi, Y., Kundaje, A., Peng, Y., Wiley, L., Segler, M., Boca, S., Swamidass, S., Huang, A., Gitter, A. & Greene, C. 2018. Opportunities and obstacles for deep learning in biology and medicine. *Journal of The Royal Society Interface*, 15(141).
18. Churko, J., Mantalas, G., Snyder, M. & Wu, J., 2013. Overview of high throughput sequencing technologies to elucidate molecular pathways in cardiovascular diseases. *Circulation Research*, 112(12):1613-1623.
19. Colanna, L., 2013. A taxonomy and classification of data mining. *Science and Technology Law Review*, 16(2):309-315.
20. Cook, J. 2005. Ethics of Data Mining. [Electronic] RIT Scholar Works. Available: http://scholarworks.rit.edu/article/441?utm_source=scholarworks.rit.edu%2Farticle%2F441&utm_medium=PDF&utm_campaign=PDFCoverPage [2020, 20 Jan].
21. Confessore, N. 2021. *Cambridge Analytica and Facebook: The Scandal and the Fallout So Far (Published 2018)*. [Online] NYtimes.com. Available: https://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html [2021, 2 August].
22. Congressional Research Service (CRS). 2020. *Artificial Intelligence and National Security.* CRS. 1-13.
23. Connected Health. n.d. *Policy Principles for Artificial Intelligence in Health.* [Online]. Available: https://actonline.org/wp-content/uploads/Policy-Principles-for-AI.pdf [2021, 12 August].
24. Dajose, L. 2019. Biology and Big Data: How computational biology is shaping the future of health and privacy. *Caltech Magazine* [Online]. Available: https://magazine.caltech.edu/post/biology-and-big-data.
25. Deo, R. 2015. Machine learning in medicine. *Circulation*, 132(20):1920-1930.
26. Department of Defense (DoD) & Defense Innovation Board (DIB), n.d. *AI Principles: Recommendations on the Ethical Use of Artificial Intelligence*.
27. Department of Homeland Security & Director of National Intelligence, 2018. *AI: Using Standards to Mitigate Risks.* 13-14.
28. Department of homeland Security (DHS) & Director of National Intelligence. 2018. *AI: Using Standards to Mitigate Risks.* 13-19.
29. Edwards, J. 2021. *Predictive analytics: Transforming data into future insights.* [Online] CIO. Available: https://www.cio.com/article/3273114/what-is-predictive-analytics-transforming-data-into-future-insights.html [2021, 4 August].

30. Evans, N. & Inglesby, T. 2019. Biosecurity and public health ethics issues raised by biological threats, in ANNA C. Mastroianni, Jeffrey P. Kahn, and Nancy E. Kass (eds.). *The Oxford Handbook of Public Health Ethics.* [Online]. 773-785. Available: https://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780190245191.001.0001/oxfordhb-9780190245191-e-67 [2021, 11 August].
31. Fabian, B., Ermakova, T. & Lentz, T. 2017. Large-scale readability analysis of privacy policies. In: *IEEE/WIC/ACM International Conference on Web Intelligence.* Leipzig, Germany.
32. Gaffney, J., Tibebu, R., Bart, R., Beyene, G., Girma, D., Kane, N., Mace, E., Mockler, T., Nickson, T., Taylor, N. & Zastrow-Hayes, G., 2020. Open access to genetic sequence data maximizes value to scientists, farmers, and society. *Global Food Security*, 26:100411.
33. Gerke, S., Minssen, T. & Cohen, G. 2020. Ethical and legal challenges of artificial intelligence-driven healthcare, in Adam Bohr & Kaveh Memarzadeh (eds.). *Artificial intelligence in Healthcare*, 295-336. doi: 10.1016/B978-0-12-818438-7.00012-5.
34. Goodman, K. 2001. Toward striking a balance in bioinformatics. *AMA Journal of Ethics*, 3(3).
35. Grundmann, O. 2014. The current state of bioterrorist attack surveillance and preparedness in the US. *Risk Management and Healthcare Policy*, [Online] 177. Available: https://www.dovepress.com/the-current-state-of-bioterrorist-attack-surveillance-and-preparedness-peer-reviewed-fulltext-article-RMHP [2021, 11 August].
36. Habeeb, A. 2017. *Introduction to Artificial Intelligence.* [Online]. Available: https://www.researchgate.net/publication/332548325_Artificial_Intelligence_Definition_Ethics_and_Standards#:~:text=Artificial%20intelligence%20(AI)%20is%20based,learning%20(Saleh%2C%202019)%20 [2021, 12 August].
37. Henig, R. 2020. Experts warned of a pandemic decades ago. Why weren't we ready? [Online]. *National Geographic*. Available: https://www.nationalgeographic.com/science/article/experts-warned-pandemic-decades-ago-why-not-ready-for-coronavirus [2021, 9 August].
38. Herzog, C. 2019. Technological opacity of machine learning in healthcare. In: *Weizenbaum Conference*. Berlin: Social Science Open Access Repository (SSOAR).
39. Howard, P., Rainie, L. & Jones, S. 2001. Days and nights on the Internet: The impact of a diffusing technology. *American Behavioral Scientist*, 45(3):383-404.
40. Imberman, S. 2001. *Effective Use of the KDD Process and Data Mining for Computer Performance Professionals.* [Online]. New York: City University. Available: https://www.researchgate.net/publication/221445402_Effective_Use_of_the_KDD_Process_and_Data_Mining_for_Computer_Performance_Professionals [2021, 11 August].
41. Institute of Medicine (US) Forum on Microbial Threats. 2007. *Ethical and Legal Considerations in Mitigating Pandemic Disease*. Washington, DC: National Academies Press (US).
42. Isaksson, M. 2019. *Artificial Intelligence: The implications of current technological developments for Harry Collins epistemological theory and his critique of the possibilities of a general AI.* Master's Thesis. Goteborgs Universitet: INSTITUTIONEN FÖR FILOSOFI, LINGVISTIK OCH VETENSKAPSTEORI.

43. Kaliyappan, K., Palanisamy, M., Govindarajan, R. & Duraiyan, J., 2012. Microarray and its applications. *Journal of Pharmacy and Bioallied Sciences*, 4(6):310.
44. Kernchen, R. 2020. *Coping with Complexity in Biological Threat Assessment*. Social Science Research Network (SSRN) [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3634621 [2021, 12 August].
45. Khorsheed, O., Madbouly, M. & Guirguis, S. Search engine optimization using data mining approach. *International Journal of computer Engineering and Applications,* IX(VI).
46. Klar, R. & Lanzerath, D. 2020. The ethics of COVID-19 tracking apps – challenges and voluntariness. *Research Ethics.* [Online]. 16(3-4):1-9. Doi: 10.1177/1747016120943622
47. Klingler, C., Silva, D., Schuermann, C., Reis, A., Saxena, A. & Strech, D. 2017. Ethical issues in public health surveillance: a systematic qualitative review. *BMC Public Health*, 17(1):4-5.
48. Koh, J. & Brusic, V. 2005. Database Warehousing in bioinformatics. *Bioinformatics Technologies.* 45-62.
49. Kumar, V. & Garg, M. 2018. Predictive analytics: A review of trends and techniques. *International Journal of Computer Applications,* 182(1):31-37.
50. Loike, J. & Fischbach, R. 2013. Ethical challenges in biodefense and bioterrorism. *Journal of Bioterrorism & Biodefense.*
51. Luscombe N.M., Greenbaum D. & Gerstein M. 2001. What is bioinformatics? A proposed definition and overview of the field. *Methods Inf Med.* 40(4):346-358. PMID: 11552348.
52. Madria, S., Bhowmick, S., Ng, W. & Lim, E. 1999. Research Issues in Web Data Mining. In: *Data Warehousing and Knowledge Discovery: First International Conference*. [online], Research Collection School of Information Systems. Available: https://ink.library.smu.edu.sg/sis_research/977 [2021, 29 July].
53. Marr, B. 2018. *What is Deep Learning AI? A Simple Guide With 8 Practical Examples.* Available: https://www.forbes.com/sites/bernardmarr/2018/10/01/what-is-deep-learning-ai-a-simple-guide-with-8-practical-examples/?sh=1c4460848d4b [2021, 11 August].
54. Muller, V. 2020. Ethics of Artificial Intelligence. In: A. Elliot (eds). *The Routledge Social Science Handbook of AI.* [Online] 8(2). Available: https://policyreview.info/articles/analysis/technology-autonomy-and-manipulation [2021, 12 August].
55. NATO Parliamentary Assembly. 2021. *Biological Weapons: Technological Progress and The Spectre of Bioterrorism in The Post-COVID-19 ERA.* NATO.
56. Norval, C. & Henderson, T. 2017. Contextual consent: Ethical mining of social media for health research. *ArXiv, abs/1701.07765.*
57. Nsubuga, P., White, M., Thacker, S., Anderson, M., Blount, S., Broome, C., Chiller, T., Espitia, V., Imtiaz, R., Sosin, D., Stroup, D., Tauxe, R., Vijayaraghavan, M. & Trostle, M. 2006. Public health surveillance: A tool for targeting and monitoring interventions. *Disease Control In Developing Countries*. 2nd ed. World Bank.
58. Nyholm, S. & Gordon, J. 2021. Ethics of Artificial Intelligence. [Online]. Research Gate. Available: https://www.researchgate.net/publication/349467117_Ethics_of_Artificial_Intelligence [2021, 12 August].

59. Organization for Economic Co-Operation and Development (OECD). 2009. *OECD Guidelines on Human Biobanks and Genetic Research Databases*. OECD Publishing. 4-7.
60. O'Sullivan, D. 2021. *Half a billion Facebook users' information posted on hacking website, cyber experts say*. [Online]. CNN. Available: https://edition.cnn.com/2021/04/04/tech/facebook-user-info-leaked/index.html [2021, 4 August].
61. Panch, T., Mattie, H. & Atun, R., 2019. Artificial intelligence and algorithmic bias: implications for health systems. *Journal of Global Health*, 9(2):2.
62. Paschen, U., Pitt, C. & Kietzmann, J. 2020. Artificial intelligence: Building blocks and an innovation typology. *Business Horizons*, 63(2):147-155.
63. Qi, X. & Zong, M. 2012. An overview of privacy-preserving data mining. *Procedia Environmental Sciences*, 12:1341-1347.
64. Rani, S. & Rao, S. 2021. Study and analysis of noise effect on Big Data Analytics. *International Journal of Management, Technology and Engineering*, 8(7):5841.
65. Riahi, Y. & Riahi, S. 2018. Big Data analytics: Concepts, types and technologies. *International Journal of Research and Engineering*, 5(9):524-528.
66. Riedel, S., 2004. Biological Warfare and Bioterrorism: A Historical Review. *Baylor University Medical Center Proceedings*, 17(4):400-406.
67. Robbins, S. 2019. A misdirected principle with a catch: Explicability for AI. *Minds and Machines*, 29(4):495-514.
68. Rodrigues, R. 2015. *Principles and Approaches in Ethics Assessment: Dual-use in research.* [Online]. Available: https://satoriproject.eu/media/1.g-Dual-use-in-research.pdf [2021, 12 August].
69. Sahu, H. Shrma, S. & Gondhalakar, S. 2021. A brief overview of data mining survey. *Journal of Computer Sciences and Applications*, 1(3):114.
70. Saleem, K., Derhab, A., Al-Muhtadi, J. & Shahzad, B. 2015. Human-oriented design of secure Machine-to-Machine communication system for e-Healthcare society. *Computers in Human Behavior*, 51:977-985.
71. Saleh, Z. 2019. *Artificial Intelligence Definitions.* [Online]. Available: https://www.researchgate.net/publication/332548325_Artificial_Intelligence_Definition_Ethics_and_Standards#:~:text=Artificial%20intelligence%20(AI)%20is%20based,learning%20(Saleh%2C%202019)%20. [2021, 12 August].
72. Sarkar, R., Abbass, H. & Newton, C. 2002. *Heuristics and Optimization for Knowledge Discovery*. Hershey: Idea Group Pub:1-2.
73. Schelter, S., Biessmann, F., Januschowski, T., Salinas, D., Seufert, S. & Szarvas, G., 2018. On Challenges in machine learning model management. *IEEE Data Eng. Bull.*, 41:1.
74. Searle, J. 1980. Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3):417-424.
75. Shachar, C., Gerke, S. & Adashi, E. 2020. AI Surveillance during Pandemics: Ethical Implementation Imperatives. *Hastings Center Report*, 50(3):18-21.
76. Schafer, A. 2011. *Privacy: A Philosophical Overview*. 6-9
77. Shah, A. Gulati, R. 2016. Privacy-preserving data mining: techniques, classification and implications – A Survey. *International Journal of Computer Applications*, 137(12):40-46.

78. Shahid, N., Rappon, & Berta, W. 2019. Applications of artificial neural networks in health care organizational decision-making: A scoping review. *PLOS ONE*, 14(2), p.e0212356.
79. Sharma, A. 2021. Artificial intelligence in healthcare. *International Journal of Humanities, Arts, Medicine and Science*, 5(1).
80. Surden, H. & Williams, M. 2016. Technological opacity, predictability, and self-driving cars. *SSRN Electronic Journal*, 38.
81. Susser, D., Roessler, B. & Nissenbaum, H. 2019. Technology, autonomy, and manipulation. *Internet Policy Review*. [Online]. 8(2). Available: https://policyreview.info/articles/analysis/technology-autonomy-and-manipulation [2021, 4 August].
82. Taylor-Sakyi, K. 2016. *Big Data: Understanding Big Data*. [Online]. Research Gate. Available: https://www.researchgate.net/publication/291229189_Big_Data_Understanding_Big_Data/citations [2021, 29 July].
83. The Great Hack [film]. 2019. Park city, Utah: Netflix.
84. Totschnig, W. 2020. Fully autonomous AI. *Science and Engineering Ethics*, 26(5):2473-2485.
85. Tsamados, A., Aggarwal, N., Cowls, J., Morley, J., Roberts, H., Taddeo, M. & Floridi, L. 2021. The ethics of algorithms: key problems and solutions. *AI & SOCIETY*.
86. United states. Executive Office of the President. 2018. National Biodefense Strategy. [Online]. 8-9. Available: https://www.hsdl.org/?view&did=815921
87. Van Wel, L. & Royakkers, L. 2004. Ethical issues in web data mining. *Ethics and Information Technology*, 6(2):129-140.
88. Vayena, E., Blasimme, A. & Cohen, I. 2018. Machine learning in medicine: Addressing ethical challenges. *PLOS Medicine*, 15(11).
89. Vayena, E. & Madoff, L. 2019. Navigating the ethics of Big Data in public Health. *The Oxford Handbook of Public Health Ethics*, 353-367.
90. Vogel, K. 2019. Big Data and Biodefense: Prospects and pitfalls. *Defense Against Biological Attacks*, 1:297-315.
91. Weinbaum, C., Landree, E., Blumenthal, M., Piquado, T. & Gutierrez, C. 2019. *Ethics in scientific research*. Santa Monica, California: RAND Corporation:5-29.
92. World Health Organization (WHO). 2017. *WHO Guidelines on Ethical Issues in Public Health Surveillance*. Geneva, Switzerland: World Health Organization, 10-14.
93. Wood, A. 2014. Coercion, Manipulation, Exploitation. *Manipulation*, 17-50.
94. Wooley, J.C. & Lin, H.S. (eds.) 2005. On the nature of biological data. *Catalyzing Inquiry at the Interface of Computing and Biology* [Electronic]. Washington (DC): National Academies Press (US). Available: https://www.ncbi.nlm.nih.gov/books/NBK25464/
95. Yu, P. 1999. Data Mining and Personalization Technologies. In: *Database Systems for Advanced Applications*. Hsinchu, Taiwan: IEEE.
96. Zayegh, A. & Al Bassam, N., 2018. Neural network principles and applications. *Digital Systems*.

97. Zestcott, C., Blair, I. & Stone, J. 2016. Examining the presence, consequences, and reduction of implicit bias in health care: A narrative review. *Group Processes & Intergroup Relations*, 19(4):528-542.
98. Zou, D., Ma, L., Yu, J. & Zhang, Z. 2015. Biological databases for human research. *Genomics, Proteomics & Bioinformatics*, 13(1):55-63.