



**Erasmus
Mundus**

Artificial Intelligence and Human Security:

AI Strategy Analysis

July, 2019

Student IDs

University of Glasgow: 2247793

Dublin City University: 17116155

Charles University: 22035308

Presented in partial fulfilment of the requirements for the

Degree of

International Master in Security, Intelligence and Strategic

Studies

Word Count: 21401

Supervisor: Dr. Eamonn Butler

Date of Submission: 25/7/19



CHARLES UNIVERSITY

Abstract

The capabilities of Artificial Intelligence are growing quickly while the technology widely proliferates. Many risks and benefits are emerging as it becomes more powerful. This thesis aims to identify the set of risks and benefits emerging from AI that affect Human Security and the implications of how they are being acknowledged and addressed in government strategies. A definition of Human Security was chosen after a review of the literature. Due to its individual country agnosticism and useful split into 7 dimensions (Economic, Food, Health, Environmental, Personal, Community, and Political Security) the 1994 UN Human Development report's definition was chosen.

Inductive document analysis was performed using the 7 dimensions of Human Security as an analytical framework. Twenty-four documents from 18 different governments were identified, and then keyword analysis was implemented on each to highlight their relevancy to each of the 7 dimensions. This allowed for the creation of a binary set of data showing the coverage of those 7 dimensions in each document and the overall coverage of each dimension across all documents. The relevancies that were identified were then qualitatively and quantitatively analysed by addressing the trends seen across the documents and each document's relevancy to the 7 dimensions of Human Security. This provides a picture of how the threats and opportunities of AI are being acknowledged and addressed by governments, shedding light on the unique insights of each document and trends across governments.

The use of the 7 dimensions of Human Security as an analytical framework has potential for analysis of other controversial topics. This research also shows the development of both traditional and emerging security dilemmas. The traditional dilemma is tied to the potential of a lethal autonomous weaponry arms race spiralling out of control. The emerging dilemma requires a shift in mind-set – the 7 dimensions of Human Security can be at odds with one another; the securing of one dimension may cause insecurity in another, creating tension between them. Politicians and technologists with an interest in the creation of beneficial AI may find this data and analysis useful because of its in-depth breakdown of AI trends in government strategies relative to Human Security risks and benefits. AI must be built with fairness, accountability, and transparency in mind to ensure it avoids disparate impacts and benefits everyone equitably.

Table of Contents

Abstract.....	2
Table of Contents	3
Acronyms	6
1. Introduction.....	7
1.1 A Brief History of Artificial Intelligence	7
1.2 AI Definitions.....	8
1.2.1 AI Power Descriptors	9
1.2.2 Some Important Types of AI Systems.....	10
1.3 So What?	11
2. Literature Review	13
2.1 Human Security Literature Review	13
2.2 Conclusion.....	18
3. Theory and Research Design	22
3.1 Introduction	22
3.2 Theoretical Grounding	22
3.2.1 Pragmatism.....	22
3.2.2 Human Security as an Analytical Framework.....	23
3.3 Research Design	24
3.3.1 Qualitative Document Analysis.....	24
3.3.2 Search Strategy	26
3.3.3 Reliability	27
3.3.4 Limitations, Generalizability, and Ethics	27
4. Data	29
4.1 Introduction	29

4.2 AI Policy and Strategy Data.....	29
4.2.1 United States (US).....	29
4.2.2 Canada.....	32
4.2.3 Japan.....	33
4.2.4 Singapore.....	35
4.2.5 China.....	37
4.2.6 United Arab Emirates (UAE).....	39
4.2.7 Finland.....	40
4.2.8 Taiwan.....	42
4.2.9 Italy.....	42
4.2.10 France.....	44
4.2.11 European Union (EU).....	48
4.2.12 United Kingdom (UK).....	52
4.2.13 South Korea.....	53
4.2.14 Sweden.....	54
4.2.15 India.....	55
4.2.16 Mexico.....	58
4.2.17 G7.....	59
4.2.18 Germany.....	60
4.3 Governments Developing Still.....	62
4.3 Tables and Graphs of Relevancies.....	63
4.3.1 Main Data.....	63
4.3.2 Table and Graph of Variance.....	64
5. Analysis of AI Impacts on Human Security.....	65
5.1 Human Security Impacts.....	65

5.1.1 Economic Security.....	65
5.1.2 Food Security.....	68
5.1.3 Health Security	69
5.1.4 Environmental Security	71
5.1.5 Personal Security	74
5.1.6 Community Security.....	78
5.1.7 Political Security	81
5.2 Summary and Discussion of AI Strategy Findings	85
6. Conclusion	89
6.1 Summary of Research and Outcomes	89
6.2 Implications and Future Recommendations.....	91
6.2.1 AI and Human Security – Traditional and Emerging Security Dilemmas	92
6.3 Final Thoughts.....	92
References.....	94

Acronyms

AGI – Artificial General Intelligence

AI – Artificial Intelligence

AIaaS – AI-as-a-Service

AISG – AI Singapore

ANI – Artificial Narrow Intelligence

ASI – Artificial Superintelligence

AWS – Autonomous Weapon Systems

CIFAR – Canadian Institute For Advanced Research

CIHR – Canadian Institutes of Health Research

HIC – human-in-command

HITL – human-in-the-loop

HOTL – human-on-the-loop

IDRC – International Development Research Centre (of Canada)

R&D – Research and Development

UN HDR – United Nations Human Development Report

1. Introduction

Artificial Intelligence (AI) used to be a subject reserved for Science Fiction stories, but today it surrounds us, many times unseen or unrecognized. It is a powerful technology that has changed our world already in many ways. Some of these changes have impact in trivial ways, such as cell phone helpers, while others are making decisions that affect lives and livelihoods, like the algorithms that assist in deciding who gets parole from predicted recidivism rates. Many of these algorithms have issues with bias leading to disparate impacts upon differing groups of people. As AI grows in strength, it has the potential to have a great positive effect on or to cause intense harm to Human Security if development is not properly managed with multifaceted concerns beyond economic and scientific gain in mind.

A brief history of AI and some definitions within the AI ecosphere will be laid out first in order to establish an understanding of where the technology has come from and where it is now. Then the more contentious debate surrounding the concept of Human Security will be analysed in the Literature Review. Following this, a review of current governmental AI strategies will be made in the Data chapter, with specific cases and applicability to Human Security highlighted in the Empirical Analysis chapter. Finally, this thesis will conclude and highlight the addition that it has made to the overall body of literature and theory related to AI strategy and Human Security.

1.1 A Brief History of Artificial Intelligence

In order to sufficiently forecast the possible futures surrounding AI's impact on Human Security, we must first understand the past and present of the technology. The field of Artificial Intelligence has become mainstream and made great strides in the last two decades, but its origins go back to the mid-1900s. A couple of the outstanding names of those times include Alan Turing and John McCarthy. Turing published a ground breaking paper in the psychology and

philosophy journal *Mind* that delved into the possibility of creating machines that think, and even produced an idea for a game that machines could play to prove their capabilities (aptly now called the “Turing Test”).¹ On the other hand, McCarthy is often called the “Father of AI” and coined the term “Artificial Intelligence” with his organization of the Dartmouth Summer Research Project on Artificial Intelligence in 1956.²

There were periods of great strides in AI throughout the decades, often called AI Waves; these alternated with periods of stagnation and setbacks – AI Winters. The 1st Wave began in 1956 with the Dartmouth Project, and ended around 1974 with the start of the 1st AI Winter. The bright optimism of that first wave had led to high expectations that were not met. The 2nd Wave spanned the beginning to late 1980s with the emergence of Expert Systems – complex programs coded with specific domain knowledge to help in very narrow, specific areas. The limitations of these systems led to the 2nd Winter, which was overcome as we entered the 3rd Wave in the early 90s when some of the original goals of the 1st Wave were finally achieved. This is the Wave that we are still in today.³ The opinions vary wildly right now on whether there is another Winter coming soon,⁴ or if this Wave will continue for the foreseeable future.⁵

1.2 AI Definitions

“Artificial Intelligence” is really an umbrella term for a set of technologies that attempt to act and make decisions autonomously. AI is not robotics, though they are often used together to create systems that can interact with the physical world (like autonomous vehicles). It is akin to the difference between the mind and the body in humans. In this section a basic, non-exhaustive overview of some

¹ Turing, 1950

² McCarthy et al., 1955

³ dudeperf3ct, 2018

⁴ Chishti, 2018; Piekniowski, 2018

⁵ Knight, 2016; Faggella, 2019

of the most prevalent terms within the AI ecosystem will be presented with the intention of giving the reader a baseline of understanding about the subject as it currently stands.

1.2.1 AI Power Descriptors

Artificial Narrow Intelligence (ANI):

ANI can match or beat a human in a specific domain area. This involves examples like AlphaGo, an artificial system that managed to beat the world champion of the extremely complex game of Go in 2016; not only did AlphaGo win, it won in such ways that made professionals who study the game rethink how the game should be played best.⁶

Artificial General Intelligence (AGI):

This is the next level of AI development, when it is as smart as a human across all domains. This involves general capabilities in areas like language and image processing, as well as reasoning and understanding. We have not yet reached this point with any AI.

Artificial Superintelligence (ASI):

Nick Bostrom, one of the leaders in AI thought, describes superintelligence as “an intellect that is much smarter than the best human brains in practically every field, including scientific creativity, general wisdom and social skills.”⁷ Superintelligence can be anything from a little smarter than humans to thousands of times more intelligent. Imagine how we as humans can manipulate animals less intelligent than us, and you can begin to see some of

⁶ DeepMind, n.d.

⁷ Bostrom, 1997

the issues with creating an intelligence higher than our own, and how it is important to have its goals aligned with those of humanity. There is an entire area of study dedicated to this called the AI Alignment Problem.

Weak/Strong AI: Alternative names for ANI and AGI, respectively.

1.2.2 Some Important Types of AI Systems

Supervised Learning*:

AI that is trained on a subset of training data, to be able to extrapolate what it learns to the full set of data afterwards.

Unsupervised Learning*:

AI that naturally finds similarities in the data and classifies things via those perceived correlations.

Genetic Algorithm:

Programs written that mimic genetic evolution by producing different possible algorithmic solutions, rating how well they fit to the desired outcome, and then “mating” the best solutions together for potentially better outcomes (and repeating this until an acceptable solution is met).

Neural Network:

A simulated system mimicking how the brain’s neurons process information by strengthening and weakening pathways between those neurons. There are many subtypes of these networks, and the term ‘deep learning’ is often paired with them.

*There are also hybrids of supervised and unsupervised learning systems.

1.3 So What?

For many decades AI remained primitive; scientists were more concerned with hitting new milestones and optimistic outcomes than what potential negative effects AI might have in the long run. The scientists' efforts were aimed to prove if the creation of true intelligence was possible in machines, not on whether it should be done or how it could affect humanity as a whole. This is understandable, as the field was in its adolescence, and the future seemed bright with possibilities. Each AI Winter has tempered the overenthusiastic optimism with a dose of reality; this has helped to bring more of the negative what-ifs to the discussions.

We have only recently achieved ANI, and while these AI threaten jobs, they are more limited in their risks. The emergence of AGI is inevitable on the path we are on. ASI could be quick upon us after that, far in the future, or even never (though the author personally believes that it is when, and not if, it will come to pass). Because of the technical nature of these topics, many modern policymakers and government officials don't truly understand that which they aim to control. Much of the current landscape of AI strategies concentrate heavily on pumping money into the AI ecosystem with broad goals, leaving the details to the researchers that are contracted to do the work.

Even the risk debates surrounding AI today are often limited in scope, usually concentrating on narrow areas like Autonomous Weapon Systems or the replacement of humans in the workforce. This piece of work aims to bridge that gap by taking a more holistic view of the impacts of AI upon Human Security. It intends to raise the concerns of Human Security while tempering them with the good that is also being done. This dissertation seeks to address the intersectionality of Human Security and Artificial Intelligence within the context of government strategies and projects by analysing their content and the landscape of variances among them.

This research will synthesize and analyse the ways that governmental strategies are guiding AI away from risk to and towards benefits for Human

Security. It aims to answer the question: what is the set of risks and benefits emerging from AI that can affect Human Security, and what are the implications of how they are being acknowledged and addressed in government strategies?

This will be done through a series of objectives:

1. To summarize the history of developments in AI and Human Security;
2. To review and analyse cases of relevant risks and benefits to Human Security that have emerged with the proliferation of AI;
3. To identify the commonalities and outliers in the implications and implementations of governmental AI strategies in the light of Human Security.

By the conclusion of this thesis, a broad understanding of the intersectionality of Human Security, AI, and government strategies will be developed.

2. Human Security Literature Review

2.1 Definitions of Human Security

The notion of Human Security originates from the 1994 United Nations Human Development Report (UN HDR) as an intentional widening and deepening of the security agenda away from the narrow nation-state oriented National Security paradigm that had been standard for so long. The report describes Human security as a universal concern with interdependent components that is people-centred and easier to ensure through prevention rather than intervention. Human Security in this light is broken down into two areas, “freedom from fear and freedom from want.”⁸ In more recent years the concept has been expanded to include “freedom from indignities,” the idea that Human Security also includes protecting people from humiliating circumstances, by S. Tadjbakhsh in various literature.⁹ The UN breaks human security down into 7 intertwined facets: Economic, Food, Health, Environmental, Personal, Community, and Political Security. These different categories impact each other in many ways. For example, Food Security can be undermined by a lack of Economic Security, which can lead to issues in Health Security.

While the UN definition of Human Security may have been groundbreaking, other definitions have arisen and deviated from it. Two of the first cases of this are the Canadian and Japanese styles of Human Security, which correspond to narrower and broader conceptions of the idea respectively. Canada, for example, concentrates mainly on the ‘freedom from fear’ aspect. This can be seen looking back at the Canadian championing of the ‘Ottawa Treaty’ (more properly known as the Convention on the Prohibition of the Use, Stockpiling, Production and Transfer of Anti-Personnel Mines and on their Destruction), and their argument for humanitarian intervention in support of Human Security over State Sovereignty in the case of extreme abuse such as genocide.¹⁰ Some, such as

⁸ United Nations, 1994, p. 24

⁹ Tadjbakhsh, 2010, p. 17; Tadjbakhsh, 2014, p. 1

¹⁰ Bosold and Werthes, 2005

Global Health and Human Rights academics Spiegel and Huish, criticize this Canadian approach to Human Security, claiming that the state is putting military and trade interests over humanitarian goals.¹¹

The Swiss take on Human Security is close to the Canadian philosophy with some minor variances. Switzerland's Human Security Division breaks their system down into peace, human rights, humanitarianism, and migration policy. The first three concentrate on protection from violence in various ways (such as resolution of violent conflicts, the abolition of the death penalty, and dissolution of illicit arms trading), while the last concentrates more on human dignity and combating human trafficking.¹²

Norway's Human Security outlook originated in a similar vein, which is unsurprising considering that both Canada and Norway helped to found the Human Security Network (AKA the Lysøen Group) – an assembly of countries and non-governmental organizations interested in advancing the Human Security agenda.¹³ In more recent times, Norway has expanded its conceptualization of human security to include campaigning for the inclusion of education via calling for the end of violence in schools with contributions to the “Education Cannot Wait” fund and promotion of the “Safe Schools Declaration.”¹⁴

Japan, in comparison, followed closer to the original 1994 United Nations definition of Human Security in the need for both physical protection from violence as well as protection from health, economic, and political threats. The Japanese Human Security outlook stresses the importance of not only people's survival, but also their dignity and ability to thrive, and tries to enforce this by addressing structural issues that lead to human insecurity.¹⁵ One of the ways that Japan has pursued these goals is through instigating a Trust Fund for Human

¹¹ Spiegel and Huish, 2011

¹² Swiss Federal Department of Foreign Affairs, 2018

¹³ Alkire, n.d., p. 21

¹⁴ Juul, 2019

¹⁵ Bosold and Werthes, 2005

Security through the UN; others include policy dialogues and conferences, bi- and multi-lateral conversations with other countries, and establishing groups that aim to spread the concept and implementation of Human Security.¹⁶

Under this wider and deeper conception of Human Security, many disparate aspects of our world come together because of their impact on individuals and communities. One such aspect is climate change. Climate change has the potential to affect all seven of the various categories of security identified by the aforementioned 1994 UN HDR. For a comprehensive analysis on the intertwining of climate change and Human Security, see Adger et al.¹⁷ Climate change has been worsening, making its importance for the Human Security agenda even greater. For the 25th anniversary of the 1994 report that introduced the concept of Human Security to the international agenda, UN member states gathered together to discuss the continued relevance of the philosophies and their relevance to the UN Sustainable Development Goals. Climate change was but one of the subjects discussed.¹⁸

2.2 Criticisms of Human Security

There has been a healthy amount of criticism aimed at Human Security. One of the main critiques of Human Security is that it is simply too broad of a concept to be practically useful for policy or analysis. For example, Roland Paris believes the ideas behind Human Security work well as a rallying cry to unite diverse stakeholders and as a political campaign for specific goals such as Canada's championing of the ban on land mines; however, Paris thinks that Human Security is too vague to be meaningful for policymakers and academics to use for guidance.¹⁹ Similarly, Khong feels that Human Security is a nice, idealistic notion, but overall not very useful to those individuals it is claiming to

¹⁶ Ministry of Foreign Affairs of Japan, 2009

¹⁷ Adger et al., 2014

¹⁸ Risse, 2019

¹⁹ Paris, 2001

secure when they are in bad situations within non-democratic countries. Rather, Khong thinks that Human Security is nothing but a false hope in those cases.²⁰

Amid complaints of the vagueness and imprecision of the concepts of Human security, an attempt to quantify the ideas behind it into something measurable emerged with the Human Security Index. This is made up of 3 main components: Economic, Environmental, and Social Fabric Indexes. These are further broken down into specific assessable indicators such as “equality of income distribution,” “environmental sustainability,” and “education and info empowerment.”²¹ This quantification of Human Security has not been used for much as of yet, however.

The arbitrariness of the application of Human Security to real world issues is another critique of the concept. As Alkire asks, “When the potential set of critical and pervasive threats is so wide, by what criteria is a small subset of these chosen for consideration?”²² Alkire calls to attention that often the Human Security idea is just used as an excuse to further the already standing agendas of states rather than actually being used against human insecurity.²³ How states could even rank variables such as likeliness of the length of suffering, potentiality of successful intervention, or which suffering holds the greatest threat is not made clear, however. Perhaps the aforementioned Human Security Index could be of use here.

Another line of criticism against Human Security comes from a feminist perspective. This critique develops by challenging the typically taken-for-granted definition of “human” within the ideas of Human Security. Marhia raises the problem of historical systemic separation and mistreatment of women which has led to the need for such categories as “women’s rights” separate from “human rights,” implying that human rights historically has not been inclusive of women

²⁰ Khong, 2001

²¹ Human Security Index, n.d.

²² Alkire, n.d., p. 22

²³ Ibid.

as truly human.²⁴ Similar arguments can be made for things like “black rights” or “queer rights” not having been included in the “human” of Human Security due to the history of oppression and mistreatment of those categories of individuals.

A post-colonial critique of Human Security can also be made. Because of the evolution of the concept primarily in the Global North with predominantly Western ideals in mind, Human Security does not take into account the customs and culture of many of the societies of the Global South.²⁵ This can lead to a neo-colonial effect that stems from imposing Western values upon those non-Western countries in the name of securing individual humans. This ties into the criticisms that stem from the worry that Human Security is overly Euro-centric and infringes too easily on the rights to self-sovereignty of non-Western states. McCormack, for example, believes that the implementation of Human Security actually disempowers the citizens of less developed or weaker states rather than empowering them by reinforcing the existing power inequalities.²⁶

Buzan, Wilde, and Wæver unsurprisingly have critical opinions of Human Security, considering their history of dislike for securitization and advocacy for desecuritization into normal politics.²⁷ Buzan has been the most directly critical; he is of the mind that conceptually Human Security adds no analytical value; instead he believes that the agenda is better served by the already existent human rights conceptualization. Under this outlook, Human Security is simply reductionist and serves no real purpose other than as a buzz-phrase meant to grab more attention and power from spheres of influence that it doesn't belong in.²⁸ This falls in line with the arguments from the post-colonial and sovereignty critics and their belief that Western states are using Human Security in a way that is overstepping in their grab for power.

²⁴ Marhia, 2013

²⁵ Shani, 2017

²⁶ McCormack, 2008

²⁷ Buzan, Wilde, and Wæver, 1997

²⁸ Buzan, 2004

As a final line of criticism, Christie believes that Human Security has simply become the new orthodoxy despite its origins as a critical, anti-orthodoxy ideal. When it first emerged Human Security was a shining beacon for the critical security school of thought. However, Christie thinks “the moment where human security might have served as a heterodox challenge to mainstream security practices has passed, and the concept has itself become a new orthodoxy.”²⁹ He fears that the militarization of aid that has developed through the Human Security narrative may be irreversible. However, Christie still thinks that Human Security has a lot of positive potential and is being used in many good ways for narrow goals despite the loss of its initial promise as a critique of the international Global North based, state-oriented, patriarchal security regime status quo.³⁰

2.3 Conclusion

Human Security is a widely contested concept. It is closely tied to the ideas of human development and human rights. Across the different ideas that have emerged, there are some aspects that appear to have become generally accepted by the various entities that are using the concept. One of these is that Human Security involves a paradigm shift away from states as the referent objects of security and towards the security of individuals and society on a humanistic, and ethically based level while questioning the traditional state sovereignties. It is also commonly acknowledged that Human Security is a multilateral issue that needs a multilateral response. Another typically accepted reality is that Human Security applies universally across our planet and various ethnic populations. The various world powers’ interpretations of the concept also tend to have a component involving physical safety and protection from violence, while only some extend this to include such things as economic and health factors. Despite the ambiguity of the concept, it has still managed to have a strong impact in many ways.

²⁹ Christie, 2010, p. 170

³⁰ Christie, 2010

The criticisms that emerge in contestation to the concept, though different in substance, are similar at their core. They all attack the application of Human Security as a political tool; they believe that it's being misused in some way, or that it's useless. The 1994 UN HDR's Human Security facets work as an analytical framework that can be used to evaluate the impact of other areas on Human Security; it does this by breaking down the high-level concept into more manageable (and more useful!) separate categories. This is also why the author here chose to use the least politicized version of human security for analysis.

The usefulness of the concept of Human Security, as well as what it should involve, is a widely debated topic. Human Security can be seen as a securitizing speech act upon the concepts of Human Development. Securitization theory states that issues can end up anywhere along the range of non-politicized to politicised to securitized, and that typically those that make it to the extreme securitization end of the spectrum are framed as existential or literal physical threats in one manner or another.³¹ The bulk of securitization theory was developed by Ole Weaver in order to explain the widening of the traditional security agendas beyond the traditional Realist, state-oriented conceptions, and offer a process of de-securitization down into normal politics.³² Some, such as Buzan and Wæver,³³ believe this to be a negative thing; in this case they would say that too many people are pulling concepts that belong in the areas of Human Rights and Development into exceptional politics. However, securitization can be a good thing in some cases, such as the following.

Human Security has direct roots in the ideas of Human Development, and is also tied to Human Rights. This is most evident from its origins from the 1994 UN HDR and the overlap in the applicable subject areas.³⁴ What sets Human Security apart, however, is that “in addition to its emphasis on human well-being,

³¹ Buzan, Wilde, and Wæver, 1997, p. 23-24

³² Wæver, 1995

³³ Buzan, Wilde, and Wæver, 1997

³⁴ United Nations, 1994

human security is driven by values relating to security, stability and sustainability of development gains.”³⁵

Despite criticisms surrounding its politicisation and use to entrench status quo state oriented politics, there is a lot of potential for Human Security’s use as a theoretical analytical framework with which to assess various other intersectional topics, such as the impact of Artificial Intelligence; this is examined more in 3.2.3. There is a wealth of literature arguing over the conception of Human Security. Similarly, there are various presentations of potential and current risks posed by AI. However, there is a distinct gap in the literature where their direct intersection occurs. This dissertation aims to bridge that gap by providing a way to directly link the two. With this explicit link established, government AI strategies can be examined within the light of Human Security to see how governments are acknowledging the threats that are being created and solved with AI development.

A white paper surrounding AI and Human Development published by the Canadian International Development Research Centre (IDRC) concentrates mostly on how AI is and will be impacting low and middle income countries, and the Global South as a whole. Their conclusions surrounding risks and benefits closely mirror those presented in the Analysis section 5.1. However, they also go on to raise some recommendations surrounding the optimisation of AI’s impact. Their ideas include things that will support inclusive, rights-based, and ethical AI systems through regulations and policies that have been well developed and created with multifaceted input.³⁶

The development of policies and strategies such as those being advocated for by the IDRC are key to managing this emerging technology properly. This dissertation will continue on to apply Human Security as framework to analyse current government strategies in relation to AI, rather than just AI directly. AI is

³⁵ United Nations: Human Security Unit, 2009

³⁶ International Development Research Centre, 2018

still in the throes of development, and it is yet to be seen whether or not it will have a net negative or positive impact upon humanity. The application of AI in our modern world is rife with many current and potential risks and benefits – in the end the result will depend on how we develop and manage these technologies. Governmental AI strategies can act as guidance to help keep the risks of AI in check while maximizing the benefits that may be achieved. The 7 facets of Human Security will be applied to the overall risks and benefits of AI, and more specifically to how governments are using their AI strategies in ways that impact Human Security, in the Data (4) and Analysis (5) chapters, in order to answer the research question posed in section 1.3.

3. Theory and Research Design

3.1 Introduction

Artificial Intelligence has so much potential to do good in our world, but it also has a great potential for damage in many ways as well. We are at a critical point for the advancement of this technology. The manner in which AI development is managed in this time will lead to either it becoming a positive or negative revolution in the long term. As such, it is essential that we analyse the ways in which various actors are strategizing around AI, as well as the merits and pitfalls of these strategies. The researcher aimed to do this through pragmatic analysis. Additionally, a definition of Human Security was developed through the securitization of Human Development, then used as an analytical framework with which to look at AI. At a practical level, the method behind the actual process was inductive qualitative document analysis – the benefits of which are discussed in Section 3.3.1.

3.2 Theoretical Grounding

3.2.1 Pragmatism

The researcher rejects the Realist/Relativist dichotomy in favour of a more complex view of the world in which there is both objective truths as well as subjective, that are entangled in ways that cannot be truly separated in practicality. The roots of these beliefs are found in the influence of John Dewey's Pragmatism, and particularly his words in *Chapter X: The Construction of Good* in "The Quest for Certainty."³⁷ Dewey delves into the necessary destruction of the facts/values dichotomy in this chapter; he believed that fact and value, the material and ethereal, were concretely intertwined, and as such that empirical and psychological experience are connected (ergo also practice and theory) by the logical action of inquiry.

³⁷ Dewey, 1929

Human Security is not just fact based, but also value based. Choosing to analyse AI in the light of Human Security, rather than International or State Security, involves valuing humanity's protection from fear/violence, want, and indignities over the protection of states' integrity. It requires a change in the referent object beyond the Realist militaristic, state-centric view. Human Security also is in conflict with Relativism, though, because of its basis on a lack of objective truth and ethical non-absolutism. The ideas of Pragmatism tie in well with the concept of Human Security as used in this work, with its beliefs that the subjective and objective are intertwined allowing for the establishment of objective human truths and values that need to be secured while also going beyond the traditional Realist perspective. Additionally, the dissolution of dualisms found in Dewey's words can also apply to the need for dissolution of the racial in and out groups to truly address Human Security equitably across the world.

3.2.2 Human Security as an Analytical Framework

For this analysis, definitions for two important concepts had to be established: Artificial Intelligence and Human Security. Artificial Intelligence is complex and heavily technical, but well defined. However, Human Security on the other hand is very fuzzy and heavily debated in the literature, so its definition involved more choice in perspective (see Section 2.1 for that discussion). Then with the chosen definitions in mind, a collection of current major AI players' strategic documents on AI development had to be amassed and analysed in light of the Human Security framework, pulling insights from the current and potential risks and benefits surrounding AI. Both individual state governments and collectives such as the EU were considered as relevant AI actors.

In order to analyse the intersectionality of AI policy and Human Security, a definition of Human Security had to be chosen. The 1994 UN HDR definition of Human Security was selected because of both its origination of the concept and the applicability of its 7 dimensions of Human Security as an analytical

framework with which to consider intersectional topics such as Artificial Intelligence. It is people-centred, comprehensive, multi-sectoral, easily contextualized within differing subjects, and has an emphasis on prevention and mitigation.³⁸ A narrower take on Human Security would not serve well as an analytical framework when attempting to get a comprehensive and holistic view on how AI is impacting humanity, and how governments are managing that impact.

3.3 Research Design

3.3.1 Qualitative Document Analysis

The project incorporates primarily qualitative methods of document analysis via keyword analysis. Qualitative methods were chosen due to the nature of the research question and data to be analysed – the examination of the qualities of a body of strategic texts in relation to risks and benefits. Additionally, an inductive approach was taken in order to develop an understanding of the assorted AI strategies being put into place around the world. A deductive approach would not make sense in this case, as the goal is more around developing an understanding of the trends rather than positing a hypothesis about them and trying to find the data to support it.

The main method of the research was document analysis – the act of using bodies of the written word as primary data for analysis. Some advantages of this method are its efficiency, cost-effectiveness, lack of obtrusiveness and reactivity, stability, exactness, and coverage; conversely, document analysis can also have issues with insufficient detail, low retrievability, and biased selectivity.³⁹ These detractions were nearly completely negated in this study. In most of the documents selected for this project, detail was not an issue. Due to the modernity of the topic all of the documents were from recent years and accessible online.

³⁸ Ibid.

³⁹ Bowen, 2009

Additionally, by finding all relative strategies currently produced there was no issue with biased selectivity.

Document analysis was adopted as the main method because of the inherent character of the data being looked at – governmental strategic documents. Another alternative could have been case studies as they would have allowed for similar research, however case studies would need to be constricted to a handful of countries and this would have limited the scope of the analysis to a much narrower view of the landscape of AI strategies contrary to the wishes of the researcher. Interviews could also have been an option for supplemental analysis, rather than relying on the documents themselves, however the researcher concluded that there would not be sufficient time and guaranteed access to the government representatives that would be necessary for such endeavours within the timeframe.

To complement the main qualitative analysis, 2 tables and a bar graph were created to visually show the spread of relevancies to Human Security from throughout the 24 documents that were examined. These can be found at the end of the Data chapter (4.2). Some light statistical analysis was performed on the data for comparative insights related to how individual documents were relevant to all 7 of the dimensions of Human Security, as well as how much relevancy can be found in the individual dimensions in total. A matrix was created with each document listed vertically, and the 7 dimensions listed horizontally. Where each document had a relevancy for a dimension, a green square was left to indicate this. At the end the number of green squares for each document was summed, as well as the total amount for each of the 7 dimensions. These summations were then colour coded on a scale of red-orange-yellow-green to visually indicate the levels of relevancy. The individual documents' relevancies were then counted up and a separate table and graph of their variance was developed for additional analytical value.

3.3.2 Search Strategy

Key terms were chosen for identifying relevant policy documents to be investigated. These included: artificial intelligence, AI, machine learning, algorithmic learning, neural network, and deep learning. These terms were paired with the names of important governmental AI actors' names and the words 'policy' or 'strategy.' Unsurprisingly, only 'artificial intelligence' and 'AI' turned up significant results in strategy documents titles, though the rest also appeared a few times throughout the body of the AI strategy texts. Google and FindPolicy.org were the main search engines used, supplemented by the individual search engines specific to various governmental websites. In addition to these, the Future of Life Institute's Global AI Policy indexes were examined to look for any potential holes in the set of governments being examined.⁴⁰

Shortened variations of the 7 dimensions of Human Security and relevant terms (identified through an initial reading of the documents) were also added as search qualifiers in order to find specific cases for how these governmental entities are strategizing around AI in ways relevant to Human Security. These were:

- **Economic Security:** 'income,' 'workforce,' 'employment' (employment/unemployment)
- **Food Security:** 'food,' 'agricultur' (agriculture/agricultural)
- **Health Security:** 'health' (health/healthy), 'medic' (medical/medicine), 'welfare'
- **Environmental Security:** 'environment' (environment/environmental), 'climate,' 'weather,' 'disaster'
- **Personal Security:** 'safe' (safety/safely), 'damage', 'weapon'
- **Community Security:** 'disadvantage' (disadvantage/disadvantaged), 'humanitarian,' 'divers' (diverse/diversify/diversity), 'gender', 'equality'
- **Political Security:** 'ethic' (ethics/ethical), 'privacy,' 'transparen' (transparency/transparent), 'democra' (democracy/democratic/democratisation)

⁴⁰ Future of Life Institute, n.d.

Where one of these terms was found, the passage was re-read and its relevancy to that dimension of Human Security was documented if valid. In some cases, it was not actually relevant (such as ‘environment’ turning up things relating to the ‘data environment’), and those were ignored.

3.3.3 Reliability

The documents used were primarily pulled directly from the websites of the governments that they represent. As such, it can be assumed that they are reliable at least as much as any official government communication is. In some cases, governments may be saying one thing and doing another, but there is no way to know more than their outwardly expressed intentions. For documents not already in English, as in the case of the Chinese and South Korean strategies, and the Russian publications, the Google Translate and DeepL computer assisted translation tools were both used, in order to ensure the essential core of the translations had similar meanings. The Russian and Chinese documents translated relatively clearly with these tools, but the South Korean document was much harder to decipher. This is an unfortunate, but necessary, limitation of working with documents that exist solely in a foreign language.

3.3.4 Limitations, Generalizability, and Ethics

Due to limitations on time and space, only a subset of strategic AI players were able to be examined – governments. Individual AI companies and non-governmental organizations were considered out of scope for this project. The study attempted to analyse a wide variety of states – both of the East and West. The researcher attempted to extract a set of relevant patterns from across their strategies, with the intention of highlighting the themes found across them relative to how those strategies are relevant to the risks and benefits of AI to Human Security. In an attempt to be holistic, the availability of official AI strategy and

pre-strategy documentation was the main criteria taken into consideration for selection of government samples.

In addition to that, only documents that have a significant percentage of their content focused on AI were considered. This cut out countries like Australia and Denmark whose documents aim at more general technology growth with AI only briefly mentioned, and which had no relevancy to Human Security. Australia was excluded due to its only brief mentions of AI in its document with relation to simply needing to invest.⁴¹ Denmark's document mostly talks about Big Data, AI, and the Internet-of-Things together, only mentioning AI alone when talking about AI patent numbers around the world.⁴²

As this research project does not involve any sort of participant interaction, there are no ethical concerns in that realm. However, that does not mean it is inherently free of any ethical issues. This project involves the identification of what is relevant to Human Security, and the researcher's opinion comes into play some in those decisions. Choosing which keywords to search for and the identification of relevancy is the main ethical concerns. Part of the decision to delve into the study in an inductive manner was to mitigate the influence of the researcher's opinions by reading what is in the data rather than searching for pre-idealized answers as to what should be there.

⁴¹ Australia: Department of Industry, Innovation and Science, 2017

⁴² Denmark: Ministry of Industry, Business and Financial Affairs, 2018

4. Data

4.1 Introduction

This chapter will summarize the current landscape of international governmental AI strategy documents in relation to their relevance to the framework of Human Security dimensions that were established in the 1994 UN HDR – Economic, Food, Health, Environmental, Personal, Community, and Political Security. Some of these strategies are no more than shallow websites with general goals laid out, while others are many-page documents with in-depth, detailed plans. These will be addressed mostly in chronological order; in the cases where a government has more than one relevant strategic document over a period of time they will be addressed together and the initial document’s date will be the indicator of their order. At the end 2 tables and a graph representing the data will be presented for examination.

4.2 AI Policy and Strategy Data

4.2.1 United States (US)

October, 2016 - The National Artificial Intelligence Research and Development Strategic Plan⁴³

The US was the first to publish an AI guiding strategy. In the foreword of this document, it explicitly states that it is meant only as general guidance for US agencies, while all agencies will continue to advance their individual Research and Development (R&D) priorities and manage their own budgeting; thus, it is not widely coordinated. However, the plan *is* meant to guide all Federally-funded R&D both within and without the government (such as Federally-funded academic research).

⁴³ United States: Networking and Information Technology Research and Development Subcommittee, 2016

Human Security Relevance:

Economic:

There is an emphasis on the development of effective human-AI *collaboration* systems, rather than replacing humans with AI, as a way to limit mass job displacement. A need to understand the workforce required to support AI, and find ways to fit that need, is also broached. Recommendations are made for the development of an AI R&D framework and the study of how a sustainable and healthy AI workforce might be created and promoted in the US.

Personal:

AI's safety and security impacts are briefly mentioned as a concern.

Political:

The development of understanding around the legal, societal, and ethical domains related to AI and evaluation with standardized benchmarks are mentioned. These may help limit political security concerns by pre-emptively studying issues and providing a clear framework of evaluation.

*September, 2018 - Assembly Concurrent Resolution No. 215, CHAPTER 206:
Relative to the 23 Asilomar AI Principles⁴⁴*

The state of California formally endorsed the 23 Asilomar AI Principles concentrating on building safe and beneficial AI.

⁴⁴ California, 2018

Human Security Relevance:

Personal:

This suggests AI development race avoidance is necessary to avoid cutting corners. A lethal Autonomous Weapon System (AWS) race is especially emphasized as needing to be avoided. Minimum safety requirements and failure transparency are also mentioned as means to manage risks to Personal Security.

Community:

A goal of shared benefits being spread as widely as possible across humanity is established. Additionally, the document states that superintelligence should only be developed for the common good. These principles should help to ensure equitable AI development for Community Security.

Political:

Politically related concerns include auditable transparency when AI is involved in judicial decisions, alignment with human values, limitations to effects on personal privacy and liberties, and non-subversion of social and civic processes.

*February, 2019 - Summary of the 2018 Department of Defense Artificial Intelligence Strategy: Harnessing AI to Advance Our Security and Prosperity*⁴⁵

The goals of the DOD strategy primarily fall in line with the 2016 strategic plan. The majority of the document is more relevant to state security than Human Security.

⁴⁵ United States: Department of Defense, 2019

Human Security Relevance:

Environmental:

The DOD wants to operationalize AI for disaster relief and recovery efforts. This could help to rebalance ecosystems.

Personal:

They hope to increase equipment safety through predictive maintenance and supplying, and to use AI to lower collateral damage (like civilian casualties).

Community:

They have plans to use AI for humanitarian assistance.

Political:

Developing ethical and safety AI principles for defence, advocating for a global set of guidelines, and promoting transparency in AI are all ideas presented that may help to ensure Political Security.

4.2.2 Canada

*March, 2017 - Pan-Canadian Artificial Intelligence Strategy*⁴⁶

Canada was the first to adopt a country-wide coordinated, fully funded AI strategy. The Canadian Institute For Advanced Research (CIFAR) consulted with AI researchers to help the Canadian government develop this plan, and is managing the project planning and disbursement of funds.

⁴⁶ CIFAR, n.d. (a)

Human Security Relevance:

Community:

They are developing research and training programs that “advance equity, diversity and inclusion in AI”⁴⁷ and working groups that examine the implications for AI in society.

Political:

CIFAR hopes to progress research on the implications of AI development in ethics, policy, and law.

4.2.3 Japan

*March, 2017 - Artificial Intelligence Technology Strategy*⁴⁸

As in the adoption of Human Security theory, Japan followed Canada as the second country to release a coordinated AI strategy. This strategy is more complex and explicit in its plans than the more high-level Canadian one.

Uniquely, this document looks at AI-as-a-Service (AIaaS) that can be provided rather than as product. This is laid out in a 3-step “Industrialization Roadmap.”⁴⁹ It concentrates heavily on the idea that AI is the 4th industrial revolution, and that Japan needs to capitalize on this to stay relevant in the modern, intertwined business and technological worlds.

⁴⁷ Ibid.

⁴⁸ Japan: Strategic Council for AI Technology, 2017

⁴⁹ Ibid., Attachment 1-1

Human Security Relevance:

Economic:

There is a plan to use increased health and longevity provided by AI healthcare advancements to allow elderly work longer to decrease issues with a shrinking workforce.

Health:

A multitude of beneficial plans for AI in healthcare are discussed. These included the development of customized foods for individualised health benefits, promoting longevity through early discovery of disease and preventive medicine, creation of artificial organs and sensors, and drug discovery.

Environmental:

Japan intends to build environmentally friendly travel with shared systems backed by AI transportation planning.

Personal:

They hope to remove human error from transport to increase physical safety.

Community:

It's planned to create AI-robotic home assistants for things like improved elderly care.

July, 2017 - Draft AI R&D GUIDELINES for International Discussions⁵⁰

These guidelines were released in preparation for a conference on AI networking and identified 5 philosophies and 9 principles for AI networking R&D.

Human Security Relevance:

Personal:

Japan intends to share guidelines and best practices as non-binding law internationally, and keep them in constant review; this will help to promote safe development in a wider context. Two principles were mentioned relevant to Personal Security: controllability – effective oversight and countermeasures, and safety – protection for people and their property from physical harm.

Political:

They intend to keep things human-centred with equitable benefits across society and promotion of human harmony with AI networks. The hope is to find a balance in regulation that is productive but not excessive, so the technology can flourish safely. The relevant principles for Political Security are those of transparency (particularly in systems that significantly affect people and their lives), privacy, and ethics.

4.2.4 Singapore

May, 2017 - AI Singapore

Rather than just releasing a single documented strategy on AI, Singapore developed an entire program concentrating on AI development and investment:

⁵⁰ Japan: Committee on Impact and Risk Assessment, 2017

AI Singapore (AISG). This program was founded by Singapore’s National Research Foundation with the intention to create “social and economic impacts, grow the local talent, build an AI ecosystem, and put Singapore on the world map.”⁵¹ A unique aspect of their take on AI development is the inclusion of AI makerspace resources and talent development for all ages.

Human Security Relevance:

Economic:

AI talent development programs for all ages and interests will aid in some of the development and reskilling needed to promote Economic Security.

Health:

There is an AI in Health Grand Challenge competition being run that supports high impact projects.

Personal:

Singapore wants to develop a system for proper certification/validation of AI systems.

Political:

The intentions to model norms and values in AI systems should help with balancing the ethical challenges of AI that threaten Political Security. There is an emphasis on the development of AI for public good.

⁵¹ AI Singapore, n.d. (a)

4.2.5 China

*July, 2017 - Next Generation Artificial Intelligence Development Plan*⁵²

The Chinese AI strategy is currently one of the most comprehensive. The computer-assisted English translation of the text is almost 12,000 words. The document begins by breaking down the strategic environment surrounding AI development. They follow this up with the statement that China has a solid base with AI development, but also many areas that are in need of growth and improvement. The document then delves into the requirements section, which is comprised of the ideologies, principles, strategic objectives, and deployment intentions driving China's R&D aspirations.

Human Security Relevance:

Economic:

China has a goal to develop an intelligent economy with “data-driven, human-machine synergy.” This is another instance of emphasis on collaborative AI systems rather than human replacement.

Food:

There are plans for AI agriculture and food safety systems – such as AI based sensing and control systems for crops that span the earth and sky, and an early warning system around food hazard classification and assessment. China also plans to use big data to create AI that can analyse agricultural research questions. They intend to integrate these systems in pastures, fish farms, orchards, food processing, and supply chain management.

⁵² China: State Council, 2017

Health:

Many Health Security applications for AI are laid out. These include smart hospitals, wearable sensors with AI inside, human-machine collaboration on intelligent treatment systems, image recognition, identification of pathologies, genomic studies, drug discovery, and study of epidemiological outbreaks.

Environmental:

The Chinese strategy has plans for internationally networked environmental monitoring by building intelligent systems for watching and analysing the atmosphere, water, and soil; emission predictions, prevention, and emergency responses; and AI early warning and comprehensive response platform for natural disasters of all kinds.

Personal:

The strategy hopes to achieve “coordination between incentive development and rational regulation, and maximize risk prevention.” Smart transportation systems and cities are planned with the intention of integrating AI into every aspect of the city – from infrastructure construction and planning to R&D on the operation of complex urban systems.

Community:

There are many applications for AI in Community Security mentioned. These include improving R&D for growth in social security and prosperity, developing AI in an interdisciplinary manner, extensive plans for elderly care – “smart health and retirement,” and extensive education reform.

Political:

The Chinese hope to use AI for smart social governance. AI for government service provision and decisions, applications in complex social research, risk warning, policy evaluations, and attempted prediction of public needs are all mentioned. There is also potential threat to social liberties and democracy though, with an AI platform for integrating trial, personal, dynamic monitoring, judicial disclosure, evidence collection, and legal document analysis.

4.2.6 United Arab Emirates (UAE)

October, 2017 - UAE AI 2031

There is an official website for the AI strategy, but it doesn't seem to have much depth.⁵³ It just displays a flashy YouTube video and a couple of small text blurbs about the initiative on a single home page, both in the English version and the Arabic.⁵⁴ The video highlights the UAE's goal to make their nation run smarter and more efficiently. More details can be found on a separate government website, however it still is not very substantial.⁵⁵

Human Security Relevance:

Health:

The UAE mentions AI's potential to minimise chronic disease and dangerous ailments and AI integration into medical and security services.

⁵³ UAEai, n.d.

⁵⁴ UAEai 2031, 2017

⁵⁵ UAE, 2018

Environmental:

They hope to manage renewable energy facilities, analyse and study water sources, and monitor growth of forests with AI.

Personal:

The UAE has plans to reduce accidents in transport and traffic with AI, and they are concentrating on development of safety laws.

Political:

Improving governmental processes with the use of AI is another goal.

4.2.7 Finland

*December, 2017 - Finland's Age of Artificial Intelligence - Turning Finland into a leading country in the application of artificial intelligence*⁵⁶

This is the first in a series of exploratory reports surrounding AI. The first section of the report analyses AI around the world, the possibilities for Finland in the global market, and Finland's weaknesses and strengths surrounding AI. The second section highlights the revolutionary nature of AI in both the public and private sectors. The third section is the most important strategically – it talks about 8 key recommendations for action that the working group proposed to push Finland forward in its AI development.

Human Security Relevance:

Economic:

The document questions what kinds of Economic Security mitigations would work best, and how to incentivise contribution

⁵⁶ Finland: Ministry of Economic Affairs and Employment, 2017

in work and improve employability with the threat of mass job displacement. It also concentrates on the need to secure the labour market and social security because of the unemployment threat.

Health:

There are plans to use Finland's unique healthcare data to develop global market solutions and AI for wellbeing – especially in areas like social inclusion.

Environmental:

Growth of public-private partnerships in AI for green energy is highlighted.

Personal:

Finland concentrates on using AI for “early detection of threats, preparedness and the guarantee of quick recovery”⁵⁷ in public disasters (such as terrorism). This may help to lessen the impact on people's Personal Security.

Community:

They want to use AI to reform elderly care.

Political:

There is emphasis on the need for growth in the conversation around the ethical development and use of AI systems.

⁵⁷ Ibid., p. 26

4.2.8 Taiwan

*January, 2018 - AI Taiwan*⁵⁸

There are 5 main initiatives to the Taiwanese plan: “cultivating talent, developing Taiwan’s niche AI, incubating local AI start-ups, reconciling laws for AI development and introducing AI technologies to industries.”⁵⁹ There are 5 projects in Taiwan’s AI action plan currently; these are a talent development and acquisition program, an International Innovation HUB, a Pilot Project modelled after the American DARPA and Japanese SIP, open testing “fields” of data and regulation research, and industrial innovation for the 5+2 goals.⁶⁰

Human Security Relevance:

Health:

Taiwan wants to improve biotechnology with AI development.

Environmental:

The document mentions eco-friendly energy development, creating a new business model in agriculture, and formation of a recycling based circular economy with the help of AI.

4.2.9 Italy

*March, 2018 - AI White Paper*⁶¹

One of the more unique facets of this white paper is its heavy concentration on how AI can impact and improve Public Administration. The Italian Government aims to encourage adoption of AI throughout the government with the goal of

⁵⁸ AI Taiwan, 2018

⁵⁹ Ibid.

⁶⁰ AI Taiwan, n.d.

⁶¹ Italy: Agency for Digital Italy, 2018

improving public services; they also hope to boost innovation, the economy, and improvements in the daily lives of their citizens.

Human Security Relevance:

Health:

Some of the AI applications for Health Security mentioned are smart diagnostic tools, analysis of epidemics for public health early warning systems, translation for easier foreign medical visits, and personalised medicine.

Environmental:

The document states that AI can be useful in predictive pollution modelling, developing a network of 5G wireless sensors to feed an AI algorithm that can use the various combined data to identify things like eco-crimes (such as harmful chemical spills), fires, and floods, and simulating environmental disasters before, after, or during occurrence to help with intervention and response.

Personal:

Italy plans to overhaul traffic management with AI data processing. Better management means increased safety.

Community:

There is a wide set of applications in Community Security mentioned, such as personalised citizen relations, education reform, identification of indicators for educational drop-out risk, promotion of equal access to AI benefits, use of AI to prevent and reduce various social inequalities, and support for those with disabilities and health issues.

Political:

Political Security is also highly relevant in Italy's strategy.

Development of AI Ethics through reflection on the use “in decision-making processes related to social, health and judicial issues (such as risk assessment),”⁶² focusing on data quality and neutrality, developing accountability and liability standards, integrating transparency in development, protecting privacy and spreading knowledge and guarantees of individual rights are all applicable ideas.

4.2.10 France

*March, 2018 - For a Meaningful Artificial Intelligence: Towards a French and European Strategy*⁶³

This French AI report, written by French parliament member and mathematician Cédric Villani, fed into their overall current AI strategy. At over 60,000 words it is one of the most comprehensive takes on AI currently developed.

Human Security Relevance:

Economic:

Villani emphasizes that AI should be developed with the idea of complementarity in mind – it should work with humans, not replace them, by taking advantage of the fortes of humanity such as creativity and manual dexterity. He wants to prepare people for the upcoming job transitions pre-emptively, rather than after the AI displacement of jobs happens. It's suggested that certain jobs at higher risk of more immediate automation should be given extra resources and consideration to help keep balance in the short term,

⁶² Ibid., p. 20

⁶³ Villani, 2018, pp. 80-99

and also that the ideas behind professional development need an overhaul to fit the new paradigm. He also thinks testing needs to be done on the concept of a universal basic income.

Food:

Intelligent agriculture, predictive analysis for food assistance with French Red Cross, and harvest prediction models could all benefit Food Security.

Health:

Benefits can be seen for the “4 Ps of healthcare”⁶⁴ – personalization, prevention, prediction, and participation. Earlier detection of cancer and other diseases and the elimination of ‘medical deserts’ where people are at distance from medical care are relevant major applications.

Environmental:

Attention needs to be paid to the growing energy costs of computation and how that negatively impacts our planet. There needs to be concentration on developing AI with energy efficiency in mind. Areas where AI can help with Environmental Security are energy management, environmental preservation, understanding complex ecosystems’ biodiversity, development of a platform for measuring AI’s impact on the environment, weather and climate prediction and risk management, agricultural development with less pollution, and development of emission free transport.

⁶⁴ Ibid., p. 9

Personal:

Issues with intentionally altered environments meant to mess up AI vision need to be considered, especially in the case of driverless vehicles. Lethal AWS are a direct threat and there needs to be adherence to international and humanitarian law. Establishment of an observatory for autonomous weaponry non-proliferation would be useful. There is an emphasis on keeping automation of lethal weaponry limited to having a human involved.

Community:

There is a need to manage the gender balance in AI development in order to make sure that a wider set of perspectives is taken into account when planning the technology. A couple of suggested solutions for this are to positively incentivise female enrolment for tech based courses to boost the outgoing numbers of relevant female engineers and scientists, and to create a national initiative for growth of female inclusion in technology. Other relevant suggestions are a national database on workplace gender equality, promoting transparency in the processes for promotions and hiring, and setting aside funding specifically to address diversity issues.

The creation of an administrative procedure management system run by AI may be beneficial for guaranteeing equal access to public services. Villani shows support for AI based social innovations – such as the provision of AI skills training for social actors. The effects of robots and automation in social support systems needs to be studied. Development of technologies for help in facilitating digital accessibility for disabled persons could also make a strong impact. Predictive policing can have discriminate impact issues, and needs to be managed more carefully. For

management of risk, development of a discrimination impact assessment for AI would be a good idea.

Political:

Villani states that ethics need to be implemented from the start in design, and throughout each stage of development. Training computer scientists in social sciences so they understand the impact not just the technicalities may help. Transparency, auditability, and impartiality are paramount in development. Additionally there should be creation of an AI ethics committee and development of collective data rights to deal with how AI processes mass quantities of data. Predictive policing also affects Political Security, and may increase risk of mass surveillance infringing on rights and democratic processes.

*March, 2018 - Artificial Intelligence: "Making France a leader"*⁶⁵

At the end of the AI for Humanity summit in Paris, President Macron said AI is “not only a technological but also an economic, social, ethical and [therefore] political revolution,” and the mastery of this technology will be necessary to retain sovereignty.⁶⁶

Human Security Relevance:

Economic:

France shows concerns over the eradication of jobs, but hopes that the technology itself will help to ameliorate those issues.

⁶⁵ French Government, 2018

⁶⁶ Ibid.

Health:

The National Institute for Health Data is developing a secure, anonymous health data hub to develop AI from.

Community:

AI can help improve lifelong learning via need anticipation, individualisation, etc.

Political:

The document shows ethical concerns surrounding the neutrality of algorithms, private data use, and “e-inclusion.” France plans to make all State AI algorithms public for transparency.

4.2.11 European Union (EU)

December, 2018 - Coordinated Plan on Artificial Intelligence⁶⁷

This document has an emphasis on human-centric AI development in Europe. It also encourages all EU member states to establish their own National AI strategies and budgets by mid-2019.

Human Security Relevance:

Economic:

Policymakers need to plan for managing rapid job loss and the socio-economic shifts that will come with it.

Health:

Care needs to be taken when processing healthcare data for AI to ensure compliance with GDPR. In 2020, a common database of

⁶⁷ European Union, 2018

voluntarily supplied health images will be created; initially it will be used for diagnosis and treatment of the most common cancers.

Personal:

Basic human control needs to be maintained with AWS, and they need to comply with Human Rights and International Humanitarian Laws.

Community:

There is a need for diversity and gender balance in the European AI market.

Political:

The success of an ethical and legal framework will depend on pan-European coordination. Innovation-friendly ethics guidelines need to be developed with a global outlook.

April, 2019 - Ethics Guidelines for Trustworthy AI⁶⁸

The document starts out with a breakdown of the 3 components that they believe make up Trustworthy AI: lawful, ethical, and robust (both technically and socially). The rest of the document is separated into the Foundations of Trustworthy AI (a philosophical presentation of the foundational ethical principles), Realising Trustworthy AI (a set of more direct requirements for Trustworthy AI based on the prior ethical philosophies), and Assessing Trustworthy AI (practical direction for operationalising the requirements in the prior chapter). The final chapter provides examples for evaluation of AI projects' trustworthiness, both qualitatively and quantitatively.

⁶⁸ European Union: AI High-Level Expert Group, 2019

Human Security Relevance:

Economic:

Risk of job loss/deskilling of the workforce should be assessed and counteracted. AI could help forecast this loss, allowing for preventative reskilling and education.

Health:

Threats to mental autonomy and health should be mitigated. Smarter and more targeted treatment development, more accurate and detailed analysis of complex health data, with personalised preventative measures and earlier disease detection will be possible with AI.

Environmental:

AI should be both sustainable and environmentally friendly in the entire development, deployment, and usage process. Energy consumption should be critically examined and supply chain environmental friendliness is also important. AI may be used for detecting energy needs and efficient energy management. Optimisation of energy efficient engines through AI may help with decarbonisation.

Personal:

AI needs to be designed with harm prevention in mind and human oversight must be maintained for safe development. This means “human-in-the-loop (HITL), human-on-the-loop (HOTL), or human-in-command (HIC)”⁶⁹ setups. Essentially no completely autonomous and unsupervised systems. Resilience to attack

⁶⁹ Ibid., p. 16

(physical and cyber) must be built to ensure no malicious harm is enacted. Fallback plans must be developed as part of appropriately vigorous security plans dependent on the level of risk inherent in the AI systems. Lethal AWS raise many concerns in relation to ethicality, legality, and safety. An arms race could get out of control with lack of proper safety verification and potentially near complete removal of human oversight. AI may also help reduce fatalities in car crashes.

Community:

Fairness must be ingrained in AI. Non-discrimination, equality, and solidarity must be maintained – especially with regards to vulnerable populations (women, disabled, minorities, children, etc.). Accessibility needs to be a concentration with universal design in mind. The social impact of AI should be properly monitored. AI & robotics will be useful for elderly care, too. An increase in average learning speed and education quality through personalized, adaptable education is possible.

Political:

Fundamental rights are the core of Trustworthy AI: respect for human dignity, individual freedom, law, justice, and democracy. Trustworthy AI will be human-centric, lawful, ethical, robust (both technically and socially) to prevent accidental harm, transparent, accurate, reliable/reproducible, and explicable. AI should not negatively affect citizens' rights. Privacy/data protection should be maintained throughout the entire AI lifecycle. AI makes identification and tracking of individuals much easier; this may have unexpected psychological and sociological effects. AI's impact on democracy and society at large should be considered. Covert AI systems should be avoided – any AI systems should be

explicitly identified as so. Use of AI in “citizen scoring” systems should be avoided due to potential discrimination and loss of autonomy. In narrow cases like e-learning or licensing it should be transparent and ideally have an opt-out option.

4.2.12 United Kingdom (UK)

*April, 2018 - Industrial Strategy: Artificial Intelligence Sector Deal*⁷⁰

This document is primarily about the economic side of AI, with a large concentration on building human resources, R&D investment, and partnerships between government, business, and academia. The five core pillars of the Strategy are ideas, people, infrastructure, the business environment, and places.

Human Security Relevance:

Food:

AI can be used to reduce crop disease.

Health:

The Industrial Strategy Challenge Fund is making a £210m investment in early disease diagnostics and precision medicine. Development of devices for home medical advice will also be a boon.

Personal:

The Industrial Strategy Challenge Fund is also using £93m for AI robotics in dangerous/extreme environments to support safer work for people.

⁷⁰ United Kingdom, 2018

Community:

The UK is committed to building diverse research and work teams in AI.

Political:

Establishment of a Centre for Data Ethics and Innovation, with AI as one of the advisory topics, should help to limit risk of AI adversely affecting Political Security.

4.2.13 South Korea

May, 2018 - Investment of KRW 2.2 trillion to secure world-class artificial intelligence technology⁷¹

This strategy concentrates mostly on technical and human resource investments. One of the stand out aspects of this plan is the scheduled investment of at least 1 trillion South Korean won into the development of an AI semiconductor by 2029. A major focus of this document is on forest disaster response.

Human Security Relevance:

Health:

AI will benefit accelerated health evaluation. South Korea intends to establish a learning centre for medical equipment to support tech development such as AI integration in health.

Environmental:

Pre-emptive natural disaster response planning can be done with AI. This may include development of an intelligent forest fire disaster response strategy, forecasted landslide evacuation

⁷¹ South Korea, 2018

notification system, and a forest pest occurrence, diffusion prediction, and response system.

Political:

Ethical AI needs to be built from the design phase forward.

4.2.14 Sweden

*May, 2018 - Artificial Intelligence in Swedish Business and Society*⁷²

This is a summary of a report commissioned by the Swedish government to analyse the potential for AI in Sweden.

Human Security Relevance:

Economic:

AI creates a threat of instability for employment and job loss.

Health:

AI image analysis (such as breast cancer screening) and wearable health trackers are goal projects.

Personal:

There is potential for AI in mining to replace dangerous tasks.

Political:

We need to balance privacy, ethics, and protection of the community with the need for data to fuel AI.

⁷² Vinnova, 2018

*May, 2018 - National approach to artificial intelligence*⁷³

The Swedish view maintains a strong dependence on working with other countries, as much of the ground-breaking AI work is being done outside of Sweden. Fittingly, also in May, the Swedish Government led a meeting of Ministers from the Nordic-Baltic Region that ended in a declaration of cooperation, statement of shared goals, and the intent of using the Nordic Council of Ministers to facilitate partnerships throughout the region.⁷⁴ A little relevance to Human Security can be found in the document, but it primarily concentrates on the conditions needed to build AI.

Human Security Relevance:

Political:

Sustainable AI development will be safe, secure, reliable, and transparent. Sweden will work with the GDPR to ensure privacy in AI applications.

4.2.15 India

*June, 2018 - National Strategy for Artificial Intelligence: #AIforAll*⁷⁵

While this strategy has the typical aspect of increased investment, it also has some unique features as well. Namely, the Indian strategy is extremely people-centred; this is hinted at by the hashtag in the Strategy title: #AIforAll. Rather than concentrating mainly on what AI can do for the economy and scientific growth, it focuses on using AI for positive social impact.

⁷³ Sweden: Ministry of Enterprise and Innovation, 2018

⁷⁴ Nordic Council of Ministers, 2018

⁷⁵ India: NITI Aayog, 2018

Human Security Relevance:

Economic:

India plans to use AI to improve agriculture and help bridge the gap in farmer to non-farmer income. Skilling and reskilling of the workforce is a priority due to AI disruption. They plan to create financial incentives for companies to reskill employees and open platforms for learning. Again the augmentation of the workforce with AI rather than displacement is emphasized. India plans to create a standing committee/taskforce to follow AI related changes in the employment landscape.

Food:

AI has the potential to drive a food revolution and address challenges in demand prediction, unsustainable resource usage, irrigation, fertilizer/pesticide use, detection of pests, soil and crop health monitoring, and making farming machines more efficient.

Health:

AI can make care more affordable and accessible, especially in rural areas. The combination of AI and the Internet of Medical Things will be beneficial. AI will also aid in moving from a mostly reactive approach to healthcare to an automated, proactive model via use of AI based “diagnostics, personalised treatment, early identification of potential pandemics, and imaging diagnostics.”⁷⁶ Overall it will allow for easier universal healthcare.

⁷⁶ Ibid., p. 20

Environmental:

Smart transportation needs to be developed with sustainability in mind to overcome the challenges of congestion and lack of public transport. AI may also be used to track climate change and increase efficiency of utility distribution.

Personal:

Safer transportation from AI may help alleviate road accidents with high levels traffic deaths. Smart cities can help decrease violent crime and accidents, and smart parks can improve lighting at night and monitor facility conditions. AI may also be used for help in solving human trafficking issues. India plans to develop a framework for analysing damages caused by AI.

Community:

Smart cities can help with management of urban emergency risk with things such as AI for crowd management in major events and disasters. AI based education resources are particularly helpful in a linguistically diverse population. Also, AI may help with development of tools for predictive management of students at risk for drop-out and adaptive learning and tutoring systems for individualized need.

Political:

Ethics need to be taken into account with a concentration on fairness, accountability, and transparency. There are challenges around data usage consent, identification of individuals, and bias – the goal is to actively identify biases, assess impact, and reduce them. Privacy concerns may be alleviated by establishing data protection and regulatory frameworks, spreading awareness of privacy rights, keeping up with progress compared to international

data protections, and working on mathematical ways of preserving privacy.

4.2.16 Mexico

*June, 2018 - Towards an AI Strategy In Mexico: Harnessing the AI Revolution*⁷⁷

This report was supported by the British Government's Prosperity Fund. It surveys the state of AI around the world and how Mexico fares in comparison using an Oxford Insights AI readiness index to show Mexico's rank compared to 34 other countries – its overall ranking was 22/35. Mexico was last in innovation and tech skills, and close to last in government effectiveness and digitalisation. The only areas where Mexico was in the top 10 were data capability and availability.

Human Security Relevance:

Economic:

Around 1/5 of Mexican jobs are in danger of negative impact and/or loss due to AI growth, many from business that is currently exported to Mexico being cheaper to do at home with AI. Mexico needs to develop the tools for a more resilient workforce.

Food:

AI will help with improved agricultural efficiency via things like crop, weather, and soil monitoring, and smart, weather-sensitive irrigation.

⁷⁷ Oxford Insights and C Minds, 2018

Health:

AI may improve access to healthcare with things like automated medical chat-bots.

Community:

More than 2x as many jobs at risk are occupied by men, due to a disparity in the proportion of men and women across sectors in Mexico. There is chance for social disruption without proper mitigations in reskilling. There is also a great potential for AI applications to advance social equity in areas like health and education. To help equality grow, Mexico should return productivity increases from AI into wage boosts and innovation in neglected areas and reskill when jobs are threatened/lost. They intend to promote gender equality in digital skills.

Political:

The protection of personal privacy via the Mexican Institute of Access to Information and Personal Data Protection frameworks is necessary. AI may be helpful in countering false information, especially during disasters. Mexico is recommended to create an AI Ethics Council, to set standards, and award official recognition of those that abide by them.

4.2.17 G7

*June, 2018 - Charlevoix Common Vision for the Future of Artificial Intelligence*⁷⁸

The G7 vision statement comments that “AI that fosters economic growth, societal trust, gender equality and inclusion depends on a predictable and stable

⁷⁸ G7, 2018

policy environment that promotes innovation.”⁷⁹ It states that policy that promotes innovative and human-centric AI should be supported.

Human Security Relevance:

Community:

The G7 emphasizes lifelong training and education of all, but especially minorities (such as women and at-risk populations). AI accessibility for all is also a concern, with particular interest in smaller businesses and minorities. They see the possibility of using AI for women’s economic empowerment, but need to address issues related to biases in algorithms. Some issues may be solved by involving marginalized and underrepresented populations in all stages of AI development.

Political:

Human-centric AI is a priority. The G7 will safeguard privacy through proper legal regimes, spreading information about privacy concerns and international law, and using other transformative research to protect privacy and promote transparency. They have a goal for free flow of data within the limits of proper privacy concerns. They call for human oversight in AI decisions.

4.2.18 Germany

*November, 2018 - Artificial Intelligence Strategy*⁸⁰

Germany has the intention of becoming a competitive AI hub with responsible use and development of AI technologies while keeping culture, ethics, and legalities

⁷⁹ Ibid.

⁸⁰ Germany, 2018

in mind. They see the GDPR as a solid step in the right direction for Europe in the data regulation and protection concerns of AI.

Human Security Relevance:

Economic:

Development of a broad set of workforce skill growth tools with AI is a goal for Germany. They intend to proactively design future work as changes from AI innovation are captured.

Food:

AI has applications in animal welfare, food security and safety, and sustainability.

Health:

Personalized health care risk analysis and preventative action are a couple of the benefits catalysing targeted funding support for AI in healthcare.

Environmental:

AI can be developed for climate change monitoring and mitigation, eco-friendly transport, and investigation of agricultural deforestation. There is earmarked financing for AI tech for protection of natural resources and the environment.

Personal:

Safer transport systems and emergency responses are two applications where AI may affect Personal Security in Germany.

Community:

AI robotics for elderly care assistance is a goal. There is a concentration on making sure data rights are properly addressed

for disadvantaged groups to avoid discrimination. Germany wants to account for gender and diversity issues when assessing and taking action in the AI sector, and around reskilling.

Political:

There is a need for human-centric development of AI for the common good, based on democratic principles with attention to data and privacy concerns. Ethical development should be “ethics by, in and for design.”⁸¹ People should maintain cultural and media freedoms when developing AI usage in media with awareness to potential threats to democracy. The government needs to develop proper, ethical data protection guidelines in relation to AI, and there should be transparency of decisioning processes and in the use of people’s data.

4.3 Governments Developing Still

Many other governments are in the throes of developing their AI strategies still today. Every country within the EU, for example, has been called to put together a strategy by the EU Commission’s own strategy. Surprisingly, Israel, a major AI player already, does not have a strategy put together yet, despite the Israel Innovation Authority’s warning that the country could fall behind if it doesn’t act.⁸² In a similar vein, Russia is falling behind in the creation of an AI strategy; it has more done than Israel, though. There was a conference in 2018 that led to 10 recommendations on a Russian strategy, but it has not been made official yet.⁸³ A call for submissions on possible approaches to the strategy was

⁸¹ Ibid., p. 16

⁸² Solomon, 2019

⁸³ Russia: Ministry of Defense, 2018

put out by President Putin with a deadline of the end of February, so there is at least a strategy in the works.⁸⁴

4.3 Tables and Graphs of Relevancies

4.3.1 Main Data

In the following table, each document has been listed vertically, with the 7 dimensions of Human Security horizontally. They have been assigned a green square for documents that have any relevant data for each dimension. The sums to the right simply add up the number of relevant dimensions per document, while the sums at the bottom add up the number of government documents with relevancy to each dimension. Both sums are colour coded in various intensities of red, orange, yellow, and green as their numbers increase (indicating wider coverage of the dimensions). The first line, 4.1.1.1, for example shows that this document has relevancy in Economic, Personal, and Political Security, leaving the sum at an orange three – indicating not the widest of coverage, but still better than red.

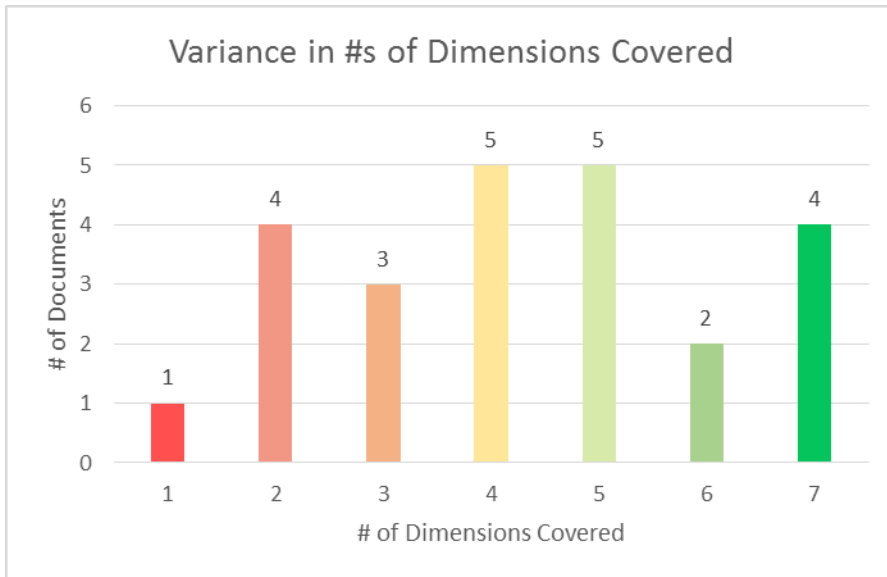
Government Entity	Reference	Relevant Dimensions of Human Security							Sum
		Economic	Food	Health	Enviornental	Personal	Community	Political	
United States	4.1.1.1	Green				Green		Green	3
US: California	4.1.1.2					Green			3
US: DOD	4.1.1.3				Green				4
Canada	4.1.2.1						Green	Green	2
Japan	4.1.3.1	Green			Green				5
	4.1.3.2					Green		Green	2
Singapore	4.1.4.1	Green							4
China	4.1.5.1	Green	Green		Green			Green	7
United Arab Emirates	4.1.6.1							Green	4
Finland	4.1.7.1	Green						Green	6
Taiwan	4.1.8.1								2
Italy	4.1.9.1							Green	5
France	4.1.10.1	Green						Green	7
	4.1.10.2	Green							4
European Union	4.1.11.1	Green							5
	4.1.11.2	Green							6
United Kingdom	4.1.12.1							Green	5
South Korea	4.1.13.1				Green				3
Sweden	4.1.14.1	Green				Green			4
	4.1.14.2								1
India	4.1.15.1	Green				Green		Green	7
Mexico	4.1.16.1	Green						Green	5
G7	4.1.17.1								2
Germany	4.1.18.1	Green				Green		Green	7
Sum	24	13	6	17	12	17	16	22	

⁸⁴ Russia, 2019

4.3.2 Table and Graph of Variance

The following table and graph represent the variance in the numbers of dimensions covered in each document. It's simply a count of how many documents represented each possible number of dimensions, from only a single dimension, to all 7.

Variance in #s of Dimensions Covered							
# of Dimensions	1	2	3	4	5	6	7
Count of Documents	1	4	3	5	5	2	4



5. Analysis of AI Impacts on Human Security

The world of AI is replete with risks to Human Security. These can be broken down into a few various categorizations: they can be short- or long-term, systemic or acute, ethical/moral, economic, or even existential. Benefits can also be seen in almost every sector of our world from healthcare, business, and the environment to transportation and war. The positives and negatives of these variables can be examined in the light of the 1994 UN HDR Human Security categories: Economic, Food, Health, Environmental, Personal, Community, and Political Security.

Although the previous chapter was handled mostly chronologically with a sub-breakdown into the 7 dimensions of Human Security, this section will be separated into the 7 dimensions directly. After this, a discussion and summary of the findings from the 24 AI strategy documents that were examined will ensue. These two parts will essentially answer the main two points of the research question that was proposed in section 1.3 by highlighting the set of risks and benefits emerging from AI that can affect Human Security and then delving into the comparisons surrounding how are they being acknowledged and addressed within the government strategies presented in chapter 4.

5.1 Human Security Impacts

5.1.1 Economic Security

One of the most acute and pressing risks that Artificial Intelligence poses to Human Security is job loss and its resulting side effects, both macroeconomically and in individuals' lives. Job loss due to automation is already affecting all economic sectors in both the goods and services markets.⁸⁵ AI is threatening an ever-growing number of jobs, even many (such as white collar and creative positions) that seemed secure from the impact of machine

⁸⁵ Rojanasakul et al., 2017

automation. This job loss catalyses even more risks, such as a rise in inequality, macroeconomic instability, and secular stagnation.⁸⁶

In the United States, productivity has been rising faster than wages, while the economic share of the elite .01% has grown. This is partially due to the fact that the technological growth in areas such as AI is enabling companies to swap out labour costs for hardware and software costs – leaving wages to stagnate while productivity rises.⁸⁷ This can affect Human Security in the Economic realm directly – the loss of income creating insecurity for those it affects through the potential for poverty to emerge if another job isn't found or proper social mitigations aren't in place. Job loss from increased automation then also impacts Food, Health, and Political Security indirectly.

If unemployment were to arise en masse, social unrest and drastic revenue drops from declined tax funding may also come to pass unless government policies properly mitigate the risks.⁸⁸ A sudden decline in tax revenue would create instability in government and might threaten Political Security. The lack of revenue would limit funds for everyday tasks and public works/assistance. As unemployment rises, the need for governmental assistance would as well; when paired with the lack of available assistance this could leave many people out in the cold both literally and figuratively.

One study estimated that global GDP is likely to rise by about \$15.7 trillion by 2030 due to AI related market changes; they also highlight that this boost will not be balanced across markets, with China likely to account for a little under half of that income and North America for about a quarter of it, leaving the other quarter to be spread around through the various world markets.⁸⁹ Markets that lack AI resources (such as in developing countries) will also be affected by an AI boom. Their populations will continue to grow, while jobs will still be lost to

⁸⁶ Frey and Osborne, 2015; Beard et al., 2017

⁸⁷ Frey and Osborne, 2015

⁸⁸ United States Department of Homeland Security, 2017

⁸⁹ PricewaterhouseCoopers, 2017

AI. A lack of proper precautions coupled with no AI income boost (to help balance out the displacement of labour) could end in global market restructuring and economic dependence in those affected areas without AI revenue.⁹⁰

Some countries with heavy manufacturing economies where others outsource work to, such as Mexico, will also have issues with losing jobs due to the outsourcers being able to use AI systems to manufacture cheaper at home.⁹¹ The income from this revolution is useful for the economic security of those entrepreneurs who are taking the initiative for its business applications. However, unless managed properly, the Economic Security of the few may become the Economic Insecurity of the many – with potential trickledown effects on other facets of Human Security.

One of the major areas for ensuring Economic Security for populations is access to education for skilling and reskilling of the population. There are many examples of this in the Data. India is planning to use financial incentives to encourage businesses' reskilling efforts.⁹² Singapore's main AI initiative has an entire section devoted to AI learning, from children to industry level.⁹³ The UK is taking a similar view on the needs of skill development and are developing a National Retraining Scheme to try to balance the needs of the changing economy.⁹⁴ To compound the reskilling initiatives, a handful of governments (such as the US, China, France, and India) also emphasized keeping the creation of AI limited to human-AI collaboration systems instead of replacing workers with the technology.⁹⁵

There is some hope that AI will help to solve the problems it is creating. For one, the EU is planning on trying to use AI to forecast job losses.⁹⁶ Similarly,

⁹⁰ Lee, 2017

⁹¹ Data: 4.2.16

⁹² Data: 4.2.15

⁹³ AI Singapore, n.d. (a)

⁹⁴ United Kingdom, 2018

⁹⁵ Data: 4.2.1, 4.2.5, 4.2.10, 4.2.15

⁹⁶ Data: 4.2.11

Germany hopes to plan for future work by predicting needs as AI develops.⁹⁷ Japan, uniquely, is considering using the health and longevity benefits of AI to enable their older citizens to work for longer if they'd like, helping to decrease the gaps in the workforce developing from an aging workforce.⁹⁸ Another potential tool for balancing income insecurity from job losses is the roll out of a universal basic income, but only France is explicitly considering testing it.⁹⁹

5.1.2 Food Security

There are a few current applications for Food Security and AI, but they're mostly limited to crop monitoring and harvesting, with few deviations. For example, AI has been widely investigated for use in agriculture in Japan. A labour force shortage has been happening there, and they see AI and robotics as a possible compliment to this shortage, as well as a possible replacement for humans in dangerous agricultural jobs.¹⁰⁰ Autonomous vehicles for harvesting already exist but aren't widely implemented due to issues with liability not being sorted out yet. However, drones have been in use for pesticides in Japan since the 80's, and with AI image recognition of plant health (growth, pest, and disease detection) from aerial photographs it is becoming more streamlined and precise. There has even been a wearable sensor for cattle created that uses AI to analyse their behaviour and diagnose potential disease.

China also has a set of plans that would impact Food Security; some of these are similar to the crop management systems that Japan is planning and running, but it is also more comprehensive – including agricultural research, food processing, and supply chain management. They also have a unique plan for using AI for food hazard classification and assessment.¹⁰¹ India has similarly

⁹⁷ Data:4.2.18

⁹⁸ Data: 4.2.3

⁹⁹ Data: 4.2.10

¹⁰⁰ Acceptable Intelligence with Responsibility, 2018, pp. 80-84

¹⁰¹ Data: 4.2.5

comprehensive plans with the intention to use AI to improve all aspects of agriculture – from irrigation to efficiency of farming machines.¹⁰² France has some unique plans as well. They have the typical plans for intelligent agricultural systems, however, they also want to use AI for predictive analysis to aid in the French Red Cross’s food assistance planning.¹⁰³ Finally, Germany has some general plans for AI and food security, but their standout plan is to help in Animal Welfare.¹⁰⁴

5.1.3 Health Security

Healthcare is an area prime to receive benefits from AI’s impact, with the potential to increase global Health Security. “The impact of AI can be considered in two main themes: first, extracting meaning from ‘Big Data’ in the research domain; and second, aiding clinicians in delivering care to patients.”¹⁰⁵ Some of the ways that healthcare can be impacted by AI include: smart scheduling, supporting diagnoses, pandemic identification and tracking, and imaging diagnostics,¹⁰⁶ medical emergency response,¹⁰⁷ robotic surgery, and modelling of assorted bodily processes.¹⁰⁸

Canada has a strong emphasis on projects for AI to improve public health. The Canadian Institutes of Health Research (CIHR) have a set of government funded projects that are being worked on currently, from a grant competition in 2017 – one of the initiatives is Artificial Intelligence for Population Health.¹⁰⁹ In addition to this, one of the main “strategic research priorities” of CIHR is Equitable (and ethical) AI.¹¹⁰ Canada is concerned that if not proactively managed

¹⁰² Data: 4.2.15

¹⁰³ Data: 4.2.10

¹⁰⁴ Data: 4.2.18

¹⁰⁵ Jones et al., 2018, p. 224

¹⁰⁶ PricewaterhouseCoopers, 2017

¹⁰⁷ Desouza et al., 2017

¹⁰⁸ Jones et al., 2018

¹⁰⁹ Canadian Institutes of Health Research, 2017

¹¹⁰ Canadian Institutes of Health Research, 2019

AI will not have equitable effects on their citizens' health. In a workshop concentrated on the application of AI to public health, a set of opportunities and challenges were identified.¹¹¹ Some of the opportunities that came to light included AI tools for supporting health interventions and healthy behaviours, using AI to understand complex societal health patterns, and outbreak management. Two of the challenges identified include the issue of data sharing with the privacy of health information and the potential endangering of certain groups of at-risk individuals if regulation and information access aren't handled correctly.

Japan also has many ideas and current cases surrounding the use of AI in healthcare. They are already using AI for assistance in diagnostic imaging support, genomic medicine, diagnosis and treatment support, and surgery support.¹¹² China's plans are quite similar to Japan's. There has been a lot of worry in the past about AI taking over Japanese doctor's jobs, so most of the cases emphasize the use of AI as a tool to help doctors rather than to replace them; although, there has been talk around the possibilities of combined AI-robotics to replace doctors as well.¹¹³ One exceptional area that Japan is considering implementing AI is with the development of customized foods to promote individualised health benefits.¹¹⁴

Singapore put out a call for an "AI in Health Grand Challenge" meant to spur innovation in the AI & Health Security realm.¹¹⁵ The main challenge question was "How can Artificial Intelligence (AI) help primary care teams stop or slow disease progression and complication development in 3H – Hyperglycemia (diabetes), Hypertension (high blood pressure) and Hyperlipidemia (high cholesterol) patients by 20% in 5 years?" The project submissions have closed and the experiments are currently underway.

¹¹¹ Canadian Institutes of Health Research, 2018

¹¹² Acceptable Intelligence with Responsibility, 2018, pp. 54-58

¹¹³ Data: 4.2.5

¹¹⁴ Data: 4.2.3

¹¹⁵ Data: 4.2.4

The EU's Horizon 2020 project has plans for the development of a common database of health images to build up AI for diagnostic imaging purposes, among many other non-AI related initiatives.¹¹⁶ The intention is to start with common forms of cancer and to see how that works out before expanding to other diseases. They deal with the issues around data and privacy by making the database anonymized and strictly voluntary. Other Horizon 2020 projects that have Health Security implications include a cancer mutation prediction AI called REVOLVER (Repeated EVOLution in cancER)¹¹⁷ and another called EMPATTICS (EMpowering PATients for a BeTTer Information and improvement of the Communication Systems) that intends to use AI to help guide and monitor patients with chronic illness's physical and mental states.¹¹⁸

There are a few outliers in the strategies' plans for AI & health. Finland stands out in their take of health; they emphasized the need to use AI for more than physical health, rather, for holistic wellbeing with things like social inclusion.¹¹⁹ Similarly, one of the EU's strategic documents stresses the need to mitigate threats to mental autonomy and health.¹²⁰ Italy's distinctive idea is to use AI for automated translation to allow for easier medical visits for foreigners who don't speak Italian.¹²¹ Using AI for basic diagnoses with hopes to help eliminate 'medical deserts' is where France is special.¹²²

5.1.4 Environmental Security

Transportation is already beginning to see the impact of AI with such technologies as driverless cars. If driverless ridesharing systems¹²³ became more widespread, there could be a lesser need for as many cars on the roads, and thus

¹¹⁶ Data: 4.2.11

¹¹⁷ Caravagna et al., 2018

¹¹⁸ EMPATTICS, n.d.

¹¹⁹ Data: 4.2.7

¹²⁰ Data: 4.2.11

¹²¹ Data: 4.2.9

¹²² Data: 4.2.10

¹²³ PricewaterhouseCoopers, 2017

decreased carbonisation. Japan is intending to build systems such as this.¹²⁴ Pairing this with the potential for AI mechanical diagnostics, the environmental impact of automobiles could be monitored more closely with notifications of when a car's environmental friendliness is lessening due to wear and tear on the parts. India is particularly interested in the use of AI to help with transportation due to the current issues they have with traffic congestion and a lack of public transport.¹²⁵

The energy sector may also see benefits from AI management algorithms. The efficiency of power grids and storage may be improved with AI analysis and the creation of smart electrical meters.¹²⁶ Additionally, AI can help tap into existing traditional and modern renewable energy resources at lower costs and higher efficiencies.¹²⁷ These improvements may save on both the costs to businesses and individuals as well as lower how much Environmental Security is being impacted by the waste of energy.

Finland wants to grow public-private partnerships surrounding AI and green energy.¹²⁸ Finnish state-owned energy company Fortum is making steps to integrate AI into the energy ecosystem; they say that “Smart and sustainable district heating holds great potential to reduce emissions.”¹²⁹ They believe that AI run district heating will play a part in helping tackle the climate challenge. Fortum has developed an AI system that attempts to predict consumer's heat demands and guides the control room's asset utilization. The company plans for a future where electrical systems become more distributed and complex, and believe that integrating AI now also helps prepare for that future. Fortum is also using AI to track where their systems need maintenance, and how it needs it, keeping the systems up and running at best efficiency to avoid wasting power as much as

¹²⁴ Data: 4.2.3

¹²⁵ Data: 4.2.15

¹²⁶ PricewaterhouseCoopers, 2017

¹²⁷ Victor, 2019

¹²⁸ Data: 4.2.7

¹²⁹ Kaasinen, 2019

possible.¹³⁰ Similar projects can be found in the BRIDGE Horizon 2020 working group concentrating on smart grids and energy storage.¹³¹

Similarly, AI has the potential to help assess and reduce the impact of climate change. As weather and climate become more unstable, these technologies may be able to help assess the option values of different adaptation strategies for various regions; they also may help in predicting where will be hit the hardest by these changes.¹³² The Chinese AI strategy lays out extensive intentions for building AI systems for monitoring and analysing the environment – particularly in relation to emissions and natural disaster responses.¹³³ The insights that AI brings to these situations may help to minimize uncertainty and the resulting ability for private interests to take advantage of others for private gains, and to help poorer areas that may not have the knowledge or means to help themselves.¹³⁴

Another EU Horizon 2020 project – enCOMPASS – is using AI to help the environment in a different manner.¹³⁵ It intends to alter people’s behaviours sociologically. They want to combine smart sensors in people’s homes with easier to understand energy usage data to help develop intelligent control systems for home energy use. In addition to this they want to gamify the process to have positive feedback encouraging people to keep up their constructive impacts.

Finally, a couple of governments had ideas that affect Environmental Security that stick out from the crowd. Taiwan is one of these; they think AI can help to smooth the process of creating a circular economy with a greater emphasis on recycling.¹³⁶ Italy is interesting because they concentrate on the development

¹³⁰ Lundmark, 2019

¹³¹ BRIDGE Horizon 2020, n.d.

¹³² Victor, 2019

¹³³ Data: 4.2.5

¹³⁴ Victor, 2019

¹³⁵ enCOMPASS, n.d.

¹³⁶ Data: 4.2.8

of an environmental sensor network for identifying eco-crimes.¹³⁷ France's pre-strategy document, though, has a particular concern that it raises: the potentially negative impact on the environment of the growing energy costs of computation (which AI keeps demanding greater amounts of).¹³⁸ The EU also emphasized that the energy needs of AI should be examined, while adding that AI supply chain environmental friendliness is also important.¹³⁹ South Korea, on the other hand, stood out because of its plans for an intelligent forest protection system with pest and fire aspects; apparently they have been having issues with deforestation and see potential for AI to aid in rectifying this.¹⁴⁰ Germany also mentions deforestation investigation, but doesn't talk about protection per se.¹⁴¹

5.1.5 Personal Security

Most of AI's threats to Personal Security exist within the power that it is given that has the potential for hurting people. Thus the primary concerns that are raised are those of safety in systems that have control over some sort of robotics, such as autonomous vehicles or AWS. AWS are the most direct threat to Personal Security within the AI ecosphere, with their ability to use force and potentiality for taking human life intentionally. Some even believe that it is morally abhorrent to give that power to a machine in the first place, calling it "a fundamental affront to human dignity, and something that can never be morally justified."¹⁴² This is a direct contradiction to the freedom from fear and indignity aspects of Human Security. A person is responsible for their actions on the battlefield, but how can a machine be held accountable for what it does? Additionally, as the state of the art in machine learning currently stands, no person can reliably foresee the choices

¹³⁷ Data: 4.2.9

¹³⁸ Data: 4.2.10

¹³⁹ Data: 4.2.11

¹⁴⁰ Data: 4.2.13

¹⁴¹ Data: 4.2.18

¹⁴² Hart, 2017

that these deadly autonomous systems will make, which may create issues with battlefield attribution of responsibility.¹⁴³

Automation bias – the issue that is created when people begin to rely too heavily on automated judgements – is also a risk for Personal Security. Automation bias can end up in false positive or false negative situations, also called errors of commission and omission respectively.¹⁴⁴ As AI complexity grows and human oversight and control lessens, the risks of these errors occurring rises. This is especially worrisome in the cases where machines may be allowed to make life or death decisions, such as autonomous weaponry.

Just as AI can cause issues with human decision making, human biases can seep into AI systems' decisioning processes. Algorithmic bias may also emerge in AWS, potentially creating a machine that will use force under misleading pretences. Because it is still essentially a computer at its core, it will do what it is programmed to do, whether or not the outcome is fair or actually what is wanted. The semi-autonomous defence systems that are currently in existence have had issues with things like friendly fire already, even with a human still in the loop and able to veto the AI's decisions.¹⁴⁵ This is called the AI goal misalignment problem when this is an accident.¹⁴⁶ The EU emphasizes the need for AWS to comply with Humanitarian and Human Rights Laws and to keep the systems from ever being fully autonomous and unsupervised.¹⁴⁷

Goal misalignment becomes ever more risky the more power that AI gets. There is a field of thought worried about the catastrophic existential risk that a superintelligent AI will end up with a goal that ends up having deadly side effects for humanity.¹⁴⁸ Though not a direct risk now at the current level of intelligence that AI has achieved, this is an emerging risk that we as a society need to prepare

¹⁴³ Winter, 2017

¹⁴⁴ Cummings, n.d.

¹⁴⁵ Cummings, n.d.; Kania, 2017

¹⁴⁶ Russell, 2015; Osoba and Welser IV, 2017

¹⁴⁷ Data: 4.2.11

¹⁴⁸ Global Challenges Foundation, 2017, pp. 60-65

for as we push forward in development. A catastrophic existential threat is the most widespread possible impact on Human Security, potentially destroying us all.

However, accidental goal misalignment is less threatening currently than the more concrete risk that AWS may actually be programmed to intentionally hunt down a specific ethnic group. If the technology can be developed well enough to identify potential adversaries accurately, it could also be reprogrammed to identify harmless groups as well. If Hitler had this kind of technology at hand, the Holocaust might have been even deadlier than it already was. If we end up in a lethal AI arms race, AWS will likely eventually end up in the hands of someone who would abuse it, potentially for genocide; this fact has caused many robotics and AI experts to call for an international ban on AWS development.¹⁴⁹ A ban is unlikely to be effective, however. France has a more realistic outlook with their desire for the establishment of a non-proliferation observatory.¹⁵⁰ Autonomous weaponry has many potential positives, nonetheless. If managed correctly it may have a net positive impact on Personal Security. Some even argue that it's a moral imperative to remove soldiers from the battlefield and save people's lives.¹⁵¹ Additionally, if AWS are developed anywhere in the world, governments are going to have to develop them just to keep up with the speed of warfare.

AI is also being developed as wartime command and control assistance, further speeding up the battlefield.¹⁵² There is a workshop planned by CIFAR that applies to this area: "Regulation of defense and security AI technologies: Options beyond traditional arms control."¹⁵³ This workshop set aims to debate past the traditional complete ban that has been proposed for Autonomous Weapon Systems. This is really the only major initiative that the researcher could find that's being governmentally sponsored currently.

¹⁴⁹ Future of Life Institute, 2015

¹⁵⁰ Data: 4.2.10

¹⁵¹ Etzioni and Etzioni, 2017

¹⁵² West and Allen, 2018

¹⁵³ CIFAR, n.d. (b)

There are some unique trends around personal security concerns. The United States document from the State of California, for one, addresses the need to avoid a lethal AWS arms race, to set minimum standards for safety, and to develop failure transparency in the event of a mishap.¹⁵⁴ In the same vein, Singapore desires the development of a validation and certification system for AI systems to help with these issues.¹⁵⁵ Also in the US, the DOD is uniquely considering how AI can help with predictive maintenance of their machines to increase equipment usage safety and the potential for using AI to lower collateral damage risk in operations.¹⁵⁶ Finland, in comparison, wants to use AI for management of public disasters like terrorism attacks.¹⁵⁷ India thinks that AI also may help to solve human trafficking issues.¹⁵⁸ The UK's unique contribution lies with plans for AI robotics for work in dangerous and extreme environments,¹⁵⁹ and Sweden addresses something similar, but limited to simply mining jobs.¹⁶⁰

Japan briefly mentions the hope that AI based transport networks will help to remove the risk of human error from the transport process.¹⁶¹ India also is concentrating on integrating AI with transportation, with the hopes that it can help to decrease their currently high levels of traffic deaths.¹⁶² France is the only one to acknowledge one of the risks of AI in transport; they bring up the threat of someone intentionally altering an environment to confuse vehicles with AI vision.¹⁶³ This is particularly worrisome, as AI vision is already hard to manage and if someone goes out of their way to create a confusing environment for an autonomous vehicle a crash is likely.

¹⁵⁴ Data: 4.2.1

¹⁵⁵ Data: 4.2.4

¹⁵⁶ Data: 4.2.1

¹⁵⁷ Data: 4.2.7

¹⁵⁸ Data: 4.2.15

¹⁵⁹ Data: 4.2.12

¹⁶⁰ Data: 4.2.14

¹⁶¹ Data: 4.2.3

¹⁶² Data: 4.2.15

¹⁶³ Data: 4.2.10

China talks about the desire to integrate AI into every aspect of the city, from infrastructure construction and planning to R&D of the operation of complex urban systems.¹⁶⁴ They also bring up the idea of using AI in informational systems for community public services and consumer intelligent homes. While this sounds good on paper, it's possible that over planning the use of AI and networking the cities together may backfire and create more issues, such as a higher vulnerability to hacking and more brittle systems. If critical infrastructure is networked, people's lives may come into danger with hacking disruptions.

5.1.6 Community Security

AI has potential to open up education, help prevent conflicts, and assist in human rights abuse responses¹⁶⁵ – all areas that help to provide Community Security, and in the case of human rights interventions, Political Security as well. Intelligent tutoring systems and personalized learning, have already begun to help individualize students' learning processes and have shown great promise.¹⁶⁶ Technology like this may make better education more accessible due to no longer needing a highly educated human resource for the job. Italy is interesting because it not only wants to enact education reform, but also to use AI to identify drop-out risk with students.¹⁶⁷

Many governments are planning and executing AI assisted education and skilling projects. India's AI strategy is inherently geared towards Community Security, as it delves into its desire to make AI equitable and for all people.¹⁶⁸ One of the ways they're trying to do this is with a huge AI education project called LeadingIndia.ai, with support from the UK's Royal Academy of

¹⁶⁴ Data: 4.2.5

¹⁶⁵ Beard et al., 2017

¹⁶⁶ Beck et al., 2005; Levesque, 2018

¹⁶⁷ Data: 4.2.9

¹⁶⁸ Data: 4.2.15

Engineering.¹⁶⁹ This project has the ambitious goal of spanning 1,000 institutions, 10,000 teachers and 1 million students; it has the potential to open up AI education for many areas in India and boost their overall AI competencies with potential for tremendous growth.

Canada is actively researching the implications of AI for society as well as how to make AI inclusive, diverse, and equitable.¹⁷⁰ They have a couple of workshops applicable to the impacts of AI on Community Security set for 2019.¹⁷¹ The first is “Indigenous protocol and AI” – using the indigenous people of Canada’s ways of relationship building to analyse ways in which AI-human ethical relationships can be developed. “Generation AI: Reducing inequality and enhancing digital inclusion via smart design and developmental science” is the other set of workshops applicable to Community Security; these are aimed at tackling issues of algorithms’ impact on children, with a special care taken to consider those in high levels of inequality and poverty. They hope to tackle the issue of AI as a catalyst for inequality’s growth. In a similar vein, the US emphasized that AI, and particularly superintelligence, needs to be developed for the common good.¹⁷²

Elderly care is an interesting domain that Japan is attempting to tackle with AI paired robotics.¹⁷³ One aspect of this is with support for care facilities and providers with AI run “communication robots” and watching devices.¹⁷⁴ An example of such systems is a robot that tracks how long a person has been in a room, has been laying or sitting, etc., and then prompts them to do some light exercise or get up and go out for a bit. Work is being done to expand this to also predictive behaviours to prevent issues rather than try to respond to them. There is some concern being expressed over possible ethical issues with constant camera

¹⁶⁹ LeadingIndia.ai, n.d.

¹⁷⁰ Data: 4.2.2

¹⁷¹ CIFAR, n.d.(b)

¹⁷² Data: 4.2.1

¹⁷³ Data: 4.2.3

¹⁷⁴ Acceptable Intelligence with Responsibility, 2018, pp. 58-64

supervision of people, though this may be avoided via using other sensors besides visual. Another area where AI and elderly care is intersecting include automated creation of care plans. China, Finland, and Germany are some of the others also working on developing smart elderly care systems.¹⁷⁵ Considering that elderly care is needed due to the increased health issues with age, there are also opportunities for AI to assist with non-age related disabilities as well. This is something that Italy and France are concentrating on.¹⁷⁶

However, Community Security is also at risk from AI. A set of risks stems from the existence of discrimination and bias within the Artificial Intelligence algorithms that are making important decisions about people's lives. The nature of the development of these algorithms is currently fundamentally opaque, which leads to issues in verifiability and legitimacy in the decision making processes that they are enacting. Bias can leech into an algorithm from being created from data that has discriminatory leanings, which is called algorithmic bias.¹⁷⁷ This can lead to situations of disparate impact (essentially non-intentional discrimination) upon disadvantaged groups such as racial minorities or the poor. This is a threat to the Community Security of those groups.

Many examples of algorithmic bias can already be found within our narrow AI decisioning systems. Some of these are simply offensive, such as Google+'s automatic photo classification identifying a bunch of African Americans as "Gorillas."¹⁷⁸ Other instances have a direct impact on the lives and liberty of individuals – such as AI systems that are used for the management of policing decisions. Some of these algorithms are being used to decide who gets parole based off of predicted recidivism risks,¹⁷⁹ while other are deciding what areas should get more police patrols.¹⁸⁰ France highlights these risks.¹⁸¹ These

¹⁷⁵ Data: 4.2.5, Data: 4.2.7, Data: 4.2.18

¹⁷⁶ Data: 4.2.9, Data: 4.2.10

¹⁷⁷ Osoba and Welser IV, 2017

¹⁷⁸ Zhang, 2015

¹⁷⁹ ProPublica, 2016; Osoba and Welser IV, 2017

¹⁸⁰ NYU Law News, 2016; O'Neil, 2016; Harris, 2017

impacts threaten both the Personal Security of the individuals at their mercy and the Community Security of the overall classes that are affected. If not properly mitigated through intentional development and policy creation, as AI systems gain power and complexity these impacts will only grow.

Managing these risks will not be easy. The literature surrounding solutions calls for candidness and intentionality in the creation of these AI systems. These systems' have issues with "lack of visibility, information asymmetry and hidden influence."¹⁸² Because of the way that machine learning works, even the data scientists creating the systems typically don't explicitly know how the AI is making its decisions. These processes need to be rebuilt with transparency, objectivity, and accountability in mind. "We need to switch the default in situations like this away from an assumption of secrecy, and toward the expectation that people deserve to know how they are rated and ranked."¹⁸³ The EU and G7 both believe that fairness must be integrated by design, with particular care taken to protect vulnerable, marginalized, and underrepresented populations.¹⁸⁴ Additionally, growing the diversity in the people who have influence in AI development may also help keep AI equitable. France in particular identified the need for balancing gender and suggested incentivising female enrolment in tech courses as a possible aid.¹⁸⁵

5.1.7 Political Security

The majority of the impacts AI may have on Political Security reside in areas of privacy, ethics, and morality. Almost all of the government strategies examined have some sort of relevancy to Political Security – it is the most widely covered dimension of Human Security. Repeated topics include the need for

¹⁸¹ Data: 4.2.10

¹⁸² Tufekci, 2015, p. 207

¹⁸³ Citron and Pasquale, 2014, p. 21

¹⁸⁴ Data: 4.2.11, Data: 4.2.17

¹⁸⁵ Data: 4.2.10

transparency, neutrality, and auditability in AI algorithms (especially in areas that affect lives such as judicial decisions and AWS), developing a global set of standards and norms (that don't stifle innovation), maintaining human centred development, and improvement of government processes.

There are some less common relevancies as well, though. For example, the US and Germany both emphasize the need to keep AI from subverting democratic processes.¹⁸⁶ Similarly, Italy wants to ensure individual rights are maintained.¹⁸⁷ In this vein, the EU insists that covert AI systems should be avoided to ensure maximum accountability.¹⁸⁸ Meanwhile Singapore's contribution lies in the need to ensure AI development is for the common good.¹⁸⁹ Finally, France's nonstandard addition to the topic is the suggestion that training AI computer scientists in the social sciences may help ensure they understand the implications of their work.¹⁹⁰

Ethics are most critical in the realms of AWS. The issue of machine morality is a hot topic in the AWS debate. Figuring out how to encode morality in and of itself is a vital yet immense challenge. Even if that technical challenge is overcome, which morality to encode still has to be decided. Legal codes, cultural values, and morals vary around the globe.¹⁹¹ The death penalty is a good example of this. Some countries still believe that it is a valid punishment, while others condemn the use of it. Morality's relativity will create issues when an actor wants to legitimize their usage of AWS. To compound these issues, laws and regulations for rapidly evolving technology tend to lag behind developmental breakthroughs – potentially creating a vacuum in which norms may emerge internally with those

¹⁸⁶ Data: 4.2.1, Data: 4.2.18

¹⁸⁷ Data: 4.2.9

¹⁸⁸ Data: 4.2.11

¹⁸⁹ Data: 4.2.4

¹⁹⁰ Data: 4.2.10

¹⁹¹ Russell, 2015; Osoba and Welser IV, 2017; Torres, 2017

working on the technologies without regards to outside moral standards or international law.¹⁹²

The future of privacy is another concern that arises with AI systems. Privacy and usability are unfortunately at odds with each other within data science; the more that data is scrubbed of personal details, the less useful it is for analysis.¹⁹³ It isn't as simple as just restricting the data that AI can use, either. AI has the potential to reconstruct restricted, sensitive fields from correlated data.¹⁹⁴ A basic example of this would be guessing someone's income range based off of the zip code that they live in. AI also can sometimes take anonymized data and de-anonymize it. One study found that all it took was gender, zip code, and birthdate to identify 63% of individuals.¹⁹⁵ That is still few enough variables that it could be done by a human easily. AI, however, is capable of much more complex calculations and inferences. Using only aggregated Facebook 'likes,' one study was able to guess with on average about 90% accuracy such attributes as the sexual orientation, race, gender, and political affiliation of users.¹⁹⁶ France in particular voices this concern quite strongly and they believe that the development of collective data rights may help to mitigate the risks.¹⁹⁷ India believes that there may even be mathematical ways to preserve privacy in these algorithms.¹⁹⁸

It could be easy for an AI system with surveillance goals to build up biases similar to those already emerging in areas where AI is assisting law enforcement. China has explicit plans for using AI in social governance and in the judicial system.¹⁹⁹ Hopefully China's extensive plans to research proper laws, regulations, and ethical norms in AI will help to keep the balance in this area. AI has the potential to negatively impact democracies and individual liberties if not

¹⁹² Bode and Huelss, 2018

¹⁹³ Lubarsky, 2017

¹⁹⁴ Osoba and Welser IV, 2017

¹⁹⁵ Golle, 2006

¹⁹⁶ Kosinski et al., 2013

¹⁹⁷ Data: 4.2.10

¹⁹⁸ Data: 4.2.15

¹⁹⁹ Data: 4.2.5

managed ethically. Canada has a workshop planned titled “AI-powered information ecosystems and democracy”;²⁰⁰ it is meant to study the potential negative impact AI may have on democratic institutions, concentrating on the polarizing dichotomy of the potentials for good and bad within AI.²⁰⁰

The UAE is also planning a social governance project, with a goal of rollout in 2020: the phase out of immigration officers with AI as a replacement.²⁰¹ They have already begun with iris scans and facial recognition; by the time 2020 comes around they hope to completely eliminate the need for human personnel with automated scanning where you just walk straight through. This does raise potential issues, however. The same bias issues could easily appear here, creating unfair or even prejudiced automated border control. There is also the risk of the government intentionally using it in a negatively discriminatory manner. Time will tell whether this National Security benefit will be a benefit for or thorn in the side of Human Security.

Singapore has released a framework for developing ethical AI.²⁰² This framework lays out 4 main factors to consider when developing AI, ending in a decision on how much human involvement should be necessary based on a risk assessment. The framework is based on the principles that AI systems and their decisions should be human-centric, fair, explainable, and transparent. They define 4 categorical factors that need to be considered with those principles in mind when creating an AI system: internal governance, risk management, operations management, and customer relationship management. Once a proper analysis has been done on these aspects, the decision of how much human intervention is ethically appropriate can be made (complete automation, human supervision and veto power, or human decision making with AI suggestions). Adhering to this framework should help companies ensure that they will be less likely to infringe upon human rights.

²⁰⁰ CIFAR, n.d. (b)

²⁰¹ Malek, 2018

²⁰² Singapore: Personal Data Protection Commission, 2019

There has been some work around using AI to identify and combat disinformation campaigns that threaten democracy by destabilizing democratic states' populations' understandings of the truth. This is a delicate issue however, as the accidental removal of valid voices can happen with automated disinformation response, which also threatens democratic foundations. A European Parliament report on “Regulating disinformation with artificial intelligence” attempted to address this controversy.²⁰³ They came to the conclusion that they recommended human judgement stay strongly in control of the review processes if AI is going to be involved in content mediation processes. Interestingly, Mexico thinks that this kind of anti-disinformation AI can also help in combating false portrayals during disasters.²⁰⁴

5.2 Summary and Discussion of AI Strategy Findings

Twenty four government strategies and pre-strategy documents from around the world were examined. There is much to both compare and contrast throughout them. A few trends can be identified. The most common is the increase in AI investment both technologically and in human resources with the goal of catalysing further growth in the industry and boosting other industries/economies. Many documents also have some aspect on skilling or reskilling to help balance potential job losses.

One of the investments that appeared quite often is the creation of a network of research institutes, and another is the development of DARPA-like organizations and AI challenges to stimulate AI growth. Another trend that can be seen is the development of open data policies with the intention of speeding up development by aiding in access to the precursor for AI: lots and lots of data. In a similar vein, many strategies emphasized the co-partnership of the public and private sectors, and in some cases even with military R&D. Less common, though

²⁰³ European Parliamentary Research Service, 2019

²⁰⁴ Data: 4.2.16

still prevalent, was the emergence of the need for AI ethics and regulation research.

There is a wide variety in the depth and breadth of what governments are doing with their AI strategy development. For example, China has by far the most comprehensive official AI strategy on the market right now (though France's pre-strategy research document is more comprehensive), while in contrast the UAE has a simple splash page and flashy YouTube video that seem to be more talk than substance currently. Some strategies, such as the Canadian, French, and Taiwanese, are fully funded plans, while others are simply guiding documents, like India, Italy, and Germany. The United States is a unique stand out because it has a guiding document, but it isn't considered strictly bonded to the overall goals of the country, rather individual government organizations have complete leeway to take things their own way with those general goals in mind; this leads to a mix of documents from different areas of the government, such as California or the DOD.

There is a wide variance in the coverage and variety of topics that the strategies are delving into. Environmental and Food Security have great potential for AI intervention, but are less tackled currently than they could be. Health Security, in contrast, is being widely taken on by many differing projects. Similarly, Community Security has a wealth of relevant projects to pull from, though with a surprisingly large concentration on elderly care. Food security, on the other hand, is rather uniform aside from a few stand out cases (such as China's desire to use AI for food hazard analysis).

Most of the strategies concentrate on the ways that AI can benefit societies without focusing on the insecurities that AI may bring. There are a few exceptions to this overly optimistic concentration, however. Namely, France, the EU, India, and Germany. There are many AI initiatives out there right now that have the potential to positively impact Human Security. Unfortunately there are also a few that have the capacity to do harm as well. The main facets of Human Security that have the biggest potential for something to go awry are Community Security,

Personal Security and Political Security; unfortunately these are also the three areas that impact people the most directly with harm. Protection for communities, and from violence and human rights violations are inherently more pressing than protection in, say, the economic or food realms (though this is not to say those aren't important too!). There also simply aren't many aspects where AI is likely to harm the other facets of Human Security, as most of the applications for AI either will be positive if they go well, or neutral if they fail.

If the Main Data table is examined in 4.2.1, other trends in the 24 documents can be seen as well. Particularly, the most well covered dimension of Human Security is Political at 22/24 documents addressing something to do with it; on the other hand, the least addressed is Food Security at 6/24. There is a lot of room for more governments to increase concentration on what AI can do for their populations' Food Security. Health, Personal, and Community security are relatively well covered, at 16-17 of the documents each. Economic and Environmental Security concerns could be addressed more though, with only around half (12-13) of the documents covering related aspects.

To examine the individual documents and their respective governments, only those with 6-7, and 1-2 dimensions will be highlighted – those with the most and least relevancy to Human Security. China, France, India, and Germany stand out as the only 4 with some relevancy to all 7 of the dimensions of Human Security. Finland and the EU come in 2nd with 6 each, both missing Food Security relevance in their documents. Japan can arguably be included in the 2nd category here as it has two relevant documents, and 6 dimensions are covered between them. Canada, Taiwan, the 2nd Japanese document, and the G7 documents are close to last with only 2 dimensions covered. Dead last is the second Swedish document, with relevance only to Political Security.

Moving to 4.2.2, the table and graph of variance, the trends in how many documents had each amount of possible relevancies to the dimensions of Human Security (from 1-7). Only 8 documents, 1/3 of the pool, were severely limited in their dimensional spread (1-3 relevant dimensions); of these only one document

was strictly limited to a single dimension, four documents had relevancies for 2 dimensions, and three were relevant to 3 dimensions. Five documents found themselves directly in the middle at 4 dimensions. There were 11 documents in the upper half with relevancies of 5, 6, and 7 dimensions; these numbered five, two, and four documents respectively. This is promising as it shows that there are distinctly more AI strategy documents that have higher or middling relevancies to the dimensions than there are with limited applicability (2/3 of the documents cover 4 or more dimensions).

6. Conclusion

This chapter brings the journey through the intersectional examination of the topics of AI, Human Security, and government AI strategies to a close. It will first provide a summary of the research and outcomes related to the research question and objectives presented in section 1.3. After this, considerations surrounding the implications of the analysis and some recommendations for further work will be provided. To close, this chapter will present the researcher's final thoughts surrounding the intersectionality of AI & Human Security.

6.1 Summary of Research and Outcomes

This dissertation began by familiarizing the reader with the general history and current state of Artificial Intelligence, in order to provide a minimum of understanding for this complex technology for those that are unfamiliar with it. In order to comprehend how AI is affecting Human Security, you must first understand AI itself. The historical pattern surrounding the exciting leaps in progress (AI Waves) and the sobering periods of restrained expectations (AI Winters) that have ensued since the 1950's was explained. Then a basic description of the divergent strengths of AI (ANI/AGI/ASI, or weak/strong AI) and some of the most common styles of AI engineering were provided.

Following that grounding in the concepts surrounding AI's history and development, the Literature Review fleshed out the differing ideas surrounding Human Security. This began with the originating 1994 UN HDR definition and its breakdown further into the 7 dimensions of Economic, Food, Health, Environmental, Personal, Community, and Political Security. Then conceptions of the other takes on the subject that have evolved from this definition were examined in the literature. The Canadian, Swiss, and Norwegian definitions were found to concentrate primarily on the 'freedom from fear' and protection from physical violence aspects of the topic, while the Japanese view is wider in scope

and closer to the originating conception with an emphasis on freedom from fear, want, and indignity.

Various criticisms surrounding the usages of Human Security theories were also addressed; these were discarded as invalid concerns with the intended use of the concept for this dissertation. Some of these criticisms included: the concept is too vague or broad to be truly useful, it is simply a discursive false hope for many cases of injustice in undeveloped countries, its application is overly arbitrary despite its roots in equal access humanitarianism, or that the idea of ‘human’ is taken for granted and has historically excluded some groups (women, racial minorities, queer communities). None of the objections to the concept particularly get in the way of its use for categorical analysis, however.

The original 1994 UN HDR definition was eventually settled on as the most fitting to use as a framework to analyse both the effects of AI in general and how the current landscape of government AI strategies are relevant to Human Security. This was because it is less politicised due to it being country-neutral and because of its useful breakdown into the 7 dimensions. This decision allowed for a more specific analysis than a more generic definition that only concentrates on some incomplete mix of ‘freedom from fear, want, and/or indignity’ could.

Once the core concepts surrounding AI and Human Security were established, a thorough investigation of all 24 government strategic documents on AI that currently exist to develop a database of which strategies cover what dimensions of Human Security. The documents were read through once and summarized, and then re-examined via keyword analysis to highlight the specific relevancies to the dimensions of Human Security that exist in each of them. This allowed for the creation of a basic set of binary data (either relevant or not relevant to each of the 7 dimensions), tables, a graph, and some light statistical analysis of the human security relevancy currently being provided in the strategies. Subsequently, the patterns of commonalities and outliers within the government AI strategy landscape were then revealed, to show what areas are being well covered by today’s strategies and where there is room for growth.

6.2 Implications and Future Recommendations

The Pragmatic philosophical basis allowed for the examination of AI from a perspective grounded in core non-subjective values while breaking out of the purely Realist, state-centric mind set, shifting the referent object of security to both humanity as a whole and the individual. This was done through a productive securitization of the concepts of human development. The findings of this study could be used to develop more mindful and holistic strategies and policies surrounding AI development and usage in the future; this would in turn help to guide AI into beneficial effects on mankind. It can also help to demystify AI from a strategic perspective. It provides a good snapshot of how governments understand and are trying to use AI currently. Policymakers could use the raw data and representative tables and graph to see what they themselves could do better or how they're stacking up against the rest of the AI strategic terrain. Additionally, this research may be useful to a technology engineering team that is concentrating on developing beneficial AI.

This research presented a process where the 7 dimensions of Human Security can be used as a framework with which holistically pursue systemic analysis of the effects of a technology. It enables development of an understanding of how intersectional topics are positively and negatively affecting humanity through the prevention and propagation of fear, want, and indignities. This dissertation was specifically concentrated on governmental AI strategies and supporting cases, but it could also be extended to analyse many other areas. For example, business development of AI and AI law could also be examined in this light. It could be particularly useful for analysis of other controversial technology applications, too. Some possible avenues for application in similar domains could be nuclear technology, individual sectors of robotics such as nanotechnology, genetic engineering, and even new societal systems such as citizen rating/social credit. Expanded use of Human Security as a framework for analysis could be of great benefit, especially for holistic policy development.

6.2.1 AI and Human Security – Traditional and Emerging Security Dilemmas

AI development that affects Human Security has implications in the traditional sense of a realist sovereign state security dilemma, just with a new technology. The threat of an AI arms race, particularly with AWS, conforms to the ideas of a traditional security dilemma. One country builds up AWS systems, and others are forced to do the same to keep up – even if the first country was only building them for defensive purposes. This threatens to spiral into a situation where actors are cutting corners to speed up development, adding more risk into the situation. Also, if someone decides to remove human control on the decisioning of the AWS, then others may be forced to do the same; this may cause a loss of control of the battlefield.

Another type of security dilemma is also emerging with AI growth – one that involves threats to Human Security rather than that of the state. In some cases, the process of trying to secure one facet of Human Security may come into conflict with other facets. For example, an attempt to increase Personal Security by employing AI for more efficient policing will likely negatively impact Community and/or Political Security; this might happen due to biases in the algorithms or a loss of privacy rights from how the AI systems are consuming data. The issue of data and privacy being at odds with the development of AI systems is one of the core dilemmas of AI development, especially in AI systems that work with sensitive data such as health or sensitive outcomes like policing and the judicial system. The potential for non-state centric security dilemma emergence is another avenue for future work.

6.3 Final Thoughts

Some applications of AI are more dangerous than others, with more possibilities for physical harm or infringement upon democratic liberties. Essentially any system that influences the life or liberty of individuals should be closely watched and managed even more delicately than purely scientific or business oriented AI. Projects such as Autonomous Weapons Systems, judicial

decision assistance, and big data fuelled surveillance are some examples of where AI is already being developed with questionable impacts.

The perspective taken here has been one concentrating on governmental strategic documents, but there are many others that also need to be managed and maintained as well. Artificial Intelligence is a controversial and revolutionary technology that is in its adolescence. There is a vast potential for beneficial AI engineering, but the more powerful the technology gets, the more careful society has to be to retain the positive effects while reining in the potential for harm. There is still a lot of development on the horizon, and as this area of technology matures, it needs to be with conscious effort to ensure the advantages that it brings are not altogether unbalanced in their spread among humanity or outweighed by negative side effects on others. An AI that is built for the benefit of the few, at the expense of the many, will be harmful to Human Security. Fairness, accountability, and transparency need to be at the forefront of all AI with any impact on human life.

References

Acceptable Intelligence with Responsibility. (2018). *Perspectives on Artificial Intelligence/Robotics and Work /Employment*. [online] Available at: http://sig-air.org/wp/wp-content/uploads/2018/07/PerspectivesOnAI_2018.pdf [Accessed 24 Apr. 2019].

Adger et al. (2014). *Human Security*. In: *Climate Change 2014: Impacts, Adaptation, and Vulnerability*. [online] Available at: https://www.ipcc.ch/site/assets/uploads/2018/02/WGIIAR5-Chap12_FINAL.pdf [Accessed 27 Feb. 2019].

AI Singapore. (n.d. a). *About AI Singapore: Our mission, vision and goals*. [online] Available at: <https://www.aisingapore.org/about-us/> [Accessed 4 Apr. 2019].

AI Singapore. (n.d. b). *AI in Health Grand Challenge*. [online]. Available at: <https://www.aisingapore.org/grand-challenges/health/> [Accessed 24 Apr. 2019].

AI Taiwan. (2018). *Cabinet plans to develop the nation's AI industry*. [online] Available at: <https://ai.taiwan.gov.tw/news/cabinet-plans-to-develop-the-nations-ai-industry/> [Accessed 11 Apr. 2019].

AI Taiwan. (n.d.) *AI Taiwan*. [online] Available at: <https://ai.taiwan.gov.tw/#actionplan> [Accessed 11 Apr. 2019].

Alkire, S. (n.d.). *A Conceptual Framework for Human Security*. [online] Available at: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.580.2805&rep=rep1&type=pdf> [Accessed 2 Mar. 2019].

Australia: Department of Industry, Innovation and Science. (2017). *Australia 2030: Prosperity through Innovation*. [government publication] Available at: <https://www.industry.gov.au/sites/g/files/net3906/f/May%202018/document/p>

df/australia-2030-prosperity-through-innovation-full-report.pdf [Accessed 29 Apr. 2019].

Beard, S. et al. (2017). *Written Evidence to the United Kingdom's House of Lords Select Committee on Artificial Intelligence*. [online] Data.parliament.uk. Available at:

<http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/artificial-intelligence-committee/artificial-intelligence/written/69655.html> [Accessed 8 Mar. 2019].

Bode, I. and Huelss, H. (2018). *Autonomous weapon systems and changing norms in international relations*, *Review of International Studies*, 44(3), pp. 393-413. doi: 10.1017/S0260210517000614.

Bosold, D. and Werthes, S. (2005). *Human Security in Practice: Canadian and Japanese Experiences*. [online] Available at:

https://www.researchgate.net/publication/237671115_Human_Security_in_Practice_Canadian_and_Japanese_Experiences [Accessed 17 Feb. 2019].

Bostrom, N. (1997). *HOW LONG BEFORE SUPERINTELLIGENCE?*

[online] Available at: <https://nickbostrom.com/superintelligence.html> [Accessed 16 May 2019].

Bowen, G. (2009). *Document Analysis as a Qualitative Research Method*.

[online] Available at:

https://www.researchgate.net/profile/Glenn_Bowen/publication/240807798_Document_Analysis_as_a_Qualitative_Research_Method/links/59d807d0a6fdc2aad065377/Document-Analysis-as-a-Qualitative-Research-Method.pdf [Accessed 4 Jun. 2019].

BRIDGE Horizon 2020. (n.d.). *Projects*. [online] Available at:

<https://www.h2020-bridge.eu/participant-projects/> [Accessed 27 Apr. 2019].

Brundage et al. (2018). *The Malicious Use of Artificial Intelligence:*

Forecasting, Prevention, and Mitigation. [online] Available at:

<https://arxiv.org/ftp/arxiv/papers/1802/1802.07228.pdf> [Accessed 21 May. 2019].

Buzan, B. (2004). *A Reductionist, Idealistic Notion that Adds Little Analytical Value*, *Security Dialogue*, 35(3), pp. 369-370. doi: 10.1177/096701060403500326.

Buzan, B., Wilde, J. and Wæver, O. (1997). *Security: A New Framework for Analysis*, Lynne Rienner Publishers, London.

California. (2018). *Assembly Concurrent Resolution No. 215, CHAPTER 206: Relative to the 23 Asilomar AI Principles*. [government publication] Available at:

http://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201720180ACR215 [Accessed 23 Apr. 2019].

Canadian Institutes of Health Research. (2017). *Artificial Intelligence for Public Health*. [online] Available at: <https://www.researchnet-recherchenet.ca/rnr16/viewOpportunityDetails.do?progCd=10822&language=E&fodAgency=CIHR&view=browseArchive&browseArc=true&org=CIHR> [Accessed 23 Apr. 2019].

Canadian Institutes of Health Research. (2018). *Application of Artificial Intelligence Approaches to Tackle Public Health Challenges – Workshop Report*. [online] Available at: http://www.cihr-irsc.gc.ca/e/documents/artificial_intelligence_approaches-en.pdf [Accessed 23 Apr. 2019].

Canadian Institutes of Health Research. (2019). *Equitable AI: activating opportunities*. [online] Available at: <http://www.cihr-irsc.gc.ca/e/51204.html> [Accessed 23 Apr. 2019].

Caravagna et al. (2018). *Detecting repeated cancer evolution in human tumours from multi-region sequencing data*, *Nat Methods*, 15(9), pp. 707–714. doi:10.1038/s41592-018-0108-x.

China: State Council. (2017). *Notice of the New Generation Artificial Intelligence Development Plan*. [online] Available at: http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm [Accessed 8 Apr. 2019].

Chishti, Z. (2018). *Artificial intelligence: winter is coming*. [online] Available at: <https://www.ft.com/content/47111fce-d0a4-11e8-9a3c-5d5eac8f1ab4> [Accessed 16 May 2019].

Christie, R. (2010). *Critical Voices and Human Security*, Security Dialogue, 41(2), pp. 169-190. doi: 10.1177/0967010610361891.

CIFAR. (n.d. a). *CIFAR Pan-Canadian Artificial Intelligence Strategy*. [online] Available at: <https://www.cifar.ca/ai/pan-canadian-artificial-intelligence-strategy> [Accessed 2 Apr. 2019].

CIFAR. (n.d. b). *First Call for AI & Society Workshops*. [online] Available at: <https://www.cifar.ca/ai/ai-society/ai-society-workshops-call-one> [Accessed 23 Apr. 2019].

Citron, D. and Pasquale, F. (2014). *The Scored Society: Due Process for Automated Predictions*. [online] Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2376209 [Accessed 12 Mar. 2019].

Cummings, M. (n.d.). *Automation Bias in Intelligent Time Critical Decision Support Systems*. [online] Available at: <http://hal.pratt.duke.edu/sites/hal.pratt.duke.edu/files/u13/Automation%20Bias%20in%20Intelligent%20Time%20Critical%20Decision%20Support%20Systems.pdf> [Accessed 12 Mar. 2019].

DeepMind. (n.d.) *The story of AlphaGo so far*. [online] Available at: <https://deepmind.com/research/alphago/> [Accessed 16 May 2019].

Denmark: Ministry of Industry, Business and Financial Affairs. (2018). *Strategy for Denmark's Digital Growth*. [government publication] Available at: https://eng.em.dk/media/10566/digital-growth-strategy-report_uk_web-2.pdf [Accessed 29 Apr. 2019].

Desouza et al. (2017). *Learning from public sector experimentation with artificial intelligence*. [online] Available at: <https://www.brookings.edu/blog/techtank/2017/06/23/learning-from-public-sector-experimentation-with-artificial-intelligence/> [Accessed 17 Mar. 2019].

Dewey, J. (1929). *The Quest for Certainty: Chapter X: The Construction of Good*. [online] Available at: <https://www.giffordlectures.org/books/quest-certainty-study-relation-knowledge-and-action/chapter-x-construction-good> [Accessed 3 Jun. 2019].

dudeperf3ct. (2018). *History, Waves and Winters in AI, Hacker Noon*. [online] Available at: <https://hackernoon.com/history-waves-and-winters-in-ai-dd5feb558e45> [Accessed 15 May 2019].

EMPATTICS. (n.d.). *About Us*. [online] Available at: <http://empattics.eu/about-us/> [Accessed 28 Apr. 2019].

enCOMPASS. (n.d.). *Objectives*. [online] Available at: <http://www.encompass-project.eu/project/objective/> [Accessed 28 Apr. 2019].

Etzioni, A. and Etzioni, O. (2017). *Pros and Cons of Autonomous Weapons Systems*. *Military Review*, pp.72-81. [online] Available at: <https://www.armyupress.army.mil/Portals/7/military-review/Archives/English/pros-and-cons-of-autonomous-weapons-systems.pdf> [Accessed 17 Mar. 2019].

European Parliamentary Research Service. (2019). *Regulating disinformation with artificial intelligence*. [government publication] Available at: [http://www.europarl.europa.eu/RegData/etudes/STUD/2019/624279/EPRS_STU\(2019\)624279_EN.pdf](http://www.europarl.europa.eu/RegData/etudes/STUD/2019/624279/EPRS_STU(2019)624279_EN.pdf) [Accessed 28 Apr. 2019].

European Union. (2018). *Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions: Coordinated Plan on Artificial Intelligence*. [government publication] Available at: <https://ec.europa.eu/digital-single-market/en/news/coordinated-plan-artificial-intelligence> [Accessed 16 Apr. 2019].

European Union: AI High-Level Expert Group. (2019). *ETHICS GUIDELINES FOR TRUSTWORTHY AI*. [government publication] Available at: <https://ec.europa.eu/futurium/en/ai-alliance-consultation> [Accessed 26 Apr. 2019].

Faggella, D. (2019). *Will There Be Another Artificial Intelligence Winter? Probably Not*. [online] Available at: <https://emerj.com/ai-executive-guides/will-there-be-another-artificial-intelligence-winter-probably-not/> [Accessed 16 May 2019].

Finland: Ministry of Economic Affairs and Employment. (2017). *Finland's Age of Artificial Intelligence - Turning Finland into a leading country in the application of artificial intelligence: Objective and recommendations for measures*. [government publication] Available at: http://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/160391/TEMrap_47_2017_verkkojulkaisu.pdf?sequence=1&isAllowed=y [Accessed 10 Apr. 2019].

French Government. (2018). *Artificial Intelligence: "Making France a leader"*. [online] Available at: <https://www.gouvernement.fr/en/artificial-intelligence-making-france-a-leader> [Accessed 13 Apr. 2019].

Frey, C. and Osborne, M. (2015). *Technology at Work: The Future of Innovation and Employment*. [online] Oxford Martin School and Citi GPS: Global Perspectives & Solutions, pp.67-77. Available at: https://www.oxfordmartin.ox.ac.uk/downloads/reports/Citi_GPS_Technology_Work.pdf [Accessed 8 Mar. 2019].

Future of Life Institute. (2015). *Open Letter on Autonomous Weapons*. [online] Available at: <https://futureoflife.org/open-letter-autonomous-weapons/> [Accessed 13 Mar. 2019].

Future of Life Institute. (n.d.). *Global AI Policy*. [online] Available at: <https://futureoflife.org/ai-policy/> [Accessed 4 Apr. 2019].

G7. (2018). *Charlevoix Common Vision for the Future of Artificial Intelligence*. [online] Available at: <http://www.g7.utoronto.ca/summit/2018charlevoix/ai-commitment.html> [Accessed 21 Apr. 2019].

Germany. (2018). *Artificial Intelligence Strategy*. [government publication] Available at: <https://www.ki-strategie-deutschland.de/home.html> [Accessed 21 Apr. 2019].

Global Challenges Foundation. (2017). *Global Catastrophic Risks 2017*, pp. 60-65. [online] Available at: <https://api.globalchallenges.org/static/files/Global%20Catastrophic%20Risks%202017.pdf> [Accessed 18 Mar. 2019].

Golle,P. (2006). *Revisiting the Uniqueness of Simple Demographics in the US Population*, Palo Alto Research Center. [online] Available at: <https://www.privacylives.com/wp-content/uploads/2010/01/golle-reidentification-deanonymization-2006.pdf> [Accessed 13 Mar. 2019].

Harris, M. (2017). *HOW PETER THIEL'S SECRETIVE DATA COMPANY PUSHED INTO POLICING*. [online] Available at: <https://www.wired.com/story/how-peter-thiels-secretive-data-company-pushed-into-policing/> [Accessed 11 Mar. 2019].

Hart, R. (2017). *The case for banning autonomous weapons rests on morality, not practicality*. [online] Bulletin of the Atomic Scientists. Available at: <https://thebulletin.org/case-banning-autonomous-weapons-rests-morality-not-practicality10707> [Accessed 13 Mar. 2019].

Human Security Index. (n.d.). *About*. [online] Available at:
http://www.humansecurityindex.org/?page_id=2 [Accessed 3 Mar. 2019].

India: NITI Aayog. (2018). *National Strategy for Artificial Intelligence: #AIFORALL*. [government publication] Available at:
http://niti.gov.in/writereaddata/files/document_publication/NationalStrategy-for-AI-Discussion-Paper.pdf [Accessed 19 Apr. 2019].

International Development Research Centre. (2018). *Artificial intelligence and human development*. [online] Available at:
<https://www.idrc.ca/en/stories/artificial-intelligence-and-human-development> [Accessed 23 May 2019].

Italy: Agency for Digital Italy. (2018). *AI White Paper*. [government publication] Available at: <https://buildmedia.readthedocs.org/media/pdf/ai-white-paper/latest/ai-white-paper.pdf> [Accessed 17 Apr. 2019].

Japan: Committee on Impact and Risk Assessment. (2017). *Draft AI R&D GUIDELINES for International Discussions*. [government publication] Available at: http://www.soumu.go.jp/main_content/000507517.pdf [Accessed 24 Apr. 2019].

Japan: Strategic Council for AI Technology. (2017). *Artificial Intelligence Technology Strategy*. [government publication] Available at:
<https://www.nedo.go.jp/content/100865202.pdf> [Accessed 3 Apr. 2019].

Jones et al. (2018). *Artificial intelligence, machine learning and the evolution of healthcare: a bright future or cause for concern?* Bone & Joint Research, 7(3), pp. 223-225. doi: 10.1302/2046-3758.73.BJR-2017-0147.R1

Juul, M. (2019). *Human Security at 25, Statement by Ambassador Mona Juul at Human Security at 25: Building on its contributions to achieve the SDGs*. [online] Available at:
<https://www.norway.no/en/missions/UN/statements/other-statements/2019/human-security-at-25/> [Accessed 3 Mar. 2019].

Kaasinen, V. (2019). *Why AI-driven district heating plays an important role in the climate challenge?*, Fortum. [online] Available at:

<https://www.fortum.com/about-us/blog/forenergy-blog/why-ai-driven-district-heating-plays-important-role-climate-challenge> [Accessed 27 Apr. 2019].

Kania, E. (2017). *The critical human element in the machine age of warfare.*

[online] Bulletin of the Atomic Scientists. Available at:

<https://thebulletin.org/critical-human-element-machine-age-warfare11277> [Accessed 13 Mar. 2019].

Khong, Y. (2001). *Human Security: A Shotgun Approach to Alleviating*

Human Misery? Global Governance, 7(3), 231-236. [online] Available at:

<http://www.jstor.org/stable/27800301> [Accessed 23 Feb. 2019].

Knight, W. (2016). *AI Winter Isn't Coming.* [online] Available at:

<https://www.technologyreview.com/s/603062/ai-winter-isnt-coming/> [Accessed 16 May 2019].

Kosinski et al. (2013). *Private traits and attributes are predictable from*

digital records of human behaviour, PNAS, 110(15), pp. 5802-5805. [online]

Available at: <https://doi.org/10.1073/pnas.1218772110> [Accessed 14 Mar. 2019].

LeadingIndia.ai. (n.d.). *About LeadingIndia.ai.* [online] Available at:

<https://www.leadingindia.ai/> [Accessed 28 Apr. 2019].

Lee, K. (2017). *The Real Threat of Artificial Intelligence.* [online]

Nytimes.com. Available at:

<https://www.nytimes.com/2017/06/24/opinion/sunday/artificial-intelligence-economic-inequality.html> [Accessed 11 Mar. 2019].

Levesque, E. (2018). *The role of AI in education and the changing US*

workforce. [online] Available at: [https://www.brookings.edu/research/the](https://www.brookings.edu/research/the-role-of-ai-in-education-and-the-changing-u-s-workforce/)

[role-of-ai-in-education-and-the-changing-u-s-workforce/](https://www.brookings.edu/research/the-role-of-ai-in-education-and-the-changing-u-s-workforce/) [Accessed 18 Mar. 2019].

Lubarsky, B. (2017). *Re-Identification of “Anonymized” Data*, Georgetown Law Technology Review, pp. 202-213. [online] Available at: <https://georgetownlawtechreview.org/re-identification-of-anonymized-data/GLTR-04-2017/> [Accessed 13 Mar. 2019].

Lundmark, P. (2019). *Artificial intelligence holds a substantial promise for the power industry*, Fortum. [online] Available at: <https://www.fortum.com/about-us/blog/forenergy-blog/artificial-intelligence-holds-substantial-promise-power-industry> [Accessed 27 Apr. 2019].

Malek, C. (2018). *UAE to phase out immigration officers in favour of AI by 2020*, The National. [online] Available at: <https://www.thenational.ae/uae/government/uae-to-phase-out-immigration-officers-in-favour-of-ai-by-2020-1.704633> [Accessed 26 Apr. 2019].

Marhia, N. (2013) *Some humans are more Human than Others: Troubling the ‘human’ in human security from a critical feminist perspective*, Security Dialogue, 44(1), pp. 19–35. doi: 10.1177/0967010612470293.

McCarthy et al. (1955). *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*. [online] Available at: <http://raysolomonoff.com/dartmouth/boxa/dart564props.pdf> [Accessed 14 May 2019].

McCormack, T. (2008). *Power and agency in the human security framework*, Cambridge Review of International Affairs, 21(1), pp. 113-128. doi: 10.1080/09557570701828618.

Ministry of Foreign Affairs of Japan. (2009). *The Trust Fund for Human Security*. [online] Available at: https://www.mofa.go.jp/policy/human_secu/t_fund21.pdf [Accessed 20 Feb. 2019].

Nordic Council of Ministers. (2018). *AI in the Nordic-Baltic region*. [government publication] Available at:

https://www.regeringen.se/49a602/globalassets/regeringen/dokument/naringsdepartementet/20180514_nmr_deklaration-slutlig-webb.pdf [Accessed 19 Apr. 2019].

NYU Law News. (2016). *Predictive policing and human rights on the table at second annual Bernstein Institute conference*. [online] Available at: <http://www.law.nyu.edu/news/predictive-policing-human-rights-analytics> [Accessed 11 Mar. 2019].

O'Neil, C. (2016). *Weapons of Math Destruction*. [book] Crown Publishing.

Osoba, O. and Welser IV, W. (2017). *An Intelligence in Our Image: The Risks of Bias and Errors in Artificial Intelligence*. [online] Santa Monica, California: RAND Corporation. Available at: https://www.rand.org/content/dam/rand/pubs/research_reports/RR1700/RR1744/RAND_RR1744.pdf [Accessed 10 Mar. 2019].

Oxford Insights and C Minds. (2018). *TOWARDS AN AI STRATEGY IN MEXICO: Harnessing the AI Revolution*. [online] Available at: https://docs.wixstatic.com/ugd/7be025_e726c582191c49d2b8b6517a590151f6.pdf [Accessed 21 Apr. 2019].

Paris, R. (2001). *Human Security: Paradigm Shift or Hot Air?* [online] Available at: <http://nghiencuuquocte.net/wp-content/uploads/2014/03/Human-security-paradigm-shift-or-hot-air.pdf> [Accessed 20 Feb. 2019].

Piekiewicz, F. (2018). *AI Winter Is Well On Its Way*. [online] Available at: <https://blog.piekiewicz.info/2018/05/28/ai-winter-is-well-on-its-way/> [Accessed 16 May 2019].

PricewaterhouseCoopers. (2017). *Sizing the Prize: What's the Real Value of AI for Your Business and How Can You Capitalise?* [online] Available at: <https://www.pwc.com/gx/en/issues/analytics/assets/pwc-ai-analysis-sizing-the-prize-report.pdf> [Accessed 16 Mar. 2019].

- ProPublica. (2016). *Machine Bias*. [online] Available at: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> [Accessed 11 Mar. 2019].
- Risse, N. (2019). *Officials Stress Relevance of Human Security in SDG Era*. [online] Available at: <http://sdg.iisd.org/news/officials-stress-relevance-of-human-security-in-sdg-era/> [Accessed 6 Mar. 2019].
- Rojanasakul, M. et al. (2017). *Is Your Job About To Disappear?: QuickTake*. [online] Bloomberg.com. Available at: <https://www.bloomberg.com/graphics/2017-jobs-automation-risk/> [Accessed 9 Mar. 2019].
- Russell, S. (2015). *Value Alignment | Stuart Russell*. [video] Available at: https://www.youtube.com/watch?v=WvmeTaFc_Qw&list=PL7m903CwFUgllnhKsy3po_13bpZfao2_U&index=7 [Accessed 12 Mar. 2019].
- Russia. (2019). *The list of instructions following the meeting of the supervisory board of the Agency for Strategic Initiatives*. [online] Available at: <http://kremlin.ru/acts/assignments/orders/59758> [Accessed 23 Apr. 2019].
- Russia: Ministry of Defence. (2018). *Conference "Artificial Intelligence: Problems and Solutions" - 2018 "*. [online] Available at: <http://mil.ru/conferences/is-intellekt.htm> [Accessed 23 Apr. 2019].
- Shani, G. (2017). *Human Security as ontological security: a post-colonial approach*, *Postcolonial Studies*, 20(3), pp. 275-293. doi: 10.1080/13688790.2017.1378062.
- Singapore: Centre for Finance, Technology and Entrepreneurship. (n.d.). *ARTIFICIAL INTELLIGENCE IN FINANCE*. [online] Available at: <http://aifinance.sg/> [Accessed 25 Apr. 2019].
- Singapore: Personal Data Protection Commission. (2019). *RESPONSIBLE AI MADE EASY: FOR ORGANISATIONS*. [government publication] Available

at: <https://www.pdpc.gov.sg/-/media/Files/PDPC/PDF-Files/Resource-for-Organisation/AI/Model-AI-Gov-Framework-Primer.pdf> [Accessed 25 Apr. 2019].

Solomon, S. (2019). *Israel needs national vision for AI or risks falling behind, tech authority says*. [online] Available at: <https://www.timesofisrael.com/israel-needs-national-vision-for-ai-or-risk-falling-behind-tech-authority-warns/> [Accessed 23 Apr. 2019].

South Korea. (2018). *Investment of KRW 2.2 trillion to secure world-class artificial intelligence technology*. [government publication] Available at: <https://www.msit.go.kr/web/msipContents/contentsView.do?cateId=mssw311&artId=1382727> [Accessed 18 Apr. 2019].

Spiegel, J. and Huish, R. (2011). *Canadian foreign aid for global health: Human security opportunity lost*. [online] Available at: <https://www.tandfonline.com/doi/abs/10.1080/11926422.2009.9673492> [Accessed 20 Feb. 2019].

Sweden: Ministry of Enterprise and Innovation. (2018). *National approach to artificial intelligence*. [government publication] Available at: <https://www.regeringen.se/4aa638/contentassets/a6488cceb6f418e9ada18bae40bb71f/national-approach-to-artificial-intelligence.pdf> [Accessed 19 Apr. 2019].

Swiss Federal Department of Foreign Affairs. (2018). *Human Security Division*. [online] Available at: <https://www.fdfa.admin.ch/eda/en/home/fdfa/organisation-fdfa/directorates-divisions/directorate-political-affairs/hsd.html> [Accessed 28 Feb. 2019].

Tadjbakhsh, S. (2010). *HUMAN SECURITY BEYOND RESPONSIBILITY TO PROTECT*. [online] Available at: https://www.academia.edu/3487405/_Human_Security_Beyond_Responsibilit

y_to_Protect_Revista_Espa%C3%B1ola_de_Desarrollo_y_Cooperaci%C3%B3n_No._26_Vol_2_2010_pp._13-37 [Accessed 1 Mar. 2019].

Tadjbakhsh, S. (2014). *Human security twenty years on*. [online] Available at: <https://www.files.ethz.ch/isn/181368/540cb240aa84ac7133bce008adcde01f.pdf> [Accessed 1 Mar. 2019].

Torres, P. (2017). *Yes, superintelligence is a threat that should be taken seriously*. [online] Bulletin of the Atomic Scientists. Available at: <https://thebulletin.org/why-superintelligence-threat-should-be-taken-seriously11219> [Accessed 13 Mar. 2019].

Tufekci, Z. (2015). *ALGORITHMIC HARMS BEYOND FACEBOOK AND GOOGLE: EMERGENT CHALLENGES OF COMPUTATIONAL AGENCY*. [online] Available at: <http://ctlj.colorado.edu/wp-content/uploads/2015/08/Tufekci-final.pdf> [Accessed 12 Mar. 2019].

Turing, A. (1950). I.— *Computing Machinery and Intelligence*, *Mind*, 59 (236), pp. 433-460. [online] Available at: <https://doi.org/10.1093/mind/LIX.236.433> [Accessed 14 May 2019].

UAE. (2018). *UAE Strategy for Artificial Intelligence*. [government publication] Available at: <https://government.ae/en/about-the-uae/strategies-initiatives-and-awards/federal-governments-strategies-and-plans/uae-strategy-for-artificial-intelligence> [Accessed 9 Apr. 2019].

UAEai 2031. (2017). *UAE Artificial Intelligence Strategy 2031*. [video] Available at: <https://www.youtube.com/watch?v=6P-5PkzC2ZI> [Accessed 9 Apr. 2019].

UAEai. (n.d.). *UAE Artificial Intelligence Strategy*. [online] Available at: <http://www.uaesai.ae/en/> [Accessed 9 Apr. 2019].

United Kingdom. (2018). *Industrial Strategy: Artificial Intelligence Sector Deal*. [government publication] Available at:

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/702810/180425_BEIS_AI_Sector_Deal__4_.pdf
[Accessed 16 Apr. 2019].

United Nations. (1994). *Human Development Report 1994*. [online] Available at:
http://hdr.undp.org/sites/default/files/reports/255/hdr_1994_en_complete_nostats.pdf [Accessed 15 Feb. 2019].

United Nations. (1994). *Human Development Report 1994*. [online] Available at:
http://hdr.undp.org/sites/default/files/reports/255/hdr_1994_en_complete_nostats.pdf

United Nations: Human Security Unit. (2009). *Human Security in Theory and Practice: Application of the Human Security Concept and the United Nations Trust Fund for Human Security*. [online] Available at:
<https://www.unocha.org/sites/dms/HSU/Publications%20and%20Products/Human%20Security%20Tools/Human%20Security%20in%20Theory%20and%20Practice%20English.pdf> [Accessed 4 Jun. 2019].

United States: Department of Defense. (2019). *Summary of the 2018 Department of Defense Artificial Intelligence Strategy: Harnessing AI to Advance Our Security and Prosperity*. [government publication]. Available at:
<https://media.defense.gov/2019/Feb/12/2002088963/-1/-1/1/SUMMARY-OF-DOD-AI-STRATEGY.PDF> [Accessed 18 Apr. 2019].

United States: Networking and Information Technology Research and Development Subcommittee. (2016). *The National Artificial Intelligence Research and Development Strategic Plan*. [government publication]. Available at: https://www.nitrd.gov/PUBS/national_ai_rd_strategic_plan.pdf [Accessed 18 Apr. 2019].

Victor, D. (2019). *How artificial intelligence will affect the future of energy and climate*. [online] Available at: <https://www.brookings.edu/research/how-artificial-intelligence-will-affect-the-future-of-energy-and-climate/> [Accessed 18 Mar. 2019].

Villani, C. (2018). *For a Meaningful Artificial Intelligence: Towards a French and European Strategy*. [government publication] Available at: https://www.aiforhumanity.fr/pdfs/MissionVillani_Report_ENG-VF.pdf [Accessed 27 Apr. 2019].

Vinnova. (2018). *Artificial Intelligence in Swedish Business and Society – Analysis of development and potential*. [online] Available at: https://www.vinnova.se/contentassets/29cd313d690e4be3a8d861ad05a4ee48/vr_18_09.pdf [Accessed 6 Jun. 2019].

Wæver, O. (1995). 'Securitization and Desecuritization', in Lipschutz, R. (ed.) *On Security*, Columbia University Press, New York.

West, D. and Allen, J. (2018). *How artificial intelligence is transforming the world*. [online] Brookings. Available at: <https://www.brookings.edu/research/how-artificial-intelligence-is-transforming-the-world/> [Accessed 17 Mar. 2019].

Winter, C. (2017). *'Killer robots': autonomous weapons pose moral dilemma*. [online] DW.COM. Available at: <http://www.dw.com/en/killer-robots-autonomous-weapons-pose-moral-dilemma/a-41342616> [Accessed 12 Mar. 2019].

Zhang, M. (2015). *Google Photos Tags Two African-Americans As Gorillas Through Facial Recognition Software*. [online] Available at: <https://www.forbes.com/sites/mzhang/2015/07/01/google-photos-tags-two-african-americans-as-gorillas-through-facial-recognition-software/#4bc6b57c713d> [Accessed 11 Mar. 2019].