DEPARTMENT OF LAW
Master of Laws Program
Master Thesis
HRO800, AT 2019
30 ECTS

Human Rights in the Contemporary Context of Artificial Intelligence influencing Human Life

ARTIFICIAL INTELLIGENCE, INTERNATIONAL HUMAN RIGHTS AND COGNITIVE LIBERTY

EMMIE NORDELL

Supervisor: Moa De Lucia Dahlbeck

Abstract

This thesis is an attempt to contribute to the understanding of the landscape of artificial intelligence, human cognition, and cognitive liberty, set in the context of contemporary life. I am exploring if there is a need to rethink the protection of human cognition, provided by the international human rights framework and if that framework is elastic enough to be able to provide sufficient protection of human cognition in the contemporary context of AI influencing human life.

I suggest that turning to a posthuman understanding of the world can provide new terminologies and new understandings that can be helpful in understanding new human living conditions in which AI constitutes integrate parts. I am, therefore, in light of a posthuman understanding of the world, analysing how relevant kinds of AI are affecting human cognition. A posthuman understanding of the world could lead to an understanding that posthuman rights are necessary. I do, however, suggest that the concept of human rights can have a desirable function in human society and am therefore not proposing posthuman rights. My suggestion is rather, that human rights, can be rethought in light of a posthuman understanding of how AI affects human cognition. Based on these explorations and analyses I suggest that the relevant kinds of AI are affecting human cognition in new ways that alter common human vulnerability and makes human cognition vulnerable. Since protection of human cognition, as a common human interest that is vulnerable and necessary for humans to be agents, can be argued to be a common human interest that should be protected by human rights, the introduction of AI entails a need for rethinking protection of human cognition in the contemporary context of AI influencing human life. I am, therefore, analysing the concept of cognitive liberty in the contemporary context of AI influencing human life to understand if the international human rights framework is elastic enough to adapt already recognized rights so that they can protect human cognition in this contemporary context, and if cognitive liberty, as a new human right, can be justified. I am arguing that already recognized rights cannot, in line with rules of interpretation, be interpreted so that they can accommodate such protection and, thus, that the international human rights framework is not elastic enough to adapt already recognized rights to the contemporary context of AI influencing human life. I am, further, suggesting that a new human right guaranteeing protection of human cognitive liberty ought to be considered by the international community.

List of abbreviations

AI- Artificial intelligence

EEG- Electroencephalography

ICCPR- The International Covenant on Civil and Political Rights

ICESCR- The International Covenant on Economic, Social and Cultural Rights

UDHR- Universal Declaration of Human Rights

UN- United Nations

UNESCO- United Nations Educational, Scientific and Cultural Organization

Table of contents

1. Introduction

"We shape our tools, and thereafter our tools shape us."[1]

1.1 Problem and purpose

This thesis is an attempt to contribute to the understanding of the landscape of artificial intelligence, human cognition, and cognitive liberty, set in the context of contemporary life. I am exploring if there is a need to rethink the protection of human cognition, provided by the international human rights framework and if that framework is elastic enough to be able to provide sufficient protection of human cognition in the contemporary context of artificial intelligence (AI) influencing human life. In this context, technological advancement is moving at a very high pace and human society is changing character for every day that passes. A new phenomenon called Artificial Intelligence has been developed.[2] Technological advancement has now reached a stage where AI has surpassed human intelligence, even if that is the case only when performing narrow tasks.[3] AI can learn through machine learning, a new technique through which systems are learning automatically when being presented with large amounts of data.[4]

AI entails substantial benefits to human society, but it also poses certain risks and may have negative impacts on, for example, democracy, the rule of law and the human being, itself.[5] AI is being used in countless different ways and is affecting human life to a large extent. Self-

[1] Culkin, J.M., "A schoolman's guide to Marshall McLuhan." Saturday Review, (1967) pp. 51-53, p. 71-72.
[2] The development of Artificial intelligence can be considered to have started with the first work that today is generally recognized as artificial intelligence, done by Warren McCulloch and Walter Pitts in 1943. (Russell, Stuart J, Norvig, Peter & Davis, Ernest, Artificial Intelligence: A Modern Approach, Global Edition (Pearson Education M.U.A., 2016) [Electronic] Available: Dawsonera, https://www-dawsonera-com.ezproxy.ub.gu.se/readonline/9781292153971 (Last accessed: 1/12-2019)) Based on this recognition, artificial intelligence has been around for 76 years. In relation to the history of human beings, AI can therefore be considered to be a new phenomenon.
[3] See. Chen, Jim X., "The Evolution of Computing: AlphaGo." Computing in Science & Engineering, vol. 18, no. 4, 2016, pp. 4–7 & Lu, Huimin, Yujie Li, Min Chen, Hyoungseop Kim, and Seiichi Serikawa, "Brain Intelligence: Go Beyond Artificial Intelligence." ArXiv.org, 2017, arXiv.org. & Hayles, N. Katherine, "Computing the Human." Theory, Culture & Society, vol. 22, no. 1, 2005, pp. 131-151, p. 132.
[4] Lexico powered by oxford, "Machine learning", https://www.lexico.com/en/definition/machine_learning (Last accessed: 18-10-2019).
[5] The EC High-Level Expert Group on Artificial Intelligence, AI HLEG, 2019, "Ethics Guidance for Trustworthy AI", https://ec.europa.eu/futurium/en/ai-alliance-consultation (Last accessed: 11/12-2019), p. 2.

driving cars are rapidly becoming reality,[6] and human bodies can communicate with chatbots.[7] Developing "artificial intelligence information communication technology",[8] has been one of the focuses of the technological development of AI. AI is, thus, now even shaping the flow of information, which is reaching human beings.

AI in personalization algorithms is filtering and thereby customizing the information flow, for every specific human being, creating what has been called filter bubbles.[9] AI in Augmented reality technology is overlying virtual content over what is being perceived in the non-virtual world.[10] Neurodevices such as consumer-based headsets, analysing brainwaves, are reading human cognition and might soon be replacing the keyboard, the touch screen, the mouse and voice command as preferred ways for human beings to communicate with technology.[11] AI can find patterns in brain data and decode brain activity to reveal aspects of human cognition. Even though there are not, yet, any algorithms that can reliably decode complex thoughts, it is possible to reveal mood and even single-digit numbers, shapes or simple words that are thought, heard or seen.[12] It has been argued that "[t]his possibility of *mining the mind* (or at least informationally rich structural aspects of the mind) can be potentially used not only to infer mental preferences, but also to prime, imprint or trigger those preferences".[13]

---

[6] Kallioinen, Noa, Pershina, Maria, Zeiser, Jannik, Nosrat Nezami, Farbod, Stephan, Achim, Pipa, Gordon & König, Peter. "Moral Judgements on the Actions of Self-driving Cars and Human Drivers in Dilemma Situations from Different Perspectives." OSF Preprints, (2019).
[7] Martinez, Rogelio. "The Power of Artificial Intelligence." Franchising World, vol. 50, no. 5, 2018, pp. 92–94.
[8] Lu, Huimin, Yujie Li, Min Chen, Hyoungseop Kim, and Seiichi Serikawa, "Brain Intelligence: Go Beyond Artificial Intelligence.".
[9] Gottron, Thomas & Felix Schwagereit. "The Impact of the Filter Bubble -- A Simulation Based Framework for Measuring Personalisation Macro Effects in Online Communities." ArXiv.org, 2016, pp. arXiv.org.
[10] Billinghurst, Mark. "Augmented Reality." The SAGE Encyclopedia of the Internet. Ed. Barney Warf. Thousand Oaks,: SAGE Publications, Inc., 2018. 35-40. SAGE Knowledge. Web. (Last accessed: 1/12 2019).
[11] Ienca, Marcello, "The Right to Cognitive Liberty." Scientific American, vol. 317, no. 2, 2017, pp. 10 & Ienca, Marcello & Roberto Andorno, "Towards New Human Rights in the Age of Neuroscience and Neurotechnology." Life Sciences, Society and Policy, vol. 13, no. 1, 2017, pp. 1–27, p. 4.
[12]Anumanchipalli, Gopala K, Chartier, Josh & Chang, Edward F., "Speech Synthesis from Neural Decoding of Spoken Sentences." Nature, vol. 568, no. 7753, 2019, pp. 493–498. & Whyte, Chelsea, "Mind-reading device uses AI to turn brainwaves into audible speech", 24/4-2019, https://www.newscientist.com/article/2200683-mind-reading-device-uses-ai-to-turn-brainwaves-into-audible-speech/ (Last accessed: 24/11-2019) & The Guardian, "Neuroscientists decode brain speech signals into written text", 30/7-2019, https://www.theguardian.com/science/2019/jul/30/neuroscientists-decode-brain-speech-signals-into-actual-sentences (Last accessed: 24/11-2019) & Greshko, Michael & Wei-Haas, Maya, "New device translates brain activity into speech. Here's how", National Geographic, Published: April 24, 2019 https://www.nationalgeographic.com/science/2019/04/new-computer-brain-interface-translates-activity-into-speech/ (Last accessed: 7/11-2019).
[13] Ienca, Marcello & Roberto Andorno, "Towards New Human Rights in the Age of Neuroscience and Neurotechnology.", p. 4.

Human beings are now living in a world where the human being is increasingly entangled with AI. AI might actually affect human cognition and human cognition is now, possibly being produced in convergence with AI. Human beings do of course affect the cognition of other human beings, all the time. Social life may even be described as an infinite web of mutual influence.[14] This condition poses the question; what are the legitimate ways of affecting someone's cognition?

Human beings are ascribed human rights and the international human rights framework sets out fundamental human rights to be universally protected. These rights are said to be inherent to every human being and are protecting different parts of human life.[15] A foundation for human rights can be described as common human vulnerability,[16] and human rights can be seen as protecting common human interests that are vulnerable and necessary for humans to be agents.[17]

In the contemporary context, when the new phenomenon AI is affecting numerous parts of human life, it is possibly also affecting the protection given by the international human rights framework. There are, already recognized, human rights connecting to human cognition. For example, human thought is protected by the international human rights framework.[18] However, there are those that argue that these recognized rights aren't adequately protecting human cognition in this contemporary context.[19] These ideas can be seen as supported by documented statements made by one of the drafters of the Universal Declaration of Human Rights (UDHR)

---

[14] Bublitz, Jan & Christoph Merkel. "Crimes Against Minds: On Mental Manipulations, Harms and a Human Right to Mental Self-Determination." Criminal Law and Philosophy, vol. 8, no. 1, 2014, pp. 51–77.
[15] United Nations, "Human rights", https://www.un.org/en/sections/issues-depth/human-rights/ (Last accessed: 17/11-2019)
[16] Turner, Bryan S., "Sociology of Human Rights", in Shelton, Dinah, (ed.) The Oxford Handbook of International Human Rights Law (Oxford: Oxford University Press 2013) [Electronic] Available: Oxford university press, https://opil-ouplaw-com.ezproxy.ub.gu.se/view/10.1093/law/9780199640133.001.0001/law-9780199640133 (Last accessed: 1/12-2019), p. 9 & This understanding will be discussed and elaborated on in the text. See. e.g. section 3.1.2.
[17] This understanding will be discussed and elaborated on in the text. See e.g. section 3.1.2.
[18] Universal Declaration of Human Rights, GA Res 217A (III), UNGAOR, 3rd Sess, Supp No 13, UN Doc A/810 (1948) 71, article 18 & International Covenant on Civil and Political Rights, GA Res 2200A (XXI), Treaty Series, vol. 999, p. 171, (1966), article 18.
[19] See Wrye, Sententia, "Neuroethical Considerations: Cognitive Liberty and Converging Technologies for Improving Human Cognition." Annals of the New York Academy of Sciences, vol. 1013, no. 1, 2004, pp. 221–228, & Bublitz, Jan & Christoph Merkel. "Crimes Against Minds: On Mental Manipulations, Harms and a Human Right to Mental Self-Determination." & The Center for Cognitive Liberty, http://www.cognitiveliberty.org (Last accessed: 17/11-2019) & Ienca, Marcello, "The Right to Cognitive Liberty." & Farahany, Nita, "Cognitive Liberty in the Era of Brain Hacking, The Aspen Institute", The Aspen Institute, Published: 11/9-2014, https://youtu.be/8CqgZ0V5pvY (Last accessed: 5/10 2019) & Ienca, Marcello & Roberto Andorno, "Towards New Human Rights in the Age of Neuroscience and Neurotechnology.".

when speaking of the right to freedom of thought, conscience and religion, that today is acknowledged in article 18 in the UDHR.[20]

> [I]t would be unnecessary to proclaim that freedom if it were never to be given an outward expression; if it were intended, so to speak, only for the use of the inner man. It was necessary however to stress the external manifestation of creeds by which expression was given to beliefs.[21]

At the time of the drafting of the UDHR, it might not have been seen as necessary to protect human cognition that was not to be given an outward expression, but is that still the case today? Is the human cognition sufficiently protected from impact, influence, modulation, and manipulation in a time where human cognition is possibly being produced in convergence with AI?

Voices have been raised to promote the idea that there is something in human vulnerability that is not protected, that the human rights protection needs to be expanded. A new concept that has been put forward in the judicial discourse is the concept, cognitive liberty. Cognitive liberty has been proposed as a conceptual update of existing human rights or as a new human right, which could serve as protection of human cognition in the contemporary context of AI influencing human life.[22]

The new phenomenon AI has great implications, and to understand these, it is necessary to understand the landscape of the contemporary context of AI influencing human life. This text is an attempt to contribute to this understanding. My interest is focused on whether there is a need to rethink the protection of human cognition that the present international human rights framework offers and if this framework is elastic enough to be able to provide sufficient protection in the contemporary context of AI influencing human life. The purpose of this essay is, therefore, to explore if the introduction of AI into human society presents a need to rethink the protection of human cognition that the international human rights framework provides, as of today, and if the framework is elastic enough to be able to provide sufficient protection of

---

[20] GA, Third Committee, Third Session, 127th Meeting (1948), A/C.3/SR.127, at 395. & GA, Third Committee, Third Session, Draft International Declaration of Human Rights: Recapitulation of Amendments to Article 16 of the Draft Declaration (E/800), (1948), A/C.3/289/REV.1.
[21] GA, Third Committee, Third Session, meeting 127, at 395.
[22] Wrye, Sententia, "Neuroethical Considerations: Cognitive Liberty and Converging Technologies for Improving Human Cognition." & Bublitz, Jan & Christoph Merkel. "Crimes Against Minds: On Mental Manipulations, Harms and a Human Right to Mental Self-Determination." & The Center for Cognitive Liberty & Ienca, Marcello, "The Right to Cognitive Liberty." & Farahany, Nita, "Cognitive Liberty in the Era of Brain Hacking, The Aspen Institute" & Ienca, Marcello & Roberto Andorno, "Towards New Human Rights in the Age of Neuroscience and Neurotechnology.".

human cognition in the contemporary context of AI influencing human life, given the understanding that human rights should be protecting common human interests that are vulnerable and that are necessary for humans to be agents.[23]

1.2 Theory and method

When a new phenomenon is introduced to any given context, there is a need to understand the new phenomenon. In order to understand new phenomena, it is sometimes necessary to employ new conceptualizations, or, as it is expressed in the Posthuman Glossary by Rosi Braidotti and Maria Hlavajova, "[w]e need new terms. And new terminologies require conceptual creativity, which means to trust in the powers of the imagination, as well as rely on academic credentials and conventions."[24] The new phenomenon, AI, has been introduced to the contemporary context and needs to be understood. Hence, new understandings are necessary. In an attempt to make questions and ideas that can contribute to this particular understanding of contemporary development visible I will, therefore, engage in a theoretical examination of the problem of protection of human cognition in a time where AI is affecting human life. This thesis is, thus, a theoretical examination of the conceptual stakes of a given new practical legal context.

I suggest that posthuman theory can provide new terminologies and new understandings that we need in order to understand the contemporary context of AI and human cognition. I am thus, examining the problem from a posthuman theoretical perspective. In Braidottis words, the defining features of posthuman theory are,

> that it rests on a neo-materialist philosophy of immanence, which assumes that all matter is one (monism); that matter is intelligent and self-organizing (autopoiesis); that the subject is not unitary but nomadic; and that subjectivity includes relations to a multitude of non-human 'others'. [25]

A posthuman understanding of the problem could, possibly, result in some kind of destabilization of human rights. I suggest that a posthuman understanding is helpful in understanding the landscape of the contemporary context of AI influencing human life. I do,

---

[23] This understanding will be discussed and elaborated on in the text. See e.g. section 3.1.2.
[24] Braidotti, Rosi & Hlavajova, Maria, "Introduction", in Braidotti, Rosi, & Hlavajova, Maria, (ed.) Posthuman Glossary (London: Bloomsbury Publishing Plc, 2018) [Electronic] Available: ProQuest Ebook Central, https://ebookcentral-proquest-com.ezproxy.ub.gu.se/lib/gu/reader.action?docID=5226228 (Last accessed: 1/12-2019), p. 10.
[25] Braidotti, Rosi, "Posthuman critical theory", in Braidotti, Rosi, & Hlavajova, Maria, (ed.) Posthuman Glossary (London: Bloomsbury Publishing Plc, 2018) [Electronic] Available: ProQuest Ebook Central, https://ebookcentral-proquest-com.ezproxy.ub.gu.se/lib/gu/reader.action?docID=5226228 (Last accessed: 1/12-2019), p. 340.

however, also suggest that the concept of human rights can have a desirable function in human society. I am therefore not proposing posthuman rights. My suggestion is rather, that human rights, can be rethought in the light of a posthuman understanding of how AI affect human cognition. This will be discussed, further, in section 4.1.

To understand the development of relevant aspects of AI that are affecting human cognition in ways that could show a need for protection of human cognition by the international human rights framework, I have been engaging in studies of texts, statements and terminology, present in the discourse connected to the development of AI. I have also studied material from contemporary discourses on cognition in order to understand how human cognition is described. To understand if/how the introduction of relevant aspects of AI is altering human vulnerability, I have further analysed how AI is affecting human cognition in light of a posthuman understanding of the world.

To understand what the international human rights framework is protecting in so far as human cognition is concerned, I have moreover studied legal sources in the international human rights framework.

To understand if the international human rights framework can provide protection of human cognition in the contemporary context of AI influencing human life it has been necessary to analyse if it is elastic enough to provide protection by rethinking of already recognized rights as well as by recognition of a new human right. To understand how elastic the international human rights framework is, in terms of rethinking already recognized rights to accommodate protection of human cognition, I have studied legal sources in the international human rights framework. To be able to analyse the necessity and possibility of recognition of cognitive liberty as a new right, I have studied legal sources and texts, statements and terminology in the legal discourse. This has been done to understand how the suggested new right has been described, and if it can be justified as a human right.

To contribute to an understanding of the landscape of AI, human cognition and international human rights, findings from the review of texts, statements, terminology, and legal sources are described in this thesis. A posthuman understanding of the world is then used to analyse how AI affects human cognition in the contemporary context of AI influencing human life. To contribute to an understanding of whether the international human rights framework is elastic enough to accommodate protection of human cognition, in the contemporary context of AI

influencing human life, I am then analysing foundations and justifications for human rights and rules for the interpretation of international human rights together with conclusions from the other parts of the thesis.

1.3 Material

The material used in this thesis are texts, statements and terminology, collected from the areas of the discourse connected to the development of AI, the contemporary discourses on cognition, the international legal discourse, and legal sources, such as treaties, case law, preparatory works and doctrine. The selection of material used in this thesis is mainly done by searches on the Gothenburg university library webpage and Google scholar. This means that AI is present in the selection of material for this thesis of AI and human cognition. Personalization algorithms, that is studied in this thesis, is therefore also affecting this thesis.[26] This will be further discussed in section 5.

To be able to achieve awareness of the personalization algorithms in my study, I have taken a few measures. I have, to some extent, used Google's option to turn off personalized searches, to minimize personalization in my search results. However, as will be discussed in section 2.2.1, it is unclear to what degree and in what ways such searches are being filtered anyway. Personalization can, therefore, be present regardless of this option being used. Further, since this makes the sorting of information much harder it has not been possible to use this option for every search. However, to analyse my own search results, I have used the option as a way of controlling my results, after I have used personalized searches. Furthermore, I have used a Virtual Private Network (VPN) and stayed logged out from my Google account when searching for information. These measures make it harder to track my IP address and location and to connect my email with my search, which minimizes personalization in the search results.[27] These different measures make some aspects of, how my search results are personalized visible.

---

[26] C.f. Curkovic, Marko., "Need for Controlling of the Filter Bubble Effect." Science and Engineering Ethics, vol. 25, no. 1, 2019, p. 323.

[27] See Giordano, Sarah, "Popping the Filter bubble", 2014, https://derekbruff.org/blogs/fywscrypto/practical-crypto/popping-the-filter-bubble/ (Last accessed: 18/11-2019) & See. A Dictionary of Computer Science 7 ed., "VPN", Butterfield, Andrew, Ekembe Ngondi, Gerard & Kerr, Anne (ed.), (Oxford University press, 2016), https://www-oxfordreference-com.ezproxy.ub.gu.se/view/10.1093/acref/9780199688975.001.0001/acref-9780199688975-e-6258?rskey=LboWvL&result=1 (Last accessed: 2/12 2019) & See Google, "Updating Our Privacy Policies and Terms of Service," Google Official Blog, 24/1-2012, https://googleblog.blogspot.com/2012/01/updating-our-privacy-policies-and-terms.html (Last accessed: 11/12-2019).

Since the development of AI is moving at a high pace, the selection of relevant material is, to a large degree, based on how recently the material was published.

The selection of legal sources is based on the sources of international law, recognized in article 38(1) the Statute of the International Court of Justice. International human rights law is a specialist regime within general public international law, the sources of international law are therefore also the recognized sources of human rights law.[28]

> **Article 38**
>
> 1. The Court, whose function is to decide in accordance with international law such disputes as are submitted to it, shall apply:
> - international conventions, whether general or particular, establishing rules expressly recognized by the contesting states;
> - international custom, as evidence of a general practice accepted as law;
> - the general principles of law recognized by civilized nations;
> - subject to the provisions of Article 59, judicial decisions and the teachings of the most highly qualified publicists of the various nations, as subsidiary means for the determination of rules of law.[29]

The recognized sources of international human rights law can be summarised as treaties, international customary law, general principles of law, case law and scholarly writings. I am using these different sources to understand how the recognized rights, connecting to human cognition, have been described and what they are considered to protect, but also to understand how elastic the international human rights framework is, in this regard.

To understand how the international human rights framework is protecting human cognition, I am in this thesis, studying the UDHR, the International Covenant on Civil and Political Rights (ICCPR) and the International Covenant on Economic, Social and Cultural Rights (ICESCR). The ICCPR and the ICESCR are treaties and are therefore recognized sources of international human rights law. The UDHR is not a treaty. It is, however, generally agreed to be the foundation of international human rights law and to reflect customary law.[30] This document, and the interpretations of it are therefore relevant for this thesis. I am studying these documents since they can be seen as the foundation of the international human rights framework. Other

[28] Besson, Samantha, "Justifications", in Moeckli, Daniel, Sangeeta Shah, Sandesh Sivakumaran, & David J. Harris, (ed.) International Human Rights Law. Third ed., (Oxford: Oxford University Press 2018), p. 63-65.
[29] Statute of the International Court of Justice, 18 April 1946.
[30] United Nations, "The Foundation of International Human Rights Law", https://www.un.org/en/sections/universal-declaration/foundation-international-human-rights-law/index.html (Last accessed: 7/11-2019).

topic-specific and/or regional documents have not been studied since the focus of this thesis is how the human cognition is protected on a general and universal level. I have found relevant human rights in the UDHR and the ICCPR.

Further, customary law, general principles, case law and, scholarly writings are used to find rules of interpretation for the declaration and the treaty. To look at the intention of the parties is, in accordance with the Vienna Convention on the Law of Treaties, considered a rule of interpretation.[31] I am, therefore, using documents from the drafting of the declaration and the treaty to determine the general interpretation of the relevant recognized provisions. To clarify the meaning of the provisions, I am also using general comments adopted by the Human Rights Committee, the body of independent experts that monitors implementation of the International Covenant on Civil and Political Rights by its State parties. The general comments are adopted in accordance with ICCPR.[32]

1.4 Delimitations

There are many different kinds of AI that may affect human cognition. Since the purpose of this thesis is to explore if the introduction of AI into human society, presents a need to rethink the protection of human cognition, provided by the international human rights framework, I am studying some relevant examples of relevant kinds of AI. The kinds of AI that are studied are personalization algorithms; augmented reality technologies; and, some forms of neurotechnology.

Further, there are many recognized human rights that can be seen as protecting parts of human cognition in different ways. Since this thesis is a study of the need to rethink the protection of human cognition, I am here discussing the relevant recognized rights in closest connection to human cognition, that could possibly offer protection of the human cognition, in the contemporary context of AI influencing human life. This will be further discussed in section 3.2.

[31] Vienna Convention on the Law of Treaties, United Nations, Treaty Series, vol. 1155, p. 331, (1969), Article 31(4) & Fitzmaurice, Malgosia, "Interpretation of Human Rights Treaties", in Shelton, Dinah, (ed.) The Oxford Handbook of International Human Rights Law (Oxford: Oxford University Press 2013) [Electronic] Available: Oxford university press, https://opil-ouplaw-com.ezproxy.ub.gu.se/view/10.1093/law/9780199640133.001.0001/law-9780199640133 (Last accessed: 1/12-2019), p. 7.
[32] International Covenant on Civil and Political Rights, Article 40, paragraph 4.

1.5 Outline

This thesis is divided into three main parts. First, AI and human cognition are discussed in section 2, second, international human rights are discussed in section 3 and third, Cognitive liberty is discussed in section 4. In the first part, I am describing relevant aspects of AI and human cognition. To understand if/how the introduction of AI can be seen as altering human vulnerability I am, further, in light of a posthuman understanding of the world, analysing how human cognition is affected in the encounter with AI. In the second part, I am describing relevant parts of the international human rights framework. To show that the history of human rights is an ongoing one, I am briefly describing the history of human rights. To present an understanding of the concept of human rights, I am describing foundations and justifications for human rights and to set a starting point for the following interpretation of substantive human rights I am describing rules of interpretation. In the third part, I am analysing the concept of cognitive liberty in the contemporary context of AI influencing human life. I am analysing rethinking of human rights in the contemporary context of AI influencing human life in relation to posthuman rights, the relevance of international human rights in the contemporary context of AI influencing human life and the need to rethink the protection of human cognition, given by the international human rights framework in the contemporary context of AI influencing human life. To understand if the international human rights framework is elastic enough to accommodate protection of human cognition in the contemporary context of AI influencing human life, I am then analysing already recognized rights in light of the contemporary context of AI influencing human life and cognitive liberty. That analysis is followed by an analysis of cognitive liberty as a new human right. The thesis is concluded with a further and summarizing analysis and conclusions, finishing discussion, and, closing reflection.

2. Human Cognition and Artificial Intelligence

2.1 Human cognition

Human cognition, which is a central concept to this thesis, can be defined in numerous different ways. It can, for example, be defined generally, as "an umbrella term for all higher mental processes".[33] It can also be defined in more detail, for example, as "the collection of mental processes and activities used in perceiving, remembering, thinking, and understanding, as well as the act of using those processes".[34] These processes are thought to have evolved as a way of controlling action.[35] The survival value in being able to perceive things is only thought to be present if the human being can respond to what is perceived.[36] There is, thus, a need to process information[37] which is perceived. Cognition has in line with this also been defined as "[t]he mental activities involved in acquiring and processing information".[38] From these ideas about cognition, the following definition can be derived: cognition is mental processes and activities, processing information and constituting perceiving, remembering, thinking, and understanding, as well as the act of using those processes. The medium for human information processing is thought to be the nervous system which consists of the brain, the peripheral nervous system, and the spinal cord. Information and motor commands are transmitted to and from the brain and the brain performs millions of computations upon the information it receives.[39] The brain and activity in the brain is thus crucial to the mental processes, processing information.[40]

---

[33] Ashcraft, Mark H, Cognition. 3rd ed. (Upper Saddle River, N.J: Prentice Hall, 2002), p. 10.
[34] Ashcraft, Mark H, Cognition, p. 11.
[35] See Glass, Arnold Lewis, Cognition (Cambridge, UK; New York: Cambridge University Press, 2016), p. 2.
[36] Glass, Arnold Lewis, Cognition, p. 2.
[37] Information can be defined in numerous different ways. I am using a definition that can be derived from Gilbert Simondons theory of individuation. Information is, defined in this way, modalities of change rather than mere attributes of entities. In-formation processes are processes in which matter gain form. (See Rodriguez, Pablo & Blanco, Javier, "Organization and Information in Simondon's Theory of Individuation." Culture and Organization, vol. 23, no. 1, 2017, pp. 34–43.) Information gives the universe its structure and complexity. (See Cesar Hidalgo: "Why Information Grows" | Talks at Google, 6/8-2015, https://www.youtube.com/watch?v=r38kK26SieE (Last accessed: 26/11-2019), at 0.27 min) Physical origins of order are therefore relevant for the definition of information. The second law of thermodynamics says that the universe has a tendency to average itself out or to strive towards equilibrium. (See Cesar Hidalgo: "Why Information Grows" | Talks at Google, at 2.00 min.) Systems that are out of equilibrium minimize the rate of entropy production. They are producing entropy, but they are producing as little entropy as possible. These systems are self-organizing to produce as little entropy as possible. It is in these systems that order emerges. (See Cesar Hidalgo: "Why Information Grows", at 4.50 min.).
[38] A Dictionary of psychology 4 ed., "Cognition", Colman, Andrew M. (ed.), (Oxford University press 2015), https://www-oxfordreference-com.ezproxy.ub.gu.se/view/10.1093/acref/9780199657681.001.0001/acref-9780199657681-e-1594?rskey=iftF5J&result=6 (Last accessed: 2/12 2019).
[39] Glass, Arnold, Cognition. (Mason, OH: Thomson Custom Pub, 2007), p. 8.
[40] See Glass, Arnold, Cognition, p.8, See Ashcraft, Mark H, Cognition, p.53-67 & See Glass, Arnold Lewis, Cognition, p. 53 & 132.

2.2 Artificial intelligence

Considering the fact that there are numerous different definitions of AI,[41] I am in this study, using a broad definition of the term, relying on the idea that AI is some kind of system that has some kind of ability that has been associated with intelligent beings.[42]

Further, AI is commonly separated into two categories, general and narrow AI. Narrow AI is defined as systems that demonstrate intelligence in a specialized area. Whereas general AI has been defined as systems that "can solve a variety of complex problems in a variety of different domains, and that controls itself autonomously, with its own thoughts, worries, feelings, strengths, weaknesses and predispositions".[43] The vast part of the AI field today, is focusing on narrow AI. [44] The focus of this thesis is, therefore, narrow AI.

AI encompasses a large variety of subfields.[45] Different kinds of AI raise different challenges,[46] and human cognition is affected in different ways in encounters with different kinds of AI. To be able to start understanding the landscape of AI and human cognition in the contemporary context of AI influencing human life, I will, in the following sections 2.2.1, 2.2.2 and 2.2.3, briefly discuss some relevant kinds of AI, namely AI in personalization algorithms, AI in augmented reality technology and AI in neurotechnology.

2.2.1 Artificial intelligence in personalization algorithms

The Internet is getting richer in information. It would, today, be impossible for a human being to sort the amount of information that is present on the internet.[47] In an attempt to make this amount of information manageable for human beings, internet actors, such as the company

[41] Dobrev, Dimiter, "A Definition of Artificial Intelligence." ArXiv.org, 2012, pp. arXiv.org. p. 2.
[42] See Encyclopaedia Britannica, "Artificial intelligence", Copeland, B.J., Britannica Online Academic Edition, 2019, https://www.britannica.com/technology/artificial-intelligence (Last accessed: 18-10-2019)& The EC High-Level Expert Group on Artificial Intelligence, AI HLEG, 2019, "Ethics Guidance for Trustworthy AI", p. 36 & C.f. Russell, Stuart J, Norvig, Peter & Davis, Ernest, Artificial Intelligence: A Modern Approach, Global Edition, p. 4-5.
[43] Pennachin, Cassio & Goertzel, Ben, "Contemporary Approaches to Artificial General Intelligence", In Goertzel Ben., Pennachin Cassio., (ed.) Artificial General Intelligence. Cognitive Technologies, (Springer-Verlag Berlin Heidelberg 2007) [Electronic] Available: Springer Link, https://link.springer.com/chapter/10.1007/978-3-540-68677-4_1 (Last accessed: 1/12-2019).
[44] Pennachin, Cassio & Goertzel, Ben, "Contemporary Approaches to Artificial General Intelligence"..
[45] Russell, Stuart J, Norvig, Peter & Davis, Ernest, Artificial Intelligence: A Modern Approach, Global Edition, p. 1.
[46] The EC High-Level Expert Group on Artificial Intelligence, AI HLEG, 2019, "Ethics Guidance for Trustworthy AI", p. 5.
[47] Krafft, Tobias, Michael Gamer, and Katharina Zweig, "What Did You See? Personalization, Regionalization and the Question of the Filter Bubble in Google's Search Engine." ArXiv.org, 2018, arXiv.org.

Google, uses algorithms to filter the information in a personalized way.[48] Personalization is enabled by algorithm-based systems, in which algorithms decide what content the human user might be interested in. This decision is determining what information is presented to the human being.[49]

In the book, *The filter bubble: what is the internet hiding from you,* Eli Pariser, who is an internet activist and entrepreneur, coined the term filter bubble.[50] It was coined to describe a situation in which online users, due to personalization algorithms, live in personalized information universes biased towards their own interests.[51] The concept has been defined as a "phenomenon whereby the ideological perspectives of internet users are reinforced as a result of the selective algorithmic tailoring of search engine results to individual users (as reflected in recorded data such as search history, click data, and location)".[52] The term filter bubble has also been described as representing the unique, personal universe of online information, in which a human being lives. The boundaries of that universe depend on personalization algorithms in search engines and social networks. To decide which information is likely to be considered relevant, the algorithm considers the human user's interaction with previously encountered information. The algorithm will show the information, that is labelled as likely to be considered relevant at higher ranks, and in some cases, it will even block out other information.[53] Pariser suggests that "[f]or an individual user this might lead to a skewed and biased perception of the world".[54]

---

[48] Google, "Updating Our Privacy Policies and Terms of Service," & Google, "Så fungerar sökalgoritmer", https://www.google.com/search/howsearchworks/algorithms/ (Last accessed: 31/10-2019).
[49] Krafft, Tobias, Michael Gamer, and Katharina Zweig, "What Did You See? Personalization, Regionalization and the Question of the Filter Bubble in Google's Search Engine.".
[50] Pariser, Eli, The Filter Bubble: What the Internet Is Hiding from You (London: New York: Viking; Penguin Press, 2011). & Gottron, Thomas & Felix Schwagereit. "The Impact of the Filter Bubble -- A Simulation Based Framework for Measuring Personalisation Macro Effects in Online Communities.". The book "The Filter Bubble: What the Internet Is Hiding from You" does not derive from a scientific context. I do however think it is necessary to use the concept coined in the book since it has become a big part of the discourse, regarding this issue. I do further, want to assert that the concept, filter bubble, has gained scientific relevance which can be seen in the fact that it is being used in different scientific discussions. (See e.g. other references in this section).
[51] Gottron, Thomas & Felix Schwagereit. "The Impact of the Filter Bubble -- A Simulation Based Framework for Measuring Personalisation Macro Effects in Online Communities.".
[52] A Dictionary of Social Media, "Filter Bubble.", Chandler, Daniel, & Munday, Rod (ed.), (Oxford University press 2016), https://www-oxfordreference-com.ezproxy.ub.gu.se/view/10.1093/acref/9780191803093.001.0001/acref-9780191803093-e-482?rskey=m5WEu0&result=1 (Last accessed: 2/12 2019).
[53] Gottron, Thomas & Felix Schwagereit. "The Impact of the Filter Bubble -- A Simulation Based Framework for Measuring Personalisation Macro Effects in Online Communities."
[54] Gottron, Thomas & Felix Schwagereit. "The Impact of the Filter Bubble -- A Simulation Based Framework for Measuring Personalisation Macro Effects in Online Communities.".

In the book *Are Filter Bubbles Real?* Axel Bruns, professor in the Creative Industries Faculty at Queensland University of Technology in Brisbane, presents a more irresolute view on the concept of filter bubbles. He points out that empirical work on filter bubbles suffer from limitations such as the lack addressing the actual experience of the users. He further points out that users who are part of an apparent filter bubble on one platform could be consuming a broader spectrum of information on other platforms.[55] Further, Kieron O'Hara, associate professor in Electronics and Computer Science and David Stevens, lecturer in the School of Politics and International Relations at the University of Nottingham, suggests that phenomena like filter bubbles can be beneficial under some circumstances, such as when enabling users to consume information that interests them and to follow news and other information at a level of complexity and detail that suits their level of information literacy.[56]

Personalization is used not only by Google but also for example by Facebook, Youtube, Yahoo News and the New York Times-funded startup News.me.[57] Pariser argues that "[t]ogether, these engines create a unique universe of information for each of us […] which fundamentally alters the way we encounter ideas and information".[58] Furthermore, Google has stated on its official blog that the "new Privacy Policy makes clear that, if you're signed in, [Google] may combine information you've provided from one service with information from other services. In short, [Google will] treat you as a single user across all [Googles] products, which will mean a simpler, more intuitive Google experience".[59] Hence, personalization is a part of many aspects of the internet.

Moreover, the reach of personalization algorithms has expanded beyond personal computers. An example of this is a billboard in Japan, that is using personalization algorithms and facial recognition on human beings passing by.[60] Of relevance is also, Ambient Intelligence, which is a term for describing a "world in which 'intelligence' is embedded in virtually everything around us".[61] It has been described as a world where everything, "the clothes you wear, the

[55] Bruns, Axel, Are Filter Bubbles Real? Digital Futures Series, (Cambridge: Polity Press, 2019), p. 11 & 15-35.
[56] O'Hara, Kieron & Stevens, David "Echo Chambers and Online Radicalism: Assessing the Internet's Complicity in Violent Extremism." Policy & Internet, vol. 7, no. 4, 2015, pp. 401–422., p. 417 & Bruns, Axel, Are Filter Bubbles Real?, p. 12.
[57] Pariser, Eli, The Filter Bubble: What the Internet Is Hiding from You, p. 8.
[58] Pariser, Eli, The Filter Bubble: What the Internet Is Hiding from You, p. 9.
[59] Google, "Updating Our Privacy Policies and Terms of Service,".
[60] The Guardian, "Advertising billboards use facial recognition to target shoppers" on& Gray, Richard, "Minority Report-style advertising billboards to target consumers" https://www.telegraph.co.uk/technology/news/7920057/Minority-Report-style-advertising-billboards-to-target-consumers.html (Last accessed: 23/10-2019).
[61] Wright, David, "The Dark Side of Ambient Intelligence." Info, vol. 7, no. 6, 2005, pp. 33–51.

paint on your walls, the carpets on your floor, the paper money in your pocket have a computer communications capability."[62] It has been called a world of smart dust and been described as the internet of things. David Wright, specialized in policy and regulatory issues relating to ambient intelligence, has predicted that this could be part of our near future.[63]

A team at Massachusetts Institute of Technology, led by Professor of Marketing John Hauser, has further developed techniques for what the team call website morphing. Morphing is when not only the content on a website but also the look and feel are personalized. In this case based on what the team has called cognitive style, which is inferred from clickstream data.[64] This technique is thus, further, personalizing not only the content but also how the content is presented.

2.2.2 Artificial intelligence in augmented reality technology

Augmented reality has been defined as "technology that seamlessly overlays virtual content over the real world so that both can be experienced at the same time".[65] The earliest examples of applications of augmented reality were probably the so-called heads-up-displays that were used in military airplanes and tanks. This was a technology that showed instrument panel-type information projected onto the same display as the one through which the pilot saw the surroundings.[66]

Today, augmented reality technology can be found in products available for everyday use, such as noise cancelling headphones with functions that amplify sounds from the surroundings or Google Glass. Google Glass is a wearable device that, according to Google itself, provides "glanceable, voice-activated assistance that is designed to be worn all day".[67] Mercedes-Benz is also using augmented reality, for navigation, in Mercedes driver assist. The technique is showing navigation and traffic information in live pictures. A camera is filming the

[62] Wright, David, "The Dark Side of Ambient Intelligence.".
[63] Wright, David, "The Dark Side of Ambient Intelligence.".
[64] Braun, Michael, Hauser, John & Urban, Glen & Liberali, G. "Website Morphing." Marketing Science: the Marketing Journal of TIMS/ORSA, vol. 28, no. 2, 2009, pp. 202–223.
[65] Billinghurst, Mark. "Augmented Reality.".
[66] Encyclopædia Britannica, "Augmented Reality", Hosch, William L., Britannica Online Academic Edition, 2019, https://www.britannica.com/technology/augmented-reality (Last accessed: 18-10-2019).
[67] Google, "Glass", https://www.google.com/glass/tech-specs/ (Last accessed: 23/10-2019).

surroundings and highlighting certain details, such as house numbers, street names or traffic lights.[68]

2.2.3 Artificial intelligence in non-invasive neurotechnology

Easy and cheap neuro-devices are, today, available to human beings. Examples of neuro-devices that are available, on the market, are the headsets Emotiv, Neurosky and Muse.[69] Other examples of neurotechnology are Freer logic performance monitor, which is a device that, to understand the performance of a human being in the situation at hand, is monitoring brain activity to determine the state of mind of the human being. This device is being used by for example National Aeronautics and Space Administration (NASA).[70] A head-based wearable technology called Mindrider, originally developed at Massachusetts Institute of Technology, analyses how movements, and location engage a human brain.[71] These kinds of neurotechnology are recording Electroencephalography (EEG) to observe electrical activity in the human brain.[72] Another neurotechnology that is available today, is so-called brain fingerprinting. Brain fingerprinting uses the EEG recordings of a human brain and puts it in relation to the human being's memory of events, as an improved polygraph. It computes a determination of if information is present or not and statistical confidence of the determination. Laboratory and field testing made by the Federal Bureau of Investigation, Central Intelligence Agency and the United States Navy has supposedly resulted in zero percent errors.[73]

2.3 Human cognition in convergence with artificial intelligence

AI is now a part of human life and human beings are part of the technological environment.[74] Human beings and human cognition are becoming connected to AI. Braidotti describes these

[68] Mercedes benz, "MBUX Augmented Reality för navigation", https://www.mercedes-benz.se/passengercars/mercedes-benz-cars/models/eqc/comfort.pi.html/mercedes-benz-cars/models/eqc/comfort/comfort-gallery/augmented-video (Last accessed: 23/10-2019).
[69] Emotiv, https://www.emotiv.com (Last accessed: 11/12-2019), Neurosky, http://neurosky.com, (Last accessed: 11/12-2019) & Muse, https://choosemuse.com, (Last accessed: 11/12-2019).
[70] Farahany, Nita, "Cognitive Liberty in the Era of Brain Hacking, The Aspen Institute" & Freerlogic, http://www.freerlogic.com/products/hardware (Last accessed: 23/10-2019).
[71] Mindriderdata, http://mindriderdata.com (Last accessed: 23/10-2019).
[72] Ienca, Marcello & Roberto Andorno, "Towards New Human Rights in the Age of Neuroscience and Neurotechnology.", p. 4.
[73] Farwell, Lawrence, "Brain Fingerprinting: a Comprehensive Tutorial Review of Detection of Concealed Information with Event-Related Brain Potentials." Cognitive Neurodynamics, vol. 6, no. 2, 2012, pp. 115–154, p. 115.
[74] C.f. Paasonen, Susanna, "Networked affect", in Braidotti, Rosi, & Hlavajova, Maria, (ed.) Posthuman Glossary (London: Bloomsbury Publishing Plc, 2018) [Electronic] Available: ProQuest Ebook Central,

altered boundaries, between what is human and what is technology, as the posthuman condition.[75] A theoretical point of departure in posthuman theory is the need to "overcome binaries and to state that matter, the world and humans themselves are not dualistic entities structured according to dialectic principles of internal or external opposition, but rather materially embedded subjects-in-process circulating within webs of relation with forces, entities, and encounters".[76] In the book "How We Became Posthuman", Katherine Hayles discuss how she has noticed two tendencies of how the human being has been subjected to alienation by AI. On the one hand, she argues that there is one type of narrative that indicates the fear of loss of humanity and loss of control, but also fear of dissolution of the human self. She argues that these narratives comprise a perception of technology as separate from the human body. On the other hand, she argues, that there are stories that suggest a contrasting perception of the human in relation to the contemporary context of AI influencing human life. She argues that by disentangling assumptions about the human as an independent entity, the possibility for the human to live in close connection with other life forms, renders available. [77] In light of this idea, I suggest that understanding human cognition as entangled with AI, such a possibility can be made available. However, I further suggest that understanding human cognition in such a way can also alter the understanding of human vulnerability.

2.3.1 Understanding human cognition as entangled with artificial intelligence

AI in personalization algorithms and in augmented reality technology is filtering and shaping the flow of information that is reaching human cognition. These technologies use data about the human being to shape the flow of information. The human being and the human cognition is thus being increasingly entangled with AI. There are, as Braidotti and Hlavajova express it "new forms of interconnection between humans and non-human factors and agents".[78]

---

https://ebookcentral-proquest-com.ezproxy.ub.gu.se/lib/gu/reader.action?docID=5226228 (Last accessed: 1/12-2019), p. 284.

[75] Braidotti, Rosi, The Posthuman (Oxford: Polity Press, 2013) [Electronic] Available: ProQuest Ebook Central, https://ebookcentral-proquest-com.ezproxy.ub.gu.se/lib/gu/detail.action?docID=1315633 (Last accessed: 1/12-2019), passim.

[76] Braidotti, Rosi & Hlavajova, Maria, "Introduction", p. 8.

[77] Hayles, N. Katherine, How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics (Chicago, Ill.: Univ. of Chicago Press, 1999), p. 235-240 & See also Pisters, Patricia, "Body without organs", in Braidotti, Rosi, & Hlavajova, Maria, (ed.) Posthuman Glossary (London: Bloomsbury Publishing Plc, 2018) [Electronic] Available: ProQuest Ebook Central, https://ebookcentral-proquest-com.ezproxy.ub.gu.se/lib/gu/reader.action?docID=5226228 (Last accessed: 1/12-2019), p. 75-76.

[78] Braidotti, Rosi & Hlavajova, Maria, "Introduction", p. 2.

In the book chapter "Where are the Missing Masses? The Sociology of a Few Mundane Artefacts" Bruno Latour argues that even very commonplace technologies can shape the decisions humans make and the way humans are moving through the world.[79] He reasons that "[t]hey persuade, facilitate and enable particular human cognitive processes, actions or attitudes, while constraining, discouraging and inhibiting others".[80] AI, in personalization algorithms and augmented reality technologies, prioritize and present information in a particular order and the selection of information a human being get to see is, therefore, affected. AI can thus be seen as shaping a human being's perception, experience, existence, and action.[81] Since human cognition, as defined in this thesis, is mental processes processing information it can be seen as persuading, facilitating and enabling particular human cognitive processes actions and attitudes. Since the selection of information, a human being doesn't get to see is also affected, AI is also constraining, discouraging and inhibiting other cognitive processes. As Pariser argues, personalization algorithms and filter bubbles make it less likely that chance encounters will happen. Pariser states that a world constructed from the familiar, by definition, "is a world in which there's nothing to learn. If personalization is too acute, it could prevent us from coming into contact with the mind-blowing, preconception-shattering experiences and ideas that change how we think about the world and our-selves".[82] In this way, Pariser claims that "the rise of the filter bubble doesn't just affect how we process news. It can also affect how we think".[83] Pariser, further, argues that a narrow filter bubble impedes creativity.[84] Studies show that human beings that have been assigned creative attributes tend to see things in many different ways and place them in wide categories.[85] When filter bubbles show results that are narrow by selection, this might constrain possible ways in which things can be seen and, therefore, constrain creativity. Since creativity can be argued to consist of mental processes, human cognition can in this way be seen as affected in the encounter with AI in personalization algorithms.

---

[79] Latour, Bruno, "Where are the Missing Masses? The Sociology of a Few Mundane Artefacts," in W. Bijker & J. Law (eds.), Shaping Technology/Building Society: Studies in Socio-Technical Change, (Cambridge, Massachusetts: The MIT press, 1992) pp. 225–258. [Electronic] Available: Nottingham Trent University, http://www.bruno-latour.fr/sites/default/files/50-MISSING-MASSES-GB.pdf, p.151, passim.
[80] Latour, Bruno, "Where are the Missing Masses? The Sociology of a Few Mundane Artefacts,", p.151, passim.
[81] See Noorman, Merel, "Computing and Moral Responsibility", The Stanford Encyclopedia of Philosophy (Spring 2018 Edition), Zalta, Edward N., (ed.), 2012 (edited 2018), https://plato.stanford.edu/archives/spr2018/entries/computing-responsibility/ (Last accessed: 2/12-2019).
[82] Pariser, Eli, The Filter Bubble: What the Internet Is Hiding from You, p. 15.
[83] Pariser, Eli, The Filter Bubble: What the Internet Is Hiding from You, p. 76.
[84] Pariser, Eli, The Filter Bubble: What the Internet Is Hiding from You, p. 99.
[85] Cropley, Arthur J., Creativity in Education and Learning: A Guide for Teachers and Educators (London: Kogan Page, 2001) e.g p. 113.

Latour presents a model for understanding what a nonhuman does or in other words what role or function it has. He exemplifies with the nonhuman, the door and asks; what work would humans have to do if they had no door? He says that we have delegated the work of reversibly solving the wall-hole dilemma. Latour suggests that "[a]s a more general descriptive rule, every time you want to know what a nonhuman does [what its role or function is], simply imagine what other humans or other nonhumans would have to do were this character not present".[86] This model can be applied to better understand what the relevant kinds of AI does in relation to human cognition. To understand what the nonhumans, the relevant kinds of AI, does we can imagine what other humans or other nonhumans would have to do if the relevant kinds of AI were not present. If AI in personalization algorithms were not present, other humans or nonhumans would have to sort through massive amounts of information to find and decide what information is relevant to the specific human being connecting to the AI, in the specific situation that the human being is in, and for the decision and cognitive process which that human being is about to undertake. It can, therefore, be derived that the AI sorts through these amounts of information, finds and decides what information that is relevant to the specific human being, in that specific situation and for the decision and cognitive process that the human being is about to undertake. If AI in augmented reality technology was not present, other humans or non-humans would have to analyse the surrounding environment and, at the same time, sort through massive amounts of information to find and decide what information is relevant to the specific human being connecting to the AI, in the specific situation, and for the decision and cognitive process that the human being is about to undertake. It can, therefore, be derived that the AI in augmented reality technology is analysing the environment, sorts through information, finds and decides what information is relevant to the specific human being, in the specific situation that the human being is in and for the decision and cognitive process that the human being is about to undertake. What AI in augmented reality technology does is similar to what it does in personalization algorithms, with the addition that AI in augmented reality technology is analysing the surrounding environment. If AI in non-invasive neurotechnology were not present, other humans or non-humans would have to reveal some of the content of human cognition. Other methods of revealing the content of human cognition would require the human being itself to actively reveal the content, through for example telling what it is. It is, therefore, likely that it would, in some instances when the human connecting to the AI is not willing to reveal the content of its cognition, require coercion or force. It can, therefore, be derived that

[86] Latour, Bruno, "Where are the Missing Masses? The Sociology of a Few Mundane Artefacts,", p. 155.

the AI in non-invasive Neurotechnology does reveal the content of human cognition. It is further likely that it, in some instances, does this through coercion or force. With these understandings, of what these kinds of AI do, as a basis, I suggest that it is evident that human cognition is being affected in the encounter with the discussed kinds of AI.

As is discussed in the Posthuman Glossary, the development of new techniques within the field of AI, such as artificial neural networks and machine learning, entails "the emergence of automated systems of knowledge".[87] In that manner, the development of AI has reduced knowledge to information, and information has itself become coinciding with large amounts of complex data that is processed, correlated and modelled by learning algorithms.[88] The introduction of AI has, thus, affected the way in which information is processed. The information that is reaching human beings is, today, often processed by AI. Since human cognition, as it is defined in this thesis, consist of mental processes that is processing information, the altered way in which information is processed is affecting the human cognition. Moreover, Hayles argues that technology presents new capacities to process massive amounts of information. She further argues that it fundamentally extends knowledge production. Technology has in that manner an impact on modes of understanding and learning.[89] Understanding and learning can, in accordance with the definition of cognition used in this thesis, be defined as constituted by human cognition and it can, therefore, be argued that human cognition is affected in the encounter with AI, that is changing the way that information is processed.

The impact of AI in personalization algorithms and so-called filter bubbles has been tested, even though, to my knowledge, not in any largescale researches.[90] One test of the implications of filter bubbles has been based on three signals for human beings in a filter bubble. These being: "(1) a reduction of the active social context, (2) a reduction of the vocabulary a user perceives and (3) users perceives messages mainly or only from the fields of her core

[87] Parisi, Luciana, "Computational turn", in Braidotti, Rosi, & Hlavajova, Maria, (ed.) Posthuman Glossary (London: Bloomsbury Publishing Plc, 2018) [Electronic] Available: ProQuest Ebook Central, https://ebookcentral-proquest-com.ezproxy.ub.gu.se/lib/gu/reader.action?docID=5226228 (Last accessed: 1/12-2019), p. 88.
[88] Parisi, Luciana, "Computational turn", p. 88.
[89] Hayles, N. Katherine, How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics, p. 29-30 & See also Parisi, Luciana, "Computational turn", p. 90.
[90] Further, standard metrics for the filter bubble effect do not seem to have been developed. (Gottron, Thomas & Felix Schwagereit. "The Impact of the Filter Bubble -- A Simulation Based Framework for Measuring Personalisation Macro Effects in Online Communities.").

interests".[91] In this study, it was found that filter bubble effects do occur when AI in personalized algorithms are being used. The human users were found to have a biased perception that was based on the interests of the user. However, the effect was, on average, not strong enough for the user not to be able to get out of the bubble.[92] Moreover, in 2012, a team of psychologists from Cornell University conducted a Facebook emotional manipulation study which involved the feeds of 689,003 Facebook users. The study was controversial, especially since it was performed without the human users' explicit informed consent. The AI in personalization algorithms selecting the content of the human users' feeds were tweaked and manipulated to show more or less positive or negative posts. This was done to examine how such manipulation of the personalization algorithms affected the human users' emotional states. It was found that emotional states can be affected by this kind of altering of personalization algorithms.[93]

It is in this context interesting to point out that some services such as Google are providing an option to turn off personalized searches as an alternative.[94] It is, however, not clear how the remaining filter algorithm is working. Other services are not providing that option. It can, however, still be argued that humans have the choice to use services that are not using personalized filtering. The invisible character of personalization, the investment in social networks and the lack of alternatives can, however, make such a decision difficult.[95] Further, Facebook, which can currently be described as a dominant social networking site has stated that it aims to cater for what has been called happy accidents. The happy accident algorithm is set to render visible unexpected information.[96] If applied, such a design could be decreasing the discussed way in which human cognition is affected in the encounter with AI in personalization algorithms.

Moreover, as pointed out by Bruns, O'Hara and Stevens the effect of personalization algorithms and filter bubbles, as of today, can be questioned and can under some circumstances be

[91] Gottron, Thomas & Felix Schwagereit. "The Impact of the Filter Bubble -- A Simulation Based Framework for Measuring Personalisation Macro Effects in Online Communities.".
[92] Gottron, Thomas & Felix Schwagereit. "The Impact of the Filter Bubble -- A Simulation Based Framework for Measuring Personalisation Macro Effects in Online Communities.".
[93] Jouhki, Jukka, Penttinen, aija, Sormanen, Niina & Uskali, Turo, "Facebook's Emotional Contagion Experiment as a Challenge to Research Ethics." Media and Communication, vol. 4, no. 4, 2016, pp. 75–85. & Paasonen, Susanna, "Networked affect", p. 284.
[94] Google, "Sökinställningar", https://www.google.com/preferences (Last accessed: 31/10-2019).
[95] Holone, Harald, "The Filter Bubble and Its Effect on Online Personal Health Information." Croatian Medical Journal, vol. 57, no. 3, 2016, pp. 298–301.
[96] Paasonen, Susanna, "Networked affect", p. 284.

described as beneficial to human beings.[97] However, the fact that personalization algorithms enable a possibility for human cognition to be affected in new ways can be seen as enough of a reason to consider protection of human cognition regardless of the extent to which the effects are present today.

2.3.2 Understanding human cognition as entangled with artificial intelligence and an altered human vulnerability

It can be argued that human beings have always been affected by nonhumans. As Hayles argues in her article "computing the human" "[a]nthropologists have long recognized that the construction of artifacts and living spaces materially affects human evolution".[98] Hayles further reasons that "[w]e need not refer to something as contemporary and exotic as genetic engineering to realize that for millennia, a two-cycle phenomenon has been at work: humans create objects, which in turn help to shape humans".[99] This does, according to Hayles not diminish the fact that AI is an object created by human beings that are now shaping human beings.

Human cognition can be seen as always having been affected by non-humans, however, this does not diminish the fact that AI is affecting human cognition. Further, the introduction of AI can be argued to have changed the ways in which human cognition can be affected. For example, humans have, according to Pariser, always consumed media that is personally interesting, and ignored other media. However, Pariser argues that the new phenomenon, AI in personalization algorithms, introduces three new dynamics. The first one is that human beings are alone in a filter bubble, as opposed to that not being the case in previous ways of consuming information where there have always been some other humans to share a frame of reference with. The second new dynamic is that the filter bubble is invisible. When consuming information from news sources, such as newspapers or television, many humans know that they are getting the information from a source with a particular political view. The agenda behind AI in personalization algorithms, for example Google's agenda, is opaquer since it isn't visible how the AI is categorizing the human or why it is showing the specific results the human being gets to see. The human being, using services such as Google, haven't chosen the criteria by which the information is filtered through, which can create the illusion that the information is unbiased. Pariser, therefore, argues that it is almost impossible to see how biased the sorting of

[97] Bruns, Axel, Are Filter Bubbles Real?, p. 11 & 15-35.
[98] Hayles, N. Katherine, "Computing the Human.", p. 131.
[99] Hayles, N. Katherine, "Computing the Human.", p. 132.

information actually is. The last new dynamic is that the human doesn't choose to enter the bubble. He argues that it is an active choice to consume newspapers and television channels, while the filter bubble comes to the human.[100]

Further, it has been argued that the case for most computer technologies is such that "[u]sers only see part of the many computations that a computer performs and are for the most part unaware of how it performs them; they usually only have a partial understanding of the assumptions, models, and theories on which the information on their computer screen is based".[101] Pariser has stated that polls show that a huge majority of human beings assume that search engines are unbiased. He argues that this may be a result of the condition that search engines in fact are being increasingly biased to share the human's own views. Pariser, therefore, claims that personalized search marked a turning point of a nearly invisible revolution in how humans consume information.[102] This invisibility of personalization algorithms marks a difference from other ways of receiving information, where the filtering of information is often more visible. Moreover, AI in personalization algorithms reaches beyond personal computers and content.[103] This makes the ways in which AI can affect human cognition far-reaching.

Another example of how AI has changed the ways in which human cognition can be affected is concerning neurotechnology. More rudimental and non-computational techniques, such as interrogation, polygraph-based lie detection, psychoactive drugs, and hypnotic inductions have been used, in ways that can be argued to affect human cognition, before cutting-edge neurotechnology was introduced to the human society. However, these techniques do not directly target neural processing but instead via speech, behaviour or physiological indices. Further, the degree of accuracy of the results of such techniques is very low. Based on such considerations it can be argued that neurotechnology is enabling a significantly higher degree of access into and manipulation of neural processes.[104]

Based on these analyses, it can be argued that the contemporary context of AI influencing human life comprises AI that is affecting human cognition in ways that have not been possible

[100] Pariser, Eli, The Filter Bubble: What the Internet Is Hiding from You, p. 9-10.
[101] Noorman, Merel, "Computing and Moral Responsibility".
[102] Pariser, Eli, The Filter Bubble: What the Internet Is Hiding from You, p. 3 & 9.
[103] See the Guardian, "Advertising billboards use facial recognition to target shoppers"
& Gray, Richard, "Minority Report-style advertising billboards to target consumers" & Wright, David. "The Dark Side of Ambient Intelligence.
[104] Ienca, Marcello & Roberto Andorno, "Towards New Human Rights in the Age of Neuroscience and Neurotechnology.", p. 9-10.

before. When AI is affecting human cognition, it is affecting *how* human beings think, perceive, remember and understand, it is changing the *how* in human cognition. Before the introduction of AI, human cognition could not be affected in these ways. After this introduction, human cognition can be affected in these ways and humans can, therefore, think in other ways. The *how* in human cognition, *how* humans think, perceive, remember, and understand, can therefore be different. Since human cognition is now exposed to these ways of being affected, I suggest that human vulnerability can be seen as altered.

Human beings have been described as biologically vulnerable as embodied agents.[105] Changes to embodiment must, thus, have implications for human vulnerability.[106] The medium for human information processing, and human cognition, is thought to be the nervous system that consists of the brain the peripheral nervous system, and the spinal cord.[107] When the human being and human cognition is being increasingly entangled with AI this can, therefore, have implications for human vulnerability. Human cognition is, as presented in section 2.1, thought to have evolved as a way of controlling action and, thus, as a way processing information, so that the human being can respond to what is perceived.[108] The survival value in being able to perceive is only thought to be present if the perceived information can be processed.[109] When AI in personalization algorithms and augmented reality technology is affecting human cognition, through deciding what information the human being is presented with, and therefore is perceiving, it is affecting the evolved way of controlling action. The perceived information can be seen as the basis for action and humans must, therefore, be seen as vulnerable to filtering of information that does not constitute a suitable basis for action. Since information can be seen as the basis for action, there is a vulnerability in the fact that information can be filtered in ways that are not suitable as a basis for action. Since the methods that AI is using for filtering information is largely unknown to a human being, it is difficult for the human to decide if the information is suitable for action or not.[110] The evolved way of processing information is therefore vulnerable when AI in personalization algorithms and augmented reality technology is filtering information. Further, AI in neurotechnology is revealing parts of the content of human cognition. It is, therefore, revealing content of the evolved way of controlling action and

[105] C.f. Turner, Bryan S., "Sociology of Human Rights", p. 6 & 9.
[106] Turner, Bryan S., "Sociology of Human Rights", p. 6 & 9.
[107] Glass, Arnold, Cognition, p. 8.
[108] See Glass, Arnold Lewis, Cognition, p. 2.
[109] See Glass, Arnold Lewis, Cognition, p. 2.
[110] C.f. Noorman, Merel, "Computing and Moral Responsibility" & C.f. Pariser, Eli, The Filter Bubble: What the Internet Is Hiding from, p. 9-10.

of processing information. As discussed in section 2.3.1, the revealing of parts of the content of human cognition can, further, be seen as done through some kind of coercion or force since some kinds of neurotechnology can reveal content of human cognition without the consent of the human being. This kind of AI makes it possible to, in real-time, observe and update the influence that is exercised on human cognition. The human being and human cognition can, therefore, be seen as vulnerable to the possibility of revealing the content of cognition without its consent. Based on this analysis human cognition can be seen as vulnerable to AI in personalization algorithms and augmented reality technology filtering information and to the possibility of revealing the content of human cognition without the human's consent. Human vulnerability can be seen as altered by the condition that human cognition is now increasingly entangled with AI and exposed to these new ways of being affected.

3. The Concept Human Rights

3.1.1 A brief history of International human rights

The idea of human rights has a long history. Going back in history, concepts related to human rights can be found in ancient civilizations, across the world. It is probably not possible to define a precise starting point for the history of human rights or to credit any specific culture, region or religion with its origins. The human rights system is complex, there are many different aspects to it, and it is impossible to identify one of these aspects as the beginning. Despite these difficulties, the Magna Carta is typically regarded as the starting point of the history of human rights.[111]

It is, however, clear that human rights history started on the domestic level. Before the 1940s, there was no real conception of human rights on the international level. During the second world war, Franklin Roosevelt and Winston Churchill spoke about human freedoms, and human rights became one of the moral bases upon which the war was fought. When the war ended the United Nations (UN) charter was framed as a starting point for the endeavour to create legal instruments protecting human rights. The UN Commission on Human rights was instituted to create an international bill of rights, a human rights instrument applicable to all states and people across the globe. The work of the commission resulted among other things in the Universal Declaration of Human rights. 1954 the Commission also completed the drafts of what would become the ICCPR and the ICESCR.[112]

Eleanor Roosevelt, who chaired the Commission, described the UDHR as a possible Magna Carta for mankind.[113] The Magna Carta was a kind of a starting point for domestic protection of human rights. In the same manner, the UDHR can be seen as a starting point for international protection of human rights.

It has been expressed that;

> the history of human rights law, both at the domestic and the international level, confirms that '[t]he vindication of human liberties does not begin with their complete and triumphant assertion at the

[111] Bates, Ed, "History", in Moeckli, Daniel, Sangeeta Shah, Sandesh Sivakumaran, & David J. Harris, (ed.) International Human Rights Law. Third ed., (Oxford: Oxford University Press 2018), p. 4-5.
[112] Bates, Ed, "History", p. 16-19.
[113] Bates, Ed, "History", p. 16-19.

> very outset'. Rather, 'it commences with their recognition in *some* matters, to *some* extent, for *some* people, against *some* organ of the State'.[114]

The history of human rights can in this manner be seen as ongoing.

3.1.2 What should be protected by human rights? Foundations and justification for human rights

The traditional view of human rights is that they are inherent to human beings and that they are to be attributed to humans simply by the fact of being human.[115] Human rights can, in line with that view, be seen as protecting universal or common human interests. A challenge with this idea is to determine interests that every human shares.[116] It poses the question of what it is that humans have in common that allows us to talk about a common world and common human interests.[117] Is there something in the human ontology that can secure a foundation for common human interests and human rights? [118] Human beings have been described as biologically vulnerable as embodied agents and, therefore, in need of social institutions in order to protect themselves from the contingencies of the world. Human rights require a wide social consensus and it has been suggested that the needed social consensus could be grounded in recognition of common human vulnerability and the need for effective social institutions.[119] Common human vulnerability can, therefore, be seen as a foundation for human rights. If human vulnerability is seen as something common in human ontology, human vulnerability entails the condition that there can be common human interests that are vulnerable. It has, further, been claimed that humans have rights only to interests necessary to be agents,[120] and that human rights should respond to common and serious threats to important goods or interests.[121] Based on these ideas,

---

[114] Bates, Ed, "History", p. 20-21
[115] van Duffel, Siegfried "Moral Philosophy", in Shelton, Dinah, (ed.) The Oxford Handbook of International Human Rights Law (Oxford: Oxford University Press 2013) [Electronic] Available: Oxford university press, https://opil-ouplaw-com.ezproxy.ub.gu.se/view/10.1093/law/9780199640133.001.0001/law-9780199640133 (Last accessed: 1/12-2019), p. 2.
[116] van Duffel, Siegfried "Moral Philosophy", p. 5.
[117] C.f. Turner, Bryan S., "Sociology of Human Rights", p. 6.
[118] C.f. Turner, Bryan S., "Sociology of Human Rights", p. 6.
[119] Turner, Bryan S., "Sociology of Human Rights", p. 5-6 & 9.
[120] van Duffel, Siegfried "Moral Philosophy", p. 15.
[121] Nickel, J., "Human Rights", The Stanford Encyclopedia of Philosophy, Zalta, Edward N., (ed.) 2014, https://plato.stanford.edu/archives/spr2017/entries/rights-human (Last accessed: 2/12-2019) & See also Ienca, Marcello & Roberto Andorno, "Towards New Human Rights in the Age of Neuroscience and Neurotechnology.", p. 9.

common human interests that are vulnerable and necessary for humans to be agents can be seen as common human interests that should be protected by human rights.[122]

Moral justifications that are commonly put forward as justifications for human rights are equality and dignity, also referred to in the preambles of the UDHR, ICCPR, and ICESCR.[123] Equality as a moral justification can also be referred to as equal moral status. Equal moral status is the idea that all people have equal worth. Equal moral status can also be seen as including the idea that there are claims that all humans are entitled to make on one another, simply based on their status as human beings. The two core ideas, in this understanding of equal moral status, are that all humans should be seen as having the same moral worth and that the equal moral worth is the basis for mutual moral claim. The concept of human rights is one phenomenon that constitutes a human's equal moral status and human rights protect interests that can give rise to entitlements that are constitutive of equal moral status.[124] Dignity, as a moral justification, can be used to refer to what is unique in human beings. This can be formulated as the possession of personhood and the capacity for rational and moral agency. Dignity can be seen as identical to the idea captured by the concept of equal moral status. It is thus unclear if dignity is an independent justification for human rights. Article 1 in the UDHR refers to human beings being born equal in dignity and rights. This implies that dignity is not a justification on its own, but rather a value that can be seen as a part of the justification equality.[125]

3.1.3 Interpretation of human rights

Interpretation of human rights is a complex undertaking. Rules for interpretation of international treaties are provided in the 1969 Vienna Convention on the Law of Treaties. This treaty was concluded in 1969, which is after the drafting of the human rights treaty, discussed and interpreted in this thesis. Even though VCLT is non-retroactive, human rights law has generally acknowledged the application of VCLT, and its rules of interpretation, as customary law.[126]

---

[122] C.f. Nickel, J., "Human Rights" & See also Ienca, Marcello & Roberto Andorno, "Towards New Human Rights in the Age of Neuroscience and Neurotechnology.", p. 9 & C.f. Turner, Bryan S., "Sociology of Human Rights", p. 5-6 & 9 & van Duffel, Siegfried "Moral Philosophy", p. 15.
[123] Besson, Samantha, "Justifications", p. 32-34 & International Covenant on Civil and Political Rights & Universal Declaration of Human Rights.
[124] Besson, Samantha, "Justifications", p. 32-34.
[125] Besson, Samantha, "Justifications", p. 34-35.
[126] Fitzmaurice, Malgosia, "Interpretation of Human Rights Treaties", p. 1.

Prior to the VCLT, there were, broadly, three schools of interpretation of human rights. The first of these laid great emphasis on the travaux préparatoires, the second was the textual approach and the third was the teleological approach, placing great emphasis on the object and purpose of the treaty. The three schools of thought were incorporated in the VCLT. [127] The provisions that provide the rules for interpretation in VCLT is Article 31 and 32;

> Article 31, GENERAL RULE OF INTERPRETATION
> 1. A treaty shall be interpreted in good faith in accordance with the ordinary meaning to be given to the terms of the treaty in their context and in the light of its object and purpose.
> 2. The context for the purpose of the interpretation of a treaty shall comprise, in addition to the text, including its preamble and annexes:
> (a) Any agreement relating to the treaty which was made between all the parties in connexion with the conclusion of the treaty;
> (b) Any instrument which was made by one or more parties in connexion with the conclusion of the treaty and accepted by the other parties as an instrument related to the treaty.
> 3. There shall be taken into account, together with the context:
> (a) Any subsequent agreement between the parties regarding the interpretation of the treaty or the application of its provisions;
> (b) Any subsequent practice in the application of the treaty which establishes the agreement of the parties regarding its interpretation;
> (c) Any relevant rules of international law applicable in the relations between the parties.
> 4. A special meaning shall be given to a term if it is established that the parties so intended.
>
> Article 32. SUPPLEMENTARY MEANS OF INTERPRETATION
> Recourse may be had to supplementary means of interpretation, including the preparatory work of the treaty and the circumstances of its conclusion, in order to confirm the meaning resulting from the application of article 31, or to determine the meaning when the interpretation according to article 31:
> (a) Leaves the meaning ambiguous or obscure; or
> (b) Leads to a result which is manifestly absurd or unreasonable.[128]

These articles are not considered to provide step-by-step formulas for interpretation in every case, but rather to indicate elements that should be taken into consideration.[129]

Rules that are of certain interest in interpreting human rights can be found in Article 31(1), (2), (3) & (4). Article 31(1) provides that interpretation shall be made in accordance with the ordinary meaning to be given to a provision in the context and in light of its object and purpose.[130] Since human rights are usually formulated with abstract concepts or general terminology, it may prove especially difficult to find an acceptable ordinary meaning in the

[127] Fitzmaurice, Malgosia, "Interpretation of Human Rights Treaties", p. 5-6.
[128] Vienna Convention on the Law of Treaties.
[129] Fitzmaurice, Malgosia, "Interpretation of Human Rights Treaties", p. 7.
[130] Vienna Convention on the Law of Treaties.

field of human rights.[131] The contemporaneity is also problematic since the ordinary meaning of a term is developing with time. Moreover, the object and purpose can be hard to establish.[132] Article 31(1) is incorporating the principle of integration that can be seen as widening the scope of ordinary meaning. Article 31(2) is further, extending the concept "context" to include other related agreements. Article 31(3) extend the scope of interpretation even further to include important additional matter. Article 31(3)(c) more specifically permits any relevant rules of international law to be taken into account.[133]

Moreover, article 31(4) provides that a "special meaning shall be given to a term if it is established that the parties so intended".[134] Given the growing importance of multilateral treaties, interpretation based on common intentions of the parties has become harder and harder. This shift, from bilateral to multilateral agreements have thus led to an interpretive methodology concentrated on, what have been called, more objective and ascertainable principles. The intention of the parties does, nevertheless, remain an important factor for interpretation.[135]

The rules of interpretation in VCLT provide a starting point for interpreting human rights. Interpretation of human rights does, however, also require consideration of other practices, that have been introduced by human rights courts and human rights treaty bodies. Human rights bodies have frequently asserted that their interpretation is consistent with the rules in VCLT. However, positions that are hard to reconcile with the VCLT provisions have been adopted and these are generally viewed as, at least, expanding on traditional methods or even as introducing new methods. Further, the provisions in VCLT are not entirely clear and in addition to that, human rights are often general, vague and subjective in character.[136] It is, therefore, generally acknowledged that an expansive attitude towards the interpretation of human rights is needed.[137]

[131] Mcgrogan, David, "On the Interpretation of Human Rights Treaties and Subsequent Practice." Netherlands Quarterly of Human Rights, vol. 32, no. 4, 2014, pp. 347-78 & Fitzmaurice, Malgosia, "Interpretation of Human Rights Treaties", p. 8.
[132] Fitzmaurice, Malgosia, "Interpretation of Human Rights Treaties", p. 8.
[133] Vienna Convention on the Law of Treaties & Fitzmaurice, Malgosia, "Interpretation of Human Rights Treaties", p. 8.
[134] Vienna Convention on the Law of Treaties, Article 31(4) & Fitzmaurice, Malgosia, "Interpretation of Human Rights Treaties", p. 7.
[135] Fitzmaurice, Malgosia, "Interpretation of Human Rights Treaties", p. 7.
[136] Fitzmaurice, Malgosia, "Interpretation of Human Rights Treaties", p. 1-2.
[137] Fitzmaurice, Malgosia, "Interpretation of Human Rights Treaties", p. 8.

Human rights share certain specific characteristics, that may influence the interpretation. One shared characteristic is the subject matter which gives rise to an approach to interpretation sometimes referred to as the pro homine approach, or the notion that international human rights should be centred around the human being.[138] A second shared characteristic is the so-called constitutional and the non-reciprocal nature of human rights. This may lead to a reduction of the importance of the actual text of a treaty, in relation to other factors, in particular, the object and purpose of a treaty.[139] Further, by giving considerable weight to preambles of conventions and even to human rights declarations, to which the preambles often refer, a strong teleological approach to interpretation has been adopted by human rights fora.[140]

An interpretative tool based on the notion of object and purpose is evolutive interpretation.[141] UN human rights bodies have, relied upon the concept of dynamic interpretation of treaties, interpreting treaties according to contemporary standards. The Human Rights Committee has for example, expressly stated that the ICCPR should be interpreted as a living instrument and that the provisions of that treaty should be interpreted in the context and in light of present-day conditions.[142]

Another tool of interpretation that has been used in human rights courts and tribunals is the object and purpose test, which relates to the doctrine of effectiveness. This doctrine is captured by the maxim that a treaty is presumed to have a definite force and effect. In order to ensure that all provisions of a treaty have an independent and non-superfluous meaning, a treaty may be interpreted expansively.[143]

Moreover, to interpret and clarify the meaning of provisions, the Human Rights Committee does in accordance with ICCPR adopt general comments on articles in the covenant.[144]

---

[138] Mazzuoli, Valerio de Oliveira & Dilton Riberio, "The Pro Homine Principle as a Fundamental Aspect of International Human Rights Law/O Principio Pro Homine Como Um Aspecto Fundamental Do Direito Internacional Dos Direitos Humanos." Meridiano 47, vol. 17, no. 47, 2016, p. 5 & Fitzmaurice, Malgosia, "Interpretation of Human Rights Treaties", p. 3.
[139] Fitzmaurice, Malgosia, "Interpretation of Human Rights Treaties", p. 3.
[140] Fitzmaurice, Malgosia, "Interpretation of Human Rights Treaties", p. 21.
[141] Fitzmaurice, Malgosia, "Interpretation of Human Rights Treaties", p. 9.
[142] Fitzmaurice, Malgosia, "Interpretation of Human Rights Treaties", p. 21 & Roger Judge v. Canada, CCPR/C/78/D/829/1998, UN Human Rights Committee (HRC), 13 August 2003, para 10.3. Further, the CERD Committee made similar statements (Hagan v. Australia, Comm. 26/2002, U.N. Doc. A/58/18, at 139 (2003), para 7.3), as did the CAT Committees (V.X.N. and H.N. v. Sweden, CAT/C/24/D/130/1999; CAT/C/24/D/131/1999, UN Committee Against Torture (CAT), 2 September 2000).
[143] Fitzmaurice, Malgosia, "Interpretation of Human Rights Treaties", p. 10.
[144] International Covenant on Civil and Political Rights, Article 40, paragraph 4.

3.2 Recognized substantive human rights in connection to human cognition

With the complexity of human cognition in mind, it is possible to find connections to many of the recognized international human rights. Freedom of thought, conscience and religion is perhaps the one most commonly referred to when discussing the protection of human cognition. Another right that is closely connected to human cognition is freedom of opinion and expression. In the contemporary context of AI influencing human life other rights, such as the right to privacy, the right to property, the right to a fair trial and the principle against self-incrimination, can be relevant for the protection of human cognition. This thesis is, however, focusing on the rights in closest connection to human cognition. The following sections 3.2.1 and 3.2.2 are thus focusing on how these two rights have been described and interpreted. The other mentioned rights can possibly only protect parts of human cognition, such as privacy of human cognition, data about human cognition as property and parts of human cognition that can be self-incriminating. The study of these rights is therefore not necessary to understand if human cognition is sufficiently protected in the international human rights framework.

3.2.1 Freedom of thought, conscience and religion

> Article 18
>
> Everyone has the right to freedom of thought, conscience and religion; this right includes freedom to change his religion or belief, and freedom, either alone or in community with others and in public or private, to manifest his religion or belief in teaching, practice, worship and observance.

- Universal Declaration of Human Rights[145]

> Article 18

1. Everyone shall have the right to freedom of thought, conscience and religion. This right shall include freedom to have or to adopt a religion or belief of his choice, and freedom, either individually or in community with others and in public or private, to manifest his religion or belief in worship, observance, practice and teaching.
2. No one shall be subject to coercion which would impair his freedom to have or to adopt a religion or belief of his choice.
3. Freedom to manifest one's religion or beliefs may be subject only to such limitations as are prescribed by law and are necessary to protect public safety, order, health, or morals or the fundamental rights and freedoms of others.

- International Covenant on Civil and Political Rights[146]

[145] Universal Declaration of Human Rights.
[146] International Covenant on Civil and Political Rights.

The right to freedom of thought, conscience and religion is recognized in the UDHR and in the ICCPR. It has been discussed as an essential human right, and some of the drafters of the UDHR even called it the most important right in the declaration and the basis and origin of all other rights.[147]

The right to freedom of thought can be seen as connected to human cognition, since human cognition is mental processes that, as it is defined in this thesis, is constituting thought.[148] The wording of the article in which the right is recognized does, thus, entail a connection between the right and human cognition.

However, the wording of the right is unclear as to what is included in the protection and in what way human cognition can be protected. Complaints to supervisory international organs have relatively seldom regarded the right in question.[149] Since the right has been seen as having an unproblematic character, there are, moreover, few clarifications to the meaning of the right. It has been said in the literature that the:

> difficulties start when we come to the right to express one's conviction, the right to organize as a community in order to promote a religion or belief and the right to act in accordance with one's conscience even in cases where a domestic legal system seems to require uniform behaviour irrespective of the different convictions held by individuals. The real problems concerning freedom of thought, conscience and religion do not concern the nucleus of the right itself, the freedom of the *forum internum* or an inner state of mind, but issues that relate also to other human rights. The interdependence of the right under discussion with other human rights, including freedom of expression, freedom of assembly and freedom of association, is evident.[150]

At the time of the drafting of the UDHR one of the drafters stated that it would not be necessary to protect human thought that was not to be given an outward expression;

> it would be unnecessary to proclaim that freedom if it were never to be given an outward expression; if it were intended, so to speak, only for the use of the inner man. It was necessary however to stress the external manifestation of creeds by which expression was given to beliefs.[151]

This statement seems to suggest that the protection of the forum internum is actually not what has been seen as central to the protection given by this human right. However, the nucleus of

[147] GA, Third Committee, Third Session, 127th Meeting, at 395. & GA, Third Committee, Third Session, Draft International Declaration of Human Rights: Recapitulation of Amendments to Article 16 of the Draft Declaration, p. 397-401 & Scheinin, Martin, "Freedom of Thought, Conscience and Religion." Studia Theologica - Nordic Journal of Theology, vol. 54, no. 1, 2000, pp. 5–18, p. 8.
[148] See section 2.1.
[149] Scheinin, Martin, "Freedom of Thought, Conscience and Religion.", p. 5-6.
[150] Scheinin, Martin, "Freedom of Thought, Conscience and Religion.", p. 5-6.
[151] GA, Third Committee, Third Session, 127th Meeting, at 395.

the right to freedom of thought, conscience and religion cannot be derogated from, even in time of public emergency. This indicates a fundamental character of the right.[152] Article 18 ICCPR, further, distinguishes the freedom of thought, conscience, religion or belief from the freedom to manifest religion or belief. Limitations are not permitted to the freedom of thought and conscience or to the freedom to have or to adopt a religion or belief. The freedom to manifest religion or belief can, however, be subject to limitations that are "prescribed by law and are necessary to protect public safety, order, health, or morals or the fundamental rights and freedoms of others".[153] The absolute character of the freedom of what was called the inner state of mind was also stressed during the drafting of the UDHR.[154] At the drafting of the ICCPR, it was agreed that no restrictions of legal character can be imposed on "man's inner thought or moral consciousness, or his attitude towards the universe or its creator; only external manifestations of religion or belief might be subject to legitimate limitations".[155] This implies that protection of what has been called the forum internum should be seen as central and absolute.

The committee has, further, pointed out that the right includes protection that means that no one shall be compelled to reveal thoughts or adherence to a certain religion or belief.[156]

Moreover, the predominant focus of the right has been interpreted to be the right to freedom of religion.[157] In 1993, the Human Rights Committee adopted the General comment No. 22: Article 18 (Freedom of thought, conscience or religion) that was supposed to clarify the meaning of the right. The Human Rights Committee stated, that it wished to draw attention to the fact that freedom of thought and freedom of conscience are protected equally with freedom of religion and belief.[158] The importance of the freedom of thought in relation to the freedom of religion was also discussed during the drafting of UDHR.[159] The delegation from Uruguay

[152] International Covenant on Civil and Political Rights, Article 4.2 & HRC, General Comment 22. HRI/GEN/1/Rev9 (VolI).
[153] International Covenant on Civil and Political Rights, Article 18.3 & The inclusion of possible limitations of the freedoms in question in Universal Declaration of Human Rights, Article 29.
[154] Scheinin, Martin. "Freedom of Thought, Conscience and Religion.", p. 7.
[155] UN Secretary-General, Annotations on the Text of the Draft International Covenants on Human Rights, 1 July 1955, A/2929 at para 99, p. 48.
[156] HRC, General Comment 22.
[157] McGoldrick, Dominic, "Thought, Expression, Association, and Assembly", in Moeckli, Daniel, Sangeeta Shah, Sandesh Sivakumaran, & David J. Harris, (ed.) International Human Rights Law. Third ed., (Oxford: Oxford University Press 2018), p.210-21.
[158] HRC, General Comment 22.
[159] GA, Third Committee, Third Session, meeting 127, at 395. & GA, Third Committee, Third Session, Draft International Declaration of Human Rights: Recapitulation of Amendments to Article 16 of the Draft Declaration, (1948), A/C.3/289/REV.1. p. 391, 395, 397 & 401.

did, for example, put forward two observations. “First, he thought that it laid too much stress on religious freedom; actually, and especially at that very time, freedom of thought should be extended particularly to the realms of politics and science.”[160]

UDHR explicitly expresses protection of the freedom to change one’s religion or belief as well as to publicly manifest religion or belief. This protection is not explicitly expressed in connection to freedom of thought and conscience. However, it is not clear what is included in the term belief and to what extent that term could include the complexity of human cognition. The terms are not defined in the instruments, but it has been stated in the literature that the terms, religion and belief together, cover “all possible attitudes of the individual toward the world, toward society, and toward that which determines his fate and the destiny of the world, be it a divinity, some superior being or just reason and rationalism, or chance”.[161] The wording “change religion or belief”,[162] was largely discussed during the drafting of the UDHR. For example, the Saudia Arabia delegation pointed at the history of missionaries and efforts to convert human beings as a risk, present in the formulation.[163]

In the ICCPR, the corresponding protection is formulated as a freedom to have or to adopt a religion or belief of one’s choice.[164] This formulation was put forward in order to provide protection, both for the right to change religion and for a right against proselytizers and missionaries.[165] The Human Rights Committee has stated that the freedom, to have or to adopt a religion or belief, necessarily entails the freedom to choose a religion or belief and that it thus includes the right to replace a current religion or belief with another, or to adopt an atheistic view, but also to retain one’s religion or belief. It has, further, stated that the formulation forbids coercion that would impair the right to have or adopt a religion or belief. This prohibition includes policies or practices that would have the same intention or effect as coercion.[166] Further, the ICCPR provides a provision that explicitly says that “no one shall be subject to coercion which would impair his freedom to maintain or to change his religion or belief”.[167]

---

[160] GA, Third Committee, Third Session, meeting 127, at 395. & GA, Third Committee, Third Session, Draft International Declaration of Human Rights: Recapitulation of Amendments to Article 16 of the Draft Declaration, p. 40.
[161] Scheinin, Martin, “Freedom of Thought, Conscience and Religion.”, p. 5-6.
[162] Universal Declaration of Human Rights, Article 18.
[163] GA, Third Committee, Third Session, meeting 127, at 395. & GA, Third Committee, Third Session, Draft International Declaration of Human Rights: Recapitulation of Amendments to Article 16 of the Draft Declaration, p. 391.
[164] International Covenant on Civil and Political Rights.
[165] Scheinin, Martin, “Freedom of Thought, Conscience and Religion.”, p. 5.
[166] HRC, General Comment 22.
[167] International Covenant on Civil and Political Rights, Article 18.

During the drafting it was said that the term coercion should not be understood as "applying to moral or intellectual persuasion, or to any legitimate limitation of freedom to manifest one's religion or belief".[168] The boundary between coercion and intellectual persuasion is not further defined.

The protection to "have or to adopt" in the ICCPR is provided with respect to religion or belief. This implies that beliefs other than religious are protected by this part of the protection.[169] The Human Rights Committee has stated that the wording: religion or belief shall be interpreted broadly and that it should not be limited to traditional religions.[170] The Committee has also stated that theistic, non-theistic and atheistic beliefs as well as choosing not to confess to any belief should be included in the protection.[171] The terms "belief" and "religion" are thus, in accordance with the comments to be broadly construed. The committee further, stated that the article:

> is not limited in its application to traditional religions or to religions and beliefs with institutional characteristics or practices analogous to those of traditional religions. The Committee therefore views with concern any tendency to discriminate against any religion or belief for any reason, including the fact that they are newly established, or represent religious minorities that may be the subject of hostility on the part of a predominant religious community.[172]

Hence, it is clear that not only traditional religions or beliefs are protected. There are, however, limits to what have been seen as protected. In a case, the Committee found that the cultivation and worship of a narcotic drug could not be protected as a religion or belief under the right in ICCPR.[173]

In the drafting of the ICCPR, the two terms thought, and belief were discussed with regards to how they relate to each other and if they were intended to be different concepts. It was said by one of the drafters that a distinction between the terms seemed to have been contemplated and that it would be desirable to clarify the meaning of the two terms. The term thought was also discussed in relation to the term, opinion, in article 19.[174] There does, however, not seem to

[168] UN Secretary-General, Annotations on the Text of the Draft International Covenants on Human Rights, p. 48.
[169] International Covenant on Civil and Political Rights, Article 18.
[170] HRC, General Comment 22.
[171] HRC, General Comment 22.
[172] HRC, General Comment 22.
[173] MAB, WAT and JAYT v Canada, CCPR/C50/D570/1993 (25 April 1994). C.f. Prince v South Africa, CCPR/C/91/D/1474/2006 (14 November 2007).
[174] UN Secretary-General, Annotations on the Text of the Draft International Covenants on Human Rights, p. 48, Summary record of the 10th meeting, held at Lake Success, New York, on Wednesday, 18 June 1947 : Commission on Human Rights, Drafting Committee, 1st session, E/CN.4/AC.1/SR.10, 26, Commission on

exist any clarification of any common intent to give these different concepts any special meaning.

The Human Rights committee has stated that the freedom of thought, conscience and religion is far-reaching and profound and that it "encompasses freedom of thought on all matters, personal conviction and the commitment to religion or belief, whether manifested individually or in community with others".[175] This comment implies that the freedom is to be interpreted broadly.

3.2.2 Freedom of opinion and expression

> Article 19
>
> Everyone has the right to freedom of opinion and expression; this right includes freedom to hold opinions without interference and to seek, receive and impart information and ideas through any media and regardless of frontiers.

- Universal Declaration of Human rights[176]

> Article 19
>
> 1. Everyone shall have the right to hold opinions without interference.
>
> 2. Everyone shall have the right to freedom of expression; this right shall include freedom to seek, receive and impart information and ideas of all kinds, regardless of frontiers, either orally, in writing or in print, in the form of art, or through any other media of his choice.

- International Covenant on Civil and Political Rights[177]

Freedom of opinion and expression is recognized in UDHR and ICCPR. The right at hand provides protection for human beings, to hold opinions without interference. Human beings have different opinions in all sorts of questions and the right has been described as a right for people to think what they like.[178] Since human cognition is consisting of mental processes that, as it is defined in this thesis, is constituting thought and remembering, it is clear that this right protects some dimensions of human cognition. The right does further include a right to inform

---

Human Rights, 8th session: summary record of the 319th meeting, E/CN.4/SR.116, 117, 319, Drafting Committee, First session, 10th Meeting (1947), E/CN.4/82/Rev.l, &
Official Records of the General Assembly, Fifth Session, Third Committee, Draft first international covenant on human rights and measures of implementation (A/1374, A/C.3/534, A/C.3/535 and E/1681), 289th meeting. A/C.3/SR.289.

[175] HRC, General Comment 22.
[176] Universal Declaration of Human Rights.
[177] International Covenant on Civil and Political Rights.
[178] McGoldrick, Dominic, "Thought, Expression, Association, and Assembly", p. 217-219.

oneself; to be informed and to access information. Since human cognition is consisting of mental processes that, as it is defined in this thesis, are processing information, this part of the right too is relevant with respect to a potential protection of human cognition.

The wordings of the provisions that formulate this right are not clear as to what parts and to what extent human cognition is being given protection. The Human Rights Committee has adopted a general comment to the article. This comment is, however, very short and does not develop the interpretation of the wording very much. The committee briefly states that the right cannot be restricted and that the committee welcomes information from state parties, concerning the first paragraph of the article.[179] In a publication, the UN writes that it is logical that the covenant does not permit exceptions or restrictions to the right since it is impossible to control what goes on in a person's mind.[180] This statement implies, on the one hand, that the right is important since no restrictions are permitted, but on the other hand, that the scenario, where "control"[181] over the "mind"[182] of a human being is possible, has not been considered by the UN.

At the drafting of the UDHR, several delegations pointed at the importance of the right. For example, Mrs. Roosevelt (United States of America) stated that "no human rights were more fundamental than freedom of opinion and expression".[183] Another example is that "Mr. Corominas (Argentina) emphasized the fact that article 17 was perhaps the most important of the whole declaration".[184]

Mr. Corominas said that "without it the individual would be deprived of freedom of expression, which was his most effective weapon for defending democratic institutions and the very principles on which the declaration was based".[185] This statement can be interpreted as implying that this delegation thought the freedom of expression to be the central aspect of the article.

[179] HRC, General Comment 10.
[180] United Nations, "Manual on Human Rights for Judges, Prosecutors and Lawyers", Published: 2003. https://www.un.org/ruleoflaw/files/training9chapter12en.pdf (Last accessed: 24/11-2019).
[181] United Nations, "Manual on Human Rights for Judges, Prosecutors and Lawyers".
[182] United Nations, "Manual on Human Rights for Judges, Prosecutors and Lawyers".
[183] GA, Third Committee, Third Session, meeting 127, at 395. & GA, Third Committee, Third Session, Draft International Declaration of Human Rights: Recapitulation of Amendments to Article 16 of the Draft Declaration, p. 413.
[184] GA, Third Committee, Third Session, meeting 127, at 395. & GA, Third Committee, Third Session, Draft International Declaration of Human Rights: Recapitulation of Amendments to Article 16 of the Draft Declaration, p. 419.
[185] GA, Third Committee, Third Session, meeting 127, at 395. & GA, Third Committee, Third Session, Draft International Declaration of Human Rights: Recapitulation of Amendments to Article 16 of the Draft Declaration, p. 419.

Further, the focus in the discussion of the article seems to have been on the right to freedom of expression. This could imply that freedom of expression was seen as the part in certain need of discussion and clarification, but it could also imply that it was seen as the more important part of the right. The delegation from Panama proposed separation between the two freedoms of opinion and expression.[186] The delegation from the United Kingdom did, on the other hand, state that it considered such a separation unwise since the two ideas were inalterably linked.[187] At the drafting of the ICCPR the relation between the two freedoms was discussed again;

> it became clear that freedom of opinion and freedom of expression were not of the same character: the former was purely a private matter, belonging as it did to the realm of the mind, while the latter was a public matter, or a matter of human relationship, which should be subject to legal as well as moral restraint. Although it was claimed that a person was invariably conditioned or influenced by the world, it was generally agreed that no law could regulate his opinion and no power could dictate what opinion he should or should not entertain.[188]

It was during the drafting of ICCPR, in contrast with the drafting of UDHR, decided that the covenant should treat the freedom of opinion and the freedom of expression separately.[189] This separation of the two freedoms makes it clear that what was called the realm of the mind; i.e., the holding of opinions, should be protected independently of the public matter of outward expression.

Another question that was discussed at the drafting of ICCPR was if there was any distinction between the freedom of opinion and the freedom of thought, as stated in the covenant's article 18. "One comment was to the effect that the words 'thought' and 'opinion', though not identical, were very close to each other in meaning; another that the two words were not mutually exclusive but complementary to each other; a third that 'freedom to hold any opinions without interference' was a truism and therefore superfluous."[190] These discussions can, however, not be seen as establishing a common intent to any special meaning of the words and can thus not be seen as giving any distinct clarification to what should be protected by the respective provisions.

---

[186] GA, Third Committee, Third Session, meeting 127, at 395. & GA, Third Committee, Third Session, Draft International Declaration of Human Rights: Recapitulation of Amendments to Article 16 of the Draft Declaration, (1948), A/C.3/289/REV.1. p. 427.
[187] GA, Third Committee, Third Session, meeting 127, at 395. & GA, Third Committee, Third Session, Draft International Declaration of Human Rights: Recapitulation of Amendments to Article 16 of the Draft Declaration, p. 411.
[188] UN Secretary-General, Annotations on the Text of the Draft International Covenants on Human Rights, p. 51.
[189] UN Secretary-General, Annotations on the Text of the Draft International Covenants on Human Rights, p. 51.
[190] UN Secretary-General, Annotations on the Text of the Draft International Covenants on Human Rights, p. 51.

4. Cognitive Liberty

Cognitive liberty has been described as "a concept that deals with privacy and self-determination and autonomy with respect to what is happening in human brains".[191] It has also been called mental self-determination.[192] The concept has been argued to be of fundamental character since "the right and freedom to control one's own consciousness and electrochemical thought processes is the necessary substrate for just about every other freedom".[193]

The right to cognitive liberty can be seen as resembling the recognized international human right to freedom of thought. Wrye Sententia, one of the legal scholars who coined the term cognitive liberty, presented it as a conceptual update to the freedom of thought. The update was presented as way of taking into account the contemporary context of AI influencing human life and the new and increasing power to monitor and manipulate cognitive functions.[194] Sententia has, further, argued that cognitive liberty is a fundamental right to use the full spectrum of the mind and that the right concerns the "ethics and legality of safeguarding one's own thought processes, and by necessity, one's electrochemical brain states".[195] She has also stated that humans and not corporate or government interests should have jurisdiction over control and modulation of mental processes and human brain states.[196] Sententia has also stated that "the right and freedom to control one's own consciousness and electrochemical thought processes is the necessary substrate for just about every other freedom".[197]

[191] Farahany, Nita, "Cognitive Liberty in the Era of Brain Hacking, The Aspen Institute"da.
[192] Ienca, Marcello & Roberto Andorno, "Towards New Human Rights in the Age of Neuroscience and Neurotechnology.", p. 10.
[193] Wrye, Sententia, "Neuroethical Considerations: Cognitive Liberty and Converging Technologies for Improving Human Cognition." & Ienca, Marcello & Roberto Andorno, "Towards New Human Rights in the Age of Neuroscience and Neurotechnology.", p. 10.
[194] Wrye, Sententia, "Neuroethical Considerations: Cognitive Liberty and Converging Technologies for Improving Human Cognition.".
[195] Wrye, Sententia, "Neuroethical Considerations: Cognitive Liberty and Converging Technologies for Improving Human Cognition.".
[196] Wrye, Sententia, "Neuroethical Considerations: Cognitive Liberty and Converging Technologies for Improving Human Cognition".
[197] Wrye, Sententia, "Neuroethical Considerations: Cognitive Liberty and Converging Technologies for Improving Human Cognition.".

The concept is further seen as multi-dimensional. Three recognized interrelated dimensions are:

> (i) the liberty to change one's mind or to choose whether and by which means to change one's mind; (ii) the protection of interventions into other minds to protect mental integrity, and (iii) the ethical and legal obligation to promoting cognitive liberty.[198]

The concept, cognitive liberty, has been discussed mostly in connection to neuroscience and neurotechnology.[199] My suggestion is that the concept can be relevant, in the context of this thesis, not only in connection to neurotechnology but also to the other kinds of AI that I have discussed in this thesis, which can affect, modulate or manipulate human cognition. Thus, similar arguments that have been put forward concerning neurotechnology can also be put forward concerning these other technologies, since these too, as I have argued in section 2.3, are affecting human cognition in new ways.

4.1 Human rights in the contemporary context of artificial intelligence influencing human life or posthuman rights?

I am in this thesis exploring what I call human rights in the contemporary context of AI influencing human life. I suggest that human rights can be rethought in the light of a posthuman understanding of how the relevant kinds of AI affect human cognition.

Others have discussed, what has been called, posthuman rights.[200] Posthuman rights have, in the Posthuman Glossary, been described as "[s]uch a praxis of rights as posthuman celebrates the creation of a new thinking and praxis of rights, one which is cut loose from the ordered and majoritarian thinking of rights in neoliberal modernity. Posthuman rights embody the claims of transversal assemblages of individuals who do not see a binary cut between thought and action, life and death, environment and humanity, or animality and humanity".[201] Posthuman rights can be and have been criticized. It has been said that individuals understood according to a posthuman scheme do not share a common ontology and that they might, therefore, not share a

---

[198] Ienca, Marcello & Roberto Andorno, "Towards New Human Rights in the Age of Neuroscience and Neurotechnology.", p. 10-11.
[199] C.f., e.g. Wrye, Sententia, "Neuroethical Considerations: Cognitive Liberty and Converging Technologies for Improving Human Cognition.", In its whole. & C.f., e.g. Ienca, Marcello & Roberto Andorno, "Towards New Human Rights in the Age of Neuroscience and Neurotechnology.", Passim.
[200] See, e.g. Hanafin, Patrick, "Posthuman rights, a micropolitics of", in Braidotti, Rosi, & Hlavajova, Maria, (ed.) Posthuman Glossary (London: Bloomsbury Publishing Plc, 2018) [Electronic] Available: ProQuest Ebook Central, https://ebookcentral-proquest-com.ezproxy.ub.gu.se/lib/gu/reader.action?docID=5226228 (Last accessed: 1/12-2019), p.352-355.
[201] Hanafin, Patrick, "Posthuman rights, a micropolitics of", p. 354-355 & C.f. Braidotti, Rosi, The Posthuman, p. 8 & 167.

common set of human rights. Posthuman rights embody the claims of assemblages of individuals who do not see a binary cut between thought and action, life and death, environment and humanity, or animality and humanity, assemblages of individuals which have been described as not sharing a common ontology. Ideas connecting to posthuman understandings and posthuman rights have also been claimed to be a threat to democracy which depends on shared biological and cultural foundations as a ground for human equality.

The introduction of AI that is affecting human cognition can, as discussed in section 2.3, be seen as altering human vulnerability and could, therefore, possibly be seen as posing a risk to the condition comprising a shared common set of human rights. With regards to the problem, of protection of human cognition in a time where AI is affecting human life, posed in this thesis, existential risks can, therefore, arise from the technological advancements. AI can be described as altering human vulnerability and could, therefore, also be described as challenging the idea of a shared human ontology.

As discussed in section 2.3, Hayles argues that there are narratives concerning AI and humans that express a fear of loss of humanity and loss of control, but also fear of dissolution of the human self. [202] She, further, argues that by disentangling assumptions about the human as an independent entity, the possibility for the human to live in close connection with other life forms renders available.[203] In light of this idea, I suggest that understanding human cognition as entangled with AI, renders such a possibility available. Human cognition, a human function that is thought to have evolved as a way of controlling action, must, be seen as important for human beings in general and could, therefore, be considered to be common. Since human cognition could be considered to be common, I suggest that acknowledging that there is a convergence of human cognition and AI instead of a dissolution of the human self, can help maintain an understanding of a common human ontology and, therefore, help manage existential risks that can arise from the technological advancements. I, therefore, suggest that the altered human vulnerability that emerges when viewing human cognition as entangled with the relevant kinds of AI, as discussed in section 2.3, does not pose a risk to the condition of a shared common set of human rights. I do rather, suggest that it can be possible to manage existential risks that can arise from technological advancements, with regards to the problem

[202] Hayles, N. Katherine, How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics, p. 235-240 & See also Pisters, Patricia, “Body without organs”, p. 75-76.
[203] Hayles, N. Katherine, How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics, p. 235-240 & See also Pisters, Patricia, “Body without organs”, p. 75-76.

posed in this thesis, by acknowledging that there is a convergence of human cognition and AI. I, therefore, suggest that rethinking human rights in light of a posthuman understanding of how the relevant kinds of AI affect human cognition could help manage such risks. By rethinking human rights in this context my suggestion is, therefore, that human rights can maintain a desirable function in human society.

4.2 International human rights in the contemporary context of artificial intelligence influencing human life

Radical developments in technology put pressure on the legal system to accommodate new knowledge about how human cognition works.[204] As Wrye Sententia, puts it:

> What new forms of law will be necessary to cope with rapid advances in cognitive intervention and monitoring? What common concepts will require re-evaluation based on new models in our understanding of the brain and its functional properties? How will these emerging technologies, with an enhanced capacity to monitor and control cognitive function, be restricted or applied? How will the law cope with discoveries and revelations from brain science that call for a revision of some of its most basic core assumptions of human autonomy and freedom? These questions, and others like them, point to areas by which our shared cultural systems may be woefully unprepared to incorporate coming revolutions in new material ways of thinking about thinking.[205]

In, for example, criminal law, attention is being devoted to neurotechnological applications. However, little focus has been directed to the implications for human rights law.[206] This is applicable to the dimensions of AI discussed in this text in general. The development and deployment of AI and its impact does not stop at national borders. Hence, global solutions are required for the global challenges and risks posed by AI.[207] The implications for human rights law is, therefore, of particular interest, since the universal nature of human rights law could be suitable for providing a solid foundation for the emerging jurisprudence of technology.[208]

Moreover, in the development of AI, ethical codes for this development have been presented. These codes do often advocate respect for human rights. To give some examples, the future of life institute, an organization working to ensure that AI is developed in a way that is beneficial

[204] C.f. Wrye, Sententia, "Neuroethical Considerations: Cognitive Liberty and Converging Technologies for Improving Human Cognition.", p. 226.
[205] Wrye, Sententia, "Neuroethical Considerations: Cognitive Liberty and Converging Technologies for Improving Human Cognition.", p. 226.
[206] Ienca, Marcello & Roberto Andorno, "Towards New Human Rights in the Age of Neuroscience and Neurotechnology.", p. 7.
[207] The EC High-Level Expert Group on Artificial Intelligence, AI HLEG, 2019, "Ethics Guidance for Trustworthy AI", p. 5
[208] C.f. Ienca, Marcello & Roberto Andorno, "Towards New Human Rights in the Age of Neuroscience and Neurotechnology.", p. 7.

to humans,[209] includes a principle of respect for fundamental human rights in their code of ethics.[210] It is formulated as follows: "AI systems should be designed and operated so as to be compatible with ideals of human dignity, rights, freedoms, and cultural diversity."[211] United Nations Educational, Scientific and Cultural Organization (UNESCO) has, in the UNESCO Courier, published that AI must be ensured to serve humanity with respect to human dignity and human rights.[212] Google has written, concerning its principles regarding AI, that AI will not be designed or deployed in the area of "[t]echnologies whose purpose contravenes widely accepted principles of international law and human rights".[213] The European Commission for the efficiency of justice has included a principle of respect for human rights in the European Ethical Charter on the use of AI in judicial systems and their environment.[214] The European Commission has also set up the High-level expert group on AI which have published ethics guidelines for trustworthy AI where it is stated that " [i]t is through Trustworthy AI that we, as European citizens, will seek to reap its benefits in a way that is aligned with our foundational values of respect for human rights, democracy and the rule of law".[215] This shows an existing focus within ethical frameworks in the field of AI, on developing AI in a way that is aligned with respect for human rights. Human rights can thus provide promising foundations for recognizing abstract ethical principles and values for the context of AI.[216]

4.2.1 Need to rethink the protection of human cognition given by the international human rights framework in the contemporary context of artificial intelligence influencing human life

The history of human rights can, as discussed above in section 3.1.1, be seen as an ongoing one. Human rights can, therefore, not be seen as definitive and should thus be rethought when

[209] Future of life institute, "Who we are", https://futureoflife.org/team/ (Last accessed: 16/12-2019).
[210] More ethical codes on artificial intelligence can be found on: Algorithmwatch, "AI Ethics Guidelines Global Inventory", https://algorithmwatch.org/en/project/ai-ethics-guidelines-global-inventory/ (Last accessed: 11/12-2019).
[211] Future of life institute, "Asilomar AI Principles", https://futureoflife.org/ai-principles/?cn-reloaded=1 (Last accessed: 24/10-2019).
[212] UNESCO. "Artificial intelligence: the promises and the threats", https://unesdoc.unesco.org/ark:/48223/pf0000265211 (Last accessed: 24/11-2019).
[213] Google, "AI at Google: our principles", https://www.blog.google/technology/ai/ai-principles/ (Last accessed: 31/10-2019).
[214] EUROPEAN COMMISSION FOR THE EFFICIENCY OF JUSTICE, "European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment", https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c (Last accessed: 24/10-2019).
[215] The EC High-Level Expert Group on Artificial Intelligence, AI HLEG, 2019, "Ethics Guidance for Trustworthy AI", p. 4.
[216] C.f. The EC High-Level Expert Group on Artificial Intelligence, AI HLEG, 2019, "Ethics Guidance for Trustworthy AI", p. 9.

needed, in a new context. Hayles suggests that "[w]hatever our future, it will almost certainly include human interventions in biological processes, which means that 'human nature' will at least in part be what humans decide it should be".[217] Similarly, AI is created by human beings which means that humans will, at least in part, be part in deciding the future of humans and AI. Hayles states that "[i]t is likely that our future will be increasingly entwined with intelligent machines, but this only deepens and extends the necessity for principled debate, for their futures too cannot be envisioned apart from the primary concern for ethics that should drive these discussions".[218] As discussed in section 2.3, AI is affecting human cognition and it can be argued that human cognition is increasingly entwined with AI. This does, in line with Hayles's statement, deepen and extend the necessity for an initiated debate, since the future of AI cannot be envisioned as separated from ethical concerns. Discussions about the future of human rights, AI and human cognition can, in line with these thoughts, be argued to be necessary.

As I have discussed in section 2.3, AI is affecting human cognition in ways that have not been possible before the introduction of AI. The evolved way of processing information is, in the contemporary context of artificial intelligence influencing human life, as discussed in section 2.3, vulnerable to AI in personalization algorithms and augmented reality technology filtering information and to the possibility of revealing the content of cognition without the human's consent. Since human cognition is, as presented in section 2.1, thought to have evolved as a way of controlling action and, thus, as a way of processing information, so that the human being can respond to what is perceived,[219] protection of human cognition must be seen as necessary for humans to be agents. Protection of human cognition must, therefore, be seen as a common human interest that is vulnerable and necessary for humans to be agents. Since the protection of common human interests that are vulnerable and necessary for humans to be agents can be seen as common human interests that should be protected by human rights, I further suggest that international human rights need to be rethought in light of the contemporary context of artificial intelligence influencing human life.

Such a rethinking of international human rights, to protect human cognition in the contemporary context of AI influencing human life, needs to be justified in the human rights system.[220] A moral justification put forward for human rights is generally equality, or, basic moral equality.

[217] Hayles, N. Katherine, "Computing the Human.", p. 133.
[218] Hayles, N. Katherine, "Computing the Human.", p. 148.
[219] See Glass, Arnold Lewis, Cognition, p. 2.
[220] See section 3.1.2.

Human rights can be seen as one phenomenon that constitutes human beings' equal moral status. Human rights can, in line with this thought, protect interests that can give rise to entitlements that are constitutive of equal moral status.[221] The discussed protection is concerning a human function that is thought to have evolved as a way of controlling action, a function that must be seen as important for human beings in general and common to humans. Protection of human cognition as a common human interest that is vulnerable and necessary for humans to be agents could, thus, be constitutive for the idea that all persons should be seen as having the same moral worth and could, therefore, be the basis for mutual moral claim. I, therefore, suggest that protection of human cognition as a common human interest that is vulnerable and necessary for humans to be agents can give rise to entitlements that are constitutive of equal moral status. Dignity, as a part of the moral justification equality, can be formulated as the possession of personhood and the capacity for rational and moral agency.[222] Since human cognition is thought to be a function evolved to control action and a way of processing information it can be seen as crucial to the human capacity for rational and moral agency. I, therefore, suggest that protection of human cognition as a common human interest that is vulnerable and necessary for humans to be agents is necessary to protect human dignity.

4.3 Rethinking of recognized human rights in the contemporary context of artificial intelligence influencing human life

To understand if the international human rights framework is elastic enough to provide protection by rethinking of already recognized rights I will in this section, in accordance with rules of interpretation, analyse how and if the relevant recognized rights can be interpreted to accommodate protection of human cognition in the contemporary context of AI influencing human life.

As discussed in section 3.1.3, the rules of interpretation that can be found in VCLT is considered to be customary law, applicable to human rights law. The articles in VCLT are not considered to be step-by-step formulas for interpretation but rather indications of elements that should be taken into consideration.[223] The rules of interpretation that can be relevant for this context will, therefore, be taken into consideration when analysing if the recognized rights can accommodate protection of the human cognition in the contemporary context of AI influencing human life.

[221] See section 3.1.2.
[222] See section 3.1.2.
[223] Fitzmaurice, Malgosia, "Interpretation of Human Rights Treaties", p. 7.

Further, other interpretative tools that are used in the interpretation of international human rights, will be taken into consideration.

The terms in the two relevant rights are abstract concepts. With this in mind, it can be difficult to find an acceptable ordinary meaning of the right.[224] Human rights are often general, vague and subjective.[225] These rights can be described with those terms and it is therefore motivated to take an expansive attitude towards the interpretation of the right.[226] In line with this, the Human Rights Committee also has implied that the right to freedom of thought, conscience and religion is to be interpreted broadly.[227] Since it is generally acknowledged that a more expansive attitude towards interpretation is needed for human rights, I will approach the interpretation of the rights with an expansive attitude.[228]

The ordinary meaning should according to VCLT be given in the context. Further, the ordinary meaning of a term develops over time.[229] The ordinary meaning of the terms in the rights should, therefore, be interpreted in the contemporary context of AI influencing human life. Further, in accordance with the interpretative tool evolutive interpretation, which UN human rights bodies have relied upon, the rights can be interpreted in light of present-day conditions.[230] Present-day conditions are, as suggested in this thesis, altering common human vulnerability. The altered common human vulnerability should, therefore, be taken into account when interpreting the rights.

VCLT, further, provides that if it can be established that the parties intended a special meaning to a term this should be the meaning given to the term in interpretation.[231] The intentions of the parties are, therefore, an important factor for interpretation. The documents from the drafting of the ICCPR and UDHR are therefore important sources of information in this interpretation of the rights. Given the number of parties to the UDHR and the ICCPR, it can, however, be difficult to find common intentions.[232]

---

[224] C.f. Fitzmaurice, Malgosia, "Interpretation of Human Rights Treaties", p. 8.
[225] C.f. Fitzmaurice, Malgosia, "Interpretation of Human Rights Treaties", p. 1-2.
[226] Fitzmaurice, Malgosia, "Interpretation of Human Rights Treaties", p. 8.
[227] HRC, General Comment 22.
[228] See section 3.1.3 & Fitzmaurice, Malgosia, "Interpretation of Human Rights Treaties", p. 8.
[229] Fitzmaurice, Malgosia, "Interpretation of Human Rights Treaties", p. 8.
[230] Fitzmaurice, Malgosia, "Interpretation of Human Rights Treaties", p. 21 & Roger Judge v. Canada, para 10.3. Further, the CERD Committee made similar statements (Hagan v. Australia, Comm, para 7.3), as did the CAT Committees (V.X.N. and H.N. v. Sweden).
[231] Fitzmaurice, Malgosia, "Interpretation of Human Rights Treaties", p. 7.
[232] C.f. Fitzmaurice, Malgosia, "Interpretation of Human Rights Treaties", p. 7.

4.3.1 Freedom of thought, conscience and religion in the contemporary context of artificial intelligence influencing human life

Article 31(1) VCLT states that interpretation shall be made in accordance with the ordinary meaning of a provision.[233] The freedom of thought, conscience and religion is, by the wording of the provisions where this right is recognized, connected to human cognition.[234] Human cognition is, as it is defined in this thesis, mental processes, processing information, that is constituting thinking and remembering. The wording of the articles, in which the right is recognized does, therefore, entail a connection between the right to freedom of thought and human cognition. Further, the freedom of religion and belief can also be seen as connected to human cognition since having, and adopting a religion or belief includes mental processes, processing information. However, it is not clear what aspects of human cognition that is protected by the right.

The documentations from the discussions of the drafting of the provisions that recognize the right at hand, together with the wording of the provisions in their entirety, show a tendency to focus on the outward expression of thoughts and conscience. [235] However, the tendency is not strong enough to established that such meaning of the wording was the common intent of the parties. Further AI is, as discussed in section 2.3, entailing possibilities to affect human cognition in new ways. I suggest that this bring the so-called nucleus of the right into a situation where human cognition can possibly be affected in other ways than imagined at the time of the drafting of the right. I therefore argue that the so-called nucleus of the right can no longer be seen as so unproblematic as it was seen to be at the time of the drafting. A focus on the outward expression of thoughts and conscience can, therefore, reasonably not be given to the provision when interpreted in light of present-day conditions. The freedom of thought, conscience and religion cannot be derogated from and have been described as fundamental.[236] This, further, suggests that the right to freedom of thought, conscience and religion should enjoy broad protection.

---

[233] Vienna Convention on the Law of Treaties.
[234] International Covenant on Civil and Political Rights, Article 18 & Universal Declaration of Human Rights, Article 18.
[235] See GA, Third Committee, Third Session, 127th Meeting, at 395.
[236] International Covenant on Civil and Political Rights, Article 18(3), the inclusion of possible limitations of the freedoms in question in Universal Declaration of Human Rights, Article 29, Scheinin, Martin, "Freedom of Thought, Conscience and Religion.", p. 7 & UN Secretary-General, Annotations on the Text of the Draft International Covenants on Human Rights, p. 48.

It has been suggested that the focus of the right is the freedom of religion. As discussed in section 3.2.1, it has, though, also been argued that the different parts of the right should be given equal weight. The right should, thus, not be interpreted to provide weaker protection concerning freedom of thought and conscience.

The ICCPR provides a prohibition of coercion which would impair freedom to maintain or change religion or belief.[237] The term coercion has been interpreted as not "applying to moral or intellectual persuasion".[238] Since the boundary between coercion and persuasion is not further defined it is, however, difficult to decide if the ways in which AI is affecting human cognition could be seen as coercion prohibited by the right or persuasion interpreted to be in accordance with the right. Interpreted in light of present-day conditions in which human beings, as discussed in section 2.3.2, can be seen as vulnerable to the contemporary possibility to constantly and unconsciously reveal the content of cognition, without the consent of the human being, some of the ways in which AI is affecting human cognition could perhaps be argued to be coercion. This interpretation of the right appears, however, to be rather far-fetched and as stretching the interpretation process a bit too far. Further, such protection against coercion is only recognized concerning religion and belief and not concerning thought and conscience, even though it has been stated that the term belief shall be interpreted broadly.[239]

Moreover, the ordinary meaning should be given in light of the object and purpose of the provisions. The specific nature of human rights may, in fact, lead to enhanced importance of the object and the purpose of the treaty while the importance of the text might be reduced.[240] A teleological approach to interpretation has in line with this been adopted in human rights law. Hence, much weight is to be given to preambles of treaties and declarations. The preamble of the UDHR states that the declaration should be a common standard for all peoples and nations.[241] It further proclaims that "the inherent dignity and of the equal and inalienable rights of all members of the human family is the foundation of freedom, justice and peace in the world".[242] The preamble of the ICCPR is, also, mentioning dignity and equality.[243] This implies that the purpose and the object of the documents are dignity and

[237] International Covenant on Civil and Political Rights, Article 18.
[238] UN Secretary-General, Annotations on the Text of the Draft International Covenants on Human Rights, p. 48.
[239] See. HRC, General Comment 22.
[240] Fitzmaurice, Malgosia, "Interpretation of Human Rights Treaties", p. 3.
[241] Universal Declaration of Human Rights, Preamble.
[242] Universal Declaration of Human Rights, Preamble.
[243] International Covenant on Civil and Political Rights.

equality. The object and purpose of the documents can thus be argued to be, to foster freedom, justice, and peace through protecting human dignity and equal rights of all human beings. Since the discussed protection is regarding a human function that is thought to have evolved as a way of controlling action, a function that must be seen as important for human beings in general, protection of human cognition as a common human interest that is vulnerable and necessary for humans to be agents could be constitutive for the idea that all persons should be seen as having the same moral worth and could, therefore, be the basis for mutual moral claim. I, therefore, suggest that protection of human cognition as a common human interest that is vulnerable and necessary for humans to be agents can give rise to entitlements that are constitutive of equal moral status. Further, since human cognition is thought to be a function evolved to control action and a way of processing information it can be seen as crucial to the human capacity for rational and moral agency. I, therefore, suggest that protection of human cognition as a common human interest that is vulnerable and necessary for humans to be agents is necessary to protect human dignity. The right could, therefore, in light of the object and purpose of the documents be interpreted to protect human cognition in the contemporary context of artificial intelligence influencing human life.

When considering the so-called 'pro homine' or human-centred approach to human rights, it can be argued that it is central to protect human cognition against non-humans such as AI.

Since the Human Rights Committee has pointed out that the right to thought and religion includes protection against being compelled to reveal one's thoughts and/or adherence to a certain religion or belief, this right can be interpreted to accommodate protection of some parts of human cognition as a common human interest that is vulnerable and necessary for humans to be agents, such as protection against AI in neurotechnology that can reveal content of human cognition.[244]

Based on these analyses this human right can possibly be interpreted to accommodate protection of some parts of human cognition. I do, however, suggest that the wording of the right to freedom of thought, conscience and religion implies that the right is protecting the *what* in cognition and not the *how.* It is the thought, conscience and religion that is protected and, for example, not *how* to think. The right does, however, include that no one shall be subject to coercion which would impair freedom to have or adopt religion or belief. This part of the right

[244] C.f. HRC, General Comment 22.

can be seen as protecting the *how* in cognition. This part of the provision mentions religion and belief but not conscience or thought. The term belief shall be interpreted broadly, but the fact that belief and thought are separated does, nevertheless, imply that there is a difference between the two concepts. This possible protection of the *how* can therefore not be interpreted to include thought and conscience.

When AI is affecting human cognition, it is affecting *how* human beings think, perceive, remember and understand, it is affecting the *how* in human cognition. Based on these reflections and analyses, I argue that even though it is motivated to interpret the right broadly and in light of the contemporary context of AI influencing human life, an interpretation that would allow the right to accommodate adequate protection of human cognition in this context is, in line with the rules of interpretation, not achievable.

4.3.2 Freedom of opinion and expression in the contemporary context of artificial intelligence influencing human life

Article 31(1) VCLT states that interpretation shall be made in accordance with the ordinary meaning of a provision.[245] Human cognition is, as it is defined in this thesis, mental processes that are constituting thinking and remembering. Human beings have different opinions in all sorts of questions and this right has been described as a right for people to think what they like.[246] Thus, the wording of the articles, in which the right is recognized entails a connection between the right and human cognition.[247] However, it is not clear what aspects of human cognition is protected.

Further, the right includes a right to inform oneself, to be informed and to access information.[248] Since human cognition, as it is defined in this thesis, is mental processes, processing information, and since I suggest that humans are vulnerable, when AI in personalization algorithms and augmented reality technology is filtering information, this part of the right could be protecting parts of human cognition in the contemporary context of AI influencing human life. However, the right to information oneself is, in ICCPR, recognized as included in the freedom of expression and not in the freedom of opinion.[249] This could be seen as implying that

[245] Vienna Convention on the Law of Treaties.
[246] McGoldrick, Dominic, "Thought, Expression, Association, and Assembly", p. 217-219.
[247] International Covenant on Civil and Political Rights, Article 19 & Universal Declaration of Human Rights, Article 19.
[248] McGoldrick, Dominic, "Thought, Expression, Association, and Assembly", p. 217-219.
[249] International Covenant on Civil and Political Rights, Article 19.

the freedom to inform oneself is protected as a part of the freedom to expression and not as a part of the freedom to opinion.

The UN statement which is discussed in section 3.2.2, indicates that a scenario, where "control" over the "mind" of a human being is possible, has not been considered by the UN. This could suggest that the common intent of the parties is not to give the provision the meaning that it should protect the human cognition from "control".[250] However, when interpreting the provision in light of present-day conditions, in which AI has enabled new ways of affecting human cognition it can be argued that protection of human cognition from being affected in certain ways that could be categorized as "control" could be seen as included in the right.

Further, the discussions seen in documents from the drafting of the ICCPR seem to focus on the right to freedom of expression.[251] It is, however, not clear if the reason for this was that freedom of expression was seen as more important or, as more in need of discussion for clarification. Therefore, any common intent, to give the wording of the provision any special meaning, in this regard, cannot be established. Moreover, the ICCPR treats freedom of opinion separately from freedom of expression.[252] Freedom of opinion is therefore, clearly, protected independently from freedom of expression. This separation shows that, what was called, the realm of the mind in the drafting of the convention should be protected, independent of what was called the public matter of outward expression.

Based on these analyses this human right can possibly be interpreted to accommodate protection of some parts of human cognition. The wording of the right does, however, show that it is the holding of an opinion that is protected. This implies that it is not the mental process of acquiring or changing an opinion that is protected. It is, therefore, reasonable to interpret the right as protecting the *what* in opinion but not the *how*. When AI is affecting human cognition, it is affecting *how* human beings think, perceive, remember and understand, it is affecting the *how* in human cognition. I do, therefore, argue that even though it is motivated to interpret the right broadly, an interpretation that would allow the right to accommodate adequate protection of human cognition in the contemporary context of AI influencing human life is, in line with the rules of interpretation, not achievable.

[250] United Nations, "Manual on Human Rights for Judges, Prosecutors and Lawyers".
[251] See discussion section 3.2.2.
[252] International Covenant on Civil and Political Rights & UN Secretary-General, Annotations on the Text of the Draft International Covenants on Human Rights, p. 51.

The right does, however, also include a right to inform oneself. The right could, therefore, be seen as protecting some relevant aspects of human cognition. The wording of that part of the right implies that it is the right to seek, receive and impart information that is protected. This implies that it is not access to ways of seeking, receiving and imparting of information that is protected. It is, therefore, also in this part, reasonable to interpret the right as protecting the *what*, but not the *how* in cognition. I, therefore, suggest that this part of the right can only protect part of human cognition in the contemporary context of AI influencing human life.

4.3.3 Protection of human cognition in the contemporary context of artificial intelligence influencing human life by recognized human rights

The terms thought, belief and opinion and the question of how they relate to each other have been discussed, for example, at the drafting of the ICCPR.[253] However, there are, to my knowledge, no clarifications of any common intent of how the terms relate to each other. In accordance with the interpretative tool object and purpose test, the rights can be interpreted expansively to ensure that the provisions of the treaty have an independent or non-superfluous meaning. [254] When interpreting the terms to ensure that the provisions of the documents have an independent and non-superfluous meaning the terms can, together with the two terms religion and belief, protect a broad spectrum of human cognition.

Even though the recognized rights can be interpreted to accommodate protection of parts of the human cognition in the contemporary context of AI influencing human life, I suggest that the wording of the rights imply that they cannot be interpreted so they can accommodate the *how* in cognition, such as the freedom of *how* to think. When AI is affecting human cognition, it is affecting *how* human beings think, perceive, remember and understand, it is affecting the *how* in human cognition. I do, therefore, suggest that the international human rights framework cannot be seen as elastic enough to provide protection by rethinking of already recognized rights.

[253] Scheinin, Martin, "Freedom of Thought, Conscience and Religion.", p. 5-6 & 16.
[254] Fitzmaurice, Malgosia, "Interpretation of Human Rights Treaties", p. 10.

4.4 Recognition of new human rights in the contemporary context of artificial intelligence influencing human life

When discussing recognition of new human rights, it is relevant to once again point out that the history of human rights can be seen as ongoing.[255] Human rights can thus not be seen as definitive, and recognition of new rights can, therefore, be made when it is necessary.

When new rights are considered for recognition, an objection commonly raised is that it can lead to rights inflation, which has been described as the tendency to label everything "that is morally desirable as a 'human right'".[256]

The international law scholar Philip Alston has formulated a justificatory test to manage this risk of rights inflation. He framed a list of criteria, that a given claim should satisfy in order to qualify as an international human right, the proposed new human right should, "reflect a fundamentally important social value"[257]; "be relevant, inevitably to varying degrees, throughout a world of diverse value systems"[258]; "be eligible for recognition on the grounds that it is an interpretation of UN charter obligations"[259]; "be consistent, but not merely repetitive, of the existing body of international human rights law"[260]; "be capable of achieving a very high degree of international consensus"[261]; "be compatible or at least not clearly

[255] See. Section 3.3.1.
[256] Ienca, Marcello & Roberto Andorno, "Towards New Human Rights in the Age of Neuroscience and Neurotechnology.", p. 9.
[257] Alston, Philip. "Conjuring up New Human Rights: A Proposal for Quality Control." American Journal of International Law, vol. 78, no. 3, 1984, p. 607 & See also Ienca, Marcello & Roberto Andorno, "Towards New Human Rights in the Age of Neuroscience and Neurotechnology.", p. 9.
[258] Alston, Philip. "Conjuring up New Human Rights: A Proposal for Quality Control." p. 607 & See also Ienca, Marcello & Roberto Andorno, "Towards New Human Rights in the Age of Neuroscience and Neurotechnology.", p. 9.
[259] Alston, Philip. "Conjuring up New Human Rights: A Proposal for Quality Control.", p. 607 & See also Ienca, Marcello & Roberto Andorno, "Towards New Human Rights in the Age of Neuroscience and Neurotechnology.", p. 9.
[260] Alston, Philip. "Conjuring up New Human Rights: A Proposal for Quality Control.", p. 607 & See also Ienca, Marcello & Roberto Andorno, "Towards New Human Rights in the Age of Neuroscience and Neurotechnology.", p. 9.
[261] Alston, Philip. "Conjuring up New Human Rights: A Proposal for Quality Control.", p. 607 & See also Ienca, Marcello & Roberto Andorno, "Towards New Human Rights in the Age of Neuroscience and Neurotechnology.", p. 9.

incompatible with the general practice of states"[262]; and 'be sufficiently precise as to give rise to identifiable rights and obligations".[263]

Another justificatory test that has been proposed is formulated as follows;

> it could be required that a proposed human right not only deal with some very important good but also respond to a common and serious threat to that good, impose burdens on the addressees that are justifiable and no larger than necessary, and be feasible in most of the world's countries.[264]

To understand if the suggested new right, cognitive liberty, could lead to rights inflation or if it can be justified as a new right in the international human rights framework, I am, in the following section 4.4.1, performing these two tests.

4.4.1 Cognitive liberty as a new right in the contemporary context of artificial intelligence influencing human life

According to the first justificatory test, the following criteria need to be met for the right to be justified:

1. it should reflect a fundamentally important social value
2. it should be relevant, inevitably to varying degrees, throughout a world of diverse value systems
3. it should be eligible for recognition on the grounds that it is an interpretation of UN charter obligations, a reflection of customary law rules or a formulation that is declaratory of general principles of law
4. it should be consistent, but not merely repetitive, of the existing body of international human rights law
5. it should be capable of achieving a very high degree of international consensus
6. it should be compatible or at least not clearly incompatible with the general practice of states

[262] Alston, Philip. "Conjuring up New Human Rights: A Proposal for Quality Control.", p. 607 & See also Ienca, Marcello & Roberto Andorno, "Towards New Human Rights in the Age of Neuroscience and Neurotechnology.", p. 9.
[263] Alston, Philip. "Conjuring up New Human Rights: A Proposal for Quality Control.", p. 607 & See also Ienca, Marcello & Roberto Andorno, "Towards New Human Rights in the Age of Neuroscience and Neurotechnology.", p. 9.
[264] Nickel, J., "Human Rights" & See also Ienca, Marcello & Roberto Andorno, "Towards New Human Rights in the Age of Neuroscience and Neurotechnology.", p. 9.

7. it should be sufficiently precise as to give rise to identifiable rights and obligations.[265]

Regarding the first criteria, it has been argued that;

> cognitive life, although in various forms and degrees, is inherent in all human beings, cognitive liberty is consistent with a definition of human rights as inalienable fundamentals rights 'to which a person is inherently entitled simply because she or he is a human being'[…] regardless of their nation, location, language, religion, ethnic origin or any other status.[266]

Since human cognition, in the contemporary context of AI influencing human life, is exposed to new ways of being affected, I suggest, as discussed in section 2.3 and 4.2.1, that protection of human cognition must be seen as a common human interest that is vulnerable and necessary for humans to be agents. Protecting common human interests that are vulnerable and necessary for humans to be agents can, as discussed in section 3.1.2, be seen as a common human interest that should be protected by human rights. I do, thus, suggest that protection of human cognition in the contemporary context of AI influencing human life can be seen as a fundamentally important social value.

Regarding the second criteria, it can be argued that, since human cognition is thought to be a function evolved to control action and process information, it is an important function to human being's in general. Since human cognition is exposed to new ways of being affected, I suggest, as discussed in section 2.3 and 4.2.1, that protection of human cognition must be seen as a common human interest that is vulnerable and necessary for humans to be agents. Protection of human cognition could, thus, be relevant throughout a world of diverse value systems.

Regarding the third criteria, it can be argued that the right can be seen as an interpretation of UN charter obligations. Article 2(2) in the UN charter states that the organization and the members shall act in accordance with the principle that "[a]ll Members shall settle their international disputes by peaceful means in such a manner that international peace and security, and justice, are not endangered".[267] This, in pursuit of the purposes stated in Article 1. One of which states that the purposes of the United Nations are "[t]o develop friendly relations among nations based on respect for the principle of equal rights and self-determination of peoples, and

[265] Alston, Philip. "Conjuring up New Human Rights: A Proposal for Quality Control.", p. 607 & See also Ienca, Marcello & Roberto Andorno, "Towards New Human Rights in the Age of Neuroscience and Neurotechnology.", p. 9.
[266] Ienca, Marcello & Roberto Andorno, "Towards New Human Rights in the Age of Neuroscience and Neurotechnology.", p. 11.
[267] United Nations, Charter of the United Nations, 24 October 1945, 1 UNTS XVI.

to take other appropriate measures to strengthen universal peace".[268] A UN charter obligation can, therefore, be argued to be for the organization and the members to respect the inherent dignity and the equal and inalienable rights of all members of the human family, to preserve international peace and security, and justice. This obligation can also be seen as reflected in the preamble of the UDHR, in the statement "the inherent dignity and of the equal and inalienable rights of all members of the human family is the foundation of freedom, justice and peace in the world".[269] Since I, as discussed in section 4.2.1, suggest that protection of cognitive liberty can be justified by the justifications equality and dignity, I further suggest that the right can be seen as an interpretation of UN charter obligations.

Regarding the fourth criteria, it has been argued that "[b]eing the neurocognitive substrate of all other liberties, cognitive liberty cannot be reduced to existing rights, hence [it] is immune to the risk of rights inflation".[270] It has, further, been argued that "its integration into the human rights framework would enable the protection of constitutive features of human beings that are not being entirely protected by existing rights".[271] In the discussions in section 4.3, regarding already recognized rights, I suggest that the recognized rights cannot be interpreted as accommodating protection of the *how* in human cognition. Since cognitive liberty, as a new human right, can be discussed as the right and freedom to control one's consciousness and electrochemical thought processes, the new right could protect the *how* in human cognition. It can, therefore, not be seen as merely repetitive of existing rights. The nature of cognitive liberty can, further, be seen as consistent with the existing body of international human rights law, since it would be protecting what can be argued to be a common human interest that is vulnerable and necessary for humans to be agents.

Regarding the fifth criteria, it is difficult to confidently decide how an international consensus would emerge. However, the fact that cognition is a function that can be seen as important for human beings in general, and that has been described as inherent in all human beings regardless of nation or location, religion or ethnicity makes it possible that protection of cognitive liberty would be capable of drawing international consensus.[272] However, the described common

[268] United Nations, Charter of the United Nations, 24 October 1945
[269] Universal Declaration of Human Rights.
[270] Ienca, Marcello & Roberto Andorno, "Towards New Human Rights in the Age of Neuroscience and Neurotechnology.", p. 11.
[271] Ienca, Marcello & Roberto Andorno, "Towards New Human Rights in the Age of Neuroscience and Neurotechnology.", p. 11.
[272] C.f. Ienca, Marcello & Roberto Andorno, "Towards New Human Rights in the Age of Neuroscience and Neurotechnology.", p. 11.

human vulnerability is an issue that can be politically and religiously controversial, since the so-called inner sphere of a human being can be and has been seen as outside the scope of legal regulation. This implies that it might be difficult to achieve a high degree of international consensus. Nonetheless, when studying ethical codes developed to handle AI, by different actors from different locations in the world, some kind of international consensus regarding the view that AI needs to be regulated in terms of how it is affecting human beings and human dignity can be seen.[273] It is, therefore, possible that a high degree of international consensus could be achieved.

Regarding the sixth criteria, it can be argued that, since the international community has agreed upon protecting human thought and opinion, that as discussed in section 3.2.1, 3.2.2, 4.3.1 and 4.3.2 can be connected to human cognition, it cannot be seen as incompatible with general practices of states to protect human cognition to some degree. Further, the possibility to affect the *how* in human cognition can be seen as a new phenomenon.[274] It is, therefore, not likely that general practices of states have developed in such a way that protection of this part of human cognition would be incompatible with the general practice of states.

Regarding the last criteria, it can be argued that cognitive liberty could give rise to an identifiable right to choose *how* to think. To be able to choose *how* to think it is necessary to be informed about possible ways of *how* to think and to not be forced into certain ways of *how* to think. This could, therefore, give rise to an identifiable obligation to inform about deployment of certain kinds of AI, in adequate ways, and in some instances to refrain from deploying certain kinds of AI. The suggested dimensions of the right, "(i) the liberty to change one's mind or to choose whether and by which means to change one's mind; (ii) the protection of interventions into other minds to protect mental integrity, and (iii) the ethical and legal obligation to

[273] List of ethical codes: Algorithmwatch, "AI Ethics Guidelines Global Inventory" & e.g. The EC High-Level Expert Group on Artificial Intelligence, AI HLEG, 2019, "Ethics Guidance for Trustworthy AI", p. 4 & Laskai, Lorand & Webster, Graham, "Translation: Chinese Expert Group Offers 'Governance Principles' for 'Responsible AI'", Published: 17/6- 2019 https://www.newamerica.org/cybersecurity-initiative/digichina/blog/translation-chinese-expert-group-offers-governance-principles-responsible-ai/ (Last accessed: 2/12-2019) & National Research Council Canada, "Advisory statement on human ethics in artificial intelligence and big data research", 2017, https://nrc.canada.ca/en/corporate/values-ethics/research-involving-human-participants/advisory-statement-human-ethics-artificial-intelligence-big-data-research-2017 (Last accessed: 24/11-2019) & OECD, "Recommendation of the Council on Artificial Intelligence", 22/5-2019, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449 (Last accessed: 2/12-2019).

[274] See discussion in section 2.3.

promoting cognitive liberty", [275] further shows identifiable rights that can be derived from the right.

According to the second justificatory test, the following criteria need to be met for a right to be justified:

1. it should deal with some very important good
2. it should respond to a common and serious threat to that good
3. it should impose burdens on the addressees that are justifiable and no larger than necessary
4. it should be feasible in most of the world´s countries.[276]

Regarding the first criteria, it can be argued that, since human cognition is thought to have evolved to control action and process information it is fundamental for the function of the human being. Protecting the human cognition can, therefore, be seen to be a very important good.[277]

Regarding the second criteria, I suggest, as discussed in section 2.3.2, that since AI can reveal parts of the content of human cognition without the human's consent and is filtering information, the considered kinds of AI affect human cognition in new ways that makes human cognition vulnerable. This implies that there is a serious threat to human cognition. Since human cognition is an important function to humans in general, [278] this threat is common to humans.

Regarding the third criteria, human cognition is thought to have evolved to control action and process information and is thus fundamental for the function of the human being.[279] This important function is, as discussed in section 2.3.2, vulnerable to AI in personalization algorithms and augmented reality technology filtering information and to the possibility of revealing the content of cognition without the human's consent. It can, thus, be argued that it

[275] Ienca, Marcello & Roberto Andorno, "Towards New Human Rights in the Age of Neuroscience and Neurotechnology.", p. 10-11.
[276] Ienca, Marcello & Roberto Andorno, "Towards New Human Rights in the Age of Neuroscience and Neurotechnology.", p. 9.
[277] See. Section 2.2.
[278] See Glass, Arnold Lewis, Cognition, p. 2.
[279] See Glass, Arnold Lewis, Cognition, p. 2.

can be justifiable to impose the burden to inform of and possibly, in some instances, to refrain from the deployment of certain kinds of AI, that affects human cognition in certain ways.

Regarding the last criteria, I suggest, as discussed in section 4.2.1, that protection of human cognition can be seen as a common human interest that is vulnerable and necessary for humans to be agents. Human cognition is, further, a function that should be seen as important for human beings in general, and that has been described as inherent in all human beings, regardless of nation or location, religion or ethnicity.[280] Protection of human cognition in the contemporary context of AI influencing human life can, therefore, be relevant throughout a world of diverse value systems.

These justificatory tests show that it could be possible to justify cognitive liberty as a new human right and that the international human rights framework could, therefore, possibly be elastic enough to provide protection of human cognition in the contemporary context of AI influencing human life through recognition of this new human right.

[280] Ienca, Marcello & Roberto Andorno, "Towards New Human Rights in the Age of Neuroscience and Neurotechnology.", p. 11.

5. Further and Summarizing Analysis and Conclusions

In the contemporary context of artificial intelligence influencing human life, human cognition is connected with AI. As Hayles argues, there are narratives concerning AI and humans that express fear of loss of humanity and fear of dissolution of the human self.[281] By disentangling assumptions about the human as an independent entity, the possibility for the human to live in close connection with other life forms renders available. [282] In light of this idea, I suggest that by understanding human cognition as entangled with AI, such a possibility can be made available. Such an understanding can help us understand human cognition in the contemporary context of artificial intelligence influencing human life. Human cognition is common to humans and understanding human cognition in this new context can, thus, maintain an understanding of humans, entangled with AI instead of dissolved.

To understand humans as entangled with AI it is helpful to understand what AI does. To understand what AI does we can based on Latour's general descriptive rule discussed in section 2.3, imagine what other humans or other nonhumans would have to do if the relevant kinds of AI were not present. [283] Based on such analysis it can be understood that AI in personalization algorithms and augmented reality technology is shaping the flow of information and is deciding what information, the human being is presented with. It can be understood that AI in the discussed kinds of neurotechnology is revealing the content of human cognition. It can, further, be understood that it is likely that it, in some instances, does this through coercion or force. Human cognition is, thus, affected in the encounter with AI. As discussed in section 2.3, Latour argues that technologies can shape the decisions humans make.[284] As discussed in section 2.2, AI in personalization algorithms and augmented reality technologies, prioritize and present information in a particular order and the selection of information a human being gets to see is, therefore, affected. AI can thus be seen as shaping a human being's perception, experience, existence, and action. [285] It can be seen as persuading, facilitating and enabling particular human cognitive processes actions and attitudes. Since AI is also affecting which information a human

[281] Hayles, N. Katherine, How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics, p. 235-240 & See also Pisters, Patricia, "Body without organs", p. 75-76.
[282] Hayles, N. Katherine, How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics, p. 235-240 & See also Pisters, Patricia, "Body without organs", p. 75-76.
[283] Latour, Bruno, "Where are the Missing Masses? The Sociology of a Few Mundane Artefacts,", p. 155.
[284] Latour, Bruno, "Where are the Missing Masses? The Sociology of a Few Mundane Artefacts,", p.151, passim.
[285] See Noorman, Merel, "Computing and Moral Responsibility"

being doesn't get to see, it is also constraining, discouraging and inhibiting other cognitive processes.

Since human beings have been described as biologically vulnerable as embodied agents,[286] changes to embodiment must have implications for human vulnerability.[287] The medium for human information processing, and human cognition, is thought to be the nervous system that consists of the brain the peripheral nervous system, and the spinal cord.[288] When the human being and human cognition is being increasingly entangled with AI this can, therefore, alter human vulnerability. As discussed in section 2.3.2, human cognition can be seen as vulnerable to AI in personalization algorithms and augmented reality technology filtering information, and to the possibility of revealing the content of human cognition without the human's consent. Human vulnerability can be seen as altered by the condition that human cognition is now increasingly entangled with AI and exposed to these new ways of being affected. The introduction of these kinds of AI is, further, making human cognition vulnerable in new ways.

The possibilities of monitoring and affecting human cognition enabled by the introduction of AI have not been possible before this introduction. When AI is affecting human cognition in the discussed ways it is, as discussed in section 2.3.2, changing *how* human beings think, perceive, remember and understand. It is changing the *how* in human cognition.

It has been suggested that the possibility of the emergence of an altered human vulnerability is an argument against the idea that the, for human rights required, social consensus could be grounded in human vulnerability.[289] As discussed in section 4.1, I suggest that the altered human vulnerability that emerges when viewing human cognition as entangled with the relevant kinds of AI does not pose a risk to the condition of a shared common set of human rights. I do rather, suggest that it can be possible to manage existential risks that can arise from technological advancements, with regards to the problem posed in this thesis, by acknowledging that there is a convergence of human cognition and AI. I, therefore, suggest that rethinking human rights in light of a posthuman understanding of how the relevant kinds of AI affect human cognition could help manage such risks. The altered common human vulnerability

[286] C.f. Turner, Bryan S., "Sociology of Human Rights", p. 6 & 9.
[287] Turner, Bryan S., "Sociology of Human Rights", p. 6 & 9.
[288] Glass, Arnold, Cognition, p. 8.
[289] C.f. Turner, Bryan S., "Sociology of Human Rights", p. 9.

does, however, require a flexible view of human rights in which the new altered common human vulnerability can provide a basis for rethinking recognized rights or recognition of new rights.

As I have discussed in section 2.3, AI is affecting human cognition in ways that have not been possible before the introduction of AI. The evolved way of processing information is, in the contemporary context of artificial intelligence influencing human life, as discussed in section 2.3, vulnerable to AI in personalization algorithms and augmented reality technology filtering information and to the possibility of revealing the content of cognition without the human's consent. Since human cognition is, as presented in section 2.1, thought to have evolved as a way of controlling action and, thus, as a way of processing information, so that the human being can respond to what is perceived,[290] protection of human cognition must be seen as necessary for humans to be agents. Protection of human cognition must, therefore, be seen as a common human interest that is vulnerable and necessary for humans to be agents. Since the protection of common human interests that are vulnerable and necessary for humans to be agents can be seen as common human interests that should be protected by human rights, I further suggest that the international human rights framework need to be rethought so that it can accommodate protection of human cognition in the contemporary context of artificial intelligence influencing human life. Rethinking of the protection of human cognition can also, as discussed in section 4.2.1, be justified by the moral justifications equality and dignity.

The international human rights framework could possibly be elastic enough to provide protection of human cognition in the contemporary context of AI influencing human life by rethinking of already recognized rights or by recognition of a new human right.

As discussed in section 3.2, recognized rights that could be rethought to accommodate protection of human cognition as a common human interest that is vulnerable and necessary for humans to be agents is the right to freedom of thought, conscience and religion and the right to freedom of opinion and expression. Other rights could only protect parts of human cognition and could thus not offer protection sufficient to protect human cognition in the contemporary context of AI influencing human life. As to rethinking these recognized rights so that they can accommodate protection of human cognition in the contemporary context of artificial intelligence influencing human life, I suggest, as discussed in section 4.3, that even though these recognized rights can, possibly, be interpreted to accommodate protection of parts of the

[290] See Glass, Arnold Lewis, Cognition, p. 2.

human cognition in the contemporary context of AI influencing human life, I suggest that the wording of the rights imply that they cannot be interpreted so that they can accommodate the *how* in cognition, such as the freedom of *how* to think. When AI is affecting human cognition, it is affecting *how* human beings think, perceive, remember and understand, it is affecting the *how* in human cognition. I do, therefore, suggest that the international human rights framework cannot be seen as elastic enough to adapt already recognized rights to the contemporary context of AI influencing human life regarding protection of human cognition.

As discussed in section 4.4.1, I suggest that protection of human cognitive liberty can be argued to pass justificatory tests, justifying new human rights. I do, therefore, suggest that the international human rights framework can be argued to be elastic enough for recognition of cognitive liberty as a new human right. Cognitive liberty is described as a fundamental right to use the full spectrum of the mind, the right concerns the "ethics and legality of safeguarding one's own thought processes, and by necessity, one's electrochemical brain states",[291] and the concept is seen as multi-dimensional comprising the three recognized interrelated dimensions:

> (i) the liberty to change one's mind or to choose whether and by which means to change one's mind; (ii) the protection of interventions into other minds to protect mental integrity, and (iii) the ethical and legal obligation to promoting cognitive liberty.[292]

For a human being to be able to use the full spectrum of the mind it is necessary that the human is able to decide *how* to perform its mental processes. For a human being to be able to safeguard its own thought processes, and its electrochemical brain states, the human need to be able to protect *how* those processes are performed. For a human being to be able to change its mind or to choose whether and by which means to change its mind the human has to be able to decide *how* to change its mind. For a human to be protected from interventions into its mind it needs to be protected from interventions that change *how* its mental processes are performed. Cognitive liberty as a new human right could, therefore, be argued to accommodate protection of the *how* in human cognition. When AI is affecting human cognition, it is affecting *how* human beings think, perceive, remember and understand, it is affecting the *how* in human cognition. Cognitive liberty as a new human right can, thus, be argued to accommodate protection of human cognition in the contemporary context of AI influencing human life.

[291] Wrye, Sententia, "Neuroethical Considerations: Cognitive Liberty and Converging Technologies for Improving Human Cognition.".
[292] Ienca, Marcello & Roberto Andorno, "Towards New Human Rights in the Age of Neuroscience and Neurotechnology.", p. 10-11.

My summarized conclusions are, based on these analyses, that the relevant kinds of AI affect human cognition in new ways that can be argued to alter common human vulnerability and that makes human cognition vulnerable. Since protection of human cognition as a common human interest that is vulnerable and necessary for humans to be agents can be argued to be a common human interest that should be protected by human rights, the introduction of AI entails a need for protection of human cognition in the contemporary context of artificial intelligence influencing human life. Since the already recognized rights cannot, in accordance with interpretative rules, be interpreted to accommodate protection of the *how* in human cognition, the international human rights framework cannot be seen as elastic enough to provide protection of human cognition in the contemporary context of artificial intelligence influencing human life, by rethinking of already recognized rights. However, protection of cognitive liberty can possibly be justified as a new human right, and the international human rights framework might, therefore, be elastic enough to provide protection of human cognition in the contemporary context of artificial intelligence influencing human life by recognition of a new human right. I, therefore, propose that a new human right guaranteeing protection of human cognitive liberty ought to be considered by the international community.

6. Discussion

6.1 Artificial intelligence in personalization algorithms in scientific research

I have, in the research for this thesis, used ways of collecting information where personalization algorithms have sorted information.[293] This may to some extent have affected my thesis and the thesis must, therefore, be read with that in mind. I dare say that this is probably true as regards a large part of scientific studies in the contemporary context of AI influencing human life. I would therefore like to put forward that this is an important discussion to address.

Scientific research is based on an idea of transparency and since personalization algorithms are often invisible or hard to understand it is important to take measures to make visible the personalization algorithms in such researches. As discussed in section 1.3 I have taken some measures to minimize or make visible the personalization. I suggest that such measures can be a possible way of handling personalization algorithms in scientific research since it can enable a possibility of understanding how personalization algorithms have affected the research. I, therefore, want to suggest that, in the scientific contemporary context of AI influencing human life, such measures are important to consider or to use to analyse gathered information.

Moreover, it is important to bear in mind that personalization algorithms are present in the everyday life of human beings and do possibly affect a large part of a human being's conception of the world. Since personalization algorithms affect me in my everyday life, and since it is inevitable that my conception of the world is visible in my thinking, personalization algorithms may not only affect my search results and gathering of material and information but also my analysis and reflections.

[293] See section 1.3.

7. Closing reflection

As a closing reflection, I want to point out that large technology companies seem to get increasing opportunities to collect data, about more and more aspects of human beings. Since data is used in AI that is affecting human cognition, this can possibly lead to AI affecting human cognition in more and more encompassing ways. To give an example, the news that Google has bought the company Fitbit was released during the writing of this thesis.[294] Fitbit is a company that makes and sells fitness trackers and smartwatches.[295] This company can, therefore, collect massive amounts of data about human beings. The purchase could, thus, increase, Google's possibilities to collect different kinds of data about human beings. However, in an announcement, Google has stated that "Fitbit health and wellness data will not be used for Google ads".[296] Independent of this statement, this example implies increasing possibilities for big technology companies to deploy kinds of AI that are affecting human cognition.

Since this could lead to AI affecting human cognition in more and more encompassing ways addressing the problem, of protection of human cognition in a time where AI is affecting human life, in the legal discourse is more and more necessary.

[294] Google, "Helping more people with wearables: Google to acquire Fitbit", https://www.blog.google/products/hardware/agreement-with-fitbit (Last accessed: 7/11-2019) & Phelan, David, "Google Buys Fitbit For $2.1 Billion: Here's What It Means", 1/11-2019, https://www.forbes.com/sites/davidphelan/2019/11/01/google-buys-fitbit-for-21-billion-heres-what-it-means/ (Last accessed: 7/11-2019).
[295] Fitbit, https://www.fitbit.com/se/home (Last accessed: 18/12-2019).
[296] Google, "Helping more people with wearables: Google to acquire Fitbit".

8. Bibliography

8.1 Regulation, Preparatory work and Case law

Commission on Human Rights, 8th session: summary record of the 319th meeting, E/CN.4/SR.116, 117, 319

Drafting Committee, First session, 10th Meeting (1947), E/CN.4/82/Rev.l

GA, Third Committee, Third Session, Draft International Declaration of Human Rights: Recapitulation of Amendments to Article 16 of the Draft Declaration (E/800), (1948), A/C.3/289/REV.1

GA, Third Committee, Third Session, 127th Meeting (1948), A/C.3/SR.127

Hagan v. Australia, Comm. 26/2002, U.N. Doc. A/58/18, at 139 (2003)

HRC, General Comment 22. HRI/GEN/1/Rev9 (VolI)

International Covenant on Civil and Political Rights, GA Res 2200A (XXI), Treaty Series, vol. 999, p. 171, (1966)

MAB, WAT and JAYT v Canada, CCPR/C50/D570/1993 (25 April 1994)

Official Records of the General Assembly, Fifth Session, Third Committee, Draft first international covenant on human rights and measures of implementation (A/1374, A/C.3/534, A/C.3/535 and E/1681), 289th meeting. A/C.3/SR.289

Prince v South Africa, CCPR/C/91/D/1474/2006 (14 November 2007)

Roger Judge v. Canada, CCPR/C/78/D/829/1998, UN Human Rights Committee (HRC), 13 August 2003

Statute of the International Court of Justice, 18 April 1946

Summary record of the 10th meeting, held at Lake Success, New York, on Wednesday, 18 June 1947: Commission on Human Rights, Drafting Committee, 1st session, E/CN.4/AC.1/SR.10, 26

United Nations, Charter of the United Nations, 24 October 1945, 1 UNTS XVI

Universal Declaration of Human Rights, GA Res 217A (III), UNGAOR, 3rd Sess, Supp No 13, UN Doc A/810 (1948) 71

UN Secretary-General, Annotations on the Text of the Draft International Covenants on Human Rights, 1 July 1955, A/2929

Vienna Convention on the Law of Treaties, United Nations, Treaty Series, vol. 1155, p. 331, (1969)

V.X.N. and H.N. v. Sweden, CAT/C/24/D/130/1999; CAT/C/24/D/131/1999, UN Committee Against Torture (CAT), 2 September 2000

8.2 Books

Ashcraft, Mark H, Cognition. 3rd ed. (Upper Saddle River, N.J: Prentice Hall, 2002)

Bates, Ed, "History", in Moeckli, Daniel, Sangeeta Shah, Sandesh Sivakumaran, & David J. Harris, (ed.) International Human Rights Law. Third ed., (Oxford: Oxford University Press 2018)

Besson, Samantha, "Justifications", in Moeckli, Daniel, Sangeeta Shah, Sandesh Sivakumaran, & David J. Harris, (ed.) International Human Rights Law. Third ed., (Oxford: Oxford University Press 2018)

Braidotti, Rosi & Hlavajova, Maria, "Introduction", in Braidotti, Rosi, & Hlavajova, Maria, (ed.) Posthuman Glossary (London: Bloomsbury Publishing Plc, 2018) [Electronic] Available: ProQuest Ebook Central, https://ebookcentral-proquest-com.ezproxy.ub.gu.se/lib/gu/reader.action?docID=5226228 (Last accessed: 1/12-2019)

Braidotti, Rosi, "Posthuman critical theory", in Braidotti, Rosi, & Hlavajova, Maria, (ed.) Posthuman Glossary (London: Bloomsbury Publishing Plc, 2018) [Electronic] Available: ProQuest Ebook Central, https://ebookcentral-proquest-com.ezproxy.ub.gu.se/lib/gu/reader.action?docID=5226228 (Last accessed: 1/12-2019)

Braidotti, Rosi, The Posthuman (Oxford: Polity Press, 2013) [Electronic] Available: ProQuest Ebook Central, https://ebookcentral-proquest-com.ezproxy.ub.gu.se/lib/gu/detail.action?docID=1315633 (Last accessed: 1/12-2019)

Bruns, Axel, Are Filter Bubbles Real? Digital Futures Series, (Cambridge: Polity Press, 2019)

Cropley, Arthur J., Creativity in Education and Learning: A Guide for Teachers and Educators (London: Kogan Page, 2001)

Fitzmaurice, Malgosia, "Interpretation of Human Rights Treaties", in Shelton, Dinah, (ed.) The Oxford Handbook of International Human Rights Law (Oxford: Oxford University Press 2013) [Electronic] Available: Oxford university press, https://opil-ouplaw-com.ezproxy.ub.gu.se/view/10.1093/law/9780199640133.001.0001/law-9780199640133 (Last accessed: 1/12-2019)

Glass, Arnold, Cognition. (Mason, OH: Thomson Custom Pub, 2007)

Glass, Arnold Lewis, Cognition (Cambridge, UK; New York: Cambridge University Press, 2016)

Hanafin, Patrick, "Posthuman rights, a micropolitics of", in Braidotti, Rosi, & Hlavajova, Maria, (ed.) Posthuman Glossary (London: Bloomsbury Publishing Plc, 2018) [Electronic] Available: ProQuest Ebook Central, https://ebookcentral-proquest-com.ezproxy.ub.gu.se/lib/gu/reader.action?docID=5226228 (Last accessed: 1/12-2019)

Hayles, N. Katherine, How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics (Chicago, Ill.: Univ. of Chicago Press, 1999)

Latour, Bruno, "Where are the Missing Masses? The Sociology of a Few Mundane Artefacts," in W. Bijker & J. Law (eds.), Shaping Technology/Building Society: Studies in Socio-Technical Change, (Cambridge, Massachusetts: The MIT press, 1992) pp. 225–258 [Electronic] Available: Nottingham Trent University, http://www.bruno-latour.fr/sites/default/files/50-MISSING-MASSES-GB.pdf

McGoldrick, Dominic, "Thought, Expression, Association, and Assembly", in Moeckli, Daniel, Sangeeta Shah, Sandesh Sivakumaran, & David J. Harris, (ed.) International Human Rights Law. Third ed., (Oxford: Oxford University Press 2018)

Paasonen, Susanna, “Networked affect”, in Braidotti, Rosi, & Hlavajova, Maria, (ed.) Posthuman Glossary (London: Bloomsbury Publishing Plc, 2018) [Electronic] Available: ProQuest Ebook Central, https://ebookcentral-proquest-com.ezproxy.ub.gu.se/lib/gu/reader.action?docID=5226228 (Last accessed: 1/12-2019)

Pariser, Eli, The Filter Bubble: What the Internet Is Hiding from You (London: New York: Viking; Penguin Press, 2011)

Parisi, Luciana, “Computational turn”, in Braidotti, Rosi, & Hlavajova, Maria, (ed.) Posthuman Glossary (London: Bloomsbury Publishing Plc, 2018) [Electronic] Available: ProQuest Ebook Central, https://ebookcentral-proquest-com.ezproxy.ub.gu.se/lib/gu/reader.action?docID=5226228 (Last accessed: 1/12-2019)

Pennachin, Cassio & Goertzel, Ben, “Contemporary Approaches to Artificial General Intelligence”, In Goertzel Ben., Pennachin Cassio., (ed.) Artificial General Intelligence. Cognitive Technologies, (Springer-Verlag Berlin Heidelberg 2007) [Electronic] Available: Springer Link, https://link.springer.com/chapter/10.1007/978-3-540-68677-4_1 (Last accessed: 1/12-2019)

Pisters, Patricia, “Body without organs”, in Braidotti, Rosi, & Hlavajova, Maria, (ed.) Posthuman Glossary (London: Bloomsbury Publishing Plc, 2018) [Electronic] Available: ProQuest Ebook Central, https://ebookcentral-proquest-com.ezproxy.ub.gu.se/lib/gu/reader.action?docID=5226228 (Last accessed: 1/12-2019)

Russell, Stuart J, Norvig, Peter & Davis, Ernest, Artificial Intelligence: A Modern Approach, Global Edition (Pearson Education M.U.A., 2016) [Electronic] Available: Dawsonera, https://www-dawsonera-com.ezproxy.ub.gu.se/readonline/9781292153971 (Last accessed: 1/12-2019)

Turner, Bryan S., “Sociology of Human Rights”, in Shelton, Dinah, (ed.) The Oxford Handbook of International Human Rights Law (Oxford : Oxford University Press 2013)[Electronic] Available: Oxford university press, https://opil-ouplaw-com.ezproxy.ub.gu.se/view/10.1093/law/9780199640133.001.0001/law-9780199640133 (Last accessed: 1/12-2019)

van Duffel, Siegfried “Moral Philosophy”, in Shelton, Dinah, (ed.) The Oxford Handbook of International Human Rights Law (Oxford: Oxford University Press 2013) [Electronic] Available: Oxford university press, https://opil-ouplaw-com.ezproxy.ub.gu.se/view/10.1093/law/9780199640133.001.0001/law-9780199640133 (Last accessed: 1/12-2019)

8.3 Articles

Alston, Philip. “Conjuring up New Human Rights: A Proposal for Quality Control.” American Journal of International Law, vol. 78, no. 3, 1984, p. 607

Anumanchipalli, Gopala K, Chartier, Josh & Chang, Edward F., “Speech Synthesis from Neural Decoding of Spoken Sentences.” Nature, vol. 568, no. 7753, 2019, pp. 493–498

Billinghurst, Mark. "Augmented Reality." The SAGE Encyclopedia of the Internet. Ed. Barney Warf. Thousand Oaks,: SAGE Publications, Inc., 2018. 35-40. SAGE Knowledge. Web. (Last accessed: 1/12 2019)

Braun, Michael, Hauser, John & Urban, Glen & Liberali, G. “Website Morphing.” Marketing Science: the Marketing Journal of TIMS/ORSA, vol. 28, no. 2, 2009, pp. 202–223

Bublitz, Jan & Christoph Merkel. “Crimes Against Minds: On Mental Manipulations, Harms and a Human Right to Mental Self-Determination.” Criminal Law and Philosophy, vol. 8, no. 1, 2014, pp. 51–77

Chen, Jim X., “The Evolution of Computing: AlphaGo.” Computing in Science & Engineering, vol. 18, no. 4, 2016, pp. 4–7

Culkin, J.M., “A schoolman’s guide to Marshall McLuhan.” Saturday Review, (1967) pp. 51-53, 71-72

Curkovic, Marko., “Need for Controlling of the Filter Bubble Effect.” Science and Engineering Ethics, vol. 25, no. 1, 2019, p. 323

Dobrev, Dimiter, “A Definition of Artificial Intelligence.” ArXiv.org, 2012, pp. arXiv.org

Farwell, Lawrence, “Brain Fingerprinting: a Comprehensive Tutorial Review of Detection of Concealed Information with Event-Related Brain Potentials.” Cognitive Neurodynamics, vol. 6, no. 2, 2012, pp. 115–154

Gottron, Thomas & Felix Schwagereit. “The Impact of the Filter Bubble -- A Simulation Based Framework for Measuring Personalisation Macro Effects in Online Communities.” ArXiv.org, 2016, pp. arXiv.org

Hayles, N. Katherine, “Computing the Human.” Theory, Culture & Society, vol. 22, no. 1, 2005, pp. 131-151

Holone, Harald, “The Filter Bubble and Its Effect on Online Personal Health Information.” Croatian Medical Journal, vol. 57, no. 3, 2016, pp. 298–301

Ienca, Marcello & Roberto Andorno, “Towards New Human Rights in the Age of Neuroscience and Neurotechnology.” Life Sciences, Society and Policy, vol. 13, no. 1, 2017, pp. 1–27

Ienca, Marcello, “The Right to Cognitive Liberty.” Scientific American, vol. 317, no. 2, 2017, pp. 10

Jouhki, Jukka, Penttinen, aija, Sormanen, Niina & Uskali, Turo, “Facebook’s Emotional Contagion Experiment as a Challenge to Research Ethics.” Media and Communication, vol. 4, no. 4, 2016, pp. 75–85.

Kallioinen, Noa, Pershina, Maria, Zeiser, Jannik, Nosrat Nezami, Farbod, Stephan, Achim, Pipa, Gordon & König, Peter. “Moral Judgements on the Actions of Self-driving Cars and Human Drivers in Dilemma Situations from Different Perspectives.” OSF Preprints, (2019)

Krafft, Tobias, Michael Gamer, and Katharina Zweig, “What Did You See? Personalization, Regionalization and the Question of the Filter Bubble in Google's Search Engine.” ArXiv.org, 2018, arXiv.org

Lu, Huimin, Yujie Li, Min Chen, Hyoungseop Kim, and Seiichi Serikawa, “Brain Intelligence: Go Beyond Artificial Intelligence.” ArXiv.org, 2017, arXiv.org

Martinez, Rogelio. “The Power of Artificial Intelligence.” Franchising World, vol. 50, no. 5, 2018, pp. 92–94

Mazzuoli, Valerio de Oliveira & Dilton, Riberio, “The Pro Homine Principle as a Fundamental Aspect of International Human Rights Law/O Principio Pro Homine Como Um Aspecto Fundamental Do Direito Internacional Dos Direitos Humanos.” Meridiano 47, vol. 17, no. 47, 2016

Mcgrogan, David, “On the Interpretation of Human Rights Treaties and Subsequent Practice.” Netherlands Quarterly of Human Rights, vol. 32, no. 4, 2014, pp. 347-78

O'Hara, Kieron & Stevens, David “Echo Chambers and Online Radicalism: Assessing the Internet's Complicity in Violent Extremism.” Policy & Internet, vol. 7, no. 4, 2015, pp. 401–h422.

Rodriguez, Pablo & Blanco, Javier, “Organization and Information in Simondon's Theory of Individuation.” Culture and Organization, vol. 23, no. 1, 2017, pp. 34–43

Scheinin, Martin, “Freedom of Thought, Conscience and Religion.” Studia Theologica - Nordic Journal of Theology, vol. 54, no. 1, 2000, pp. 5–18

Wright, David, “The Dark Side of Ambient Intelligence.” Info, vol. 7, no. 6, 2005, pp. 33–51

Wrye, Sententia, “Neuroethical Considerations: Cognitive Liberty and Converging Technologies for Improving Human Cognition.” Annals of the New York Academy of Sciences, vol. 1013, no. 1, 2004, pp. 221–228

8.4 Dictionaries

A Dictionary of Computer Science 7 ed., “VPN”, Butterfield, Andrew, Ekembe Ngondi, Gerard & Kerr, Anne (ed.), (Oxford University press, 2016), https://www-oxfordreference-com.ezproxy.ub.gu.se/view/10.1093/acref/9780199688975.001.0001/acref-9780199688975-e-6258?rskey=LboWvL&result=1 (Last accessed: 2/12 2019)

A Dictionary of psychology 4 ed., “Cognition”, Colman, Andrew M. (ed.), (Oxford University press 2015), https://www-oxfordreference-com.ezproxy.ub.gu.se/view/10.1093/acref/9780199657681.001.0001/acref-9780199657681-e-1594?rskey=iftF5J&result=6 (Last accessed: 2/12 2019)

A Dictionary of Social Media, “Filter Bubble.”, Chandler, Daniel, & Munday, Rod (ed.), (Oxford University press 2016), https://www-oxfordreference-com.ezproxy.ub.gu.se/view/10.1093/acref/9780191803093.001.0001/acref-9780191803093-e-482?rskey=m5WEu0&result=1 (Last accessed: 2/12 2019)

Encyclopaedia Britannica, “Artificial intelligence”, Copeland, B.J., Britannica Online Academic Edition, 2019, https://www.britannica.com/technology/artificial-intelligence (Last accessed: 18-10-2019)

Encyclopaedia Britannica, “Augmented Reality”, Hosch, William L., Britannica Online Academic Edition, 2019, https://www.britannica.com/technology/augmented-reality (Last accessed: 18-10-2019)

Lexico powered by oxford, “Machine learning”, https://www.lexico.com/en/definition/machine_learning (Last accessed: 18-10-2019)

8.5 Internet sources

Algorithmwatch, “AI Ethics Guidelines Global Inventory”, https://algorithmwatch.org/en/project/ai-ethics-guidelines-global-inventory/ (Last accessed: 11/12-2019)

Cesar Hidalgo: “Why Information Grows” Talks at Google, 6/8-2015, https://www.youtube.com/watch?v=r38kK26SieE, (Last accessed: 26/11-2019)

Emotiv, https://www.emotiv.com, (Last accessed: 11/12-2019)

EUROPEAN COMMISSION FOR THE EFFICIENCY OF JUSTICE, “European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment”, https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c (Last accessed: 24/10-2019)

Farahany, Nita, ”Cognitive Liberty in the Era of Brain Hacking, The Aspen Institute”, The Aspen Institute, Published: 11/9-2014, https://youtu.be/8CqgZ0V5pvY (Last accessed: 5/10 2019)

Fitbit, https://www.fitbit.com/se/home (Last accessed: 18/12-2019)

Freerlogic, http://www.freerlogic.com/products/hardware (Last accessed: 23/10-2019)

Future of life institute, "Asilomar AI Principles", https://futureoflife.org/ai-principles/?cn-reloaded=1 (Last accessed: 24/10-2019)

Future of life institute, "Who we are", https://futureoflife.org/team/ (Last accessed: 16/12-2019)

Giordano, Sarah, "Popping the Filter bubble", 2014, https://derekbruff.org/blogs/fywscrypto/practical-crypto/popping-the-filter-bubble/ (Last accessed: 18/11-2019)

Google, "AI at Google: our principles", https://www.blog.google/technology/ai/ai-principles/ (Last accessed: 31/10-2019)

Google, "Glass", https://www.google.com/glass/tech-specs/ (Last accessed: 23/10-2019)

Google, "Helping more people with wearables: Google to acquire Fitbit", https://www.blog.google/products/hardware/agreement-with-fitbit (Last accessed: 7/11-2019)

Google, "Så fungerar sökalgoritmer", https://www.google.com/search/howsearchworks/algorithms/ (Last accessed: 31/10-2019)

Google, "Sökinställningar", https://www.google.com/preferences (Last accessed: 31/10-2019)

Google, "Updating Our Privacy Policies and Terms of Service," Google Official Blog, 24/1-2012, https://googleblog.blogspot.com/2012/01/updating-our-privacy-policies-and-terms.html (Last accessed: 11/12-2019)

Gray, Richard, "Minority Report-style advertising billboards to target consumers" https://www.telegraph.co.uk/technology/news/7920057/Minority-Report-style-advertising-billboards-to-target-consumers.html, (Last accessed: 23/10-2019)

Greshko, Michael & Wei-Haas, Maya, "New device translates brain activity into speech. Here's how", National Geographic, Published: April 24, 2019 https://www.nationalgeographic.com/science/2019/04/new-computer-brain-interface-translates-activity-into-speech/ (Last accessed: 7/11-2019)

Laskai, Lorand & Webster, Graham, “Translation: Chinese Expert Group Offers 'Governance Principles' for 'Responsible AI'”, Published: 17/6- 2019 https://www.newamerica.org/cybersecurity-initiative/digichina/blog/translation-chinese-expert-group-offers-governance-principles-responsible-ai/ (Last accessed: 2/12-2019)

Mercedes benz, “MBUX Augmented Reality för navigation”, https://www.mercedes-benz.se/passengercars/mercedes-benz-cars/models/eqc/comfort.pi.html/mercedes-benz-cars/models/eqc/comfort/comfort-gallery/augmented-video (Last accessed: 23/10-2019)

Mindriderdata, http://mindriderdata.com (Last accessed: 23/10-2019)

Muse, https://choosemuse.com (Last accessed: 11/12-2019)

National Research Council Canada, “Advisory statement on human ethics in artificial intelligence and big data research”, 2017, https://nrc.canada.ca/en/corporate/values-ethics/research-involving-human-participants/advisory-statement-human-ethics-artificial-intelligence-big-data-research-2017 (Last accessed: 24/11-2019)

Neurosky, http://neurosky.com, (Last accessed: 11/12-2019)

Nickel, J., ”Human Rights”, The Stanford Encyclopedia of Philosophy, Zalta, Edward N., (ed.) 2014, https://plato.stanford.edu/archives/spr2017/entries/rights-human (Last accessed: 2/12-2019)

Noorman, Merel, "Computing and Moral Responsibility", The Stanford Encyclopedia of Philosophy (Spring 2018 Edition), Zalta, Edward N., (ed.), 2012 (edited 2018), https://plato.stanford.edu/archives/spr2018/entries/computing-responsibility/ (Last accessed: 2/12-2019)

OECD, “Recommendation of the Council on Artificial Intelligence”, 22/5-2019, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449 (Last accessed: 2/12-2019)

Phelan, David, “Google Buys Fitbit For $2.1 Billion: Here’s What It Means”, 1/11-2019, https://www.forbes.com/sites/davidphelan/2019/11/01/google-buys-fitbit-for-21-billion-heres-what-it-means/ (Last accessed: 7/11-2019)

The Center for Cognitive Liberty, http://www.cognitiveliberty.org (Last accessed: 17/11-2019)

The EC High-Level Expert Group on Artificial Intelligence, AI HLEG, 2019, "Ethics Guidance for Trustworthy AI", https://ec.europa.eu/futurium/en/ai-alliance-consultation (Last accessed: 11/12-2019)

The Guardian, "Advertising billboards use facial recognition to target shoppers", 7/9-2010, https://www.theguardian.com/media/pda/2010/sep/27/advertising-billboards-facial-recognition-japan (Last accessed: 23/10-2019)

The Guardian, "Neuroscientists decode brain speech signals into written text", 30/7-2019, https://www.theguardian.com/science/2019/jul/30/neuroscientists-decode-brain-speech-signals-into-actual-sentences (Last accessed: 24/11-2019)

UNESCO. "Artificial intelligence: the promises and the threats" https://unesdoc.unesco.org/ark:/48223/pf0000265211 (Last accessed: 24/11-2019)

United Nations, "Human rights", https://www.un.org/en/sections/issues-depth/human-rights/ (Last accessed: 17/11-2019)

United Nations, "Manual on Human Rights for Judges, Prosecutors and Lawyers", Published: 2003. https://www.un.org/ruleoflaw/files/training9chapter12en.pdf (Last accessed: 24/11-2019)

Whyte, Chelsea, "Mind-reading device uses AI to turn brainwaves into audible speech", 24/4-2019, https://www.newscientist.com/article/2200683-mind-reading-device-uses-ai-to-turn-brainwaves-into-audible-speech/ (Last accessed: 24/11-2019)

United Nations, "The Foundation of International Human Rights Law", https://www.un.org/en/sections/universal-declaration/foundation-international-human-rights-law/index.html (Last accessed: 7/11-2019)