116-0002
2-20-7 Arakawa Arakawaku Tokyo Japan

ISSN 2435-3345

Robotics & AI Ethics

2021 6(1)

<Index>

J-INSTITUTE

2021 6(1) 1-6 Submission: 2021/02/10, Peer review: 2021/02/20, Accepted: 2021/02/25, Published: 2021/03/30

Robotics & AI Ethics

Publisher: J-INSTITUTE
ISSN: 2435-3345

Website: www.j-institute.jp/ai/
Editor: ai@j-institute.jp

Corresponding author
E-mail: j.choi3935@gmail.com

dx.doi.org/10.22471/ai.2021.6.1.01

Current and Future CONSIDERATIONS for the Use of Artificial Intelligence by the United States' Department of Defense

Jina Choi

Ewha Womans University, Seoul, Republic of Korea

Abstract

Purpose: *The "Narrow" AI technologies and its application to different sectors in the U.S. have brought tremendous advantages to the country and its citizens. This paper attempts to examine the opportunities and risks of AI's potential application to the military health care sector. Yet, some human rights concerns remain. This paper further argues that the new Biden administration should pursue policies that reinforce a human-centric AI to be more supportive of human rights.*

Method: *This paper employs the politics of responsibility theory to examine the present and future of AI responsible governance in the Biden era. Based on resources from various AI related research by the U.S. government, scholars, and experts, this paper further attempts to examine the opportunities and challenges in the ethical area that will likely be generated by applying AI to the American system of military health care.*

Results: *This paper finds that the application of AI in military health care would likely provide immense opportunities, especially in times of COVID-19. However, there are risks posed by the application of AI in various sectors as well as in military health care. Meanwhile, the new Biden administration's future policymaking stage should emphasize ways in which AI may remain human-centric to be more supportive of human rights.*

Conclusion: *In incorporating AI into virtually every economic sector and into daily life, the U.S. government, together with all sectors of society, should seek ways to maximize opportunities and minimize the risk of over-dependence or unregulated exposure to an intelligence far beyond the human level. The new Biden administration should be at the forefront in promoting responsible AI policymaking to be more supportive of human rights.*

[Keywords] ***The Politics of Responsibilities, The Biden Administration, The U.S. Department of Defense(DOD), Human Rights Challenges, Potential AI Applications to Military Health Care***

1. Introduction

The high-tech industry of Artificial Intelligence(AI) is one of the fastest growing ones in the United States and beyond. Recent data indicates that It will likely increase GDP by 15.7 trillion dollars globally, by the year 2030[1]. The U.S. has been one of the world leading nations in regards to the AI research and development; China and Russia are potential competitors and are making significant investments. Still, the U.S. has managed to incorporate AI into its economy most intensely and its defense sector is using AI in its weapons and surveillances systems. The American public health care sector recently adopted AI upon the COVID-19 outbreak and most likely, AI will be further expanded in years ahead.

The so-called "Narrow" AI technologies and its application to different sectors have brought tremendous advantages to the country and its citizens. Yet, practical and human rights or ethical concerns remain. There have been increased human rights and legal concerns over the AI applications in arms race such as "killer robots"[2][3][4]. Moreover, some computer programmers mishandled AI algorithms, resulting in criticism and debate. The U.S. Department of Defense(DOD) has also recognized a potential AI bias, again leading to mismanagement[5]. Furthermore, the DOD's ongoing problem with racial discrimination requires structural change, and thus far AI may have been applied in a way that did not ameliorate concerns, including when it involved patients' pre-existing conditions. This paper attempts to examine how AI presents certain opportunities and risks for the sector of military health care. This paper further argues that the new Biden administration should pursue human-centric AI policymaking to be more supportive of human rights.

2. The Politics of Responsibility in Emerging AI Technology and Human Rights

Since the end of the Cold War, there has been the shift from focusing on traditional to non-traditional security. The non-traditional security issues range from human rights concerns such as infectious diseases, COVID-19 pandemic inequality, racism, rights concerns generated from emerging AI technology like cyber security, privacy rights and climate change. These are imminent threats to both national security and the international community. A Professor at Harvard University, Dr. Sikkink, argues that there is an urgent need to address the human rights threats posed by the crises individuals face today. In her relatively recent book, Dr. Sikkink considers forward-looking responsibility theory along with new rights such as digital privacy[6]. The new forward-looking approach to norm change is distinct from the "backward-looking" ones, which are limited in that they only address a small part of the problem[6]. Dr. Sikkink further argues that the U.S. government should assimilate these norms into policymaking and practice, particularly for the protection of digital privacy, thereby strengthening civil liberties[6]. Given the growing ethical and human rights concerns in AI development, experts from the Brookings Institution[7], argues that there is a need to build upon "responsible AI governance" to minimize discrimination when applying AI to various operational situations[7].

The second wave of AI has mainly been focused on Narrow AI. Narrow AI is being broadly incorporated into various operational situations for transportation, education, agriculture, finance, law, space science, defense, health care, humanitarian purposes[8], and manufacturing such as intelligence vehicles[9]. In health care, AI has helped improve efficiency and cost-effectiveness in treating patients especially during the COVID-19 pandemic in the U.S. when new challenges have exacerbated the system's structural inefficiency and high cost. To be specific, AI has been used in cases of pneumonia, respiratory illnesses, and virus transmission[7]. Experts expected that up to 18 billion dollars would likely be saved by applying AI applications to the U.S. public health care sector[7]. AI applications are also being widely incorporated into defense purposes by the DOD[10]. AI applications in U.S. military operational situations are changing its landscape of the U.S. military defense and warfare[11], and the DOD is further considering diverse AI applications. Numerous medical malpractice cases have been reported within the military, largely due to professional negligence [12]. Regulation involving victim compensation has many loopholes and the Government Accountability Office(GAO) identified bureaucratic inconsistencies[12]. To avoid cases of malpractice, practitioners in the military tend to be overly cautious and their practices have resulted in increased costs but without a commensurate increase in the quality of care[12]. In the public health care sector, AI or surgery has led to improved treatment. Furthermore, a study by MIT indicates that AI algorithms have led to fewer errors in terms of pharmaceutical dosage[7]. While more than 15 million people have been affected by COVID-19 worldwide[13], as of January 2021, approximately 50,000 people in the U.S. military have tested positive for COVID-19, and the numbers of infected are expected to be rise[14]. According to the DOD officials, the DOD has "the most comprehensive healthcare dataset in the world," and that might suggest that AI is a good match for the system of military health care[7]. The application of AI to military health care system, especially during pandemics, provides an opportunity for transforming it into a human-centric system to be more supportive of human rights.

The implementation of AI in military health care system raises some concerns. Computer programmers have not yet understood how to adequately train all of the data so that the interface avoids problems of bias and discrimination – problems that have caught the eye of the U.S. Congress.

Upon closer inspection, the original problem does not lie in AI but rather in how medical professionals approach the technology. Indeed, the issue of algorithmic bias is one that rests on the training of data[15]. More specifically, bureaucratic procedure is such that it often hamstrings the computer and leaves it with no option other than making a medical-related decision based on data that is partial or incomplete. The DOD's lack of a robust database leads to unintentional bias in terms of gender, age, race, and income[7]. Also, AI algorithms for facial recognition have been criticized for lacking diversity[5]. Here, too, the cause may well be that the DOD tasked AI with assigning only a handful of ethnic or racial categories to a full spectrum of faces, which could suggest that race is more of a social construction than a scientific one. Nevertheless, this bias has been observed during the process of medical diagnosis[16][17]. The U.S. military is susceptible to this problem because it is the most racially diverse institution in the country. The data from the survey conducted in 2017 and 2020 concluded that there is serious level of racial discrimination observed within the DOD and defense sector. According to the survey, 31.2 percent of black, 23.3 percent of Asian, 21 percent of Hispanic troops and its members have suffered from racial discrimination[18], and those human rights conditions seem to remain unchanged[19].

3. AI and Responsible Government Policy in the Biden Era

Prior to the Biden era, the Obama and Trump administrations followed somewhat different policymaking paths when elaborating strategic plans for AI. The most fundamental plans were outlined during the Obama administration, which still serve as guidelines. In late 2016, the Obama administration announced a seven-part "The National Artificial Intelligence Research and Development Strategic Plan"[20]. The seven strategies consisted of the following:

1. Make Long-Term Investments in AI Research;
2. Develop Effective Methods for Human-AI Cooperation;
3. Understand and Address the Ethical, Legal, and Societal Implications of AI;
4. Ensure the Safety and Security of AI System;
5. Develop Shared Public Datasets and Environments for AI Training and Testing;
6. Measure and Evaluate AI Technologies Through Standards and Benchmarks;
7. Better Understand the National AI R&D Workforce Needs [20].

The plans address issues ranging from its innovation to ethics, as well as inclusion of key regulatory and societal measures to address AI technology. More specifically, the Obama administration's strategic plans has sought to enhance AI's capability and reliability in manufacturing, medicine, agriculture, consumer products, and even healthcare, in the interests of both national security and ethics[20].

While the Obama administration had introduced the nation to human friendly AI, the Trump administration's efforts were limited in scope. It did not develop its own new AI strategy and federal funding had been sharply reduced. Across two consecutive years the Trump administration allocated 2.9 percent below the previous budget for the DOD's health care program in 2020[12], and 3.8 percent below the amount needed in 2021[12]. This budget reduction has largely been counterproductive for the application of AI in the defense department.

AI technology has rapidly been proliferated in virtually every sector of the nation and society of the U.S. in recent decades. AI has widely adopted in many sectors especially during the times of COVID-19[21][22][23]. The U.S. Department of Defense in particular, has been putting utmost efforts in investing AI, including more than 900 million dollars up until 2020[5]. The Joint Artificial Intelligence Center(JAIC) under the U.S. Department of Defense currently runs over 600 AI related projects[24], investing over 15 million dollars[25]. The JAIC has been putting efforts in employing AI technology in areas such as automated weapons system, cyberspace, and even in humanitarian aid and relief[5]. The JAIC has also made decision to invest over 800 million dollars in 2021 on a relatively new project "Maven" under the JAIC[5]. Together with the JAIC, The U.S. Department of Defense has also sought to implement strategic guidance on the laws and ethics related to AI operations[26]. Despite these DOD and JAIC efforts, the years ahead should include a careful review of policymaking.

The Biden administration announced that it was a priority to invest in AI even before it took office in 2021[27]. President Biden signed multiple executive orders and then reversed the Trump administration's policies on climate change, racial inequality, immigration, employment, and pandemic response[28]. This effort included restoring the Presidential Council on Science and Technology[29]. Also, President Biden's "Build Back Better" proposal implies that the new administration will embrace AI technology more fully than any other past administration. In fact, the "Build Back Better" plan is seeking the ways to promote innovation and responsibility in policies involving AI[30]. The new administration has already been committed to investing 300 billion dollars in AI research and development[31]. Furthermore, following Executive Order 13859, the Biden administration attempts to include tougher regulations and measures to protect some of the most contending AI related concerns regarding racial bias, cyber security[32], privacy rights[33][34][35][36], and the reform of the Section 230 of the Communications Decency Act[32]. Regardless of recent progress, these gray areas remain a concern for the DOD. The Biden administration has demonstrated its political will by restoring global leadership in AI research and development by promoting responsible and human-centric AI policymaking.

4. Conclusions and Implications

As analyzed in earlier sections of this paper, the new Biden administration will likely prioritize AI investment and strengthen AI related regulations in years ahead. The Section 230 of the Communications Decency Act(1996), for instance, seems somewhat outdated. As examined in the earlier sections, applying AI to military health care shows immense promise while simultaneously suggesting that ethical and human rights concerns need to be resolved. The new Biden administration should promote responsible AI policies in accordance with a human-centric approach. The government, along with governmental sectors, universities, experts, schools, and other individuals should work altogether in promoting responsible AI governance, particularly to maximizing the benefits and minimizing the risk of using AI technology.

5. References

5.1. Journal articles

[2] Maas M. Innovation-proof Global Governance for Military Artificial Intelligence?. *Journal of International Humanitarian Legal Studies*, 10, 129-158 (2019).

[3] Himmerlreich J. Responsibility for Killer Robots. *Ethical Theory and Moral Practice*, 22, 731-747 (2019).

[4] Burri T. The Politics of Robot Autonomy. *European Journal of Risk Regulation*, 7(2), 341-360 (2016).

[9] Li Y & Cho H & Park G. The Advent of AI and Its Some Implications to Military Affairs. *International Journal of Military Affairs*, 5(1), 38-47 (2020). [Article]

[11] Kim H & Park G. Ethical Issues on AI Equipped Combat Robots. *Robotics & AI Ethics*, 5(2), 1-7 (2020). [Article]

[13] Park G & Kim H & Li Y. Virtue for Post Covid-19 and AI Technology. *Robotics & AI Ethics*, 5(2), 8-18 (2020). [Article]

[16] Obermeyer Z & Powers B & Vogeli C & Mullainathan S. Dissecting Racial Biases in an Algorithm Used to Manage the Health of Populations. *Science*, 366(6464), 447-453 (2019).

[17] Panch T & Mattie H & Atun R. Artificial Intelligence and Algorithmic Bias: Implications for Health Systems. *Journal of Global Health*, 9(2), 1-5 (2019).

[21] Park G & Lee W & Lim Y. Covid-19 and Some Implications to AI. *Robotics & AI Ethics*, 5(1), 16-24 (2020). [Article]

[22] Harmon S & Sanford T & Xu S & Turkbey B & Roth H & Xu Z & Yang D & Myronenko A & Anderson V & Amalou A & Blain M & Kassin M & Long D & Varble N & Walker S & Barci U & Lerardi A & Stellato E & Plensich G & Franceschelli G & Girlando C & Irmici G & Labella D & Hammoud D & Malayeri A & Jones E & Summers R & Choyke P & Xu D & Flores M & Tamura K & Obinata H & Mori H & Patella F & Cariati M & Carrafiello G & An P & Wood B & Turbey B. Artificial Intelligence for the Detection of COVID-19 Pneumonia on Chest CT Using Multinational Datasets. *Nature Communications*, 11, 4080-4080 (2020).

[23] Scott I & Coiera E. Can AI Help in the Fight against Covid-19? Period. *The Medical Journal of Australia*, 2013(10), 439-441 (2020).

[32] Vought R. Memorandum for the Heads of Executive Departments and Agencies: Guidance for Regulations of Artificial Intelligence Applications. *Executive Office of the President: Office of the Management and Budget*, 2106, 1-16 (2020).

[33] Li Y & Park G. AI Ethics and Privacy Right. *Robotics & AI Ethics*, 5(2), 27-33 (2020). [Article]

[34] Bialek A. What's the Big Deal about Privacy?. *The National Law Review*, 10(30), 1-6 (2020).

[35] Rodrigues R. Legal and Human Rights Issues of AI: Gaps, Challenges and Vulnerabilities. *Journal of Responsible Technology*, 4, 1-12 (2020).

[36] Saetra H. A Shallow Defense of a Technocracy of Artificial Intelligence: Examining the Political Harms of Algorithmic Governance in the Domain of Government. *Technology in Society*, 62, 1-10 (2020).

5.2. Books

[6] Sikkink K. The Hidden Face of Rights toward a Politics of Responsibilities. Yale University (2020).

[7] West D & Allen J. Turning Point: Policymaking in the Era of Artificial Intelligence. The Brookings Institution (2020).

[15] Eubanks V. Automating Inequality: How High-tech Tools Profile, Police, and Punish the Poor. St. Martin's (2018).

5.3. Additional references

[1] Anand SR & Verweij G. Sizing the Prize: What's the Real Value of AI for Your Business and How Can You Capitalize? PricewaterhouseCoopers (2017).

[5] Sayler K. Emerging Military Technologies: Background and Issues for Congress. Congressional Research Service (2020).

[8] Harris L. Overview of Artificial Intelligence. Congressional Research Service (2017).

[10] Sayler K. Artificial Intelligence and National Security. Congressional Research Service (2020).

[12] Mendez B & Lewis K. Military Medical Malpractice and the Feres Doctrine. Congressional Research Service (2019).

[14] https://militarybenefits.info/ (2021).

[18] https://www.reuters.com/ (2021).

[19] https://www.reuters.com/ (2020).

[20] Baru C & Daimler E & Ferguson R & Forbes N & Harder E & Kenneally E & Kim D & Korelsky T & Kuehn D & Langendoen T & Lyster P & Morris KC & Munoz H & Rindflesch T & Schlenoff C & Sofge D & Spengler S. The National Artificial Intelligence Research and Development Strategic Plan. National Science and Technology Council: Networking and Information Technology Research and Development Subcommittee, 1-48 (2016).

[24] McCain J & John S. McCain National Defense Authorization Act for Fiscal Year 2019, Section 2, Division A, Title II, Authenticated U.S. Government Information (2018).

[25] Work R. Department of Defense: Artificial Intelligence, Big Data, and Cloud Taxonomy. Govini (2017).

[26] Defense Innovation Board. AI Principles: Recommendations on the Ethical Use of Artificial Intelligence. The U.S. Department of Defense (2019).
[27] https://www.brookings.edu/ (2020).
[28] https://www.bbc.com/(2021).
[29] https://www.whitehouse.gov/(2021).
[30] https://www.brookings.edu/ (2020).
[31] https://predictiontechnology.ucla.edu/ (2020).

6. Appendix

6.1. Authors contribution

	Initial name	Contribution
Author	JC	-Set of concepts ☑ -Design ☑ -Getting results ☑ -Analysis ☑ -Make a significant contribution to collection ☑ -Final approval of the paper ☑ -Corresponding ☑ -Play a decisive role in modification ☑ -Significant contributions to concepts, designs, practices, analysis and interpretation of data ☑ -Participants in Drafting and Revising Papers ☑ -Someone who can explain all aspects of the paper ☑

2021 6(1) 7-17 Submission: 2021/02/10, Peer review: 2021/02/20, Accepted: 2021/02/25, Published: 2021/03/30

Robotics & AI Ethics

Publisher: J-INSTITUTE
ISSN: 2435-3345

Website: www.j-institute.jp/ai/
Editor: ai@j-institute.jp

Corresponding author
E-mail: han3761297@naver.com

dx.doi.org/10.22471/ai.2021.6.1.07

The Preliminary Elementary School Teachers' Perceptions and Attitudes on the Future English Education Using AI Technology: Focusing on the Students Majoring in English Education

Daeun Han

Gwangju National University of Education, Gwangju, Republic of Korea

Abstract

Purpose: *The purpose of this study is to investigate preliminary elementary school teachers' perceptions and attitudes about the upcoming future English education, namely the fourth industrial revolution era.*

Method: *For the research, 91 preliminary elementary school teachers were surveyed. As for a questionnaire form, this study defines preliminary elementary school teachers' attitudes in the following categories; 1)interests in English education using AI technology, 2)innovation resistance and anxiety about English education using AI technology.*

Results: *The results of the survey data were as follows. First of all, the preliminary elementary school teachers' interest in the impact of AI technology on English education was high, but their interest remain at the 'personal'(1-2 stages) level. Second, the level of the preliminary elementary school teachers' innovation resistance was not high, but the anxiety level was above average. They showed willingness to take AI technologies as instructional tools and positive attitudes toward AI technology-embedded future classroom. On the one hand, they were anxious about a lack of knowledge about new teaching method and skills related with utilizing AI technology into their classroom.*

Conclusion: *The curriculum should be the tailored curriculum to meet their level and needs, which was required in the fourth industrial revolution through the use of AI technologies. It is to show the possibility of using AI technology to learn English so that we can maintain the current positive perception of the English education using AI that the preliminary school teacher showed. Also, in order to be open to any paradigm changes related to English education in the era of the 4th industrial revolution, concerns and proposals for a new method of English education should be actively discussed. Finally, the university of education, which is a teacher training institution, need to introduce to preliminary elementary school teachers what AI technologies can be combined with English education and what are their characteristics, and to guide how AI technology is being applied to English education*

[Keywords] ***Preliminary Elementary School Teacher, Artificial Intelligence(AI), Technology, Future English Education, Fourth Industrial Revolution***

1. Introduction

With the 4th Industrial Revolution emerging as a keyword in the whole society, and access to a wide variety of knowledge is growing more than ever, the demand for changes to the traditional paradigm of school education, which focuses on teacher-centered knowledge transfer, is increasing[1][2]. The case of English education is no exception, and the rapid development of artificial intelligence-based translation technology or the release of similar applications in Korea has caused a great repercussion in the English education world. Therefore, it can be said that now is the time when the search for the future direction of English education becomes inevitable.

In particular, the prediction that the importance of foreign language subjects in public education will change with the expectation that artificial intelligence, neural network translation technology, and voice recognition technology, which are emerging as major new technologies in the era of the 4th industrial revolution, will reduce the difficulty of foreign language communication[3]. The foreign language translation technology recently provided by portal sites such as Google and Naver provides qualitative results, and in the future, even if you are not good at foreign languages, especially English, it is expected that there will be no inconvenience in communicating with foreigners at home or abroad. As a result, there are many concerns from the media and experts who question whether foreign language education is really necessary in the future education field[4].

As such, in a situation where the public's awareness of the importance and necessity of foreign language teachers and foreign language-related subjects is gradually changing due to new technologies of the 4th industrial revolution, it is necessary to check the level of awareness of preliminary elementary school teachers who will teach English in the era of the 4th industrial revolution and the knowledge and attitudes that prepare them for the future aspects of English education. This is because foreign languages teachers, especially English teachers, are the subjects who should lead this change in a desirable direction and establish a new paradigm of English education.

Therefore, now at the threshold of the new era that the 4th industrial revolution will bring, this study aims to survey pre-primary elementary teachers who will directly select and utilize the 4th industrial revolution technology in the educational field, and to analyze to what extent they are aware of and prepare for the upcoming change and their attitude toward change. Research questions are as follows.

First, how much interest do preliminary elementary school teachers have in English classes using AI technology in the era of 4th industrial revolution?

Second, how are the preliminary elementary school teachers' innovation resistance and anxiety about English class using AI technology in the era of 4th industrial revolution?

2. Theoretical Background

2.1. The fourth industrial revolution technology for English education: artificial intelligence (AI)

Artificial intelligence(AI) technology, one of the core technologies of the 4th industrial revolution, is affecting the fields of education as well as language, economy, and society[5]. By embodying human perception, reasoning, and learning skills, computers have reached a level similar to human thinking[6]. In particular, the driving force behind the recent rapid development of AI is the learning ability of AI. Through the method of neural network machine learning, it not only recognizes and processes vast amounts of data as meaningful and useful data, but also learns and accumulates knowledge more efficiently if it undergoes similar processing in the future[6][7].

Neural network technology has been introduced in various fields such as image and voice recognition, neural network construction, natural language processing, expert systems, and proof of theories. Among these factors, the integration of neural network construction technology and natural language processing is the subject of interest of language education-related workers with artificial neural network translation such as Google's neural network translation and Naver ‘Papago’[8]. Neural network translation goes beyond the previous word unit or phrase unit translation, and enables translation by sentence unit based on a vast amount of learned big data, resulting in a much more accurate and natural translation than the existing translation method[9]. Nevertheless, on the one hand, it is pointed out that the translation of long texts, natural paraphrase, literary translation with human emotion, and discourse translation of oral language are still inexperienced[10]. However, the Naver 'Papago' development team noted that through continuous updates, it will ultimately enable translation in a manner similar to the human brain structure[8]. In addition, as many researchers are working hard to improve the performance of AI-based translation, this expectation seems highly feasible.

Meanwhile, in the field of English education, research using AI robots as a tool to expand students' English communication opportunities is also increasingly actively taking place. AI chatbot refers to a robot that chats, and is also called a chatterbot or talkingbot. It refers to an artificial intelligence program that can promptly respond to user's questions like a human according to a set response rule and can have human-like conversations[11]. It is predictable that chatbots will be able to play a role as conversational agents that are almost similar to humans in the future, but chatbots that can communicate the same as humans have not yet appeared. However, it is expected that it will serve as an excellent English learning tool for EFL learners, who have difficulty in learning English communication other than exchanging conversations in a fixed dialog format with someone other than native speakers.

2.1. Attitudes of preliminary elementary school teachers on future English education

2.1.1. Interest

First, interest is a concept derived from the 'Stage of Concern' referred to in the Concern-Based Adoption Model(CBAM)[12]. The CBAM's stages of concern framework(see <Table 1>) may be a useful heuristic for school-based consultants interested in assessing the attitudes and feelings teachers have towards an innovation, whether it is an individual intervention or a systemic school reform effort[13]. The stage of concern describe "the affective dimension of change: how people feel about doing something new or different, and their concerns as they engage with a new program or practice"[14].

Therefore, the CBAM can be an instrument that educational leaders use to evaluate innovations; it shows them how the individuals most affected by change react to the implementation of these innovations. Also, it can be a tool that can diagnose the current state of the inmate in the process of change. It is a state that considers the emotions, thoughts, prejudices, and considerations of an individual in the process of change.

In order to continuously accumulate the highly specialized knowledge required of the preliminary elementary school teachers according to new changes, constant attention and efforts of them are required. The CBAM is used to evaluate the level of interest(Stages of Concern) and the level of use, including individual expectations for innovation and change in education[3]. It is said that the process can be explained by dividing it into 0-6 steps. Accordingly, it is judged that preliminary elementary school teachers who will lead elementary English education in the future have some interest in future English education using AI technology, one of the technologies of the 4th industrial revolution, and it is necessary to find out what the level is.

Table 1. The CBAM stages of concern framework[3][15].

Interest			Examples
Impact	Stage 6	Refocusing	I know a more effective way.
	Stage 5	Collaboration	I want to collaborate with a colleague in using the new method.
	Stage 4	Consequence	How will the use of the new method affect students and teachers?
Task	Stage 3	Management	It seems that most of my time is spent preparing materials.
Self	Stage 2	Personal	What will be the effect on me if I use the new method?
	Stage 1	Informational	I want to know a new way.

Unrelated	Stage 0	Awareness	I have no interest in the new option.

2.1.2. Innovation resistance and anxiety

Teacher factors that are closely related to interest are the innovation resistance or anxiety that they have about the technology itself and the class using technology. Resistance is one of the major teacher factors that can naturally arise in innovative change in education. The degree of resistance of teachers, a key member, is important to successfully accepting change, and it can sometimes be a serious obstacle[16]. Innovation resistance is defined as a series of actions to maintain against the pressure to change the current state[17]. The higher the degree of change, the higher the degree of innovation, and also factors such as the characteristics of the innovation perceived by the user, the user's own characteristics, and the distribution path of the innovation influence the resistance degree[17][18].

On the one hand, along with the teacher's innovation resistance, the teacher's anxiety affects the overall performance of work including classes[3]. The factors of anxiety felt when adopting a new teaching method or application of new technology are described as a teacher's knowledge and training experience, teacher's role, class burden, management ability of equipment, negative evaluation, and comparison with other subjects[19]. As a result of using the ‘Foreign Language Class Anxiety Scale(FLCAS),’ which can measure the anxiety that non-native teachers feel when they teach a foreign language, it is said that the factors of teacher anxiety can affect the quality of education. Therefore, the anxiety becomes a major variable in preparing for changes in educational methods in the future era of the 4th industrial revolution and can be said to be an element that must be studied[20].

Thus, now that the 4th industrial revolution is having an enormous impact on our society as a whole, it can be said that it is necessary to explore the degree of anxiety felt by pre-primary elementary teachers who should lead classes in the future era of the 4th industrial revolution and how to overcome this anxiety.

3. Research Method

3.1. Research subject

This study targeted 91 preliminary elementary school teachers majoring in English education intensively attending a local university of education, and their basic information is shown in the following <Table 2>. The online survey program 'Naver Office Form' was used.

Table 2. Research subject.

Item		Number	Proportion (%)
Gender	Male	25	27.5
	Female	66	72.5
	Total	91	100.0
Grade	1st grade	27	29.7
	2nd grade	29	31.9
	3rd grade	21	23.1
	4th grade	14	15.4
	Total	91	100.0

3.2. Research tools

This study conducted a survey to find out the perceptions of pre-service teachers about the 4th Industrial Revolution and the use of AI, one of the latest technologies, in English classes. The questionnaire was revised and supplemented by reflecting the field and the latest in elementary English education based on the questionnaire items used in previous studies. The content of the questionnaire is as follows.

First of all, the questionnaire included personal information(2 questions) such as gender and grade to investigate the personal characteristics of preliminary elementary school teachers. To find out the degree of interest, innovation resistance, and anxiety of the preliminary elementary school teachers in future English education using AI technology in the era of the 4th industrial revolution, the questionnaire(22 questions) used in the study of Kim HS and Kim HY were modified to suit the characteristics and purpose of this study[3].

The questionnaire items were composed of multiple choice, narrative, and 5-point Likert scale. The online survey was conducted using the Naver questionnaire. In order to test the reliability of the questionnaire, the reliability was confirmed by calculating the Cronbach's alpha coefficient, which measures the internal consistency between items, and the results are shown in <Table 3>. In addition, to verify the validity of the content of the questionnaire, the questionnaire was reviewed by a researcher, two current English teachers, and one professor in the Department of English Education, and as a result, it was verified that there was no problem with the validity.

Table 3. Reliability.

Item	Reliability(cronbach’s alpha)
Interest in English education using AI technology	.923
Innovation resistance to English education using AI technology	.865
Anxiety about English education using AI technology	.906

3.3. Data collection and data analysis

This study was conducted in the stages of study preparation and design-survey-organization for about 6 months from May to October 2020. First, in May, research topics were selected and prior research analysis was conducted. In the months of June and July, the area and contents of the questionnaire were composed, revised and reviewed based on the research design and prior research. In August, a survey was conducted, and the online survey, which became the main data of this study, was currently conducted for undergraduate students in the 1st, 2nd, 3rd, and 4th grades majoring in English education at the University of Education located in the province. For the online survey, the Naver Office form questionnaire was created, and the link was shared in KakaoTalk's private chat room for each grade to obtain responses from various characters. In September and October, the response results of preliminary elementary school teachers were analyzed and interpreted using the SPSS program.

4. Research Results and Discussion

4.1. Preliminary elementary school teachers' interest in English education using AI technology

First, the degree of interest in AI technology related to the 4th industrial revolution was examined by preliminary elementary school teachers. The questionnaire items related to interest and the average of responses to each item are as follows.

The average of the eight questions asking the level of interest in AI technology of the 4th industrial revolution of preliminary elementary school teachers was 3.60, indicating that they showed interest above average. The item with the highest interest was 'Influence on my future job'(M=3.92), and interest in both 'Students' attitude'(M=3.82) and 'Influence on students(M=3.81)' were also relatively high.

It was found that preliminary elementary school teachers are showing considerable interest in how students will show attitudes and reactions when applying AI technology to English classes, and whether AI technology can have a positive effect on students. The next question of high interest was 'Influence on future English education'(M=3.63), and it was found that preliminary elementary school teachers are showing more than average interest in what changes the AI technology will bring to the future English education. This can be said to correspond to the 'Personal', which is the second stage of the CBMA's concern[3]. In other words, it can be said that preliminary elementary school teachers are at a stage where they simply have 'interest' for new technologies without specific goals or willingness to implement them without in-depth exploration of how new technologies will affect individuals and what changes they will bring[12].

On the other hand, 'Interaction with other people'(M=3.38) and 'Discussion on applicability'(M=3.40), which are stage 5 'Collaboration,' were relatively low in interest. The interest in the 'Development of new teaching methods'(M=3.44), which is the stage 6 'Refocusing', was also lower than the average. Through this, preliminary elementary school teachers are generally interested in future English education, the 4th industrial revolution, and the latest technologies, but without active introduction and specific action plans, they can be seen to be passively reacting to the upcoming changes.

4.2. Innovation resistance and anxiety of preliminary elementary school teachers in English education using AI

4.2.1. Innovation resistance

As a result of analyzing the results obtained through the five questions asked about the innovation resistance of preliminary elementary school teachers, it was found that they generally did not have much resistance(M=2.43). Preliminary elementary school teachers responded that they would not(M=2.41) to the question "I will adhere to the existing English education method rather than the English education method using AI technology." As for the question "I don't think that English education using AI technology will continues.", they answered that it was not(M=2.52). In particular, the question "I am willing to object other people's English education using AI technology"(M=2.18) showed a relatively lower degree of resistance than other items.

However, in the case of the question "I have a criticism of English education using AI technology", the degree of resistance was not great, but the average response was higher than other items. This shows the thought of preliminary elementary school teachers that even if the latest technologies, including AI technologies due to the 4th industrial revolution, become commonplace, there may be limitations or problems in applying them to the educational field. Therefore, it is judged that not only technological advancement but also research and support beyond the existing efforts and costs are necessary for the change of English classes in which the latest technologies related to the 4th industrial revolution are introduced.

4.2.2. Anxiety

Preliminary elementary school teachers' anxiety about future English education, which will change due to AI technology, averaged 3.17, which was above average. The most prominent item that showed anxiety was "I am worried that I will lack knowledge of teaching theory and method for teaching English using AI technology."(M=3.54). It is necessary to develop their own teaching methods or knowledge in the process of incorporating AI technology into English classes, and it is found that preliminary elementary school teachers are concerned that they will lack knowledge of related teaching theories or methods. Next, the question “I am worried that I do not have the skills required to take charge of English education using AI technology”(M=3.41) showed a relatively high level of anxiety. Even in the question “I am worried that I will have difficulty in preparing and operating English classes using AI technology”(M=3.37), the anxiety of preliminary elementary school teachers was relatively high. However, it was found that the anxiety from the evaluation of English proficiency by AI robots in the intelligent information society(M=2.75) and anxiety from future job changes(M=2.91) were lower than that of other items.

In general, preliminary elementary school teachers were at a low level of anxiety about the introduction of AI technology, but their lack of knowledge about the teaching theory and methods required for English class using AI technology, and their worries about the ability to use technology were relatively high. This can be presumed that preliminary elementary school teachers had fewer opportunities to learn about English education using AI technology due to the curriculum problem that does not reflect the latest trends in the current curriculum and adheres to the existing class content. Subsequently, responses were somewhat high that there would be many difficulties in using AI technology in class. This is generally contrary to the general expectation that AI and the latest technology will enable more efficient and faster work processing based on fast processing speed and automated systems. It can be seen that it is expected that a lot of time and effort will be required before acquiring and becoming familiar with the related knowledge and using it freely in order to introduce and apply new teaching methods to the classroom environment that changes with development. In addition, it can be speculated that the anxiety about the efforts for such ‘introduction’ could slow the change in the educational field by sticking to the existing method and creating a passive attitude toward change. Therefore, it is judged that the university of education, an educational institution that trains future elementary school teachers, needs to experience the latest technology including AI and a new form of English education that combines it and teaching practice.

5. Conclusion and Suggestion

The purpose of this study is to find out what kind of perceptions and attitudes preliminary elementary school teachers have in English education using AI technology in the era of the 4th industrial revolution. Based on this, it was intended to examine the direction of English education suitable for the era of the 4th industrial revolution by discussing what direction the future English education should change and proceed, and what measures are needed for this change. For the purpose of this study, a questionnaire survey was conducted on 91 preliminary elementary school teachers who majored in English education at a local university of education, and the results are as follows.

5.1. Preliminary elementary school teachers' concern about English education using AI technology

As shown in the previous results, preliminary elementary school teachers were very interested in English education using AI technology(M=3.60). This is in line with the research of Kim SY & Kim HY and Park SJ that pre-service teachers have a positive perception of English education using new technologies in the 4th industrial revolution. The results of the response to the level of interest are analyzed and summarized as follows.

First, it was found that the level of interest of preliminary elementary school teachers still remained at the level of individual related to them. They showed higher interest in “I want to know about ~” which is the question of 'self' corresponding to stage 1-2(informational, personal) of CBAM's stages of concern, such as the impact on my future career(M=3.92), impact on the students(M=3.81), and impact on future English education(M=3.63). On the other hand, the level of interest in applying teachers themselves more directly(stage 3 ‘management'), and in earnest discussion or knowledge sharing(stage 5 'collboration') was relatively low.

In other words, preliminary elementary school teachers are generally interested in the changing trends of the times and teaching methods, but they only have an interest in the changes that will affect themselves at the level of personal interest and curiosity. It is interpreted that this is not an active implementation step enough to specifically apply it to English education. As such, if the level of interest of preliminary elementary school teachers continues to remain at the level of personal interest, they will not be able to escape from the existing teacher-centered class and the auxiliary use of multimedia technology. It is judged that it is necessary to provide information on specific changes and application method to preliminary elementary school teachers, such as how AI technology can be applied to English education and what changes will be made to the future English education.

5.2. Preliminary elementary school teachers' innovation resistance and anxiety about English education using AI technology

Preliminary elementary school teachers who participated in this study were not very resistant to innovation in new English education using AI technology(M=2.43), but they had some anxiety(M=3.17). First of all, in the case of innovation resistance, they did not deny or criticize the changes in English education("I am willing to object to other people's English education using AI technology." M=2.18; "I have a criticism of English education using AI technology." M=2.79). They had the idea that this could be a sustainable change rather than a temporary change(M=2.52). They also showed an attitude to accept changes in new English teaching methods using AI technology(M=2.41).

The ability to use AI technology(M=3.41), difficulty in operating English classes using AI technology(M=3.37), the teaching theory required for English class using AI technology It can be seen that there is a higher anxiety about difficulties in practical application such as lack of knowledge about the method(M=3.54). These findings indicate that pre-primary elementary teachers are not greatly agitated by the changing educational paradigm in the future, but are concerned about the educational use of AI technology. Therefore, before going to the educational field, it is judged that it is necessary to practice for preliminary elementary school teachers to experience and demonstrate AI technology use cases.

Next, looking at the questions about anxiety one by one, the most essential problems predicted by the new English education paradigm in the era of the 4th industrial revolution, namely the fundamental things such as anxiety about the 'Development of AI technology beyond teachers' English proficiency'(M=2.75), and the 'Disappearance of job as English teachers'(M=2.91), were relatively low. On the other hand, it can be seen that there is a higher anxiety about difficulties in practical application such as 'AI technology utilization ability'(M=3.41), 'Difficulty in operating English classes using AI technology'(M=3.37), and 'Lack of knowledge about teaching theory and methods required for English class using AI technology'(M=3.54). These findings indicate that preliminary elementary school teachers are not greatly agitated about the changing educational paradigm in the future, but are concerned about how to use AI technology in English education. Therefore, it is judged that it is necessary to practice for preliminary elementary school teachers to experience AI technology use cases and demonstrate them before working at the educational field.

5.3. Suggestions

Based on the analysis of the results of this study, the suggestions of this study are as follows to improve teacher expertise or to reorganize the university curriculum. First, it is to show the possibility of using AI technology to learn English so that we can maintain the current preliminary elementary school teachers' positive perception of the English education using AI. It is necessary to tell how far the level of AI robot or automatic translation technology has reached, how much it can overcome the difficulties of speaking and writing in

English, or how diverse and abundant teachers' classes can be with voice recognition technology or English corpus data built with big data. Therefore, if the latest technologies such as artificial intelligence provide preliminary elementary school teachers with teaching methods and practical examples that enhance the interest and motivation of elementary school students and make English classes effective, the level of interest in English education using AI technology is beyond the individual level. It is expected to expand to interest in applying to and to full-scale discussion and knowledge sharing.

Second, in the 21st century modern society entering the era of intelligent information, the learning paradigm is changing according to the rapidly developing 4th industrial revolution technology. Accordingly, in order to be open to any paradigm changes related to English education in the era of the 4th industrial revolution, concerns and suggestions for a new method of English education should be actively discussed. From the results of this study, it can be seen that in the future society, there is a certain degree of criticism about English education using AI technology compared to other items, so it can be seen that the current preliminary elementary school teachers have some doubts about English education using AI. Therefore, it is necessary to present examples of innovative changes in English teaching and learning so far so that they can naturally accept the rapid changes in classroom environment and teaching methods.

Finally, in this study, it was found that a large number of preliminary elementary school teachers felt anxiety about the lack of knowledge acquisition and technical competence. Therefore, the University of Education, which is a teacher training institution, needs to introduce to preliminary elementary school teachers what AI technologies can be combined with English education and what are their characteristics, and to guide how AI technology is being applied to English education.

It is difficult to generalize this study because it is a limited survey of 91 preliminary elementary school teachers. In addition, since the speed of development of the latest new technology is very fast and its influence is also great, the perception of preliminary elementary school teachers may change over time. Therefore, it is not desirable to define the results of this study as fixed facts. However, it is judged that it is meaningful to examine the current level of awareness of preliminary elementary school teachers who are the subject to lead the right direction while experiencing the rapid change in English education in the era of 4th industrial revolution. In the future society, it is required to cultivate high-level thinking ability to give instructions, evaluate, and make decisions to AI, so the goal of English education will change from cultivating existing communication skills, and it will be necessary to establish a complex and multi-dimensional setting. Therefore, it is important to understand the level of perception of preliminary elementary school teachers who will lead such a changing English education. Accordingly, in the future studies, it is necessary to examine how much preliminary elementary teachers are prepared for the constantly changing educational environment and teaching and learning methods, and in what direction the improvement of teacher expertise should be made.

6. References

6.1. Journal articles

[1] Kang YD. A Design on Teaching and Learning Method for Creative Talent in the Fourth Industrial Revolution. *International Journal of Human & Disaster*, 5(1), 1-9 (2020). [Article]

[2] Kang YD. English Teaching Method Using Flipped Learning in the Artificial Intelligence Era. *Robotics & AI Ethics*, 4(2), 14-21 (2019). [Article]

[3] Kim HS & Kim HY. A Study of Korean English Teachers' Future Readiness for the Fourth Industrial Revolution. *Multimedia-assisted Language Learning*, 20(3), 179-205 (2017).

[4] Cho SS. The 4th Industrial Revolution' and Future Education. *Media & Education*, 6(2), 152-185 (2016).
[5] Baek MJ. The Study on the Criminal Subject and Liability of AI Robots. *International Journal of Justice & Law*, 2(2), 15-21 (2017). [Article]
[8] Kim YS. Elementary School Teachers' and Teacher Educators' Ideas of English Education in the 4th Industrial Society. *The Journal of Education*, 37(3), 123-150 (2017).
[12] Park SJ & Ihm HJ. Elementary English Teachers' Perception toward Future of English Education. *The Journal of Education*, 39(4), 123-144 (2019).
[13] Roach AT & Kratochwill TR & Frank JL. School-based Consultants as Change Facilitators: Adaptation for the Concerns-based Adoption Model(CBAM) to Support the Implementation of Research-based Practices. *Journal of Educational and Psychological Consultation*, 19, 300-320 (2009).
[14] Horsley DL & Loucks-Horsley S. CBAM Brings Order to the Tornado of Change. *Journal of Staff Development*, 19(4), 17-20 (1998).
[16] Kim HY. Teachers as a Barrier to Technology-integrated Language Teaching. *English Teaching*, 57(2), 35-64 (2002).
[17] Ram S. A Model of Innovation Resistance. *Advances in Consumer Research*, 14, 208-212 (1987).
[18] Park JH. Validation of a Model to Measure Teacher's Resistance to Change. *Andragogy Today: International Journal of Adult & Continuing Education*, 14(3), 1-31 (2011).
[20] Horwitz EK & Horwitz MB & Cope J. Foreign Language Classroom Anxiety. *The Modern Language Journal*, 70(2), 125-132 (1986).

6.2. Thesis degree

[11] Han DE. Effects of AI Chatbot on Korean EFL Learners' Speaking Ability and Affective Factors. Chonnam National University, Doctoral Thesis (2020).

6.3. Books

[6] Choi YS. Futurist's Artificial Intelligence Scenario. Korea.com (2016).
[7] Matsuo Y. Do Artificial Intelligences Outreach Human? Thinking of Deep Learning. Dongamnb (2015).
[15] Hall GE & Hord SM. Implementing Change: Patterns, Principles, and Potholes. Allyn and Bacon (2006).
[19] Hativa N. Technology and the Classroom Teacher. In Anderson LW International Encyclopedia of Teaching and Teacher Education Pergamon (1995).

6.4. Conference proceedings

[9] Park JH & Yoon SR. Comparative Analysis of Word/Subword/Character Level Korean-English Neural Machine Translation. Proceedings of Symposium of the Korean Institute of Communications and Information Sciences (2016).

6.5. Additional references

[10] https://m.biz.chosun.com/ (2017).

7. Appendix

7.1. Authors contribution

	Initial name	Contribution
Author	DH	-Set of concepts ☑ -Design ☑ -Getting results ☑ -Analysis ☑ -Make a significant contribution to collection ☑ -Final approval of the paper ☑ -Corresponding ☑ -Play a decisive role in modification ☑ -Significant contributions to concepts, designs, practices, analysis and interpretation of data ☑ -Participants in Drafting and Revising Papers ☑ -Someone who can explain all aspects of the paper ☑

Submission: 2021/02/10, Peer review: 2021/02/20, Accepted: 2021/02/25, Published: 2021/03/30

Robotics & AI Ethics

Publisher: J-INSTITUTE
ISSN: 2435-3345

Website: www.j-institute.jp/ai/
Editor: ai@j-institute.jp

Corresponding author
E-mail: david1@daum.net

dx.doi.org/10.22471/ai.2021.6.1.18

AI-Based CYBERSECURITY: Benefits and Limitations

Sangsoo Lee

Korea National Defense University, Nonsan, Republic of Korea

Abstract

Purpose: *This study intends to examine the efficiency and limitations of AI-based cyber defense systems by applying game theory and to explore the direction of the development of an AI-based national cyber defense system. In cybersecurity attackers and defenders can choose strategies and make a decision based on their resources, to attain the rewards, while anticipating the actions from opposing players.*

Method: *For better analysis, this article use Game theory an analytical framework to identify the position between the attacker and defender in the Cyber domain. Game theory has been used to observe competition among multiple competitors(players) fighting under pre-set rules. Game theory is appropriate for analyzing cyber interaction between defenders, attackers, and users and what happens as a result.*

Results: *To respond to infringement incidents, first, for artificial intelligence to more efficiently detect and respond quickly to security threats, it is necessary to analyze vast amounts of security data with human experience and knowledge and enter accurate data for artificial intelligence to learn. Second, it is necessary to strengthen the self-learning ability of security solutions to apply signature analysis, behavior analysis technology, and machine learning technology to enable automated detection and response to AI-based cyber attacks.*

Conclusion: *Although there are many limitations and risks of artificial intelligence, it is important to concern about how to use artificial intelligence usefully for the welfare and prosperity of mankind. First, cybersecurity is directly related to national security, and the government has to enhance an AI-based cybersecurity system. Second, to cope with increasingly diverse and evolving external cyber-attacks. Third, it is necessary to develop R & D investment and professional human resources to build an AI-based cybersecurity system. Fourth, it is required to overhaul related legal systems to strengthen AI-based cybersecurity. Although there are many limitations and risks of artificial intelligence, it is important to concern about how to use artificial intelligence usefully for the welfare and prosperity of mankind.*

[Keywords] ***AI, Cyber Security, Benefits, Limitations, National Security***

1. Introduction

Amid the growing importance of cybersecurity in the international community, artificial intelligence technology has begun to gain attention in cybersecurity. The application of artificial intelligence(AI)-based cybersecurity is increasing to cope with ever-increasing cyber-attacks. With the spread of untact culture in the COVID-19 pandemic, digital transformation, which expands from the real world to the digital world, is expected to expand further. Cybersecurity is a very important area of national security because cyberattacks can paralyze the nation's major infrastructures of power plants, transportation, and water supply.

For this reason, major countries are trying to establish national cybersecurity strategies against cyber attacks. The United States map out its national cyber strategy in 2018 and announced the nation's top cybersecurity R&D strategic plan in December 2019. Following the announcement of the national cybersecurity strategy in November 2016, the UK also announced plans to double its cybersecurity budget(2.8 trillion won) by 2020. China has announced an emergency response plan for a cyberattacks in June 2017. In September 2019, Korea also announced the National Cyber Security Basic Plan jointly with related ministries in the changing cybersecurity environment.

Cyber protection refers to overall activities to protect information assets from malicious electronic attacks on computers, servers, mobile devices, journalist systems, a network of vehicle, smart home, router data, etc. The size of the cybersecurity market is expected to grow further as digital transformation expands in the future. At the same time, digital transformation, represented by AI, big data, and cloud, is expected to increase the number of cyberattacks.

This study intends to identify the efficiency and limitations of AI-based cyber defense systems by applying game theory. It explores the direction of building AI-based cyber control systems. To this end, chapter 2, deals with the literature review and analysis framework, and chapter 3, analyzes the benefits and limitations of applying AI-based cyber control systems. Chapter 4 explores a building strategy of an AI-based cyber control system. Chapter 5 will summarize the existing review and present policy alternatives for building an AI-based cyber control system.

Amid the growing importance of cybersecurity in the international community, artificial intelligence technology has begun to gain attention in cybersecurity. The application of artificial intelligence(AI)-based cybersecurity is increasing to cope with ever-increasing cyber-attacks. With the spread of untact culture in the COVID-19 pandemic, digital transformation, which expands from the real world to the digital world, is expected to expand further. Cybersecurity is a very important area of national security because cyberattacks can paralyze the nation's major infrastructures of power plants, transportation, and water supply.

For this reason, major countries are trying to establish national cybersecurity strategies against cyber attacks. The United States map out its national cyber strategy in 2018 and announced the nation's top cybersecurity R&D strategic plan in December 2019. Following the announcement of the national cybersecurity strategy in November 2016, the UK also announced plans to double its cybersecurity budget(2.8 trillion won) by 2020. China has announced an emergency response plan for a cyberattacks in June 2017. In September 2019, Korea also announced the National Cyber Security Basic Plan jointly with related ministries in the changing cybersecurity environment.

Cyber protection refers to overall activities to protect information assets from malicious electronic attacks on computers, servers, mobile devices, journalist systems, a network of vehicle, smart home, router data, etc. The size of the cybersecurity market is expected to grow further as digital transformation expands in the future. At the same time, digital transformation, represented by AI, big data, and cloud, is expected to increase the number of cyberattacks.

This study intends to identify the efficiency and limitations of AI-based cyber defense systems by applying game theory. It explores the direction of building AI-based cyber control systems. To this end, chapter 2, deals with the literature review and analysis framework, and chapter 3, analyzes the benefits and limitations of applying AI-based cyber control systems. Chapter 4 explores a building strategy of an AI-based cyber control system. Chapter 5 will summarize the existing review and present policy alternatives for building an AI-based cyber control system.

2. Literature Review and Analysis Framework

The paper co-authored by Sangmin Park, Kyungho Lee, and Jongin Lim, Strategy Making Model Operational Intelligence Using The Game Theory in Cyber Attacks: In the Case study of KHNP(Korea Hydro & Nuclear Power Co., Ltd.) Hacking deals with the use of an open policy decision model in the process of strategic decision-making for the cyber terror response. The study pointed out that cyber warfare appears as a combination of hacking and psychological warfare and emphasizes the importance of strategic decision-making of the control tower for effective response[1].

An Analysis of Cyber Defense Information System for Utilization of AI Technologies, co-authored by Yeongwol Moon, Jangyong Park, Jinha Lee, and Jaeseung Song, analyzed information protection technology using artificial intelligence and proposed a phased development strategy of intelligent information protection system[2]. Artificial Intelligence-based Security Control Construction and Countermeasures, co-authored by Junhyeok Hung·byongyup Lee, calls for the establishment and operation of an artificial intelligence-based integrated control system that can analyze vast amounts of data by the evolution of cyber-attack methods and respond preemptively in a short period[3]. In his paper Rethinking Cybersecurity in the AI and Blockchain Age, Yoon Jung-hyun argues that the blockchain method is the emergence of an open security paradigm in cybersecurity, and suggests that it can reduce attacks by restoring resilience based on distributed networks and identifying traces of hidden attackers[4]. Bae Jae-kwon proposed the technical and institutional elements necessary for the construction of Enterprise Security Management System(ESMS) in his paper, A Study on the Establishment of Enterprise Security Management System Based on Artificial Intelligence and Big Data Analysis[5]. The main service areas of ESMS are digital security consulting, IT control service, and security consulting service.

This research agrees with the policy suggestions of the existing research and suggests realistic policies for establishing an AI-based cybersecurity system. In this paper, to resolve the unbalanced confrontation structure of defenders in cybersecurity, it was suggested to strengthen the ability to track and respond to the origin of the attacker's computer by strengthening the AI-based security solution. It also emphasized R & D investment and training of professionals as part of efforts to compensate for the instability of artificial intelligence. Also, it emphasized the expansion of cyber capabilities to prevent attacks, the improvement of early warning system performance, and the enhancement of counter-attack capabilities to cyber attackers.

As cyber-attacks have recently become more advanced, it is difficult to respond accurately, and artificial intelligence is recognized as an alternative to effectively respond to these security threats. The advantage of the AI-based method is that using AI technologies such as machine learning and natural language processing, analysts can respond confidently and quickly to threats[6]. First, on the learning side, AI learns using billions of data from both formal and unstructured sources(e.g. blogs and news). At AI, machine learning and deep learning technologies improve knowledge to "understand" cybersecurity threats and cyber risks. Second, on the analytics side, AI collects insights and uses inferences to identify the relationship between malicious files, suspicious IP addresses, or threats. These kinds of analyses are performed in seconds or minutes, allowing security analysts to respond to threats up to 60 times faster. Third, using AI's function of augmented reality eliminates the need to conduct time-consuming research tests and provides managed risk analysis. This reduces the time it takes for security analysts to make important decisions, adjust them to address threats, and respond quickly.

However, even if artificial intelligence is applied to security, it is necessary to use a supervised learning technique that mimics human cognitive ability to automatically classify and identify the attacker's intention. Despite these shortcomings, artificial intelligence technology has upgraded the current security technology stage, which not only guarantees the continuity of a stable business environment but also has the advantage of efficient manpower management. However, this AI provides an equal opportunity not only for defense but also for attack.

In terms of defense, just as artificial intelligence prepares countermeasures by automatically classifying and analyzing the attack types of various malicious codes through machine learning techniques, the attack side continues to detect and attack security vulnerabilities of the current system to identify loopholes insecurity[7]. It can also be used to effectively disguise attackers. The surprising fact is that artificial intelligence learns and evolves on its own. This means that at some point, artificial intelligence may develop into areas that humans have never thought of and create a new attack pattern that transcends human experience or imagination. AI is a useful tool for maximizing human understanding and behavior through a series of processes that automatically prepare data using augmented analytics, discover and interpret insights from data.

It should be noted that the development of artificial intelligence technology can trigger the evolution of attack systems along with the strengthening of cybersecurity systems. "The Malicious Use of Artificial Intelligence," written at the " AI Workshop" involving academics, civic groups, and industry, including the University of Oxford, UK, in February 2017, presented six AI attributes that have not developed enough to exceed human capabilities <Table 1>.

Table 1. Characteristics that artificial intelligence[8].

Characteristics	Main content
Duality	· Artificial intelligence can be used for civilian, military, etc., and is not specified to be used for specific purposes
Efficiency and scalability	· Certain tasks can be performed faster and cheaper cost than manpower, and at the same time, systems can be replicated to perform many tasks
Excellence	· Beyond the basic environment that humans have, they can perform better than humans in the water and at night
Anonymity	· Reducing areas of communication and face-to-face response with others
Supply ability	· AI-related technologies can be easily applied and abused differently from the original purpose
Vulnerability	· Many vulnerabilities have not been solved yet, such as attack techniques that cause errors by learning artificial intelligence with abnormal data or attack techniques and defects in autonomous systems that cause errors by adding noise to data used in image recognition algorithms

There are many vulnerabilities that have not been solved yet, such as attack techniques that cause errors by AI's learning of abnormal data which cause errors by adding noise to data used in image recognition algorithms. Major security threats using artificial intelligence are shown in Table 2. To solve these problems, people involved in AI should cooperate to understand and prepare for the abuse of AI, and actively respond in advance. It is necessary to coordinate stakeholders and experts in various fields that can prevent risks related to the abuse of artificial intelligence.

Table 2. Major security threats using artificial intelligence[8].

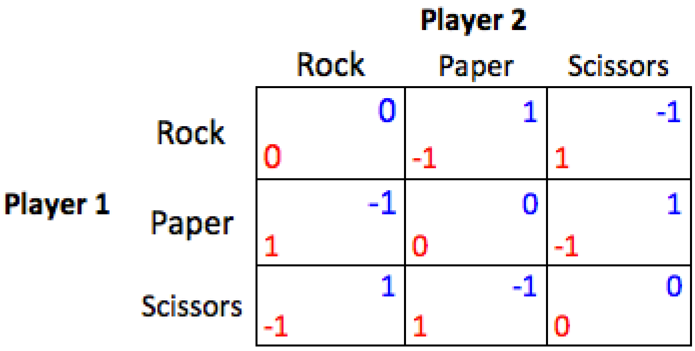
Attributes of AI	Main contents
Adversarial patch	· Stickers containing images cause artificial intelligence algorithms to malfunction · In the event of a problem with artificial intelligence applied to autonomous vehicles, there is a high possibility of accidents that are directly related to life
Spear phishing	· The form in which an attacker collects and enters information related to an attack target in advance and performs an attack · Existing phishing required skilled workers, but using artificial intelligence can reduce the proficiency used in attacks and expand the scope of attacks
Imitation	· Although it does not mimic the human voice, it is possible to imitate the voice using algorithms, and it is possible to impersonate and spread false information with synthetic technology · It secretly instructs hacking commands by giving voice recognition systems commands of white noise that AI recognizes but humans cannot decipher · Sing Deep Fake technology(Deep Learning + Fake), fake news can be used for social anxiety, and it is becoming increasingly difficult to determine the authenticity of the video
Automation of social engineering attacks	· It is a technique that targets vulnerabilities that humans have, and it is automated and managed in large quantities by sending messages by automatically creating mail links, etc. that the subject may be interested in
Sophisticated and automated swarm attacks	· Swarms that attack various vulnerabilities, access points, and devices are self-learning, exchanging information with each other, attacking multiple victims, and easing responses
Extracting black- box model	· Extracting parameters from the black box model, allowing malicious users to access the underlying technology
Leakage of personal	· The latest digital information devices are equipped with artificial intelligence technology,

information	and they have the potential to abuse personal information.

2.1. Game theory and cybersecurity

Game theory is an analytical framework used to observe competition among multiple competitors(players) fighting under pre-set rules. Game theory is appropriate for analyzing cyber interaction between defenders, attackers, and users. In this game, whether the player participating in the game is an individual or a team, the interaction minimizes the loss or gains rewards. Rock-paper-scissors, for example, two players have the same strategy and a 50/50 chance of winning <Figure 1>. A prerequisite is that participants in the game make reasonable decisions. Players can always play unpredictable games because they can't make the best choice under any circumstances.

Figure 1. Game theory[9].



		Player 2		
		Rock	Paper	Scissors
	Rock	0 0	1 -1	-1 1
Player 1	Paper	-1 1	0 0	1 -1
	Scissors	1 -1	-1 1	0 0

In cybersecurity attackers and defenders can choose strategies and make a decision based on their resources, to attain the rewards, while anticipating the actions from opposing players.

2.2. Cybersecurity is an imbalanced game

Defenders are always at a disadvantage in cyberspace. Even if all cyber safety technologies are used to defend, the defenders do not know when and where they may be attacked. Attackers always take advantage of cyber warfare. If cyber hacking occurs, it affects attackers, defenders, and everyday network users. They have different reasons to succeed. On the attacker's side, it takes advantage of the weakness of the attacker's computer to steal, tamper with, or paralyze the cyber network. Defenders, on the other hand, try to keep their information assets, data, and networks functioning normally. Legitimate network users want normal operations to handle their tasks. Attackers have the dominant strategy and always get a more substantial payoff than defenders and legitimate users. Attackers don't have to spend many resources, and still can maximize their interest if they can successfully infiltrate the network. Arms race are likely between attacker and defender in the cyber domain[10].

3. Benefits and Limitations for Defender

3.1. Benefits for defender

3.1.1. AI-based real-time autonomous cyber defense

Cyber domains are like a tilted playground for defenders. Efforts are being made to defend against the unexpected increase in cyber attacks, but there is a limit to preventing various cyber-attacks that are difficult to predict. As hackers evolve into various forms of attack, security paradigms continue to develop[11]. Attempts are underway to apply it to integrated security controls using artificial intelligence to analyze exponentially increasing data due to limitations in time and available resources. By utilizing artificial intelligence, it is easy to respond to security threats by quickly analyzing vast amounts of data[12]. An intelligent security control system can process real-time infringement, and reduce time by collecting and analyzing various digital device information using big data on one platform and applying it to protect and detect unknown threats. To improve this detection capability, it is necessary to develop security control algorithms based on neural networks <Table 3>.

Table 3. Integrated control system algorithm based on the neural network[5].

Division	Learning techniques	Algorithm	Characteristic
Event prediction	Guidance learning	XG boost algorithm linear booster algorithm	It improves the detection rate and reduces the error rate
Anomaly prediction	Unsupervised learning	Reconstructive tied-weight auto encoder, isolation forest & parametric stats	Accident prediction, abnormal behavior detection

Advanced countries like the U.S. and U.K. are launching software using artificial intelligence in cybersecurity to create a safe working environment against cyber threats. IBM of the United States operates 'X Force', a security solution that monitors global events in real-time and detects threats. In South Korea, many companies are already operating cyber protection systems using artificial intelligence to tackle real-time threats.

3.1.2. AI-based identification of unknown threats

Network-based Intrusion Detection(NIDS) refers to dealing with unauthorized access to information resources or illegal infringement of information resources. However, the current firewall-based border security method requires special experts and maintenance costs to analyze various attacks to create patterns. The STS(Statistics-Based Anomaly Detection) technique, which sets statistics on normal network users but there are limitations in identifying and responding to threats on a real-time basis due to the recent diversification of cyberattacks. If artificial intelligence is applied to solve these problems, normal network packets and various abnormal network packets are collected and various machine learning algorithms previously explained about these packets can be applied to determine whether real-time packets are normal or abnormal based on AI's learned knowledge[13].

3.1.3. AI-based preventive analytics for cybersecurity

Malware is a popular method of cyber-attack using viruses, worms, Trojan horses, exploits, botnet, retroviruses[14]. Most malicious code creators are creating and distributing various variants of code to avoid vaccines. The existing 'Signature Vaccine' analyzes and diagnoses samples or patterns of previously collected malicious codes, identifies treatment methods, and adds them to the anti-virus database. This method can respond quickly if a malicious code attack is added to the database but cannot respond to a new type of attack. To solve this problem, artificial intelligence can be used to quickly analyze new malicious code and update it to the database 10 times faster than humans. The AI-applied vaccine determines the abnormality of malicious code by using intelligent information such as various abnormal behavior, black IP list, Indicator of Code(Ioc), and C & C IP list based on the threat model of malicious code <Table 4 >.

Table 4. Main branches of cybersecurity applications adopting AI techniques[14].

Cyber application of AI-base methods			
Malware detection	Network intrusion	Phishing/spam detection	Other
· PC malware · Android malware	· Intrusion · Anomaly detection	· Web phishing detection · Mail phishing detection · Spam on social networks ·S pam mail	· Countering advanced persistent threats(APTs) · Identify domain names generated by domain

			generation algorithms (DGAs)

Recently, as intelligent malicious code is increasing rapidly and hacking technology advances, Endpoint Detection and Response(EDR) solutions are emerging to respond to them through detection and analysis on a real-time basis.

3.2. Limitations for defender

Although AI is a powerful tool, it is a relatively new technology and has limitations in cybersecurity. Security solutions have to be re-trained to keep up the function as new threats emerge. As various cyberattacks evolve, the type of cyber attack must be learned by machine learning to respond to the same type of attack[15][16][17]. AI can be also useful for cyber attackers. Through machine learning, they disguise their attack and learn what the AI-based security system is looking for or erase traces of the attack as if nothing had happened <Table 5>.

Table 5. The use of AI for malicious activities in cybersecurity[18].

Malicious use of AI	
Autonomous intelligent threats	Tool for attacking AI
· Strengthening malware · Social engineering	· Adversarial inputs · Poisoning training data · Model extraction

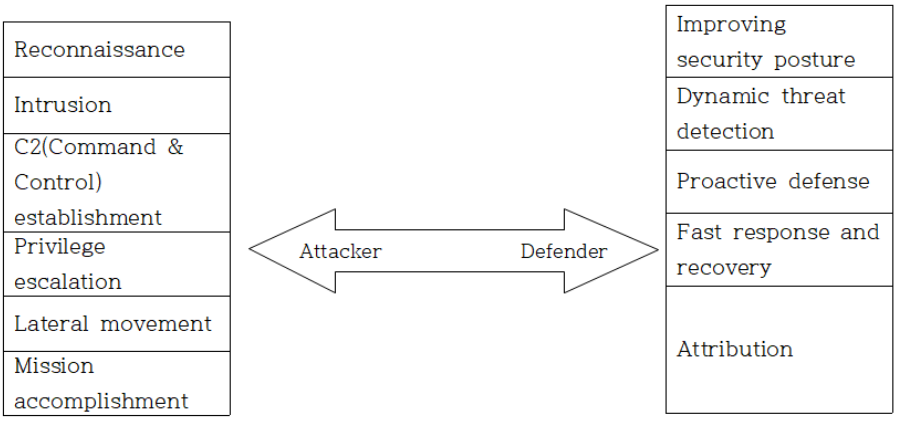
The artificial intelligence system has not developed enough to distinguish 100% between malicious and normal acts. Therefore, artificial intelligence should be continuously learned similar to human learning. Artificial intelligence quickly detects many cybersecurity threats and informs human analysts[19][20]. And AI shortens the time for cybersecurity officers to analyze threats. However, artificial intelligence is still not complete enough to completely replace dedicated IT experts. Cyber attackers equipped with artificial intelligence attack the opponent's computer, but infiltration is allowed if the opponent's artificial intelligence defense solution is poor, but it is blocked if the defense network is tight. Internet security issues can be understood as a series of unpredictable games of spears and shields between attackers and defenders in the cyber domain.

4. A Building Strategy of an AI-Based Cyber Control System

The cyber defense process of the AI-based security control platform goes through the process of threat data collection → analysis → response → management, and can effectively respond to cyber attacks[21]. The combination of artificial intelligence and cyber-attacks increases the scope and frequency of attacks, improves attack accuracy, and enables bypass attacks. Cyber attacks using artificial intelligence are attacked automation systems that can speed up cyberattacks or expand the scope of attacks without the help of experts. It also utilizes deep learning analysis to enable more sophisticated attacks. Using AI, various attacks that evolve malicious code for security control bypass can be executed. Cyber attacks go through five stages: reconnaissance, infiltration, Lateral movement, spread, and attack. Reconnaissance and penetration is a process in which a hacker detects and penetrates the vulnerable target. ZeroFOX, a security company, demonstrated this possibility in 2016. As a result of a cyber-attack using AI, AI wrote 819 while hackers wrote 200 comments and hackers potentially infected 49 people, whereas AI infected 275(2021).

Using artificial intelligence, AI can learn patterns of cyber mail security systems through machine learning and bypass them to attack. Also, IBM's AI malicious code Deep Locker can use AI-based image technology to identify the attack target and execute the Ransomware attack when the target uses the video app <Figure 2>.

Figure 2. Balance of AI attacker and defender[22].



Cyber attacks using AI are threatening, but AI is also used for prevention and detection in cyber defense. Prevention is a step that 'strengthens attack detection by analyzing cyber-attack patterns and strengthens security by analyzing internal vulnerabilities, and the major technology used in this area is 'Threat Hunting'. 'Threat hunting' is the process of finding threats and complementing vulnerabilities. In other words, it is to increase the level of security detection by analyzing the attack patterns. Detection is a step in preventing cyberattack attempts, and security officers can automatically analyze malicious code using AI. Tracing threat technology refers to Security Order Automation and Response(SOAR) security operating system and it is a platform that supports people and machines to cooperate organically to cope with various security threats in the computer system. The definition and merits of SOAR's technology are as follows <Table 6>.

Table 6. Definition and merit of major SOAR technologies[5].

Core technology	Definition	Advantages
Response to security incidents platform(SIRP)	· Technology to plan, manage, track, harmoniously manipulate and respond to security incidents	· It is possible to integrate different security tools, third-party threat alerts and, IT data, and increase the visibility of security events so that the importance of threat response can be classified and responded to
Threat intelligence platform(TIP)	· Measures and works on vulnerabilities, reporting collaboration tool shaping technology	· Security incidents can be reviewed and evaluated based on best practices and quarantine remedial measures are possible, and threat response and problem resolution are possible even in a work environment with excessive work and insufficient resources
Security operation automation and orchestration(SOA)	· Task, process, policy execution, and reporting automation and orchestration technology	· By automating actions that frequently occur in transactions, security engineers can focus on high-level threat analysis such as threat hunting

By automating actions that frequently occur in transactions, security engineers can focus on high-level threat analysis such as threat hunting. The response phase is a step in which a security controller quickly responds to cyber-attacks. It is a technology that categorizes the history of abnormal symptoms into the same things and shows them by time using AI's context technology. These programs allow controllers to quickly identify and respond to cyber-attacks. For cybersecurity, there is a need to use game theory to come up with a strategy in responding to Hacker' AI-based attacks. In particular, AI utilization in cybersecurity covers the areas of prediction, detection, and response <Table 7>.

Table 7. Status of artificial intelligence utilization by cybersecurity level[23].

Usage status	AI security area with high potential
Prediction	Analysis of user/machine behavior
Detection	Intrusion detection, fraud detection, and malicious code detection
Respond	Network risk analysis

The status of AI utilization by the cybersecurity stage of AI predicts behavioral analysis of users and machines. It is used for network intrusion, malicious code detection, and abnormal activity detection. And network risk analysis allows preventive measures to be taken to respond to cyber-attacks. There is the know-how to operate the latest security equipment and solutions to defend against cyberattacks, but the damage is also increasing due to failure to prevent evolving attacks on a real-time basis. Therefore, the next-generation artificial intelligence-based cybersecurity system will go through an organic cycle of prevention, control, response, and analysis <Table 8>.

Table 8. Security control process[21].

Division	Major contents
Prevention	· Strengthen security response level of service/system/security policy · Attack forecasting service based on new threat information collection system · Providing threat assessment analysis services
Control	· Strengthen security level of the security response system · Strengthening accident prevention functions by strengthening access control · Creating new security policies and reflecting changes in authority policies
Response	· Normal-based abnormal behavior detection and potential threat detection services · Priority and response to cyber threats · Isolation and blocking of exposed threats
Analysis	· Infringement precise analysis service using external threat intelligence · Providing emergency security policy & Push function of automation of security policy application

To respond to cyber-attacks, first, for artificial intelligence to more efficiently detect and respond quickly to security threats. It is necessary to analyze vast amounts of security data with human experience and knowledge and enter accurate data for artificial intelligence to learn[24]. Second, it is necessary to strengthen the self-learning ability of security solutions to apply signature analysis, behavior analysis technology, and machine learning technology to enable automated detection and response to AI-based cyber attacks. Although there are many limitations and risks of artificial intelligence, it is important to consider how to use artificial intelligence usefully for the welfare and prosperity of mankind.

5. Conclusion

As shown above, AI-based cybersecurity is a two-way unpredictable game between attackers and defenders. The preemptive attacker is always in an advantageous position, so for cybersecurity, the defender should continuously improve the function of the AI-based security solution and sometimes use a trick to analyze the attacker's behavior and take action in advance to maintain the resilience of the network so that the user can use it normally.

The AI-based cybersecurity system's advantage in its defense position is that it can utilize artificial intelligence to quickly process massive data analysis and cope with security threats in real-time. Second, the AI intelligence control system detects threats that are difficult to detect in advance.

Third, samples or patterns of collected malicious code can be analyzed and diagnosed to find out how to treat them and update them quickly to the database.

In contrast, the AI-based cybersecurity system has the demerit for the defender that it can not respond to a new type of attack as various cyberattacks evolve. Second, attackers can also utilize artificial intelligence. Cybercriminals disguise their attacks through machine learning what AI-based security systems are looking for or erase attack traces as if nothing had happened. Third, artificial intelligence shortens the time for cybersecurity officers to analyze threats. However, an AI-based cyber control system is still not complete enough to replace dedicated IT experts. Therefore, training skilled IT professionals is the best alternative for cybersecurity.

In the future, cybersecurity will emerge as a major agenda for national security, and AI technology is a requirement for security in response to advanced cyber attacks. AI is likely to lead to cybersecurity soon. The government should cooperate with private companies to support the development of AI-based cyber defense systems.

As a result of the study, AI-based cybersecurity development measures are as follows. First, cybersecurity is directly related to national security, and the government has to enhance the capability of an AI-based cybersecurity system. Second, to cope with increasingly diverse and evolving external cyber-attacks. The public and private sectors should cooperate organically and individuals, organizations, businesses, public institutions, and government agencies should form an integrated cyber crisis management system that suits each level. This is because computers operated by civilians are organically connected to government and military circles. So private computers infected with malicious codes infect those computers. Third, it is necessary to promote R & D investment and professional human resources to build an AI-based cybersecurity system that may give early warning to prevent cyber attacks. At the same time, the capability to control and reject attackers from a distance has to be enhanced. Fourth, it is also required to overhaul related legal systems to strengthen AI-based cybersecurity. Although there are many limitations and risks of artificial intelligence, it is important to concern about how to use artificial intelligence usefully for the welfare and prosperity of mankind.

6. Reference

6.1. Journal articles

[1] Park S & Lee K & Lim J. Strategic Decision Making Model Among Collective Intelligence Using the Game Theory in Cyber Attacks: Case Study of KHNP Hacking. *Journal of the Korea Institute of Information Security & Cryptology*, 26(1) 237-246 (2016).

[2] Moon JW & Park JY & Lee JH & Song JS. Analysis of Cyber Defense Information System for Utilization of AI Technologies. *Journal of Digital Contents Society,* 21(5), 891-900 (2021).

[3] Hong JH & Lee BY. Artificial Intelligence-based Security Control Construction and Countermeasures. *The Journal of the Korea Contents Association*, 21(1), 532-540 (2021).

[4] Yoon J. Rethinking Cybersecurity in the AI and Blockchain Age. *The Korean Journal of International Studies,* 59(4), 75-77 (2019).

[5] Bae JK. A Study on the Establishment of Enterprise Security Management System Based on Artificial Intelligence and Big Data Analysis. *Logos Management Review,* 18(2), 151-166 (2020).

[6] Lee JK & Han CH. Future Warfare and Military Artificial Intelligence Systems. *The Journal of Korean Institute of Communications and Information Sciences*, 44(4), 782-790 (2019).

[7] Nah S & Jasmine M & Kim JH & Joo J. Communicating Artificial Intelligence(AI): Theory, Research, and Practice. *Communication Studies,* 71(3), 369-372 (2020).

[10] Robert J. Cooperation Under the Security Dilemma. *World Politics*, 30(2), 210-211 (1978).

[11] Coker C. Artificial Intelligence and the Future of War. *Scandinavian Journal of Military Studies*, 2(1), 55-60 (2019).

[12] Yi L & Cho HJ & Park GY. The Advent of AI and Its Some Implications to Military Affairs. *International Journal of Military Affairs*, 5(1), 38-47 (2020). [Article]

[13] Cuzzolin F & Cîstea B & Sahakian BJ. Knowing Me, Knowing You: Theory of Mind in AI. *Psychological Medicine*, 50(7), 1057-1061 (2020).
[14] Thanh CT & Quoc BD & Ivan Z. Artificial Intelligence in the Cyber Domain: Offense and Defense. *Symmetry*, 12(410), 1-24 (2020).
[15] Park BS. Enterprise Security Management for IoT Services Based on Event Correlation in the Republic of Korea. *International Journal of Crisis & Safety*, 3(1), 17-24 (2018). [**Article**]
[16] Jo S. Terrorism Crisis on Northeast Asia. *International Journal of Crisis & Safety*, 5(1), 35-42 (2020). [**Article**]
[17] Lee J & Dong L. Suggestions for Using AI in Preparation for a Super-aging Society. *Robotics & AI Ethics*, 5(2), 57-64 (2020). [**Article**]
[18] Hanh CT & Quoc BD & Ivan Z. Artificial Intelligence in the Cyber Domain: Offense and Defense. *Symmetry*, 12(410), 14-15 (2020).
[19] Shabbir J & Tarique A. Artificial Intelligence and its Role in Near Future. *Journal of Later Class Files*, 14(8), 1-11 (2015).
[20] Lim Y & Lee M. Implications of Emotional Coaching and Integrated Art Therapy Teaching Method on Leadership Education in the AI Era. *Robotics & AI Ethics*. 5(2), 42-49 (2020). [**Article**]
[21] Hong J & Lee B. Artificial Intelligence-based Security Control Construction and Countermeasures. *The Journal of the Korea Contents Association*, 21(21), 532-533 (2021).
[24] Yoon J. Rethinking Cybersecurity in the AI and Blockchain Age. *The Korean Journal of International Studies*, 59(4), 65-66 (2019).

6.2. Additional References

[8] National Information Society Agency. The Abuse of AI Its Threats and Alternatives, National Information Society Agency Special Report (2018).
[9] https://threatpost.com/ (2021).
[22] WEF. Future Series: Cybersecurity, Emerging Technology, and Systemic Risk. Insight Report (2020).
[23] Gapgemini Research Institute. Reinventing Cybersecurity with Artificial Intelligence. Report (2019).

7. Appendix

7.1. Authors contribution

	Initial name	Contribution
Author	SL	-Set of concepts ☑ -Design ☑ -Getting results ☑ -Analysis ☑ -Make a significant contribution to collection ☑ -Final approval of the paper ☑ -Corresponding ☑ -Play a decisive role in modification ☑ -Significant contributions to concepts, designs, practices, analysis and interpretation of data ☑ -Participants in Drafting and Revising Papers ☑ -Someone who can explain all aspects of the paper ☑

2021 6(1) 29-34

Submission: 2021/02/10, Peer review: 2021/02/20, Accepted: 2021/02/25, Published: 2021/03/30

Robotics & AI Ethics

Publisher: J-INSTITUTE
ISSN: 2435-3345

Website: www.j-institute.jp/ai/
Editor: ai@j-institute.jp

Corresponding author
E-mail: hans.kim@pusan.ac.kr

dx.doi.org/10.22471/ai.2021.6.1.29

Suggestion of Building the AI Code of ETHICS through Deep Learning and Big Data Based AI

Hyunsoo Kim

Pusan National University, Pusan, Republic of Korea

Abstract

Purpose: *Current social and technical issues related AI Ethics take many different forms. Therefore, efforts to make ethical guidelines to cope with these issues are actively being developed. However, most kinds of ethical guidelines present general ethical principles and take a deductive method of solving individual problems in accordance. The purpose of this study is to propose creating ethical guidelines through an inductive method of deriving ethical principles based on ethical judgements on each individual AI-related cases.*

Method: *At first, most representative cases of AI ethics-related guidelines would be investigated in domestic and international level, with collecting documents and literature review. After that, examine the commonalities and differences between cases as these basic data through comparative research methods. Accordingly, it would be revealed that each case is constituted by a deductive method. Finally, as an alternative to these methods, presenting the merits of establishing ethical principles related to AI through inductive cases and specific examples.*

Results: *Most of the representative AI Code of Ethics that currently exist have the form of suggesting principles and then solving ethical problems by applying the principles to the actual events accordingly. This type of approach corresponds to the method of ethics which based on moral principles. However, complex and unpredictable problems are likely to arise when it comes to AI ethics. In order to solve these problems, it is necessary to extract the principles of the AI Code of Ethics by establishing and presenting ethical principles through researching and analyzing various individual events related to AI Ethics using deep learning and Big Data Based AI.*

Conclusion: *The following effects can be achieved by using deep learning techniques and Big Data Based AI that contains Ethical Issues together with sound and desirable Moral Judgement on each case, to derive the principles of the AI code of ethics. First, it is possible to extract and secure Big Data as basic resource of presenting Ethical Directions on various ethical problems arising in connection with the development of AI technology. Second, since the Ethical Principles as AI Code of Ethics are established based on empirical data, the validity of the principles can be secured. On the other hand, the AI Code of Ethics derived through deep learning based on such Big Data is likely to result in multiple tyranny or errors of majority due to certain limitations. So evaluation, verification and correction by Human Ethics Experts are essential to prevent these kinds of fault.*

[Keywords] ***AI Ethics, AI Code of Ethics, Deep Learning, Big Data, Human Ethics Experts***

1. The AI code of Ethics as a Widespread Trend

Today's issues related to AI ethics are unfolding in a wide variety of forms. Among them, the most recent case dealing with remarkable AI ethics related to recent scientific and technological developments such as Research on Future War and AI equipped Combat robots[1], Regarding the Ethical Principles required in the Post COVID-19 Era and AI Technology, a study that extracted four virtues of Transparency, Trust, Endurance, and Prudence by approaching from the perspective of three parties of Policy Maker, Researcher, and Citizens[2], A study that reviewed the definition and development process of the concept of privacy rights, and proposed the protection of regulation and the change of algorithm[3], Suggestions for setting the direction of AI ethics related to securing research grants in the Bioethics area[4], and such are the like exist as representative themes.

Meanwhile, in the field of health care and education where AI is actively used, discussions are being actively conducted not only on the use of AI technology itself, but also on ethical issues related to the use of AI. For example, suggestions related to the complex use of AI technology in the super-aging society[5][6], A study in which discussions on Emotional Coaching and Art Theraphy were developed by paying attention to the emotional aspects of humans requested in the AI era[7], Research on innovative teaching methods using AI in untact situations[8], A study in which AI-based Character Educations tried to approach the Rubrics and Schoolwide Approach by referring to six benchmarks of the Character Education Framework[9], and such could be said as recent research trends related to AI Ethics.

These studies can be said to be efforts to establish AI Ethics in common, or to reflect this in real-world issues. However, these studies show a pattern to be implemented in the form of trying to suggest a certain code of conduct in AI Ethics. Originally, efforts to establish principles and guidelines for AI-related issues began from the academic perspective. And after that it is gradually being actively deployed in research institutions, IT companies and other industries till now. Some representative examples of this Code of Ethics can be identified as follows: First, at the government level, there are EU Commission's Ethics Guidelines for Trustworth AI, Ten Principles of UK NHS AI Code of Conduct, and US Depart of Defense's AI ethical principles that encompassing five major areas, and so on. Second, at the enterprise level, there are Google AI Principles, KAKAO AI algorithm Code of Conduct which is one of the first AI ethics applied in the IT tech field in Korea, and such are worth to review. Third, in the Public Domain and Civil Society, there are a wide variety of forms of AI Code of Ethics are being proposed.

This phenomenon can be called as an explosive increase in AI Code of Ethics. And the reason that various subjects propose such Ethics of AI can be said that it is the starting point for making Ethical AI in the end[10]. At this time, Ethical AI is a concept that assumes a subject that interacts ethically with humans and society. In addition, since these subjective actions should based on criteria for judgment, it can be said that the Ethical Code of Conduct that AI must comply with is requested.

These norms are basically presented in the form of a declarative presentation of the principle of action. There are examples of this trend such as the results of a recent study of AI code of conduct or guidelines on Ethical AI corpus mapping and analysis, and when various ethical standards of conduct are synthesized, the five principles of transparency, justice and fairness, non-maleficence, responsibility, and privacy are extracted and presented[11]. And such kinds of AI Code of Ethics could be implemented in a form applied to situations where actual ethical judgment is requested. Such a mechanism is composed in the form of creating and presenting a declarative sentence that can be applied to a specific subject or area based on confirming the ethical norm that is generally applicable. And again it is used in the form of interpreting the contents of these declarative sentences and applying them to individual ethical issues whenever they arise.

On the other hand, there are studies that emphasize the importance of human judgment in attempts to incorporate the correlation between AI and Ethics[12]. According to this, it is necessary to re-interpret the concept of “autonomous”, which is one of the unique properties of AI. This study points out that AI-equipped machines or robotics such as driverless cars cannot themselves be completely ethical or moral agency. This also means that a new form of ethical choice made by human beings should be at the core, in a new period in connection with AI which is related to the development of science and technology.

It can be seen that such various discussions do not entirely object to the need for AI code of Ethics. Then, it is necessary to examine the commonalities of these perspectives and the extent to which the extension can be expanded in the future.

2. The Necessity of Establishing AI Code of Ethics Using Deep Learning and Big Data

There are lots of AI code of Ethics, but most of them are merely a listing or declaration of certain virtues. For example, In the United States, Future of Life Institute adopted Asilomar AI principles. In Japan, Ministry of Internal Affairs and Communications Japan drafts AI R&D Guidelines. In China, Ministry of Science and Technology suggests the Governance Principles for a New Generation of AI which is drafted by the National New Generation AI Governance Expert Committee. In Korea, Korea Communications Commission presents AI Charter of Ethics, etc.

The common features of various types of AI code of Ethics discussions that have been conducted so far are summarized in the following two aspects. First, there is a broad spectrum of AI's ethical behavior. One aspect is that the autonomous behavior of AI and its resulting ethical judgments are recognized only at a very low level, and only Ethical Judgments done by Human being are worth to regard. The other side of aspect is that acknowledging the high level of AI's autonomous behavior and requesting the AI's own Code of Ethical Conduct for that autonomous behavior. Second, this spectrum of ethical considerations means that value judgment by Human being must be intervened as a whole. On the one hand, it can be in the form of human teaches AI the ethical norms, and on the other hand, it can be in the form of AI learning ethical norms by itself.

Efforts to establish the AI code of ethics discussed above are the AI code of conduct that humans establish and let AI learn it, or AI that responds to unexpected ethical situations based on the AI Code of Ethics established by humans. So it is analyzed as a form of evaluating the behavior of a person. In reality, however, concentrated elaborating of multi-stakeholder and each political subjects organizations is required in relation to AI's Ethical Judgment[13]. And although not in full form, this view can be said to be an acknowledgment of moral characters in artificial moral agents at a certain level. In particular, this perspective can be understood as being in context with the approach of virtue ethics that focuses on the characteristics of the actor[14].

Meanwhile, problematic or facing a dilemma situations that are the basis of ethical judgments requested in everyday life have the characteristics of Big Data. These kinds of situations have combinations of situations that are visually recognized and situations that are communicated verbally. In order to make moral judgments in any given situation, very complex information is required. In this regard, it is worth referring to the following research results such as a study applying deep learning technology for analysis in the visual aspect[15], A study suggesting an approach to core linguistic processing issues by applying deep learning technology to natural language processing[16], etc. These studies correspond to the thing which could be accessed from the AI side in relation to the various data collected and used by humans in making moral judgments.

On the other hand, moral decision-making performed in human life has a characteristic that considers a wide variety of factors in a complex manner. It would be impossible to reduce these characteristics to complete quantitative data, but it is considered that it is possible to categorize and understand these characteristics to some degree. As the categories of data related to moral judgment become more subdivided and increased, it can be said to have the characteristics of big data for decision making. Data related to this decision-making can be used to make ethical choices in conflict situations as the ultimate goal pursued by Deep Learning.

Summarizing the above discussion, efforts to establish the AI Code of Ethics made up to now consist of a deductive procedure that declaratively presents ethical principles first and then confirms individual issues accordingly. However, humans use very diverse and complex data to make ethical judgments, and this has the characteristics of Big Data. If so, the design of an inductive mechanism that enables AI to make ethical decisions using Deep Learning techniques using such enormous data as basic data is considered to have some validity. If so, it is necessary to review the process of exploring the Code of Ethics using Deep Learning and Big Data.

3. Procedure of Exploring Code of Ethics Using Deep Learning and Big Data

As discussed above, the topics related to AI Ethics tend to be explored in connection with the perspective of virtue ethics related to human behavior. However, not only this, but there are also studies that understand today's society from the view of the smart community together with the perspective of moral development. This study sets up a topic tree related to ethical issues, explores the possibility that AI can act as an artificial moral agent, and suggests that the Moral Competence Test should be applied to this. This research also extracts key issues and its related topics in Fourth industrial revolution and mapping it into eight subjects[17]. The key issues presented in this study are significant in that they can be presented as initial conditions for Deep Learning.

With this, there is a research which use the method of Cognitive Computing in relation to solving the refugee problem[18]. The research provides certain implications for the correction method of the results of Deep Learning through the intervention of Human Ethics Experts in establishing AI Ethics based on Deep Learning and Big Data. This study approaches the stages of Cognitive Computing by dividing it into Observation, Understanding, Analysis, and Decision Making. According to this model, basic resources are observed at the Big Data as a raw level, understanding and analyzing these through AI technology, and human intervention is performed in the final decision making stage.

Applying these points, the procedure in which the expression learned by the Deep Learning algorithm appears has similar characteristics to that of ethical decision making. First, for ethical decision making, various related norms established through the tradition of ethics can be set in each layer of Layered Representations Learning or Hierarchical Representations Learning. Next, Autonomous Machine Learning will be performed based on nonlinear and qualitative data that is requested for this ethical decision making. Finally, these results will be reviewed by Human Ethics Experts.

In the above process, most of the learning about ethical decisions will be done by AI, but the most important is the role of Human Ethics Experts. When making an ethical judgment on a real issues or problem, there is a risk of bias due to external effects. In addition, in some cases, if an ethical judgment requires a choice contrary to its own interests, there is a risk of falling into self-contradiction and justifying the wrong choice. In response to these problems, it is the task of Human Ethics Experts to modify the judgment mechanism or, if inevitable, to set up a device that can manual override and to play the role of the final supervising or executive manager. Of course, the result of this correction will be integrated into Deep Learning Mechanism and AI function, thereby elaborating the model of AI Code of Ethics establishment.

The above mechanism is the basic mechanism for collecting various data to be considered in a given situation for individual ethical decision making in the form of Big Data and making ethical choices based on this. And the Dig Data set, which gathers content related to each of these ethical decisions, can also be incorporated into higher-level ethical principles. Meanwhile, the essential consideration when using such Big Data is related to privacy as a human virtue[19]. Various information is requested from the judgment mechanism for ethical decision-making, and this judgment contains extremely personal judgment on sensitive issues. In some cases, such information may lead to a judgment on the value of the person who made such a decision, so basic data on the decision making itself can be used, but personal information that could confirm the identity of the decision maker must be protected, and clear action for protecting each decision maker is needed to make use of Big Data possible.

4. Conclusion

All the AI practitioner should maintain and attribute to AI's ethical responsibility for building AI ethics at all stages of conception, production, distribution, and usage of AI code of Ethics. Ai Ethics is a kind of collective work of sound members of the human societies[20]. The AI Code of Ethics is weakly persuasive when approached only declaratively. We need to explore what millions of people around the world think about AI ethics, and create an AI code of ethics based on this, so that more people can sympathize and participate.

The following effects can be achieved by using deep learning techniques and Big Data based AI that contains Ethical Issues together with sound and desirable Ethical Judgement on each case, to derive the principles of the AI Code of Ethics. First, it is possible to extract and secure Big Data as basic resource of presenting Ethical Directions on various ethical issues or problems arising in connection with the development of AI technology. Second, since ethical principles are established based on empirical data, the validity of the principles can be secured. The big data related to ethical decision making collected in this way becomes data that enables general people as a majority of the public to know the results and trends of ethical decision making encountered in daily life. And it can be said that the result of learning based on such a large population is representative of the ethical decision making pattern made by human being at a general level.

On the other hand, the Code of Ethics derived through Deep Learning based on such Big Data is likely to result in multiple tyranny or errors of majority vote due to certain limitations of multiple judgments, so it must be evaluated, verified and corrected by Human Ethics Experts. This is because the direction the majority chooses in any ethical conflict cannot necessarily said to be ethical. Human Ethics Experts verifying the results of AI decisions on ethical issues based on the Ethical Principles learned by Deep Learning techniques in relation to any ethical decision making, solving ethical dilemma situations. It should be used not only as a prescription for the problems as ethical treatment, but also as basic data to analyze the problem when ethical choices do not occur. Further research related to this will be developed in the future.

5. References

5.1. Journal articles

[1] Kim H & Park G. Ethical Issues on AI Equipped Combat Robots. *Robotics & AI Ethics*, 5(2), 1-7 (2020). [Article]

[2] Park G & Kim H & Li Y. Virtue for Post Covid-19 and AI Technology. *Robotics & AI Ethics*, 5(2), 8-18 (2020). [Article]

[3] Li Y & Park G. AI Ethics and Privacy Right. *Robotics & AI Ethics*, 5(2), 27-33 (2020). [Article]

[4] Schuklenk U. On the Ethics of AI Ethics. *Bioethics*, 34(2), 146-147 (2020).

[5] Seo E & Park G. The Super-aged Multiculturalism in South Korea and the Necessity of Wearable AI Ethics. *Robotics & AI Ethics*, 5(2), 34-41 (2020). [Article]

[6] Lee J & Liang D. Suggestions for Using AI in Preparation for a Super-aging Society. *Robotics & AI Ethics*, 5(2), 57-64 (2020). [Article]

[7] Lim Y & Lee M. Implications of Emotional Coaching and Integrated Art Therapy Teaching Method on Leadership Education in the AI Era. *Robotics & AI Ethics*, 5(2), 42-49 (2020). [Article]

[8] Choi J. Instructor Competency for Innovative Teaching Methods in the Untact Era. *Robotics & AI Ethics*, 5(2), 50-56 (2020). [**Article**]

[9] Kim Y & Park G. Rubrics and Schoolwide Approach to the Character Education and Some Implications to AI-Based Character Education. *Robotics & AI Ethics*, 5(2), 19-26 (2020). [**Article**]

[10] Siau K & Wang W. Artificial Intelligence(AI) Ethics: Ethics of AI and Ethical AI. *Journal of Database Management*, 32(1), 74-87 (2020).

[11] Jobin A & Lenca M & Vayena E. The Global Landscape of AI Ethics Guidelines. *Nature Machine Intelligence*, 1(1), 389-399 (2019).

[12] Etzioni A & Etzioni O. Incorporating Ethics into Artificial Intelligence. *Journal of Ethics*, 21(1), 403-418 (2017).

[13] Piano SL. Ethical Principles in Machine Learning and Artificial Intelligence: Cases from the Field and Possible Ways Forward. *Humanities and Social Sciences Communications*, 7(9), 1-7 (2020).

[14] Gamez P & Shank DB & Arnold C & North M. Artificial Virtue: The Machine Question and Perceptions of Moral Character in Artificial Moral Agents. *AI & Society*, 35(1), 795-809 (2020).

[15] Garcia-Garcia A & Orts-Escolano S & Oprea S & Villena-Martinez V & Martinez-Gonzalez P & Garcia-Rodriguez J. A Survey on Deep Learning Techniques for Image and Video Semantic Segmentation. *Applied Soft Computing*, 70(1), 41-65 (2018).

[16] Otter DW & Medina JR & Kalita JK. A Survey of the Usages of Deep Learning for Natural Language Processing. *IEEE Transactions on Neural Networks and Learning Systems*, 32(2), 604-624 (2020).

[17] Kim TW & Park GY & Seo ES. IR4.0 and Ethical Tasks of AI. *Robotics & AI Ethics*, 4(2), 6-13 (2019). [**Article**]

[18] Kim HS & Park GY. AI-based Migrant Crisis Management. *Robotics & AI Ethics*, 5(1), 1-7 (2020). [**Article**]

[19] Puaschunder JM. Big Data Ethics. *Journal of Applied Research in the Digital Economy*, 1(1), 55-75 (2019).

[20] Orr W & Davis JL. Attributions of Ethical Responsibility by Artificial Intelligence Practitioners. *Information, Communication & Society*, 23(5), 719-735 (2020).

6. Appendix

6.1. Authors contribution

	Initial name	Contribution
Author	HK	-Set of concepts ☑ -Design ☑ -Getting results ☑ -Analysis ☑ -Make a significant contribution to collection ☑ -Final approval of the paper ☑ -Corresponding ☑ -Play a decisive role in modification ☑ -Significant contributions to concepts, designs, practices, analysis and interpretation of data ☑ -Participants in Drafting and Revising Papers ☑ -Someone who can explain all aspects of the paper ☑

6.2. Funding agency

This study was supported by the Fund for Humanities & Social Studies at Pusan National University 2020.

2021 6(1) 35-44

Submission: 2021/02/10, Peer review: 2021/02/20, Accepted: 2021/02/25, Published: 2021/03/30

Robotics & AI Ethics

Publisher: J-INSTITUTE
ISSN: 2435-3345

Website: www.j-institute.jp/ai/
Editor: ai@j-institute.jp

Corresponding author
E-mail: pgy556r@daum.net

dx.doi.org/10.22471/ai.2021.6.1.35

Future Oriented SMART SOCIETY and Ethical Issues

Taewoong Kim[1]

Gyeongsang National University, Jinju, Republic of Korea

Gyunyeol Park[2*]

Gyeongsang National University, Jinju, Republic of Korea

Abstract

Purpose: *This study explores the ethical issues that arise in building a community where humans can use smart technology to lead happier lives, a smart society.*

Method: *Technology in a smart society is not unrelated to the human variable, no matter how advanced its level is. As humans pursue value, smart societies also need to be explored from an ethical perspective. This study provides the need for an ethical approach that emphasizes humanity in the process of building and maintaining a smart society. The research was conducted by contents based approach.*

Results: *This research showed that the smart society was clarified by the Fourth Industrial Revolution, and got two characteristics as follows: First, AI, which is the core of smart society in particular, fosters human nature and potentialility through automatic self-realization, human representation, personal and social skills, and social cohesion. Second, smart AI ethical standards include beneficence, non-maleficence, autonomy, justice, and explicability.*

Conclusion: *This paper presented a normative and substantial directions to what preparations should be made for human-centered ethical coexistence in a future-oriented smart society. The ethical virtue or a way of living could be summarized as followed: love for humans, retrospective understanding, variable dialogue, self-sustainability, tarnsparency, explainability.*

[Keywords] ***Smart Society, Ethical Approach, Smart Internet, Industrial Revolution, Smart Technology, Artificial Intelligence***

1. Introduction

Human being has pursued a desire for survival. Currently human being has met a converged information society based on ICT(Information and Communications Technology)[1][2][3][4].

Smart society is based on a smart technology which contains IoT(Internet of Things), cloud computing, big data, artificial intelligence, etc. It means that in that society we human being can survive, feel happy, and can deal with whole things through a smart technology. And it is possible to innovate the entire society that consists of industry, economy, administration, and even culture. It could be said a new industrial revolution. A smart society have met the necessity that a smart technology and its values have been devoted to develop a human-centered society[5][6][7][8].

Through the process of the industrial revolution up to now, various needs of mankind have been satisfied. And it was an overall paradigm in the fields of industry, economy, society, academics, and arts due to vigorous scientific and technological development. In particular, the shift has been accelerated by the convergence technology revolution. IoT basically can make whole objects connect through the Internet. Based on this, mankind can lay the groundwork for realizing the smart society that humanity pursues[7][8].

However, as modern science and technology advances rapidly, humans must quickly adapt to the new environment. In making decision and prosecute the process, humans could not enough time to verify the effectiveness of the technology. In this process, the smart society focused on the satisfaction of human needs. To make linkage between humans and technology, it should be dealt with deliberately in the point of ethical view[8].

In order to pursue a future smart society, humans converge in terms of science and technology. Each part has developed to the independent ethical perspective. So it is necessary to introduce and promote technologies based on the ethics.

Recently, artificial intelligence(AI) is improving its completeness by self-learning based on big data. However, the acceleration of commercialization has caused many social problems. The reason came from that AI has not been constructed on the basis of the ethics. In this ethical foundation, there are two kinds. One ethics is focused on AI itself, another is focused on the human who made AI. In a smart society, human-centered ethics should be prioritized[8].

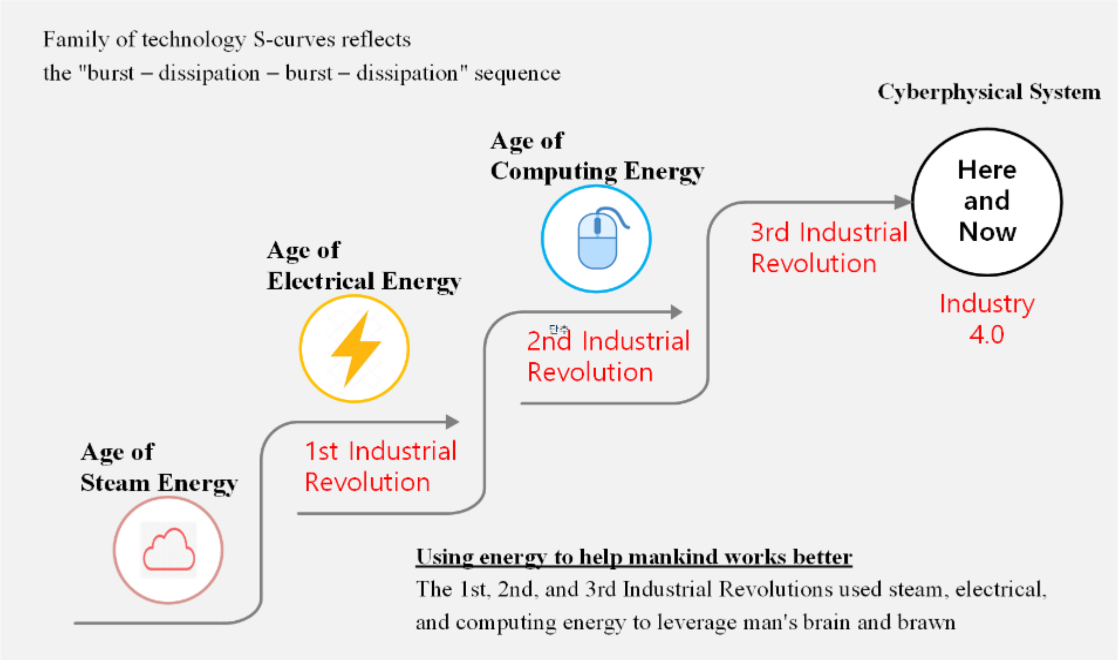
In particular, human privacy is very severe issues in the cyber physical space. Until now, humans are not connected to each other in cyberspace like the IoT, but in the smart society the connect is unpredictably high level.

Therefore, this paper aims to show the reasonable and practical plan how to prepare for coexistence between AI and real human being.

2. Human Civilization Transition Process and the Advent of AI-Based Smart Society

In the early era of human civilization, human being have got a labor-based agricultural way, steam based, electricity based, and computing/internet based industrial society[7][8], as shown in the following <Figure 1>.

Figure 1. Industrial revolution's characteristics: mechanization, mass production, automation, and adaptiveness.


Family of technology S-curves reflects
the "burst – dissipation – burst – dissipation" sequence
Cyberphysical System
Age of Computing Energy
Here and Now
3rd Industrial Revolution
Age of Electrical Energy
Industry 4.0
2nd Industrial Revolution
Age of Steam Energy
1st Industrial Revolution
Using energy to help mankind works better
The 1st, 2nd, and 3rd Industrial Revolutions used steam, electrical, and computing energy to leverage man's brain and brawn

In 1784, the First Industrial Revolution was based on a technological innovation using steam engines which led to consume goods and light mechanized production process. The Second

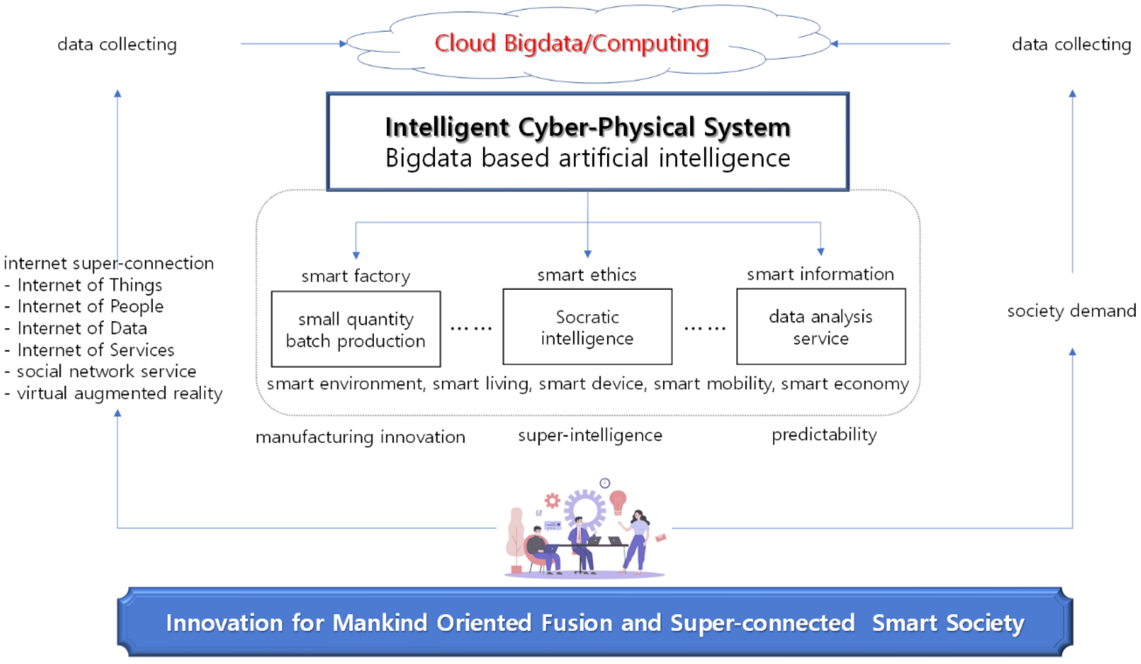
Industrial Revolution began in the 1870s. During the period, heavy chemicals, such as automation of conveyor belts and internal combustion engines of petroleum power had been implemented by the electrical energy. It brought the vitalization of industry.

The Third Industrial Revolution began in the 1960s. During this period, the main body was digitalization and automation. It comprises electronics, communication, information, and semiconductor technologies, computing, and information industry. They all had been moved by internet.

The biggest problem in the Third Industrial Revolution scenario was "the pursuit of cooperation and community" to maintain employment due to the end of the market and the decline of jobs. The Fourth Industrial Revolution is an extension of the Third Industrial Revolution. The main icon of this era is super-connectivity that connects to the network in whole dimensions. Convergence technology to connect all agents and communities is based on cloud computing, big data, AI which breaks down boundaries[7][8].

Today's civilized society is not only the extension of the Third Industrial Revolution, which is the core of the information industry. Also the scope and system shows the advent of a marked change of the Fourth Industrial Revolution. Unlike the previous industrial revolutions, it was not linear, but exponentially rapid advance[8]. And it affects most of industries. The breadth and depth of such change could lead to changes in all fields such as services, people, and information. All areas are connected like the IoT to process and store, and time and time. We have created an environment that can be controlled in real time regardless of the places, as shown in the following <Figure 2>.

Figure 2. Innovation for mankind oriented fusion and super-connected smart society.



Charles Levy and David Wong talked about "smart society" at Big Innovation 2014. They explained that a smart society leverages the power and the potential of technology to make human beings more productive; to allow us to focus our resources on activities and relationships that matter; and ultimately to improve health, wellbeing and the quality of life. And they said that a smart society could be defined "One that successfully harnesses the potential of digital technology and connected devices and the use of digital networks to improve people's lives."[11]. However, this definition goes beyond improving human life. Human factors such as participation and cooperation in the society are neglected.

A smart society is a society that is empowered and generally focuses on the scientific and technological aspects. There is a high probability that humans can be excluded. Therefore, it is necessary to consider future smart society based on human centered ethics. In social systems, the data feedback loop with detection function is an essential element for sustainability.

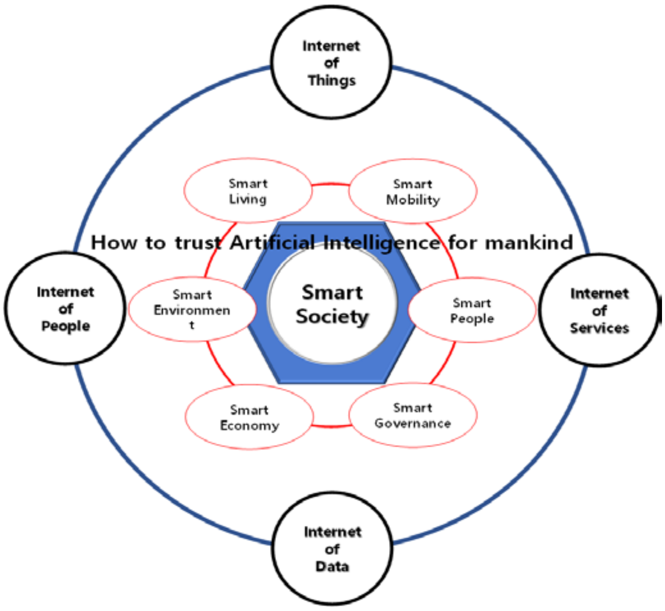
Future-oriented smart society is a governance city that helps improve the existing information society. The data from the system is integrated into a controllable system with a feedback loop function, It is possible to generate and utilize knowledge to be more successful with feedback control.

Currently, the case where AI algorithm technology based on big data is applied is an autonomous vehicle, drones, translation, investment, and games. This is a vast exponential development of computing power and the availability of vast data.

As mentioned above, the Fourth Industrial Revolution got the physical, spatial, and biologically fusion. As shown in <Figure 3>, a smart society can give challenge and opportunities to human being, at the same time it can give harmfulness also[11].

Smart society is based on the smart technology for everything such as humans, systems, information, and services, i.e. IoT and Big Data. Technology and AI are applied to easily solve difficult social problems. It can bring a happy society in the part of work, economy, economy, industry, administration. It can be said that it is a new revolution that transforms society as a whole, down to culture[11], as shown in the following <Figure 3>.

Figure 3. Core components and smart internet connection for smart society.



The Fourth Industrial Revolution is particularly prone to disrupting the labor market. It can lead to skepticism inequality and increase conflict between machines and humans. But smart factory system adopted to improve production and technologies moved to increase efficiency of management.

In the near future, competent smart humans will play a role as an important element of production. As a result, the wage system changes according to the importance of human roles, and social tensions can be increased. Such an inequality must be regarded as a social concern in the era of the 4th industrial revolution.

Due to the dissemination of digital technology and the dynamics of information sharing represented by social media, human being can get the unsatisfied and use social media platforms to connect and distribute information. In a smart society, this kind of interaction can be served for understanding and cohesion. It offers opportunities, but can create and spread unrealistic expectations of success. It can also provide an opportunity for extreme ideas and ideologies to spread.

In late 2009, as Apple's iPhone introduced, various new smart devices such as network service, cloud computing, location-based service changed collection-storage-utilization method in the name of Smart Big Bang. Nowadays personal identifiable information such as name, social security number, location information made by CCTV have been exposed. In the smart society, most important thing is to protect private information. Tracking and sharing information is an important part of the new connection, and data control. The inner impact of loss of power will be further discussed in the near future[9][10][13].

The revolution driven by biotechnology and AI needs, through overcoming limitations, to redefine the meaning of humans and ethical boundaries. And it is necessary to introduce AI ethical technology. The direction should focus on the importance of human being.

In thinking of robotization of human, the Fourth Industrial Revolution can deprive the mind and spirit of human being. But it has the potentiality to improve a shared sense of destiny, moral consciousness. That could be duty of mankind. In a future-oriented smart society, if AI cannot be controlled, those who control the AI should be ready to keep the sincere and solid human sensitivity and creativity that AI cannot have.

3. The Role of AI Smart Technology and Ethical Issues

The AI technology of future smart society has received the request such as some ethical principles. AI could not be a kind of new utility, but strong power such as smart agency. AI4People(artificial intelligence for people) should be scheduled to give help to mankind.

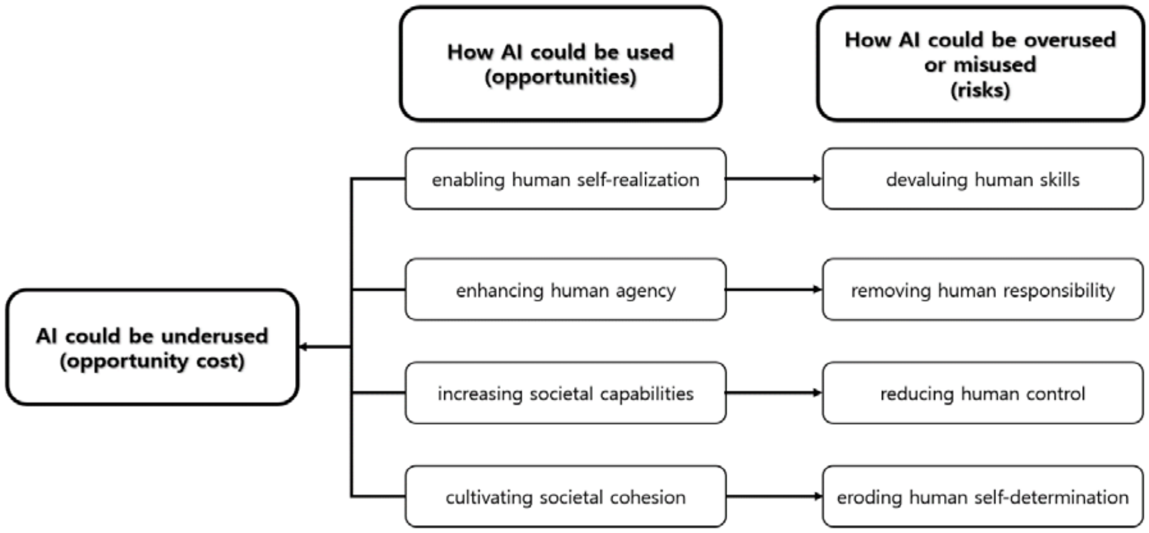
AI technologies provide opportunities and opportunities to protect human dignity and grow humans. Opportunities and risks should be moderated through nurturing human nature and potential[9].

- Autonomous self-realization
- Human representation
- Personal and social skill
- Social cohesion: to interact with each other in the world

Human being have got the fear, ignorance, false concerns, or excessive reactions. Because of these, human will not be able to make the most of AI smart technology that can be broadly and fully explained. For example, human can face the excessive or poorly regulated, underinvested, or genetically modified crops. Most of these risks are unintended consequences. It is caused by harm and is usually associated with false will. But we have the wrong incentives, greed, hostile geopolitics, or malicious intentions. The risks associated with abuse or deliberate misuse also need to be considered.

As mentioned above, the social development can get malicious impact by smart AI. Here <Figure 4> summarizes the risk and the opportunity cost of using artificial intelligence.

Figure 4. Four core opportunities by AI, four corresponding risks, and the opportunity cost of underusing AI.

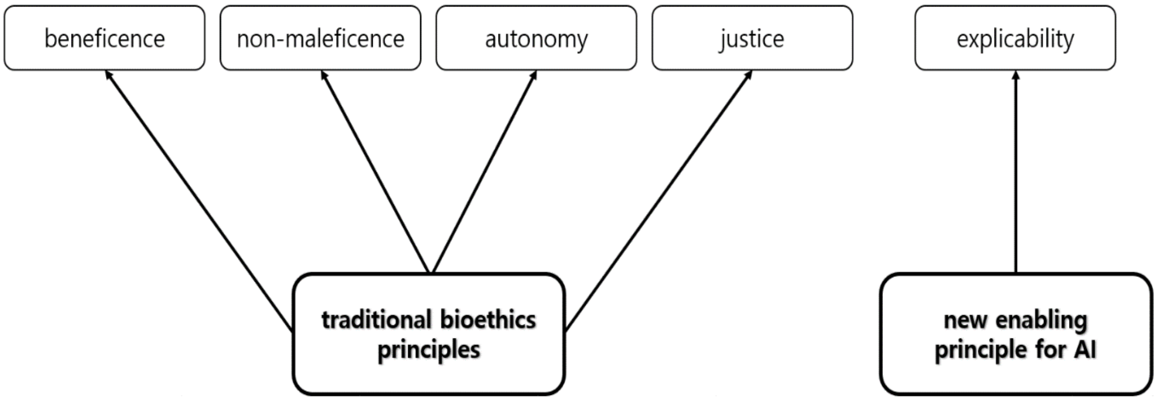

How AI could be used (opportunities)
How AI could be overused or misused (risks)
AI could be underused (opportunity cost)
enabling human self-realization
devaluing human skills
enhancing human agency
removing human responsibility
increasing societal capabilities
reducing human control
cultivating societal cohesion
eroding human self-determination

The working principle of AI is mostly invisible and understood only by high level experts. Until now, efforts for bioethics has missed the crucial part. So it is necessary to establish ethical standards for smart AI as follows[9][12].

- Beneficence: promoting well-being, preserving dignity, and sustaining the earth
- Non-maleficence: privacy, security and capability caution
- Autonomy: the power to make decision
- Justice: promoting prosperity and preserving solidarity
- Explicability: enabling the other principles through intelligibility and accountability

After confirming these principles, it is necessary to construct an AI ethics frame as shown in the following <Figure 5>.

Figure 5. Ethical framework for AI which is formed of 4 traditional principles and one new principle.


beneficence
non-maleficence
autonomy
justice
explicability
traditional bioethics principles
new enabling principle for AI

According to Schwab, AI should have the contextual, emotional, inspired, and physical considerations, further more Oosthuizen suggested to contain enterpreneurial, strategic, transdisciplinary, ecosystem, ethical considerations. <Table 1> follows their suggestions[12].

Table 1. Fourth industrial revolution intelligence framework in smart society.

Intelligence type	Key	Contents
Contextual intelligence	Mind	How we understand and apply our knowledge
Emotional intelligence	Heart	How we process and integrate our thoughts and feelings and relate to ourselves and to one another
Inspired intelligence	Soul	How we use a sense of individual and shared purpose, trust, and other virtues to effect change and act towards the coming good
Physical intelligence	Body	How we cultivate and maintain our personal health and well-being and that of those around us to be able to apply the energy required for both individual and systems transformation
Entrepreneurial intelligence	Disposition	How we recognize opportunity through synthesis of the whole and creativity combine resources
Strategic intelligence	Orientation	How we adapt to changing environments; gather, examine and disseminate intelligence of strategic valve
Transdisciplinary intelligence	Perspective	How we understand a system in relation to its larger environment, relationships and connections, integrating the information from separate disciplines
Ecosystem intelligence	Coalescence	How we grow and develop within the setting of the system of relationships that form our environmental factors have on us, and how we impact one another and our environment
Socratic intelligence	Philosophy	How we analyze ideas in terms of their opposites with the objective of creating a more enlightened synthesis

4. Conclusion

The present is the return of the past and a stopover to the future. As if there is no present without the past, the future without present couldn't be thought.

There is no future without. The smart society of the future vaguely approaching is a revolution inside whirlpools. However, through impartial reason based thinking, if human evaluate the current situation well, prepare for it, the smart society never be pessimistic it is feasible.

The smart society of the future should be a human-centered convergence and hyperconnected smart society due to the complexity of science and technology. The preparative efforts for smart society have the following sequence: visioning exercise, priority setting and selection of trends, implementation plan, implementation[14] as shown in the following <Table 2>.

Table 2. How governments can use the future possibilities framework.

	Step 1: Visioning exercise	Step 2: Priority setting and selection of trends	Step 3: Implementation plan	Step 4: Implementation
Tools	▪ Foresight exercises to identify main global trends ▪ Research and analysis ▪ Societal conclusions (e.g. townhall meetings) at different levels of government	▪ Consultations across main government entities ▪ Review of Vision ▪ Analysis of market opportunities ▪ Analysis of global trends and how they affect the country	▪ Analysis of selected trends and barriers and enables to levering associated future opportunities ▪ System maps and stakeholder mappings for each trend and the most relevant and	▪ Implementation plans ▪ Feedback loops in institutions and process ▪ Regular meetings with key stakeholders ▪ Impact metrics to measure progress

		▪ Foresight exercises	affected sectors ▪ Discussions and consultation process for each of the trends	
Key success factors	▪ Buy-in from main societal groups ▪ Solid understanding of the mail global trends and how they will affect the country ▪ Forward Looking mindset within public service	▪ Buy-in from highest level of government ▪ Buy in from across the government	▪ Buy-in from main actors on each of the trends ▪ Solid knowledge of the countries strengths and weakness in each of the transformation trends	▪ Government capacity to change and introduce new public policies and process ▪ Public innovation capacity ▪ Ability to drive consensus around the main priorities ▪ Support from stakeholders under each trend
Outcome	Vision with strong societal buy-in that provides a strong direction for the country as a whole.	Priority trends and sectors that lay the base for future Development.	Development of action plans for each of the selected transformational trends.	An ongoing process that aims at building a society and economy focused on future opportunities.

The smart society of the future is expected to move complicatedly due to the convergence of cutting-edge technologies. Some technologies are even ahead of humans. In particular, neuroethics, law, autonomous vehicles, and medicine are showing the tendency etc.[15][16][17].

To prepare for the smart society, this study shows some ethical suggestions. First, the smart society needs a love for humans. Although humans are dignified beings, it is not because they are at next higher ranks under God, but because they are the most imperfect. Because of this reason, humans can sustain their lives by eating relatively large amounts of animals and plants. If so, what would such imperfect beings do with each other? The answer is just to care for each other with pity mind. All human being should have a feeling of compassion.

Second, it needs the retrospective understanding to the original state and convergence of source technology. If ordinary people who are leaving want to make very simple things, eg. paper or pencil, it is not easy to make them without any other person and machine. As human society becomes more complex, this need becomes more and more complex.

Third, it needs the necessity of various dialogues with people in different occupations. In future smart society, AI will do much of the work for humans. When this phenomenon accelerates, the fundamental autonomy of personal thinking would be limited. Therefore, human beings have a variety of conversations with people in different professions.

Fourth, it needs to secure self-sustainability. The smart society has highly dependency to the AI. However, if AI have met emergency situation, human beings will meet the difficulties. It is necessary to secure an alternative ability to solve the problem.

Fourth, it needs to get the transparency. Even secret get the appropriate transparency. Otherwise, in the case of black out of whole process and function of information, whole the smart society can be down to the non-smart ugly society.

Sixth, it needs to get the explainability. Smart society also should be democratic. If the unexplainable dictator get whole information, even he/she can govern well, it would be very dangerous. In the smart society, monopoly for information is the most harmful enemy.

Seventh, it needs to get unlimited responsibility. In agricultural society, no matter how much human do wrong, it will reach at just ruining the farming field in the year. But one fault in a smart society has a great impact. Especially the gateholder must have a great responsibility.

5. References

5.1. Journal articles

[1] Seo E & Park G. The Super-aged Multiculturalism in South Korea and the Necessity of Wearable AI Ethics. *Robotics & AI Ethics*, 5(2), 34-41 (2020). [Article]

[2] Lim Y & Lee M. Implications of Emotional Coaching and Integrated Art Therapy Teaching Method on Leadership Education in the AI Era. *Robotics & AI Ethics*, 5(2), 42-49 (2020). [Article]

[3] Lee J & Liang D. Suggestions for Using AI in Preparation for a Super-aging Society. *Robotics & AI Ethics*, 5(2), 57-64 (2020). [Article]

[4] Li Y & Park G. AI Ethics and Privacy Right. *Robotics & AI Ethics*, 5(2), 27-33 (2020). [Article]

[8] Kim TW & Park GY & Seo ES. IR4.0 and Ethical Tasks of AI. *Robotics & AI Ethics*, 4(2), 6-12 (2019). [Article]

[9] Floridi L. AI 4 People -An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations-. *Minds and Machines*, 28, 689-707 (2018).

[10] Lee J & Jang J & Park K. Protection of Personal Information in Cyberspace. *Information Society and Media*, 12, 51-62 (2007).

[13] Jeong J & Kim D. Study on Targets and Methods of Personal Information Risk in Smart Society. *Journal of Korean Association for Regional Information Society*, 16(3), 113-136 (2013).

5.2. Books

[11] Levy C & Wong D. Towards a Smart Society. Big Innovation Centre (2014).

[14] 2020 Centennial Lab. Future Possiblities Report 2020. Ministry of Cabinet Affairs and the Future. UAE (2020).

[15] Ashley KD. Artificial Intelligence and Legal Analytics. Cambridge University (2017).

[16] Joldersman CW. Neuroscience and Education: A Philosophical Appraisal. Taylor & Francis (2016).

[17] Levy N. Neuroethics. Cambridge University (2007).

5.3. Conference proceedings

[12] Oosthuizen J. The Determinants of Fourth Industrial Revolution Leasdership Dexterity: A Proposed Framework for 4IR-intelligence and Subsequent 4IR Leadership Development. 4th International Conference on Responsible Leadership (2017).

5.4. Additional References

[5] https://www.foreignaffairs.com/ (2015).

[6] http:// intelligence.weforum.org/ (2021).

[7] https://www.ictworks.org/ (2021).

6. Appendix

6.1. Authors contribution

	Initial name	Contribution
Lead Author	TK	-Set of concepts ☑
		-Design ☑
		-Getting results ☑
		-Analysis ☑
		-Make a significant contribution to collection ☑
		-Final approval of the paper ☑
Corresponding Author*	GP	-Corresponding ☑
		-Play a decisive role in modification ☑
		-Significant contributions to concepts, designs, practices, analysis and interpretation of data ☑
		-Participants in Drafting and Revising Papers ☑
		-Someone who can explain all aspects of the paper ☑