The Role of Higher Education for the Ethical AI Society

Ji-Hun Lim, Jeong-Eun Seo and Hun-Yeong Kwon
School of Cybersecurity, Korea University, Seoul, Republic of Korea
ljh89719@naver.com, sje5279@korea.ac.kr, khy0@korea.ac.kr

Abstract
AI is creating numerous ethical issues. In the upcoming AI society, we will have to deal with the Influence of AI correctly. To this end, ethical education is needed for experts who learn and develop AI technology. Above all, this is because their ethical capabilities lead to ethical AI technologies and services. Therefore, this study sought the role and direction of AI ethics education in the field of higher education that educates AI experts.

The Coming of the Age of AI: Why Ethics is so Important
AI technology has become a tool for the prosperity of mankind by increasing the welfare of the individual and the society as well as bringing innovations (AI HLEG 2019; Villani 2018).

However, it is also true that AI technology cannot be said to have an only positive influence on our society. Already, we are faced with various problems of bias, non-transparency and unclear division of responsibility in the diverse fields where AI is being adopted, causing unforeseen issues (Lim and Kwon 2021). When taking into consideration the potential harm that AI can cause, we will need the capacity to deal with the various ethical issues that such technology can bring in advance (Leslie 2019). As Paul Goodman put it, “technology is a branch of moral philosophy” (Martin and Schinzinger 2004) and more than ever, ethics has become more important for AI.

How to make an ethical AI society?

Efforts to Prepare Ethical Principles for AI
Since the establishment of the concept of AI in the 1950s, the significance of AI has been discussed and researched from the perspective of empiricism as a tool and technological mechanism to replace human beings derived from an ontological point of view (Samuel 1960; Wiener 1960). It can be said that the discussion on AI ethics was mostly developed from a philosophical standpoint (Coeckelbergh 2019; Dignum 2019; Müller 2020).

Recently, AI ethics has been approached more and more from a policy perspective, and how to fairly distribute the benefits from AI technology in our society (Swan 2017). This is the reason why ethical AI has become the hot topic of many governments that are trying to foster the new technology sector (West et al. 2019). According to a German non-profit organization, the AlgorithmWatch, over 160 principles on ethics for AI have been announced worldwide as of April 2020 (Lim and Kwon 2021). Considering that this research did not consider the various ethical programs on AI, we can easily surmise that countless groups and institutions worldwide are proposing various ethics concerning AI.

Limits of Ethical Principles and the Need for Education in Ethics
While a multitude of discussions is proliferating on how to incorporate ethical values into the soon to be realized AI society through the establishment of solid ethical principles, it is questionable that simply laying out a set of ethical principles will be enough to realize the ethical AI society. The word “principle” means “source” in Latin and can be understood largely as the justification, or the source of a given set of rules and regulations. Ethics based on such principles have been criticized as being too abstract and thin when trying to use them in critical moments of policy decision making. Moreover, they have been pointed out to provide only “minimal morality” (Kang 2007) and will most likely lead to “moral routinization” (Cambell 1994).

In the end, we need to consider how to correctly apply the ethical principles for AI that are being proposed by the global community to truly realize ethics for AI. All the members of society that develop and make use of the AI technology need to share the ethical values of the ethical AI society and live by them. And the best way to reach this goal...
is through education. Education will transform the educatee from the current natural state to a better and ideal state. The role of education has become ever more important to realize the ideal ethical AI society.

**Methodology for the Education of Ethics for AI**

Recent AI policies differentiate the developers, suppliers, and users of AI (European Parliament 2020). The developers are the technical experts (or groups) that build and implement the AI systems. The suppliers have the decision-making power to distribute the AI systems developed and the users are the actual end-users who make use of the said AI systems (Stahl 2021). Ethics for AI should also take such classification into account.

**Ethical Engineering Education for the Developer**

The developer shares the responsibility for the end product from the initial stage of its first development to its use and its quality control management as well. The developer needs to possess a high degree of technical competency to minimize technical errors and risks from a safety and prevention point of view as well as develop the ethical aptitude needed. The ethical education for such AI developers and experts can be approached as part of ethical education for engineering. To provide the necessary knowledge and intellectual understanding for becoming a socially responsible engineer, the ethical education for engineering needs to promote the understanding of the ethical values of the influences that science and technology have on our society (Harris 2008; Pritchard 1998; Crawford-Brown 1997; Herkert 2001).

**Ethical Management Education for the Supplier**

The supplier must provide the goods produced by the developer with the public interest in mind by striking a reasonable balance between commercial profit and the public good (Lim 2017). Moreover, the supplier needs to take responsibility for the various side effects and other social issues that arise from the distribution and proliferation of AI services and goods that have been distributed by the supplier (Carroll and Buchholz 2011). In other words, the ethical education for the supplier needs to address various issues associated with the production, sales, and distribution of services and goods that make use of the AI technology from the ethical business management perspective.

**Social Ethical Education for the User**

The user of AI technology must not use the various AI technology with malicious intent and fully understand that the user must take responsibility when using the services and goods based on AI technology. The user must not use such technology beyond the given limits or arbitrarily manipulate it to infringe upon the rights of others. While the general population has become aware of the benefits and convenience provided by AI technology, researchers suggest that the general population has limited understanding of its negative aspect, the ethical issues (Ko and Leem 2021). Active promotion and education are needed to ensure that the general population is aware of the ethical issues when using AI technology. This can be achieved as part of education for digital citizenship. As a citizen of the digital society, the user must be familiar with the exact principles of conduct for using AI goods and services as well as have a strong understanding of the ethical use of AI technology.

**The Importance of Ethical Education for the AI Developer**

To realize the ethical AI society, we need ethical education for all the subjects mentioned previously. However, in considering the nature of the AI technology, the authors of this paper seek to concentrate on the ethics needed for the developer.

The operation of AI is based on its programmed learning algorithm and by itself, it does not guarantee attributes such as fairness, neutrality, etc. As such, it is up to the people involved in its creation and operation when it comes to the ethical use of AI technology. Ethical education of all the users of AI technology is the best method for the ethical use of the AI technology, but even with such education, we cannot expect every person to exhibit ethical behaviors, and leaving everything up to the end-user may make it difficult to properly respond to potential risks and threats that can come from the AI technology (Reich 2021).

We need to understand the attributes of the AI technology and have the capability to control it to fundamentally respond to such problems; it is obvious that the developer who designed the learning and decision-making process of the AI is the perfect person to deal with such issues. Providing ethics courses for AI in undergraduate or graduate curriculum for AI developers could be a good beginning (Lee 2020).

**Types of Ethical AI Education in the Higher Education**

**Class Knowledge Transfer Type Curriculum**

Many colleges worldwide have created various AI ethics classes and are making the effort to promote the ethical values for the AI. For example, Stanford University in the U.S. opened the “Artificial Intelligence - Philosophy, Ethics, and Impact” class as a curriculum for discussion on AI and ethics in 2014 (Stanford 2014). This program included knowledge transfer on the philosophy and history of AI as well as the ethical, social, legal, and economical impact of the AI technology. Since then, many colleges have created similar curriculum.

In Korea and China, many colleges include AI ethics curriculum. Theses classes contain the history of AI ethics and the basics of ethics in general focused on the current issues in AI ethics. And deals with the impact of AI on human and machines as well as diverse subjects on how to prevent
potential risks from the use of AI technology (Cho 2021; UCAS 2018-2022).

Practical Scientific-Philosophical Convergence Training Curriculum

Unlike the curriculum in the previous section, there are various types of technical classes that incorporate the philosophical aspects of the AI technology as a scientific-philosophical convergence training curriculum. For example, the Harvard University provides AI ethics education by encouraging the study of ethics within their computer science classes by making use of the Embedded EthiCS program developed for that specific purpose. Specifically, the goal is to make the students understand that ‘the work done by computer scientists include ethical reasoning’ through the practical scientific-philosophical convergence training based on the Embedded EthiCS that is designed to provide ethical reasoning (Grosz et al. 2019). Thus, the current computer science classes for AI, machine learning, O/S, programming languages and more include philosophical contents.

The Way Forward for Ethical AI Education in Higher Education

As can be seen from the examples above, the goal of ethical education for computer science experts, specifically the AI developer, is to provide the knowledge and understanding to strengthen the ethical capability needed for the developer to become a socially responsible technical expert. The social responsibilities of the technical experts were discussed in detail below.

Convergence Education for Technology and Ethics

We need educational classes on how to understand what ethical elements are part of which specific AI technology. Such a curriculum cannot be provided by philosophers with little technical knowledge and understanding but only with the collaboration between the technical expert and the philosopher with a deep understanding of ethical values. And such a curriculum should not be a separate and additional course but rather, as we have seen in the Harvard University case, provided as a convergence of education and training where ethical education is an intrinsic part of computer science courses about the AI technology.

Design Training for Autonomous Code of Ethics

Ethics for the AI developer is most akin to the professional code of conduct. Such code depends on autonomous regulation and mutual checks and balances between the expert groups most capable of taking responsibility for technology adoption in the most ethical manner as well predicting the consequences of such adoption (Michael 1991). There exist plenty of such cases in various professions like medical doctors, lawyers, and soldiers where voluntary code of ethical conduct is discussed by a few expert groups (Lim 2017). The AI developer community also needs such voluntary regulation based on a code of ethical conduct. Rather than blindly following ideal regulations made by the government or private companies, we need to strengthen the capabilities of the AI developers to take the matter into their own hands and develop their own set of regulations and guidelines. At this time, they will be able to refer to existing expert ethics codes such as “ACM Code of Ethics and Professional Conduct” and “IEEE Code of Conduct”. In addition, it is necessary to derive what development is necessary for these codes from the perspective of AI.

Ethical Dilemma Education

Ethics is comprised of three elements: cognition, emotion, and behavior. While these three elements seem separate, they are only so in concept. All the three elements come into play together and are interconnected thereby requiring the goal of ethics education to take these three elements that culminate into ethical behavior. Therefore, ethical education for AI needs to train computer scientists to make ethical decisions and put them into action. Recently, AI technology has been closely integrated into business models, thus requiring a sense of balance between demand and motivation from the developer (Wayner 2014). The dilemma discussion methodology may help to resolve this issue. The dilemma discussion methodology is a methodology based on the theory of cognitive development developed by Piaget (1932) and Kohlberg (1958) that is comprised of the following stages: 1) recognition of an ethical problem, 2) potential position statement, 3) moral reasoning and discussions, and 4) discussion of a position statement (Galbreith and Jones 1976). Through such training, the AI developer must internalize the ability to deal with countless moral dilemmas that he will face in the field through constant and unending ethical considerations.

Conclusion

For the ethical AI society, we need ethical domain experts who can analyze social changes and recognize the moral issues and execute ethical solutions to resolve these issues. Because highly specialized experts such as AI developers have a practical monopoly on knowledge and technical operation of their domain, it is extremely difficult for the public to discover and control their unethical conduct. Therefore, a high degree of ethics and morality proportional to their expertise and autonomous freedom in their domain is needed by the AI developer. Seen from this perspective, this paper is arguing that the importance of higher education institutions such as the
AI HLEG. 2019. Ethics guidelines for trustworthy AI, European Commission, Brussels.

Campbell, C. S. 1994. Experience and Moral Life: A Phenomenological Approach to Bioethics. In: Dubose, E. R.; Hamel, R. P.; and O’Connell, L. J. ed. A Matter of Principles: A Ferment in U.S. Bioethics. Valley Forge, PA: Trinity Press International.

Carroll, A. B., and Buchholtz, A. K. 2011. Business and society: Ethics and stakeholder management 7th Edition. Mason, OH: South-Western College Publishing.

Cho, J. S. 2021. Analysis the Ethical Consciousness on Artificial intelligence and Educational Demands of University students. The Digital Ethics 5(2): 29-38.

Coeckelbergh, M. 2019. Artificial Intelligence: some ethical issues and regulatory challenges. Technology and regulation 2019: 31-34.

Crawford-Brown, D. 1997. Virtue as the basis of engineering ethics. Science and Engineering Ethics 3(3): 481-489.

Dignum, V. 2019. Responsible artificial intelligence: how to develop and use AI in a responsible way. Cham, Switzerland: Springer Nature Switzerland AG.

European Parliament. 2020. Draft report with recommendations to the Commission on a framework of ethical aspects of artificial intelligence, robotics and related technologies(2020/2012(INL)). European Parliament, Committee on Legal Affairs.

Galbraith, R. E., and Jones, T. M. 1976. Moral reasoning: A teaching handbook for adapting Kohlberg to the classroom. Minneapolis, Minnesota: Greenhaven Press.

Grosz, B. J.; Grant, D. G.; Vredenburgh, K.; Behrends, J.; Hu, L.; Simmons, A.; and Waldoen, J. 2019. Embedded EthiCS: integrating ethics across CS education. Communications of the ACM 62(8): 54-61.

Harris, C. Jr. 2008. The good engineer: Giving virtue its due in engineering ethics. Science and Engineering Ethics 14: 153-164.

Herbert, J. 2001. Future directions in engineering ethics research: Microethics, macroethics and the rope of professional societies. Science and Engineering Ethics 7(3): 403-414.

Jongbloed, B.; Enders, J.; and Salerno, C. 2008. Higher education and its communities: Interconnections, interdependencies and a research agenda. High Education 56: 303–324.

Kang, J. H. 2007. An Ethical Examination of Research Ethics Guidelines. CHUL HAK SA SANG : Journal of Philosophical Ideas 24: 111-142.

Ko, Y. H., and Leem, C. S. 2021. The Influence of AI Technology Acceptance and Ethical Awareness towards Intention to Use. Journal of Digital Convergence 19(3): 217-225.

Lee, S. G. 2020. Improvement Tasks for Ethical Use of Artificial Intelligence. Issues and Perspectives (1759): 1-4.

Leslie, D. 2019. Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector. The Alan Turing Institute. Available at SSRN 3403301.

Lim, J. H., and Kwon, H. Y. 2021. A Study on the Modeling of Major Factors for the Principles of AI Ethics, In Proceedings of DG. 02021: The 22nd Annual International Conference on Digital Government Research, 208-218. Omaha, NE, USA.: Association for Computing Machinery.

Lim, S. S. 2017. Moral Education in the Age of Artificial Intelligence: From the Perspective of Consumer Ethics. Journal of Korean Ethics Studies 117: 89-116.

Martin, M. W. and Schinzinger, R. 2004. Ethics in Engineering. New York, NY: McGraw-Hill Education.

Michael D. 1991. Thinking Like an Engineer: The Place of a Code of Ethics in the Practice of a Profession. Philosophy and Public Affairs 20(2): 150-167.

Müller, V. C. 2020. Ethics of artificial intelligence and robotics. Metaphysics Research Lab, Stanford University, Stanford, CA.

Pritchard, M. S. 1998. Professional responsibility: Focusing on the exemplary. Science and Engineering Ethics 4: 215-233.

Reich, R. 2021. Embedded EthiCS for a Better Tomorrow. https://blog.ncsoft.com/ai-framework-ep06-210617/

Samuel, A. L. 1960. Some moral and technical consequences of automation-a refutation. Science 132(3429): 741-742.

Stahl, B. C. 2021. Ethical Issues of AI. In: Artificial Intelligence for a Better Future. Cham, Switzerland: Springer.

Stanford. 2014. CS122: Artificial Intelligence - Philosophy, Ethics, and Impact. https://web.stanford.edu/class/cs122/

Swan, M. 2017. Is Technological Unemployment Real? An Assessment and a Plea for Abundance Economics. In: Kevin L. G., and James J. H. ed. Surviving the Machine Age. New York, NY.: Springer International Publishing.

University of Chinese Academy of Sciences(UCAS). 2018-2022. Philosophy and Ethics of Artificial Intelligence. https://bit.ia.ac.cn/pea/

Villani, C. 2018. For a Meaningful Artificial Intelligence—Towards a French and European Strategy. AI FOR HUMANITY.

Wayner, P. 2014. 12 ethical dilemmas gnawing at developers today. InfoWorld, 2014.04.21., https://www.infoworld.com/article/2607452/12-ethical-dilemmas-gnawing-at-developers-today.html

West, S. M.; Whittaker, M.; and Crawford, K. 2019. Discriminating systems: Gender, Race and Power in AI. AlNow Institute, New York University, Microsoft Research, New York, NY.

Wiener, N. 1960. Some moral and technical consequences of automation. Science 131(3410): 1355-1358.