

## 실천력 강화를 위한 인공지능 윤리 교육 모델

배진아<sup>1,2</sup>, 이정훈<sup>1,2</sup>, 조정원<sup>3,4\*</sup>

<sup>1</sup>제주대학교 대학원 컴퓨터교육전공 대학원생, <sup>2</sup>제주대학교 지능소프트웨어교육연구소 연구원  
<sup>3</sup>제주대학교 컴퓨터교육과 교수, <sup>4</sup>제주대학교 지능소프트웨어교육연구소 소장

## An Artificial Intelligence Ethics Education Model for Practical Power Strength

Jinah Bae<sup>1,2</sup>, Jeonghun Lee<sup>1,2</sup>, Jungwon Cho<sup>3,4\*</sup>

<sup>1</sup>Graduate Student, Major in Computer Education, Graduate School, Jeju National University

<sup>2</sup>Researcher, Intelligent Software Education Research Institute, Jeju National University

<sup>3</sup>Professor, Department of Computer Education, Jeju National University

<sup>4</sup>Director, Intelligent Software Education Research Institute, Jeju National University

**요약** 인공지능 기술로 인한 사회-윤리적 문제 사례가 발생하면서 인공지능의 위험과 부작용에 대한 사회적 관심과 함께 인공지능 윤리가 주목받고 있다. 인공지능 윤리는 알고, 느끼는 것에 그치는 것이 아니라 행동과 실천으로 이루어져야 한다. 이에 본 논문은 인공지능 윤리의 실천력을 강화하기 위한 인공지능 윤리 교육 모델을 제안하고자 한다. 인공지능 윤리 교육 모델은 선행 연구 분석을 통해 교육목표와 인공지능을 이용한 문제해결 프로세스를 도출하고, 실천력 강화를 위한 교수학습방법을 적용하였으며 기존에 제안된 인공지능 교육 모델과 비교 분석하여 그 차이를 도출하였다. 본 논문에서 제안하는 인공지능 윤리 교육 모델은 컴퓨팅 사고력 함양과 인공지능 윤리의 실천력 강화를 목표로 한다. 이를 위해 인공지능을 이용한 문제해결 프로세스를 6단계로 제안하고, 인공지능 특성을 반영한 인공지능 윤리요소를 도출하여 문제해결 프로세스에 적용하였다. 또한, 인공지능 윤리 의식에 대한 사전-사후 평가와 과정 평가를 통해 인공지능 윤리 기준을 무의식적으로 확인하게 하고, 학습자 중심의 교수학습방법을 적용하여 학습자의 윤리 실천을 습관화하도록 설계하였다. 본 연구를 통해 개발된 인공지능 윤리 교육 모델이 컴퓨팅 사고력을 함양하고, 인공지능 윤리가 실천으로 이어지는 인공지능 교육이 될 수 있을 것으로 기대한다.

**키워드** : 인공지능 교육, 인공지능 윤리 교육, 컴퓨팅 사고력, 인공지능 교육 모델, 정보 교육

**Abstract** As cases of social and ethical problems caused by artificial intelligence technology have occurred, artificial intelligence ethics are drawing attention along with social interest in the risks and side effects of artificial intelligence. Artificial intelligence ethics should not just be known and felt, but should be actionable and practiced. Therefore, this study proposes an artificial intelligence ethics education model to strengthen the practical ability of artificial intelligence ethics. The artificial intelligence ethics education model derived educational goals and problem-solving processes using artificial intelligence through existing research analysis, applied teaching and learning methods to strengthen practical skills, and compared and analyzed the existing artificial intelligence education model. The artificial intelligence ethics education model proposed in this paper aims to cultivate computing thinking skills and strengthen the practical ability of artificial intelligence ethics. To this end, the problem-solving process using artificial intelligence was presented in six stages, and artificial intelligence ethical factors reflecting the characteristics of artificial intelligence were derived and applied to the problem-solving process. In addition, it was designed to unconsciously check the ethical standards of artificial intelligence through pre- and post-evaluation of artificial intelligence ethics and apply learner-centered education and learning methods to make learners' ethical practices a habit. The artificial intelligence ethics education model developed through this study is expected to be artificial intelligence education that leads to practice by developing computing thinking skills.

**Key Words** : AI education, AI ethics, AI ethics education, AI education model, AI ethical practice

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea(NRF-2021S1A5C2A04088646).

\*Corresponding Author : Jungwon Cho(jwcho@jejunu.ac.kr)

Received April 26, 2022

Revised May 16, 2022

Accepted May 20, 2022

Published May 28, 2022

## 1. 서론

지능정보사회에서는 인공지능, 빅데이터 등과 같은 기술이 사회 변화를 주도하고 발전을 가속화 한다. 이러한 변화를 이끌어 나갈 수 있는 역량을 가진 인재 양성을 목표로 국내외에서 인공지능 교육을 강화하고 있다.

이러한 기술은 인간의 한계를 극복하고 생활의 편리함을 제공하지만, 인공지능 기술 개발에 사용된 데이터나 알고리즘의 편향과 같은 문제도 함께 발생한다. 이처럼 인공지능 기술 발전은 긍정적 요소와 더불어 부정적 요소가 존재한다.

인공지능은 데이터를 기반으로 학습하는 과정에서 규칙의 역할을 하는 모델이 구축되게 되고, 그 과정에서 블랙박스와 같이 인간이 통제하거나 예측하기 어려워 데이터 편향이나 개발자가 의도하지 않은 결과로 이어질 위험이 있다[1]. 그 예로 자율주행 자동차가 특정 색으로 도색된 차를 인지하지 못하여 사고가 발생하거나, 특정 네티즌들에 의하여 왜곡된 내용을 중심으로 학습된 인공지능이 편견에 치우친 발언을 하여 세간의 비난 대상이 되기도 하였다[2]. 이는 인공지능의 특성을 고려하지 않았거나 인공지능을 이용하여 서비스를 개발하는 과정에서 인공지능으로 인해 발생할 수 있는 윤리적 문제를 충분히 고려하지 않았기에 나타난 결과이다. 이러한 윤리적 문제를 최소화하기 위해서는 인공지능 교육과 함께 인공지능 윤리 교육이 함께 이루어져야 한다.

윤리 교육의 목적은 윤리적 요소에 대한 개념적 교육을 넘어 윤리의 적용 방안을 탐색하고 계획하여 실천하는 데 중점을 두는 교육으로 진행되어야 한다[3]. 실천력을 높일 수 있는 교육이 되기 위해서는 인공지능 교육과 인공지능 윤리 교육이 통합적으로 함께 이루어져야 한다. 따라서, 본 논문에서는 컴퓨팅 사고력 함양과 인공지능 윤리의 실천력 강화를 목표로, 인공지능을 이용한 문제해결 과정을 담은 6단계 프로세스를 개발하였다. 6단계 프로세스 각각에 고려되어야 할 인공지능 윤리 요소를 도출하였다.

이를 적용하여 교육 내용을 구성하고, 컴퓨팅 사고력 함양을 위한 교수학습 방법을 제안하였다. 마지막으로 인공지능 윤리 교육의 효과성 진단을 위한 사전·사후 평가를 개발하여 적용하고, 각 프로세스에서 진행되는 과정 평가를 통해 인공지능 윤리 실천력을 강화할 수 있는 인공지능 윤리 교육 모델을 개발하였다.

## 2. 선행연구

### 2.1 인공지능 윤리

지능정보사회를 주도하는 인공지능 기술은 데이터와 결합하여 경제·사회 전 분야에서 디지털 전환을 촉진하고 있다. 하지만 인공지능 기술 확산으로 인한 다양한 사회 윤리적 문제 사례가 계속해서 발생하고 있어, 인공지능으로 인한 사회적 영향에 대한 사회적 관심과 우려가 이어지고 있다[4].

인공지능 핵심 기술인 딥러닝(Deep Learning)을 중심으로 한 기계학습(Machine Learning)은 데이터를 기반으로 학습하며 지능적인 역할을 하게 된다. 이 과정에서 인간은 인공지능이 어떻게 지능적인 결과를 도출하는지 알 수 없다[5]. 이러한 인공지능의 특성으로 인해 인공지능 학습 데이터에 따른 편향이나 개발자의 의도하지 않은 결과를 도출할 수 있다. 또한, 인공지능의 무분별한 활용은 딥페이크(Deepfake), 개인정보 및 사생활 침해 등 인공지능 기술의 오·남용과 사생활 침해로 인한 복합적 문제를 유발하는 잠재적 위험이 되기도 한다[6].

한상기(2016)는 인공지능 기술의 발전은 사회 전반에 긍정적인 역할을 하지만, 이로 인한 역기능과 윤리적인 문제들을 야기할 가능성을 가지고 있다고 주장하였다. 이러한 문제에 대처하기 위해 인공지능 연구자, 개발자 등이 인공지능으로 인한 윤리문제에 관해 관심을 기울이는 것이 중요하다고 하였다[7].

인공지능 기술의 발전은 우리 삶을 편리하게 해주고 대부분 바람직한 방향으로 전환해 줄 것으로 기대하고 있지만, 이전에는 고려하지 않았던 문제들이 새롭게 나타나고 있으며, 이로 인해 인공지능 윤리 문제에 대한 보다 적극적인 연구가 필요하게 되었다[8]. 이에 따라 인공지능 윤리 의식 함양을 위한 인공지능 윤리 기준 마련과 함께 인공지능 윤리 교육의 중요성이 커지고 있다.

### 2.2 국내외 인공지능 윤리 가이드라인 동향

우리나라를 포함한 인공지능 선도국은 인공지능 기술 발전에 따른 사회적 변화에 대비하기 위한 구체적인 방안을 Table 1과 같이 마련하고 있다[9-13]. 유럽연합, 미국, 일본, 유네스코 등 인공지능 선도국을 중심으로 'AI 윤리원칙·가이드라인'이 마련되고 있으며, 지속적으로 수정·보완이 진행되고 있다.

Table 1. Main Contents of overseas AI Ethics guidelines

Main Contents of overseas AI Ethics guidelines
<b>EURONs 13 Principles of Robot Ethics(EU. 2007)</b> ①human dignity and human rights ②equality, justice and equity ③Benefits and losses ④Respect for cultural diversity ⑤Prohibition of discrimination and stigmatization ⑥Autonomy and individual accountability ⑦Known consent ⑧Privacy ⑨Confidentiality ⑩Solidarity and cooperation ⑪Social responsibility ⑫Shared interests ⑬Responsibility to life on Earth[9][13]
<b>Some of the 23 Principles of Asiloma AI I (IEEE. 2016)</b> ①Safety ②Transparency of Failure ③Judicial Transparency ④Accountability ⑤Consistency of Values ⑥Human Values ⑦Privacy ⑧Freedom and Privacy ⑨Sharing of Profit ⑩Sharing of Prosperity ⑪Human Control ⑫Prevention of Subversion ⑬Artificial Intelligence Weapon Competition [9][10]
<b>AI Development Principles (Japan. 2018)</b> Promotion of AI Benefits and AI Networking (①Linkage), AI Risk Suppression (②Transparency ③Controllability ④Safety ⑤Security ⑥Privacy ⑦Ethics), Improving User Acceptance (⑧User support ⑨Responsibility)[9]
<b>Ethically Aligned Design (IEEE. 2019)</b> ①Human Rights ②Welfare ③Data Agency ④Efficiency ⑤Anonymity ⑥ Responsibility ⑦Right of Abuse ⑧Competence [12]
<b>UNESCO AI Ethics Recommendations (UNESCO. 2021)</b> Four values (①respect, protection and promotion of human rights, fundamental freedoms, and human dignity ②prosperity of the environment and ecosystems ③guarantee of diversity and inclusiveness ④establishment of a peaceful, just and interconnected society), Ten principles (① proportionality and harmlessness Gender ②Safety and Security ③Fairness and Nondiscrimination ④Sustainability ⑤Privacy and Data Protection ⑥ Human Supervision and Decision ⑦Transparency and Explainability ⑧ Responsibility and Duty ⑨Consciousness and Literacy ⑩Multilateral and Adjustable Governance and Collaboration[13]

Table 1에서 유럽연합은 2007년 ‘로봇 윤리 로드맵’을 통해 로봇 개발자를 위한 가이드로 사회적 책무, 차별과 낙인화 금지 등 총 13개의 원칙을 발표하였다. 미국의 IEEE는 2016년 ‘아실로마 인공지능 23원칙’에서 인공지능 연구 관련 쟁점, 윤리와 가치, 장기적 이슈 3가지 영역에서 23개의 원칙을 발표하며, 인간의 가치와 투명성, 책임성, 프라이버시, 개인정보 보호, 공공의 이익 등을 제시하였다. 또한, 2019년에는 지능적인 시스템(Autonomous and Intelligent Systems: A/IS)에 적용할 수 있는 8개의 일반원칙을 구성하여 제시하고, 인류학적, 정치적, 기술적 측면을 반영한 프레임 워크를 제공하기 위해 보편적 인류 가치(Universal Human Values), 정치적 자기 결정과 데이터 에이전시(Political Self-Determination and Data Agency), 기술적인 신뢰도(Technical Dependability) 측면으로 설계하였다. 일본은 2018년 ‘AI 개발 원칙’을 통해 개발자가 인공지능 개발 시 지켜야 할 9원칙을 편익증진, 위험 억제, 이용자 수용성 향상 3개의 기준에 따라 분류하였으며, 투명성, 책임성, 프라이버시, 보안을 강조하며 인공지능의 위험을 억제하는 데 중점을 두었다.

유네스코는 2021년 ‘유네스코 AI 윤리 권고안’을 통해 유네스코의 강점인 다수의 회원국, 다학제적 전문성, 인권에 대한 이해를 바탕으로 모든 인류, 지속 가능한 발전, 평화에 이바지할 수 있는 인공지능 개발을 위한 윤리 4대 가치와 10대 원칙을 발표하였다.

우리나라는 신뢰할 수 있는 인공지능을 실현하기 위해 정부와 공공기관은 물론 기업과 사용자 등 모두가 지켜야 할 인공지능 윤리를 Table 2와 같이 제시하고 있다.

Table 2. Main Contents of AI Ethics guidelines in S.KOREA

Main Contents of domestic AI Ethics guidelines
<b>Guideline for intelligent information society (NIA. 2018)</b> ①Publicity ②Accountability ③Control ④ Transparency [14]
<b>Human-centered ‘Artificial Intelligence (AI) Ethics Standards,’</b> 3 Basic Principles (①principle of human dignity ②principle of public good of society ③principle of purposefulness of technology) 10 Core Requirements (①guarantee of human rights ②protection of privacy ③respect for diversity ④non-infringement ⑤publicity ⑥solidarity ⑦data management ⑧accountability ⑨Safety ⑩Transparency ⑥)

Table 2를 살펴보면 한국지능정보사회진흥원(NIA)은 2018년 ‘지능정보사회 윤리 가이드라인’ 발표를 통해 지능정보기술의 특성에 기반을 둔 공통적인 원칙으로 공공성, 책무성, 통제성, 투명성과 같은 4가지 원칙에 대하여 정의하고 있다. 또한, 4개의 공통원칙 아래에 개발자, 공급자, 사용자 등 모두가 지켜야 하는 원칙을 정의하고 대상별로 지켜야 하는 세부지침도 명시하고 있다. 과학기술정보통신부가 발표한 ‘사람이 중심이 되는 「인공지능(AI) 윤리 기준」’에서는 인공지능의 사회 전반에 건전한 인공지능 의식 확산과 인공지능 윤리 교육 강화를 위한 전략으로 3대 원칙과 10대 실천 윤리를 제시하고 있다.

이와 같이 국내외 인공지능 윤리 가이드라인을 살펴보면 인공지능의 예측 불가능한 특성을 반영하고 기업, 개발자, 사용자 등 모든 사회 구성원을 고려하여 ‘다양성’, ‘투명성’, ‘책임-공공성’, ‘개인정보보호’와 같은 윤리 요소에 대해 공통으로 가이드라인을 제시한 것을 알 수 있다. 또한 인공지능의 ‘사회적 영향력’과 같은 파급력과, 알고리즘과 데이터의 중요성을 담은 지침이 주를 이루고 있다. 본 연구에서는 국내외 인공지능 윤리 가이드라인의 분석을 통해 공통 윤리 요소를 추출하고 인공지능의 사회적 영향력과 인공지능의 특성을 고려한 인공지능 윤리 요소를 도출하였다.

### 2.3 인공지능 윤리 교육

인공지능의 사회적 영향력이 커지면서 인공지능 교육과 함께 인공지능 윤리 교육의 중요성도 강조되고 있다. 미국의 AI4K12는 인공지능의 5가지 빅아이디어를 통해 인공지능이 가져올 사회적 영향(Social Impact) 내용을 포함하여 인공지능 윤리의 중요성을 강조하고 있으며, MIT Media Lab과 몬투어 교육청은 독립적인 인공지능 윤리 과목인 '디자인(AI 설계)과 가치(AI 윤리)'를 필수 과목으로 지정하고 있다. MIT 공대는 중학생 대상의 인공지능 윤리 교육인 '모두를 위한 인공지능' 과목도 개설하였다. 또한, AI4ALL은 'AI & Ethics Teaching Guide'에서 수학·프로그래밍·윤리학의 경험이 없어도 인공지능 윤리 교육이 가능한 교육과정을 마련하였고 그 예로 가상의 회사를 설립·운영하는 과정 전반에 인공지능 윤리 문제를 다루고 있다. 그 밖에도 국외의 인공지능 윤리와 관련한 교육을 Table 3에 정리하였다.

Table 3. Current status of AI Education

Current status of AI Education	
<b>Finland</b>	<ul style="list-style-type: none"> <li>• Online courses created by the University of Helsinki + Leaktor (Technology Advisory Company)</li> <li>• Introduction to AI. Building AI. Two parts.</li> <li>• The Social Impact Unit of AI: Algorithm bias, reliability, and social influence (change in the concept of personal information protection, change in occupation)</li> </ul>
<b>Canada</b>	<ul style="list-style-type: none"> <li>• ACTUA AI Training Guide.</li> <li>• Focusing on teachers rather than students.</li> <li>• AI is regarded as literacy and reflected in all subjects and subjects.</li> </ul>
<b>UK</b>	<ul style="list-style-type: none"> <li>• Recognizing that AI-related products are used.</li> <li>• The goal is to have the knowledge you need to become a confident user.</li> </ul>

핀란드는 인공지능 교육에 인공지능의 사회적 영향 단원을 편성하여 알고리즘 편향, 신뢰성, 사회적 영향력에 대해 2차시 과정으로 구성하였고, 캐나다의 'ACTUA AI 교육 안내서'에는 인공지능과 인공지능 윤리를 리터러시로 간주하고 전 과목·전 교과에 반영하고 있다. 영국의 경우 인공지능 사용자로서 갖추어야 할 역량과 윤리 의식을 교육하고 있다.

우리나라는 인공지능 시대의 교육정책 방향과 핵심 과제를 논의하며 2022 개정 교육과정을 통한 인공지능 교육과 인공지능 윤리 교육 강화 계획을 발표하였다 [15]. 또한 교육부(2020)는 '초·중고 인공지능 교육내용 기준'을 발표하고 인공지능 윤리와 관련하여 '인공지능의 사회적 영향' 영역을 Table 4와 같이 포함하였다.

Table 4. Criteria for Contents of AI Education in Elementary, Middle, and High schools

School -Level	Expectations
3-4	We look for artificial intelligence services and products that help us.
5-6	It can suggest what roles and powers to give to efficiently utilize AI. You can know how to use AI correctly and practice it in your daily life.
7-9	It can be explained by linking the relationship between the development of a AI and occupation. It is possible to explore cases where AI is not suitable for purpose or has been excessively used and to suggest preventive measures.

'인공지능의 사회적 영향'은 학교급에 따라 '인공지능 영향력', '인공지능 윤리'가 세부영역으로 나뉘고, 그에 따른 수행 기대와 성취기준이 포함되어 있다.

국내외에서 인공지능 윤리에 대한 관심이 커지면서 인공지능 윤리 교육에 대한 연구도 활발히 이루어지고 있다. 김지연(2021)은 인공지능 윤리 이슈를 분석하여 윤리 항목(공정성, 투명성, 책임성)을 선정하고, 교육과정의 관련성을 위해 성취과정을 분석하여 인공지능 윤리 내용과 융합할 수 있는 STEAM 교육을 적용하였다[16].

류미영 외(2021)는 초·중등 대상의 인공지능 소양 및 윤리 교육을 위한 인공지능 학습모델 개발을 위하여 선행 연구와 FGI(Focus Group Interview)를 통해 인공지능 교육내용과 인공지능 윤리요소를 추출하였다[17].

공은주(2022)는 학습자들에게 인공지능 개발자로서의 사회적 책임을 교육하는 프로그램 개발을 목적으로, 설계 준비 단계에서 고등학교 인공지능 기초 과목 중 인공지능의 사회적 영향 단원 내용을 분석하고 인공지능 윤리 관련 내용을 분석 및 적용하였다[18].

국내외 인공지능 윤리 교육 현황과 연구 내용을 분석해 보면 공통적 키워드로 '사회적 영향력', '인공지능 윤리', '인공지능 가치'를 뽑고 있으며, 인공지능 기술의 경험, 데이터, 사회적 이해관계자에 대한 교육을 통해 인공지능으로 인한 윤리적 이슈와 문제에 대해 참여하도록 제시하고 있다. 또한, 인공지능 교육과 함께 인공지능 윤리 교육을 강화하고 이를 적용하기 위한 연구가 활발히 이루어지고 있음을 알 수 있다. 하지만 인공지능에 대한 교육과 인공지능 윤리에 대한 교육이 통합되지 않고 분리되어 이루어지고 있다. 인공지능 윤리는 인공지능 개발부터 활용 전 범위에 걸쳐 고려되고 실천 될 때 의미가 있다. 따라서 본 논문에서는 인공지능 윤

리의 실천력 강화를 위해서 인공지능 교육과 인공지능 윤리 교육을 분리하고 하지 않고 문제해결의 각 단계마다 인공지능 윤리를 적용하는 인공지능 윤리 교육 모델을 제안하고자 한다.

### 3. 연구방법

본 연구는 인공지능 윤리 교육의 모델 개발을 위해 Table 5과 같이 4단계로 진행하였다.

Table 5. Research Method of AI Ethics Education

Step	Procedure	
	Research Progress and Methods	
	Derivation of Educational Elements	
1	<b>Analysis of Prior Literature</b>	
	·Ethics Guide for AI Ethics Education, analysis of domestic and foreign AI Ethics Education Status ·Project-Based AI Ethics Education Program ·Advancement of Educational Goals ·Education Content	
2	<b>Model Analysis and Advancement</b>	
	·Derivation of complementary points through existing model analysis (content aspect analysis) ·Derivation a problem-solving process using AI ·Teaching and Learning	
3	<b>Model Development</b>	
	·Design the content and procedure of the model for each educational element and draw connectivity  ·Educational Goals ·Education Content ·Teaching and Learning ·Assessment Methods	
4	<b>Model Comparison Analysis</b>	
	Comparative Analysis with Existing AI Education	

1단계는 배진아(2021)의 프로젝트 기반 인공지능 윤리 교육 프로그램[19]과 선행 연구 분석을 통해 윤리 교육의 교육목표 및 내용 기준을 심화 및 고도화하였다. 2단계는 학습자의 인공지능 윤리 실천력 강화를 위한 선행 연구를 분석하여 교수학습방법, 평가방법 등에 내용이 어떻게 제안되었는지 도출하였다. 또한, 프로젝트 기반 인공지능 윤리 교육 프로그램[19] 연구를 고도화하여 컴퓨팅 사고력 함양을 위한 인공지능 문제해결 프로세스를 6개로 정의하였다. 3단계는 인공지능 문제해결 프로세스에서 고려해야 할 윤리적 요소를 추출하고 실천력 강화를 위한 교수학습 방법(교육목표, 내용

기준, 교수학습, 평가방법)을 적용한 인공지능 윤리 교육 모델을 개발하였다. 4단계는 일반적인 인공지능 교육과 본 연구에서 개발된 인공지능 윤리 교육 모델을 비교, 분석하였다.

### 4. 연구결과

#### 4.1 인공지능 윤리 교육의 교육목표

인공지능 윤리 교육의 첫 번째 방향은 컴퓨팅 사고력 함양이다. 2020년 교육과정 부분 개정으로 추가된 ‘인공지능 기초’ 교과목의 성격을 보면 인공지능은 컴퓨터과학의 한 영역으로 다양한 분야와 실생활 문제해결에 활용되고 있으며, 인공지능 교육을 통해 컴퓨팅 사고력을 바탕으로 인공지능의 원리와 기술을 탐구하고 인공지능의 관점에서 문제를 바라보고 해결하는 역량을 함양해야 한다고 하였다[20].

신승기(2020)는 인공지능 교육이 기본적으로 컴퓨터를 활용하는 소양을 토대로 기존 정보교육에서 강조하고 있는 컴퓨팅 사고력을 통한 교육과정 편성이 이루어져야 한다고 주장하며 컴퓨팅 사고력 기반의 인공지능 교육 프레임워크를 제시하였다[21].

Tabesh(2017)은 21세기 핵심 역량을 컴퓨팅 사고력으로 제시하며 인공지능의 개념과 알고리즘을 바탕으로 문제를 해결하는 과정에서 컴퓨팅 사고력을 향상할 수 있다고 제시하였다[22].

이처럼 인공지능 교육은 컴퓨팅 사고력을 목적으로 이루어져야 하며, 인공지능 윤리 교육을 적용하여 윤리적인 요소를 충분히 고려한 문제 해결 방안이 결국에는 보다 효과적인 컴퓨팅 사고력으로 접근하게 될 것이다.

인공지능 윤리 교육의 두 번째 목표는 실천력 강화를 위한 인공지능 윤리 교육이다. 인공지능 기술로 인한 윤리적 문제는 현실적이고 구체적인 해결책을 요구하여 기존의 지식 전달 중심의 인공지능 윤리로는 구체적인 실천 행위에 대한 지침을 제공하지 못한다. 따라서 실천 중심의 인공지능 윤리 교육이 필요하며 지식 전달에 머무르지 않고 인공지능에 윤리를 직접 적용함으로써 자연스럽게 윤리 의식을 함양할 수 있는 교육이 이루어져야 한다[3]. 또한, 실천력 강화를 위한 인공지능 윤리 교육이 이루어지기 위해서는 인공지능 교육과 인공지능 윤리 교육이 분리되지 않고 함께 이루어지도록 해야 한다.

## 4.2 인공지능 윤리 요소와 문제해결 프로세스

### 4.2.1 인공지능 윤리 핵심 가치와 윤리 요소

인공지능을 이용한 기술이나 서비스의 핵심 요소는 알고리즘과 데이터이다. 일반적으로 문제해결을 위한 절차나 방법이 알고리즘이지만 인공지능 알고리즘은 규칙이 생성되는 과정에서 기존 알고리즘과는 차이가 있다. 기존 알고리즘은 예측 가능하고 통제 가능한 결과를 가져오도록 인간의 의도대로 설계하지만, 인공지능 알고리즘은 데이터를 기반으로 기계 학습을 진행하는 과정에서 블랙박스과 같은 모델이 구축되어 결과를 도출하기 때문에 예측과 통제가 상대적으로 어렵다.

즉, 인공지능 모델은 알고리즘과 기계 학습에 사용되는 데이터에 따라 결과를 도출하므로 이 과정에서 발생할 수 있는 예측 불가능한 문제와 윤리적 요소를 미리 고려해야 한다.

본 논문에서 제시하는 인공지능 윤리 요소는 국내외 인공지능 윤리 가이드라인에서 공통 요소를 추출하고 인공지능을 이용한 문제해결 과정에서 반드시 다루어야 하는 윤리 요소를 추출하여 Table 6과 같이 '다양성', '투명성', '책임·공공성', '개인정보보호' 4가지로 정의하였다.

Table 6. AI Ethical Elements

Ethical Standards & Definition
<b>Diversity(bias exclusion)</b> Artificial intelligence systems should be usable by everyone regardless of age, gender, race, etc., and should be treated fairly and should not have any unfair biases.
<b>Transparency</b> Decisions made by artificial intelligence systems must be understandable and traceable by humans, risk-related information can be disclosed and shared, and the entire process of personal information processing must be made transparently.
<b>Accountability &amp; Publicity</b> The AI system must be evaluable, and the responsibilities of all interested parties and influences, including users, must be clear, and stakeholders include users, developers and researchers, and suppliers.
<b>Privacy</b> Data collected and utilized by artificial intelligence systems. All personal information must be secure and private and must not be infringed or opened.

### 4.2.2 인공지능을 이용한 문제해결 프로세스

본 연구는 인공지능 윤리의 실천력 강화와 컴퓨팅

사고력 함양을 단계적으로 적용하기 위하여 인공지능을 이용한 문제해결 프로세스를 제시하였다. 기존의 프로젝트 기반 인공지능 윤리 교육 프로그램의 프로세스와 달리 응용프로그램 개발 및 적용 단계를 추가하여, 훈련된 모델 적용을 통해 인공지능 기술과 관련한 이해관계자와 사회적 영향력에 대한 이해를 도울 수 있도록 설계하였다. 인공지능을 이용한 문제해결 프로세스는 Fig. 1과 같이 문제정의 및 주제선정, 데이터 수집, 데이터 전처리, 모델 훈련 및 알고리즘, 응용프로그램 개발 및 적용, 성능 평가 및 분석 단계로 선정하고 단계마다 핵심 윤리 요소를 고려하고 적용할 수 있도록 구성하였다.

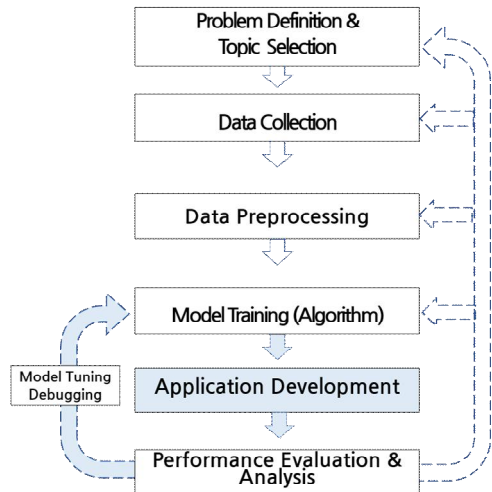


Fig. 1. Problem-solving Process for AI

### 4.3 실천력 향상을 위한 교수학습방법

인공지능 윤리는 알고 느끼는 것에 그치는 것이 아니라 행동과 실천을 통해서 그 역할을 다 할 수 있다 [23]. 따라서 인공지능 윤리 교육은 인지 교육만으로 구성하는 것이 아니라, 생활 속에서 자연스럽게 실천할 수 있는 교수 방법이 적용되어야 한다[24].

이재운(2008)은 실천력 강화를 위한 학습자 중심 평가 루브릭 모델을 개발하여 자기 평가와 동료 평가를 통한 성찰을 통해 학습자의 실천력이 강화될 수 있다는 것을 확인하였다[25].

이상우(2009)는 개인정보보호 실천력 향상을 위해 역할 놀이 교수학습 방법을 채택하여 적용하였다. 그는 역할 놀이 수업을 통해 문제 상황에 대한 이해를 돕고,

학습자의 역할 행동이 실천으로 이어질 수 있다고 하였다. 또한, 토론 활동의 과정을 거치면서 더욱 내면화할 수 있다고 하였다[26].

김완수(2012)는 학습자의 도덕성을 인지적 영역에서 탈피하기 위해 협동 학습, 역할 놀이, 체험학습 등을 제시하며 학습자의 실천력을 높이는 방안을 제시하였다[27].

이처럼 실천력 향상을 위한 교수학습 방법의 선행 연구를 살펴보면 자기 평가와 성찰 일지를 통한 자기성찰, 역할 놀이와 협동 학습 및 토론을 통해 학습자의 자율성과 행동을 습관화하는 것을 알 수 있다. 따라서 본 연구는 인공지능 윤리 실천력 강화를 위한 교수학습 방법을 Fig. 2와 같이 적용하였다.

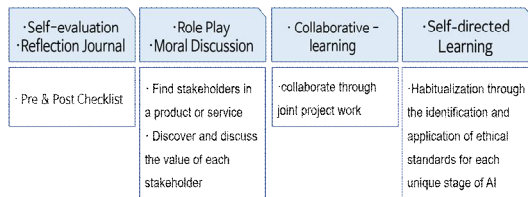


Fig. 2. Learner-Centered Teaching and Learning Method for Practical Power Strength

자기 평가와 성찰일지는 수업 전·후에 교육을 통하여 변화된 자신을 스스로 점검하는 것으로 자신의 행동을 관찰하고 판단하며 자기 반응을 통하여 스스로 조절하는 능력을 키우는 것이다[28]. 또한, 인공지능을 이용한 공동 프로젝트 작업은 협동 학습, 자기주도학습을 통해 이루어지고 프로세스마다 윤리 기준을 확인하고 적용하는 활동을 통해 윤리문제에 대한 인식을 습관화할 수 있다. 마지막으로 도덕적 토론과 역할 놀이를 통해 사회적 영향력과 이해관계자의 의미와 관계를 올바르게 이해할 수 있다.

4.4 인공지능 윤리 평가 방법

인공지능 윤리 의식 평가는 사전·사후 진단 평가와 문제해결 과정 프로세스마다 윤리 항목의 적용을 확인하는 과정평가로 이루어진다. 사전·사후 진단 평가의 목표는 스스로 현재 상태를 파악하고, 자가 진단을 통해 스스로 인공지능 윤리 기준을 무의식적으로 확인하게 하려는 목적으로 적용하였다. 진단을 위한 설문 예시는 Table 7과 같이 인공지능 핵심가치와 윤리 기준

을 반영하였다. 이러한 자가 진단 평가는 인공지능을 이용한 문제해결 프로세스마다 고려해야 할 핵심 윤리적 요소 생각하고 고려하도록 설계하였다.

Table 7. Example of a Questionnaire for Diagnosis

No	Question	Reply
1	Technology using AI always works beneficial to our humans.	12345 □□□□□
2	The results guessed by AI are important, and the process of why they guessed is not important.	12345 □□□□□
3	Personal Information can be provided at any time to make good products or services.	12345 □□□□□

인공지능 윤리에 대한 과정 평가는 Table 8과 같이 사전·사후 평가의 한계를 보완하고 인공지능 윤리를 습관화하기 위한 방법으로 적용한다. 인공지능을 이용한 문제해결 프로세스마다 관련된 인공지능 윤리 항목의 중요성에 대한 본인의 생각을 왼쪽에 체크 하도록 하고, 실제 문제해결 과정에서 어느 정도 고려하였는지 확인하도록 오른쪽에 체크 하도록 설계하였다. 이를 통해 인공지능 윤리의 중요도를 인식하고 실천 정도를 확인하여 인공지능 윤리 의식과 실천력 향상을 유도하였다.

Table 8. Example of AI Ethics Evaluation

Import-ance	Data Collection	Execut-ion
12345 □□□□□	Did you collect data that did not include other people's personal information and did not infringe on portrait rights?	12345 □□□□□
12345 □□□□□	Is the use marks and sources of copyright accurate?	12345 □□□□□
12345 □□□□□	Was copyrighted data used without copyright permission to create the desired model?	12345 □□□□□

4.5 인공지능 윤리 교육 모델

본 논문에서 제안하는 인공지능 교육 모델은 Fig. 3과 같이 인공지능 교육과 인공지능 윤리 교육이 함께 이루어지는 교육 모델이다. 인공지능 윤리 교육 모델은 컴퓨팅 사고력 함양과 인공지능 윤리의 실천력 강화를 목표로 설정한다. 인공지능을 이용한 문제해결 프로세스는 AP1~AP6 프로세스로 설정하였고, 각 프로세스마다 고려해야 할 윤리 요소를 포함한다. 윤리 요소는 다양성, 투명성, 책임·공공성, 개인정보보호의 항목에 대해 중요도에 따라 ●, ○, ◎ 순으로 구분할 수 있다. 인공지능 윤리 의식 평가는 자가 진단과 체크리스트를 통한 사전·사후 평가와 윤리 요소와 실천 윤리를 평가



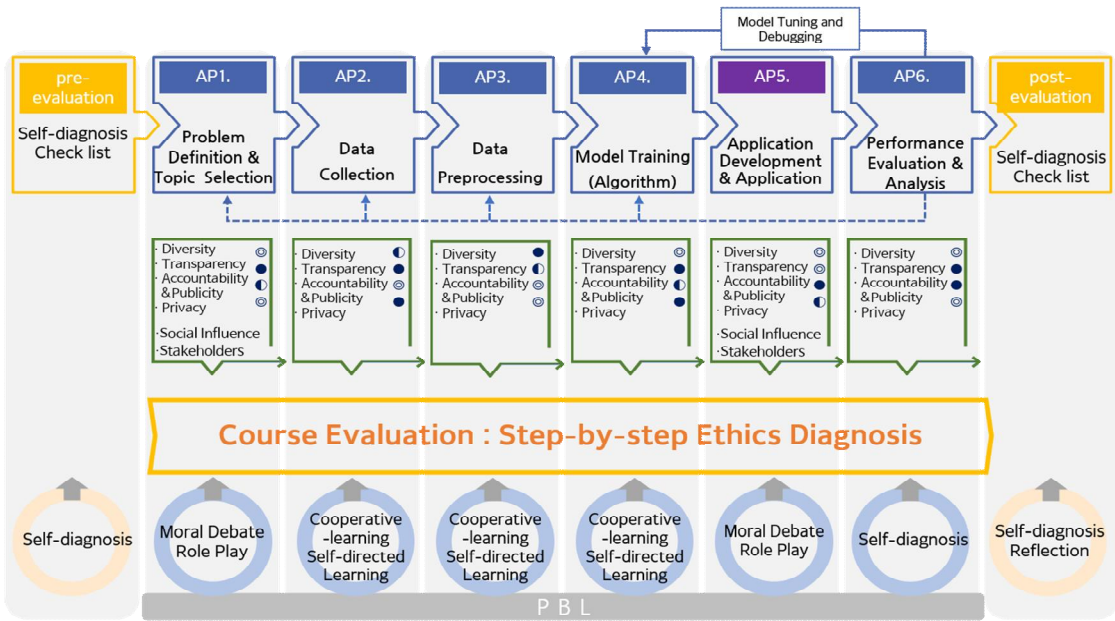


Fig. 3. AI Ethics Education Model

하는 과정평가로 진행된다. 사전 평가는 프로세스를 진행하기 전에 인공지능 윤리에 대한 현재 상태를 점검함으로써 윤리에 대한 의식을 깨우는 단계이다. 사후 평가는 프로세스 완료 후에 인공지능 윤리에 대한 인식 변화를 스스로 경험하고 피드백을 통해 윤리 의식을 강화하는 평가이다. 또한, 학습자 중심의 교수학습방법인 자기평가, 성찰, 토론, 협동 학습 등을 통해 컴퓨팅 사고력과 인공지능 윤리 실천력을 강화하였다.

#### 4.6 인공지능 교육 모델 비교 분석

류혜인(2021)은 4P(Play, Problem Solving, Product Making, Project) 기반의 K-12 대상 수준별 인공지능 교육체계 개발을 통해 AI 윤리를 공통적인 핵심으로 선정하고 인공지능 교육의 핵심개념과 연관된 윤리적 이슈와 사회적 영향력을 다루도록 Fig. 4와 같이 설계하였다[29].

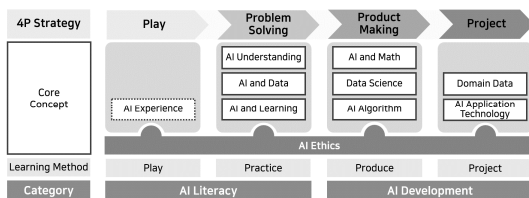


Fig. 4. AI Education System[29]

4P 기반 인공지능 교육은 인공지능 윤리 교육을 별도로 구성함으로써 본 연구에서 주장하는 인공지능 윤리가 포함된 인공지능 교육에는 한계가 있다. 또한, 인공지능의 특성을 반영한 윤리요소의 적용 및 반영이 이루어지지 않고, 지식으로 머물지 않고 실천력으로 이어질 수 있는 인공지능 윤리 교육에 적합한 교수학습방법이 제시되지 않았다.

따라서, 본 논문은 기존 인공지능 교육의 부족한 부분인 인공지능의 특성을 반영한 인공지능 윤리와 인공지능 윤리 실천력 강화를 위한 교육목표, 교육내용, 교수학습방법, 교육평가를 제안하고 인공지능 윤리 교육과 인공지능 교육을 통합해서 진행할 수 있는 인공지능 윤리 교육 모델을 제안하였다.

### 5. 결론

인공지능은 인간에게 이로운을 주는 긍정적인 목적의 기술이지만 잘못된 의도로 설계되거나 인간의 통제를 벗어난 인공지능은 인간에게 심각한 해를 끼칠 수 있다. 이러한 양면성 때문에 인공지능 기술은 인공지능 윤리와 함께 발전하고, 윤리적인 고려가 충분히 되었을 때 안전한 적용이 가능할 것이다[30].



본 논문에서는 컴퓨팅 사고력 함양과 인공지능 윤리의 실천력 강화를 목표로 인공지능 윤리 교육 모델을 개발하였다. 구체적 개발 내용은 첫 번째, 인공지능 교육을 통해 인공지능의 특성을 반영한 인공지능 윤리 요소를 도출하고, 인공지능 윤리의 실천력 강화와 컴퓨팅 사고력 함양을 목표로 인공지능 윤리 교육 모델을 설계하였다. 두 번째, 인공지능 윤리 교육 모델을 위한 문제해결 프로세스를 '문제 정의 및 주제 선정', '데이터 수집', '데이터 전처리', '모델 훈련(알고리즘)', '응용프로그램 개발 및 적용', '성능 평가 및 분석'의 6단계로 정의하였다. 세 번째, 인공지능 윤리의 실천력 강화를 위해 문제해결 프로세스에 교육목표, 교육내용, 교수학습방법, 교육평가를 위한 윤리적 요소를 모두 고려하여 적용하였다. 마지막으로 비교 연구를 통해 인공지능 윤리 교육 프로그램에서 교육적 요소와 실천력 강화 요소를 비교 분석하였다.

본 연구를 통해 인공지능 윤리의 중요성을 인식하고 인공지능 교육과 인공지능 윤리 교육이 분리되지 않고 함께 이루어지는 교육이 활발히 이루어져 기존의 지식 전달 위주의 인공지능 윤리 교육에서 벗어나 인공지능 윤리의 실천력을 강화하기 위한 연구가 활발히 이루어지기를 기대한다.

## REFERENCES

- [1] KIST. (2019. Dec). He ethics/policy/social issues of artificial intelligence. *ISSUE PAPER No. 08*.
- [2] M. J. Kim. (2017). The Necessity of Artificial Intelligence Ethics and Trends in Korea and abroad. *KICS(Information & Communications Magazine)*, 34(10), 45-54.  
DOI : 10.20972/kjee.31.3.202009.153
- [3] K. S. Hwang & S. D. Kim. (2013). *Practical Ethics : How to Live Ethically in this Era*. Gyeonggi : Yeonam Bookstore.
- [4] Jointly related Ministries and related Ministries. (2021). *Reliable Artificial Intelligence Practice Strategies for Human -centered Artificial Intelligence*. Seoul.
- [5] S. G. Han. (2019). Social Issues and Ethical Issues in Artificial Intelligence Technology. *Korea Multimedia Society*, 20(3), 41-52.
- [6] Jointly related Ministries and related Ministries. (2020). *Human-centered 'Artificial Intelligence [AI] Ethics Standards'*. Seoul.
- [7] S. G. Han. (2016). Social Issues and Ethical Issues in Artificial Intelligence Technology. *Korea Multimedia Society*, 20(3), 41-52.
- [8] E. K. Kim. (2022). *The Effect of Artificial Intelligence Ethics Education Using Moral Machine on Elementary School Students' Attitudes and Images toward Artificial Intelligence*. Master dissertation. KNUE University, Cheongju.
- [9] NIA. (2019). *Artificial Intelligence Ethics Guidelines-Focused on Japan and EU Cases*. Intelligent Information Society Legal System *ISSUE REPORT*.
- [10] Asiloma Conference. (2017). Asiloma AI Principle 23. THE FUTURE OF LIFE INSTITUTE. <https://futureoflife.org/ai-principles-korean>.
- [11] EU. (2019). *AI Ethics Guidelines*, Requirements for reliable AI realization.
- [12] IEEE. (2019). *EAD1e (ETHICALLY ALI- GNED DESIGN, First Edition) REPORT*.
- [13] NIA. (2021). *UNESCO AI Ethics Recommendation Key Contents and Implications*. IT\_& Future\_Strategy\_ 2021-10. Dague : NIA.
- [14] NIA. (2018). *Ethics Guidelines for Intelligent Information Society* . Dague : NIA.
- [15] Ministry of Education. (2021). *The Age of Artificial Intelligence, Education Policy Direction and Key Challenges*. Seoul.
- [16] J. E. Kim. & C. H. Lee. (2021). Development of STEAM Program for Artificial Intelligence Ethic Education for Elementary School Student. *Korean Association of Artificial Intelligence Education Transaction*, 2(1), 21-28.
- [17] M. Y. Ryu & J. J. Jung. (2021). Extraction of Learning Elements and Development of Teaching and Learning Models for AI Ethics Education. *Korean Association of Artificial Intelligence Education Transaction*, 2(1), 45-52.
- [18] E. J. Gong. (2022). *Development of Artificial Intelligence Ethics Education Program Reflecting Social Responsibility*. Master dissertation. KNUE University, Cheongju.
- [19] J. Bae. (2020). *Development of Project-Based Artificial Intelligence Ethics Education Program*. Master Dissertation. JNU University, Jeju.
- [20] Ministry of Education. (2020). Ministry of Education Notice No. 2020-236. Revised elementary and secondary curriculum.
- [21] S. G. Shin. (2020). *Designing the Framework of Evaluation on Learner's Cognitive Skill for*

Artificial Intelligence Education through Computational Thinking. *Journal of The Korean Association of Information Education*, 24(1), 59-69. DOI : 10.14352/jkaie.2020.24.1.59

- [22] Yahya TABESH. (2017). *Computational Thinking : A 21st Century Skill*. Olympiads in Informatics, 11, 65-70.
- [23] M. H. Lee. (2008). *Development and Application of a Teaching and Learning Model Integrating Cognitive, Affective and Conative Dimensions for the Reinforcement of Junior Elementary School Students' Practice of Information Technology Ethics*. Master dissertation. KNUE University, Cheongju.
- [24] C. S. Jang. (2002). *A Study on the effects of activity and experience-centered environmental education programme for fostering the learners' practical ability*. Master Dissertation. CJE University, Cheongju.
- [25] J. W. Lee. (2008). *The influence of student-generated rubrics on achievement and ability to put information literacy into practice in informatics subject performance assessment*. Doctoral Dissertation. KNUE University, Cheongju.
- [26] S. W. Lee. (2009). *The Effect of action to private information protection by using a case-focused and roleplaying model*. Master Dissertation. KNUE University, Cheongju.
- [27] W. S. Kim. (2013). *A study on the empathic education method for the advancement of moral behavior in elementary school students*. Master Dissertation. JNUE University, Jeonju.
- [28] S. H. Kang. (2018). *Information Ethics Education Programbased on Scratchfor Execution Power Strength*. Master dissertation. JNU University, Jeju.
- [29] H. Y. Ryu & J. Cho. (2021). Development of Artificial Intelligence Education System for K-12 Based on 4P. *Journal of Digital Convergence*, 19(1), 141-149. DOI : 10.14400/JDC.2021.19.1.141
- [30] IAAE. (2019). *The AI Ethics Charter, introduction*. Seoul : IAAE.

**배진아(Jinah Bae)**

[정회원]



- 2003년 8월 : 한국방송통신대학교 컴퓨터과학과(이학사)
- 2020년 8월 : 제주대학교 컴퓨터교육전공(교육학석사)
- 2020년 9월~현재 : 제주대학교 컴퓨터교육전공 박사과정
- 2018년 7월~현재 : 제주대학교 지능소프트웨어교육연구소 연구원
- 관심분야 : 정보·컴퓨터(SW, AI), 인공지능윤리, 디지털미디어, 디지털격차해소
- E-Mail : banggle726@jejunu.ac.kr

**이정훈(Jeonghun Lee)**

[정회원]



- 2018년 2월 : 제주대학교 전자공학학과(공학사)
- 2020년 8월 : 제주대학교 컴퓨터교육전공(교육학석사)
- 2020년 9월~현재 : 제주대학교 컴퓨터교육전공 박사과정
- 2018년 7월~현재 : 제주대학교 지능소프트웨어교육연구소 연구원
- 관심분야 : SW교육, AI교육, 인공지능, 지능형시스템
- E-Mail : 2ehdrks@jejunu.ac.kr

**조정원(Jungwon Cho)**

[정회원]



- 2004년 2월 : 한양대학교 전자통신전파공학과(공학박사)
- 2004년 9월~현재 : 제주대학교 컴퓨터교육과 교수
- 2020년 3월~현재 : 한국컴퓨터교육학회 부회장, 논문지편집위원장
- 2012년 12월~현재 : 한국정보과학회 전산교육시스템연구회 위원장
- 2018년 7월~현재 : 제주대학교 지능소프트웨어교육연구소 소장
- 관심분야 : 정보·컴퓨터(SW, AI)교육, 지능정보윤리, 지능형시스템, 멀티미디어
- E-Mail : jwcho@jejunu.ac.kr