



국내·외 인공지능 윤리교육 사례 분석과 시사점



KERIS

한국교육학술정보원
KOREA EDUCATION AND RESEARCH INFORMATION SERVICE

국내·외 인공지능 윤리교육 사례 분석과 시사점

I 서론

II 해외 인공지능 윤리교육 사례

III 국내 인공지능 윤리교육 사례

IV 결론 및 시사점

연구·집필진

김법연 연구교수 (고려대학교 BK21 정보보호학교교육연구단)

기획·검토

최미애 부장 (한국교육학술정보원)

최효재 선임연구원 (한국교육학술정보원)

본 리포트는 한국교육학술정보원의 공식의견이 아니라
연구자의 개인 견해를 밝힙니다.

차 례

I. 서 론	1
II. 해외 인공지능 윤리교육 사례	1
1. 미국	1
2. 유럽	10
3. 기타 국가	26
III. 국내 인공지능 윤리교육 사례	36
1. 공공분야 개발 사례	36
2. 민간분야 개발 사례	51
IV. 결론 및 시사점	57
참고 문헌	61

표 목 차

<표 II-1> How to Train Your Robot 세부 프로그램	7
<표 II-2> 공인 AI 윤리 및 거버넌스 교육과정 주요 내용	9
<표 II-3> Artificial Intelligence: Ethical & Societal Challenges 프로그램	12
<표 II-4> 독일 뮌헨 공과대 AI 강좌 주요 내용	14
<표 II-5> Schule macht KI 프로그램	16
<표 II-6> 핀란드 Elements of AI 프로그램 주제	18
<표 II-7> 핀란드 Elements of AI 프로그램 주제별 주요 내용	19
<표 II-8> Ethics of AI 프로그램	21
<표 II-9> 영국 Ethics of AI 세부 프로그램	23
<표 II-10> AI4T 세부 프로그램	25
<표 II-11> 태국 AiX 세부 프로그램	27
<표 II-12> 싱가포르 AI4K 교육 과정	30
<표 II-13> 싱가포르 AI4E 세부 내용	31
<표 II-14> 싱가포르 AI4I Literacy in AI 세부 내용	32
<표 II-15> 캐나다 인공지능 윤리 마이크로 자격 증명 AI 세부 프로그램	34
<표 III-1> KISDI 초등용 인공지능 윤리 교재 주요 골자	38
<표 III-2> KISDI 초등용 인공지능 윤리 교재 주제별 구성 의도	38
<표 III-3> KISDI 중등용 인공지능 윤리 교재 주요 골자	40
<표 III-4> KISDI 중등용 인공지능 윤리 교재 주요 내용	41
<표 III-5> KISDI 중등용 인공지능 윤리 교재 세부주제별 학습내용	43
<표 III-6> KISDI 일반 성인 대상 인공지능 윤리교육 기본 프레임	45
<표 III-7> KISDI 일반 성인 대상 인공지능 윤리교육 단위별 교재 타이틀	45
<표 III-8> KISDI 일반 성인 대상 인공지능 윤리교육 단계별 주요 내용(1단계)	46
<표 III-9> KISDI 일반 성인 대상 인공지능 윤리교육 단계별 주요내용(2~4단계)	46
<표 III-10> NIA 인공지능 윤리교육 프로그램의 주요 내용	49
<표 III-11> NIA 인공지능 윤리교육 프로그램의 주요 내용	49
<표 III-12> IAAE 인공지능 윤리교육의 주요 커리큘럼(일반인용)	52

<표 III-13> 네이버 인공지능 윤리교육의 주요 내용	54
<표 III-14> 네이버 생성형 인공지능 윤리교육의 주요 내용	56
<표 IV-1> 국가별 세부 인공지능 윤리교육 프로그램 및 사례 비교	57

그 립 목 차

[그림 II-1] MIT Media Lab의 AI+Ethics 홈페이지	1
[그림 II-2] MIT Media Lab의 How to Train Your Robot 홈페이지	5
[그림 II-3] 공인 AI 윤리 및 거버넌스 전문가 과정	8
[그림 II-4] AI Lund의 AI 윤리 코스 홈페이지	11
[그림 II-5] 뮌헨 공과대학교(TUM)의 무료 온라인 강좌 홈페이지	13
[그림 II-6] KI Campus의 AI 관련 강의 홈페이지	15
[그림 II-7] 핀란드의 Elements of AI 홈페이지	17
[그림 II-8] 핀란드의 Ethics of AI 온라인 강의	20
[그림 II-9] 런던 정치경제대학교(LSE)의 Ethics of AI 홈페이지	22
[그림 II-10] 유럽연합의 AI4T 홈페이지	24
[그림 II-11] 태국의 AiX 프로그램	26
[그림 II-12] LearnAI 세대별 AI 역량 개발	28
[그림 II-13] 캐나다 아타바스카 대학의 AI 윤리 마이크로 자격 증명	33
[그림 II-14] 인도네시아에서 진행된 AI 윤리 교사 연수 과정	35
[그림 III-1] 과학기술정보통신부, KISDI의 인공지능 윤리교육 교재(초·중·고등학교)	37
[그림 III-2] 방송통신위원회, NIA 디지털윤리 홈페이지	48
[그림 III-3] IAAE 인공지능 윤리교육 홈페이지	51
[그림 III-4] 소프트웨어야 놀자 홈페이지	53

I. 서 론

인공지능 기술은 효율성이나 생산성, 편의성 등을 확대해주는 이점으로 인해 활용이 점차 증가하고 있다. 인공지능 기술의 정의나 범주가 어디까지인가에 대해 명확하게 규정하기는 어렵지만 인공지능 기술을 여러 분야에서 다양한 방식과 내용으로 활용하고 있는 것은 거부할 수 없는 사실이다. 학교 현장도 예외는 아니다. 인공지능 기술의 학교 현장 적용이 확대됨에 따라 위험성에 대한 우려도 커지고 있다. 프라이버시나 저작권 침해, 개인이나 집단의 특징에 대한 사회적 편견이나 차별 이슈와 같은 권리 침해 문제에서부터, 노동력 대체, 인간의 능력에 대한 도전이나 자율성 훼손 문제, 기술사용 역량에 따른 격차 문제 등 다양한 이슈와 위험들이 나타나고 있다.

이에 인공지능 기술에 대한 두려움으로 위험성에 대한 통제가 필요하다는 논의와 함께 이를 사용하는 사람들에 대한 교육의 필요성도 제기되고 있다. 모든 기술이 그러하듯 인공지능 기술 또한 기술 자체가 위험을 유발하는 것이라기보다는 기술을 어떻게 사용하느냐가 문제이기 때문이다. 앞서 언급한 인공지능 기술의 문제와 위험들은 기술을 만들고 배포하는 사람들에 의해서도 야기될 수도 있고, 사용자에 의해서도 발생할 수 있다. 개발자가 고의적으로 위험을 야기하도록 만들 수도 있고, 사용자가 인공지능 시스템 또는 서비스의 제공자가 의도한 목적대로 사용하지 않는 경우도 발생할 수 있다. 최근 등장한 언어모델을 기반으로 하는 생성형 AI의 경우 할루시네이션을 유발하는 특성을 지니고 있는데, 이러한 특성은 사용자에게 생성형 AI의 산출물에 대한 정확성이나 허위성에 대한 판단을 요구하기도 한다. 원하는 결과를 얻기 위하여 필요한 프롬프트의 내용을 선택하는 것도 필요한데, 이는 결국 사용자의 역량을 강화시켜야 한다는 결론에 이르게 된다.

인공지능에 대한 정확한 이해를 바탕으로 기술의 편익을 누리면서도 인공지능을 우리 사회가 합의한 가치에 반하지 않게 의도된 목적으로 사용하는 것은 중요하다. 또한 타인의 권리를 침해하지 않고 사용할 수 있도록 하는 것이 중요한 국가적 목표가 되고 있다. 이러한 상황에서 인공지능 윤리교육은 목표 달성을 위한 필수 수단이라고 여겨진다. 즉, 인공지능 윤리교육은 우리 사회가 인공지능 기술이 주는 유익함과 효용을 극대화하면서 기술의 위험성을 최소화 하기

위한 필수 전제라고 할 수 있다.

인공지능 윤리교육의 필요성에 대한 공감대가 형성되면서, 구체적으로 어떻게 교육의 내용과 방법을 구성하는 것이 필요한지에 대한 논의가 활발하게 이루어지고 있다. 해외 주요국들은 국가 수준 또는 학교 및 교육 관련 단체, 민간 기업 등 개별 단위의 조직에서 인공지능 윤리교육 프로그램을 개발하고 있다. 국내에서도 이에 대한 관심과 필요성이 지속적으로 제기되면서 교육내용과 방법에 대한 논의와 연구를 실시 중에 있다.

인공지능 기술이 여러 분야에서 다양한 주체에 의해 광범위하게 활용되고 있는 상황인 만큼 인공지능 윤리교육의 목표를 실효적으로 달성하기 위해서는 다양한 교육내용과 교육방법들에 대한 이해가 필요하다. 이에 본고에서는 가능한 많은 프로그램의 내용을 비교 분석하여 적정하고 구체적인 교육 프로그램을 개발하는데 기여하고자 해외 주요 국가에서 개발·제안되고 있는 인공지능 윤리교육의 프로그램 등을 조사·분석하였다. 또한 다양한 기관과 주체들이 제시하고 있는 인공지능 윤리교육 프로그램의 비교·분석을 통해 적정하고 합리적인 인공지능 윤리교육이 실시될 수 있도록 시사점을 제공하고자 한다.

II. 해외 인공지능 윤리교육 사례

1. 미국

미국의 인공지능 윤리교육은 AI4K12¹⁾에서 제시한 국가 지침에 기반하여 MIT 소재 미디어랩(Media Lab)이 운영하고 있는 교육 프로그램이나 (주)Tonex가 운영하는 전문가 과정처럼 주로 학계나 기업 등 민간과의 협력을 통해 이루어진다. 인공지능 기술이 발달하면서 발생한 위험성을 윤리적인 측면에서 살펴 보며, 기술을 사용하는 사람으로서 지녀야 할 책임 등에 대해 교육한다.

가. MIT Media Lab의 AI+Ethics

MIT 소재 미디어랩(Media Lab)은 AI 기술의 윤리적 활용을 강조하는 중학생 대상 교육 프로그램인 AI+Ethics를 개발하여 제공하고 있다. 학생들이 인공지능과 인공지능 기술이 사회에 미치는 영향력을 이해할 수 있도록 구상되었으며 몇몇 학교에서 시범적으로 사용하고 있다.



[그림 II-1] MIT Media Lab의 AI+Ethics 홈페이지

1) 교육대상

해당 교육은 5~8학년 정도의 초·중등 학생을 대상으로 설계되고 시범 적용되었다.²⁾³⁾

1) AI4K12는 미국의 K-12용 AI교육에 대한 국가지침으로 인공지능 교육 표준모델이다. 자세한 사항은 AI4K12 웹사이트 참조.(Available at: <https://ai4k12.org/>)

2) MIT Media Lab, “An Ethics of Artificial Intelligence Curriculum for Middle School Students”, 3면.

2) 운영방안

MIT 미디어랩에서 개발한 교재에 안내된 자료를 바탕으로 오프라인 방식의 수업을 시범적으로 운영하였으나, 일부 활동은 컴퓨터를 활용하도록 하고 있다.

3) 세부 프로그램 주요 내용

활동은 다음과 같은 학습 목표 달성을 위해 마련되었으며 각 활동별로 학습 목표는 다르게 구성되어 있다.⁴⁾

학습 목표
• 인공지능 시스템의 기본적인 작동 원리(메커니즘)를 이해한다. • 모든 기술 시스템은 사회-기술 시스템(socio-technical systems)이라는 것을 이해할 수 있다. 사회-기술 시스템은 가치중립적이지 않으며, 정치적 쟁점을 가진다는 것을 이해할 수 있다. • 사회-기술 시스템에는 많은 이해관계자들이 있으며 시스템은 이러한 이해관계자들에게 각각 다른 방식으로 영향을 미칠 수 있다는 것을 인식할 수 있다. • 사회-기술 시스템의 올바른 목표를 결정하기 위해 인공지능에 대한 기술적 이해와 이해관계자에 대한 지식을 적용할 수 있다. • 기술이 세계에 미치는 영향력을 고려할 수 있다.

5가지로 분류된 학습 목표 및 목표의 하위 항목은 다음과 같다.

1. 인공지능 시스템의 기본적인 작동 원리(메커니즘)를 이해한다.
a. 전 세계의 다양한 알고리즘을 인식하고, 일상생활 속에 있는 알고리즘과 컴퓨터 알고리즘의 예를 제시할 수 있다. b. 알고리즘의 세 단계(입력, 처리, 출력)를 알 수 있다. c. 인공지능은 알고리즘의 특정 유형임을 알고, 데이터셋(dataset), 학습 알고리즘(learning algorithm) 그리고 예측(prediction)으로 구성되어 있다는 것을 알 수 있다. i. 지도 기반 머신러닝 환경에서 분류 문제를 이해할 수 있다. ii. 훈련(training)데이터의 양이 어떻게 지도 기반 머신러닝 모델의 정확성(accuracy)과 안정성(robustness)에 영향을 미치는지 이해할 수 있다. d. 일상생활 속에 있는 인공지능 시스템을 이해하고 인공지능 시스템이 예측하는 것과 인공지능 시스템이 사용하는 잠재적 데이터셋을 추론할 수 있다.

3) 한국교육학술정보원, “모두를 위한 인공지능 윤리”, 2019, 6면.

4) 이하의 학습목표와 학습목표의 하위 항목의 내용은 한국교육학술정보원, “모두를 위한 인공지능 윤리”, 2019, 11-12면에서 인용함. 원문은 MIT Media Lab, “An Ethics of Artificial Intelligence Curriculum for Middle School Students”, 2019.8., pp.7-9.을 참조.

2. 모든 기술 시스템은 사회-기술 시스템이라는 것을 이해할 수 있다. 사회-기술 시스템은 가치중립적이지 않으며, 정치적 쟁점을 가진다는 것을 이해할 수 있다.

- a. ‘최적화(optimization)’라는 용어를 이해하고 인간이 만든 사회-기술 시스템의 목표는 인간이 결정할 수 있다는 것을 인식할 수 있다.
- b. 일상생활 속에 사회-기술 시스템을 알고, 사회-기술 시스템의 외재적 목표(advertised goals)와 실제적 목표(true goals)를 구분할 수 있다.(예를 들어 YouTube의 추천 알고리즘의 경우 회사의 이익 창출을 목표로 하지만 외부적으로는 이용자를 즐겁게 하기 위한 목적으로 홍보된다.)
 - i. 현재 사용하고 있는 사회-기술 시스템의 특징과 확인된 목표(identified goals)를 연결할 수 있다.
- c. 분류 과정 속에서 발생할 수 있는 ‘알고리즘 편향성(Bias)’이라는 용어를 알 수 있다.
 - i. 훈련 데이터(training data)가 머신러닝 시스템의 정확성에 미치는 영향력을 이해할 수 있다.
 - ii. 인간이 훈련 데이터를 수집하고 구성할 수 있다는 것을 인식할 수 있다.
 - iii. 훈련 데이터의 구성이 ‘지도 기반 머신러닝 시스템의 결과’에 어떻게 영향을 미치는지 이해할 수 있다.

3. 사회-기술 시스템에는 많은 이해관계자들이 있으며 시스템은 이러한 이해관계자들에게 각각 다른 방식으로 영향을 미칠 수 있다는 것을 인식할 수 있다.

- a. 사회-기술 시스템과 관련된 이해관계자들을 식별할 수 있다.
- b. 사회-기술 시스템의 결과에 대해 각각의 이해관계자가 관심 갖는 이유를 합리적으로 설명할 수 있다.
- c. 사회-기술 시스템에서 각각의 이해관계자가 가지는 가치를 식별할 수 있다. 예를 들어 사용자의 요구를 충족하기 위해 시스템이 어떤 목표를 가져야 하는지를 설명할 수 있다.
- d. 사회-기술 시스템을 중심으로 윤리 매트릭스를 작성할 수 있다.

4. 사회-기술 시스템의 올바른 목표를 결정하기 위해 인공지능에 대한 기술적 이해와 이해관계자에 대한 지식을 적용할 수 있다.

- a. 윤리 매트릭스를 분석하고 사회-기술 시스템의 새로운 목표를 위해 분석 결과를 활용할 수 있다.
- b. 목표를 달성하기 위해 인공지능 시스템 훈련에 필요한 데이터셋을 식별할 수 있다.
- c. 사회-기술 시스템의 식별된 목표를 반영하거나 이해관계자의 가치를 반영한 기능을 설계할 수 있다.

5. 기술이 세계에 미치는 영향력을 고려할 수 있다.

- a. 기술로 인해 발생하는 2차적, 3차적 영향력을 이해하고, 기술이 다양한 이해관계자를 위해 만들어진다는 것을 알 수 있다.

수업과정의 주요 내용은 다음과 같다.

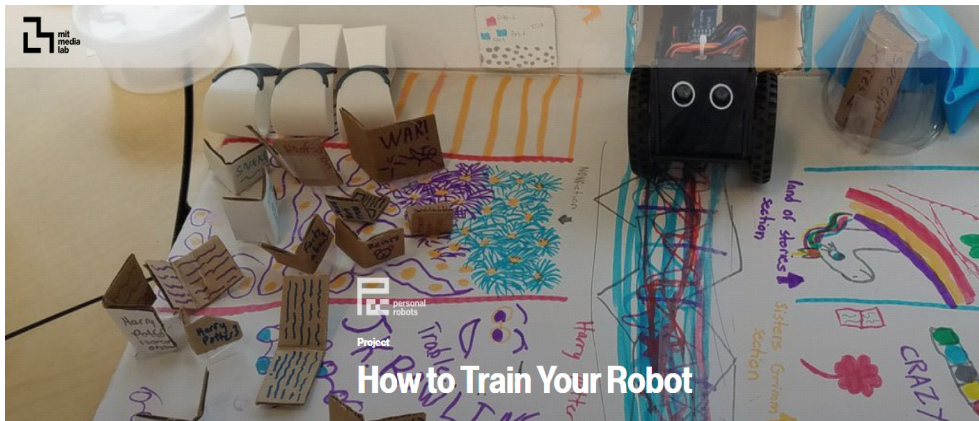
- **(AI 빙고)** 학생들에게 인공지능의 기본 개념에 대해 설명한 후 인공지능 시스템이 제시된 빙고 카드를 나누어준다. 학생들은 각 빙고 칸의 인공지능 시스템을 사용해 본 경험이 있는 학생들과 각 인공지능 시스템이 어떤 예측을 하고 어떤 데이터셋을 사용하는지 함께 살펴본다.
- **(지도학습 머신러닝과 알고리즘 편향성)** 학생들은 편향된 데이터셋을 통해 지도 기반 머신러닝에 대해 학습하게 되며, 이후 새롭게 재구성한 데이터셋으로 분류기를 다시 학습시켜 차이점을 확인할 수 있다.⁵⁾
- **(누구를 위한 알고리즘인가?)** 학생들은 ‘최고’의 땅콩버터잼 샌드위치를 만들기 위한 알고리즘을 작성해 보면서 ‘최고’의 의미에 대하여 생각해 보고, 알고리즘에 자신의 의견이 어떻게 반영되는지 살펴볼 수 있다.
- **(윤리 매트릭스)** 땅콩버터잼 샌드위치 알고리즘을 설계할 때 알고리즘에 관심이 있는 이해관계자가 있음을 파악하여 이해관계자들이 하나의 알고리즘이 가지는 여러 가치를 식별해 보고, 겹치거나 충돌하는 가치를 파악하기 위하여 윤리 매트릭스를 채워 넣는다.
- **(YouTube를 재설계하라!)** 학생들은 YouTube 추천 알고리즘 중심으로 윤리 매트릭스를 구성하며, 이를 기반으로 알고리즘의 목적을 결정한다. 학생들은 새로운 YouTube 버전에 대한 프로토타입을 작성하고, 식별된 이해관계자의 가치를 충족시키는 기능을 상상해 본다.
- **(YouTube 소크라테스식 세미나 활동)** 학생들은 YouTube 플랫폼의 광고 매칭 알고리즘, 추천 알고리즘, 댓글 분류기 등과 같은 인공지능 시스템을 살펴보고 각각의 시스템에서 알고리즘이 예측하고자 하는 바와 알고리즘이 사용하는 데이터셋을 확인해 본다.
- **(미래 모습 상상하기: 공상과학소설)** 학생들은 감정 인식 소프트웨어 및 GANs 같은 다양한 기술과 상호작용해 보며, 누가 기술에 영향을 받게 될 것이며, 기술이 미래에 야기할 득실에 대해 창의적으로 글을 작성해 본다.⁶⁾

5) 한국교육학술정보원, “모두를 위한 인공지능 윤리”, 2019, 14면.

6) 해당 교육의 교육과정과 학습활동은 한국교육학술정보원, “모두를 위한 인공지능 윤리”, 2019.에서 구체적으로 번역, 정리하고 있다.

나. MIT Media Lab의 How to Train Your Robot

How to Train Your Robot은 AI4K12에서 제시한 국가 지침에 기반하여 개발된 MIT 미디어랩의 중학교 AI 및 AI 윤리교육 프로그램이다. 5~8학년 학생들이 인공지능과 윤리를 탐구하기 위한 커리큘럼⁷⁾으로 인공지능의 기술에 대한 인식을 높이는 것을 목표로 한다.⁸⁾ 학생들은 해당 과정을 통해 인공지능이 오늘날 사회에 어떤 영향을 미치고 있는지에 대해 배우기 위해 토론과 실습, 창의적 활동에 참여하게 된다. MIT는 i2Learning⁹⁾과 협력하여 커리큘럼을 개발하였으며 매사추세츠 주변 학교의 일반 교사들에게 제공되고 있다. 해당 교육은 활동에 대한 피드백을 얻기 위하여 중학교 교사를 통해 실제 학교에서 시범적으로 적용하기도 하였다.



[그림 II-2] MIT Media Lab의 How to Train Your Robot 홈페이지

1) 교육대상

AI+Ethics의 대상인 중학생과 AI 윤리에 대해 학습하고자 하는 교사 모두를 위해 구성된 교육과정이다.

7) MIT Media Lab, Project How to Train Your Robot, Available at: <https://www.media.mit.edu/projects/ai-5-8/overview/>, Accessed: 2024.6.1.

8) Unite.AI, “New AI Curriculum Designed for Middle School Students”, 2022.12.9., Available at: <https://www.unite.ai/new-ai-curriculum-designed-for-middle-school-students/>, Accessed: 2024.6.1.

9) i2Learning은 STEM커리큘럼 및 전문성 개발 세션을 개발하는 조직이다.

2) 운영방안

해당 교육은 활동에 대한 피드백을 얻기 위하여 시범 운영을 하였는데, i2Learning의 여름 캠프에서 시범 운영한 바 있고, 매사추세츠 지역의 교사들에게 공유하고 학교에서 시범 운영하기도 하였다. 또한 Mass STEM Week에서 지역 학교의 교사들에게 2일간 20시간 이상의 교육을 통해 AI 콘텐츠를 제공할 수 있도록 하고 3개 학교 6개 교실에서 수업을 실시하였다.¹⁰⁾ 현직 중학교 교사와 두 차례의 교실 학습을 진행하기도 하였는데, 첫 번째 학습은 수업일 기준 5일 동안 교실에서 진행되었으며, 두 번째 학습은 전염병 발병으로 인해 교육과정을 전환하여 2시간 30분 정도의 다섯 개 세션으로 구성된 온라인 강좌를 통해 운영하였다.¹¹⁾

3) 세부 프로그램 주요 내용

시범 운영된 교육에서 학생들은 인공지능과 머신러닝의 다양한 윤리적, 기술적 개념을 탐구하였다. 윤리적 개념에 대해서는 AI+Ethics 프로그램에서 확인할 수 있었던 것처럼 인공지능의 긍정적인 영향과 부정적인 영향을 판별하고, 윤리 매트릭스를 사용하여 이해관계자 분석을 수행하는 방법에 대해 학습하였다. 각 세션은 이론 학습과 실습을 결합하여 학생들이 적극적으로 참여하고, 이해할 수 있도록 설계되어 있어 AI와 AI 윤리의 개념을 익히고 이를 실제로 적용해 볼 수 있다.

-
- 10) MIT News, “Bringing artificial intelligence and MIT to middle school classrooms”, 2019.12.30., Available at: <https://news.mit.edu/2019/bringing-artificial-intelligence-and-mit-middle-school-classrooms-12-30> Accessed: 2024.6.1.
- 11) 시범운영 결과에 대한 사항은 Randi Williams et.al., “Teacher Perspectives on How To Train Your Robot A Middle School AI and Ethics Curriculum”, The Thirty-Fifth AAAI Conference on Artificial Intelligence (AAAI-21), 2021.을 참조

<표 II-1> How to Train Your Robot 세부 프로그램

세션	활동	학습 목표
1	AI인지 아닌지	AI를 정의하고, 무엇이 AI인지 아닌지 이유를 설명할 수 있다.
	윤리적 딜레마 비디오	윤리적 딜레마를 해결하기 위한 전략을 사용할 수 있다.
	Scratch 및 로봇 소개	블록 기반 프로그래밍을 사용하여 로봇으로 미니 프로젝트를 완성할 수 있다.
2	Alexa 피자 배달 앱	알고리즘을 정의하고, 실제 사용을 위해 알고리즘을 설계할 수 있다.
	이미지 분류	데이터셋을 수집하여 이미지 분류기를 훈련시키고, 신경망을 이해할 수 있다.
	알고리즘 편향 비디오 및 기사	알고리즘 편향의 의미를 논의하고, 이를 완화할 방법을 이해할 수 있다.
	Teachable Machine + Scratch	맞춤형 이미지 분류 모델을 만들어 Scratch를 통해 로봇을 프로그래밍 할 수 있다.
3	윤리 매트릭스: Alexa 재설계	이해관계자 분석을 사용하여 기술 설계의 영향을 파악할 수 있다.
	단어 유사성 탐색	단어 벡터와 그들이 언어를 어떻게 인코딩하는지 이해할 수 있다.
	텍스트 분류 + Scratch	K-NN 알고리즘과 이를 머신러닝에서 사용하는 방법을 설명할 수 있다.
4	최종 프로젝트 브레인스토밍	연구를 통해 최종 프로젝트 아이디어를 생성할 수 있다.
	최종 프로젝트 계획	윤리 매트릭스를 사용하여 최종 프로젝트 아이디어를 개발할 수 있다.
	최종 프로젝트 작업 시간	열린 프로젝트에서 시간 관리를 하여 작업할 수 있으며, 동료 피드백을 주고받을 수 있다.
5	쇼케이스	한 주 동안 배운 내용을 적용하여 최종 프로젝트를 완료할 수 있다.
	발표	최종 프로젝트를 일반 청중에게 설명하기 위한 프레젠테이션을 만들 수 있다.

다. 공인 AI 윤리 및 거버넌스 전문가 과정(Certified AI Ethics and Governance Professional™, CAEGPTM)

해당 교육 프로그램은 AI와 관련된 윤리적 위험을 평가하는 방법을 학습하는 교육과정으로 기술 분야의 AI 윤리 담당자와 정책 입안자 등 이해관계자나 전문가 등을 대상으로 실시하는 교육이다(2024년 9월 24~25일 진행 예정). (주)Tonex가 운영하는 인증과정으로 인공지능으로 인한 윤리 및 거버넌스 문제를 탐색하는 데 필요한 지식과 기술을 전문가에게 제공하도록 설계된 포괄적 프로그램이다.¹²⁾ 해당 프로그램은 AI 윤리 및 거버넌스 전문가들이 AI 기술이 책임감 있고 윤리적인 방식으로 개발 및 배포되도록 하는 데 중요한 역할을 한다는 것을 강조한다. 이러한 이해관계자나 전문가들은 윤리지침을 수립하고 투명성과 책임성을 증진하며 공정성과 형평성 문제를 해결하며, 책임 있는 AI 정책을 옹호함으로써 사회 전체에 도움이 되는 방식으로 AI의 미래를 형성하는 데 기여할 수 있다고 한다. 또한 이들은 AI 시스템의 공정성, 형평성, 포용성 등의 문제를 해결할 책임이 있으므로 AI 윤리교육이 필요하다고 제안한다.¹³⁾

NICCS[®]
NATIONAL INITIATIVE FOR CYBERSECURITY CAREERS AND STUDIES

Education & Training Workforce Development Cybersecurity & Career Resources

Education & Training > NICCS Education & Training Catalog > Tonex, Inc. > Certified AI Ethics and Governance Professional (CAEGP)

Certified AI Ethics and Governance Professional (CAEGP)

Online, Instructor-Led

The Certified AI Ethics and Governance Professional (CAEGP) Certification Course by Tonex is a comprehensive program designed to equip professionals with the knowledge and skills needed to navigate the ethical and governance challenges posed by Artificial Intelligence (AI). This course delves into the intricate intersection of technology, ethics, and governance, providing participants with a holistic understanding of responsible AI practices. This CAEGP Certification Course by Tonex is a comprehensive program designed for professionals seeking to navigate the complex intersection of artificial intelligence, ethics, and governance. This course equips participants with a deep understanding of AI ethics principles, frameworks, and risk assessment. Delving into regulatory landscapes and compliance requirements, it empowers individuals to develop and implement effective AI governance strategies.

Learning Objectives

- Understand the fundamental principles of AI ethics and governance
- Explore the ethical considerations in AI development and deployment
- Examine the impact of AI on privacy and data protection
- Learn to assess and mitigate bias in AI algorithms
- Gain insights into the legal and regulatory frameworks related to AI
- Develop strategies for responsible AI decision-making
- Explore the role of transparency and explainability in AI systems
- Understand the socio-economic implications of AI on society
- Learn best practices for implementing ethical AI policies within organizations
- Acquire the skills to conduct ethical AI audits and assessments

Course Overview

Overall Proficiency Level
2 - Intermediate

Course Catalog Number
T101

Course Prerequisites
None

Training Purpose
Functional Development
Management Development

Specific Audience
All

Delivery Method
Online, Instructor-Led

More courses from this provider:
Tonex, Inc.

Contact Information
Tonex, Inc.
6275 W. Plano Parkway, Suite 500
Plano, TX 75093

Visit course page for more information on Certified AI Ethics and Governance Professional...

[그림 II-3] 공인 AI 윤리 및 거버넌스 전문가 과정

- 12) National Initiative for Cybersecurity Careers and Studies 웹페이지, “Certified AI Ethics and Governance Professional™ (CAEGPTM)”, Available at: <https://niccs.cisa.gov/education-training/catalog/tonex-inc/certified-ai-ethics-and-governance-professionaltm-caegptm>, Accessed: 2024.6.1.
- 13) TONEX 웹페이지, “Certified AI Ethics and Governance Professional™ (CAEGPTM)”, Available at: <https://www.tonex.com/training-courses/certified-ai-ethics-and-governance-professional-caegp/>, Accessed: 2024.6.1.

1) 교육대상

본 프로그램은 기술 분야의 AI 윤리 담당자, 정책 입안자 등의 전문가를 비롯하여 AI 윤리 및 거버넌스에 대한 지식을 향상시키고자 하는 개인을 대상으로 한다. AI 전문가, 데이터 과학자, 비즈니스 리더, 정책 입안자, AI 개발 및 배포 또는 의사결정에 관련된 모든 사람들을 포함한다.

2) 운영방안

본 과정은 Tonex사(社)에서 개발한 것으로 텍사스 주에서 2024년 9월 24~25일까지 2일간 진행된다. 본 프로그램을 이수한 사람들은 AI 윤리 및 거버넌스 인증서(Certificate in AI Ethics and Governance)를 발급받게 된다.

3) 세부 프로그램 주요 내용

해당 교육과정의 학습 목표는 다음과 같다.

학습 목표	
• AI 윤리 원칙 및 프레임워크에 대한 심층적인 이해를 한다. • AI 구현과 관련된 윤리적 위험을 평가하고 관리하는 방법을 배운다. • 효과적인 AI 거버넌스 전략을 개발하고 구현하기 위한 기술을 습득한다. • AI와 관련된 규제 환경 및 규정 준수 요구사항을 살펴본다. • AI의 사회적 영향과 책임 있는 배포를 위한 전략을 이해한다.	

교육과정의 주요 내용은 다음 <표 II-2>와 같다.

<표 II-2> 공인 AI 윤리 및 거버넌스 교육과정 주요 내용

모듈	구분	주요 내용
1	AI 윤리 및 거버넌스 소개	- AI 윤리의 진화 - AI 거버넌스의 기초 - 윤리적 AI의 핵심 동인 - AI 생태계에서 거버넌스의 역할 - AI 의사결정의 윤리적 고려사항 - AI 거버넌스의 업계 모범 사례

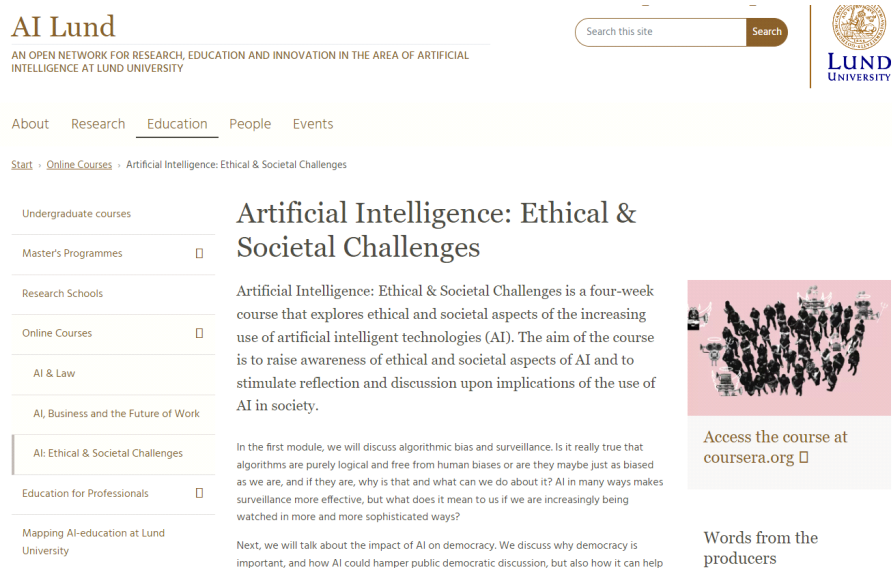
모듈	구분	주요 내용
2	AI 윤리 원칙 및 프레임워크	- AI의 핵심 윤리 원칙 - AI 윤리학의 공리주의와 의무론 - AI 개발에 윤리적 프레임워크 적용 - AI의 윤리적 딜레마에 대한 사례 연구 - 새로운 AI 윤리 기준 - 윤리적 고려사항을 AI 프로젝트 라이프 사이클에 통합
3	AI의 위험평가 및 관리	- AI 프로젝트의 윤리적 위험 식별 - AI의 편견과 공정성의 윤리적 영향 - AI 의사결정 시스템의 윤리적 과제 - 윤리적 위험 완화를 위한 전략 - AI 연구 개발의 윤리적 고려사항 - AI 프로젝트 수명 주기 전반에 걸친 윤리 지침 모니터링 및 조정
4	AI 거버넌스 전략 개발	- 효과적인 AI 거버넌스 구조 설계 - AI 윤리위원회의 설치 - AI 거버넌스를 조직 프레임워크에 통합 - AI 거버넌스를 기업 가치에 맞추기 - AI 의사결정의 책임 보장 - AI 거버넌스 전략의 지속적인 개선
5	규제 환경 및 규정 준수	- 글로벌 AI 규제 프레임워크 - AI 규정 준수에 대한 법적 및 윤리적 고려사항 - AI의 개인정보보호 및 보안 규정 탐색 - AI 시스템의 투명성 보장 - AI의 윤리적 규정 준수 감사 - AI 규정 준수의 과제와 기회
6	사회적 영향 및 책임 있는 AI 배포	- AI의 사회적 영향 분석 - 사회적 공익을 위한 AI의 윤리적 고려사항 - 책임 있는 AI 배포를 위한 전략 - AI 개발에 대한 커뮤니티 참여 - AI 마케팅 및 커뮤니케이션의 윤리적 고려사항 - AI 프로젝트의 사회적 가치 평가 및 전달

2. 유럽

유럽은 최초의 포괄적인 인공지능 규제법 제정을 통과시키며 인공지능 기술과 더불어 기술이 야기할 수 있는 사회 문제에 대해서도 이른 시기부터 활발한 논의가 이루어졌다. 인공지능에 대한 접근성을 향상시키는 동시에 이를 책임감 있게 사용할 수 있는 방안을 교육하고자 유럽연합 및 국가 차원은 물론 학교 차원에서 여러 계층을 대상으로 하는 교육을 운영하고 있다.

가. 스웨덴 : Lund 대학 AI Lund의 ‘인공지능: 윤리적 및 사회적 과제(Artificial Intelligence: Ethical & Societal Challenges)’

AI Lund는 스웨덴 룬드대학교의 인공지능 분야 연구, 교육 및 혁신을 위한 개방형 네트워크로, “인공지능: 윤리적 및 사회적 과제”를 비롯한 온라인 강의를 제공하고 있다. 해당 과정은 인공지능 기술의 윤리적 및 사회적 측면 탐구를 목표로 하여 인공지능의 윤리적 문제와 사회적 영향을 이해하고, 이에 대해 논의할 수 있도록 설계되어 있다.¹⁴⁾



[그림 II-4] AI Lund의 AI 윤리 코스 홈페이지

1) 교육대상

대학생, 대학원생, 연구자를 포함하여 AI 유관 산업에 종사하는 전문가 등 AI와 AI 윤리 문제에 관심을 가지고 있는 모든 학습자를 대상으로 교육이 진행되어 인공지능에 대한 윤리적 관점을 학습할 수 있다.

2) 운영방안

해당 교육과정은 인공지능 기술사용 증가의 윤리적, 사회적 측면을 탐구하는 4주 과정으로, 인공지능의 윤리적, 사회적 측면에 대한 인식을 높이고 사회에서 AI 사용의 의미에 대한 성찰과 토론을 하는 것을 목표로 한다. 온라인 학습 플랫폼

14) Lund대학 AI Lund 웹사이트, “Artificial Intelligence: Ethical & Societal Challenges”, Available at: <https://www.ai.lu.se/mooc/ai-ethics>, Accessed: 2024.6.1.

품인 Coursera(<https://www.coursera.org/learn/ai-ethics>)를 통해 무료로 제공되고 있으며, 탄력적인 일정으로 학습자에게 맞는 속도로 강의를 수강할 수 있다.

강의는 총 네 개의 모듈로 구성되어 있으며, 하나의 모듈에는 여러 개의 강의와 텍스트 자료가 포함된다. 각 강의는 강의에서 얻은 지식을 요약하여 제출해야 하는 과제 및 동료 학습자가 제출한 과제를 평가하는 동료 평가가 필수적으로 수반되어 적극적인 참여를 유도한다.

3) 세부 프로그램 주요 내용

알고리즘 편향 및 감시, 인공지능이 민주주의에 미치는 영향, 인공지능 이용 윤리, 책임과 통제에 네 가지 모듈로 구성된 해당 프로그램은 세부적으로 다음과 같은 내용을 다루고 있다.

<표 II-3> Artificial Intelligence: Ethical & Societal Challenges 프로그램

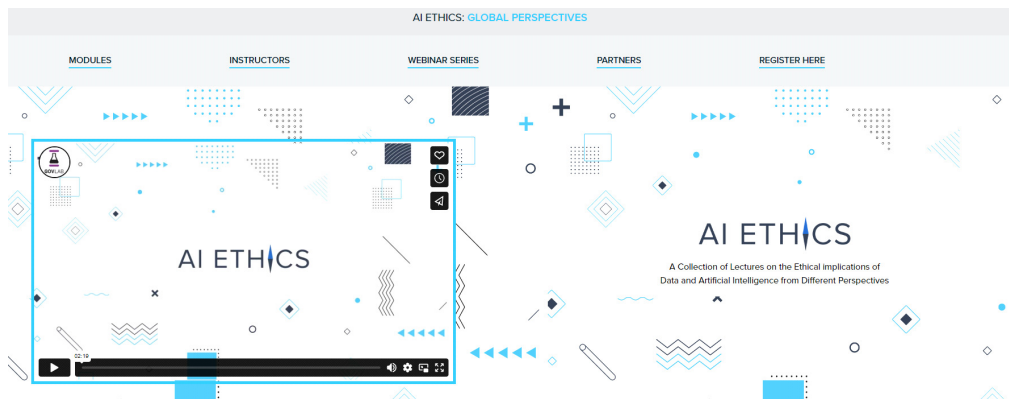
모듈	주제	내용	주요 콘텐츠
1	알고리즘의 편향 및 감시 (Algorithmic bias and surveillance)	- 인공지능 알고리즘이 편향성을 지니는 이유 - 인공지능을 활용한 감시	- 소셜 미디어 및 개인정보보호 - 개인정보보호에 신경 써야 하는 이유는 무엇인가? - 증강AI 비디오 감시와 감시 관련 문제
2	인공지능이 민주주의에 미치는 영향 (Impact of AI on democracy)	- 인공지능이 어떻게 공공 민주주의 토론을 방해할 수 있는가? - 인공지능 어떻게 민주주의 개선에 도움이 될 수 있는가?	- AI와 민주주의-왜 민주주의인가? - 빅데이터와 정치 담론 - 소셜 미디어와 정부의 과도한 영향력 - 민주주의 발전을 위한 AI 도구
3	인공지능 이용 윤리 (Ethics of using AI)	- 인공지능이 의식을 지닐 수 있는가?	- 인텔리전스란 무엇인가? - 인간과 기계의 의식 측정 - 의식과 지능의 관계 - 의식이 생성될 수 있는가? - 인공 의식 연구 - 의식 및 도덕적 상태 - 의식 창조의 윤리적 의미
4	책임과 통제 (Responsibility and control)	- 자율주행 자동차 및 자율주행 로봇이 인간과 충돌하면 책임은 누구에게 있는가? - 인공지능이 안전하고 민주적인 방식으로 개발되도록 하는 책임은 누구에게 있는가?	- 인과적 책임 - 도덕적 책임: 선택의 자유 - 도덕적 책임감: 이해능력 - 미래 지향적 책임 - 영향력 있고 책임감 있는 행위자 - 미래 지향적 책임의 근거 - 책임감 있는 AI? - 로봇공학의 법칙

출처: Coursera - Artificial Intelligence: Ethical & Societal Challenges

과정을 수료하게 되면 수료생들은 ① AI 편향 현상과 감시(surveillance)에서 AI 역할에 대한 기본적 이해, ② AI와 관련된 민주주의의 중요성, ③ 인공의식 생성에 대한 일반적 접근 방식, ④ ‘미래 지향적’과 ‘미래 지향적 책임’의 개념에 대한 기본적 이해와 이러한 개념을 AI에 적용하는 것과 관련된 문제에 대한 지식, ⑤ AI의 제어 문제에 대한 기본적 이해와 이 문제에 대하여 일반적으로 논의되는 솔루션에 대한 지식, ⑥ 이러한 문제의 윤리적, 사회적 측면에 대하여 토론하고 성찰할 수 있는 능력 등을 함양할 수 있게 된다.

나. 독일 : 뮌헨 공과대학교 AI 윤리에 관한 무료 온라인 강좌

해당 강좌는 뮌헨 공과대학교(TUM)의 인공지능윤리연구소, 뉴욕대학교, 글로벌 AI 윤리 컨소시엄이 공동으로 제작한 온라인 강의이다.



[그림 II -5] 뮌헨 공과대학교(TUM)의 무료 온라인 강좌 홈페이지

1) 교육대상

해당 강좌는 연구원, 정치인, 기업의 종사자 및 관리자 등을 대상으로 한다.

2) 운영방안

온라인 무료 강의로 공개되어 있으며, 모듈별(주제별) 약 30분가량의 강의가 업로드 되어 있으며, 강의 대본도 함께 게시되어 있다.

3) 세부 프로그램 주요내용¹⁵⁾

‘AI 거버넌스와 정부 AI’, ‘AI, 예술 및 윤리’, ‘신뢰할 수 있는 도시: 윤리적인 도시 인공지능’ 등의 주제가 모듈형식으로 공개되어 있다. 프로그램의 주요 주제와 내용은 다음과 같다.

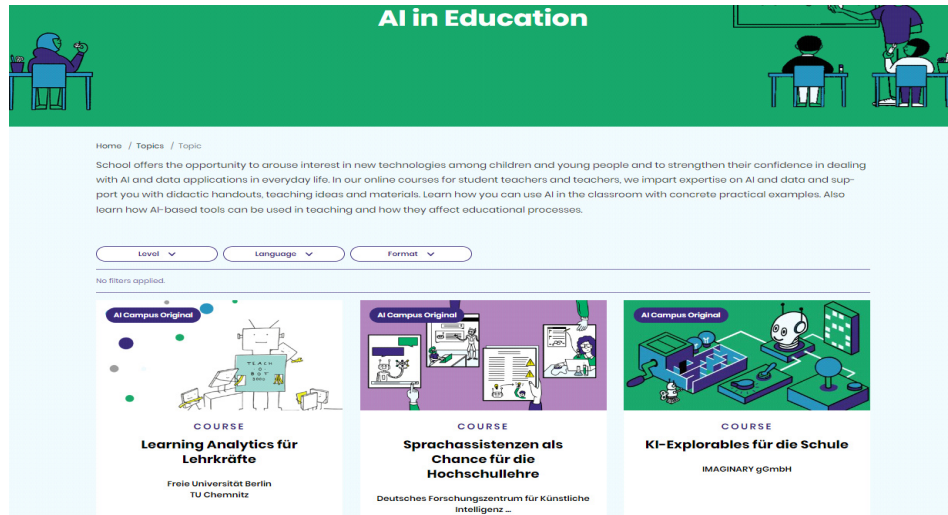
<표 II -4> 독일 뮌헨 공과대 AI 강좌 주요 내용

주요 주제	내용
AI 거버넌스 및 정부 AI	- 정부가 자체 기능에서 AI를 사용하는 방법, 이에 따르는 정책의 주요내용 - 정부에서 AI의 광범위한 사용과 관련한 문제 및 쟁점, 우려사항 등에 대한 내용
AI와 윤리	- 인공지능 윤리와 공평한 연구 관계를 위한 글로벌 행동강령인 TRUST Code의 주요내용
‘AI의, AI와 함께, AI에 의한’ 거버넌스: 공공부문에서 AI의 사용과 영향에 대한 유럽의 관점	- 공공부문에서 AI를 사용하는 유럽의 접근방식 - ① 디지털 유럽 2040 형성, ② "AI의, AI와 함께, AI에 의한" 거버넌스의 트릴레마(trilemma), ③ EU의 공공 서비스에서의 AI; 인간 중심 AI에 대한 EU의 접근 방식 등으로 나누어짐
온라인 연령 확인: 솔루션으로서의 AI?	- 온라인 연령 확인의 어려움 - 미성년자 보호를 위해 현재 존재하는 법적, 기술적 틀 검토 - AI 솔루션의 장점과 한계 분석 - 온라인 AI 기반 연령 확인 시스템을 관리하기 위한 향후 고려 사항
AI 예술 및 윤리	- 제너레이티브 아트 등 예술 분야에서 생성형 AI를 사용할 때 발생하는 윤리적 문제의 탐구
AI와 윤리: 농업	- 농식품 부문에서 AI를 사용한 농장 데이터 수집 및 사용에 대한 검토
신뢰할 수 있는 도시: 윤리적 도시 인공지능	- 도시의 문제를 해결하고 공공서비스 제공을 개선하기 위해 AI를 사용하는 방법
데이터 보호 및 AI: 윤리적 접근과 법적 접근의 관계	- 주요 국가의 다양한 법적 및 윤리적 접근방식에 대한 개요와 차이점 - 데이터보호 규정과 AI법률에 대한 주요내용

15) AI ETHICS: GLOBAL PERSPECTIVES 웹페이지, “MODULES”, Available at: <https://aiethicscourse.org/modules>, Accessed: 2024.6.1.

다. 독일 : KI Campus

인공지능 전문 디지털 학습 플랫폼인 KI Campus는 독일 정부의 AI 전략의 일환으로, 숙련된 인력의 부족 문제를 해결하고 일반 대중이 인공지능을 책임감 있게 다룰 수 있도록 장려하는 것을 목표로 한다.



[그림 II-6] KI Campus의 AI 관련 강의 홈페이지

1) 교육대상

전공과 무관하게 다양한 분야의 대학생과 직장인 및 평생 교육 학습자 등의 일반인을 대상으로 강의를 제공한다.

2) 운영방안

온라인 학습 플랫폼을 통해 무료 온라인 강의를 제공하고 있다. 챗봇, 데이터, 머신러닝, 약학, 교육, 사업 등 다양한 주제에 대한 강의를 선택해서 수강할 수 있으며, 특히 교육 분야에서는 학생과 교사를 위한 과정에서 AI 및 데이터에 대한 전문적인 지식을 제공할 수 있도록 강의 및 교육 자료를 지원한다.¹⁶⁾

16) AI Campus 웹사이트, “AI in Education”, Available at: <https://ki-campus.org/themen/schule>, Accessed: 2024.6.1.

3) 세부 프로그램 주요내용

교직 전공 대학생 및 교사 대상의 인공지능 강의인 ‘학교가 인공지능을 만든다(Schule macht KI)’는 인공지능의 기초, 인공지능과 사회, 인공지능과 언어, 인공지능과 창의성 등 각 120분짜리 전체 네 개의 모듈로 구성되어 있으며, 독일어와 영어로 학습이 가능하다.¹⁷⁾¹⁸⁾

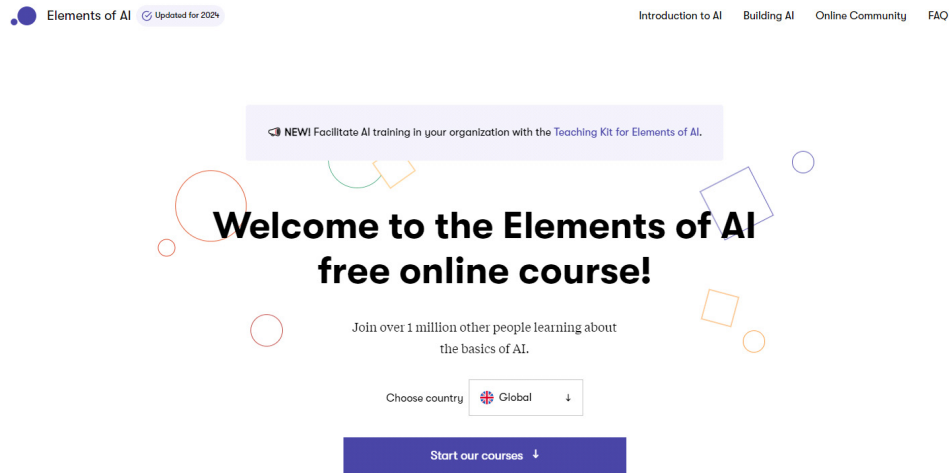
<표 II -5> Schule macht KI 프로그램

모듈	주제	내용
1	인공지능의 기초	- 인공지능의 개념 (인공지능, 머신러닝, 딥러닝) - 인공지능의 역사 - 컴퓨터의 학습 방법
2	인공지능과 사회	- 윤리적인 기계 - 수업이 가능한 기계 - 미래 인공지능의 역할에 대한 다양한 논의
3	인공지능과 언어	- 챗봇의 지능 검사 - 학교에서의 챗봇 프로그래밍
4	인공지능과 창의성	- 인공지능의 사진 수준의 스케치 능력 - 인공지능의 얼굴 인식 필터 만들기

라. 핀란드 : Elements of AI

Elements of AI 프로그램은 핀란드 헬싱키대학교와 온라인 교육 회사 MinnaLearn(당시 Reaktor Education)이 협력하여 개발한 온라인 인공지능 교육 프로그램으로, 이론과 실습이 포함된 AI 기초 및 활용 관련 교육을 제공하고 있다.¹⁹⁾

17) 배지혜, “독일의 인공지능(AI) 융합교육 현황과 시사점”,
<https://edpolicy.kedi.re.kr/firt/boardView.do?strCurMenuId=10091&pageIndex=1&pageCondition=10&nTbBoardArticleSeq=830276>
18) <https://learn.ki-campus.org/courses/kischule2020/overview>
19) Elements of AI 웹페이지, Available at: <https://www.elementsofai.com/>, Accessed: 2024.6.1.



[그림 II-7] 핀란드의 Elements of AI 홈페이지

1) 교육대상

무료 온라인 강의인 프로그램의 특성상 인터넷 환경이 구축되어 있다면 연령·지리적 위치·교육 수준에 무관하게 전 세계 어디에서나 누구든지 프로그램을 수강할 수 있다.²⁰⁾ 현재 170개국에서 100만 명이 넘는 수강생들이 프로그램에 등록하였다.

교육 프로그램이 처음 실시되었던 2018년에는 교육 대상을 핀란드 국민으로 한정하여 핀란드 인구의 1%(5,500명)에게 AI 기초 교육을 제공하는 것을 목표로 하였으나, 이를 조기 달성하여 현재는 EU 인구의 1%에 해당하는 약 5천만 명의 유럽연합 시민들로 목표를 수정하였다. 이를 위해 영어, 독일어, 프랑스어, 핀란드어, 스페인어 등의 EU 언어들을 선택하여 강의를 수강할 수 있도록 마련해 두고 있다.

2) 운영방안

Elements of AI 프로그램은 온라인으로 제공되며, 자율 학습 방식으로 진행되어 학습자는 본인의 속도에 맞추어 각 모듈을 자유롭게 학습할 수 있다. 각 모듈은 강의, 퀴즈, 실습 문제 등으로 구성되어 있어 학습자의 적극적인 참여를 유도하며, 온라인 커뮤니티(<https://community.elementsofai.com/home>)를 제공하여 학습자 간 상호작용이 가능하다.

20) 한국지능정보사회진흥원, “핀란드, AI를 활용한 사람 중심 사회 실현”, 2면, 2022.

3) 세부 프로그램 주요내용

프로그램의 교육 과정은 ‘인공지능 입문(Introduction to AI)’과 ‘인공지능 구축(Building AI)’의 두 부분(part)으로 구성되어 있으며, 부문별로 각각 여섯 개와 다섯 개의 장(chapter)이 존재한다.

1부 인공지능 입문의 제6장 시사점(Implications)은 세 개의 구간(section)으로 구성되어 있는데, 인공지능의 사회적 영향(The societal implications of AI)을 다루는 두 번째 구간은 알고리즘의 편향성, 딥페이크 등을 통한 증거 조작, 개인정보 보호, 업무 방식의 변화 등의 시사점을 언급하며 인공지능 윤리와 관련된 교육을 제공하고 있다.²¹⁾

<표 II -6> 핀란드 Elements of AI 프로그램 주제

장 / 부문	1부 인공지능 입문	2부 인공지능 구축
제1장	AI란 무엇인가?	AI 시작하기
제2장	AI 문제 해결	불확실성 조율
제3장	실생활 AI	머신러닝
제4장	머신러닝	인공 신경망
제5장	인공 신경망	결론
제6장	시사점	-

21) <https://course.elementsofai.com/6/2>

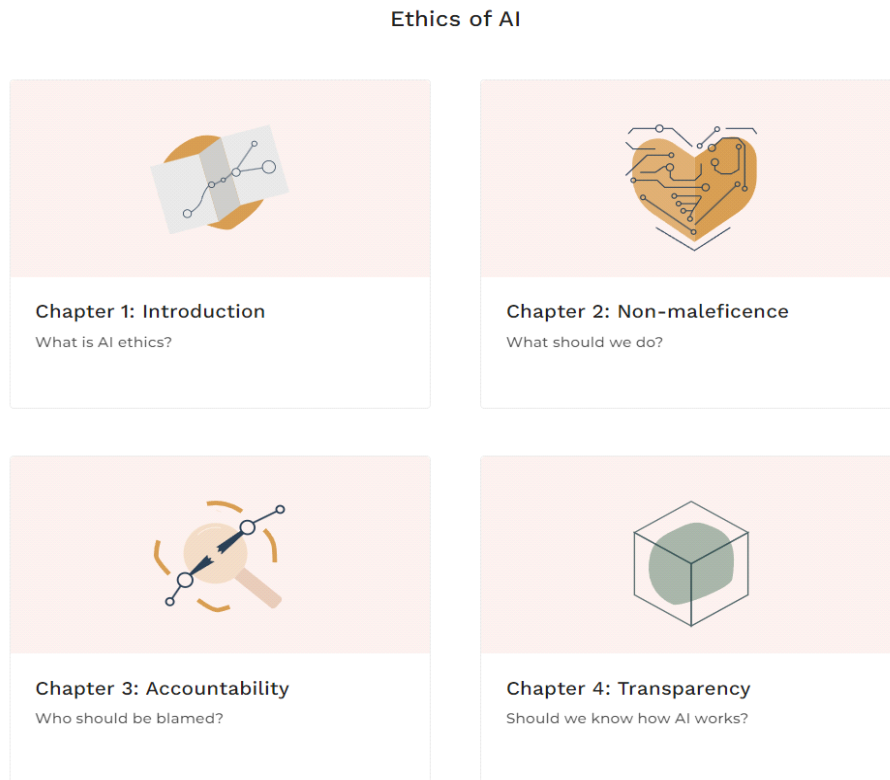
1부 및 2부의 부문별 주요 교육 주제는 다음과 같다.

<표 II -7> 핀란드 Elements of AI 프로그램 주제별 주요 내용

구분	주요 주제
AI란 무엇인가	- AI를 어떻게 정의해야 할까? - 관련 분야 - AI 철학
AI 문제 해결	- 검색 및 문제 해결 - AI로 문제 해결 - 검색 및 게임
실생활 AI	- 확률과 가능성 - 베이즈 규칙(The Bayes rule) - 나이브 베이즈 분류(Naive Bayes classification)
머신러닝	- 기계학습의 유형 - 최근접 이웃 분류(The nearest neighbor classifier) - 회귀
인공 신경망	- 신경망 기본 사항 - 신경망 구축 방법 - 고급 신경망 기술
시사점	- 미래 예측에 대하여 - AI의 사회적 영향
AI 시작하기	- AI가 중요한 이유 - 최적화 - 힐 클라이밍(Hill climbing)
불확실성 조율	- 확률의 기초 - 베이즈 법칙 - 나이브 베이즈 분류
머신러닝	- 선형 회귀 - 최근접 이웃 방법 - 텍스트 작업 - 과적합
인공 신경망	- 로지스틱 회귀 - 로지스틱 회귀에서 신경망으로 - 딥러닝
결론	- AI아이디어 - AI아이디어 갤러리

마. 핀란드 : Ethics of AI

Ethics of AI 프로그램은 인공지능 윤리교육을 제공하기 위하여 헬싱키대학교에서 2020년 개발하였다. 프로그램을 통해 AI가 윤리적으로 수용 가능한 사용 및 개발에 대한 우려를 제기하는 이유를 파악하고, 현대 AI와 연관된 윤리적 문제와 개념에 접근하며, 실습을 기반으로 윤리 이론을 AI에 적용하는 방법에 대해 학습할 수 있다.²²⁾



[그림 II-8] 핀란드의 Ethics of AI 온라인 강의

1) 교육대상

해당 교육은 핀란드 국민을 비롯하여 인공지능의 윤리적 측면에 관심이 있는 전 세계 사람들을 대상으로 한다. 현재 프로그램은 핀란드어, 영어, 스웨덴어로 제공되고 있다.

22) <https://ethics-of-ai.mooc.fi/>

2) 운영방안

Ethics of AI 프로그램은 무료 온라인 강의로, 시스템 로그인 시 진행 상황이 유지되어 학습자는 본인의 일정에 맞추어 강좌를 수강할 수 있다. 강좌는 텍스트로 된 자료 및 과제로 구성되어 있으며, 순서와 무관한 학습이 가능하나 제안된 순서대로 수강할 것을 권장하고 있다. 과제를 수행하기 전 해당하는 섹션의 텍스트 자료를 학습해야 하며, 과정을 수료하기 위해서 과제의 90% 이상을 수행해야 하며, 80%의 과제가 통과되어야 하고, 마지막 모듈의 최종 과제를 필수적으로 제출해야 한다.

3) 세부 프로그램 주요내용

프로그램의 과정은 소개, 무해성, 책임, 투명성, 인권, 공정성, AI 윤리의 실천을 포함한 일곱 개의 장으로 구성되어 있다. 학습 과정에서 코딩 및 특별한 기술이 요구되지 않으며, 학습자가 일곱 개의 장에 대한 텍스트 자료 및 과제 수행을 통하여 인공지능의 윤리적 활용에 대해 숙고해 볼 수 있는 교육이 제공되고 있다.

<표 II -8> Ethics of AI 프로그램

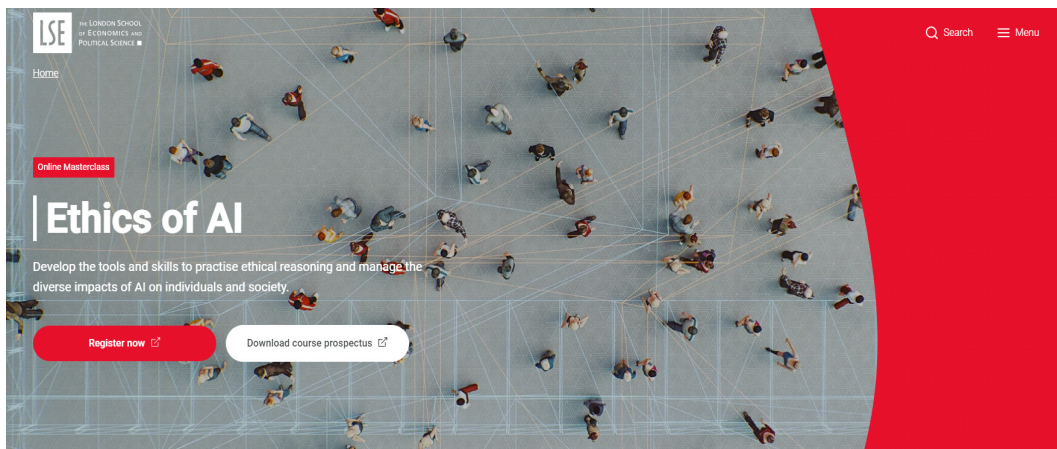
주제	내용
소개	1. AI 윤리 가이드 2. AI 윤리란 무엇인가? 3. 가치와 규범 4. AI 윤리 프레임워크
무해성	1. 우리는 무엇을 해야 하는가? 2. 공동선: 결과 계산 3. 공동선과 웰빙
책임	1. 알고리즘 및 책임 2. 책임이란 무엇인가? 3. 누구를, 무엇으로 탓할 것인가? 4. 책임 분담의 문제
투명성	1. 인공지능의 투명성 2. 투명성이란 무엇인가? 3. 투명성과 개방성의 위험
인권	1. 소개 2. 인권이란 무엇인가? 3. 인권의 예: 프라이버시, 보안, 포용 4. 인공지능과 관련된 아동의 권리

주제	내용
공정성	1. 공정성이란 무엇인가? 2. 공정성의 종류 3. 차별과 편견 4. 편견 제거
AI 윤리의 실천	1. 원칙부터 실천까지 2. 행동으로서의 윤리 3. 윤리의 발전

출처: 한국교육학술정보원, 인공지능의 교육적 역할 및 윤리적 활용; 정보통신정책연구원, 일반 인공지능 윤리교육 교재 개발 및 확산, 2023, 26면.

바. 영국 : 런던 정치경제대학교(LSE)의 Ethics of AI

영국의 런던 정치경제대학교(London School of Economics and Political Science, 이하 LSE)는 AI가 개인과 사회에 미치는 다양한 영향을 관리하고, 윤리적 사고를 실천하기 위한 도구와 기술 개발을 위하여 ‘인공지능 윤리(Ethics of AI)’라는 온라인 마스터 클래스 강좌를 개설하였다.²³⁾



[그림 II-9] 런던 정치경제대학교(LSE)의 Ethics of AI 홈페이지

23) <https://www.lse.ac.uk/study-at-lse/online-learning/courses/ethics-of-ai-masterclass#programmeContent>

1) 교육대상

인공지능이 개인과 사회에 미치는 영향에 관심이 있는 사람으로 폭넓게 교육 대상을 설정해 두었다. 그러나 AI, 데이터 및 일반적인 기술 발전과 관련된 윤리적 문제에 직면한 공공 부문, 민간 부문 및 시민 사회의 전문가와 AI 기술을 직접 구현하거나 기술 변화에 따른 사회적 영향을 관리해야 하는 조직의 리더, 경력 및 개인 성장을 이유로 AI 윤리에 대한 지식을 취득하고자 하는 업계의 전문가를 대상으로 함을 명시하고 있어 해당 프로그램의 주요 대상이 AI 업계 종사자임을 확인할 수 있다.

2) 운영방안

해당 강의는 온라인 교육 제공업체인 GetSmarter와 협력하여 운영되고 있으며, 주당 6~8시간 정도 3주 동안 진행되는 과정이다. 수강생들은 실시간 수업에 참여하거나 AI 윤리 관련 조사를 수행하는 등 적극적인 방식으로 교류할 수 있으며, 과정을 완료한다면 수료증을 취득할 수 있다.

3) 세부 프로그램 주요 내용

강의는 불평등, 공정성, 투명성 등의 도덕적 개념을 실제 상황에 적용하여 윤리적 딜레마를 이해하고, 채용 및 직원 감독을 포함한 다양한 비즈니스 상황에서 AI 기술이 어떻게 사용될 수 있는지에 대해 살펴보는 등 실생활에 밀접하게 적용할 수 있는 내용으로 구성되어 있다.

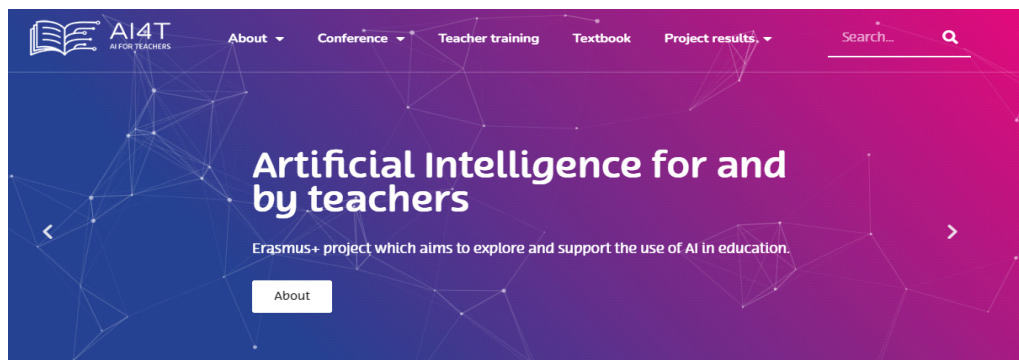
<표 II-9> 영국 Ethics of AI 세부 프로그램

모듈	주제	내용
1	AI와 국가 (AI and the State)	- 민주주의, 정당성, 투명성 - 민주주의 및 행정권 이론의 기초 - AI가 민주주의에 야기할 수 있는 문제
2	AI와 비즈니스 (AI and Business)	- 알고리즘 편향 및 불평등 - AI가 비즈니스에서 자원, 기회, 권력의 불평등을 조성하는 상황 파악
3	AI와 사회 (AI and Society)	- 과제 및 기회 - AI에 대한 기업의 사회적 책임, 위험에 따른 의사결정 등에 대한 개념 이해

출처: LSE - Ethics of AI

사. 유럽연합 : AI4T(AI for Teachers)²⁴⁾

AI4T는 프랑스, 슬로베니아, 이탈리아, 아일랜드, 룩셈부르크가 참여하는 에라스무스+ K3(Erasmus+ K3) 프로젝트의 지원을 받아 시작된 프로그램으로, 고등학교 수학, 과학 및 현대 영어 과목에서 인공지능을 활용한 교육을 할 수 있도록 교사를 지원한다. 해당 프로그램은 AI 도구를 사용하는 교사 교육 과정을 개발하고, 그 효과를 평가하여 궁극적으로 교육의 질을 향상시키는 것을 목표로 한다.



[그림 II-10] 유럽연합의 AI4T 홈페이지

1) 교육대상 및 운영방안

교사들이 전문적인 강의를 하는 과정에서 인공지능 요소를 포함하는 교육 자료를 분석할 수 있는 역량을 강화하고, 학생들에게 사용되는 데이터 및 인공지능의 특성 및 목적을 설명할 수 있도록 훈련시키는 것을 목표로 한다. 이는 유럽연합 집행위원회가 새로운 교육을 위한 디지털 아젠다를 설정한 방향과 흡사하다.

따라서 중등학교 교사들을 대상으로 하여 AI 도구 사용법에 관한 교육 과정을 구축하였으며, IAI ClassCode MOOC을 기반으로 온라인 강의를 통해 과정이 진행된다. 교육의 내용 등은 참여 국가의 필요 등에 따라 조정될 수 있다.

교사들이 교육을 수강하는 동시에 설문 조사를 수행하여 해당 프로그램이 AI 기반 도구를 사용한 교육과 관련하여 각각 교사 및 학생의 태도에 미친 영향 등을 조사하여 효과를 평가하고자 한다.

24) <https://www.ai4t.eu/about>

2) 세부 프로그램 주요내용

‘AI4T: IAI MOOC’을 통해 수강할 수 있는 “Artificial Intelligence... with intelligence!” 강의는 교사, 교육자 등의 교육 전문가를 대상으로 인공지능에 대한 기본적인 개념 및 윤리를 습득하기 위해 마련되었다.

총 세 가지의 모듈로 구성되어 있는 강의는 온라인으로 제공되어 자율적인 수강이 가능하다. 각 모듈마다 퀴즈 및 연습을 통한 자기 평가, 교육용 게임, 제출한 과제 및 발표물을 바탕으로 진행되는 동료 평가 등을 통해 학습자의 적극적인 참여를 유도하고 있다.

<표 II-10> AI4T 세부 프로그램

모듈	주제	내용
1	인공지능이란 무엇인가? (Did you say AI? What is artificial intelligence?)	1. 소개 2. 누가 AI를 두려워하는가? 3. 첫 번째 프로그램 테스트 4. AI란 무엇인가? 5. (토론) AI의 모든 측면 6. 이론 관련 퀴즈 7. 실습 관련 질문
2	머신러닝 및 데이터 이해 (Boosted with AI! Machine learning and data)	1. 소개 2. AI의 모든 측면 3. 데이터 다루기 4. 머신러닝 및 데이터 5. (토론) AI, 교육을 위한 도구 6. 이론 관련 퀴즈 7. 실습 관련 질문
3	인간과 AI (Humans and AI: Artificial intelligence at our service?)	1. 소개 2. 인간이 쓰레기 처리장에 좋은가? 3. AI로 창작해 보자 4. 우리가 제공하는 AI 5. (토론) AI와 인간 6. 이론 관련 퀴즈 7. 실습 관련 질문

출처: Inria LearningLab - AI4T: synopsis of the IAI MOOC

3. 기타 국가

급격하게 발달한 인공지능 기술의 학교 현장 도입 및 올바른 활용을 위해 담당자, 이해관계자들을 대상으로 인공지능 교육을 제공하고 있다. 특히, 인공지능 기술에 대한 역량을 증진시키기 위한 목적으로 인공지능 개념 이해 및 윤리 역량 강화를 목표로 하는 교육 프로그램을 운영하고 있다.

가. 태국 : AI Executive Program in Digital Healthcare

1) 교육대상

디지털 헬스케어의 AI 임원 프로그램(AiX)은 태국의 전자상거래개발청(ETDA) 산하 AI 거버넌스 클리닉(AIGC)에서 개발한 교육 프로그램으로, 의료 등의 공공 서비스에서 AI 기술이 효율적이고 고품질의 서비스를 제공할 수 있도록 공공 및 민간 부문에서의 지식과 이해를 증진하고자 마련되었다.²⁵⁾

2) 운영방안

해당 교육은 AI 임원 프로그램인 만큼 인공지능을 사용하거나 사용에 관심이 있는 공공 및 민간 부문의 임원 및 의사 또는 의료 전문가, 의료 조직에서 네트워크를 구축하는 전문가 등을 대상으로 한다.



[그림 II-11] 태국의 AiX 프로그램

25) AI EXECUTIVE PROGRAM DIGITAL HEALTHCARE 웹페이지, “Program Overview” Available at: <https://dgt.eta.or.th/aix/>, Accessed: 2024.6.1.

오프라인으로 진행되는 교육 과정은 실생활의 사례를 분석하고, 학습자 간 교류를 장려하며, 해외 학술여행을 포함하는 등 소수의 인원으로 구성하여 적극적인 참여를 유도하는 방식으로 구성되었다.

3) 세부 프로그램 주요 내용

4개 파트, 7개 모듈로 구성되어 있는 프로그램은 헬스케어 분야에서의 인공지능에 대하여 사례, 서비스, 해결책 등 다양한 항목을 다루었으며 AI 윤리와 연관된 교육도 제공하고 있다.

<표 II-11> 태국 AiX 세부 프로그램

파트	모듈	주제	내용
헬스케어 분야의 현황과 사례	1	AI 기초 (AI Foundation)	- AI의 개념 - AI의 사용 사례 - AI의 윤리적 문제와 사회적 의미 - 의료 분야의 AI 사용 - 생성형 AI의 실무 경험
임원을 위한 AI 거버넌스 및 도구 키트	2	AI 윤리 및 거버넌스 (AI Ethics and Governance)	- 글로벌 AI 윤리 원칙 및 프레임워크 - 태국의 AI 윤리 거버넌스 및 국가 전략 - 임원을 위한 AI 윤리 지침
헬스케어를 위한 AI 솔루션 및 서비스	3	AI를 위한 데이터 (Data for AI)	- AI 및 생물학, 정보학 - 생의학 연구 데이터 - AI 및 의료 사용을 위한 데이터 공유 - 국제 데이터 공유 및 데이터 표준 - 개인정보보호법 개요
	4	AI 위험관리 (AI Risk Management)	- 의료 분야의 AI 솔루션 - 글로벌 AI 위험관리 프레임워크 - AI 위험 평가
실제 AI 구현 및 활용 사례	5	실무에서의 AI 헬스케어 (AI Healthcare in practice)	- 의료에 적합한 포괄적인 AI 전략 및 로드맵 - 현재 의료 실무에 AI가 적용된 사례 - 디지털 헬스케어 혁신의 기회 및 과제
	6	AI 학술탐방 (AI Study Visit)	- 싱가포르의 AI 거버넌스 탐색 - 싱가포르의 헬스케어 부문 AI 정책 - 책임감 있는 AI의 실제 사례
	7	AI 프로젝트 및 거버넌스 (AI Project and Governance)	- 프로젝트 발표

출처: ETDA - AI Executive Program Digital Healthcare

나. 싱가포르 : 국립연구재단의 ‘LearnAI’

싱가포르는 국립연구재단(National Research Foundation)이 지원하고, 싱가포르 국립대학교(National University of Singapore)가 주관하는 국가 인공지능 이니셔티브인 ‘AI Singapore’ 활동의 하나로 ‘LearnAI’ 프로그램을 시작하였다. 이 프로그램은 인공지능 분야에서 국가적 역량을 강화하고, 국내 인재를 육성하며, 공공 및 민간 부문 조직과 협력하여 세대별 AI 역량을 개발함으로써 AI 생태계를 구축하는 것을 목표로 한다.²⁶⁾²⁷⁾

Students (7 to 12 Years)	Students (Above 13 Years)	Working Professionals
AI for Kids (AI4K)® ● AI4K® Storybook ‘Daisy and her AI friends’ ● AI4K® Online Learning Course (Beginner & Intermediate)	Student Outreach Programme ● AI4I® Literacy in AI certification ● AI4I® Foundations in AI certification ● Google Professional ML Engineer Certificate ● Internship opportunities AI for Everyone (AI4E)® ● AI4E® Online Course ● AI4K® Multilingual Edition	AI for Industry (AI4I)® ● AI4I® Literacy in AI certification ● AI4I® Foundations in AI certification ● AI4I® Proficiency in AI certification

[그림 II-12] LearnAI 세대별 AI 역량 개발

출처: AI Singapore - LearnAI Handbook, 2023.

싱가포르 정부는 LearnAI의 평생학습체계를 바탕으로 세대별 AI 역량을 개발하기 위하여 7세~12세의 학생, 13세 이상의 학생, 전문가로 교육의 대상을 구분하여 AI4K·AI4E·AI4I 과정을 운영하고 있다.

26) AI Singapore 웹사이트, “#PLUSskill”, Available at: <https://learn.aisingapore.org/>, Accessed: 2024.6.1.
27) AI Singapore, “LearnAI Handbook”, 2023, p.4.

1) 교육대상

LearnAI는 [그림 II-12]를 통해 확인할 수 있듯 연령 및 인공지능 관련 배경 지식의 수준 등과 무관하게 모든 학습자를 포용하기 위하여 연령별, 수준별로 구분한 온라인 프로그램을 제공한다. LearnAI의 교육 대상은 다양한데, 어린이를 대상으로 하는 AIK4 과정, IT 또는 소프트웨어 공학 전공이 아닌 일반인으로 대상하는 AI4E 강의, 전문가를 대상으로 하는 AI4I로 나눌 수 있다.

2) 운영방안

어린이를 대상으로 하는 AI4K 과정은 유아·초등학생들에게 지도학습, 컴퓨터 비전, 자연어 처리와 같은 기본적인 머신러닝 개념에 흥미롭게 접근할 수 있도록 설계되었다. 또한, 'Daisy and her AI Friends'라는 스토리 북 발간을 통해 AI 리터러시를 함양하는 방식으로 윤리교육을 진행하고 있다.

IT 또는 소프트웨어 공학 전공이 아닌 일반인을 대상으로 하는 AI4E 강의의 경우, 인공지능의 기본 개념 및 기술의 윤리적 사용을 촉진하는 것을 목표로 과정이 설계되어 있으며, 전문가를 대상으로 진행되는 AI4I는 AI 윤리에 대해 학습할 수 있는 'AI4I - Literacy in AI' 강의를 제공한다.

3) 세부 프로그램 주요 내용

가) AI4K (AI for Kids)

AI4K는 아동이 기본적인 머신러닝 개념을 습득할 수 있도록 혼합 학습(Blended Learning)과 실습 워크숍을 통해 진행되는 프로그램으로, 이를 통해 기본적인 프로그래밍 개념과 최신 AI 기술의 응용을 학습할 수 있다. 대부분의 과정에서는 개념 이해 및 기술적인 부분에 중점을 두고 있으나, 중급 교육과정인 Grade 3에서 AI 윤리에 대해 학습하며 인공지능 기술이 인류에게 가져올 수 있는 이익과 해악을 평가하는 방법에 대해 습득할 수 있다.

<표 II-12> 싱가포르 AI4K 교육 과정

대상 연령		내용
7~8세		Storybook (Daisy and her AI Friends)
9~12세	초급 (Beginner Curriculum)	Grade 1 (온라인 과정 / 약 9시간 소요) 1. 기계와 컴퓨터 소개 (Introduction to machines & computers) 2. 문제해결과 알고리즘 (Problem Solving & Algorithms) 3. 시퀀스와 루프 (Sequences & Loops) 4. Selection Statements (Selection Statements) 5. 코드와 기능 (Code & Functions) 6. AI 소개 (Introduction to AI) Grade 2 (온·오프라인 혼합 과정 / 약 3시간 소요) 7. 당신 주변의 모든 것 (All Around You) 8. 컴퓨터 비전 (Computer Vision) 9. 머신러닝 소개 (Introduction to Machine Learning) 10. 머신러닝에 대한 실험 (Experiments on Machine Learning) 11. AI 게임 만들기 (Building an AI Game)
9~12세	중급 (Intermediate Curriculum)	Grade 3 (온·오프라인 혼합 과정 / 약 6시간 소요) 12. AI 알고리즘 언플러그드 (AI Algorithm Unplugged) 13. 의사결정나무 실습 (Decision Tree in Action) 14. 자연어 처리 - 음성 인식 (NLP-Speech Recognition) 15. 자연어 처리 - 감성 분석 (NLP - Sentiments Analysis) 16. 자연어 처리 - 챗봇 (NLP-Chatbot) 17. AI 윤리 Part 1 (AI Ethics Part 1) Grade 4 (온·오프라인 혼합 과정 / 약 6시간 소요) 18. 데이터 품질 (Data Quality) 19. 컴퓨터 비전 - 안면검출 (Computer Vision-Facial Detection) 20. 컴퓨터 시각 - 안면인식 (Computer Vision-Facial Recognition) 21. 컴퓨터 시각 - 동작인식 (Computer Vision-Pose Recognition) 22. 추천 엔진 (Recommendation Engine) 23. 언플러그드 학습 강화 (Reinforcement Learning Unplugged)

출처: AI Singapore, AI for Kids (AI4K) Grade 1~4²⁸⁾

28) 백수은, 한나, “싱가포르 AI 평생학습체계 분석 및 시사점”, 차세대컨버전스정보서비스기술 논문지, 647면.

나) AI4E (AI for Everyone)

5개의 모듈로 구성되어 있는 온라인 강의의 형태로, 각 모듈을 통해 인공지능의 다양한 측면을 다루고 있다. 각 모듈은 이십 분 정도의 짧은 분량으로 인공지능에 대한 기초 지식을 제공하고 있다.

<표 II-13> 싱가포르 AI4E 세부 내용

모듈	주제	내용
1	AI란 무엇인가 (What is AI?)	- 인공지능의 기본 개념 및 원리 - 인공지능 작동 방식
2	AI는 마법이 아니라 수학과 컴퓨터과학이다 (AI is not Magic, AI is just Maths)	- 인공지능의 기반이 되는 수학적 및 컴퓨터과학의 원리
3	AI는 무엇을 할 수 있는가? (What can AI do?)	- 인공지능 기술의 적용 분야 및 가능성 - 인공지능이 해결 가능한 문제 및 한계
4	책임감 있는 AI (Responsible AI)	- 인공지능의 윤리적 문제 및 책임 있는 사용 - 알고리즘 편향, 개인정보 보호, 투명성 등
5	AI와 일자리, 그리고 나 (AI, Jobs and Me)	- 인공지능이 기존 직업에 미치는 영향 - 인공지능이 새로운 기회를 만들 수 있는 방안

출처: AI Singapore, AI for Everyone(AI4E)

네 번째 모듈인 책임감 있는 AI를 통해 AI 윤리에 대한 기본적인 소양을 기를 수 있도록 교육 프로그램을 제공하고 있다. 현재 제공되고 있는 강의는 2021년에 제작된 버전 3으로, 생성형 AI를 포함한 버전 4를 제작하고 있으며, 발달하는 인공지능 기술 및 윤리를 반영하여 지속적으로 프로그램 콘텐츠를 업데이트하고 있다.

2024년 8월부터 인공지능의 책임, 투명성, 인간 중심, 개인정보 보호 등의 내용을 포함하고 있는 ‘AI 윤리(AI Ethics)’ 강의를 제공할 예정으로, 이를 통해 학습자가 AI 시스템의 설계, 개발 및 배포 과정에서 윤리적인 문제를 식별하고 해결할 수 있도록 지원하고자 한다.²⁹⁾

29) <https://learn.aisingapore.org/courses/ai-ethics/>

다) AI4I (AI for Industry)

AI4I는 인공지능에 대해 기초~중급 이상의 지식을 갖춘 직장인 및 학생을 대상으로 마련된 심화 교육과정으로, AI 문해력(Literacy in AI), AI 기초(Foundations in AI), AI 심화(Proficiency in AI) 등으로 구성되어 있다. AI 기초와 AI 심화는 각각 100시간 이상의 과정으로, 수강을 통해 학습자는 파이썬으로 기본 AI 모델 및 데이터 애플리케이션을 구축할 수 있다. 두 개의 강의를 수강하기 위해서는 AI 문해력 강의가 반드시 선행되어야 한다.

AI 문해력 강의는 약 5시간 분량의 강의로, 이를 통해 일반적인 머신러닝 및 딥러닝의 개념과 방식을 학습하고, 일반적인 AI 프로젝트에서 발생할 수 있는 윤리 문제를 탐구한다.

<표 II-14> 싱가포르 AI4I Literacy in AI 세부 내용

모듈	주제	내용
개요 (Overview)		
1	AI 기초 (AI Basics)	- 모두를 위한 머신러닝 (Machine Learning For Everyone)
2	AI 윤리 (AI Ethics)	- 소개 (Introduction) - 윤리적 AI를 위한 지침 및 원칙 (Guiding Principles for Ethical AI) - 제품 및 모델 설계의 윤리 문제 (Ethics Issues In Product and Model Design) - 데이터 수집 및 접근에 관한 윤리 문제 (Ethics Issues in Data Collection and Access) - 모델 평가의 윤리 문제 (Ethics Issues in Model Evaluation) - 맺음말 (Closing Words)
실습 (선택 사항)		

출처: AI Singapore, AI4I - Literacy in AI

다. 캐나다 : Artificial Intelligence Ethics Micro-Credential

캐나다의 온라인 대학인 아타바스카 대학(Athabasca University)은 인공지능 윤리에 대한 교육 및 컨설팅 등을 제공하는 기업인 Ethically Aligned AI와 협력하여 캐나다 최초의 AI 윤리 분야 마이크로 자격 증명(micro-credential) 프로그램을 출시하였다.



[그림 II-13] 캐나다 아타바스카 대학의 AI 윤리 마이크로 자격 증명

1) 교육대상

기업들이 AI 지원 솔루션 구축을 서두르면서 책임감 있는 선택이 이루어져야 했으나, 업계의 전문가들이 인공지능 윤리에 대한 교육 배경이 없었음을 고려하여 해당 과정을 마련한 것이라고 한다.³⁰⁾ 따라서 본 교육의 대상은 업계의 전문가와 이해관계자들이다.

2) 운영방안

과정은 2주 안에 수강하도록 설계되었으나, 온라인 강좌의 특성을 고려했을 때 유연한 학습이 가능하다. 가상의 시나리오 기반으로 AI 캐릭터와 상호작용하고, 새롭게 개발된 기술을 적용하여 과제를 진행하는 체험 시뮬레이션을 통해 이론 습득부터 실습까지 다양한 형태의 참여가 가능하다.

3) 세부 프로그램 주요 내용³¹⁾

윤리 마이크로 자격 증명(Artificial Intelligence Ethics Micro-Credential)은 AI 시스템의 윤리적 개발과 관련된 문제를 해결하는 데 필요한 기본적인 기술을 제공하는 네 개의 과정으로 구성된 프로그램으로, 소개, 데이터, 머신러닝 모델, 로봇 윤리의 내용이 포함된다.

30) <https://www.athabasca.ca/news/announcements/powered-by-athabasca-university-and-ethically-aligned-ai-launch-canadas-first-micro-credential-in-ai-ethics>

31) <https://learninglab.inria.fr/en/ai4t-synopsis-mooc-iai/>

<표 II-15> 캐나다 인공지능 윤리 마이크로 자격 증명 AI 세부 프로그램

과정	내용
AI 윤리: 소개 (AI Ethics: An Introduction)	- AI의 기본적인 작동 방식 - AI의 일반적인 오해 - AI의 응용 및 한계 - AI의 윤리적 문제를 해결하기 위한 철학적인 접근 방식
AI 윤리: 데이터 (AI Ethics: Data)	- 데이터 정의 및 주권을 포함한 데이터 관련 윤리적 문제 탐색 - 머신러닝 모델 훈련에서의 데이터의 역할 파악 - 데이터셋에 반영되는 편향성 및 이를 방지하기 위한 전략 - 데이터 보안
AI 윤리: 머신러닝 모델 (AI Ethics: Machine Learning Models)	- 머신러닝의 적용 및 한계 - 머신러닝과 관련된 윤리적 문제 탐색 - AI 시스템이 부당한 결과를 생성하는 데 데이터 편향성이 미치는 영향
AI 윤리: 로봇 (AI Ethics: Roboethics)	- 기존 로봇 및 자율 로봇의 응용과 한계 - 로봇 시스템의 개발 및 사용과 관련된 윤리적 문제 - 인간과 로봇의 상호작용

출처: PowerED - Artificial Intelligence Ethics Micro-Credential

라. 유네스코 : 윤리교사 연수과정 (Ethics Teacher Training Courses, ETTC)

유네스코는 윤리교사 연수과정을 마련하고 윤리교육 방법론에 대한 교사를 양성하고 있다. 이를 윤리교사 연수과정(Ethics Teacher Training Courses, ETTCs)³²⁾이라고 하는데, 최근 Gadjah Mada 대학(UGM)과 협력하여 AI 윤리에 중점을 둔 윤리교사 연수 과정을 진행한 바 있다. UGM 철학 학부의 교수진이나 강사 등을 대상으로 실시되었고 약 50여 명의 강사가 참여하였다. 유네스코는 본 과정에 대해 교수진과 학계를 참여시켜 ChatGPT와 같은 범용AI의 출현과 관련된 우려와 과제를 논의하는 것을 목표로 한다고 하였다.³³⁾ 해당 과정은 AI 기술 동향, 기회 및 AI 윤리 철학 등 윤리교육에 대한 글로벌 관점을 소개하여 생명윤리, 과학, 기술 등 다양한 학문 분야에 속한 교사의 AI 윤리 역량 강화를 목표로 하는 커리큘럼으로 구성되어 있다.

32)

UNESCO 웹사이트. “Ethics Education Programme”, Available at: <https://www.unesco.org/en/ethics-science-technology/education>, Accessed:2024.6.1.

33)

Universitas Gadjah Mada 웹사이트, “UGM Faculty of Philosophy Collaborates With UNESCO to Organize Training on Drafting AI Ethics Guidelines”, 2023.8.29., Available at: <https://ugm.ac.id/en/news/ugm-faculty-of-philosophy-collaborates-with-unesco-to-organize-training-on-drafting-ai-ethics-guidelines/>, Accessed: 2024.6.1.

UGM Faculty of Philosophy Collaborates With UNESCO to Organize Training on Drafting AI Ethics Guidelines

Faculty Corner, News Report ● 29 August 2023, 14.41 By : Salma



[그림 II-14] 인도네시아에서 진행된 AI 윤리 교사 연수 과정

Ⅲ. 국내 인공지능 윤리교육 사례

1. 공공분야 개발 사례

국내 공공분야가 주도하는 윤리교육 운영의 경우, 인공지능 기술을 수용하기 위해 학생들이 갖추어야 하는 인공지능 윤리 역량 함양을 목표로 하는 교육 프로그램이 제공되고 있다. 또한, 개발자나 운영자 등 인공지능의 설계와 밀접하게 관련된 성인이나 학생의 진로를 고려하여 인공지능 기술의 전 과정에 거친 책임 의식 향상을 목표로 하는 프로그램을 제공한다.

가. 과학기술정보통신부, 정보통신정책연구원의 인공지능 윤리교육

과학기술정보통신부와 정보통신정책연구원(KISDI)은 초·중·고등학생들의 인공지능 역량 강화를 목적으로 2022년 인공지능 윤리교육 교재를 개발하였다. 2020년 과학기술정보통신부가 발표한 ‘인공지능(AI) 윤리기준’의 3대 원칙과 10대 요건에 맞추어 인공지능 시대에 살아가는 시민으로서 갖추어야 하는 인공지능 윤리 역량을 함양하는 것을 목표로 개발한 것이다. 학교급별로 1종씩 총 3종으로 구성되어 있으며, 각 교재마다 학습용 학습교재와 교사용 지도교재 총 6권을 정보통신정책연구원 홈페이지를 통해 제공하고 있다.³⁴⁾

또한 2023년에는 일반 성인을 대상으로 하는 교재 개발도 시작하였다. 일상에서의 인공지능 활용이 확산되면서 개별 사용자의 인공지능 활용 및 대응 능력에 대한 필요성이 대두되고 있기 때문이다. 인공지능의 편향된 데이터 학습 결과에 대한 판단과 이의제기 역량, 인공지능 결과물에 대한 비판적 판단과 주체적 활용 등이 요구되므로 이를 위하여 일반 시민에 대한 인공지능 윤리교육을 개발한 것이다.³⁵⁾

34) 정보통신정책연구원 홈페이지, “KISDI, 초·중·고 학교 현장에서 활용 가능한 ‘인공지능 윤리교육 교재’ 공개”, 원문: <https://www.kisdi.re.kr/bbs/view.do?bbsSn=113985&key=m2101113056011&pageIndex=1&sc=&sw=>, 접속일: 2024.6.1.

35) 조성은 외, 일반 인공지능 윤리교육 교재 개발 및 확산, 정보통신정책연구원, 2023, 13-14면.

1) 교육대상

해당 교육은 초등, 중등, 고등, 일반시민 대상 4가지로 구성되어 있다.



[그림 III-1] 과학기술정보통신부, KISDI의 인공지능 윤리교육 교재(초·중·고등학교)

2) 운영방안

개발된 교육교재는 초·중·고등학교 교육 자료의 경우 인공지능 윤리교육이 필요한 경우 활용할 수 있도록 온라인을 통해 공개되어 있으며, 이를 활용하고자 하는 교사 및 강사를 대상으로 한 온라인 설명회 자료도 공개되어 있다.

3) 세부 프로그램 주요 내용

가) 초등 인공지능 윤리 교재

해당 교재는 학생의 인지 수준에 맞추어 놀이 중심으로 인공지능 윤리의 개념을 구성하고 체험의 기회를 제공하고 있다. ‘인공지능 윤리기준’의 3대 원칙과 10대 핵심 요건을 다루도록 고안되었고, 이를 바탕으로 학교 내외에서 선택적으로 활용할 수 있도록 모듈형 프로그램으로 구성하고 있다.³⁶⁾

36) 과학기술정보통신부·정보통신정책연구원, 초등학교 교사용 인공지능 윤리 놀이중심, 2022, 인공지능 윤리 교재 구성 페이지.

<표 III-1> KISDI 초등용 인공지능 윤리 교재 주요 골자

구분	주제	세부 주제
1단원	인간의 존엄성 지키기	- 인권보장: 인공지능이 꽃피는 인공지능 세상을 만들어요 - 프라이버시 보호: 우리의 사생활은 안전한가요? - 다양성 존중: 인공지능에 편견과 차별을 없애요 - 침해금지: 피해를 주지 않아요
2단원	목적에 맞게 기술 사용하기	- 데이터관리: 올바르게 데이터를 관리해요 - 책임성, 투명성: 책임을 갖고 투명하게 사용해요 - 안전성: 안전한 인공지능을 만들어요
3단원	함께 좋은 세상 만들기	- 공공성: 모두가 행복한 세상을 만들어요 - 연대성: 지속 가능한 미래 마을을 만들어요

출처: 과학기술정보통신부·정보통신정책연구원, 초등학교 학생용 인공지능 윤리 놀이중심, 2022.

<표 III-2> KISDI 초등용 인공지능 윤리 교재 주제별 구성 의도

주제	세부 주제	구성 의도 주요 내용	
인간의 존엄성 지키기	인권보장	1차시	인권을 위한 인공지능 기술을 알아보고 인공지능 사용으로 인권이 보장받지 못하는 문제를 생각한다. 인권의 개념과 인공지능 사용으로 인권이 보장받지 못한 사례를 알아본 뒤, 이어서 인권을 존중하는 인공지능 원칙을 만드는 활동을 한다.
	프라이버시 보호	2차시	보호받아야 될 프라이버시와 개인정보를 알아보고 인공지능 사용으로 생활 속에서 사생활이 침해된 다양한 사례를 조사한다. 조사한 사례를 친구들과 나누며 인공지능 사용으로 사생활이 침해된 사례가 많음을 공감한다.
		3차시	실제 다른 나라에서 있었던 사례를 보며 우리 교실에서도 이 사례와 같은 일이 생긴다면 어떤 기분이 들지 학생들과 충분히 이야기를 나눈다. 인공지능 기반 CCTV와 기존 CCTV를 비교하고, 사생활이 보호될 수 있도록 CCTV 설치 안내문을 친구들과 토의하며 작성한다.
	다양성 존중	4차시	실제로 편견과 차별이 드러난 대화형 인공지능 챗봇 내용을 각색하여 편견과 차별을 줄이는 대화를 학생들이 직접 만든다. 일상생활에서 사용하는 인공지능이 계속 편견과 차별을 드러낸다면 어떤 문제가 생길 수 있는지 친구들과 함께 고민해 보고 토의하며 작성한다.
		5차시	우리 주변의 편견과 차별을 줄이고 다양성을 존중하기 위해 만들어진 발명품들에는 어떤 것이 있는지 조사해 보고, 그를 바탕으로 학생들이 인공지능 기술과 접목하여 다양성을 존중한 발명품을 고안해 본다.

주제	세부 주제	구성 의도 주요 내용	
	침해금지	6차시	인공지능을 사용하는 과정에서 의도치 않은 위험과 침해 사례가 발생할 수 있음을 알아보고 인공지능 활용 상황에서 발생하는 위험을 직접 찾아본다. 인공지능의 개발 목적과 위험 요소를 찾는 메모리 카드 게임 또는 자율 주행 드론의 위험에 대한 질문을 만들고 친구의 의견을 물어보는 가상 인터뷰 활동 등(선택활동)을 통해 인공지능의 위험 요소에 대해 깊이 생각해 본다.
		7차시	인공지능 기술의 위험 요소를 찾고 해당 위험요소를 줄이기 위한 방법을 찾아 발표한다. 자율 주행 드론을 안전하게 사용하기 위한 규칙을 만들어 직접 인공지능의 위험 요소를 줄이기 위한 노력을 실천해 본다.
목적에 맞게 기술 사용하기	데이터 관리	1차시	데이터를 수집한 후 수집한 데이터에 따라 인공지능의 결과가 달라질 수 있음을 이해한다. 데이터를 직접 수집하고 데이터로 인공지능 결과를 예측해 보는 활동을 한다.
		2차시	올바른 데이터란 무엇인지 생각해 보고, 꽃 분류 프로그램을 만들기 위한 올바른 데이터를 수집한 후 꽃 분류 프로그램을 만드는 활동을 한다.
	책임성, 투명성	3차시	인공지능 사용으로 발생하는 다양한 문제 상황에서 인공지능의 개발자, 공급자, 사용자가 각각 어떤 책임을 지고 노력을 해야 할지 생각해 본다.
		4차시	인공지능이 결론을 내리는 과정과 이유가 투명하게 제공되어야 함을 이해한다. 설명 가능한 인공지능의 장점을 찾아보고 인공지능을 사용하는 과정에서 투명성이 필요한 이유를 찾아 인공지능 개발자에게 편지를 써 본다.
	안전성	5차시	안전을 위협하는 인공지능의 다양한 사례와 원인을 이해하고 안전한 인공지능을 위해 해야 할 일을 안다.
		6차시	안전한 자율 주행 자동차 프로그램을 직접 만드는 활동을 하고, 안전한 자율 주행 자동차를 만들기 위한 인공지능 개발자, 공급자, 사용자의 역할을 이해한다.
함께 좋은 세상 만들기	공공성	1차시	공익을 위한 인공지능 기술에 대해 알아보고 인공지능으로 인해 공공성이 침해되는 문제를 생각한다. 이어서 인공지능의 공공성의 의미와 공공성이 확보되지 않았을 때 초래할 수 있는 위험, 부작용에 대해 알아보고 인공지능의 공공성의 중요성에 대해 이해한다.
		2차시	인공지능이 사회에 미치는 영향에 대해 토론하며 인공지능이 사회에 미치는 순기능을 극대화하고 역기능을 최소화하는 방안이 필요함을 이해한다. 인공지능의 공공성을 유지하기 위한 방법을 생각하며 모두를 위한 인공지능 캠페인 도구를 만들어 캠페인 활동을 한다.

주제	세부 주제	구성 의도 주요 내용	
	연대성	3차시	환경을 위한 인공지능 기술에 대해 알아보고 인공지능 활용 자체가 탄소를 많이 배출하고 인공지능을 활용하지 않는 나라의 사람들에게 피해를 주는 문제를 생각한다. 이어서 국제 협력과 지속 가능 발전 목표에 대해 알아보고 인공지능 활용으로 지속 가능 발전 목표에 기여할 수 있음을 이해한다.
		4차시	미래 세대를 배려하기 위한 지속 가능한 발전 목표에 대해 알아보고 관심 있는 지속 가능한 목표를 선택하여 발생한 문제를 해결하기 위해 인공지능 기술을 활용한 아이디어를 생각하고 발표한다.

출처: 과학기술정보통신부·정보통신정책연구원, 초등학교 교사용 인공지능 윤리 놀이중심, 2022

나) 중등 인공지능 윤리 교재

해당 교재는 주 사용자로 학생을 타겟팅하여 구성되었지만, 향후 개발자나 관리자로 성장할 수 있다는 점을 감안하여 인공지능을 활용하는 전 과정에서 갖추어야 할 인공지능 윤리에 대해 생각하고 실천할 수 있도록 교육하는 것에 방점을 두어 개발되었다. ‘인공지능 윤리기준’의 10대 핵심 요건 중 관련 있는 요건들을 한 챕터로 묶어 중, 고등학생들이 체험활동을 통해 인공지능 윤리와 관련된 가치들을 확인할 수 있도록 되어있으며, 하나의 주제가 약 4차시로 구성된다.³⁷⁾

<표 III -3> KISDI 중등용 인공지능 윤리 교재 주요 골자

구분	내용
1차시	- 10대 핵심 요건과 관련된 문제 상황을 제시하고, 핵심 요건을 이해한 후 인공지능 프로그램 개발 과정에서 개발자와 사용자의 역할 이해
2차시	- 인공지능 윤리와 연계한 인공지능 프로그램 개발
3차시	- 2차시에서 개발한 인공지능 프로그램의 사용 과정에서 발생할 수 있는 윤리적 문제를 파악하고, 대응 방안을 탐색하는 과정에서 개발자와 사용자의 관점 차이 분석, 관점의 차이를 좁히기 위한 합의의 필요성 인식
4차시	- 인공지능 윤리 실천을 위한 활동

출처: 과학기술정보통신부·정보통신정책연구원, 중학교 인공지능 윤리 체험중심, 2022

37) 과학기술정보통신부·정보통신정책연구원, 중학교 인공지능 윤리 체험중심, 2022, 머리말 페이지

<표 III -4> KISDI 중등용 인공지능 윤리 교재 주요 내용

구분	주제	세부주제	학습목표
1장	인공지능 윤리, 왜 필요할까?	인공지능 윤리, 꼭 알아야 할까?	- 인공지능 윤리의 필요성과 중요성과 필요성을 깨달을 수 있다. - 인공지능 기술이 주는 가능성과 위험성을 분석하고 예방 및 대처 방안을 윤리적 측면에서 계획할 수 있다.
		인공지능 윤리, 왜 알아야 할까?	- 인공지능 윤리에 대한 관점의 변화를 살펴보고, 윤리적 태도를 갖추어야 하는 주체를 설명할 수 있다. - 인공지능 윤리 가이드라인/권고안을 살펴보고, 인공지능 윤리의 중요성을 설명할 수 있다.
2장	인공지능은 내가 한 일을 얼마나 알고 있을까?	인공지능 개발을 위한 개인정보 다루기	- 인공지능 모델을 활용한 문제 해결을 위하여 개인정보를 수집하는 과정과 방법을 익히고 필요한 정보를 수집·이용할 수 있다.
		프라이버시 보호, 침해금지 이해를 위한 친구 빨리 사귀기 인공지능 앱 만들기	- 친구 빨리 사귀기 인공지능 앱 개발을 통해 프라이버시 보호와 침해금지에 대한 의견을 제시할 수 있다.
		개발자와 사용자는 어떤 생각을 할까?	- 개발자와 사용자의 관점에서 개인정보를 다루는 의견의 차이를 분석할 수 있다. - 인공지능을 개발하고 활용하는 과정에서 프라이버시 침해 및 개인 정보를 오·남용할 때 발생하는 문제를 파악하고 이에 대한 대응 방안을 탐색할 수 있다. - 인간을 위해 개발된 인공지능 기술의 위험 요소를 최소화할 방안을 탐색할 수 있다.
		프라이버시 보호, 침해 금지 실천을 위한 활동	- 인공지능을 활용한 편의 제공과 프라이버시 보호 사이에서 발생하는 논란에 대한 자신의 생각을 표현할 수 있다.
3장	인공지능은 편견이 없을까?	인공지능은 편견 없이 판단할까?	- 편향을 줄일 수 있는 데이터를 추가하는 활동을 통하여 인공지능 데이터의 편향을 이해할 수 있다. - 성별을 판별하는 활동을 통하여 데이터 편향성을 이해하고 줄일 수 있는 노력을 할 수 있다.
		편향성 이해를 위한 옷을 추천하는 인공지능 프로그램 만들기	- 성별을 판별하기 위한 데이터를 수집하여 인공지능 모델을 만들 수 있다. - 성별 분류 인공지능 모델을 활용하여 옷을 추천해 주는 프로그램을 제작할 수 있다. - 인공지능 프로그램 개발 과정에서 개발자의 편견이 반영될 수 있다는 점을 이해하고 이를 고치기 위해 노력할 수 있다.

구분	주제	세부주제	학습목표
4장	인공지능 판단 믿어도 될까?	개발자와 사용자는 어떤 생각을 할까?	- 인공지능을 개발하고 활용하는 과정에서 편향성이 발생하는 문제를 파악하고 이에 대한 대응 방안을 제시할 수 있다. - 개발자와 사용자 입장에서 데이터 관리의 중요성을 파악하고 문제가 발생했을 때 대응 방안을 제시할 수 있다.
		다양성 존중과 데이터 관리 실천을 위한 활동	- 데이터를 라벨링하여 인공지능 모델을 만들어 편향성을 줄이는 문제를 해결할 수 있다. - 편향적일 수 있는 인공지능을 활용한 채용 도입에 대한 자신의 생각을 표현할 수 있다.
		인공지능 결과 믿어도 될까?	- 인공지능 윤리에서 강조되는 책임성과 투명성, 안전성의 개념을 이해하고 설명할 수 있다. - 의사 결정 트리 활동을 통해 인공지능 시스템의 투명성이 필요함을 이해할 수 있다. - 자율 주행 자동차 사고 사례를 통해 인공지능 윤리 책임성과 안전성을 이해할 수 있다.
		투명성 이해를 위한 범죄 형량 판단 인공지능 프로그램 만들기	- 범죄 형량 판단을 위한 올바른 판단 기준을 세울 수 있다. - 인공지능 기술을 활용하여 인공지능 범죄 형량 판단 프로그램을 제작할 수 있다.
5장	인공지능은 모든 사람들을 배려하고 있을까?	개발자와 사용자는 어떤 생각을 할까?	- 개발자와 사용자의 관점에서 인공지능의 투명성과 안전성, 책임성에 대한 의견의 차이를 분석할 수 있다. - 인공지능을 개발하고 활용하는 과정에서 투명성과 책임성, 안전성을 고려할 때 발생할 수 있는 문제를 파악하고 이에 대한 대응 방안을 제시할 수 있다.
		책임성, 투명성, 안전성 실천을 위한 활동	- 인공지능 윤리에서 책임성과 투명성, 안전성의 개념을 설명할 수 있다. - 안전하고 신뢰할 수 있는 인공지능 개발을 위한 윤리원칙을 세울 수 있다.
		인공지능 활용과 사용자 유형 이해하기	- 인공지능 윤리의 공공성과 연대성의 개념을 이해하고 설명할 수 있다. - 인공지능 발달로 인한 실생활의 변화를 인식하고, 사용자 유형을 분석할 수 있다. - 인공지능 도입으로 달성 가능한 지속가능발전목표를 이해하고 이를 위한 방안을 설명할 수 있다.

구분	주제	세부주제	학습목표
		공공성과 연대성 이해를 위한 재활동품 분리수거 인공지능 프로그램 만들기	- 인공지능 공공서비스 시스템 개발을 위한 지구촌 환경 문제 프로젝트를 계획할 수 있다. - 인공지능 공공 서비스 시스템 개발을 위한 다양한 모델링을 할 수 있다. - 인공지능 윤리 의식을 함양하기 위해 모든 사람을 배려하여 편리하게 사용할 수 있는 인공지능 프로그램을 개발할 수 있다.
		개발자와 사용자는 어떤 생각을 할까?	- 인공지능 활용에 있어 사회의 공공성에 대한 개발자와 사용자 관점을 비교 분석할 수 있다. - 인공지능 활용에 있어 공공성 및 연대성 훼손을 최소화하면서 활용하는 방안을 모색할 수 있다. - 인공지능 프로그램 제작과 사용에 있어 개발자와 사용자의 관점에서 공공성을 다루는 의견을 설명할 수 있다.
		공공성과 연대성 실천을 위한 활동	- 인공지능 윤리에서 공공성과 연대성 개념을 설명할 수 있다. - 인공지능 윤리에서 연대성과 관련된 논란에 대한 자신의 생각을 표현할 수 있다.

출처: 과학기술정보통신부·정보통신정책연구원, 중학교 인공지능 윤리 체험중심, 2022.

<표 III -5> KISDI 중등용 인공지능 윤리 교재 세부주제별 학습내용

구분	세부주제	학습단원
1장	인공지능 윤리, 꼭 알아야 할까?	- 해결해야 할 문제는 무엇일까? - 인공지능 윤리 3대 원칙과 핵심 요건 - 인공지능 기술은 어떤 기능을 하고 있을까? - 인공지능 활용에 있어 제일 중요한 것은 무엇일까? - 인공지능 윤리는 필요한 것일까?
	인공지능 윤리, 왜 알아야 할까?	- 인공지능에 윤리가 필요할까? - 인공지능 윤리 얼마나 중요한 걸까?
2장	인공지능 개발을 위한 개인정보 다루기	- 해결해야 할 문제는 무엇일까? - 인공지능 윤리 핵심 요건 - 프라이버시 보호 및 침해금지-개인정보 이해 - 개발자는 무엇을 해야 할까? - 사용자는 어떻게 해야 할까?
	프라이버시 보호, 침해금지 이해를 위한 친구 빨리 사귀기 인공지능 앱 만들기	- 개발 계획 세우기 - 인공지능 모델 만들기 - 인공지능 프로그램 만들기
	개발자와 사용자는 어떤 생각을 할까?	- 개발자 관점에서 한 걸음 더 나아가기 - 사용자 관점에서 한 걸음 나아가기 - 개발자, 사용자 관점을 비교하여 의견 정리하기

구분	세부주제	학습단원
	프라이버시 보호, 침해 금지 실천을 위한 활동	- 프라이버시 보호 신호등 만들기 - 토론하기
3장	인공지능은 편견 없이 판단할까?	- 해결해야 할 문제는 무엇일까? - 인공지능 윤리 핵심 요건 - 얼굴 이미지 속 정보로 분류하기 - 성별을 구분할 수 있는 데이터 추가로 편향성 줄이기
	편향성 이해를 위한 옷을 추천하는 인공지능 프로그램 만들기	- 개발 계획 세우기 - 인공지능 모델 만들기 - 인공지능 프로그램 만들기
	개발자와 사용자는 어떤 생각을 할까?	- 개발자 관점에서 한 걸음 더 나아가기 - 사용자 관점에서 한 걸음 더 나아가기 - 개발자, 사용자 관점을 비교하여 의견 정리하기
	다양성 존중과 데이터 관리 실천을 위한 활동	- 체험을 통한 편향성 줄이기 실습하기 - 토론하기
4장	인공지능 결과 믿어도 될까?	- 해결해야 할 문제는 무엇일까? - 인공지능 윤리 핵심 요건 - 인공지능 투명성-의사 결정 트리 만들기 - 인공지능의 책임성, 안전성-자율 주행 자동차
	투명성 이해를 위한 범죄 형량 판단 인공지능 프로그램 만들기	- 개발 계획 세우기 - 인공지능 모델 만들기 - 인공지능 프로그램 만들기
	개발자와 사용자는 어떤 생각을 할까?	- 개발자 관점에서 한 걸음 더 나아가기 - 사용자 관점에서 한 걸음 더 나아가기 - 개발자, 사용자 관점을 비교하여 의견 정리하기
	책임성, 투명성, 안전성 실천을 위한 활동	- 인공지능 윤리 원칙 만들기 - 토론하기
5장	인공지능 활용과 사용자 유형 이해하기	- 해결해야 할 문제는 무엇일까? - 인공지능 윤리 핵심 요건 - 공공성: 사용자 유형 이해하기 - 연대성: 지속가능발전목표(SDGs)알아보기
	공공성과 연대성 이해를 위한 재활용품 분리수거 인공지능 프로그램 만들기	- 개발 계획 세우기 - 인공지능 모델 만들기 - 인공지능 프로그램 만들기
	개발자와 사용자는 어떤 생각을 할까?	- 개발자 관점에서 한 걸음 더 나아가기 - 사용자 관점에서 한 걸음 더 나아가기 - 개발자, 사용자 관점을 비교하여 의견 정리하기
	공공성과 연대성 실천을 위한 활동	- 내 삶을 바꾸는 인공지능 월드컵 만들기 - 토론하기

출처: 과학기술정보통신부·정보통신정책연구원, 중학교 인공지능 윤리 체험중심, 2022.

다) 성인용 인공지능 윤리 교재

해당 교재의 내용은 일반 성인의 인공지능 윤리 의식 및 리터러시를 함양할 수 있도록 구성되었다. 일반 시민이 인공지능 기기 및 서비스를 경험할 기회를 제공하고, 인공지능 윤리기준에 따라 비판적이고 주체적인 태도로 신기술을 사용할 수 있는 역량을 갖추게 하는 것을 목적으로 하고 있다.³⁸⁾

교재는 교수학습모델로 David Kolb(1984)의 ‘경험 학습 주기’ 모델을 교재 기본 프레임 구성으로 활용하였다. 이는 ‘구체적 경험’, ‘성찰적 관찰’, ‘추상적 개념화’, ‘적극적 실험’으로 하고 있는데 이에 따라 탐색하기, 경험하기, 개념화하기, 되돌아보기로 구조화되어 있다.

<표 III-6> KISDI 일반 성인 대상 인공지능 윤리교육 기본 프레임

과정	내용
1단계 탐색하기	[구체적 경험과 성찰적 관찰] - 일상의 인공지능 서비스 + 인공지능 윤리 이슈 경험
2단계 경험하기	[구체적 경험과 성찰적 관찰] - 일상의 인공지능 서비스 + 인공지능 윤리 이슈 경험
3단계 개념화하기	[추상적 개념화] - 경험 및 관찰한 것을 연결하여 ‘인공지능 윤리기준’ 학습
4단계 되돌아보기	[적극적 실험] - 탐색, 경험 및 관찰한 것을 다시 살펴보고 체험하면서 새로운 경험 획득
더 알아보기	- 본문에서 다루지 못한 이슈 및 정보 제공

출처: 조성은 외, 일반 인공지능 윤리교육 교재 개발 및 확산, 정보통신정책연구원, 2023, 86면.

<표 III-7> KISDI 일반 성인 대상 인공지능 윤리교육 단위별 교재 타이틀

구분	타이틀 내용
1장	생성형 인공지능으로 살펴보는 인권보장과 침해금지
2장	개인 맞춤형 서비스를 통해 본 프라이버시 보호
3장	챗봇으로 경험하는 다양성 존중
4장	윤리 매트릭스로 생각하는 공공성
5장	티처블 머신으로 살펴보는 데이터 관리
6장	자율 주행차로 알아보는 책임성과 안전성
7장	SNS 알고리즘으로 알아보는 투명성
8장	지속 가능한 발전을 위한 연대성

출처: 조성은 외, 일반 인공지능 윤리교육 교재 개발 및 확산, 정보통신정책연구원, 2023, 87면.

38) 조성은 외, 일반 인공지능 윤리교육 교재 개발 및 확산, 정보통신정책연구원, 2023, 13-14면.

<표 III-8> KISDI 일반 성인 대상 인공지능 윤리교육 단계별 주요 내용(1단계)

구분	1단계 탐색하기 주요 내용	
1	생성형 인공지능과 윤리 이슈	가짜뉴스와 인권 존중 문제, 콘텐츠 생성과정에서의 인권침해와 저작권 이슈 고찰
2	인공지능 맞춤형 광고로 보는 윤리 이슈	SNS 광고 및 위치기반 서비스 경험에서 생각해 볼 수 있는 이슈
3	인공지능 챗봇과 윤리 이슈	챗봇 활용에서 경험하는 편견과 다양성 존중 이슈
4	데이터 기반 인공지능과 윤리 이슈	데이터 품질을 조정하지 않았을 때 발생하는 사례 고찰
5	자율주행차와 윤리 이슈	자율주행차 관련 책임소재와 안전성 논의 사례 돌아보기
6	SNS 알고리즘과 윤리 이슈	SNS가 선별해 보여주는 콘텐츠 논란 사례 고찰
7	공공성을 고려하는 인공지능	도시의 여러 문제들과 해결을 위한 노력 살펴보기
8	국제사회의 협력과 필요한 이슈	환경 이슈와 연관되는 인공지능 기술과 국제사회 논의

출처: 조성은 외, 일반 인공지능 윤리교육 교재 개발 및 확산, 정보통신정책연구원, 2023, 89면.

<표 III-9> KISDI 일반 성인 대상 인공지능 윤리교육 단계별 주요내용(2~4단계)

구분	2단계 경험하기	3단계 개념화하기	4단계 되돌아보기
1	뤼튼(wrtn)으로 이미지를 생성해 보면서 인권과 기타 다른 침해 이슈 경험	① 인권보장 ② 침해금지 ③ 인권보장 및 침해금지를 위한 노력	① 새로운 문제 상황에 윤리 핵심 요건 도입하기 ② 다양한 상황으로 확장하여 살펴보기 ③ 체크리스트
2	자신의 SNS 광고 및 위치기반 서비스(구글 타임라인)를 돌아보면서 프라이버시 이슈 경험	① 프라이버시 보호, ② 프라이버시 보호를 위한 노력	① 마인드맵으로 정리하기 ② 프라이버시 보호를 위한 노력 ③ 체크리스트
3	AskUp(아숙업)과 뤼튼 스튜디오로 직접 챗봇 서비스를 구현하면서 다양성 존중 원칙 경험	① 다양성 존중 ② 다양성 존중이 왜 필요할까? ③ 다양성을 존중하지 않은 사례에는 어떤 것들이 있을까? ④ 다양성 존중을 위한 노력	① 사례를 찾아 정리하기 ② 주요 용어 정리하기 ③ 체크리스트
4	티처블 머신 (Teachable Machine)으로 이미지 구분 인공지능 모델을 만들면서 데이터의 중요성 경험	① 공공성 ② 공공성을 위한 노력	① 공공성 마인드맵으로 정리하기 ② 이해관계자 요구 분석하기 ③ 체크리스트

구분	2단계 경험하기	3단계 개념화하기	4단계 되돌아보기
5	모럴 머신 선택하기를 해보면서 자율주행차 알고리즘의 위험 요소 생각해보기	① 데이터 관리 ② 데이터 관리를 위한 노력	① 새로운 문제 상황에 데이터 관리 핵심 요건 도입하기 ② 다양한 상황으로 확장하여 살펴보기 ③ 체크리스트
6	YouTube와 메타의 작동 원리를 살펴보면서 인공지능 투명성의 필요를 경험	① 책임성 ② 안전성 ③ 문제 해결을 위한 자율적 노력	① 이해관계자 역할 정리하기 ② 인공지능에 의한 피해를 최소화하기 위한 노력 ③ 체크리스트
7	대중교통 노선 최적화 구현 과정에서 다양한 이해관계자 간 합의점 찾아가기 경험	① 투명성 ② 투명성 확보를 위한 노력	① 인공지능 투명성 팩트 시트 만들기 ② 투명성 실천 자율 점검표 확인하기 ③ 체크리스트
8	에너지 소비 데이터를 직접 분석하면서 에너지 소비와 온실가스 배출량의 상관성 경험	① 연대성 ② 연대성 실천을 위한 노력	① 국가 간 합의와 협력 사례 정리하기 ② 디지털 탄소 발자국을 줄이기 위한 나의 실천 ③ 체크리스트

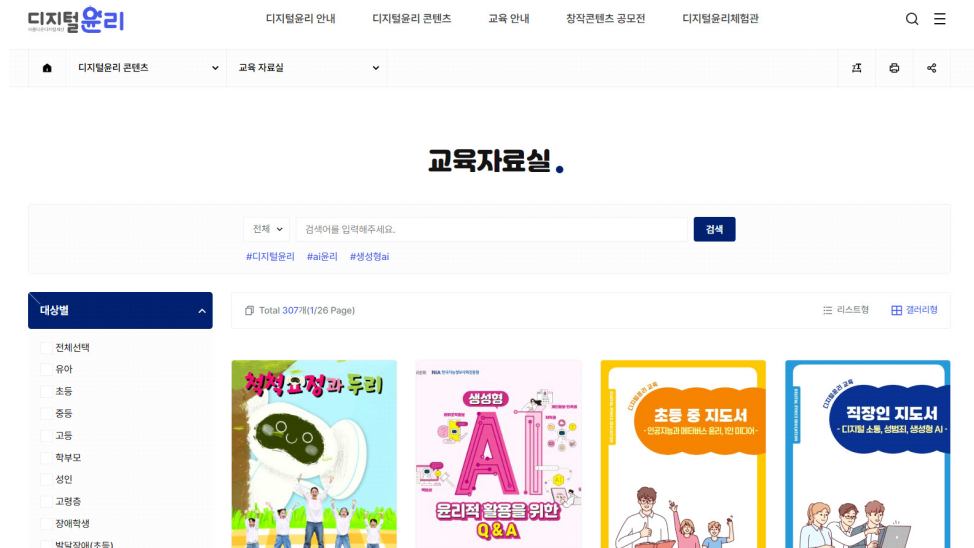
출처: 조성은 외, 일반 인공지능 윤리교육 교재 개발 및 확산, 정보통신정책연구원, 2023, 91~94면.

나. 방송통신위원회, 한국지능정보사회진흥원의 인공지능 윤리교육

방송통신위원회와 한국지능정보사회진흥원(NIA)은 청소년들이 인공지능 기술을 이해하고 이로 인한 긍정 및 부정적 영향을 인식하여 인공지능 기술 활용의 올바른 태도를 함양하기 위한 프로그램으로 '실습형 인공지능 윤리교육 프로그램'을 만들어서 온라인을 통해 공개하고 있다. 해당 교육 프로그램은 교수학습 자료로 제공되고 있으며, 강의안, 교사용지도서, 학생용 활동지와 데이터세트 등이 함께 공개되어 있다.

본 교육프로그램은 “① 인공지능으로 변화되는 세상에서 의사결정 도구로써 인공지능을 올바르게 활용하기 위한 태도를 함양한다. ② 편향된 인공지능이 가져올 역기능을 이해하고, 이를 최소화하기 위한 노력을 살펴본다. ③ 인공지능이 제공하는 정보를 주도적으로 선택하고, 인공지능을 윤리적으로 활용 및 개발할 수 있는 윤리 의식을 함양한다.”를 교육 목표로 하고 있다.³⁹⁾

39) 방송통신위원회·한국지능정보사회진흥원, 2023년 실습형 인공지능윤리교육, 2023, 6면.



[그림 III-2] 방송통신위원회, NIA 디지털윤리 홈페이지

1) 교육대상

해당 교육은 초·중등학생을 대상으로 한다.

2) 운영방안

1차시 기준 40분 수업 분량으로 총 3차시로 구성되어 있다. 현장 수업에서 활용될 수 있으며, 수업 상황에 따라 활용 가능한 모듈형으로 구성되어 있다.

3) 세부 프로그램 주요 내용

가) 초중등용 인공지능 윤리교육

본 프로그램은 학생들이 디지털 네이티브로서 정보의 무분별한 수용을 지양하고 비판적 사고력과 정보 분별력을 길러 균형 있는 시각과 능동적 소통을 가능하도록 하는 것을 목표로 삼아 구성되었다. 해당 프로그램은 사례 중심으로 개념과 특징을 이해할 수 있도록 구성하고, 기술이 미치는 영향을 이해할 수 있도록 하고 있다. 또한 기술의 긍정적·부정적 영향을 모두 파악할 수 있도록 하고, 스토리 기반으로 문제를 제기하여 학습자들에게 동기를 유발하고 관련된 실습을 진행하도록 하고 있는 것이 특징이다. 프로그램의 주요 내용은 다음과 같다.

<표 III-10> NIA 인공지능 윤리교육 프로그램의 주요 내용

구분	차시	세부 내용
초등	1차시 인공지능도 완벽하지는 않아!	- 생활 속 편리한 AI를 알아보고 개념 이해하기 - 독버섯 구분 실습을 통해 AI 예측 결과 살펴보기 - 의사결정 도구로써 AI 슬기롭게 활용하기
	2차시 AI도 편견이 있다는 것, 알고 있니?	- 인공지능 편향 사례를 살펴보고 개념 이해하기 - 산불감시 AI 실습을 통해 편향 살펴보기 - 편향 최소화 위한 노력 방안 알아보기
	3차시 인공지능은 우리의 좋은 친구!	- 생활 속 인공지능 살펴보기 - 인공지능을 지혜롭게 이용할 방안 생각하기 - 나를 돕는 인공지능 친구 만들기
중등	1차시 AI 다 믿어도 되는 걸까?	- 생활 속 편향 사례를 살펴보고 개념 이해하기 - AI가 생성한 얼굴 분류하는 실습하기 - 의사결정 도구인 AI를 올바른 목적으로 사용하기
	2차시 공정한 AI는 우리 손으로!	- 인공지능 편향 사례를 살펴보고 개념 이해하기 - 합격판단 AI 실습을 통해 편향된 AI 살펴보기 - 공정한 인공지능을 위해 필요한 노력 알아보기
	3차시 인공지능과 함께 하는 우리 미래	- 생활 속 인공지능과 생성형 AI 살펴보기 - 인공지능을 지혜롭게 이용할 방안 생각하기 - 우리를 돕는 인공지능 로봇 만들기

출처: NIA, 2023 인공지능 윤리교육 프로그램, 2023, 7면.

나) 초·중·등용 생성형 AI 윤리수업

해당 프로그램은 생성형 AI를 집중 주제로 하는 생성형 AI 윤리수업 프로그램을 제공하고 있다. 초·중·등학생을 대상으로 1차시 15분 분량으로 총 4차시 수업 교재가 제공되고 있고, 오프라인 및 온라인 수업에서 모두 활용 가능하다. 생성형 AI 윤리수업은 “① 인공지능 개념과 영향력을 알고 인공지능 윤리의 중요성을 안다. ② 인공지능 윤리의 중요성을 인식하고 정보 분별력과 비판적 사고력을 기른다. ③ 생성형 인공지능에 대해 이해하고 활용 윤리를 탐색할 수 있다”를 교육목표로 하고 있다.

<표 III-11> NIA 인공지능 윤리교육 프로그램의 주요 내용

구분	주제	주요 내용	학습 목표
초등	인공지능 개념과 영향력	- 인공지능 시대 이해 - 인공지능이 일상에 미치는 영향력	- 우리 주변에서 다양한 기능을 수행하고 있는 인공지능의 사례를 설명할 수 있다. - 인공지능의 적용 사례를 설명하고 우리 일상에 미치는 영향력에 대해 말할 수 있다.

구분	주제	주요 내용	학습 목표
	인공지능 윤리의 중요성	- 인공지능 윤리의 중요성 인식 - 세계의 다양한 인공지능 윤리 가이드	- 인공지능 윤리의 3대 원칙을 설명할 수 있다. - 인공지능 윤리의 필요성을 알고 사용자로서의 윤리적 태도를 갖출 수 있다.
	인공지능과 정보 편향	- 인공지능 정보편향 문제 탐색 - 정보 분별력과 비판적 사고력	- 데이터 편향성이 무엇인지 알고 데이터 입력의 중요성을 말할 수 있다. - 인공지능의 데이터 편향성을 줄이기 위한 바람직한 자세에 대해 말할 수 있다.
	생성형 인공지능 활용 윤리	- 생성형 인공지능 이해 - 생성형 인공지능 윤리의 중요성	- 생성형 인공지능이 무엇인지 말할 수 있다. - 생성형 인공지능이 제시한 정보가 정확한지 파악하는 일이 필요함을 알고 올바른 활용 자세에 대해 말할 수 있다.
중등	생성형 인공지능의 이해	- 생성형 인공지능 이해 및 분야별 활용 사례 - 생성형 인공지능 윤리의 중요성	- 생성형 인공지능의 개념을 이해할 수 있다. - 생성형 인공지능을 사용한 활용사례를 이해할 수 있다.
	생성형 인공지능과 저작권	- 디지털 콘텐츠의 저작권과 소유권 이해 - 생성형 인공지능의 저작권 문제 탐색	- 생성형 인공지능과 저작권 문제에 대해 이해할 수 있다. - 생성형 인공지능과 저작권 문제에 대한 자신의 생각을 표현할 수 있다.
	책임감 있게 활용하는 생성형 인공지능	- 생성형 인공지능의 오용 문제 탐색 - 생성형 인공지능의 올바른 활용 방안	- 생성형 인공지능의 오용 문제를 탐색할 수 있다. - 책임감 있게 생성형 인공지능을 활용하는 방안을 제시할 수 있다.
	생성형 인공지능 윤리 문제 예방과 대응	- 생성형 인공지능 시대의 윤리 문제 - 세계의 인공지능 관련 법과 제도	- 생성형 인공지능의 올바른 활용을 위한 법과 제도를 이해한다. - 생성형 인공지능의 올바른 활용을 위한 윤리 서약서를 작성할 수 있다.

출처: NIA, 짧게 끝내는 생성형 AI윤리 수업(초등), 2023, 6면.; NIA, 짧게 끝내는 생성형 AI윤리 수업(중등), 2023, 6면.;

2. 민간분야 개발 사례

연령, 성별 등과 무관하게 전 국민을 대상으로 하는 교육 프로그램이 운영되고 있다. 기술을 직접 다루지 않는 국민 역시 인공지능에 대한 기본적인 개념을 습득하고, 기술이 야기하는 문제를 파악하고 해결 방안을 도출하는 과정을 통해 인공지능 윤리 의식을 함양하는 것을 목표로 한다.

가. 국제인공지능 & 윤리협회

국내 민간단체인 사단법인 국제인공지능윤리협회(IAAE)는 비영리 사단법인으로 인공지능기술과 윤리교육, 인증, 거버넌스 및 정책 등의 사업을 실시하는 비영리 기관이다. 해당 기관에서는 학생부터 일반 시민까지 모두를 대상으로 인공지능 윤리교육 및 강의 프로그램을 마련하여 온라인을 통해 공개하고 있다.



협회홈 AI윤리최고경영자과정 AI리더스포럼 협회소개 사업안내 회원가입안내 ...

AI활용 교육/강의 안내
AI기술 교육/강의 안내
AI윤리 교육/강의 안내
AI교양 강의 안내
교육 시행 현황

AI윤리 교육/강의 안내

협회에서 진행하는 AI윤리 교육과 강의 안내입니다.
문의 및 신청 : 02-701-1237, laae@laae.ai

인공지능 윤리 교육 및 강의 프로그램

1. 교육 프로그램명 : 4차산업혁명의 핵심, 인공지능과 인공지능 윤리
2. 교육 수행 기관 : IAAE 국제인공지능&윤리협회
3. 교육 대상 : 초·중·고등학생, 대학생, 대학원생, 공무원, 교사, 교수, 시민 등

[그림 III-3] IAAE 인공지능 윤리교육 홈페이지

1) 교육대상

해당 교육은 초, 중·고등학생, 대학생, 대학원생, 공무원, 교사, 교수, 시민 등을 대상으로 한다.

2) 운영방안

협회에서 교육 및 강의를 제공하고 있으며, 온라인 및 오프라인으로 모두 참여 가능하다. 강의의 경우 30분에서 2시간, 교육은 2시간에서 12시간의 프로그램

램으로 구성되어 있고 선택적으로 교육 및 강의를 받을 수 있다. 교육이 종료되면 교육생들은 협회에서 발행하는 수료증을 발급받을 수 있다.

3) 세부 프로그램 주요 내용

해당 교육과정의 교육 목표는 “① 4차 산업혁명 시대의 인간 삶의 긍정적, 부정적 변화를 이해한다., ② 인공지능의 기본 개념과 주요 기술을 학습한다., ③ 인공지능 윤리의 개념과 주요 내용, 사례와 이슈들을 통해 AI 윤리를 학습한다., ④ 전 세계 인공지능 윤리 가이드라인의 주요 내용을 이해할 수 있다., ⑤ 인공지능 윤리 문제의 해결 방안을 제시할 수 있다.”이며, 주요 커리큘럼은 아래와 같다.

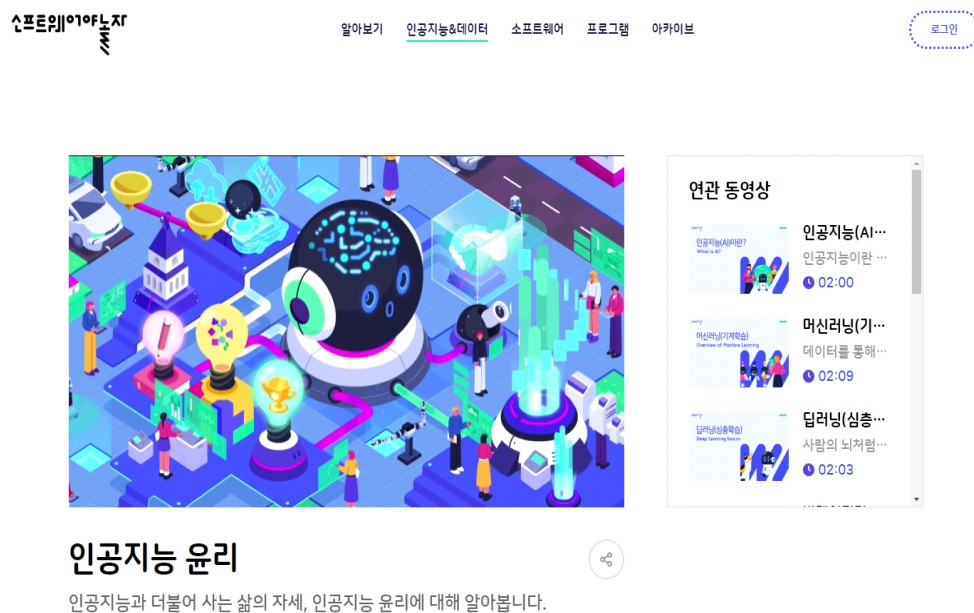
<표 III-12> IAAE 인공지능 윤리교육의 주요 커리큘럼(일반인용)

주제	타이틀 내용
인공지능 윤리의 철학적 배경 및 인문학으로 본 AI 윤리	• 인공지능 윤리의 철학적 배경 • 전통적 윤리관과 인공지능 윤리 • 포스트휴머니즘에서 트랜스휴머니즘까지 • 인문학으로 본 인공지능 윤리 • 인공지능, 윤리적으로 고민하기
인공지능 기술과 인간 삶의 변화	• 4차 산업혁명시대란? • 4차 산업혁명시대, 인간 삶의 긍정적, 부정적 변화 • 인공지능의 개념과 역사 • 인공지능의 핵심 기술과 원리 • 인공지능 최신 기술의 주요 사례
인공지능 윤리 개념과 사례	• 인공지능 윤리의 개념과 주요 내용 • 인공적 도덕행위자(AMA)의 개념과 사례 - AI 윤리 5대 주제 - AI의 편향 - AI의 오류 및 안전 - AI의 악 - 개인정보 및 저작권 침 - 킬러 로봇
인공지능 윤리 가이드라인	• 전세계 인공지능 윤리 가이드라인 • 국내 인공지능 윤리 가이드라인 • 인공지능 윤리 문제의 해결 방안 • AI와 인간의 바람직한 공존 방안
인공지능 윤리 해커톤	• AI 윤리 주제 토론과 워크샵

출처: IAAE 웹사이트, AI 윤리교육, 강의안내
원문: <https://iaae.ai/aiethicsedu>, 접속일: 2024.6.23.

나. 네이버의 인공지능 윤리교육

네이버 및 네이버커넥트 재단은 (사)한국정보교육학회, 프리월, 초등컴퓨팅교사협회, 한국정보과학교육연합회 등과 협력하여 ‘소프트웨어야 놀자’라는 웹페이지를 만들어 인공지능, 데이터, 소프트웨어 등과 관련한 지식을 배울 수 있도록 온라인을 통해 교재, 영상 등의 자료를 제공하고 있다. 인공지능에 대해서는 인공지능 윤리, 생성형 인공지능 윤리에 대해 수업 가이드와 강의안을 공개하고 있다.⁴⁰⁾



[그림 III-4] 소프트웨어야 놀자 홈페이지

1) 교육대상

해당 교육은 학생들에게 인공지능과 데이터 과학을 가르치는 교사를 위한 교재를 공개하고 있다.

2) 운영방안

온라인을 통해 수업교안, 워크시트 등을 공개하고 있다.

40) 소프트웨어야 놀자 웹페이지, 인공지능&데이터 교재,
원문: https://www.playsw.or.kr/artificial/textbook/ai_data_web_textbook_ai, 접속일: 2024.6.24.

3) 세부 프로그램 주요내용

해당 교육 과정은 ① 인공지능의 신뢰성, ② 인공지능의 편향성, ③ 인공지능의 악용가능성, ④ 인공지능의 책임성 총 4개의 챕터로 구성되어 있다.

<표 III-13> 네이버 인공지능 윤리교육의 주요 내용

주제	학습 목표	교수학습 활동
인공지능의 신뢰성	- 데이터 기반 인공지능의 신뢰성에 대해 알 수 있다. - 학습 데이터와 테스트 데이터를 이용하여 기계학습 모델을 훈련하고, 모델의 정확도를 확인 및 향상시킬 수 있다. - 인공지능의 신뢰성에 따른 사회적 영향에 대해 이야기 할 수 있다.	- 데이터 기반 인공지능 - 인공지능에 어떤 데이터가 사용되는지 이해하기 - 기계학습의 개념 알아보기 - 인공지능 모델의 정확도란 - 학습 데이터와 테스트 데이터 - 활동을 통해 데이터 기반 인공지능 모델의 정확도 확인하기 - 데이터 기반 인공지능 모델의 정확도가 100%가 될 수 없는 이유 생각해보기
		- 인공지능의 정확도 확인하기 - 학습 데이터를 이용하여 인공지능 모델 학습시키기 - 테스트 데이터를 이용하여 인공지능 모델의 정확도 확인하기
		- 인공지능의 정확도 높이기 - 데이터를 여러 방법으로 바꿔보며 정확도 올리기 - 인공지능 모델의 정확도를 올리기 위한 방법 공유하기
		- 인공지능의 신뢰성에 따른 사회적 영향 - 인공지능의 판단이 틀리는 경우 생길 수 있는 피해 예상하기 - 인공지능의 신뢰성 문제를 보완할 수 있는 방법 생각하기
인공지능의 편향성	- 데이터 편향성이 무엇인지 알고, 데이터 편향성이 생기는 이유에 대해 알 수 있다. - 다양성을 고려하지 못한 학습 데이터로 인공지능을 학습시켜 데이터 편향성을 확인해보고, 이를 해결하기 위한 방법을 생각할 수 있다. - 데이터 중 문제가 될 수 있는 속성을 이용하여 인공지능을 학습시켜 데이터 편향성을 확인해 보고, 이를 해결하기 위한 방법을 생각	- 데이터 편향성 - 데이터 편향성이란 - 데이터 편향성이 생기는 원인
		- 고인돌과 바위를 구분하는 인공지능 만들기 - 체험을 통해 다양성을 고려하지 않은 학습 데이터로 인공지능 모델을 학습시켰을 때의 문제점 파악하기 - 다양성을 고려한 학습 데이터로 재구성하여 편향성 문제 해소하기

주제	학습 목표	교수학습 활동
	할 수 있다. 인공지능의 편향성에 따른 사회적 영향에 대해 이야기할 수 있다.	- 학급 회장을 선출하는 인공지능 만들기 - 민감한 속성을 사용했을 때의 문제점 파악하기 - 적절한 속성이 무엇인지 탐색하기 - 인공지능의 편향성과 사회적 영향 - 편향된 인공지능으로 인한 차별 사례 - 인공지능의 편향성 문제를 해소할 수 있는 방안 - 인공지능 활용에 대한 마음 다지기
인공지능의 악용가능성	- 인공지능이 할 수 있는 일에 대해 이해하고, 생성 모델 인공지능이 무엇인지 이야기할 수 있다. - 생성 모델 인공지능을 체험해보고, 생성 모델이 가져올 수 있는 좋은 점과 악용될 가능성에 대해 이야기할 수 있다. - 인공지능이 악용된 사례를 살펴보고, 인공지능의 악용 가능성을 예방하는 방안을 생각해 볼 수 있다.	- 생성하는 인공지능 - 인공지능이 할 수 있는 일 - 생성 모델이란 - 생성 모델의 활용 사례 - 생성하는 인공지능 체험하기 - 클로바 보이스 - 바흐 두들 (Bach Doodle) - 어떤 게 진짜 얼굴일까요? (Which Face Is Real?) - 인공지능의 악용 가능성과 사회적 영향 - 인공지능의 악용 가능성 예상하기 - 기타 악용 가능성이 있는 인공지능 사례 - 인공지능 악용 가능성을 예방하는 방법
인공지능의 책임성	- 자율주행 자동차와 관련된 윤리적 딜레마 상황에 대한 자신의 생각을 이야기할 수 있다. - 자율주행 셔틀 버스의 윤리적 설계 가이드라인을 만들어 보고, 인공지능의 책임성에 대해 이야기할 수 있다.	- 인공지능과 윤리적 딜레마 - 여러 가지 윤리적 딜레마 상황(모럴 머신)체험하기 - 생각 나누기 - 사회적 합의와 인공지능 활용의 책임성 - 자율주행 셔틀 버스의 윤리적 설계 가이드라인 만들기 - 인공지능 개발 및 사용의 책임성

출처: 소프트웨어야 놀자 웹페이지, 인공지능 윤리,

원문: https://www.playsw.or.kr/artificial/textbook/ai_data_web_textbook_ai, 접속일: 2024.6.23.

<표 III-14> 네이버 생성형 인공지능 윤리교육의 주요 내용

구분	학습 목표	주요 내용
챕터 1	생성형 인공지능을 이해하고 인공지능 윤리의 필요성을 설명할 수 있다.	• 생성형 인공지능과 인공지능 윤리 - 생성형 인공지능 알아보기 - 튜튼 AI스토어로 대화형 인공지능 만들기 - 인공지능 윤리의 필요성
챕터 2&3	생성형 인공지능의 프라이버시 문제를 탐색하고, 프라이버시 보호 실천 방법을 안다.	• 나와 나를 지키는 인공지능 윤리 - 일상 속 나의 프라이버시 인식하기 - 엔트리로 고민상담 챗봇 개발하기 - 개발한 챗봇에 윤리 더하기 - 생성형 인공지능과 프라이버시 보호
챕터 4	생성형 인공지능의 인권 침해 문제를 탐색하고, 인권 존중 방안을 안다.	• 모두의 인권을 위한 인공지능 - 생성형 인공지능의 원리 알아보기 - 엔트리로 얼굴 합성 인공지능 체험하기 - 생성형 인공지능과 인권존중
챕터 5	다양성을 존중하는 인공지능을 위해 필요한 노력을 설명할 수 있다.	• 다양성을 존중하는 인공지능 - 이미지 생성 인공지능 탐색하기 - 이미지 생성 인공지능의 다양성 존중 탐색하기 - 인공지능의 다양성 존중을 위해 나아가기
챕터 6	공공성을 높이는 인공지능을 위해 노력할 점을 설명할 수 있다.	• 보다 더 많은 사람을 위한 인공지능 - 수화를 해설하는 인공지능 만들기 - 학교의 공공성을 높이는 인공지능 떠올리기
챕터 7	생성형 인공지능으로 다양한 목소리를 생성하고, 투명성과 안전성의 필요성을 설명할 수 있다.	• 위험을 예방하여 더 안전한 인공지능 - 더 나은 목소리를 내는 방법 - 생성형 인공지능으로 교내 방송 만들기 - 안전하고 투명한 인공지능
챕터 8	인공지능 윤리 10가지와 관련된 질문을 정하고 생활 속 실천 태도를 기른다.	• 우리가 생각해야 할 10가지 질문 - 함께 토의하며 게임하기 - 10가지 질문을 담은 윤리 네 컷 완성하기

출처: 소프트웨어야 놀자 웹페이지, 생성형 인공지능 윤리,
원문: https://www.playsw.or.kr/artificial/textbook/ai_data_web_textbook_ai, 2024.6.23.

IV. 결론 및 시사점

이상으로 국내·외 인공지능 윤리교육 프로그램 및 실제 사례를 살펴보았다. 각국의 다양한 사례에서 발견할 수 있는 특징을 바탕으로 시사점을 정리해 보면 다음 <표 IV-1>과 같다.

<표 IV-1> 국가별 세부 인공지능 윤리교육 프로그램 및 사례 비교

국가	주요 사례		특징 및 시사점
	사례명	내용	
국 외	미 국	MIT Media Lab의 AI+Ethics	학계나 기업 등 민간과의 협력을 통해 교육이 운영되며, 인공지능 기술이 발전하면서 발생한 위험을 윤리적인 측면에서 파악하여 기술을 사용하는 사람이 지녀야 할 책임 등 윤리의식 함양을 목표로 함
		MIT Media Lab의 How to Train Your Robot	
		공인 AI 윤리 및 거버넌스 전문가 과정 (CAEGP)	
	스 웨 덴	AI Lund의 인공지능: 윤리적 및 사회적 과제	유럽에서 최초의 포괄적인 인공지능 규제법이 통과되면서 성별, 연령, 직업 등에 상관없이 전 유럽 국민이 인공지능 기술에 대한 접근성을 향상시키며 이를 책임감 있게 운용할 수 있는 방안을 교육하고자 함
	독 일	뮌헨 공과대학교의 AI 윤리	
	독 일	KI Campus	

핀 란 드	Elements of AI	EU 전체 인구의 1%가 인공지능의 개념 및 구축 등에 대하여 기본적인 지식을 교육받을 수 있도록 과정 구상	
핀 란 드	Ethics of AI	AI가 윤리적으로 수용 가능한 사용 및 개발에 대해 우려가 제기되는 이유 파악 및 현대 AI와 연관된 윤리적 문제 탐색	
영 국	LSE의 Ethics of AI	AI가 개인과 사회에 미치는 영향 관리 및 윤리적 사고를 실천하기 위한 도구와 기술 개발을 목표	
유 럽 연 합	AI4T	고등학교 수학, 과학 및 현대 영어 과목에서 인공지능을 활용한 교육을 할 수 있도록 교사를 대상으로 한 AI 도구 사용법 및 윤리 문제에 대한 교육과정 구축	
태 국	AI Executive Program in Digital Healthcare	디지털 헬스케어의 AI 임원 프로그램(AiX)은 의료 등의 공공 서비스에서 인공지능 기술이 효율적이고 고품질의 서비스를 제공할 수 있도록 공공 및 민간 부문에서의 이해 증진 및 인식 제고를 목표로 과정이 운영됨	AI 기술이 실무에 도입되었음에도 이해관계자들의 기술에 대한 인식 및 윤리 역량이 미흡하여 이를 증진하기 위한 노력의 일환으로 교육 프로그램 운영
싱 가 포 르	LearnAI	국가 인공지능 이니셔티브인 'AI Singapore'의 일환으로 세대별로 커리큘럼을 구분하여 AI 역량을 개발함으로써 안전하고 책임감 있는 AI 활용 장려	
캐 나 다	아타바스카 대학의 AI Ethics Micro-Credential	기업이 시대의 흐름에 따라 AI를 도입하면서 업계의 전문가들이 인공지능 윤리교육에 대한 필요성을 절감하여 마련된 교육과정	
유 네 스 코	AI 윤리 교사 연수 과정 (ETTC)	생성형 AI 및 범용 AI의 출현과 관련된 우려와 과제를 논의하며 다양한 학문 분야에 속한 교사의 인공지능 윤리 역량 강화 목표	
국 내	공 공	과학기술정 보통신부, 정보통신정 책연구원의 인공지능 윤리교육 교재 개발	사용자, 개발자 등에게 요구되는 책임 의식 함양을 위하여 인공지능 전 과정과 밀접하게 연관된 윤리교육 제공
	방 송 통 신 위 원 회	청소년들이 인공지능 기술을 이해하고 이로 인한 영향을 인식하여 기술 활용의	

	한국지능정보사회진흥원의 인공지능 윤리교육	올바른 태도를 함양할 수 있는 프로그램 마련	
민 간	국제인공지능 & 윤리협회의 AI 윤리교육	인공지능 기술의 발달로 야기된 사례와 문제를 중점으로 인공지능 윤리를 학습하며 윤리 문제의 해결 방안 제시	전 국민을 대상으로 운영되며 인공지능에 대한 기본적인 개념 습득과 더불어 인공지능의 윤리적 딜레마에 대한 논의 진행
	네이버 소프트웨어 야 놀자	인공지능, 데이터, 소프트웨어 등과 관련한 지식을 축적하는 동시에 윤리적 딜레마를 해결할 수 있는 방안을 탐색할 수 있는 커리큘럼 제공	

해외 인공지능 윤리교육의 경우 다양한 기관이나 주체에서 프로그램을 개발하고 보급하고 있는 것으로 보인다. 인공지능 기술이 광범위하게 여러 주체들에 의해서 활용된다는 측면에서 보면 여러 관점으로 프로그램이 개발된다는 점은 매우 고무적이라 할 수 있다. 인공지능 윤리교육은 특정 집단이나 초·중·고등학교 학생뿐만 아니라 대학생, 일반인 등은 물론 인공지능 시스템과 서비스를 개발하고 공급하는 주체들에게도 인공지능 윤리교육이 실시되어야 한다는 점에서 다양한 내용과 교육방법이 개발되는 것이 필요하기 때문이다. 이에 해외 주요 국가의 정부, 대학, 민간 교육기관, 협회 등에서 개발하는 교육 프레임워크나 교육 프로그램의 구성 내용, 교육 방법 및 도구 등에 대해 지속적으로 확인할 필요가 있다. 또한 해당 인공지능 윤리교육의 효과가 높다고 평가되거나 교육 대상에 따라 적합한 교육 내용이라 판단되는 경우 적극적으로 반영하여야 할 것이다.

한편, 해외 주요 국가들의 인공지능 윤리교육의 또 다른 특징은 교사를 대상으로 하는 교육 프로그램도 함께 개발·제안되고 있다는 점이다. 학교 교육에서 인공지능 교과서를 활용하기로 한 상황에서 교사들에게도 인공지능 기술에 대한 기본적 이해는 물론 적정하고 합리적으로 기술을 활용하기 위한 기초적 지식을 쌓을 수 있도록 고려하는 것이 필요하다. 나아가서는 교사를 대상으로 하는 교육을 포함하여 다양한 집단과 주체를 고려한 인공지능 윤리교육 프로그램이 지속적으로 개발되어야 할 것이다.

현재 한국의 경우 초·중·고등학생과 일반시민을 대상으로 교육 커리큘럼이 개발되었고, 인공지능 기술 개발자를 대상으로 하는 교육도 개발 예정에 있다. 일

반 시민의 경우에도 디지털 소외계층 등 인공지능 기술을 활용하는 주체의 역량에 따라서도 교육내용의 차이를 두는 배려가 필요하며, 인공지능 기술 활용이 지속적으로 확산되고 있다는 점에서 분야별 인공지능 윤리교육 프로그램이 개발되는 것이 바람직할 것이다. 예를 들어 공공분야의 경우 인공지능 기술의 활용을 적극 장려하도록 정책이 추진되고 있는 상황으로 공무원 및 공공기관 종사자 등을 대상으로 하는 교육이 개발되어야 할 것이다. 공공분야에서의 인공지능 기술의 활용은 국민을 대상으로 한다는 점에서 프라이버시 침해 등의 문제가 발생할 경우 침해의 정도나 내용이 민간영역에서 기술 활용으로 인해 발생하는 문제와는 차이가 존재하기 때문이다.

분야별 특성을 고려하고 반영한 인공지능 윤리교육 프로그램이 지속적으로 개발될 수 있도록 법제도적으로 기반을 마련해두는 것도 좋은 방법이 될 수 있을 것이다. 이외에도 민간기관에 의한 자격증이나 인증시험 등을 활용하는 것도 인공지능 윤리교육을 확산시키는 데 기여할 수 있을 것이다. 이 또한 법제도적으로 실시 근거를 마련하는 것이 필요하다.

인공지능 윤리교육의 특징은 기술에 대한 이해를 포함하여 기술의 위험성과 유익함을 모두 파악할 수 있고 이에 더하여 기술을 사회가치에 맞게 활용할 수 있도록 종합적으로 커리큘럼이나 프로그램이 구성된다는 데 있다. 이는 인공지능 윤리교육 프로그램이나 커리큘럼을 개발하는 과정에서는 다양한 분야의 전문가나 이해관계자 등이 함께 참여하여 내용과 방법들을 개발하는 것이 필요하다는 것을 의미한다.

향후 국내 인공지능 윤리교육 프로그램 개발 및 적용 시, 교육 내용과 목표, 방향성을 설정하기 위해 우리 사회가 추구하는 지향점이 무엇인가에 대한 고민은 필수적이다. 또한 인공지능 기술이 만연해진 시대에 인간의 존엄성과 자율성 등의 가치는 어떻게 지키고 구현할 것인지, 기술로 인한 변화를 어떻게 긍정적으로 사회에서 수용할 수 있도록 할 것인지 등에 대한 고민이 필요하다. 이는 다양한 주체에 의한 숙고와 합의가 필요한 사항이라고 할 수 있다. 이러한 고민을 통해 제공되는 인공지능 윤리교육을 통해 우리 사회에 올바른 인공지능 기술 활용 기반이 마련될 수 있을 것이다.

참고 문헌

- 과학기술정보통신부·정보통신정책연구원(2022). 초등학교 교사용 인공지능 윤리
놀이중심
- 과학기술정보통신부·정보통신정책연구원(2022). 중학교 인공지능 윤리 체험중심
방송통신위원회·한국지능정보사회진흥원(2023). 2023년 실습형 인공지능윤리교육
배지혜(2021). 독일의 인공지능(AI) 융합교육 현황과 시사점. 교육정책네트워크
정보센터 해외교육동향 제396호
- 백수은·한나(2022). 싱가포르 AI 평생학습체계 분석 및 시사점. 차세대컨버전스
정보서비스기술논문지 제11권 제6호
- 조성은 외(2023). 일반 인공지능 윤리교육 교재 개발 및 확산. 정보통신정책연구
원 정책 연구
- 한국교육학술정보원(2019). 모두를 위한 인공지능 윤리
- 한국지능정보사회진흥원(2022). 핀란드, AI를 활용한 사람 중심 사회 실현
- AI Singapore(2023). LearnAI Handbook
- Blakeley H. Payne(2019). An Ethics of Artificial Intelligence Curriculum for
Middle School Students. MIT Media Lab
- Randi Williams et.al.(2021). Teacher Perspectives on How To Train Your
Robot A Middle School AI and Ethics Curriculum. The Thirty-Fifth
AAAI Conference on Artificial Intelligence (AAAI-21).
- 소프트웨어야놀자 <https://www.playsw.or.kr>
- 정보통신정책연구원 <https://www.kisdi.re.kr>
- 스웨덴 Lund대학 <https://www.lunduniversity.lu.se/>
- 영국 정치경제대학교 <https://www.lse.ac.uk/>
- 캐나다 아타바스카대학 <http://www.athabascau.ca>
- AI4T <https://www.ai4t.eu/>
- AI Campus <https://ki-campus.org/themen/schule>
- AI EXECUTIVE PROGRAM DIGITAL HEALTHCARE
<https://dgt.etcha.or.th/aix/>
- AI Singapore <https://learn.aisingapore.org/>
- Elements of AI <https://www.elementsofai.com/>,
- Ethics of AI <https://ethics-of-ai.mooc.fi/>
- IAAE <https://iaae.ai/aiethicsedu>

LearnAI <https://learn.aisingapore.org/>

MIT Media Lab <https://www.media.mit.edu/>

MIT News (2019). National Initiative for Cyber security Careers and Studies
(NICCS CAEGP) <https://niccs.cisa.gov/>

Unite.AI <https://www.unite.ai/>

UNESCO Ethics Education Programme

<https://www.unesco.org/en/ethics-science-technology/education>

연구자료 RM 2024-10

국내·외 인공지능 윤리교육
사례 분석과 시사점

발 행 일 2024년 8월 23일
발 행 인 정 제 영
발 행 처 **한국교육학술정보원**
(www.keris.or.kr)
주 소 ☎ 41061 대구광역시 동구 동내로 64
전화 : (053)714-0114
등 록 제22-1584호(1999년 7월 3일)
인 쇄 처 참 디자인 (053)716-6005
I S B N 979-11-6555-442-2 (93370)

본 내용의 무단 복제를 금함.

<비매품>



이 저작물은 "공공누리" 출처표시-상업적이용금지-변경금지 조건에 따라 이용할 수 있습니다.

* 한국교육학술정보원 : www.keris.or.kr

* 본 연구자료에는 집필진 및 제작사의 저작권이 포함되어 있습니다.

