

AI 거버넌스와 공공윤리

: 법과 윤리라는 **공공 가치**에 대한 공감을 이끌어내고, 이를 바탕으로
학생들이 주체적으로 **창의적·사회혁신적** 아이디어를 이끌어냄

시 의 성

1. 생성형 AI가 등장한 이후, AI의 이점을 누리면서도 정보 유출은 피하고 싶은 수요가 대거 등장하고 있다.

✓ 주식회사 파수, 중요 정보 유출 차단하는 AI-R DLP 출시

AI에 질문을 입력하는 과정에서의 민감정보 유출 문제가 지적됨에 따라 많은 기업들이 AI 서비스 사용을 금지하거나 제한하고 있음을 반영하여 AI의 이점을 살릴 수 있는 솔루션 개발.

머신러닝+AI=온디바이스 AI... 강력한 정보보호 '감장' (g-enews.com)

✓ AI 업계의 새로운 화두로 온디바이스 AI 부상

인터넷 없이 로컬에서 데이터를 처리해 민감한 정보가 제3자에게 노출될 위험을 줄여줌.

파수, 생성형 AI 통한 민감정보 유출 방지하는 'AI-R DLP' 출시 < 보안 > 뉴스 < 기사본문 - 아이티데일리 (itdaily.kr)

→ 이러한 수요에 따른 다양한 대안이 등장하는 2023-2 시점에서 AI와 공공 윤리의 연관성을 이해하고 실제로 연계 프로젝트를 기획해보는 것은 **가장 시의적절한 교육**으로 평가받을 수 있음.

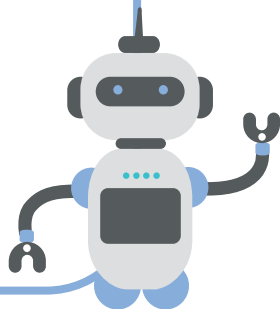
2. AI 규제에 대한 다양한 사항들이 활발히 업데이트되고 있다.

"AI 안전해야 지속 가능" ... 스스로 속도조절 나선 AI 기업

생성 AI 기술에 대한 규제론이 힘을 받기 시작하자, 기업들 스스로 속도조절을 시작했다. 지난해 하반기 미국, 유럽에서 정부 주도로 AI 규제에 대한 논의가 본격화됐다. AI 기술이 안전한지가 AI 사업 지속성에 연동되기 시작한 것. 유럽연합(EU) 27개 회원국 대표는 지난해 12월 벨기에 브뤼셀에서 'AI 법'(AI Act)으로 알려진 규제 법안에 합의했다. "범용 AI(GPAI, 다양한 용도로 사용할 수 있는 강력한 AI 모델)를 개발하는 기업은 AI 모델의 학습 과정을 보고해야 한다"는 내용까지 포함한 강력한 법이다.

국내도 '미국식 AI 규제안' 채택...다음달 자율규제 가이드라인 마련

'차세대 생성 AI 기술개발'에는 2024년부터 2027년까지 220억원을 투입할 계획이다. 이는 편향성, 불투명성 등 기술적 한계와 환각 등 위험요인에 대응하기 위해 기존 AI의 한계를 극복하고 신뢰성을 확보하기 위한 새로운 기술을 개발한다는 내용이다. 더불어 AI가 생성한 결과물에 대한 워터마크 도입의 제도화를 검토하고, 고위험 AI에 대한 해설서를 마련하는 등 신뢰성 확보를 위한 제도 정립 과제를 추진할 방침이다.



연세대학교 창립 정신과 고등교육혁신원 인재상과의 연관성



① 연세대학교 건학 정신

연세대학교는 '너희가 내 말에 거하면 참 내 제자가 되고 진리를 알지니 진리가 너희를 자유케 하리라'는 성경말씀(요한복음 8:31~32)을 바탕으로 진리와 자유의 정신을 체득한 지도자를 양성하고 있다. 즉, 이를 통해 연세대학교가 **단순 지식을 향유하는 것**에서 그치지 않고 해당 지식을 바탕으로 **새로운 가치를 창조함**으로써 **자유**를 만들어 나갈 수 있는 인재를 추구함을 파악할 수 있다. 따라서 **실천적 지식**을 기반으로 한 **창조적 자유**를 누릴 수 있도록 **학생들을 독려하는 교과목**이 필요하다.



② 고등교육혁신원 인재상

고등교육혁신원은 공동체 정신을 지닌 **혁신적 리더**를 양성하는 것을 목표로 삼고 있다. 더불어 해당 인재가 가져야 할 핵심 가치로는 **도전과 선도, 창의와 혁신, 공존과 혁신**이 강조된다. 이러한 배경을 고려하여 볼 때, **인공지능**이라는 새로운 기술이 도래한 상황에서 **기존 사회 구조와 어떻게 맞물릴 수 있는지**를 고민한다는 점에서 **공존**의 가치를, 학생들이 **주체적으로 문제에 대한 솔루션을 기획**한다는 점에서 **창의와 혁신**의 가치를 엿볼 수 있다. 마지막으로 해당 교과목을 수강함으로써 학부생들은 인공지능이 주요 산업이 될 현시대의 **선도적인 인재**로 기능할 수 있을 것이다.

1. AI 인력 양성의 중요성 부각

과학기술정보통신부가 중심이 되어 글로벌 시장 경쟁력 확보를 위한 기술, 인재, 인프라 분야 주요 추진 과제와 로드맵을 제시하며 기술력 확보를 위한 융복합 인재를 양성하기 위해 주력하고 있으나 국외 AI 기술력과 전문 고급 인재의 양적, 질적 수준부족과 보완이 시급한 상황이다.

한국 최고AI 인력 수는 **405명**으로 인도보다 낮은 수준, 세계 1위 **미국의 1/3** 수준이다. AI 기술 혁신 인프라 보완이 시급하다. **기존 AI대학원은 응용 프로그램 활용 교육과 융합적 사고를 하기 어렵다는 한계점을 지닌다.** 커리큘럼은 컴퓨터 공학적 기초 요소뿐 아니라 타과의 융합적 교육 프로그램으로 재편성되어야 한다.

<https://www.junggi.co.kr/article/articleView.html?no=31268>
<https://www.donga.com/news/Inter/article/all/20230707/120119104/1>
<https://news.mt.co.kr/mtview.php?no=2023053110092090625>
<https://nsp.nanet.go.kr/plan/subject/detail.do?nationalPlanControlNo=PLAN0000040915>
<https://www.edaily.co.kr/news/read?newsId=02820806628918048&mediaCodeNo=257&OutLnkChk=Y>



2. AI 윤리 문제 발생

미국 기업은 99%가 채용에 ai 기술을 활용하고 있다. AI는 입사 지원서를 검토해 지원 자격에 미달하는 서류를 거르거나 채용 담당자가 보다 꼼꼼히 살펴볼 후보군을 추천한다. 지난해 **골드만삭스는 인턴 채용에 AI를 활용해 지원자 23만6000명 중 1.5%에 해당하는 3700명을 선발했다.**

그러나 채용 등의 영역에서 **AI의 공정성**을 두고 논란이 여러 차례 불거진 바 있다. 구글 감원 논란, 아마존 채용 성차별 논란 등이 그 대표적 사례이다. 이로 인해 **뉴욕시가** 세계 처음으로 채용 과정에서 AI 프로그램을 활용하는 기업에 의무적으로 **편향성을 따져보게 하는 규제**를 도입했다. 최근 국내 공공기관과 기업들이 채용 과정에 AI 기술을 도입하는 사례 또한 늘고 있다.

한국자산관리공사는 지난 4년간 신입사원과 인턴 등을 뽑는 **24번의 전형에 AI 채용**을 도입했다. 이와 같이 AI를 단순히 유용히 활용하는 것뿐 아니라 **윤리적 측면에서의 신뢰성과 공정성을 갖춘 AI의 활용이 점점 더 중요해지고 있으며 강조되고 있다.**

3. AI 개인정보 이슈 부상

사회적 약자의 사진, 인상착의 등을 바탕으로 cctv를 검색해 이동경로를 찾는 솔루션과 같이 공익적 가치를 지닌 인공지능이 제안되고 있으나, 개인의 얼굴, 신체적 특징 등을 활용해야하기 때문에

현행 개인정보에서의 많은 법적 이슈들을 내포하고 있다.

개인정보보호위원회(이하 '개인정보위')는 8월 3일 **「인공지능 시대 안전한 개인정보 활용 정책방향」**을 발표하였다. AI에 대한 기대와 우려 사이에서 프라이버시 침해 위험은 최소화하면서 AI 혁신 생태계 발전에 꼭 필요한 데이터는 안전하게 활용할 수 있도록 하기 위해 정책방향을 수립했다고 밝혔다.

→ AI 활용함에 있어서 이러한 정책적, 법적 지식이 선행되어야 하며 AI 인재 육성을 위한 커리큘럼에 포함될 필요가 있다.

SDGs와의 연관성



SDG 3: 건강과 웰빙을 달성하는 데에 기여한다. AI가 유익하고 안전하며, 신뢰성과 공정성을 바탕으로 차별 없이 모두의 건강한 생활을 돕는 방향으로 쓰일 수 있도록 장려한다.

SDG 4: 품질 좋은 교육을 실현함으로써 미래 지식과 기술에 대한 학생들의 역량을 강화할 수 있다.

SDG 9: 산업 혁신 기반 구축의 목표를 달성할 수 있다. AI와 지식재산권을 다룸으로써 AI 기술 및 모델 발명을 보호하고 장려하여 동기를 부여할 수 있다.

국내외 유사 사례 분석 및 차별점

해외 사례

요약

해외 대학들의 경우 본격적으로 AI윤리 과목을 가르치기 시작했다. 특히 보스턴대는 '책임있는 AI, 법, 윤리학&사회(Responsible AI, Law, Ethics & Society)' 과정을 개설했다. 영국 케임브리지 대학은 '책임있는 AI' 관련 석사과정을 개설하기도 했다.

① Harvard University

"Data Science and Artificial Intelligence: Ethics, Governance, and Laws"

Description - 데이터 과학과 인공 지능은 기업의 의사 결정, 생산성 향상에 새로운 기회를 제공하면서 동시에 윤리적 정보보호에 대한 우려도 일으키고 있다. 해당 강좌는 **AI 엔지니어가 데이터 수집 및 분석의 수행 중에 발생할 수 있는 부작용을 인지하지 못하고 알고리즘을 개발했을 때의 책임에 중점을 두고 있다.** 수강생은 데이터 수집, 알고리즘 개발의 과정을 이해하기 위한 프레임워크를 제공받으며, 케이스 스터디와 시뮬레이션을 통해 데이터 과학, ai 윤리, 거버넌스 맥락에서 학습, 비판적 사고, 토론 등 문제 해결력을 향상시키는 것을 목표로 한다.

② Georgia institute of technology

"AI, Ethics, and Society"

Description - 모델 학습에 사용되는 대규모 데이터를 효과적으로 처리하면서 그 결과가 남용되지 않도록 하는 **알고리즘을 어떻게 디자인할지**에 대해 다룬다. 대량의 데이터로부터 학습의 **잠재적 남용과 오용을 막기 위한** 다양한 Ai, ml 기술을 살펴며, 이러한 기술이 개인, 조직 및 사회에 미치는 **영향에 중점을 두고, 컴퓨팅 전문가로서의 책임에 주목한다.**

차별점 - 위 두 강의는 알고리즘에 집중하며 AI 엔지니어로서의 소양과 책임을 강화하는 데 초점이 맞추어져 있다. 반면 본 강의는 **전반적인 AI의 활용에서 나타날 수 있는 문제점들을** 엔지니어의 입장이 아닌, AI를 활용하는 모두의 입장에서 다루고 생각해보는 **융합적이고 다학제적인 교양 수준의 강의**라는 점에서 차별화된다.

국내 사례

① [교양] 인공지능의 이해와 활용 – 사회 혁신 역량 교과목

머신러닝, 딥러닝, 강화학습에 대한 이론을 간단한 예제들을 통해 실습하는 방식으로 인공지능의 기초적인 개념들을 이해하도록 한다. **각자의 분야에서의 응용**과 추후의 심도 있는 학습을 위한 토대를 마련하도록 한다. 더불어 인공지능을 도구로 활용하여 다양한 사회문제 해결에 도전한다.

② [전공] 인공지능과 지식재산 – 공과대학 공통과목

인공지능 분야에 적용되는 특허권, 저작권, 상표권, 영업비밀 등 **지식재산권 시스템 전반**을 이해하고, 이를 어떻게 보호하고 이익을 창출할 수 있는지 탐구하는 것을 목표로 한다.

③ [대학원] 인공지능과 법

인공지능대학원 재학생 대상 수업이나 교수자는 모두 **법학전문대학원** 소속이다. 인공지능과 관련된 법적 쟁점을 다양한 법학 분야의 관점에서 팀티칭하여, **인공지능의 개발 및 실용화에 필요한 기초적인 법적 지식**을 제공한다.

④ [전공] 인공지능 시대의 윤리

신학과 과목으로, 인공지능이 어디까지 인간의 삶을 바꾸어 놓을 것이며 인간을 대신할 수 있을 것인가? 와 같은 **철학적, 신학적 관점**에서 윤리 문제들을 다룬다.

차별점- 강의를 크게 두 가지 결로 나뉘볼 수 있다.

첫째로는 AI의 활용과 관련하여 "법"에 초점을 맞추는 강의들이다.(3,4번) 해당 과목들은 본 제안 과목에 비해 윤리적 측면에 집중하여 추상적인 경향이 짙으며, **실전적인 프로젝트의 경험과 윤리 문제 발생의 과정을 실습을 통해 체험해볼 수 없다**는 점에서 본 강의와 결을 달리한다. 본 강의는 **윤리적 측면의 심도 있는 분석보다는 실전적 과정을 경험해보고, 경험해본 바를 통해 느낀 윤리적 문제점들에 대한 답을 스스로 찾아가는 과정**을 제공하는 데에 중점을 둔다.

두 번째로는 AI자체에 대한 이해에 좀 더 초점을 맞추는 강의이다. 1번 강의의 경우에는 인공지능에 대한 이해를 높이는 과목이지, 그 목적에 **거버넌스와 윤리적 활용**을 두고 있지 않다.

또한 2번 강의의 경우에는 **인공지능 모델 자체를 보호하고, 인공지능 개발자의 권익과 관련된 지식재산권을 다룬다**는 점에서, 지식재산권뿐 아니라 **공공 윤리 전반**에 관심을 두는 본 강의와 범위가 다르다고 볼 수 있다.

강의안 구성

교과목명: AI 거버넌스와 공공윤리

비고: 사회 혁신가 인증 교과목, 수료증 부여

핵심역량: 융합적 사고력 50% 창조적 사고력 30% 비판적 사고력 20%

수강대상: **정보와 기술 카테고리를 이미 수강한 학생. 인공지능과 코딩과목에 대한 아주 기초적인 이해가 있으며 AI가 만들어낸 산출물을 윤리적으로 사용하는 것에 흥미가 있는 학생, 법적 윤리에 대해 다양한 의견을 들어보고 싶은 학생, 기술 개발에 관심이 있는 학생.

수강인원: 최대 25명(소규모 진행)

평가방식: 출석 10% 중간고사 30% 기말대체 포트폴리오 40% 실습 20%

수업방법: PBL(Problem Based Learning) 방식

중간고사 이전까지는 기술적 지식을 학습 후 수업 내용에서 다룬 법적·윤리적 문제를 보완한 인공지능 아이디어를 기획하고 간단한 프로토타입을 시연하는 프로젝트를 진행.

수업체계: 중간고사(~8주차)까지는 이론강의 위주, 이후 팀프로젝트 및 발표 진행

기대효과:

- 1. 사회문제 해결에 관심이 있는 학부생들의 이론적 지식 및 윤리적 의식 함양
- 2. 시의성 있는 주제에 대한 입장을 논리적으로 정리하고 연구 및 실무에 적용 가능한 인사이트 습득
- 3) 팀 프로젝트를 통한 협력 역량 강화

특이사항:

- 1. 본과목은 고등교육혁신원 사회 혁신가 인증 교과목에 해당하며, 수업 과정 중 사회 혁신 및 사회 문제 해결과 관련한 내용의 강의 및 실습이 함께 포함됩니다.
- 2. AI 이론 및 실습을 진행하고 이를 바탕으로 이를 바탕으로 법적인 문제를 인공지능 학습 관점으로 검토하는 프로젝트를 진행합니다.
- 3. 프로젝트 진행이 아주 중요한 부분이기때 Python 및 데이터사이언스에 대한 이해는 필수입니다. 프로그래밍 및 텍스트 등 비정형화된 데이터를 이용한 프로젝트 경험이 수강하는 데 아주 큰 도움이 됩니다.

- 1주차
- 2주차
- 3주차
- 4주차
- 5주차
- 6주차
- 7주차
- 8주차
- 9주차
- 10주차
- 11주차
- 12주차
- 13주차
- 14주차
- 15주차
- 16주차

개괄적 이해, 강의소개		
Machine learning pipe-line + 데이터 수집 기법(1)		
Machine learning, Diving to Deep learning; pipe-line(2)		
NLP		
Computer Vision		
Privacy issue 1 (발생 원리 기술적 검토, 발생 가능한 ML pipe-line 위치 살펴보기)		
Privacy issue 2 (문제 사례, 윤리적 토의, 해결방안, 관련 기술 살펴보기 unlearning 등)		
실습진행주간	중간고사	프로젝트 진행 주간
토론 & Team Formation		
Infringement of intellectual property rights 1 (발생 원리의 기술적 검토, purpose, mechanism)		
Infringement of intellectual property rights 2 (문제사례, 윤리적 토의, 해결방안, 방어/보안 기술 살펴보기; watermark)		
Bias in Models 1 (발생 원리의 기술적 검토)		
Bias in Models 2 (문제사례, 윤리·법적 검토, 해결 방안, 관련 기술 살펴보기)		
기말 프로젝트 (교수자 검토 + 다른 팀과 겹치지 않는지 주제 확인)		
기말 프로젝트 발표		
기말 포트폴리오 제출 (팀별 진행이나 포트폴리오는 개인으로 제출, 2장 이내)		