

AI 위험인식이 규제 유형 선호에 미치는 영향에 관한 탐색적 연구: 머신러닝 방법을 통한 탐색*

김 나 리**

유 한 별***

국문요약

AI(Artificial Intelligence) 기반 기술이 고도화될수록 이와 관련하여 발생할 수 있는 부작용들에 대한 우려도 증가하고 있다. AI로 인해 야기되는 문제들은 개인과 사회 일각의 직접적인 피해로 이어져 강력한 규제 요구를 불러올 수 있지만 AI의 초국가적 성격, 복잡성, 개발의 분산성 등으로 인해 정부의 직접규제로 쉽게 통제할 수 없는 특징이 있다. 이에 공동규제, 독립기관규제, 자율규제 등 대안적 규제방안이 함께 논의되고 있으나 관련한 실증연구는 제한적으로 이뤄져 왔다. 이러한 배경하에 본 연구는 AI 위험인식과 규제 유형의 선호 간의 관계를 살펴보는 데 연구의 목적이 있다. 특히, AI로 인해 발생할 수 있는 사회윤리적 위험 및 기술적 위험에 대한 인식과 정부규제, 공동규제, 독립기관규제, 자율규제 간의 관계를 탐색적으로 살펴보기 위하여 머신러닝 분석과 t-test를 통한 목표 변수 기반 집단 비교를 실시하였다. 분석결과, 먼저 정부규제-공동규제 모형에서 사회윤리적 위험인식 중 기술 악용, 일자리 감소, 개인정보침해와 기술적 위험인식 중에서 시스템 통제 실패 인식이 규제 선호에 영향이 있는 것으로 나타났다. 둘째, 정부규제-독립기관규제 모형에서 일자리 감소 인식과 보안사고, 부정확한 결과에 대한 인식에 따라 규제가 전환될 수 있음을 확인하였다. 마지막으로 정부규제-자율규제 모형에서는 개인정보 침해, 불투명성 인식이 규제 선호에 영향을 미칠 수 있는 것으로 나타났다.

주제어: AI, 위험인식, 정부규제, 공동규제, 독립기관규제, 자율규제

I. 서론

AI와 같은 알고리즘 기반 기술은 점점 우리 삶의 모든 영역에 침투하여 다양한 사회적 변화를 일으키고 있다(Smuha, 2021). 특히 ChatGPT와 같은 대형 언어 모델(LLM)은 의사소통, 창작 등의 방식을 빠르게 변화시키며, 비즈니스, 의학, 교육, 연구, 엔터테인먼트 및 예술에 이르기까지 사회 전반에 영향을 미치고 있다(Hacker et al., 2023). 이러한 변화는 세계적인 디지털화 추세, 빅데이터

* 본 논문의 발전과 출판에 기여해 주신 익명의 심사위원님과 한국지방정부학회 관계자님께 진심으로 감사드립니다.

** 제1저자

*** 교신저자

의 가용성, 통계적 기계학습 및 계산 능력의 향상에 의해 더욱 가속화되는 중이다(Holzinger, 2021).

AI의 적용과 사용의 확산은 양면의 영향을 미치고 있다. 업무 생산성 향상, 자율주행, 사기 및 범죄 감지 등 사회 전반에 다양한 혁신을 불러올 것이란 기대가 높으나(Fast & Horvitz, 2017), 편견, 차별, 개인정보 침해, 통제력 상실, 가짜 뉴스, 딥페이크, 사이버 공격, 자율무기, 일자리 상실 등 AI의 본래 설계 의도에서 벗어난 부작용에 대한 우려가 공존한다(de Almeida et al., 2021; Papyshv, 2024).

이러한 부작용을 예방 및 완화하기 위하여 미국 등 개별 국가와 EU 등에서는 법적, 윤리적 문제의 예방 및 해결에 대한 논의가 이어지고 있으며(Erdélyi & Goldsmith, 2018), 대체로 인공지능 윤리규범을 가이드라인 혹은 지침 형식으로 제시하는 자율적 규제방식을 채택하고 있다(박혜성 외, 2021; 김나래, 2022). 이러한 배경에는 AI 사용과 영향의 초국가적 성격, 시스템의 전문성과 불투명성 등의 본질적 특성, 기술 개발의 분산성 등으로 인해 선불리 정부의 강력한 직접규제를 적용하기 어려운 한계가 존재한다.

하지만 가이드라인, 지침 등 자율적 규제방식이 효과적으로 적용되지 않고, AI 기술로 인해 야기되는 사회문제가 심각해질수록 정부는 더 많은, 더 강력한 규제 압력을 받게 된다(Stone et al., 2022). 이에 정부가 규제를 현명하게 적용하여 규제의 부정적인 영향을 최소화하고 유익한 혁신을 촉진할 수 있도록(Smuha, 2021), 법에 의해 강력히 통제되는 정부의 직접규제보다 유연한 규제(flexible regulation)로 볼 수 있는 공동규제, 독립기관규제, 자율규제 등의 대안적 규제방식에 대한 논의가 함께 이뤄지는 중이다(고학수 외, 2020; 김경동, 2021; 방정미, 2021; 선지원, 2022 등). AI 규제를 위한 접근 방안에 대한 탐색은 AI 규제 방향을 설정하는데 주요한 영향을 미친다(김나래, 2022). 이에 정부는 직접규제를 포함하여 최적의 규제 방안을 모색할 필요가 있다.

지금까지 AI 규제와 관련된 다양한 연구가 진행되었으나, 법적, 윤리적, 국제비교적 측면에서 AI의 문제점들을 논하고, 규제 방안을 제시하는 연구들이 주를 이루었다. 반면 AI의 사용과 확산에 직접적인 영향을 받고 여론을 형성하여 정부의 규제 논의에 압력을 가할 수 있는 일반 국민들을 대상으로 한 연구는 아직 충분히 진행되지 못하였다. 기술 발전은 국민들이 지녀온 기존의 사회적 가치 및 이데올로기와의 충돌, 기존 이해관계자와의 갈등 등 사회적 합의가 필요한 문제들에 대한 규제 이슈를 불러올 수 있다(이종한 외, 2022). 이에 AI 도입과 확산에 따라 발생할 수 있는 위험에 대한 우려 없이 국민의 신뢰를 얻는 것이 중요하다. 부작용과 신뢰 하락에 기인한 여론의 규제 압력은 정부와 정치인의 결정에 영향을 미치며 규제 논의에서 중요한 역할을 할 수 있는바(Kelley et al., 2021), 일반 국민들이 체감하는 AI의 위험과 선호하는 규제기관 유형에 대한 연구의 필요성이 존재한다.

이러한 배경하에 본 연구는 AI 위험인식과 규제 유형 선호 간의 관계를 탐색적으로 살펴보는 데 연구의 목적이 있다. 특히 AI로 인해 발생하는 사회윤리적 및 기술적 위험에 대한 인식과 정부규제, 공동규제, 독립기관규제, 자율규제의 선호 간의 관계를 살펴보기 위한 분석을 시도한다. AI로 인한 잠재적인 위험과 규제 유형 선호에 미치는 영향을 규명하면, 향후 정부와 국회가 관련 정책

형성 시 주요한 함의를 제공할 수 있을 것이다.

II. 이론적 및 개념적 배경

1. AI

AI(인공지능)는 효율성 향상과 비용 절감을 위해 최근 빠르게 성장하고 있는 기술 현상이다(Hassani et al., 2020). AI는 1950년대 후반부터 오랫동안 연구되어 왔지만 아직까지 명확히 합의된 정의가 없으며, 학자마다 조금씩 다르게 정의하고 있다. Ryan(2020)은 AI를 복잡한 작업을 용이하게 하고 인간과 유사한 방식으로 정보를 처리할 수 있도록 설계한 시스템으로 정의하며, Wirtz et al.(2019)은 인식, 이해, 행동 및 학습 등 인간의 특정 핵심 역량과 유사한 지능적인 행동을 하는 시스템으로 정의하고 있다. AI는 컴퓨터 시스템을 통해 인간의 지능 프로세스를 시뮬레이션함으로써 구현되며, 이러한 프로세스는 학습(정보 획득 및 사용 규칙), 추론(결론에 도달하기 위한 규칙 사용) 및 자기교정능력(self-correction)이 포함된다(Gillath et al., 2021). 최근 AI는 향상된 알고리즘과 컴퓨팅 성능 및 방대한 양의 데이터 융합과 음성-텍스트 변환, 텍스트 이해, 이미지 해석 등 기계학습 방법이 접목되어 인간과 유사한 수준의 인식 기능을 제공하고 있다(Rossi, 2018).

현실에서 AI는 자율주행차, 희귀질환 의료 진단 도구, 제품 추천 시스템, 자동 감시 기술, 예측 분석, 금융 사기 및 범죄 감지, 예측 치안 등을 위해 사용되고 있으며(Erdélyi & Goldsmith, 2018), 로봇, 스마트 시스템, 지능형 시스템, 지능형 에이전트, AI 에이전트, AI 알고리즘, 지능형 알고리즘, 자율 시스템 등 AI 기술에 따라 관련 용어들도 다양하게 확장되고 있다(de Almeida et al., 2021). 이와 함께 최근 몇 년 사이 LLM(대형 언어 모델) 기반 서비스인 ChatGPT 등의 출현으로 사용자가 급증하며 AI의 대중화가 이뤄지고 있다(Skjuve et al., 2023).

AI 기술의 활용으로 로봇, 자동화를 통한 업무 생산성 향상, 자율주행, 신약 개발, 맞춤형 의료 및 교육 지원 등 다양한 이로운 변화가 사회 곳곳에 나타나고 있다(Fast & Horvitz, 2017). 이에 다수의 국가가 AI를 전략적 목표로 선언하고, 기업들은 이를 거대한 성장 동력으로 전망하며 개발에 더욱 박차를 가하는 중이다(Holzinger, 2021).

하지만 AI로 인해 이미 발생했거나 예상되는 각종 사회 문제들도 적지 않다. AI의 빠른 작업속도와 비용 효율적인 특성으로 인해 기존 일자리에 영향을 미쳐 일자리 상실, 숙련도 저하 등의 문제를 발생시킬 수 있으며(Fast & Horvitz, 2017), 프로세스 처리과정을 알 수 없는 블랙박스인 투명성 부족(Kieslich et al., 2022), AI 시스템 통제력 상실(Fast & Horvitz, 2017), 새로운 데이터에 따른 부정확한 결과(박혜진, 2022) 등 AI 시스템 내에서 발생하거나 알고리즘에 의해 도출된 결과로 인해 문제가 발생하기도 한다. 또한 기본권, 인간의 존엄성, 생명윤리를 침해하는 문제들인 로봇 군인, 킬러 드론 등의 군사 응용(Fast & Horvitz, 2017), 가짜뉴스, 정보조작, 대량감시 등의 기술 악용(Lockey et al., 2021), 딥페이크 등을 통한 인권침해, 차별을 야기하는 편향성(de Almeida et

al., 2021; Smuha, 2021) 등의 윤리적 문제가 발생하기도 한다.

일부 문제는 여러 문제를 중복적, 단계적으로 발생시킨다. 가령, 빅데이터의 수집, 처리 과정에서 발생할 수 있는 개인정보 유출 및 예측, 알고리즘의 편향적 결정(Kieslich et al., 2022) 등은 초기에 기술 및 설계 문제로 시작되나 이후에는 사생활 침해, 소수자 차별 등 윤리적 문제로 이어진다.

이처럼 AI 부작용과 이에 영향을 받는 사람들과 범위는 매우 광범위하며(Fast & Horvitz, 2017), 기존 데이터를 학습해 새로운 콘텐츠를 생성하는 생성형 AI의 거듭되는 발전으로 인해 문제는 더욱 확산될 것으로 전망된다(de Almeida et al., 2021; Smuha, 2021). 이러한 문제들은 AI의 불완전성, 인공물의 자율성, 데이터 및 추론 프로세스의 가정, 알고리즘 불투명성, 문제 발생 시 불확실한 책임소재 등으로 인해 발생되며, 점점 더 심각해지고 있다(Erdélyi & Goldsmith, 2018; Clarke, 2019).

이에 AI의 도입으로 인해 발생하는 이점을 넘어 의문이 동시에 제기되고 있으며(de Almeida et al., 2021), AI로 인한 위험이 가시화됨에 따라 정부가 합법적이고 윤리적인 사용을 보장하기 위해 적절한 규제를 도입해야 한다는 요구가 커지고 있다(Smuha, 2021). Seth(2024)가 설문조사¹⁾를 기반으로 AI에 대한 대중의 인식(각국 1000명 대상)을 분석한 연구 결과, AI에 대한 엄격한 규제가 필요하다고 응답한 사람은 한국이 84%로 중국(90%) 다음으로 높은 것으로 나타났다(조사대상 10개국 평균 77%). 세부적으로는 남성 84%, 여성 45%, 55세 이상 92%가 엄격한 규제가 필요하다고 응답했으며, 취업자 87%, 미취업자 77%로 취업자의 응답률이 더욱 높았다.

이미 유럽연합은 AI 개발자와 사업자에게 AI의 특정 용도에 대한 요구사항과 의무를 적시하고, AI에 의해 생성된 위험 해결, 허용할 수 없는 위험 금지, AI 시스템의 상용화 전 적합성 평가 등의 내용을 담은 AI법²⁾을 2024년 3월 채택했으며(EU Commission, 2024), 미국은 알고리즘 차별 보호, 개인정보 보호 등의 원칙을 담은 AI 권리장전을 발표한 바 있다(The White House, 2023).

우리나라는 세계적인 추세 및 규제 요구와 함께 제21대 국회에서는 13건의 인공지능 제정법이 발의된 바 있으며,³⁾ 지난 5월 제22대 국회 개원 이후 3개월만에 6건(2024년 8월 기준)의 제정법이 발의되었다. 하지만 혁신과 국민 보호 두 가지 목표를 다 이뤄야 하는 AI의 규제의 복잡성으로 인해 정부는 AI법 제정 논의에 본격적으로 나서지 않고 있으며, 국가인공지능위원회를 구성하여 규제 발굴과 개선, 인공지능 윤리원칙의 확산 등의 업무를 추진하는 등 공공과 민간이 함께 방안을 찾아가는 거버넌스적 접근을 꾀하고 있다(과학기술정보통신부, 2024).

1) 3M State of Science Insights

2) Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence

3) 모두 임기만료 폐기되었다.

2. AI와 규제기관

AI는 사회를 변화시킬 수 있는 엄청난 잠재력을 가지고 있다(Kieslich et al., 2022). 하지만 이 잠재력이 어느 방향으로 향할지 정확히 예측할 수 없기에 예상되는 혹은 이미 발생한 부작용을 해결 및 완화하기 위한 규제 요구가 존재한다. AI 관련 규제는 위험을 최소화하며, AI의 활용을 촉진하고 법적 확실성을 높이는 데 기여할 수 있다(Smuha, 2021). 하지만 국가 경계를 초월하는 초국가적인 규제 범위와 성격(Erdélyi & Goldsmith, 2018), AI 시스템의 높은 복잡성과 전문성으로 인한 평가의 어려움(Alalawi et al., 2024), 개발의 분산성 및 기반시설의 불요(양종모, 2016) 등으로 인하여 AI 규제는 그 효과가 실제로 나타나기 어려우며, 효과적인 규제 형태와 주체가 명확하지 않은 한계가 있다(Alalawi et al., 2024). 이에 정부는 AI로 인해 발생하는 위험과 이점을 고려한 다양한 규제방안을 모색할 필요가 있다(Prainsack & Forgó, 2024).

1) 정부규제

규제는 법률과 입법 절차라는 규범적 기반을 갖춘 특정 유형의 공공 활동을 뜻한다(Christensen & Lægreid, 2007). 규제기관은 규제대상의 행동을 통제하기 위해 권한을 행사하는 실체를 말하며, 정부 행정 기관과 독립위원회, 증권거래소 등의 정부위임기관이 해당된다(Clarke, 2019). 일반적으로 규제기관은 사회적으로 바람직한 특정 가치를 지키고 법을 준수하는 자를 타인의 위법행위와 부당한 경쟁으로부터 보호하기 위한 목적을 가진다(Christensen & Lægreid, 2007). 규제대상은 규제 체계의 적용을 받는 실체자인 기업, 행정 기관, 민간기구, 협회, 개인 등이며(Clarke, 2019), 법률의 도입에 의한 법적 규제, 규제 완화, 재규제(법률 개정), 연성법 등을 통해 규제가 가해진다(Smuha, 2021). 규제는 성격과 목적에 따라 경제적 및 사회적 규제 구분된다. 경제적 규제는 시장 통제나 경제와 관련이 있으며, 사회적 규제는 시장 통제나 경제와 직접적으로 관련되지 않고 산업화의 부작용으로부터 사람이나 환경을 보호하는 규제 형태를 말한다(Hawkins & Hutter, 1993; Gunningham & Rees, 1997).

규제는 일반적으로 국가가 법률, 행정 명령 및 행정 규칙 등 다양한 수단을 통해 규정을 명시하고, 관련 기관의 행동을 감시하고 집행하는 정부의 직접규제를 일컫는다(Bartle & Vass, 2007; Castro, 2011). 정부의 직접규제는 시장실패가 발생, 지속되어 피해를 줄이기 위한 정부의 개입이 불가피할 때 주로 이뤄지며(김종석 · 김태운, 2004), 규제대상 집단의 자유와 행위를 강제적, 공식적으로 제약하는 효과가 있다(이혁우, 2009). 이렇듯 정부의 규제는 공신력이 있으며, 집행력이 강력하고, 규제의 실효성이 높은 특성이 있다(최지수 · 윤석민, 2019). 하지만 기술의 발전과 사회의 변화로 인해 전통적 규제수단인 정부의 직접규제의 효과성에 대한 비판이 증가하면서, 새로운 규제 방안에 대한 요구가 제기되고 있다(심우현 외, 2021). 정부규제 외의 대안적 규제 방안은 법적 지위, 권한, 피규제자와의 관계 등에 따라 공동규제, 독립기관규제, 자율규제 등으로 구분된다(Black, 2001; Thatcher & Sweet, 2002).

2) 공동규제

공동규제(Co-regulation)⁴⁾는 기업과 정부 등 정책 이해관계자가 함께 규제 프로세스를 관리하는 것을 말한다(Castro, 2011). 공동규제의 핵심은 공공과 민간 행위자 간의 협력이다(Eijlander, 2005). 공동규제는 규제에 영향을 받는 이해관계자의 참여로 정책을 더욱 잘 파악하기 위한 목적 하에(Saldaña & Rosón, 2015), 정부, 규제기관, 기업, 민간 이해관계자 등 관련 행위자들 사이의 상호작용, 협상, 논쟁 등의 과정을 통해 규제가 논의되고, 형성된다(Kjaer & Vetterlein, 2018). 또한 규제대상에 참여 유인을 제공하여 이들이 직접 규칙을 개발하도록 하며, 공동규제 프로세스 내에서 일반인, 사회단체 등 민간 이해관계자들의 의견과 이해가 반영된 규제가 형성될 수 있는 특징이 있다(최지수·윤석민, 2019). 이밖에 공동규제는 대화, 평가 및 검토가 반복되는 지속적인 과정으로써 불확실성 하에서 시간이 지남에 따라 결과를 도출하도록 설계된 학습 과정으로도 볼 수 있다(Finck, 2018). 이러한 과정을 통해 정부는 관련 산업의 지식과 전문성을 활용하고, 비용 효율적이고 실행 가능하며 혁신적인 제도를 만들 수 있다(Rubinstein, 2016). 즉, 공동규제는 정부가 주요 행위자로 역할을 하는 규제 정책과정이자 피규제자, 민간 이해관계자 사이의 상호활동을 포함하는 광범위한 개념으로 볼 수 있다(이종한 외, 2022).

공동규제는 정부규제와 자율규제의 특성을 결합한 접근방식으로도 볼 수 있기에, 규제 논의의 결과는 협정, 협약, 입법 등 다양한 형태로 도출된다(Eijlander, 2005). 특히, 규정의 제정과 적용, 위반 시 처벌 부과에 정부가 직접적으로 관여하는 점에서 정부규제와 일부 특성을 공유하며(Saldaña & Rosón, 2015), 참여를 기반으로 유연하게 운영되는 자율규제의 특성도 반영되어 있다(Eijlander, 2005). 하지만 공동규제는 정부와 민간의 역할, 감독과 규제 집행의 범위와 정도 등의 기준이 모호한 한계를 지닌다(최지수·윤석민, 2019).

3) 독립기관규제

독립규제기관(IRA, independent regulatory agencies)은 정부 기관과는 별도로 전문적인 공적 권한을 소유하고 행사하며, 선출된 정치인이나 공무원에 의해 직접 관리되지 않는 기관을 말한다(Thatcher, 2002; Thatcher & Sweet, 2002).

독립규제기관은 재정적, 기능적으로 규제대상 산업, 시민단체, 일반 대중은 물론 정부, 의회, 정당, 정치인으로부터 독립된 기관이 더 많은 권한과 책임에 기반하여 효율적, 생산적, 경제적, 객관적인 규제 업무를 수행할 것이라 전제하에 운영된다(Hanretty & Koop, 2012; Nunes et al., 2015). 이와 함께 정치인의 영향에서 분리되어 기존의 규제정책의 지속가능성을 확보하고(Gilardi, 2005), 규제의 전문성을 강화 및 지속하기 위한 목적도 존재한다(김종석·김태운, 2004).

독립규제기관의 주요 업무는 규칙 제정, 경제 행위 평가 및 조사, 규정 위반에 대한 제재 적용 등이며, 기관의 독립성 수준은 구성원의 인사권, 정치인의 영향 정도, 조직의 법적 지위 등에 따라

4) 본 논문에서는 공동규제와 규제 거버넌스 용어를 혼용하여 사용한다.

결정된다(Hanretty & Koop, 2013). 독립규제기관의 활동은 규제권한(지침, 규정 등 마련), 감독권한(승인, 지시, 경고, 이행 여부 감독 등), 조사 및 제재 권한(정보 요청, 제재 절차 개시, 잠정 조치 채택 및 제재 적용 등) 등에 따라 수행된다(Nunes et al., 2015).

특히 정부는 독립규제기관에 규제 권한을 위임하고, 간섭하지 않음으로써 객관적인 규제 관리와 실행을 보장할 수 있다(Cetin et al., 2016). 이에 Thatcher(2005)는 독립규제기관을 정치인과 별도의 행위자 그룹을 구성한다는 점에서 제3의 세력이라 지칭한바 있다.

4) 자율규제

자율규제는 개별 조직, 기업, 산업이 자체적으로 규제를 수립하고, 준수 여부를 모니터링하며, 구성원의 이행 여부에 따라 직접 제재를 가하는 규제 체계를 말한다(Eijlander, 2005; Bartle & Vass, 2007). 자율규제는 기업이 다른 기업과 독립적으로 스스로를 규제하는 기업 자율규제와 산업 수준에서 업계 내 기업의 행동과 관련된 규칙 및 표준을 설정하는 산업 자율규제로 구분된다(Gunningham & Rees, 1997). 자율규제는 정부의 개입이 적고, 기업과 협회가 자체적으로 규정을 마련하고, 이행을 모니터링하며, 미이행에 대한 조치를 한다는 점에서 비교적 기업 친화적인 규제 제도로 여겨진다. 반면, 자율규제는 정부규제 등 다른 규제 형식과 달리 규정과 규칙의 준수에 강제성이 따르지 않고, 미이행시 처벌이 없거나 권고 등 약한 수준으로 제재가 가해지는 한계가 있어(최지수 · 윤석민, 2019), 의도적인 규제 미이행, 경영진의 도덕적 해이 등 규정준수에 문제가 발생했을 시 정부가 개입하여 정부규제로 전환될 수 있는 한계가 존재한다(Clarke, 2019).

자율규제는 정부의 개입 정도에 따라 다양한 형태로 구분된다. Bartle & Vass(2007)는 법률에 자율규제 계획을 명시함으로써 국가가 참여하는 의무적 자율규제와 법령에 명시되어 있지 않은 비의무적 자율규제로 구분하였으며, Black(2001)은 보다 세부적으로 자발적(Voluntary), 강제적(Coerced), 제재된(Coerced), 의무적(mandated) 자율규제로 분류하였다.

자율규제는 규제 완화나 비규제와는 다르며, 규제 책임을 정부가 아닌 다른 당사자에게 재분배하는 것으로 볼 수 있다(Cohen & Sundararajan, 2015). Ogas(1995)는 해당 활동이 특정 형태의 시장 실패, 특히 외부 효과나 정보 비대칭에 영향을 받거나, 기존 법률이 시장 실패를 바로잡기에는 부적절하거나 비용이 너무 많이 들 때, 혹은 자율규제가 기존의 정부규제보다 문제를 해결하는 더 나은 방법일 때 자율규제 도입의 공익적 정당성이 있다고 판단한다.

Benbear & Coglianese(2012)는 규제 목표를 달성하면서 비용을 최소화하고 자율성이 높은 유연한 규제(flexible regulation) 개념을 논하며, 규제기관(규제 도입 및 실행 주체), 규제대상(피규제자), 규제명령(행동에 대한 지시), 규제에 인한 예상결과(규제 준수 또는 불이행에 따른 결과)라는 네 가지 요소에 따라 규제 유연성 수준이 정해진다고 보았다. 특히, 규제기관과 관련하여 규제대상과 규제기관 간의 거리에 따라 규제의 유연성이 달라진다고 가정하였다.

가령 AI 개발 및 서비스 제공 기업과 기관을 규제대상으로 보고, 앞서 논의한 네 가지 유형의 규제기관을 비교해보면, AI 자율규제는 규제대상과 규제기관(AI 자율규제기구) 간의 거리가 가까워

보다 유연한 규제이며, AI 정부규제는 정부에 의한 직접적인 규제와 제재가 가해지기 때문에 규제 대상과 규제기관의 거리가 멀어 유연성이 낮은 편이다. 공동규제는 AI 기업이 규제거버넌스에 참여하여 일정 정도 영향을 미칠 수 있기에 자율규제보다 유연성이 낮지만 정부규제보다는 유연성이 높은 편이며, 독립기관규제는 정부가 설립한 기구에서 독립된 권한과 책임하에 운영되기에 정부규제보다는 유연성이 높다고 간주할 수 있다. 즉, AI규제와 관련하여 자율규제와 정부규제는 규제 유연성 측면에서 양측 끝단에 존재하며, 독립기관규제와 공동규제 순으로 정부규제보다 유연성이 높다고 볼 수 있다.

규제의 유연성과 기관의 특징에 따라 각 규제 유형들은 AI 규제에 있어 각기 장단점을 지니고 있다. 유연성이 낮은 정부규제는 구속력이 있어 AI로 인한 위험과 기본권 침해를 강하게 대처할 수 있으나(이준복, 2023), 규제대상에 대한 전문성이 낮고 대응 속도가 느려 AI 기술 발전을 사전에 차단할 수 없으며, 개발이 분산적으로, 특별한 기반시설 없이 이뤄지고, 규제보다 혁신의 속도가 빠른 규제지체가 발생할 수 있어 규제 적용이 어려운 한계가 있다(양종모, 2016; 김정화 외, 2023; 이성엽, 2023). 독립규제는 AI 정책 선점을 피하거나, 여론의 의견에 민감한 정부와 정치인의 직접적인 영향에서 벗어나 규제를 집행할 수 있다는 이점이 있어 정부규제보다는 규제 대상에 집중된 정책을 발굴 및 집행할 수 있지만(Gilardi, 2005), 정부규제와 마찬가지로 규제대상과 거리가 멀어 규제 대응 및 피해 방지가 빠르게 이뤄지지 않는 한계가 있다. 공동규제는 정부가 개입하여 일정 정도의 구속력이 있으면서 규제대상이 거버넌스에 참여하여 협력의 가치가 강조되어 정부규제와 독립기관규제보다 유연성이 높지만(이효진, 2023), 거버넌스 내의 AI 이해관계자의 규제 인식에 대한 이념적 특성, 가치관, 규제 목적과 목표 등이 다양하여 구체적인 합의가 제시되기 어려울 가능성이 높다(백수원, 2021). 자율규제 방식은 규제기관과 규제대상의 거리가 가까운 유연한 규제방식으로 변화 속도가 빠른 기술에 적용시 융통성 있게 대응할 수 있는 장점이 있으나(이준복, 2023) 기업의 도덕적 해이가 발생하거나 규정 미준수 시 이를 강제할 수 없기에 실천적 해결 방안 도출이 어려운 측면이 존재한다(김나래, 2022).

이처럼 AI 규제를 하는데 있어 규제기관들은 각각의 한계를 지니고 있다. 하지만 각 규제기관의 특성은 AI 규제 방향을 설정하는데 주요한 영향을 미치기에 효과적인 규제를 적용하기 위한 규제기관의 탐색 노력이 계속해서 경주될 필요가 있다.

3. 선행연구 분석

효과적인 AI 규제 방안을 모색하기 위하여 학계에서는 효과적인 AI 규제 도입을 위한 다양한 연구가 진행되어 왔다. 관련 연구는 해외 규제 입법 동향(박혜성 외, 2021; 김법연, 2023; 유지연·정나영, 2024), 인공지능의 문제점과 규제 방안(고학수 외, 2020; 김경동, 2021; 방정미, 2021; 선지원, 2022; 김성준·홍승현, 2023; 김윤명, 2023; 이준복, 2023; 이효진, 2023; 정남철, 2023), 국제협력(이길원, 2021), 전쟁 자동화(임예준, 2019), 공공 안전(이준복, 2023), AI 스피커 음성정보 규제(전영규·김현경, 2020) 등 세분화된 규제가 주를 이루고 있으며, 최근에는 생성형 AI 관련 규제 연

구(김정화 외, 2023; 양은영, 2023; 이성엽, 2023)가 진행되었다.

다수의 연구 중 본 연구의 연구 주제와 관련된 AI 문제점과 규제방안을 다룬 연구를 살펴보면, 먼저 AI로 인해 야기되는 문제점으로는 통제가능성(임예준, 2019; 박혜성 외, 2021; 박혜진, 2022; 박현석, 2023), 개인정보 침해(전영균·김현경, 2020; 박혜성 외, 2021; 김경동, 2023; 양은영, 2023; 정남철, 2023), 노동력 상실(김경동, 2023; 이성엽, 2023), 불투명성(박혜성 외, 2021; 김경동, 2023; 이준복, 2023), 편향성(박혜성 외, 2021; 김승현 외, 2022; 양은영, 2023; 이성엽, 2023; 이준복, 2023; 정남철, 2023), 허위정보 생성(김정화 외, 2023), 인권침해(김경동, 2023; 김정화 외, 2023; 정남철, 2023), 기술 악용(김정화 외, 2023; 양은영, 2023), 숙련도 저하(이성엽, 2023), 잘못된 정보(이성엽, 2023) 등 AI로 인해 현재 발생하고, 발생할 것으로 전망되는 문제들이 다수 제시되어 있다.

AI로 인해 발생하는 문제들의 규제방안으로는 입법에 의하여 강제적으로 추진되는 정부규제 보다 AI의 특성을 반영한 대안적인 규제방안이 다수 제안되었으며, 규제 거버넌스, 즉 공동규제 제안이 다수를 차지하였다(고학수 외, 2020; 김경동, 2021; 방정미, 2021; 선지원, 2022; 김성준·홍승헌, 2023; 김윤명, 2023; 이준복, 2023; 정남철, 2023). 일부 문헌을 세부적으로 살펴보면, 선지원(2022)과 이준복(2023)은 자율규제와 정부규제 각각의 한계를 언급하며, 이들의 긍정적 측면을 결합한 공동규제 도입을 제시하였으며, 김경동(2021)은 중앙정부와 지방정부, 기업과 민간, 정책전문가집단 등이 협업하는 거버넌스 체계의 구성과 이해관계자의 정책참여 필요성을 역설하였다. 김성준·홍승헌(2023)은 AI 규제 시 발생가능한 위험에 기반한 위험기반규제, 공동규제, 국제기준과의 동조화를 고려할 필요가 있음을 강조하였다. 이를 위해서 위험의 확인, 심각성 평가, 대처방안 고안, 규제기관의 가용자원 분배 등의 과정이 요구된다. 김윤명(2023)은 알고리즘 거버넌스를 제안하며, 이를 위해 기술적 적법절차, 책무성과 투명성의 확보, 규제 정합성 등이 실현되어야 하며, 설명의무의 구체화, 모니터링과 감사제도, 전문기관 설립, 영향평가제도 등이 필요하다고 주장하였다.

독립기관에 의한 규제의 경우, 김광수(2021)는 인공지능 업무와 분쟁이 늘어나게 될 경우 이를 독립적, 효율적으로 관리할 규제기관이 설치되어야 한다고 강조하였으며, 김나래(2022)는 투명한 관리 및 감독 체계를 구축하기 위해 독립된 인공지능 규제기관을 설립하면, AI 관련 분쟁을 조정하고 투명성 여부를 판단하는 등 보다 전문성 있는 감독이 이뤄질 것이라 주장하였다. 알고리즘 거버넌스를 논한 김윤명(2023) 또한 AI, 알고리즘 등 기술적인 역량을 갖춘 인력 구성과 이를 뒷받침할 수 있는 공적 독립 조직 설치의 필요성을 언급하였다.

양은영(2023), 이효진(2023) 등은 자율규제 도입의 필요성을 논하였다. 양은영(2023)은 생성AI의 생성단계와 이용단계에서 발생하는 데이터 편향, 저작권 문제, 악용 가능성 등을 언급하며, 생성 AI 특성상 생성과정에서 발생하는 문제규명의 어려움으로 인해 개발자와 관련 기업이 강력한 자율규제를 할 수 있도록 강제할 필요가 있다고 강조하였다. 이효진(2023)은 자율규제의 실효성을 확보할 수 있도록 법적으로 규율할 부분의 범위를 정하고, 협력을 주도할 독립적인 조직을 마련하여 지속가능한 규제와 실질적 구속력을 가질 수 있도록 해야 한다고 주장하였다. AI 스피커 음성

정보 관련 문제를 제시한 전영균·김현경(2020)은 기술 발전에 유연한 대응이 가능한 기업 주도의 자율규제를 우선적으로 적용해야 한다고 강조하였다.

이처럼 선행연구들은 AI로 인해 발생하거나 예상되는 문제점을 지적하며, 정부규제의 대안으로 공동규제, 자율규제, 독립기관 규제에 대해서 논의하고 있다. 하지만 AI로 인해 발생하거나, 발생할 수 있는 문제들을 광범위하게 언급하고 이에 대한 규제방안을 논의하면서도 AI로 인해 발생한 혹은 발생할 수 있는 문제들의 특성에 따른 규제방안에 대한 논의는 제한적으로 진행되었다. 또한 대부분의 연구는 법학 영역에서 기존의 법을 분석하고 새로운 법과 규정을 제안하는 방식으로 이뤄졌으며, 실제 AI 도입과 확산으로 인한 부작용에 직접적인 영향을 받고, 규제여론을 형성할 수 있는 사람들을 대상으로 한 연구가 이뤄지지 못한 한계가 있다. 이에 본 연구는 AI로 인해 야기될 수 있는 다양한 위험인식과 규제 유형 선호 간의 관계를 살펴보는데 연구의 차별성이 있다.

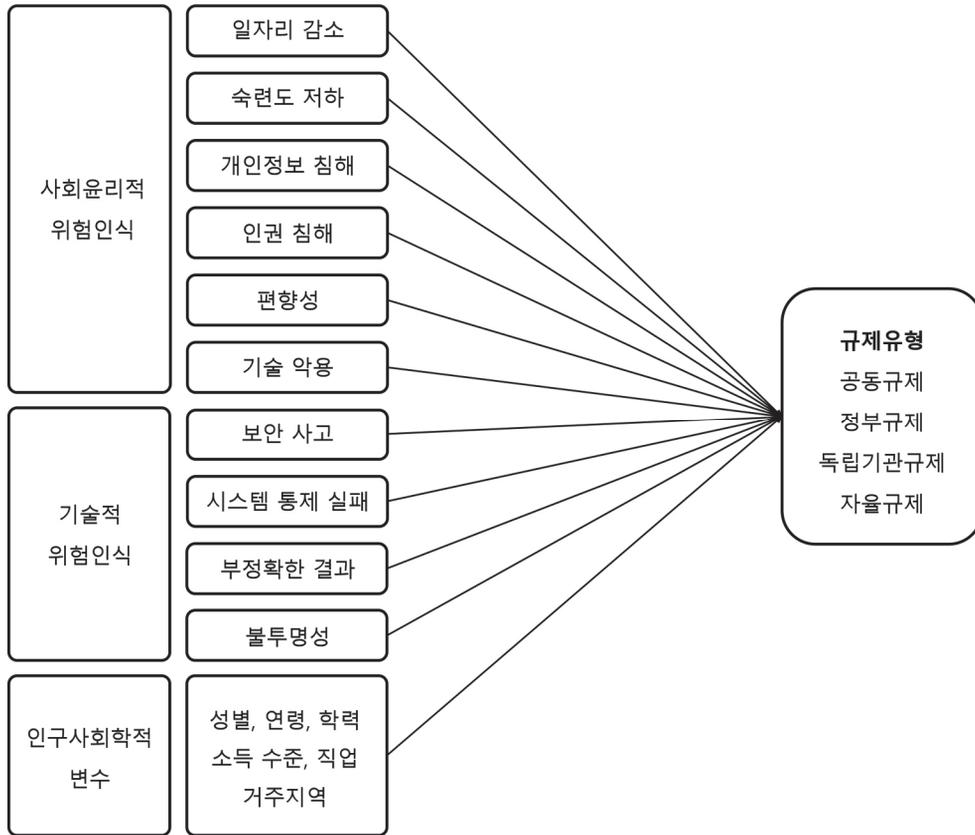
Ⅲ. 연구의 설계 및 분석방법

1. 자료수집

본 논문은 사회적가치연구원의 ‘한국인이 바라본 사회문제(2023)’ 데이터를 활용하여 수행되었다. 해당 조사는 한국인이 인식하는 심각한 사회문제와 개인 삶에 미치는 영향력을 진단하고 해결 방법과 주체 등을 조사하기 위해 사회적가치연구원과 트리플라잇이 공동으로 시행한 반복횡단자료이다. 본 데이터는 인공지능의 잠재적인 부작용, 인공지능과 사회문제에 대한 인식 등의 문항을 포함하고 있다. 해당 조사는 전국 만 20세 이상 성인남녀 1,000명을 대상으로 2023년 5월 4일부터 11일까지 할당표집 방법을 사용하여 온라인과 모바일 조사로 실시되었으며, 한국사회과학자료원(KOSSDA)에서 데이터 제공하고 있다. 본 연구에서는 분석에서 활용한 변수에 유효한 응답을 한 984명의 표본이 분석에 사용되었다.

2. 분석 모형과 가설 설정

본 연구는 AI에 대한 위험인식이 규제 유형 선호에 미치는 영향을 파악하기 위한 분석을 진행한다. 이를 위한 연구모형은 <그림 1>과 같다.



〈그림 1〉 분석틀

우리나라의 신기술과 신서비스 규제는 시민들과 이해관계자들의 영향을 받아왔다. 가령 민간 택시 업계의 움직임이 플랫폼 택시(타다) 규제 도입 과정에 큰 영향을 미쳤으며(유한별, 2020), 온라인 플랫폼 사업자들의 쇼핑, 배달, 웹툰, 뉴스 등의 사업들도 골목상권 침해, 유해 콘텐츠, 가짜 뉴스 등 국민의 일상에 부정적인 영향을 미치는 이유로 규제 논의가 이뤄져 왔다(김나리·오서은, 2024). AI 기술과 기반 서비스들은 아직 사회 전반에 적용 및 확산되고 있지 않지만, 정부가 약 9,000억 원의 예산을 투입하여 전국민 AI 일상화를 추진하고(과학기술정보통신부, 2023), 생성형 AI의 발전속도가 기존 일자리를 위협하는 등 AI가 국민의 일상에 영향을 미치기 시작하면서 가까운 미래에 AI의 영향과 결과에 대한 사회적 논의가 거세질 가능성이 높다. 실제 Seth(2024)의 연구에서는 우리나라 응답자 84%가 AI에 대한 엄격한 규제가 필요하다고 인식하였다.

선행연구에서 살펴보았듯이, AI로 인해 발생하고 있거나 발생할 수 있는 문제들은 통제가능성, 개인정보 침해, 노동력 상실, 불투명성, 편향성, 허위정보 생성, 인권침해, 기술 악용, 숙련도 저하, 잘못된 정보 등으로 매우 다양하다(임예준, 2019; 박혜성 외, 2021; 박혜진, 2022; 박현석, 2023, 전영균·김현경, 2020; 김승현 외, 2022; 김경동, 2023; 양은영, 2023; 이성엽, 2023; 이준복, 2023; 정남철, 2023 등). 이러한 문제들은 사회윤리적 위험과 기술적 위험으로 구분할 수 있는데, 사회윤리

적 문제는 AI의 사용과 확산으로 인해 일자리 감소, 인권침해, 개인정보 침해, 편향성(de Almeida et al., 2021) 등의 사회윤리적 성격의 부작용이 결과적으로 나타나며, 일반 국민들에게 영향이 가해지는 것을 의미한다. 기술적 문제는 AI 기술의 개발과 적용 과정에서 발생하는 문제들로 보안사고, 잘못된 정보(Cerf & Waytz, 2023) 등이 해당될 수 있으며, 이는 주로 기술과 시스템상에서 발생하는 문제로서 기업 및 개발자 등의 통제 범위 안에 있는 위험들로 구성된다.

사회윤리적 문제와 기술적 문제는 원인과 문제 발생 지점, 문제로 인해 영향을 받는 이해관계자 등이 각기 다르다. 이는 AI로 인한 문제들을 획일화된 한 가지 방안이나, 특정 규제방안만을 적용하여 풀어낼 수 없음을 의미한다. 이에 선행연구들과 연구자들도 AI 규제에 대하여 각기 다른 방안을 제시하고 있다. 가령 고태수 외(2020)는 AI 윤리 문제를 규제 거버넌스로 접근하였으며, 김광수(2021)은 사생활 보호, 차별금지 등의 분쟁 해결을 위하여 독립적인 전담기관을 설치하여야 한다고 주장하였다. 양은영(2023)은 개인정보 침해, 데이터 편향 등을 규제하기 위하여 자체 감독 시스템 구축 등을 통한 자율규제의 필요성을 언급한 바 있다. 하지만 이러한 접근방식의 한계는 AI로 인해 파생된 문제들을 단일 기관에서 모두 규율하는 비교적 포괄적인 접근을 하고 있는 점이다.

반면, 국민들이 인식하고 경험하는 사회윤리적 위험과 기술적 위험의 위험성과 그의 영향, 해결 방안 등은 각기 다를 수 있다. 가령 데이터의 프라이버시 및 보안에 대해 위험인식을 갖고 있을 경우 AI가 강한 규제를 통해 보호되어야 한다고 생각하는 경향이 있다(Machleidt et al., 2024).

이처럼 AI의 도입과 확산에 따라 직접적인 영향을 받는 주체인 국민들은 AI로 인해 발생하는 각각의 문제를 해결하기 위해 필요하다고 여겨지는 규제의 강도와 유연성에 따라 선호하는 규제 유형이 다를 수 있다. 이에 본 연구는 다음과 같은 가설을 설정하였다.

〈가설 1〉 AI에 대한 사회윤리적 위험인식은 규제 유형 선호에 영향을 미칠 것이다.

- 〈1-1〉 AI로 인한 일자리 감소에 대한 위험인식은 규제 유형 선호에 영향을 미칠 것이다.
- 〈1-2〉 AI로 인한 숙련도 저하에 대한 위험인식은 규제 유형 선호에 영향을 미칠 것이다.
- 〈1-3〉 AI로 인한 개인정보 침해에 대한 위험인식은 규제 유형 선호에 영향을 미칠 것이다.
- 〈1-4〉 AI로 인한 인권침해에 대한 위험인식은 규제 유형 선호에 영향을 미칠 것이다.
- 〈1-5〉 AI 편향성에 대한 위험인식은 규제 유형 선호에 영향을 미칠 것이다.
- 〈1-6〉 AI 기술 악용에 대한 위험인식은 규제 유형 선호에 영향을 미칠 것이다.

〈가설 2〉 AI에 대한 기술적 위험인식은 선호하는 규제 유형에 영향을 미칠 것이다.

- 〈2-1〉 AI 보안 사고에 대한 위험인식은 규제 유형 선호에 영향을 미칠 것이다.
- 〈2-2〉 AI 시스템 통제 실패에 대한 위험인식은 규제 유형 선호에 영향을 미칠 것이다.
- 〈2-3〉 AI의 부정확한 결과에 대한 위험인식은 규제 유형 선호에 영향을 미칠 것이다.
- 〈2-4〉 AI 불투명성에 대한 위험인식은 규제 유형 선호에 영향을 미칠 것이다.

3. 머신러닝 방법을 위한 변수의 측정과 분석방법

1) 목표변수(Target Variable; 종속변수)

본 연구의 종속변수는 규제 유형이다. 규제 유형은 다수의 문헌(고학수 외, 2020; 김경동, 2021; 방정미, 2021; 선지원, 2022; 김윤명, 2023; 등)에서 AI 규제 방안으로 언급되고 있는 공동규제, 정부규제, 독립기관규제, 자율규제를 명목형 변수로 설정하였다. 해당 변수는 'AI에 대한 규제를 누가 해야한다고 생각하십니까'라는 문항을 활용하였으며, 해당 문항은 정부 혹은 정부 산하 규제/감독 기관에서 주도해야 한다(1), 기업, 정부, 시민사회 등 모든 주체가 동등하게 참여해야 한다(2), 정부, 산업계로부터 독립적인 규제기관에서 전담해야 한다(3), AI를 개발하고 사용하는 산업과 기업에서 주도해야 한다(4)로 구성되어 있다.

2) 특성변수(Feature Variable; 독립변수, 통제변수)

본 연구의 특성변수는 사회윤리적 위험인식과 기술적 위험인식으로 구분하여 구성하였다. 위험인식은 일자리 대체, 개인정보 보호 문제, AI가 악의적인 목적으로 사용될 가능성 등 AI와 관련된 위험에 대한 사람들의 태도를 뜻한다(Gerlich, 2023). 먼저 사회윤리적 위험인식은 일자리 감소, 숙련도 저하, 개인정보 침해, 인권침해, 편향성, 기술 악용(김경동, 2023; de Almeida et al., 2021; Díaz-Rodríguez et al., 2023; Papyshv, 2024 등)으로 구성하였으며, 기술적 위험인식은 보안 사고, 시스템 통제 실패, 부정확한 결과, 불투명성(양종모, 2016; Buiten, 2019; 박현석, 2023; Cerf & Waytz, 2023 등)을 포함하였다. 해당 변수들은 'AI로 인한 잠재적인 위험 중 가장 걱정되는 위험'에 대한 문항을 활용하였다. 잠재적인 위험인식이 있다는 응답은 1, 없다는 응답은 0으로 코딩하였다.

또한 본 연구에서는 성별, 연령, 지역, 학력, 소득수준, 직업, AI 사용경험을 인구나사회학적 변수로 설정하였다. 성별은 남성(0), 여성(1)로 코딩하였다. 연령은 20~69세까지 연속형 변수로 설정하였다. 학력은 고등학교 졸업(1), 대학교 졸업(2), 대학원 재학 이상(3)으로, 소득수준은 200만 원 이하(1), 201~400만 원 이하(2), 401~600만 원 이하(3), 601만 원 이상(4)으로 코딩하였다. 직업은 자영업(1), 사무직외(2), 사무직(3), 전문관리직(4), 대학생(5), 주부·무직(6), 기타(7)로 설정하였다. 자영업에는 농업, 임업, 어업, 축산업자, 자영업자(소규모 장사 및 개인택시 등), 프리랜서가 포함되며, 판매서비스기능직은 판매, 영업, 서비스직, 기능직, 작업직이 포함된다. 사무직에는 사무직, 교사 등이 포함되며, 전문관리직은 고위공무원, 기업 고위직, 대학교수, 의사, 변호사 등이 포함된다. AI 서비스 경험 유무는 사용해 봤다(1), 사용해 보지 않았다(2), 잘 모르겠다(3)로 구분된다. 거주 지역은 수도권(0)과 비수도권(1)으로 구분하였다.

〈표 1〉 변수의 측정과 지표

구분	변수		측정
목표변수 (Target Variable)	규제 유형		1 정부규제(준거) 2 공동규제 3 독립기관규제 4 자율규제
특성변수 (Feature Variable)	사회윤리적 위험인식	일자리 감소	0 아니다 1 그렇다
		숙련도 저하	0 아니다 1 그렇다
		개인정보 침해	0 아니다 1 그렇다
		인권 침해	0 아니다 1 그렇다
		편향성	0 아니다 1 그렇다
		기술 악용	0 아니다 1 그렇다
	기술적 위험인식	보안 사고	0 아니다 1 그렇다
		시스템 통제 실패	0 아니다 1 그렇다
		부정확한 결과	0 아니다 1 그렇다
		불투명성	0 아니다 1 그렇다
	인구사회학적 변수	성별	0 남성 1 여성
		연령	1 20대 2 30대 3 40대 4 50대 5 60대
		학력	1 고등학교 졸업 2 대학교 졸업 3 대학원 이상
		소득 수준	1 200만 원 이하 2 201~400만 원 이하 3 401~600만 원 이하 4 601만 원 이상
		직업	1 자영업 2 판매서비스기능직 3 사무직 4 전문관리직 5 대학생 6 주부·무직 7 기타
거주지역		0 수도권 1 비수도권	

4. 분석방법

본 연구는 Python 프로그램을 이용하여 리코딩과 데이터클리닝을 거친 후 분석을 실시하였다. 먼저, 목표변수와 특성변수의 특성을 파악하기 위해 교차분석을 실시하였다. 또한 특성변수가 종속변수에 미치는 영향을 분석하기 위하여 머신러닝 분석에 기반한 데이터 분석을 실시하였다. 머신러닝 분석 방법은 데이터를 train data set과 test data set으로 나누어 분석하는 방법으로 본 연구에서는 층화 무작위 추출(stratified random sampling)에 기반하여 데이터를 train, test 데이터로 분류한 후 머신러닝 데이터 분석을 적용하였다. 해당 머신러닝 데이터 분석은 로지스틱 회귀(LR), 의

사결정나무(DT), 서포트 벡터 머신(SVM), 인공 신경망(NN), K-최근접 이웃(KNN), 랜덤 포레스트(RF), 배깅(Bagging), 부스팅(Boosting) 방법을 사용하였다.

상기 모형에 대한 간단한 설명을 진행하면 다음과 같다. 먼저, 로지스틱 회귀분석은 분석 대상이 두 집단 혹은 그 이상의 집단일 경우 개별 관측치들이 어느 집단에 포함될 수 있는지를 분석하는데 사용되는 분석방법이다. 로지스틱 회귀분석은 종속변수가 명목척도로 측정된 범주형 변수일 때 사용되며, 사건이 일어날 경우(1)와 일어나지 않을 경우(0)로 구분되며, 종속변수의 값은 확률의 개념을 갖게 된다. 독립변수에는 연속형 변수와 범주형 변수를 모두 사용할 수 있다(노형진, 2016). 로지스틱 회귀분석 중 다항 로지스틱 회귀분석은 종속변수가 비연속적 변수이면서 범주가 세 개 이상이어야 하며, 독립변수는 연속형, 범주형 변수를 사용할 수 있다(채구목, 2014).

다음으로, 의사결정나무는 의사결정을 나무 형태로 구성한 모델로, 각 노드에서 특정 특징을 기준으로 데이터를 분할하고, 이를 반복하여 최종적으로 잎 노드에서 분류 결과를 출력한다(Quinlan, 1986). 또한 서포트 벡터 머신은 결정 경계를 최대화하는 방식으로 데이터를 분류하는 모델이다. SVM은 선형 및 비선형 데이터 모두에 적용 가능하며, 특히 고차원 데이터 분류에 효과적이다. 인공 신경망 (Artificial Neural Network, NN)은 인간 뇌의 신경망 구조를 모방한 모델로, 딥러닝 기술의 발전과 함께 자연어 처리, 이미지 인식 등 다양한 분야에 활용되고 있으며, 합성곱 신경망(CNN), 순환 신경망(RNN), 트랜스포머(Transformer) 등 다양한 구조의 신경망 모델이 개발되어 특정 데이터 유형 및 문제 해결에 특화하여 적용할 수 있다(LeCun et al., 2015). K-최근접 이웃 (K-Nearest Neighbors, KNN)은 새로운 데이터 포인트와 가장 가까운 K개의 이웃을 기반으로 분류하는 모델이다(Cover & Hart 1967).

마지막으로 모형의 성능을 보팅(voting)과 같은 방법을 통해 통합적으로 활용하는 앙상블(ensemble) 방법은 배깅(bagging)과 부스팅(boosting) 방법이 있는데, 배깅 방법 중 하나로 랜덤 포레스트(random forest)를 포함하며, 랜덤 포레스트는 여러 개의 의사결정나무를 생성하고, 각 나무의 예측 결과를 종합하여 최종 분류 결과를 출력하는 앙상블 모델이다(Breiman, 2001). 더하여, 배깅은 여러 개의 학습 데이터 부분 집합을 생성하고, 각 부분 집합에서 학습된 모델의 예측 결과를 종합하여 최종 분류 결과를 출력하는 앙상블 모델이며, 부스팅은 여러 개의 약한 학습기를 순차적으로 학습하고, 이전 학습기의 오류를 보완하는 방식으로 강력한 학습기를 생성하는 앙상블 모델이다(Breiman, 1996; Freund & Schapire, 1997). 본 연구에서는 부스팅 방법 중 하나인 XGBoost를 활용한다. XGBoost의 경우 Gradient Boosting 알고리즘을 확장한 기법이며, 최적화를 통해 학습 속도와 성능을 향상시킨 모델이다(Chen & Guestrin, 2016).

이러한 다수의 모형을 구축하여 비교·분석하는 이유는 각 모형의 성능이 다양한 모형의 특성에 기반한 학습을 통해 다를 뿐 아니라 데이터 특성에 맞는 최적의 모형을 선별하여 특성변수와 목표변수의 추정과 설명에 가장 적절한 모형을 선택하기 위함이다. 또한 본 연구가 AI와 관련한 위험의 인식이 규제 유형 선호에 미치는 영향을 탐색하기 위함을 목표로 하기 때문에 여러 모형을 통한 데이터 분석 및 예측과 설명력 향상에 초점을 맞춰 탐색에 대한 해석의 일반화를 목표로 하기 때문이다.

본 연구에서는 상술한 바와 같은 이유로 머신러닝 분류기(classifier)를 활용하여, 특성변수가 목표변수에 미치는 영향을 추정하여 목표변수의 분류값을 예측하는 모형을 구축하여, 목표변수의 예측값($Y_{predict}$)과 목표변수의 실제값(Y_{test})을 비교하여 분류기의 정확도(accuracy)와 F1 score를 도출하여 가장 예측의 정확도가 높은 모형의 estimator(feature importance, beta, etc.)를 구하여 해당 특성변수가 목표변수에 미치는 영향을 탐색한다.

IV. 분석결과

1. 표본의 일반적 특성

본 연구의 분석 대상인 표본의 일반적 특성은 다음과 같다. 성별은 남성 506명(51.4%), 여성 478명(48.6%)을 차지하고 있다. 연령은 50대가 233명(23.7%)로 가장 많은 수를 차지하며, 20대 165명(16.8%), 30대 176명(17.9%), 40대 205명(20.8%), 50대 233명(23.7%), 60대 이상 205명(20.8%)으로 구성되어 있다. 학력은 대학교 졸업이 599명(60.9%)로 가장 높은 빈도를 보이며, 고등학교 졸업 293명(29.8%), 대학원생 이상 92명(9.3%)이 뒤를 잇는 것으로 나타났다. 소득 수준은 201만 원 이상 400만 원 이하가 333명(33.8%)로 다수를 차지하고 있으며, 다음으로 600만 원 이상이 271명(27.5%), 401만 원 이상 600만원 이하 251명(25.5%), 200만 원 이하 129명(13.1%)로 구성되어 있다. 직업은 사무직이 332명(33.7%)으로 가장 많은 수를 차지하며, 주부·무직 215명(21.8%), 판매서비스기능직 128명(13%), 자영업 97명(9.9%), 전문관리직 96명(9.8%) 등의 순으로 이어진다. 거주지역은 수도권 521명(52.9%), 비수도권 463명(47.1%)로 수도권이 다소 높은 수준이다.

〈표 2〉 표본의 일반적 특성

구분		빈도	퍼센트(%)
성별	남	506	51.4
	여	478	48.6
연령	20대	165	16.8
	30대	176	17.9
	40대	205	20.8
	50대	233	23.7
	60대 이상	205	20.8
	학력	고등학교 졸업	293
	대학교 졸업	599	60.9
	대학원 이상	92	9.3

구분		빈도	퍼센트(%)
소득수준	200만 원 이하	129	13.1
	200~400	333	33.8
	400~600	251	25.5
	600만 원 이상	271	27.5
직업	자영업	97	9.9
	사무직외	128	13.0
	사무직	332	33.7
	전문관리직	96	9.8
	대학생	56	5.7
	주부무직	215	21.8
	기타	60	6.1
거주지역	수도권	521	52.9
	비수도권	463	47.1

2. 변수 간 관계

독립변수, 통제변수와 종속변수 간 관계를 검토한 결과는 아래 <표 4>과 같다. 먼저 통제변수인 일반적 특성을 살펴보면, 성별, 연령, 학력, 소득수준, 직업의 응답자는 공동규제를 정부규제, 독립규제, 자율규제보다 선호하는 것으로 나타났다. 다만 두 번째로 선호하는 규제유형은 연령별, 학력별, 소득수준별, 직업별 차이가 있는 것으로 나타났다. 가령 20대는 공동규제 다음으로 자율규제를 선호하는 반면 30대는 정부규제를, 40대는 독립규제를, 50대와 60대는 정부규제를 선호하는 것으로 드러났다. 거주지역은 수도권에서는 정부규제를 가장 선호하였으며, 비수도권에서는 공동규제를 가장 선호하는 결과가 도출되었다.

독립변수인 AI 위험인식과 관련하여, 응답자는 공동규제를 가장 선호하는 것으로 나타으며, 위험인식이 있는 경우와 없는 경우 선호의 차이는 존재하였다. 두 번째로 선호하는 규제 또한 변수에 따라, 또한 위험기각이 있고, 없음에 따라 정부규제, 독립기관규제, 자율규제가 다르게 나타났다. 가령 일자리 감소와 숙련도 저하 위험인식이 없을 경우 독립기관규제를 두 번째로 선호하였으나, 위험인식이 있을 경우에는 정부규제를 두 번째로 선호하는 것으로 나타났다.

〈표 3〉 변수 간 관계

구분		N	공동	정부	독립	자율	x2
일자리 감소	없음	984	273(49.7%)	96(17.5%)	97(17.7%)	83(15.1%)	10.076*
	있음		260(59.8%)	64(14.7%)	58(13.3%)	53(12.2%)	
숙련도 저하	없음	984	371(52.8%)	112(15.9%)	119(16.9%)	101(14.4%)	3.735
	있음		162(57.7%)	48(17.1%)	36(12.8%)	35(12.5%)	
개인정보 침해	없음	984	373(54.9%)	115(16.9%)	107(15.7%)	85(12.5%)	3.546
	있음		160(52.6%)	45(14.8%)	48(15.8%)	51(16.8%)	
인권 침해	없음	984	438(54.8%)	128(16%)	126(15.8%)	107(13.4%)	1.052
	있음		95(51.4%)	32(17.3%)	29(15.7%)	29(15.7%)	
편향성	없음	984	467(54.1%)	137(15.9%)	136(15.7%)	124(14.4%)	2.172
	있음		66(55%)	23(19.2%)	19(15.8%)	12(10%)	
기술 악용	없음	984	274(50.6%)	105(19.4%)	84(15.5%)	78(14.4%)	10.422*
	있음		259(58.5%)	55(12.4%)	71(16%)	58(13.1%)	
보안 사고	없음	984	395(56.4%)	107(15.3%)	109(15.6%)	89(12.7%)	5.907
	있음		138(48.6%)	53(18.7%)	46(16.2%)	47(16.5%)	
시스템 통제 실패	없음	984	295(52.9%)	87(15.6%)	95(17%)	81(14.5%)	2.533
	있음		238(55.9%)	73(17.1%)	60(14.1%)	55(12.9%)	
부정확한 결과	없음	984	405(54%)	123(16.4%)	118(15.7%)	104(13.9%)	.060
	있음		128(54.7%)	37(15.8%)	37(15.8%)	32(13.7%)	
불투명성	없음	984	472(55.2%)	139(16.3%)	129(15.1%)	115(13.5%)	3.773
	있음		61(47.3%)	21(16.3%)	26(20.2%)	21(16.3%)	
성별	남	984	276(54.5%)	86(17%)	85(16.8%)	59(11.7%)	4.618
	여		257(53.8%)	74(15.5%)	70(14.6%)	77(16.1%)	
연령	20대	984	66(40%)	19(11.5%)	31(18.8%)	49(29.7%)	52.940***
	30대		97(55.1%)	34(19.3%)	26(14.8%)	19(10.8%)	
	40대		120(58.5%)	33(16.1%)	36(17.6%)	16(7.8%)	
	50대		138(59.2%)	40(17.2%)	32(13.7%)	23(9.9%)	
	60대 이상		112(54.6%)	34(16.6%)	30(14.6%)	29(14.1%)	
학력	고등학교 졸업	984	164(56%)	39(13.3%)	38(13%)	52(17.7%)	16.612*
	대학교 졸업		326(54.4%)	96(16%)	102(17%)	75(12.5%)	
	대학원 이상		43(46.7%)	25(27.2%)	15(16.3%)	9(9.8%)	
소득 수준	200만 원 이하	984	74(57.4%)	13(10.1%)	21(16.3%)	21(16.3%)	8.999
	200~400		170(51.1%)	55(16.5%)	56(16.8%)	52(15.6%)	
	400~600		137(54.6%)	46(18.3%)	41(16.3%)	27(10.8%)	
	600만 원 이상		152(56.1%)	46(17%)	37(13.7%)	36(13.3%)	
직업	자영업	984	55(56.7%)	11(11.3%)	17(17.5%)	14(14.4%)	13.408
	판매서비스기능직		70(54.7%)	23(18%)	19(14.8%)	16(12.5%)	
	사무직		175(52.7%)	56(16.9%)	55(16.6%)	46(13.9%)	
	전문관리직		53(55.2%)	14(14.6%)	15(15.6%)	14(14.6%)	
	대학생		27(48.2%)	6(10.7%)	10(17.9%)	13(23.2%)	
	주부무직		119(55.3%)	37(17.2%)	29(13.5%)	30(14%)	
	기타		34(56.7%)	13(21.7%)	10(16.7%)	3(5%)	
거주 지역	수도권	984	28(53.7%)	92(17.7%)	75(14.4%)	74(14.2%)	2.779
	비수도권		253(54.6%)	68(14.7%)	80(17.3%)	62(13.4%)	

*p<.05, **p<.01, ***p<.001

3. 머신러닝 분석결과

AI 위험인식이 규제유형 선호에 미치는 영향을 파악하기 위하여 머신러닝 분석 방법을 통한 분류 분석을 실시하였다. AI 정부규제가 가지는 한계들로 인하여 공동규제, 독립기관규제, 자율규제 등 다양한 유형의 규제방안이 제안되고 있는 바 다양한 유형의 규제방안을 정부규제와 비교하기 위하여 준거집단은 정부규제로 설정하였다. 다양한 머신러닝 분류 모델을 사용하여 규제변수(정부규제: 1, 공동규제: 2, 독립기관규제: 3, 자율규제: 4)의 클래스 전환(1에서 2로, 1에서 3으로, 1에서 4로)을 예측하고 분석하였다. 또한, 데이터의 특성상 규제 클래스 불균형 문제를 해결하기 위해 SMOTE(Synthetic Minority Over-sampling Technique)를 사용하였다. 각 모델의 성능 평가를 위하여 혼동 행렬, 정확도, F1 스코어를 활용하였으며, 가장 정확도가 높은 모형의 특성 중요도를 도출하여 어떠한 특성(feature)가 규제 클래스 변화에 영향을 미치는지 파악하였다. 여기서 혼동행렬이란, 혼동 행렬은 분류 모델의 성능을 파악하는 방법 중 하나로, 모형이 실제 클래스(y test)와 모델이 예측한 클래스(y predict) 간의 관계를 나타낸다. F1 스코어는 분류 모델의 성능을 평가하는 지표 중 하나로, Precision(정밀도)⁵⁾과 Recall(재현율)⁶⁾의 조화 평균이다. 정확도는 모든 예측 중에서 얼마나 많은 예측이 올바른지를 나타내는 지표이다. 머신러닝 모형을 구축하고, 이러한 분석을 진행하는데 python 프로그램에 기반하였으며, 구글(google)의 Colaboratory(pro)를 활용하여 분석을 진행하였다. 먼저, 모형의 성능은 다음 <표 4>와 같다.

<표 4> 머신러닝 모델의 성능

Target Transition	Model	Accuracy	F1 Score
1-2	Logistic Regression	0.644231	0.668937
	Decision Tree	0.649038	0.65704
	SVM	0.658654	0.638585
	Neural Network	0.658654	0.651902
	KNN	0.495192	0.534742
	Random Forest	0.721154	0.674962
	Bagging	0.677885	0.671513
	Boosting	0.701923	0.679212

5) 모델이 Positive로 예측한 것 중 실제 Positive의 비율

6) 실제 Positive 중 모델이 Positive로 올바르게 예측한 비율

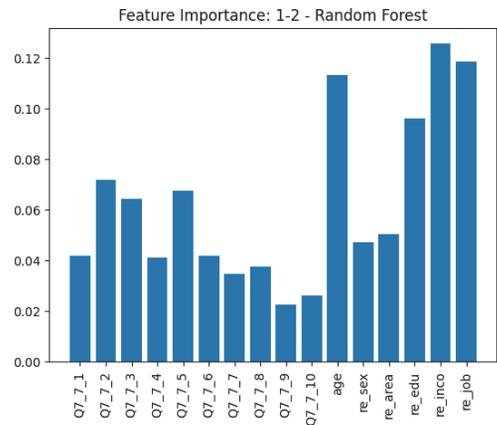
Target Transition	Model	Accuracy	F1 Score
1-3	Logistic Regression	0.494737	0.494401
	Decision Tree	0.473684	0.467182
	SVM	0.463158	0.457744
	Neural Network	0.410526	0.410526
	KNN	0.463158	0.463158
	Random Forest	0.473684	0.471222
	Bagging	0.505263	0.503612
	Boosting	0.452632	0.452267
1-4	Logistic Regression	0.595506	0.596018
	Decision Tree	0.516854	0.510725
	SVM	0.539326	0.535772
	Neural Network	0.438202	0.429442
	KNN	0.460674	0.461492
	Random Forest	0.561798	0.553565
	Bagging	0.595506	0.593224
	Boosting	0.58427	0.576459
algorithms (python)	"Logistic Regression": LogisticRegression(), "Decision Tree": DecisionTreeClassifier(), "SVM": SVC(probability=True), "Neural Network": MLPClassifier(max_iter=300), "KNN": KNeighborsClassifier(), "Random Forest": RandomForestClassifier(), "Bagging": BaggingClassifier(), "Boosting": XGBClassifier(eval_metric='mlogloss')		

1(정부규제)-2(공동규제) 분류 모형의 경우, Random Forest와 Boosting 모형이 상대적으로 높은 정확도와 F1 스코어를 보였다. 특히, Random Forest는 정확도 0.721과 F1 스코어 0.675로 가장 우수한 성능을 보였다. KNN 모형은 가장 낮은 성능을 보였으며, 이는 정확도 0.495와 F1 스코어 0.535로 나타났다. 1(정부규제)-3(독립기관규제) 분류모형의 경우, 모든 모형의 성능이 상대적으로 낮게 나타났다. Bagging 모형이 정확도 0.505와 F1 스코어 0.504로 가장 높은 성능을 보였으나, 전반적으로 모형의 성능은 만족스럽지 못했다. 1(정부규제)-4(자율규제)의 경우, Logistic Regression과 Bagging 모형이 정확도 0.596으로 가장 높은 성능을 보였으며, F1 스코어도 각각 0.596과 0.593으로 나타났다. 반면, Neural Network는 정확도 0.438과 F1 스코어 0.429로 가장 낮은 성능을 보였다. 각 모형의 정확도와 f1 score에 기반하여 1-2에서 랜덤포레스트(random forest), 1-3에서 배깅(bagging), 1-4에서 로지스틱 회귀모형의 세부 성능(classification report)과 특성중요도(feature importance)를 다음 <표 5~7>과 같이 도출하였다.

〈표 5〉 Random Forest 분석결과(목표변수: 정부규제-공동규제)

Classification Report for 1-2 - Random Forest				
	precision	recall	f1-score	support
1	0.25	0.104167	0.147059	48
2	0.771277	0.90625	0.833333	160
accuracy	0.721154	0.721154	0.721154	0.721154
macro avg	0.510638	0.505208	0.490196	208
weighted avg	0.650982	0.721154	0.674962	208

F1 Score for 1-2 - Random Forest:
0.6749622926093515

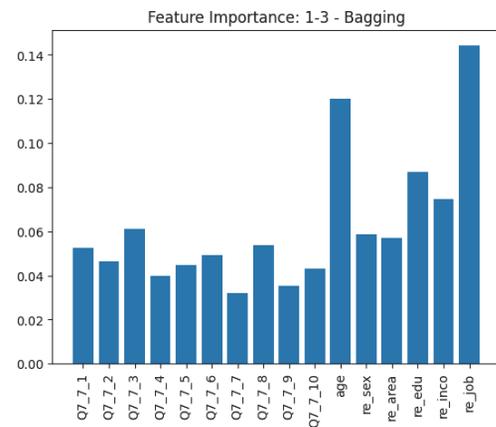


목표변수가 정부규제(1)-공동규제(2) 모형에서 가장 성능이 좋았던 Random Forest 모형은 공동규제(Class 2)를 예측함에 있어 Precision(정밀도)과 Recall(재현율) 모두 높은 값을 기록하였다 (Precision: 0.771277, Recall: 0.90625). 이 결과는 Random Forest 모델로 종속변수의 클래스 2를 정확하게 예측하고 있음을 보여준다. 반면, 정부규제(Class 1)를 예측한 경우, 매우 낮은 Precision과 Recall을 보였다. Precision은 0.25로, 모델이 클래스 1로 예측한 샘플 중 실제로 클래스 1인 샘플의 비율이 낮다는 것을 의미한다. Recall은 0.104167로, 실제 클래스 1인 샘플 중 모델이 올바르게 클래스 1로 예측한 비율이 매우 낮음을 나타낸다. 이는 모형이 클래스 1 샘플을 정확하게 예측하는 데 어려움이 있었다는 점을 확인할 수 있다. 모델의 전체적인 정확도(Accuracy)는 72.1%로 비교적 높은 편이나 클래스 1에서의 성능 저하는 전체 F1 스코어에 부정적인 영향을 미쳤다. 해당 모형의 특성중요도를 파악하면, 사회윤리적 위험인식 특성에서 기술 악용, 일자리 감소와 기술적 위험인식 특성에서 시스템 통제 실패가 클래스 예측에 영향을 크게 미쳤다. 나이, 소득 수준, 직업 등이 인구사회학적 변수에서는 높은 영향이 있었다.

〈표 6〉 bagging 분석결과(목표변수: 정부규제-독립기관규제)

Classification Report for 1-3 - Bagging				
	precision	recall	f1-score	support
1	0.509434	0.5625	0.534653	48
3	0.5	0.446809	0.47191	47
accuracy	0.505263	0.505263	0.505263	0.505263
macro avg	0.504717	0.504654	0.503282	95
weighted avg	0.504767	0.505263	0.503612	95

F1 Score for 1-3 - Bagging:
0.5036120170266583

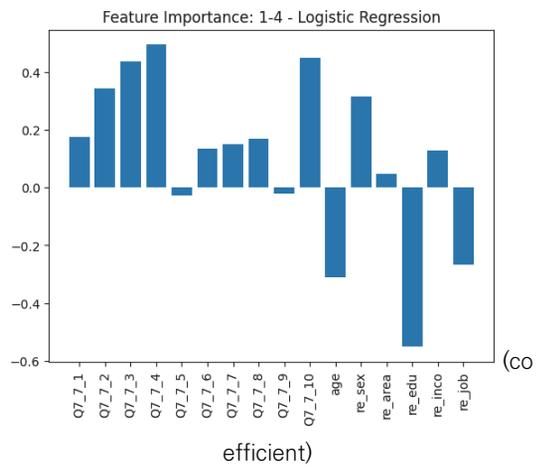


다음으로, 목표변수가 정부규제-독립기관규제 모형에서 Bagging 모형의 성능을 분석한 결과는 다음과 같다. Bagging 모형은 Precision(정밀도)과 Recall(재현율) 모두에서 비교적 일관된 성능을 보였다. 정부규제(Class 1)에서의 Precision은 0.509434로, 이는 모형이 클래스 1으로 예측한 샘플 중 실제로 클래스 1인 샘플의 비율이 약 51%임을 의미한다. Recall은 0.562500으로, 이는 실제 클래스 1인 샘플 중 약 56%를 모형이 올바르게 클래스 1으로 예측했음을 나타낸다. 독립기관 규제(Class 3)에서는 Precision이 0.500000로, 모형이 클래스 3으로 예측한 샘플 중 절반이 실제 클래스 3인 샘플이었다. Recall은 0.446809로, 실제 클래스 3인 샘플 중 약 45%를 모형이 정확히 예측했다. 모형의 전체적인 정확도(Accuracy)는 50.5%로, 이는 데이터셋의 샘플을 절반 정도 정확하게 예측했음을 의미한다. F1 스코어는 전체적으로 0.503612로, 모형의 성능을 나타내므로 전체적인 예측 성능은 매우 높지는 않았지만, 해당 목표변수를 활용한 모형에서는 가장 성능이 좋았기 때문에 해당 모형의 특성중요도를 도출하였다. 여기서 가장 높은 특성 중요도를 보인 특성은 사회윤리적 위험인식에서 일자리 감소와 보안사고 등이었으며, 기술적 위험인식에서는 부정확한 결과 등이 영향을 미치고, 본 모형에서도 인구사회학적 특성에서 직업과 나이, 교육 수준이 영향을 미쳤다.

〈표 7〉 Logistic Regression 분석결과(목표변수: 정부규제-자율규제)

Classification Report for 1-4 - Logistic Regression				
	precision	recall	f1-score	support
1	0.630435	0.604167	0.617021	48
4	0.55814	0.585366	0.571429	41
accuracy	0.595506	0.595506	0.595506	0.595506
macro avg	0.594287	0.594766	0.594225	89
weighted avg	0.59713	0.595506	0.596018	89

F1 Score for 1-4 - Logistic Regression:
0.596017895563676



마지막으로, 정부규제-자율규제 모형에서 로지스틱 회귀(Logit) 모형의 성능을 분석한 결과 목표 변수를 예측하는 데 있어서 비교적 고른 성능을 보였다. 정부규제(Class 1)의 Precision은 0.630435로, 이는 모델이 클래스 1로 예측한 샘플 중 실제로 클래스 1인 샘플의 비율이 높음을 의미하며, 클래스 1의 Recall은 0.604167로, 실제 클래스 1인 샘플 중 약 60%를 정확하게 예측하였다. 반면, 자율규제(Class 4)의 Precision은 0.55814, Recall은 0.585366로 클래스 1보다 약간 낮은 성능을 보였다. 모델의 전체적인 정확도는 59.55%로, 이는 데이터셋의 샘플을 예측하는 데 있어서 중간 수준의 성능을 나타내고, 전체 F1 스코어는 0.596018로, 모델이 전반적으로 균형 잡힌 성능을 보였음을 시사한다. 그러나, 본 모형 또한 전반적 성능은 객관적으로 높지 않은 편이었다. 다만, 해당 모형이 정부규제-자율규제 모형에서 가장 높은 성능을 보였으므로, 해당 모형에서 특성중요도(coefficient)를 도출하였다. 여기서 클래스 1에서 4로 변화하는데 가장 영향을 많이 미친 변수는 사회윤리적 위험인식 특성에서 개인정보 침해 문제였으며, 기술적 위험인식 특성에서 불투명성 또한 영향이 높게 측정되었다. 이외 교육수준, 나이, 성별, 직업 등이 인구사회학적 특성에서는 영향을 높게 미침으로 파악되었다.

정리하면, 정부규제-공동규제 모형에서 모형의 클래스 예측에 기술 악용, 일자리 감소, 개인정보 침해와 기술적 위험인식 특성에서 시스템 통제 실패, 나이, 소득 수준, 직업이 영향을 미치는 것

으로 나타났으며, 정부규제-독립기관규제 모형에서 일자리 감소, 보안사고, 부정확한 결과, 직업, 나이, 교육 수준이, 정부규제-자율규제 모형에서 개인정보 침해 문제, 불투명성, 교육수준, 나이, 성별, 직업이 영향을 미친다고 보인다. 실제 목표값(target value)의 클래스에 따라 차이를 검정하고, 규제정책적인 함의점을 이끌어 내기 위하여, 영향이 있는 사회윤리적 위험인식 변수와 기술적 위험인식 변수 별 t-test를 진행하여 보고한 결과는 다음 <표 8>과 같다.

〈표 8〉 목표값 클래스 기반 t-test

규제 유형 목표값	t-test 결과
1-2	기술 악용: t-statistic = -3.2719, p-value = 0.0012** 평균비교: 평균 (정부규제): 0.3438, 평균 (공동규제): 0.4859
	일자리 감소: t-statistic = -1.9737, p-value = 0.0494* 평균비교: 평균 (정부규제): 0.4000, 평균 (공동규제): 0.4878
	개인정보침해: t-statistic = -0.4639, p-value = 0.6431 시스템 통제 실패: t-statistic = 0.2160, p-value = 0.8291
1-3	일자리 감소: t-statistic = 0.4688, p-value = 0.6395 보안사고: t-statistic = 0.6576, p-value = 0.5113 부정확한 결과: t-statistic = -0.1556, p-value = 0.8764
	개인정보 침해: t-statistic = -1.7095, p-value = 0.0885 [†] 평균비교: 평균 (정부규제): 0.2812, 평균 (자율규제): 0.3750
	불투명성: t-statistic = -0.5644, p-value = 0.5730

*p<.05, **p<.01, ***p<.001, [†]p<.10

상기 결과에 따라, 기술 악용에 대한 인식은 대체로 정부규제에서 낮으며, 공동규제일 때 유의미하게 증가함을 확인할 수 있었다. 이는 공동규제 체계 하에서 기술 악용에 대한 경각심이 더 높아질 수 있음을 시사한다. 일자리 감소에 대한 우려도 정부규제가 낮고, 공동규제에서 높은 것으로 나타났다. 정부규제-공동규제의 t-test 결과는 공동규제가 규제 논의에 민간 부문이 공식적으로 참여하고, 이들에 책임을 부여하지만 정부의 직접적인 규제가 가해질 경우보다는 기술악용과 일자리 감소에 대한 우려가 높음을 시사한다.

이와 함께, 개인정보 침해에 대한 우려는 정부규제에서 자율규제로 전환될 때 유의미하게 증가함을 확인할 수 있었다. 이는 정부규제를 할 때보다 자율규제가 도입됐을 시 개인정보 침해 우려가 더 높을 수 있음을 의미한다. 연구결과에 따른 가설 검증 결과는 <표 9>와 같다.

〈표 9〉 가설 검증 결과

가설		검증결과
가설 1	AI 사회윤리적 위험인식 → 규제 유형 선호	일부 채택
1-1	일자리 감소 위험인식 → 규제 유형 선호	정부규제-공동규제 정부규제-독립기관규제
1-2	숙련도 저하 위험인식 → 규제 유형 선호	기각
1-3	개인정보 침해 위험인식 → 규제 유형 선호	정부규제-공동규제 정부규제-자율규제
1-4	인권침해 위험인식 → 규제 유형 선호	기각
1-5	편향성 위험인식 → 규제 유형 선호	기각
1-6	기술 악용 위험인식 → 규제 유형 선호	정부규제-공동규제
가설 2	AI 기술적 위험인식 → 규제 유형 선호	채택
2-1	보안 사고 위험인식 → 규제 유형 선호	정부규제-독립기관규제
2-2	시스템 통제 실패 위험인식 → 규제 유형 선호	정부규제-공동규제
2-3	부정확한 결과 위험인식 → 규제 유형 선호	정부규제-독립기관규제
2-4	불투명성 위험인식 → 규제 유형 선호	정부규제-자율규제

V. 결론 및 제언

전 세계 국가들은 AI를 전략적 목표로 선언하고, 기업은 거대한 성장 동력으로 전망하고 있으며, 많은 사람들이 AI로 파생된 상품과 서비스의 영향을 받고 있다(Holzinger, 2021). AI의 성능이 향상될수록 일자리 대체, 개인정보 침해, 편향이 담긴 결과도출 등 다양한 문제점들이 양산되고 있지만, AI 규제는 초국가적 성격, 복잡성, 분산성 등으로 인해 선불리 도입하기 어려운 특성이 있다. 이에 정부는 다양한 규제수단을 검토하여 어느 수준의 규제와 개입이 필요한지 신중하게 판단할 필요가 있다(Smuha, 2021).

이에 본 연구는 AI로 인한 사회윤리적 및 기술적 위험인식이 규제 유형 선호에 미치는 영향을 탐색적으로 살펴보기 위한 목적으로 분석을 수행하였다. 본 연구의 분석결과를 정리하면 다음 〈표 10〉와 같다.

〈표 10〉 분석결과 종합

모형	모형의 클래스 예측에 영향을 미치는 주요 변수
정부규제- 공동규제 모형	사회윤리적 위험인식: 기술 악용, 일자리 감소, 개인정보침해 기술적 위험인식: 시스템 통제 실패 인구사회학적 변수: 나이, 소득 수준, 직업
정부규제- 독립기관규제 모형	사회윤리적 위험인식: 일자리 감소, 보안사고 기술적 위험인식: 부정확한 결과 인구사회학적 변수: 직업, 나이, 교육 수준
정부규제- 자율규제 모형	사회윤리적 위험인식: 개인정보 침해 기술적 위험인식: 불투명성 인구사회학적 변수: 교육수준, 나이, 성별, 직업

분석 결과, 공동규제의 시행 시 정부규제보다 일자리 감소, 기술악용, 시스템 통제 실패 위험에 대한 위험인식이 높고, 정부규제에서는 이러한 위험인식이 낮은 것으로 나타났다. 이는 일자리 감소, 기술 악용의 사회윤리적 위험과 시스템 통제 실패의 기술적 위험인식이 있을 경우, 정부규제-공동규제 간 선호에 규제 전환이 나타날 가능성이 있음을 의미한다. 특히 t-test 결과 정부규제보다 공동규제에서 일자리 감소와 기술악용에 대한 위험인식이 높은 것으로 나타나, 공동규제 체계 하에서 이러한 문제들에 대한 경각심이 더 높아질 수 있음을 드러낸다.

둘째, 사회윤리적 위험인식에서 일자리 감소, 보안사고, 기술적 위험인식에서 부정확한 결과 인식은 정부규제에서 독립기관규제로의 변화 가능성을 내포하고 있는 것으로 나타났다. 하지만 통계적 분석(t-test) 결과 그 집단 간 유의미한 통계적 위험인식 차이는 나타나지 않았다.

셋째, 자율규제 시 개인정보 침해, 불투명성의 변화에 대한 위험인식이 정부규제의 경우보다 자율규제일 때 더 높은 것으로 드러났다. 특히 t-test 결과, 개인정보침해 위험인식의 평균이 자율규제보다 정부규제에서 더 낮은 것으로 나타났다.

연구 결과에 따른 함의는 다음과 같다. 먼저, 개인정보 침해에 대한 우려는 정부규제에서 자율규제로 전환될 때 유의미하게 증가함을 확인할 수 있었다. 이는 자율규제 하에서 개인정보 침해에 대한 위험이 더욱 높아질 수 있음을 의미하는바, 개인정보 보호와 관련된 정책은 민간 기업의 자율성에 기반하여 운영하기보다 안전한 보안 체계를 구축하도록 강제하고, 개인정보 유출시 강력하게 처벌하는 정부규제가 보다 적절할 수 있음을 함의한다.

둘째, 〈표 8〉의 t-test 결과에 따라 기술 악용, 일자리 감소와 같은 사회윤리적 위험인식이 없을 때 정부규제보다 공동규제를 선호하는 것은 AI와 같은 신기술 규제에 정부의 직접적인 개입보다는 협력에 기반한 유연한 접근을 원하는 것을 의미한다. 이는 동시에 해당 위험인식이 있을 경우에는 정부규제를 더욱 선호할 가능성이 높은 것을 뜻한다. 공동규제는 정부와 민간의 협력 체계에 기반한 비교적 유연한 규제으로써 기술 발전을 크게 저해하지 않으면서, 민간의 역할과 책임이 논의되고, 규정 준수 여부에 대한 상호 감시와 견제가 이뤄지는 측면에서 초기 규제정책 도입 시 검토해 볼 만하다. 하지만 본 연구결과, 위험이 없을 것이라는 가정하에 강력한 정부규제보다 공동규제를 선호하는 바 공동규제를 원하는 이해관계자들은 AI 적용에 따른 위험을 미리 점검하고, 이를

완화하는 방안을 마련할 필요가 있다.

셋째, 그럼에도 불구하고 정부는 공동규제, 자율규제 등의 대안적 규제가 결과적으로 사회질서를 무너뜨리고, 사회문제를 지속적으로 양산할 경우 정부규제를 도입을 검토해야 하며, 상시적으로 AI의 위협으로부터 사람들을 보호하는 조치를 채택해야 한다(Smuha, 2021). Floridi et al.(2021)는 AI로 인해 예상하지 못한 위험이 닥칠 수 있고, AI가 예측하지 않는 부분이 있을 수 있기 때문에 정부는 이러한 상황들을 미리 파악하고 대안적인 접근 방법을 고려해야 한다고 강조한다. 즉, AI로 인해 피해를 입은 사람들이 지원을 받을 수 있는 피해 완화 조치가 필요하며, 피해의 원인에 대한 책임소재를 가릴 수 있는 규정이 마련되는 등 피해 완화 매커니즘이 도입되어야 한다(Prainsack & Forgó, 2024).

이밖에 각 사회윤리적 및 기술적 위험인식에 따라 선호되는 규제유형이 다른바, AI가 일으킬 수 있는 실질적 위험을 파악하여 관련한 규제를 시행하거나, 규제 샌드박스의 운영에도 신중을 기할 필요가 있다. 이와 함께, 인간 중심적이고 사회 전체에 이익이 되는 것을 목표로 하는 윤리적인 인공지능의 개발이 필요하다(Kieslich et al., 2022).

본 연구에서는 인구사회학적 변수들의 규제 선호 영향성을 머신러닝 결과로 확인하였으나, 연구의 목적상 이에 대한 세부적인 통계적 검증(t-test, ANOVA, 조절효과, 매개효과 검증 등)을 확인하지 않았다. 다만 이러한 변수들이 규제의 선호에 영향을 미칠 수 있는 가능성이 있으므로 이에 대한 후속연구가 필요하다.

또한 본 연구는 설문조사 자료의 한계가 존재한다. 본 조사가 활용한 사회적가치연구원의 '한국인이 바라본 사회문제(2023)' 원자료를 바탕으로는 조사대상자들이 규제 유형에 대한 명확한 이해에 기반하여 응답을 할 수 있었는지는 확인할 수 없으며, 표본이 천명에 불과하여 빅데이터 분석 방법에 적용되는 머신러닝 방법이 적절하지 않다는 의견이 있을 수 있다. 하지만 본 연구는 규제 유형에 따라 유의미한 차이를 보이는 인식 변수들이 존재함을 확인하였으며, SMOTE 등의 방법으로 클래스간 불균형을 해결함과 동시에 목표변수의 예측에 영향을 미치는 변수를 찾기 위한 탐색적인 연구를 목적으로 하였으므로 연구의 의의가 있다.

본 연구는 AI의 사회윤리적, 기술적 위험인식에 따라 정부규제, 공동규제, 독립기관규제, 자율규제가 필요한 경우를 각기 다르게 인식할 수 있는바, 사회 전반에 영향을 미칠 수 있는 AI의 다양하고 실질적인 위협을 미리 파악하고, 피해의 원인과 규모, 해결 방안을 파악함과 동시에 규제기관들의 특성을 고려한 적절한 규제의 도입이 필요함을 시사한다.

참고문헌

- 고학수 · 박도현 · 이나래. (2020). 인공지능 윤리규범과 규제 거버넌스의 현황과 과제. *경제규제와 법*, 13(1), 7-36.
- 과학기술정보통신부. (2023). 디지털 대항해 시대, 글로벌 모범국가를 향한 대한민국 초거대 인공지능 도약 추진. 보도자료.
- 과학기술정보통신부. (2024). 인공지능 G3 국가 도약 위해 민간 최고전문가와 정부가 '원팀' 구성한다!. 보도자료.
- 김경동. (2021). 4 차 산업혁명과 윤리규범을 위한 AI 알고리즘 규제연구 국가행정과 지방자치체를 위한 지능정보화의 방향. *한국지방자치연구*, 23(2), 23-48.
- _____. (2023). 새로운 노멸, 인공지능 (AI) 과 윤리규범에 관한 소고. *법학연구*, 90, 25-50.
- 김광수. (2021). 인공지능 규제의 법체계. *토지공법연구*, 93, 217-243.
- 김나래. (2022). 인공지능 규제에 관한 소고. *동아법학*, 97, 213-252.
- 김나리 · 오서은. (2024). 옹호연합모형 (ACF) 을 통한 온라인 플랫폼 규제 정책변동과정 분석-플랫폼 자율규제를 중심으로. *지방정부연구*, 27(4), 245-276.
- 김법연. (2023). 인공지능 통제수단으로서 주요국 규제 입법의 동향과 시사점. *유럽헌법연구*, 42, 253-307.
- 김성준 · 홍승현. (2023). 인공지능 규제. *공공선택학저널*, 2(2), 35-57.
- 김승현 · 김시원 · 안정민. (2022). AI 추천 알고리즘 편향성과 규제에 관한 연구. *정보통신정책연구*, 29(2), 111-144.
- 김윤명. (2023). 알고리즘 권력화와 규제 거버넌스. *인권과 정의*, 515, 103-128.
- 김정화 · 임동민 · 차호동. (2023). 생성형 인공지능 (Generative AI) 기술의 규제 방향에 대한 입법론적 고찰-ChatGPT 등 인공지능 시스템 생성물에 대한 표시 · 고지의무를 중심으로. *형사법의 신동향*, (80), 245-283.
- 김종석 · 김태운. (2004). 독립규제위원회의 발전방향. *한국경제연구원 연구보고서*, 2004(20), 1-195.
- 노형진. (2016). SPSS를 활용한 회귀분석. *경기:지필미디어*.
- 박현석. (2023). 인공지능의 군사적 이용에 대한 국제적 규제-국제규범에 관한 논의의 최근 동향. *홍익법학*, 24(3), 27-59.
- 박혜성 · 김법연 · 권현영. (2021). 인공지능 규제에 대한 연구-유럽연합의 입법안을 중심으로. *공법연구*, 49(3), 349-374.
- 박혜진. (2022). 의료 인공지능의 활용을 둘러싼 법적 과제 규제의 진화 및 책임의 배분을 중심으로. *비교사법*, 29(4), 217-251.
- 방정미. (2021). 인공지능 알고리즘 규제거버넌스의 전환-최근 미국의 알고리즘 규제와 인공지능 윤리원칙을 중심으로. *공법연구*, 49(3), 375-406.
- 백수원. (2021). 헌법상 투명성원칙에 기반한 인공지능 규제의 방향. *성균관법학*, 33(2), 589-622.
- 선지원. (2022). 인공지능 생애주기의 관점에서 본 규제와 거버넌스. *연세법학*, 39, 1-31.

- 심우현 · 원소연 · 이종한. (2021). 언택트 디지털경제의 규제 개선방안. *기본연구과제*, 2021, 1-326.
- 양은영. (2023). 생성형 AI 의 개발 및 이용에 관한 규제의 필요성-대규모 언어모델에 기반한 대화형 인공지능 서비스 (LLMs AI) 를 중심으로. *성균관법학*, 35(2), 293-325.
- 양종모. (2016). 인공지능의 위험의 특성과 법적 규제방안. *홍익법학*, 17(4), 537-565.
- 유지연 · 정나영. (2024). 인공지능(AI) 윤리원칙을 중심으로 한 규제 내용 분석·비교. *IT와 법연구*, 28, 137-210.
- 유한별. (2020). 규제 정책 도입 과정에서 민간 이해관계자의 행태에 관한 연구: 규제혁신형 플랫폼 택시 도입에서 업계 간 갈등을 중심으로. *규제연구*, 29(2), 155-195.
- 이길원. (2023). AI 규제를 위한 국제협력 방안-국제기구의 역할과 기능에 관한 국제법적 검토를 중심으로. *미국헌법연구*, 34(3), 75-103.
- 이성엽. (2023). 생성형 인공지능 (AI) 규제원칙에 관한 연구. *고려법학*, 110, 149-180.
- 이종한 · 원소연 · 심우현 · 임현철 · 상민정. (2022). 규제기관 거버넌스 개선방안. *기본연구과제*, 2022, 1-355.
- 이준복. (2023). 공공안전 측면에서 인공지능 (AI) 의 활성화와 규제의 실질적 조화를 위한 법적 고찰 유럽연합 및 우리나라의 입법례를 중심으로. *한국테러학회보*, 16(2), 121-140.
- 이혁우. (2009). 규제의 개념에 관한 소고. *행정논총*, 47(3), 335-358.
- 이효진. (2023). 인공지능 (AI) 단일법 제정 필요성과 행정법학의 과제. *법학논총*, 35(3), 9-57.
- 임예준. (2019). 인공지능 시대의 전쟁자동화와 인권에 관한 소고-국제법상 자율살상무기의 규제를 중심으로. *고려법학*, 92, 265-302.
- 전영균 · 김현경. (2020). AI 스피커 음성정보의 합리적 규제 방안에 대한 연구. *미국헌법연구*, 31(1), 161-198.
- 정남철. (2023). 행정의 디지털 전환에 따른 인공지능 규제의 최근 동향과 시사점-인공지능에 대한 인격권 보호의 문제를 겸하여. *미디어와 인격권*, 9(2), 1-36.
- 채구묵. (2014). SPSS와 AMOS를 이용한 고급통계학. *경기:양서원*.
- 최지수 · 윤석민. (2019). 가짜뉴스 거버넌스 정부규제, 자율규제, 공동규제 모형에 대한 비교를 중심으로. *사이버커뮤니케이션학보*, 36(1), 127-180.
- Alalawi, Z., Bova, P., Cimpeanu, T., Di Stefano, A., Duong, M. H., Domingos, E. F., ... & Zimmaro, F. (2024). Trust AI Regulation? Discerning users are vital to build trust and effective AI regulation. *arXiv preprint arXiv:2403.09510*, 1-18.
- Bartle, I., & Vass, P. (2007). Self-regulation within the regulatory state Towards a new regulatory paradigm. *Public Administration*, 85(4), 885-905.
- Benbear, L. S., & Coglianese, C. (2012). Flexible environmental regulation, 1-40.
- Black, J. (2001). Decentring regulation Understanding the role of regulation and self-regulation in a 'post-regulatory' world. *Current legal problems*, 54(1), 103-146.
- Breiman, L. (2001). Random forests. *Machine learning*, 45, 5-32.
- Breiman, L. (1996). Bagging predictors. *Machine learning*, 24, 123-140.

- Buiten, M. C. (2019). Towards intelligent regulation of artificial intelligence. *European Journal of Risk Regulation*, 10(1), 41-59.
- Castro, D. (2011). Benefits and limitations of industry self-regulation for online behavioral advertising. *The Information Technology & Innovation Foundation*, 1(1), 1-14.
- Cerf, M., & Waytz, A. (2023). If you worry about humanity, you should be more scared of humans than of AI. *Bulletin of the Atomic Scientists*, 79(5), 289-292.
- Çetin, T., Sobacı, M. Z., & Nargeleçekenler, M. (2016). Independence and accountability of independent regulatory agencies the case of Turkey. *European Journal of Law and Economics*, 41, 601-620.
- Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794).
- Christensen, T., & Lægreid, P. (2007). Regulatory agencies—The challenges of balancing agency autonomy and political control. *Governance*, 20(3), 499-520.
- Clarke, R. (2019). Regulatory alternatives for AI. *Computer Law & Security Review*, 35(4), 398-409.
- Cohen, M., & Sundararajan, A. (2015). Self-regulation and innovation in the peer-to-peer sharing economy. *U. Chi. L. Rev. Dialogue*, 82, 116.
- de Almeida, P. G. R., dos Santos, C. D., & Farias, J. S. (2021). Artificial intelligence regulation a framework for governance. *Ethics and Information Technology*, 23(3), 505-525.
- Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE transactions on information theory*, 13(1), 21-27.
- Díaz-Rodríguez, N., Del Ser, J., Coeckelbergh, M., de Prado, M. L., Herrera-Viedma, E., & Herrera, F. (2023). Connecting the dots in trustworthy Artificial Intelligence: From AI principles, ethics, and key requirements to responsible AI systems and regulation. *Information Fusion*, 99, 101896, 1-24.
- Eijlander, P. (2005). Possibilities and constraints in the use of self-regulation and co-regulation in legislative policy: Experiences in the Netherlands-Lessons to be learned for the EU?. *European Journal of Comparative Law*, 9(1), 1-8.
- Erdélyi, O. J., & Goldsmith, J. (2018). Regulating artificial intelligence Proposal for a global solution. In *Proceedings of the 2018 AAAIACM Conference on AI, Ethics, and Society*. 95-101.
- EU Commission. (2024). AI Act (<https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>).
- Fast, E., & Horvitz, E. (2017). Long-term trends in the public perception of artificial intelligence. In *Proceedings of the AAAI conference on artificial intelligence*, 31(1), 963-969.

- Finck, M. (2018). Digital co-regulation designing a supranational legal framework for the platform economy. *European law review*, 1-29.
- Floridi, L., Cowls, J., King, T. C., & Taddeo, M. (2021). How to design AI for social good Seven essential factors. *Ethics, Governance, and Policies in Artificial Intelligence*, 125-151.
- Freund, Y., & Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of computer and system sciences*, 55(1), 119-139.
- Gerlich, M. (2023). Perceptions and acceptance of artificial intelligence A multi-dimensional study. *Social Sciences*, 12(9), 502, 1-24.
- Gilardi, F. (2005). The institutional foundations of regulatory capitalism: the diffusion of independent regulatory agencies in Western Europe. *The Annals of the American Academy of Political and Social Science*, 598(1), 84-101.
- Gillath, O., Ai, T., Branicky, M. S., Keshmiri, S., Davison, R. B., & Spaulding, R. (2021). Attachment and trust in artificial intelligence. *Computers in Human Behavior*, 115, 106607, 1-10.
- Gunningham, N., & Rees, J. (1997). Industry self-regulation an institutional perspective. *Law & policy*, 19(4), 363-414.
- Hacker, P., Engel, A., & Mauer, M. (2023). Regulating ChatGPT and other large generative AI models. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*. 1112-1123.
- Hanretty, C., & Koop, C. (2012). Measuring the formal independence of regulatory agencies. *Journal of European Public Policy*, 19(2), 198-216.
- Hanretty, C., & Koop, C. (2013). Shall the law set them free The formal and actual independence of regulatory agencies. *Regulation & Governance*, 7(2), 195-214.
- Hassani, H., Silva, E. S., Unger, S., TajMazinani, M., & Mac Feely, S. (2020). Artificial intelligence (AI) or intelligence augmentation (IA) what is the future. *Ai*, 1(2), 8, 143-155.
- Hawkins, K., & Hutter, B. M. (1993). The response of business to social regulation in England and Wales an enforcement perspective. *Law & Policy*, 15(3), 199-217.
- Holzinger, A. (2021). The next frontier AI we can really trust. In *Joint European conference on machine learning and knowledge discovery in databases*. 427-440.
- Kelley, P. G., Yang, Y., Heldreth, C., Moessner, C., Sedley, A., Kramm, A., ... & Woodruff, A. (2021). Exciting, useful, worrying, futuristic: Public perception of artificial intelligence in 8 countries. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, 627-637.
- Kieslich, K., Keller, B., & Starke, C. (2022). Artificial intelligence ethics by design. Evaluating public perception on the importance of ethical design principles of artificial intelli-

- gence. *Big Data & Society*, 9(1), 1-15.
- Kjaer, P. F., & Vetterlein, A. (2018). Regulatory governance rules, resistance and responsibility. *Contemporary Politics*, 24(5), 497-506.
- Lockey, S., Gillespie, N., Holm, D., & Someh, I. A. (2021). A review of trust in artificial intelligence Challenges, vulnerabilities and future directions. *Proceedings of the 54th Hawaii International Conference on System Sciences*, 5463-5472.
- Machleidt, P., Mráčková, J., & Mráček, K. (2024). Perception of the risks inherent in new AI technologies. *TATuP-Zeitschrift für Technikfolgenabschätzung in Theorie und Praxis*, 33(2), 42-48.
- Nunes, R., Nunes, S. B., & Rego, G. (2015). A new governance model for independent regulatory agencies. *Theoretical Economics Letters*, 5(01), 4, 4-13.
- Ogus, A. (1995). Rethinking self-regulation. *Oxford J. Legal Stud.*, 15, 97-108.
- Papyshev, G. (2024). Governing AI through interaction situated actions as an informal mechanism for AI regulation. *AI and Ethics*, 1-12.
- Prainsack, B., & Forgó, N. (2024). New AI regulation in the EU seeks to reduce risk without assessing public benefit. *Nature Medicine*, 1-3.
- Quinlan, J. R. (1986). Induction of decision trees. *Machine learning*, 1, 81-106.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, 521(7553), 436-444.
- Rossi, F. (2018). Building trust in artificial intelligence. *Journal of international affairs*, 72(1), 127-134.
- Rubinstein, I. (2016). The future of self-regulation is co-regulation. *The Cambridge Handbook of Consumer Privacy, From Cambridge University Press*, 1-23.
- Ryan, M. (2020). In AI we trust ethics, artificial intelligence, and reliability. *Science and Engineering Ethics*, 26(5), 2749-2767.
- Saldaña, M. M., & Rosón, V. G. I. (2015). The Importance of Self-regulation and Co-regulation in the New Digital Audiovisual Market. *Observatorio (OBS)*, 9(1), 101-114.
- Seth, J. (2024). Public Perception of AI Sentiment and Opportunity. *arXiv preprint arXiv:2407.15998*, 1-65.
- Skjuve, M., Følstad, A., & Brandtzaeg, P. B. (2023). The user experience of ChatGPT: Findings from a questionnaire study of early users. In *Proceedings of the 5th International Conference on Conversational User Interfaces*, 1-10.
- Smuha, N. A. (2021). From a 'race to AI' to a 'race to AI regulation' regulatory competition for artificial intelligence. *Law, Innovation and Technology*, 13(1), 57-84.
- Stone, P., Brooks, R., Brynjolfsson, E., Calo, R., Etzioni, O., Hager, G., ... & Teller, A. (2022). Artificial intelligence and life in 2030: the one hundred year study on artificial intelligence. *arXiv preprint arXiv:2211.06318*, 1-52.
- Thatcher, M. (2002). Regulation after delegation independent regulatory agencies in Europe.

- Journal of European public policy, 9(6), 954-972.
- Thatcher, M., & Sweet, A. S. (2002). Theory and practice of delegation to non-majoritarian institutions. *West European Politics*, 25(1), 1-22.
- Thatcher, M. (2005). The third force Independent regulatory agencies and elected politicians in Europe. *Governance*, 18(3), 347-373.
- The White House. (2023). Blueprint for an AI Bill of Rights (<https://www.whitehouse.gov/ostp/ai-bill-of-rights/>).
- Wirtz, B. W., Weyerer, J. C., & Geyer, C. (2019). Artificial intelligence and the public sector—applications and challenges. *International Journal of Public Administration*, 42(7), 596-615.

김나리: 런던정경대(LSE)에서 Management of Information Systems and Digital Innovation 석사과정을 졸업하고, 연세대학교 행정학과 박사과정을 수료하였다. 주요 연구분야는 디지털 보건·복지·사회정책, 데이터 정책, 전자정부 등이다. 최근 연구로는 “공공부조 집행 과정에서 자동화된 의사결정의 역할과 한계(2023)”, “우호연합모형(ACF)을 통한 온라인 플랫폼 규제 정책변동과정 분석(2024)”, “보건의료 개인정보 유출 및 민영화 우려가 데이터 공유 인식에 미치는 영향 연구(2024)” 등이 있다(kimnari2014@gmail.com).

한별: 연세대학교에서 행정학 박사학위를 취득하고, 현재 선문대학교 행정·공기업학과 조교수로 재직하고 있다. 주요 관심 분야는 공공관리, 환경관리, 규제, AI, Bigdata, Machine Learning 등이며, 최근 논문으로는 “재난 및 위급상황 발생시 장애인의 재난대응 수행에 따른 잠재계층 전이 양상 및 영향요인 검증”(2023), “머신러닝을 활용한 지방정부 재정건전성 영향요인 탐색”(2023) 등이 있다(hanbyeolyoo@sunmoon.ac.kr).

Abstract

An exploratory study on the impact of AI risk perception on regulation type preference: Exploration through machine learning methods

Nari Kim
Hanbyeol Yoo

As AI-based technologies become more sophisticated, concerns about the side effects that may arise from AI are increasing. Problems caused by AI can lead to direct damage to individuals and society, leading to strong demands for regulation. However, risks posed by AI cannot be easily controlled through government regulation, alternative regulatories such as co-regulation, independent agency regulation, and self-regulation are being discussed, but related empirical research has been limited. The purpose of this study is to examine the relationship between AI risk perception(social and ethical risks and technical risks) and preference for regulation types(government regulation, co-regulation, independent agency regulation and self-regulation).

As a result of the analysis, it was found that in the government and co-regulation model, there can be a shift from government regulation to preference for co-regulation due to risk awareness of technology abuse, job loss, personal information infringement, and system control failure.

Second, in the government and independent agency regulation model, there can be a shift from government regulation to preference for independent agency regulation due to the perception of job loss, security incidents and inaccurate results. Lastly, in the government and self-regulation model, personal information infringement and opacity were found to influence the transition from government regulation to self-regulation.

Keywords: AI, risk perception, government regulation, co-regulation, independent agency regulation, self-regulation