

# 논술고사 문제지

(인문계열) : 120분

모집단위		전형유형	논술(논술우수자)
수험번호		성명	

## ■ 일반 유의사항

1. 시험시간은 120분, 배점은 [문항 1]이 60점, [문항 2]가 40점입니다.
2. 답안을 구상할 때 문제지의 여백이나 문제지 내의 연습장을 사용하십시오.
3. 답안을 작성할 때 반드시 흑색 필기구만 사용하십시오(연필, 샤프 사용 가능, 사인펜 불가).
4. 답안을 정정할 때 원고지 사용법에 따르십시오(수정 테이프, 지우개 사용 가능, 수정액 불가).
5. 답안은 반드시 해당 문항의 답란에 작성하고, 답란 밖에는 작성하지 마시오.

※ 답안지는 절대 교체할 수 없습니다.

## ■ 답안 작성 유의사항

1. 제목은 쓰지 마시오.
2. 제시된 분량을 지키시오.
3. 제시문의 문장을 그대로 옮기지 마시오.
4. 서론은 쓰지 말고 본론을 중심으로 작성하십시오.
5. 수험번호, 성명 등 신상에 관련된 사항을 답란이나 답안지의 여백에 드러내지 마시오.





# 논술고사(인문계열)

[문항 1] “인공지능의 판단이 인간의 판단보다 공정하다”라는 주장에 대한 지지와 반론의 논거를 제시문 (가)~(마)를 모두 활용하여 작성하시오. 또한 반론에 대해 새로운 논거를 제시하며 재반론하시오(단, 제시문 (가)는 지지와 반론에 모두 활용할 것).

(1,000자±100자, 60점)

(가) 인공지능은 인간의 학습·추론·인식을 계산적 절차로 재현해 정보를 처리하도록 만들어진 시스템이다. 이러한 시스템은 순수한 연산 구조를 바탕으로 작동하며, 주어진 목표에 맞춰 통계적 패턴을 식별하고 그 결과를 예측치로 제시한다. 이를 구현하는 주요 방법인 머신러닝은 데이터를 활용한 알고리즘을 통해 모델의 성능을 최적화하는 기법이다. 머신러닝 기반 인공지능은 인간이 처리하기 어려운 방대한 데이터를 분석할 수 있는 능력을 갖추고 있기에 다양한 상황에서 인간의 판단을 돕는 기술 장치로 폭넓게 활용되고 있다.

인공지능 개발의 추세를 고려하면 가까운 미래에는 인공지능이 인간의 판단을 보조하는 것을 넘어 대체할 만큼의 역량을 갖추어 나갈 수 있다. 인공지능이 독자적 판단 주체가 되기 위해서는 무엇보다 그 판단의 공정성을 확보하는 것이 중요하다. 왜냐하면 판단이 필요한 현실 세계가 희소성의 조건 아래 놓여 있기 때문이다. 사회·경제적 자원이 희소한 상황에서 인간의 이타성은 제한적이기에, 누가 무엇을 얼마나 가지게 되는지를 둘러싸고 첨예한 갈등이 발생한다. 이러한 갈등을 다루는 과정에서 중요한 것이 바로 공정성이다.

오늘날 우리 사회는 공정성 훼손으로 인한 불필요한 균열과 갈등이 고조되어 상당한 사회적 비용을 치르고 있다. 그렇다면 판단이 공정하기 위한 조건은 무엇인가? 공정하기 위해선 현실을 왜곡 없이 파악하고, 합당한 기준을 설정하며, 이 기준을 일관되게 적용해야 한다. 기준의 일관된 적용을 위해 필요한 첫 번째 조건은 이해 당사자의 배제이다. 기준을 바꾸면 자신에게 이익이 돌아오는 상황에서 판단 주체는 그 기준을 임의로 변경하려고 할 수 있다. 이와 같은 상황을 방지하기 위해서는 이를테면 토지 개발사업의 인허가 과정에 이해관계가 있는 공무원의 참여가 배제될 필요가 있다.

합당한 기준의 일관된 적용을 위한 두 번째 조건은 책임 부과를 제도화하는 것이다. 자의적으로 기준을 적용하면 그에 대한 책임을 져야 한다는 점은 판단 주체가 적합한 기준을 일관되게 적용하는 강력한 동기가 된다. 책임 부과 절차는 판단 주체에게 제재를 가하는 방식으로 작동하며, 제재의 본질은 주체에게 불이익을 주는 것이다. 벌금을 물리거나 교정시설에 가두는 것이 그 예이다. 따라서 합당하게 설정된 정당한 기준을 사적 이익을 위해 자의적으로 적용하는 경우, 그에 상응하는 불이익을 확실히 부과해야 한다. 이때 비로소 합당한 기준의 임의적 적용을 효과적으로 저지할 수 있다.

고등학교 『사회·문화』, 『생활과 윤리』, 『윤리와 사상』 활용

(나) 캐럴(Lewis Carroll)의 『이상한 나라의 앨리스』는 넌센스 문학 전통을 바탕으로 언어유희와 논리적 전복을 통해 인간 사고의 허술함을 풍자한 작품이다. 이러한 특성은 소설 마지막에 등장하는 재판 장면에서 잘 드러난다. 이상한 나라의 법정은 넌센스적 요소가 극적으로 펼쳐지는 공간이고, 재판 과정은 임의적 규칙, 즉흥적 결정, 불완전한 정보, 모순된 논리의 연속으로 그려진다. 하트 여왕이 만든 타르트가 사라진 사건이 발생하여, 카드 잭이 범인으로 지목되자 재판이 열리게 된다. 배심원과 증인들은 무질서하게 의견을 개진하고, 여왕은 재판이 마음에 들지 않을 때마다 끊임없이 “목을 쳐라!”라고 외친다. 증인으로 등장한 하얀 토끼는 엉뚱하게도 앨리스의 신체 크기가 자꾸 바뀐다는 이상한 증언을 한다. 재판장인 왕은 증거를 읽기도 전에 유죄를 선고하려고 하고, 여왕은 선고를 먼저 하고 평결은 나중에 하라고 주장한다. 재판 내용에 이의를 제기하는 앨리스에게도 여왕은 “목을 쳐라!”라고 명령하지만, 앨리스는 눈 한 번 깜빡하지 않고 “너희는 모두 한 팩의 카드일 뿐이야!”라며 이 재판이 넌센스의 세계임을 지적한다. 앨리스는 이상한 나라의 재판이 형식적 절차를 따르는 것처럼 보이지만, 사실은 무질서와 모순, 우연한 요소가 끼어드는 상황임을 알아챈 것이다.

인지심리학자이자 행동경제학자인 카너먼(Daniel Kahneman)은 판단과 의사결정에서의 불합리한 특성을 ‘노이즈(noise)’라는 개념으로 설명한다. 그는 전통적인 관점과 달리 인간이 제한된 합리성 속에서 판단을 내린다고 보며, 이때 생기는 판단의 원치 않는 변동성을 노이즈라고 불렀다. 같은 정보와 조건이 주어져도, 시간, 환경, 감정 상태, 개인적 편향 등에 따라 사람들의 판단은 얼마든지 달라질 수 있다. 나아가 동일한 판단자라 하더라도 시간, 기분, 환경에 따라 같은 사례조차 다르게 판단하는 임의적 변동성을 보인다. 특히 재판, 채용, 의료 진단과 같은 분야에서 발생하는 노이즈는 사회적 비용을 초래할 뿐 아니라 결과에 대한 예측 가능성과 신뢰성을 떨어뜨리는 문제를 낳는다. 노이즈를 줄이기 위해 표준화된 절차, 구조화된 규칙, 일률적 평가 체계 등이 사용되지만, 이러한 방법에는 한계가 있다. 지나치게 규칙화하면 상황에 따른 유연한 판단이나 정당한 재량이 발휘되지 못하고, 판단 기준 자체가 설계자의 주관이나 임의적 선택에 좌우될 수 있기 때문이다. 따라서 판단의 노이즈는 단순히 오류가 아니라, 인간 판단 자체가 갖는 본질적인 변동성과 한계를 드러내는 지표로 이해될 수 있다. 캐럴의 넌센스적 재판 장면은 이러한 판단의 취약성과 불확실성을 상징적으로 보여주는 사례라 할 수 있다.

고등학교 『통합사회』, 『경제』 활용

(다) 인간이 언어를 사용하는 것은 단순한 정보 전달을 넘어, 세계를 해석하고 타인과 관계를 형성하는 복합적인 활동이라고 할 수 있다. 전 세계에는 수천 개의 언어가 존재하는데, 동일한 언어 안에서도 지역·문화·세대·사회적 맥락에 따라 다양한 표현 방식이 나타난다. 이러한 다양성으로 인해 언어 발화를 둘러싼 상황을 정확하게 이해하지 못하면 그 말의 의미를 제대로 파악하기 어렵다. 같은 문장도 화자의 의도, 관계의 성격, 분위기, 발화의 맥락에 따라 전혀 다른 의미를 갖기 때문이다. 예를 들어, 직장에서 상사가 부하 직원에게 “오늘 회의 준비는 잘 되어 있나요?”라고 물었을 때, 이 발화는 발화자의 음성, 이전 대화의 흐름, 상사와 부하 간의 관계 등의 맥락 속에서 격려가 되기도, 반대로 압박이 되기도 한다. 언어학에서 화용론(Pragmatics)은 이처럼 문장 자체의 의미가 아니라 언어가 실제 사용되는 상황 속에서 어떻게 해석되고 기능하는지를 연구한다. 물론 맥락적 이해가 항상 정확하게 잘 구현되는 것은 아니다. 서로 다른 삶의 경험과 해석의 틀이 충돌할 때 상대방의 말을 이해하는 데 간극이 생기기도 한다. 오해는

의사소통의 단순한 실패가 아니며, 다양한 삶의 방식과 세계관이 맞닿는 지점에서 발생한다.

언어를 둘러싼 이러한 관점은 철학에서도 중요한 주제로 다뤄져 왔다. 비트겐슈타인(Ludwig Wittgenstein)은 언어의 의미가 단어에 고정된 것이 아니라 언어가 사용되는 실제 맥락 속에서 형성된다고 본다. 말의 의미는 사전에 적힌 정의가 아니라 공동체가 만들어낸 규칙과 관습 속에서 지속적으로 조정된다는 것이다. 또한 흄(David Hume)은 인간 본성의 여러 성질 가운데 가장 주목할 만한 것으로 공감을 들었는데, 그의 공감 개념은 다른 사람들에 대해 인류애나 이타심의 형태로 표현되는 감정을 말한다. 공감은 타인의 감정을 추론하여 자기 안에서 재현하는 능력에 달려 있으며, 이러한 재현 능력이 사회적 삶을 가능하게 한다고 설명한다. 이처럼 인간은 언어의 표면적 의미만을 해석하는 것이 아니라, 그 언어 안에 담긴 감정·동기·상황을 함께 읽어내고, 이를 통해 공감이라는 심리적 연결을 형성한다.

인간의 우수성은 이러한 맥락적 판단 능력을 언어 자체의 이해를 넘어 폭넓은 사고력으로 연결할 수 있다는 점이다. 우리는 말뿐 아니라 목소리의 떨림, 표정, 몸짓, 침묵과 같은 비언어적 신호까지 결합해 상대방의 상태를 추론한다. 여기에는 과거 경험, 사회적 규범, 관계의 역사 등이 함께 작용한다. 이러한 과정은 단순한 논리적 분석을 넘어서는 인간 고유의 통합적 인지 능력에서 나온다. 공감은 타인의 감정을 단순히 감지하는 것이 아니라, 그 사람이 처한 상황과 내면적 맥락을 폭넓게 파악할 수 있는 능력에 기반하며, 바로 이러한 깊고 넓은 이해 능력이 사회적 관계를 지속할 수 있는 중요한 토대가 된다.

고등학교 『국어』, 『화법과 작문』, 『윤리와 사상』 활용

**(라)** 일기예보는 뜻밖의 재난을 대비하고, 일상 및 산업 활동을 영위하는 데 필수적인 기상정보다. 기상청의 예측이 자주 빗나가지, 국내 기상정보를 노르웨이나 미국 등 해외 기상청 애플리케이션을 통해 찾아보는 ‘기상 망명자’도 생겨났다. 이는 일기예보가 우리 삶에 미치는 영향을 잘 보여준다. 한국 기상청은 2020년 ‘수치예보 모델’을 도입하여 일기예보의 부정확성을 줄여서 기후 위기에 전전공공하는 국민의 생활 안정성을 도모하고자 했다. 이 기상예보 장치는 미래의 대기 상태와 움직임을 계산한 값을 활용해 날씨를 예측하는 프로그램이다. 기상청은 이전에는 영국의 통합 모델을 빌려 썼으나 이 예보 장치의 도입을 통해 한국의 지형과 기상 특성을 정밀하게 반영한 기후 관측과 예보를 실행하고자 했다. 이에 따라 기상 예측과 예보 적중률이 얼마간 향상되었으나 해당 예보 장치의 제약성을 크게 피하지는 못했다. 왜냐하면 수치예보 모델은 대기의 운동, 온도, 기압, 습도 등을 물리학 방정식에 의존하여 계산하는 전통적인 기상예보 모델이기 때문이다. 기존의 방정식은 매우 복잡한 기후 현상과 변화 과정을 완전한 수식으로 표현할 수 없다. 또한 기존 기후 모델의 경우 계산해야 하는 관측 자료의 방대함과 시간의 급속한 증가 및 비용의 폭등 때문에 신속하고 정확한 예보를 제공하는 데 어려움을 겪는다. 이렇듯 기존의 기상예보 모델로는 갈수록 잦아지고 심각해지는 기후 위기를 예측하고 대비하는 데 역부족일 수밖에 없다.

이러한 한계를 극복하기 위해 국내외 여러 기관과 업체는 인공지능 기반의 기후예보 모델을 다수 개발했다. 한국의 초단기 강수 예측 모델인 ‘나우알파(NowAlpha)’나, 구글의 딥마인드(DeepMind)가 개발한 ‘그래프캐스트(GraphCast)’ 등이 대표적 사례다. 인공지능 기상예보 모델은 위성과 레이더, 지상관측 데이터를 모두 통합하여 계산하고 수치화하는 등 방대한 관측 자료를 매우 빠른 속도로 처리한다. 또한 과거에서 현재까지 축적된 대규모 데이터를 학습함으

로써 기존 모델이 간과한 신호나 미세한 패턴까지 수집, 활용하는 능력을 발휘한다. 인공지능 기상예보 모델은 관측 데이터가 부족한 지역에서도 비슷한 패턴을 기반으로 기후 현상과 변화를 객관적으로 추정할 수 있다. 그리고 예보된 값과 실제 값의 차이를 학습하여 예보 오차를 스스로 줄여 정확한 예측을 가능하게 한다.

고등학교 『통합사회』, 『한국지리』, 『생활과 윤리』 활용

(마) 나날이 진화 중인 인공지능 시스템은 자신의 창조자인 인간, 사회의 각종 기반 시설, 복잡한 외부 세계로부터 독립하여 온갖 지식과 정보를 스스로 학습하고 구조화할 줄 아는 비실체적 두뇌라는 판타지의 주인공으로 성장했다. 이를테면 우리는 아침에 출근해서 컴퓨터를 켜면 어제 검색했던 장소나 사물 이미지를 화면에서 다시 만난다. 이것은 인터넷상의 인공지능이 사용자의 여러 행동 정보, 예컨대 검색어, 머문 시간, 클릭한 기사나 영상을 바탕으로 그의 습관이나 욕망의 패턴과 흐름을 찾아내어 맞춤형 정보로 담아낸 것이다. 이러한 맞춤형 정보 제공은 인공지능이 누군가의 내면과 개성을 충분히 고려할 줄 아는 따스한 맥락의 소유자인 것처럼 보이게 한다.

인공지능은 사용자가 검색하는 단어나 이미지가 무엇인지는 알 수 있어도 그가 그것에 관심을 표하는 진정한 까닭은 알지 못한다. 그래서 인공지능은 데이터에 대한 통계적 패턴 인식을 바탕으로 체계화할 수 없는 것을 체계화하고, 사회적이며 역사적인 것을 규격화하며, 무한히 복잡하고 변화하는 세계를 기계가 읽을 수 있는 산술적·약호적 질서로 환원하는 장치로 취급되기도 한다. 이 상황은 인공지능이 대상물의 일부 특징을 식별하고 명명하면서 오히려 수많은 다른 특징을 무시하거나 얼버무려왔음을 암시한다. 그럴수록 특정 정보를 선택하고 제공하는 알고리즘도 의도적이든 비의도적이든 관련 데이터에 대한 분류나 해석에서 왜곡과 단절, 오류와 차별을 피할 수 없게 된다.

예컨대 미국의 한 연구진은 온라인상의 수많은 텍스트로 훈련을 받은 생성형 인공지능 언어 모델인 글로브(GloVe)가 편견을 재생산하는지를 탐구하기 위해 인간의 편견을 측정하는 ‘암묵적 연관 검사’ 결과와 비교하였다. 암묵적 연관 검사란 피검사자가 컴퓨터 모니터 앞에 앉아 어떤 특징이나 사회적인 집단과 연관성이 있는 단어와 이미지의 나열을 보면서 버튼을 눌러 모든 단어 및 이미지를 되도록 빠르게 범주화하는 검사다. 글로브는 아프리카계 미국인처럼 들리는 이름보다는 유럽계 미국인처럼 들리는 이름을 선호했으며, 여성의 이름은 직업보다 가족과의 연관성이 높다고 판단했다. 이는 인공지능이 다양한 문화공동체가 축적해 온 방대한 양의 지식을 사용하여 필연적으로 그것이 긍정적이든 부정적이든 인간의 편견을 어떤 성찰 활동이나 맥락에 대한 고려 없이 그대로 학습하거나 굴절시킨다는 것을 뜻한다. 요컨대 글로브 언어모델은 현실 세계를 직접 조사하거나 경험하지 않고도 인간의 역사적 또는 선택적 편견을 그대로 물려받아 판단 오류의 가능성을 높이는 경향을 보였다.

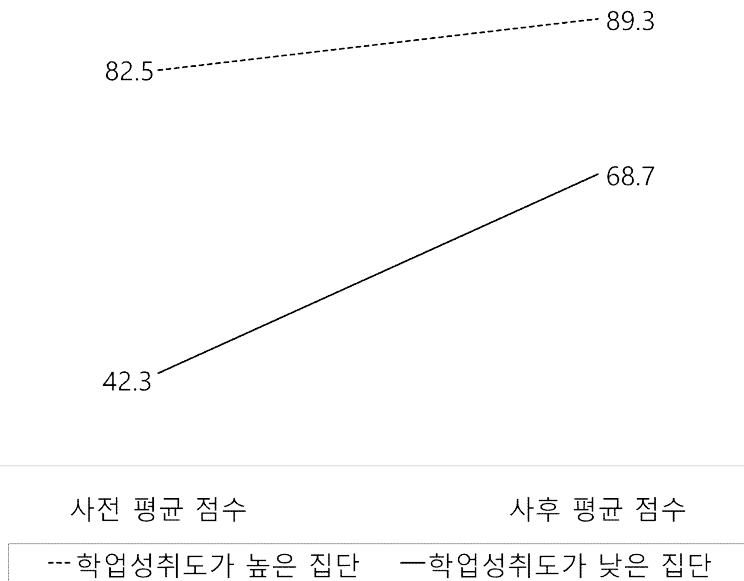
고등학교 『통합사회』, 『사회·문화』, 『윤리와 사상』 활용

[문항 2] “인공지능은 사회적 불평등을 심화시킨다”라는 주장에 대해 (자료 1)~(자료 6)을 활용해 그 정당성을 논리적으로 서술하고, 선택하지 않은 나머지 자료를 모두 활용해 반론을 제시하시오.

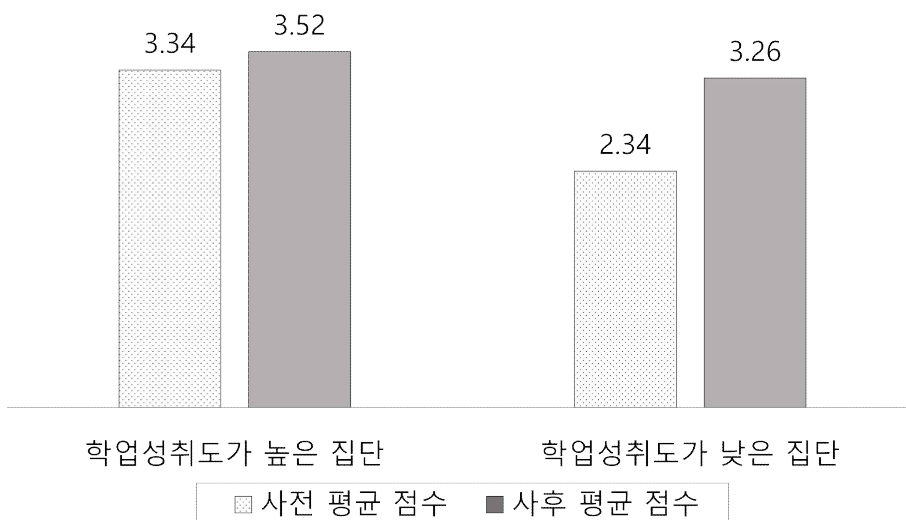
(600자±60자, 40점)

(자료 1) A초등학교는 학생들의 수학 과목 학습 지원을 위해 인공지능 학습기를 제공하였다. <그림 1>과 <그림 2>는 학업성취도가 높은 집단과 낮은 집단의 인공지능 학습기 사용 전과 후의 수학 평균 점수 및 학업적 자기효능감의 변화를 나타낸 그래프다.

<그림 1> 수학 평균 점수 변화



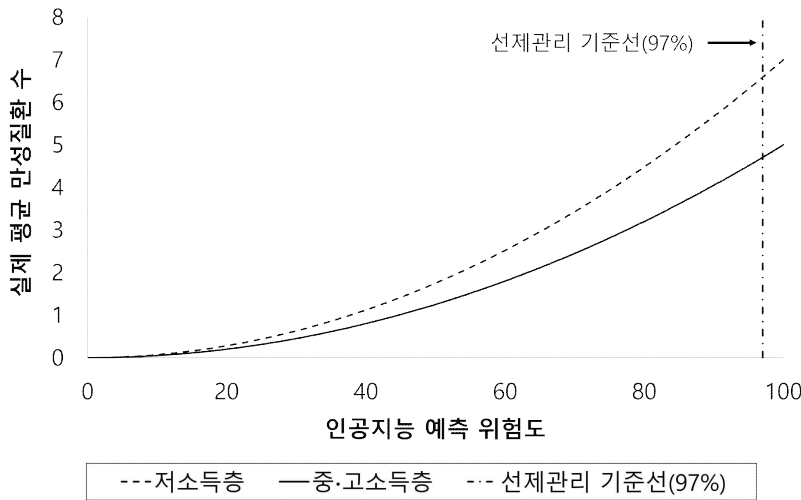
<그림 2> 학업적 자기효능감 변화



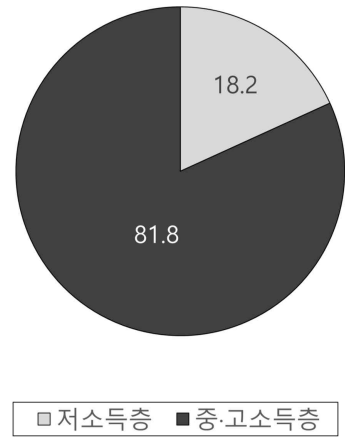
\* 자기효능감: 자신이 수행하려는 과제를 해낼 수 있는 능력이 있다고 믿는 신념

(자료 2) A국 보건당국은 만성질환 위험군을 조기에 파악하고 그에 대응하기 위해 '건강위험군 선제관리 프로그램'을 운영하고 있다. 인공지능 알고리즘은 의료비 지출과 진료·검진·처방 데이터를 바탕으로 위험도를 산출한다. 위험도가 97% 기준선을 넘으면 보건당국은 상담·모니터링 등 건강관리를 위한 선제관리 프로그램을 제공한다. <그림 3>은 소득 계층별 실제 평균 만성질환 수와 인공지능이 예측한 위험도를, <그림 4>는 선제관리 프로그램 수혜자의 소득 계층별 비율을 보여준다.

<그림 3> 소득 계층별 실제 평균 만성질환 수와 인공지능이 예측한 위험도(%)



<그림 4> 선제관리 프로그램 수혜자의 소득 계층별 비율(%)



- \* 중·고소득층은 저소득층에 비해 의료비 지출이 많음.
- \* 인공지능 예측 위험도: 의료비 지출 및 의료 데이터를 종합해 산출한 개인별 질병 위험도 지표
- \* 선제관리 기준선(97%): 위험도가 높은 개인을 선제관리 프로그램 수혜자로 자동 연계하기 위해 설정한 기준값

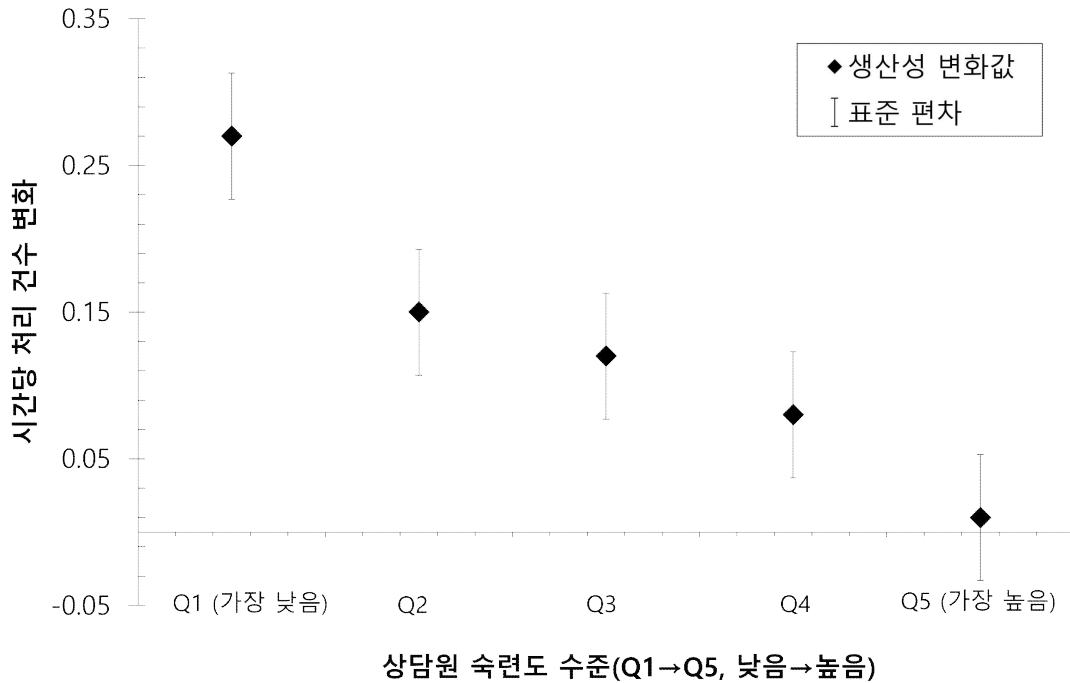
(자료 3) A대학 연구진은 유권자용 안내 책자를 핵심만 추려 설명하고, 사용자의 추가 질문에 따라 정보를 확장하여 제공하는 상호작용형 인공지능 챗봇을 개발하였다. 해당 챗봇의 효과를 검증하기 위해 정보접근성이 다른 두 지역의 주민을 각각 1,000명씩 선정한 후, 챗봇을 활용하여 정책·쟁점 정보를 탐색하도록 하였다. <표 1>은 지표별 두 지역 간 사전 점수 차이와 사후 점수 차이를 제시한 것이다.

<표 1> 정보·참여 지표별 두 지역 간 사전 점수 차이와 사후 점수 차이

지표	내용	두 지역 간 사전 점수 차이	두 지역 간 사후 점수 차이
정보 이해도	정보 해석과 이해의 정확성 정도	58	11
정보 처리 시간	정보를 이해·판단하는 데 소요된 시간의 정도	-34	-6
정보 탐색 활동	추가 정보를 자기 주도적으로 찾아본 정도	70	23
정보 탐색 범위	추가 탐색 과정에서 확인한 정보 범위와 다양성 정도	75	17
투표 참여 의향	향후 투표에 참여하려는 의지와 선택 확신의 정도	12	3

(자료 4) <그림 5>는 A기업의 고객지원센터 상담원 500명의 업무 데이터를 분석한 것으로, 생성형 인공지능 도입 이후 상담원의 숙련도별 시간당 처리 건수가 어떻게 달라졌는지를 보여준다. A기업은 생산성 향상분을 상담원의 인사 평가·승진·성과급 등 보상 체계에 반영하였다.

<그림 5> 생성형 인공지능 도입에 따른 숙련도별 시간당 처리 건수 변화

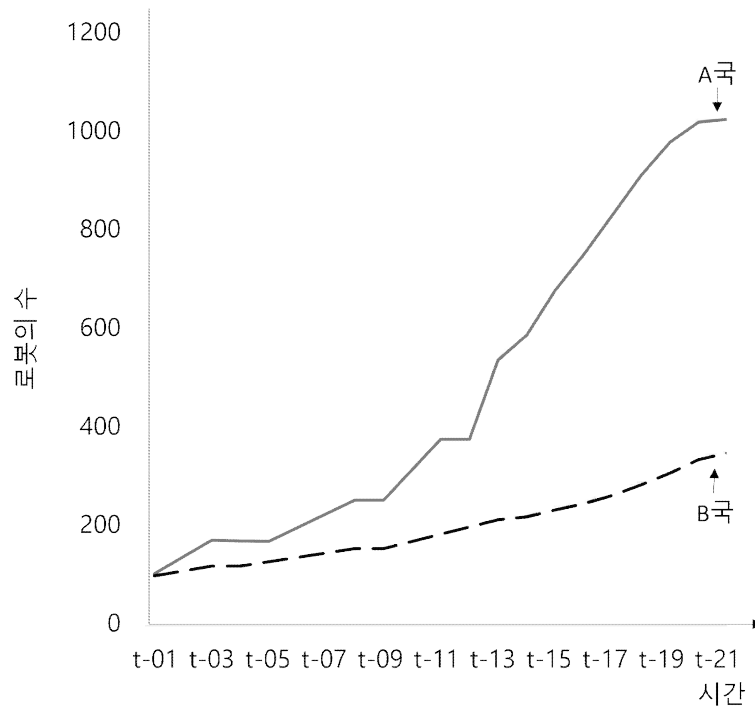


\* 시간당 처리 건수 변화: 생성형 인공지능 도입 전후 상담원이 단위 시간 동안 처리한 업무량이 얼마나 변했는지를 로그값의 변화로 나타냄. 0은 변화 없음, 양수는 증가, 음수는 감소를 의미함.

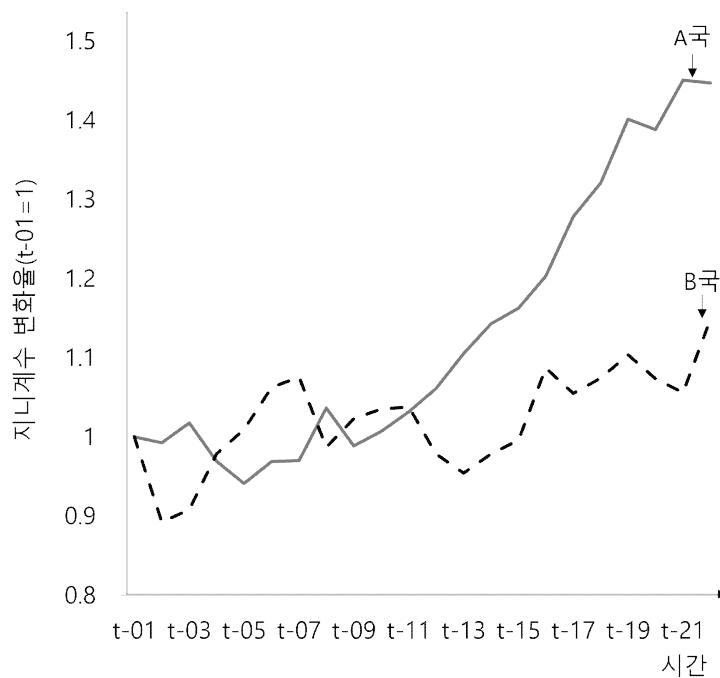
<뒷면에 계속>

(자료 5) <그림 6>은 시점 t-01부터 t-21까지 A국과 B국의 노동자 10,000명당 사용하는 로봇의 수를 나타낸 그래프다. 여기서 로봇은 물리적 행동을 수행하는 능력을 갖춘 인공지능으로 간주한다. <그림 7>은 같은 기간 이들 국가의 소득 불평등 수준을 나타내는 지니계수가 시점 t-01을 기준( $t-01=1$ )으로 얼마나 변화했는지를 보여주는 그래프다.

<그림 6> 노동자 10,000명당 로봇의 수

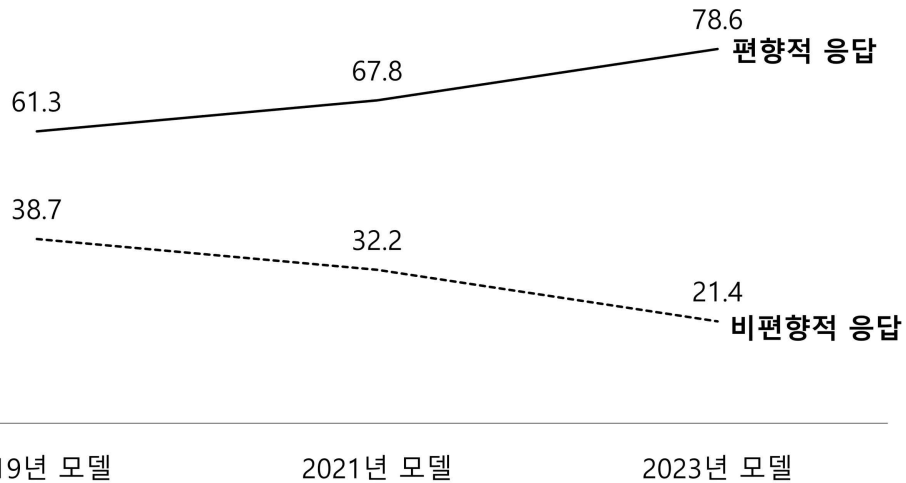


<그림 7> 지니계수 변화율



(자료 6) <그림 8>은 A연구소에서 개발한 연도별 인공지능 챗봇 모델의 응답 편향성 정도를 나타낸 그래프다. <표 2>는 2023년 B플랫폼의 직업별 구인 광고가 남성과 여성에게 어떠한 비율로 노출되었는지를 보여주는 도표다. B플랫폼의 광고 노출 알고리즘의 순수 효과를 검증하기 위해, 모든 직업 구인 광고는 투입 예산, 광고 전략, 게재 시간, 노출 기간, 목표 대상 등 외부 변인을 통제하여 동일한 조건에서 게시하였다.

<그림 8> 생성형 인공지능 챗봇 모델별 편향적 응답 비율



\* 편향적 응답은 고정관념을 보여주는 응답이고, 비편향적 응답은 고정관념에 반하는 응답임.

<표 2> 2023년 B플랫폼의 직업별 구인 광고의 성별 노출 비율(%)

직업	성별	
	남성	여성
건축 엔지니어	78.8	21.2
인공지능 엔지니어	72.8	27.2
법률가	59.2	40.8
간호사	11.8	88.2
비서	9.3	90.7
가사도우미	12.3	87.7
슈퍼마켓 점원	27.8	72.2









