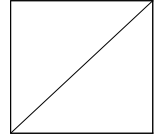


공 개



의안번호	제 5 호	보 고 사 항
보 고 연 월 일	2024. 12. 27. (제 66 회)	

2024년도 기술영향평가 결과(안) - 안전 · 신뢰 AI 기술 -

국가과학기술자문회의
심의회의 운영위원회

제 출 자	과학기술정보통신부장관 유상임
제출 연월일	2024. 12. 23.

1. 보고주문

- 「2024년도 기술영향평가 결과(안)」을 별지와 같이 보고함

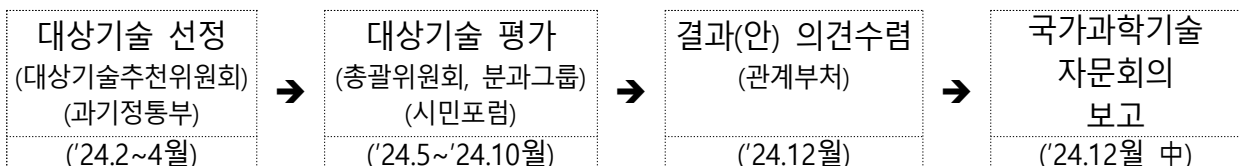
2. 제안이유

- 「과학기술기본법 제14조」 및 「동법 시행령 제23조」에 따라 「2024년도 기술영향평가 결과(안)」을 보고하고자 함

3. 주요내용

□ 평가 개요

- **(배경)** 새로운 과학기술 발전이 우리 사회 전반에 미치는 영향은 복잡하고 광범위해지고 있어, 이에 대한 선제적 예측 및 대응 필요
 - ※ 「과학기술기본법」 제14조에 의거하여 매년 실시(2003년부터)
- **(목적)** 과학기술 발전이 우리 사회 전반에 미치는 영향을 사전에 평가*하고 바람직한 발전 방향을 유도하기 위한 핵심과제 등 제시
 - * 「과학기술기본법」 시행령 제23조 제5항에 따라 민간 전문가 및 시민 등 참여 확대
- **(대상기술)** 2024년 평가 대상기술로 「안전·신뢰 AI 기술」을 선정
 - 대상기술추천위원회에서 전문가·시민 제안을 바탕으로 후보기술을 추천하고 부처 선호도 조사를 반영하여 과기정통부에서 선정(4.8)
 - ※ 전문가·시민제안('24.2~3), 부처수요조사('24.2~3), 추천위('24.3), 선호도조사('24.3)
- **(경과)** 전문가로 구성된 대상기술추천위원회, 총괄위원회(3개 분과)와 시민으로 구성된 시민포럼을 통해 기술영향평가 결과(안)을 마련
 - ※ (총괄위원회) 9명, 총 3회 개최 / (분과그룹) 3개 분과 19명, 총 9회 개최



□ 기술 분석

- **(정의)** 안전·신뢰 AI 기술은 인류의 보편적 규범에 부합하고 법적 요구에 충실하며 기술적으로 견고한 인공지능 기술

※ 국가전략기술 인공지능 임무 중심 전략로드맵(2023.10.31.)

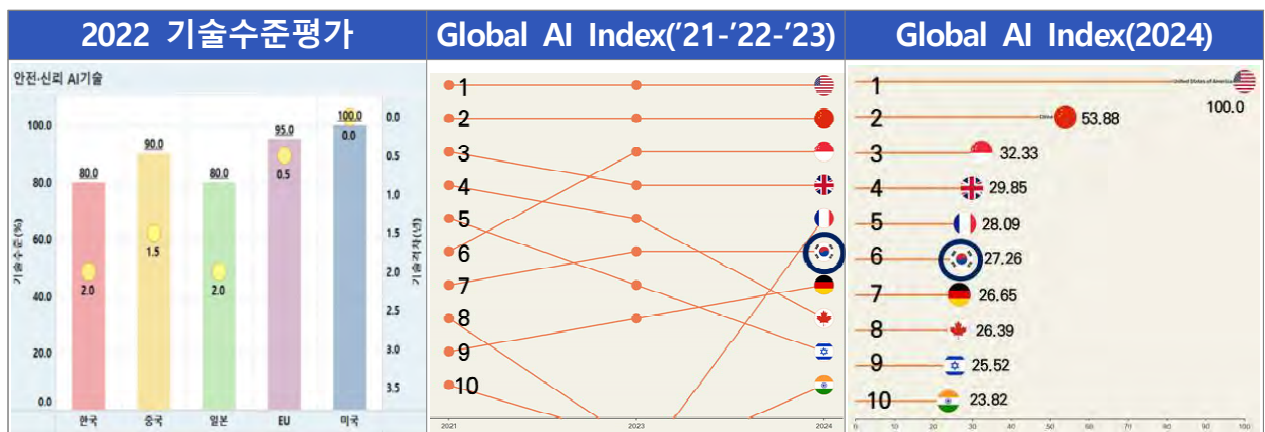
핵심기술

- ▶ 개인정보 및 저작권 침해를 방지하기 위한 **데이터 편향성 탐지 및 공정성 검증** 기술
- ▶ AI 모델의 동작원리를 정확하게 이해하기 위한 **설명가능한 AI(XAI)** 기술
- ▶ AI 모델의 안정성을 확보하기 위한 **시스템 오용 및 오작동 방지** 기술

- **(기술수준)** 한국은 안전·신뢰 AI 기술이 최고기술 보유국인 미국에 대비하여 80%로 평가*되었고, 글로벌 AI지수**에서 6위를 차지

* 2022년도 기술수준평가(2024.2.29.)

** 영국의 토터스 미디어, Global AI Index 2024(2024.9.19.)



□ 평가내용 및 방법

- **(평가범위)** 총괄위원회에서 AI 관련된 타 평가제도와 중복을 지양하면서, 상호보완적인 평가가 될 수 있는 기술 분야 선정*

* ▲헬스케어 AI, ▲휴머노이드 AI, ▲브레인 AI 등 3개 분야 기술에 대한 평가

- **(평가방법)** ①이슈 트렌드 분석*, ②전문가·시민포럼 평가 통한 주요 이슈 및 과제도출, ③설문조사** 통한 핵심이슈 및 과제 도출·분석

* 빅데이터 분석으로 키워드 사전 구축 및 웹 반응 분석실시('24.7~9)

** 도출된 이슈 및 과제에 대해 전문가 125명, 시민 1,000명 대상 정책 분석('24.11)

□ 기술영향평가 결과

- **(트렌드분석)** 이슈의 흐름이 개인정보 보호 및 정부 정책 중심에서 ChatGPT 등장 이후 안전, 신뢰, 윤리와 기업 중심으로 전환

※ 언론보도를 통해 핵심 키워드를 분석하고 전문가·시민포럼에게 정보제공 목적

※ 2019년부터 '개인정보'와 '보안'과 관련된 언론의 관심이 지속적으로 증가

- **(주요이슈)** 전문가·시민포럼이 제시한 79개(전문가 58개/시민 21개)의 이슈를 분석*하여 14개의 주요 이슈 도출

* 이슈 간 중복성 제거, 핵심 우려 사항 명확화, 평가 및 분석이 용이한 구조 확보, 전문가와 시민 관점의 균형적 반영을 고려

※ Claude 3.5 Sonnet, Perplexity 등 생성형 AI를 활용하여 초안 도출 후, 총괄위원회 검토

① AI 수집 개인정보의 무단 활용과 유출 위험	② AI 첨단 서비스의 사회경제적 접근성 격차	③ AI 사고와 오류에 대한 책임소재 불명확성
④ AI 의사결정 과정의 불투명성과 알고리즘 편향	⑤ AI 의존으로 인한 인간 자율성과 사회성 약화	⑥ AI 시스템 간 상호운용성 부재와 표준화 문제
⑦ AI 건강관리 시스템의 부작용과 오류 위험	⑧ 프리미엄 AI 의료서비스의 양극화와 건강 격차	⑨ 휴머노이드 로봇의 물리적 위험과 오작동 가능성
⑩ 휴머노이드로 인한 일자리 대체와 사회구조 변화	⑪ 뇌-컴퓨터 인터페이스의 기술적 오류와 안전성 위험	⑫ 인공적 인지능력 증강의 윤리적 한계와 부작용
⑬ 뇌 활동 데이터의 유출과 프라이버시 침해 위험	⑭ 뇌-컴퓨터 인터페이스의 중독성과 의존도 심화	-

- **(주요과제)** 전문가·시민포럼이 이슈 대응을 위해 제시한 총 100개 (전문가 78개/시민 22개)의 과제를 분석*하여 17개의 주요 과제 도출

* 실행가능한 정책수단 여부, 과제 간 유사성 등을 고려

※ Claude 3.5 Sonnet, Perplexity 등 생성형 AI를 활용하여 초안 도출 후, 총괄위원회 검토

① 미래기술 안전성 기본법	② 윤리안전 가이드라인	③ 민감정보 기반 기술 거버넌스
④ 첨단기술 품질인증·평가체계	⑤ 미래기술 안전성 혁신 네트워크	⑥ 미래기술 안전 전문가 양성
⑦ 첨단기술 안전 리터러시	⑧ 미래기술 안전성 R&D	⑨ 헬스케어 AI 임상안전 제도
⑩ 의료 AI 데이터 표준화	⑪ 의료 AI 실증 플랫폼	⑫ 로봇 물리안전 기준
⑬ 로봇-인간 상호작용 제도	⑭ 로봇협업 안전관리 제도	⑮ 브레인 인터페이스 안전기준
⑯ 인지증강 기술 관리제도	⑰ 뇌신호 처리 신뢰성 제도	-

- **(핵심이슈)** 총 14개의 주요이슈에 대해서 발생가능성 및 영향력*이 타 과제에 비해 상대적으로 높은 핵심이슈 5개**를 도출

* 전문가 125명, 시민 1,000명 대상으로 설문조사를 통해 주요이슈의 발생가능성, 영향력 점수의 평균값을 중심으로 영역을 구분하여 상대적 우선순위를 평가

** ①개인정보, ③사고/오류 책임소재, ⑤인간 자율성/사회성 악화, ⑩휴머노이드로 인한 일자리/ 사회변화, ⑭ BCI 중독/의존(※전문가, 시민 5개 이슈 모두 동일)



- **(핵심과제)** 총 17개의 주요과제 중 시급성 및 중요성이 타 과제에 비해 상대적으로 높은 핵심과제 9개*를 도출

* ①미래기술 안전성 기본법, ②인간친화 기술 윤리안전 가이드라인, ③민감정보 기반 기술 거버넌스, ⑧미래기술 안전성 R&D, ⑨헬스케어AI 임상안전제도, ⑩의료AI 데이터 표준화, ⑪의료AI 실증플랫폼, ⑫로봇 물리안전 기준, ⑭로봇협업 안전관리제도 (※전문가, 시민 8개가 동일하며, ⑫로봇 물리안전 기준은 전문가만 선택)



[참고] 9개 핵심과제 관련 정책 시범 분석

○ 9개 핵심과제와 관련된 정책을 법령 및 정책 DB 및 생성형 AI 등을 활용하여 분석

※ Claude 3.5 Sonnet, Perplexity 등 생성형 AI를 활용하여 초안 작성 후, 국가 법령정보센터, 입법예고, 행정예고, NTIS, K2Base 등 검색하여 검증

※ 본 분석은 전문가의 검증을 거치지 않은 내용으로 향후 구체적인 환류 방안 탐색을 위해 시범 분석 수행

핵심과제	관련 정책(사례)	개선방향		소관
		신규	보완	
미래기술 안전성 기본법	- 지능정보화 기본법 제21조: "지능정보기술 안전성·신뢰성" - AI 산업 발전 및 신뢰 기반 조성에 관한 법률안(2023) - 입법예고(2024.1): AI 산업 발전 및 신뢰 기반 조성에 관한 법률안 - 생체정보 활용 AI 안전기준 - 위험등급별 관리체계	- 생체-기계 인터페이스 조항 - 실시간 위험관리 기준	- 안전성 심사위원회 - 사고배상 체계	과기 정통부
인간친화 기술 윤리안전 가이드라인	- AI 윤리기준(2020): "사람중심 AI" - 신뢰할 수 있는 AI 개발 가이드라인 (2021)	- 인체개입 윤리기준 마련 - 자율성 한계 설정	- 분야별 세부지침 - 평가인증 체계	과기 정통부
민감정보 기반 기술 거버넌스	- 개인정보보호법 제23조: "민감정보 처리 제한" - 생명윤리 및 안전에 관한 법률	- 초민감정보의 범위 등 분류체계 마련 - 민감정보의 실시간 수집기준 마련	- 통합관리 체계 - 보안기준 수립	개보위 복지부
미래기술 안전성 R&D	- AI 기술개발 사업(2023) - AI안전성강화R&D프로그램	- 연구분야 확대 - 실증 강화	- 통합R&D체계 - 성과관리제도	과기 정통부
헬스케어 AI 임상안전 제도	- 의료기기법 - AI의료기기허가·심사 가이드라인	임상검증 강화	- 사고예방체계 - 책임규정마련	복지부 식약처
의료 AI 데이터 표준화	- 보건의료 데이터 표준화 로드맵(2022) - 의료데이터중심병원지원사업	- 표준화 범위 확대 - 품질기준 구체화	- 통합관리체계 - 인증제도도입	복지부
의료 AI 실증 플랫폼	국가전략기술 특화연구소 데이터플랫폼에 대한 실증특례	다기관 임상시험 지원 체계 구축	의료 AI 실증을 위한 규제 샌드박스 확대	복지부
로봇 물리안전 기준	지능형 로봇 개발 및 보급 촉진법	-	- 로봇의 물리적 안전성에 대한 구체적 기준 마련 - 안전성 평가 및 인증 체계 구축	산업부
로봇협업 안전관리 제도	-	인간-로봇 협업 환경의 안전성 보장을 위한 종합 안전관리 체계 마련	-	산업부

[참고] 2024 기술영향평가 결과 요약

단계	방법	내용
이슈 트렌드 분석		- 빅데이터 분석으로 이슈 키워드를 파악하고 전문가, 시민에게 참고자료 제공을 위해 실시 - 언론의 이슈는 chatGPT 이후 이슈가 개인 정보보호 → 안전, 신뢰, 윤리 등으로 확대
주요 이슈/과제 도출		- 전문가들은 58개 이슈, 78개 과제를 제시 - 시민포럼들은 21개 이슈, 22개 과제를 제시 - KISTEP에서 유사 과제 통합, 정책수단 연계 등 고려→주요이슈 14개, 주요과제 17개 도출 (Claude 3.5 Sonnet 등 생성형 AI 활용)
핵심 이슈/과제 도출		설문조사를 통해 주요이슈를 발생가능성, 영향력으로, 주요과제를 중요성, 시급성으로 분석 → 5개 핵심이슈, 9개 핵심과제 도출 (전문가 125명, 시민 1,000명 설문조사)
정책분석 및 환류		9개 핵심과제와 관련 정책 시범분석 (법령 및 정책 DB, 생성형 AI 활용) - 관련 정책별로 신규/보완 등 개선과제 도출 - 정책협의체 등 지속적 추적/환류 방안 마련

핵심이슈(5개)		핵심과제(9개)	
공통 이슈	1. 개인정보 위험	1. 미래기술 안전성 기본법	공통 과제
	3. 사고/오류 책임소재	2. 인간친화 기술 윤리안전 가이드라인	
	5. 인간자율성/사회성 약화	3. 민감정보 기반 기술 거버넌스	
특 화 이 슈	10. 휴머노이드 일자리/사회변화	8. 미래기술 안전성 R&D	특 화 과 제
	14. BCI 중독/의존	9. 헬스케어시 임상안전제도	
		10. 의료시 데이터 표준화	
		11. 의료시 실증플랫폼	
		12. 로봇 물리안전 기준	
		14. 로봇협업 안전관리제도	

4. 시사점 및 개선방향(안)

□ 시사점

- 기술영향평가 결과의 정책반영률이 지속적으로 하락*함에 따라 환류 체계의 실효성이 저하

* ('21,자율주행) 77.8% → ('22,합성생물학) 71.4% → ('23,양자과학기술) 68.8%

⇒ 기술영향평가의 역할과 기능, 환류체계를 재점검 및 개편하여 기술영향평가의 실효성 제고 필요

□ 개선방향(안)

과제	AS-IS	TO-BE
기본 방향	① 미래기술 담론 형성에 초점 : 미래기술 관련 시민 담론 형성에 초점 → 정책과제화, 환류 등의 연계에 한계	① 미래기술이슈 종합대응체제로 개편 : 이슈탐지→정책과제 분석→정책환류 3단계 미래기술이슈 대응 체제로 개편
대상 기술	② 장기 미래기술 평가로 환류 곤란 : 장기적으로 국민생활에 파급효과를 미칠 기술을 평가하고 있으나, 긴 시차로 정책환류가 불가	② 단·중기 확산 기술 대상 평가 : 단기 및 중기에 사회적으로 확산될 기술, 연구개발이 진행될 기술 대상 평가 실시
핵심 이슈	③ 평가항목별 기계적 이슈 도출 : 경제사회문화윤라환경 등 상행량 기준을 기계적 적용하여 핵심이슈 파악 곤란	③ 핵심이슈 중심으로 이슈 탐지 : 발생가능성, 영향력 등 고려하여 핵심 이슈 중심으로 도출
정책 분석	④ 추상적 정책 제언 : 담론적·추상적인 정책제언으로 평가결과 환류방안의 구체성 미흡	④ 핵심과제 도출·목록화 : 국가사회적 필요 기술분야에 대한 구체적 방향성을 제시할 수 있도록 핵심과제를 법령, 계획, 사업 등 유형별로 목록화(관련부처 협의 필요)
환류	⑤ 단기·1회성 정책 환류 : 평가결과 반영점검이 1회만 진행되고, 부처의 참여 유인도 낮아 환류의 지속성 부재	⑤ 정책협의체 구성·추적관리 : 평가단계에서 대상기술별 관련 기관이 참여하여 최소 3년간 주요정책과제 조정·관리 등 환류체계 구축

2024년도 기술영향평가 결과(안)

- 안전 · 신뢰성 기술 -

2024. 12.



과학기술정보통신부



목 차



I. 2024년 기술영향평가 개요	1
II. 2024년 기술영향평가 결과	6
1. 기술 개요	6
2. 평가 내용 및 방법	15
3. 기술영향평가 결과	21
① 이슈 트렌드 분석	21
② 주요 이슈 및 과제 분석	27
③ 핵심 이슈 · 과제 도출	48
III. 시사점 및 개선방향[안]	58
IV. 향후 일정	64
<붙임1> 연도별 기술영향평가 대상기술 및 주요 결과	65
<붙임2> 2023년 기술영향평가(양자과학기술) 정책 반영 실적	68
<붙임3> 2024년 기술영향평가위원회 및 시민포럼 명단	74
<붙임4> 빅데이터 분석	77

I. 2024년 기술영향평가 개요

□ 추진 배경

- 새로운 과학기술이 우리 사회에 미치는 영향은 복잡하고 광범위해지는 양상을 보이므로, 이에 대한 선제적 예측 및 대응 필요
 - 과학기술의 발전은 경제적 부가가치 창출 등의 긍정적인 효과도 가져 오지만, 환경·윤리 문제 등 예기치 못한 부작용도 초래
 - 과학기술은 다양한 분야 및 계층을 대상으로 사용되기 때문에 성별 등 사용자 특성에 대한 분석이 개발 단계에서부터 필요
 - 사회구성원이 참여한 기술영향평가(TA, Technology Assessment)를 통해 미래 신기술이 초래할 결과를 예측하여 사전 대응 방안을 마련
- ※ 연도별 기술영향평가 대상기술 및 주요 결과는 붙임 1 참고

기술영향평가 추진 근거

과학기술기본법 제14조 (기술영향평가 및 기술수준평가)

- ① 정부는 새로운 과학기술의 발전이 경제·사회·문화·윤리·환경 등에 미치는 영향을 사전에 평가(이하 "기술영향평가"라 한다)하고 그 결과를 정책에 반영하여야 한다.
- ② 정부는 과학기술의 발전을 촉진하기 위하여 국가적으로 중요한 핵심 기술에 대한 기술수준을 평가(이하 "기술수준평가"라 한다)하고 해당 기술 수준의 향상을 위한 시책을 세우고 추진하여야 한다.
- ③ 정부는 기술영향평가를 실시하는 경우 대상기술의 성격을 고려하여 성별 등 특성 분석이 반영될 수 있도록 하여야 한다.
- ④ 기술영향평가와 기술수준평가의 범위 및 절차 등에 관하여 필요한 사항은 대통령령으로 정한다.

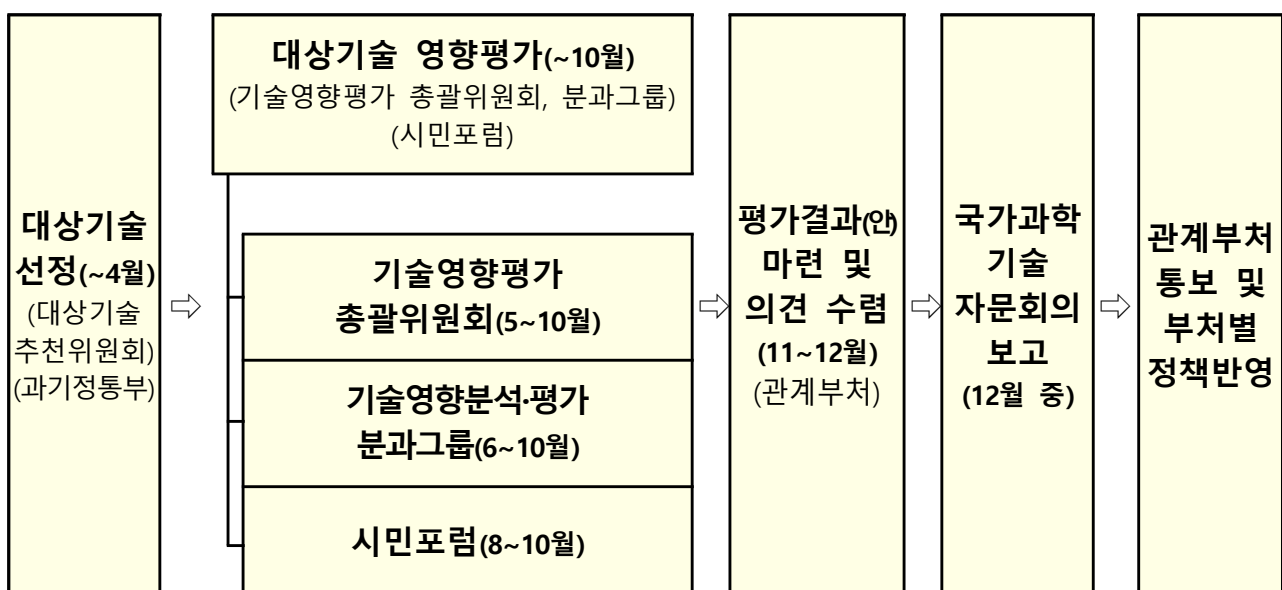
□ 평가 내용

- 새로운 과학기술의 발전이 경제·사회·문화·윤리·환경 등에 미치는 영향을 사전에 평가(과학기술기본법 제14조)
 - ※ 「과학기술기본법」 시행령 제23조 제5항에 따라 민간 전문가 및 시민 등 참여 확대
- 긍정적 영향은 극대화하고, 부정적 영향은 최소화하기 위한 정책적 고려사항을 도출하여 정책 수립부터 바람직한 발전 방향을 유도
- 관계부처는 기술영향평가 결과를 향후 소관 분야 국가연구개발사업의 연구기획에 반영하거나 부정적 영향을 최소화하기 위한 대책을 추진 (과학기술기본법 시행령 제23조 제7항)

기술영향평가의 주요 내용 (과학기술기본법 시행령 제23조 제2항)

- 국민 생활의 편의 증진 및 관련 산업 발전에 미치는 영향
- 새로운 과학기술이 경제·사회·문화·윤리·환경에 미치는 영향
- 부작용 초래 가능성이 있는 경우 이를 방지 할 수 있는 방안
- 기술의 성격과 파급효과가 성별 등 특성에 미치는 영향

□ 추진 경과



○ **기술영향평가 대상기술 선정('24.3~4월)**

미래의 신기술 및 기술적·경제적·사회적 영향과 파급효과 등이 큰 기술

- ① 현재 등장하고 있는 신기술로서 가까운 미래의 국민 생활에 파급효과가 클 것으로 예상되는 기술
 - ② 사회적 관심도가 높아 기술영향평가의 유용성과 시의 적절성이 높을 것으로 기대되는 기술
 - ③ 개념과 범주가 명확한 중간규모*의 기술
- * 과학기술표준분류상의 중분류 또는 소분류 수준으로서 과제(project) 수준 기술은 제외
- ④ 과거에 평가를 실시한 기술 중에서도 환경변화 및 기술 발전 추이를 고려해 재검토가 필요하다고 인정되는 기술

- 해외 영향평가, 유망기술, 전문가 및 시민 제안 등 반영하고, 정책적 시의성 및 파급효과 등을 고려하여 **20개의 후보기술(안)을 도출**
- 20개 후보기술 중에서 특성, 파급효과의 크기, 사회적 합의의 필요성, 결과의 활용 가능성을 평가하여 **8개*** 기술 추천
- * 디지털 트윈, 안전·신뢰 AI, 지능형반도체, 기후변화예측·대응, 수소생산·저장·운송, 도심항공교통, 하이퍼튜브, 정신건강치료 등 8개 기술
- 추천된 8개 기술 중 **'안전·신뢰 AI 기술' 최종 결정**

대상기술 선정 사유

- 2024년 기술영향평가 대상기술인 **'안전·신뢰AI'**는 2022년 국가전략기술로 선정되었으며, **2023년 10월 전략로드맵이 마련됨**
- **'데이터 공정성 확보', 'XAI', 'AI시스템 견고성'**이 요소기술이며, 주요국의 자국 중심 AI 규범 마련에 따라 신뢰성 관련 기술 주도가 필요
- 인공지능 알고리즘 설계의 **기술적 복잡성 증가**로 인해 딥페이크, 데이터 편향성 등의 **예기치 않은 오류의 인지나 신뢰 이슈** 등에 대한 **영향평가 필요**

○ 기술영향평가 대상기술 예비 분석('24.5~6월)

- (총괄위원회) 예비분석을 통해 기술 특성과 현재 기술 개발·활용 동향 등을 검토하고, 주요 평가 이슈를 도출

※ 과학기술 전문가, 인문사회 전문가 등 다양한 분야 전문가로 구성(총 9명)

- '안전·신뢰AI기술'에 대한 정의와 평가 범위 등을 확정하고, 대상 기술의 현재 동향과 발전 전망에 대한 사전 논의

○ 전문가 분석('24.7~10월)

- (전문가 분과그룹) 전문가 관점에서 기술 분야의 파급효과 분석 및 정책성을 고려하여 이슈 및 대응과제 제시

※ 헬스케어, 휴머노이드, 브레인 기술 관련 산·학·연 및 언론 등 전문가로 구성(총 19명)

○ 시민포럼('24.8~10월)

- (시민포럼) 일반시민의 관점에서 기술에 관한 생각을 표현하고 파급효과 및 권고사항을 고려하여 이슈 및 대응과제 제시

※ 성별, 연령, 직업, 전공, 학력 등을 균형 있게 배분하고 역량 등을 고려하여 선발(총 15명)

○ 기술영향평가 결과(안) 마련('24.11월)

- 전문가 및 시민포럼에서 제시한 이슈 및 과제를 토대로 유사 과제 통합 및 정책수단 연계를 통해 주요이슈 및 과제를 도출

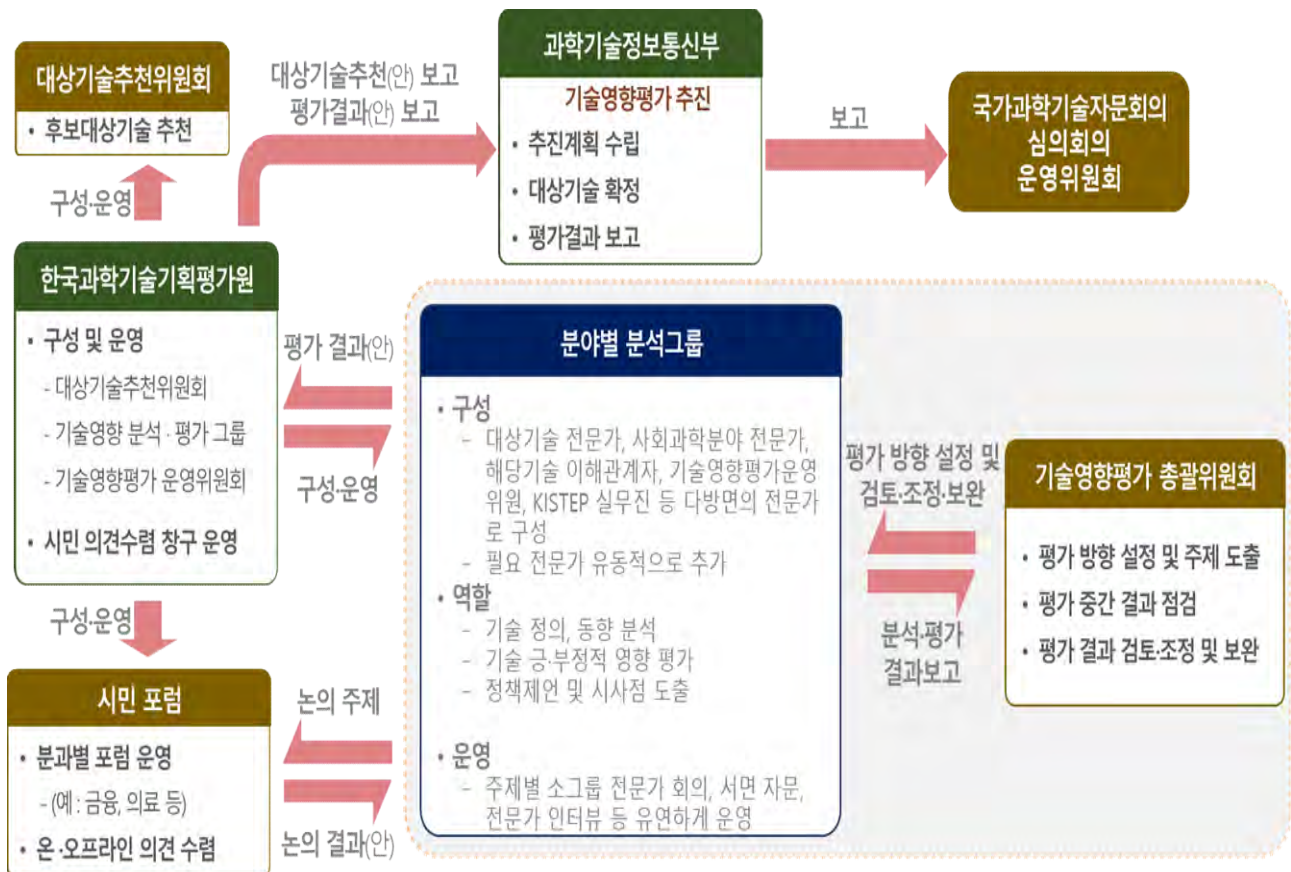
- 주요이슈 및 과제를 바탕으로 전문가·시민 대상 설문조사 통해 핵심 이슈 및 과제를 도출하고, 핵심과제와 관련 정책을 분석

※ 결과(안)를 마련하여 관계부처에 의견회람(12월)

○ 국가과학기술 자문회의 결과보고('24.12월)

- 과학기술기본법 및 국가과학기술자문회의법에 따라 결과 보고

□ 추진 체계



- (과학기술정보통신부) 기술영향평가 총괄 추진
- (대상기술추천위원회) 후보 기술군을 검토하여 후보 대상기술 추천
- (기술영향평가 총괄위원회) 평가 방향 설정 및 주제 도출, 평가 결과를 검토·조정 및 보완
- (분야별 분과그룹) 기술 정의 및 동향 분석, 기술의 긍·부정적 영향 평가, 이슈 및 대응과제 제시
- (시민포럼) 시민 관점에서 이슈 및 대응과제 제시
- (한국과학기술기획평가원) 과학기술기본법에 따라 과기정통부에서 위탁 받아 기술영향평가 업무 수행

II. 2024년 기술영향평가 결과 : 안전 · 신뢰 AI기술

1 기술 개요

□ 기술의 정의

- **(안전 · 신뢰AI)** 인류의 보편적 규범 가치에 부합하고 법적 요구사항에 충실하며 기술적으로 안전하고 견고한 신뢰할 수 있는 인공지능 기술*

* 국가전략기술 「인공지능」 임무 중심 전략로드맵(2023.10.31.)

- AI 지속적인 발전과 AI 역기능 문제를 최소화하기 위해, 결론·도출 과정 등에 대한 신뢰성 확보를 위한 설명가능한 AI 기술개발 필요

핵심기술

- ▶ 개인정보 및 저작권 침해를 방지하기 위한 데이터 편향성 탐지 및 공정성 검증 기술
- ▶ AI 모델의 동작원리를 정확하게 이해하기 위한 설명가능한 AI(XAI) 기술
- ▶ AI 모델의 안정성을 확보하기 위한 시스템 오용 및 오작동 방지 기술



※ 출처 : 조원영(2022), 'AI 신뢰 제고를 위한 연구 동향' 재구성

- 국제적으로는 신뢰성 AI(Trustworthy AI, EU/OECD), 신뢰성·책임성 있는 AI(Trustworthy and Responsible AI, NIST) 등으로 통용 → 국가별 차이가 있으나, 견고성, 투명성, 공정성 등 주요 윤리원칙을 요건으로 제시

※ 국가전략기술 「인공지능」의 임무 중심 전략 로드맵(2023.10.31.)에는 주로 신뢰성 AI (truthworthy AI)를 정의하고 있으나, 생성형 AI 등장 이후 안전(safety) 개념 강화

[표] 안전·신뢰AI의 다양한 정의

국가/기구	출처	윤리기준
EU (2019)	Ethics guidelines for trustworthy AI	▲ Lawful AI, Ethical AI, Robust AI
한국 (2020)	국가 인공지능 윤리기준	▲ (3대원칙) 인간 존엄성 원칙, 기술의 합목적성 원칙, 사회의 공공선 원칙 ▲ (10대 핵심요건) 인권보장, 프라이버시 보호, 다양성 존중, 침해 금지, 공공성, 연대성, 데이터 관리, 책임성, 안전성, 투명성
UNESCO (2021)	the Ethics of Artificial Intelligence	▲ Proportionality and Do No Harm, ▲ Safety and security, ▲ Fairness and non-discrimination, ▲ Sustainability, ▲ Right to Privacy, and Data Protection, ▲ Human oversight and determination, ▲ Transparency and explainability, ▲ Responsibility and accountability ▲ Awareness and literacy, ▲ Multi-stakeholder and adaptive governance and collaboration
미국 (NIST, 2023)	Artificial Intelligence Risk Management Framework (AI RMF 1.0)	▲ Validity and Reliability, ▲ Safety, ▲ Security and Resiliency, ▲ Accountability and Transparency, ▲ Explainability and Interpretability, ▲ Privacy, ▲ Fairness with Mitigation of Harmful Bias

□ 기술 동향

○ 인공지능 모델의 기술발전 속도는 양과 질적인 측면에서 폭발적 증가

- '23년에는 총 149개의 파운데이션 모델*이 출시되어 '22년 비해 두 배 이상 증가하였고, 깃허브의 오픈소스 프로젝트**도 '23년 한해 59.3% 증가

* 대규모 데이터와 계산 자원을 활용해 학습된 인공지능 모델을 지칭

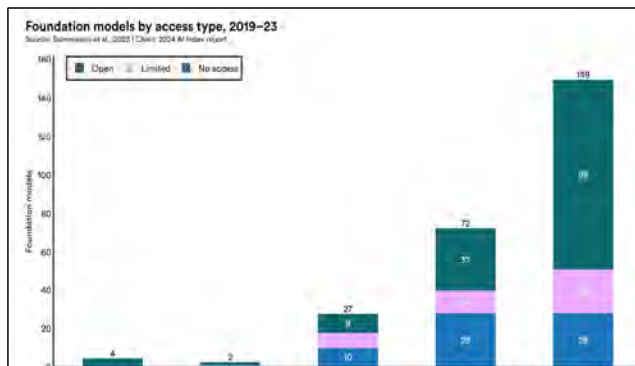
** 다양한 개발자와 조직이 협업하여 공유가 가능한 소프트웨어를 개발하고 공유하는 플랫폼으로, 누구나 소스 코드를 확인하고 기여할 수 있는 프로젝트 지칭

- '22년의 언어 모델*에 이어, '23년에는 멀티모달 모델**이 출현

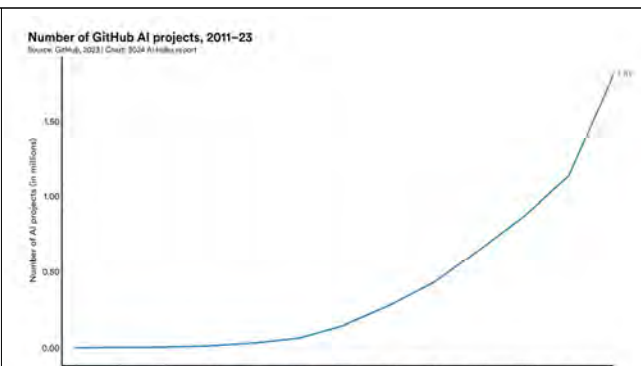
* 주어진 단어나 문장의 확률 분포를 예측하는 모델로, 주로 자연어 처리에 사용됨

** 텍스트, 이미지, 오디오, 비디오 등 다양한 유형의 데이터(모달리티)를 함께 고려하여 서로의 관계성을 학습 및 처리하는 인공지능

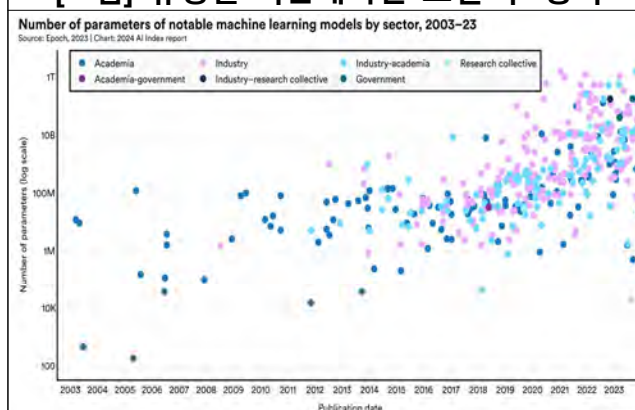
- OpenAI는 2018년 GPT-1 출시 이래, 거의 매년 새로운 GPT 모델을 출시하고 있으며, 최근에는 추론형 모델 **o1**을 출시('24.12월)



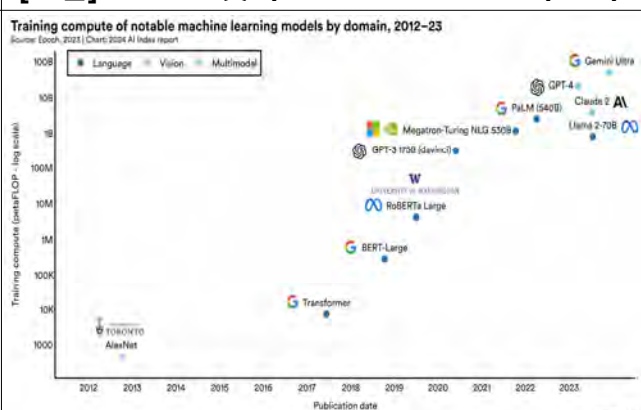
[그림] 유형별 파운데이션 모델 수 증가



[그림] 연도별 깃허브 오픈소스 프로젝트 수



[그림] 머신러닝 모델의 파라미터 수(로그스케일)



[그림] 주요 모델의 성능 발전

※ 출처 : AI Index 2024

[표] GPT 모델 업데이트 경과

버전	업데이트	주요 업데이트 내용
GPT-1	2018년 6월	▶ 최초의 GPT 모델, 트랜스포머 아키텍처를 사용하여 언어 모델링 성능 개선
GPT-2	2019년 2월	▶ 15억 개의 파라미터로 모델 크기 확장, 다양한 텍스트 생성 능력 향상
GPT-3	2020년 6월	▶ 1750억 개의 파라미터로 모델 대폭 확장, 자연어 이해 및 생성 능력 대폭 향상
GPT-3.5	2021년 11월	▶ GPT-3의 업그레이드 버전, 문맥 이해력과 응답 일관성 향상
GPT-4	2023년 3월	▶ 멀티모달 기능 도입 (텍스트 및 이미지 처리), 더 높은 정확도와 응답 일관성
GPT-4-turbo	2023년 11월	▶ GPT-4의 고속/고효율 버전, 더 낮은 비용으로 더 빠른 응답 제공
GPT-4o	2024년 5월	▶ GPT-4o는 오디오, 시각, 텍스트를 실시간으로 추론할 수 있는 OpenAI의 새로운 플래그십 모델
o1	2024년 9월	▶ 구조화된 검색 기능, 실시간 데이터 처리, 멀티턴 대화 내역을 통합하는 등 조직 인텔리전스에 초점

- 최근 미국의 Anthropic사가 AI 내부의 작동과정의 블랙박스를 이해할 수 있는 기술을 발표(24.5.21)하는 등 안전·신뢰 관련 연구도 활발하게 진행 중

[표] Anthropic의 인공지능 블랙박스 연구



※ 출처 : <https://www.anthropic.com/news/mapping-mind-language-model>

- 논문을 통해 연구분야별로 살펴보면 안전·신뢰AI 연구는 견고성 > 투명성 > 공정성 순으로 진행되고 있음을 보여줌



※ 출처 : 조원영(2022), AI 신뢰 제고를 위한 연구 동향

□ 기술수준

- OpenAI, Alphabet, Anthropic 등 미국 빅테크 기업의 Closed Model*이 MMLU** 등 주요 벤치마킹에서 타국에 비해 우수한 성능 제공

* Open Model(개방형 모델)과 대비되는 개념으로, 사용자가 모델에 직접적으로 접근하거나 수정할 수 없고, 특정 기업이나 조직에서만 제어하는 형태의 모델

** Massive Multitask Language Understanding(대규모 다중작업 언어 이해) : 15,000여 개 질문으로 구성된 AI의 지식과 문제 해결 능력을 평가하는 벤치마크 테스트

- 파운데이션 모델의 수도 미국이 압도적인 수준(AI INDEX 2024)

※ AI INDEX : 스탠퍼드 AI 연구소에서 주관에서 매년 발행하는 AI 관련 지수



- 한국의 안전·신뢰AI 기술수준은 최고기술 보유국인 미국 대비 80%로 평가 ('22 기술수준평가)

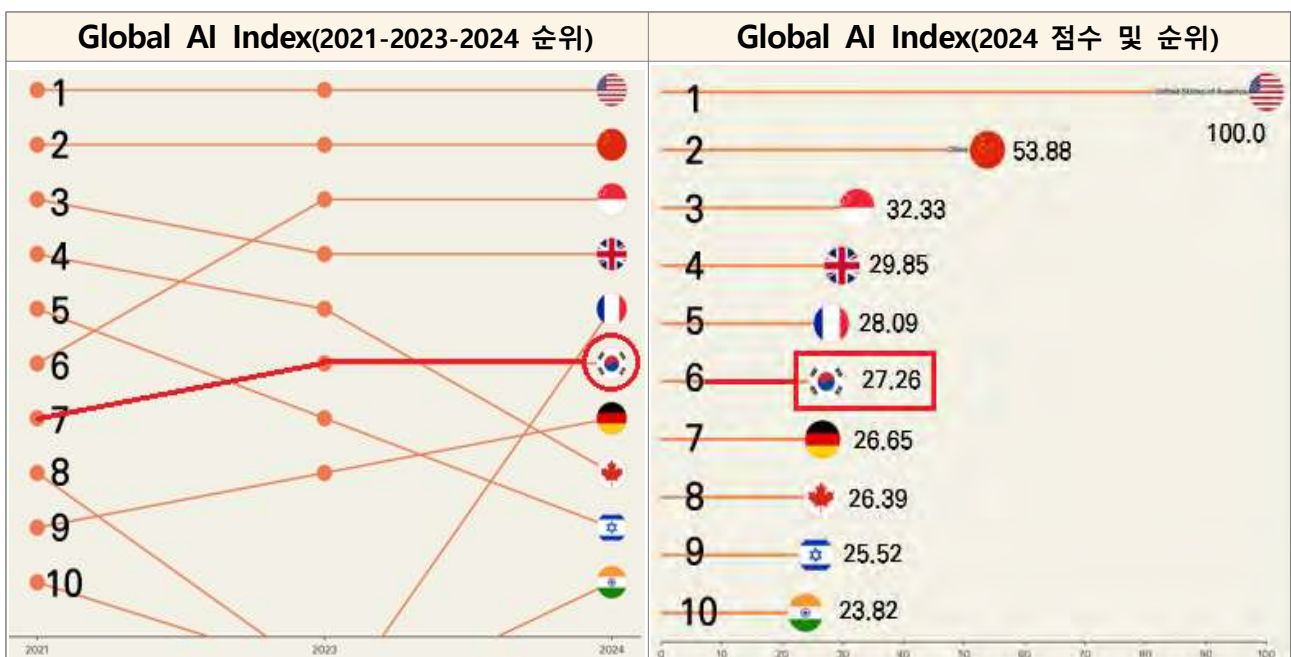
[표] 안전·신뢰AI 기술(2022 기술수준평가 결과)



※ 출처 : 과학기술정보통신부, 한국과학기술기획평가원(2024)

- 영국의 데이터 분석 업체 토터스 미디어가 발표한 글로벌 AI 지수*에서 한국은 2024년 전년에 이어 6위를 차지

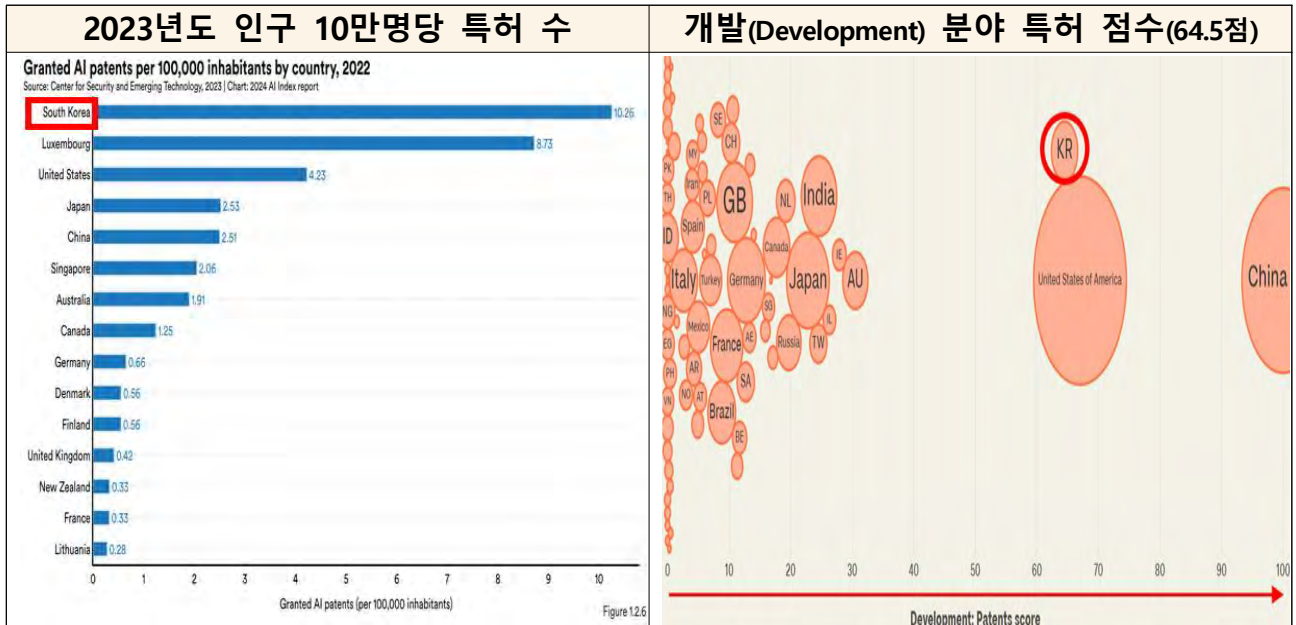
* Global AI Index 2024(2024.9.19.) : 영국의 Tortoise Media에서 제공하는 AI 기술 관련된 국가별 경쟁력 평가지표



- 한국의 AI 관련 특허성과의 경우, '23년도 인구 10만 명당 특허 수는 세계 1위*였으며, Global AI 지수에서도 3위 수준(64.5점)

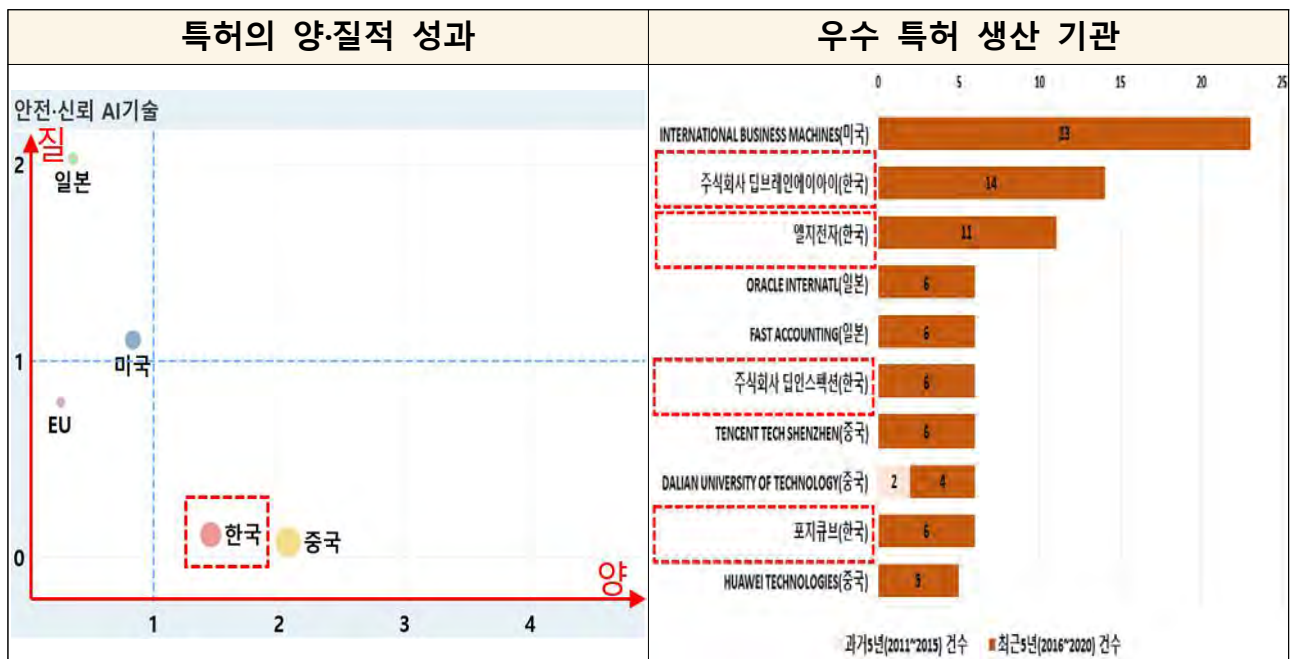
* Stanford Institute for Human-centered Artificial Intelligence, AI Index 2024(2024)

- 2022 기술수준평가 결과에도 특허는 양적인 측면에서 우위에 있으며, Top 10 특허생산기관에도 한국의 기관이 상당수 포함되어 있음



[그림] AI Index 24(좌), Global AI Index 2024(우)에 나타난 한국의 특허 역량

※ 출처 : <https://www.tortoisemedia.com/intelligence/global-ai>



[그림] 2022 기술수준평가에서 특허 분석 결과

□ 정책 동향

- 글로벌 측면에서 생성형 AI의 부상 이후 안전과 신뢰는 핵심정책 과제이며, 각국은 규제 강화, 국제공조 등을 통해 정책 대응 중
 - (EU AI Act, '24.5) AI의 위험도를 4단계*로 분류하고, 대부분의 안전·신뢰의무를 고위험 AI 제공자에 부여 등 규제 지향적 정책 전개
 - * 금지된 위험(Prohibited Risk), 고위험(High-Risk), 제한된 위험(Limited Risk), 최소 위험(Minimal or Low Risk) 등 4단계
 - (미 백악관 행정명령, '23.10) 생성형 AI의 위험이 인식됨에 따라, NIST*가 안전 및 보안을 위한 새로운 표준을 수립하고, 국토안보부, 상무부, 국무부 등이 협력하여 국제 프레임워크를 구축
 - * 국립표준기술연구소(National Institute of Standards and Technology)
 - (서울 AI 정상회의, '24.5) AI의 안전성, 혁신, 포용성을 중심으로 국제 협력을 강화하는 등 글로벌 AI 거버넌스의 방향을 제시
 - * AI 안전연구소 국제네트워크 구축, 기업 참여, 위험평가체계 개발 합의 등 성과
- 한국은 글로벌 AI 지수에서 정부전략 측면은 88.5점으로 높은 점수를 보이고 있으나, 입법 관련 점수는 50.9점으로 상대적으로 저조

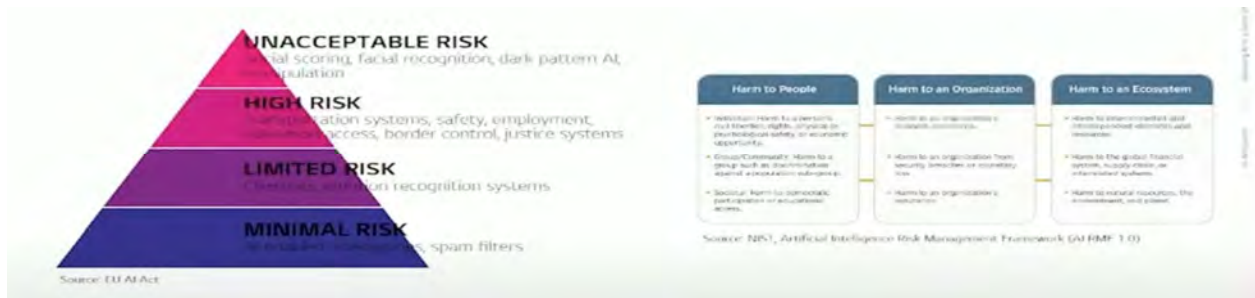


※ 출처 : Global AI Index 2024

각국의 정책 대응

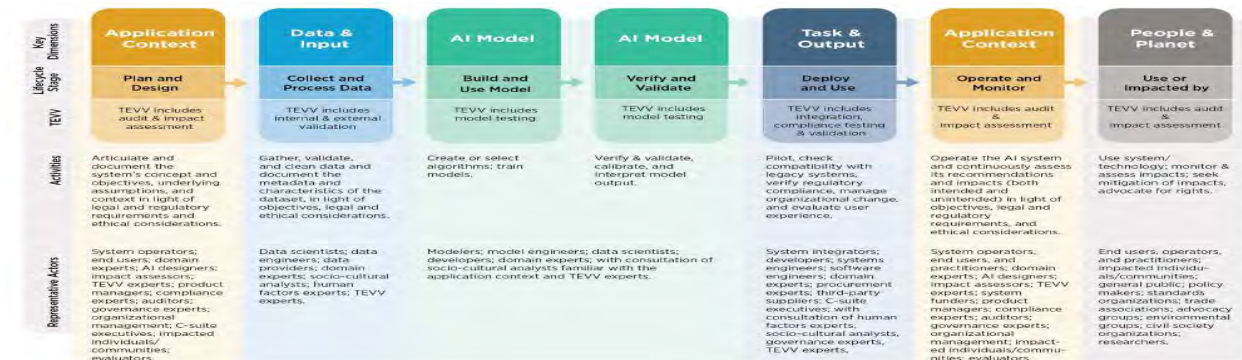
○ EU AI Act는 인공지능의 위험도에 따라 인공지능을 분류

- 허용할 수 없는 위험(예: 채점 시스템 및 조작가능한 AI) 금지하며, 최소한의 위험은 규제되지 않음



○ 미 백악관 행정명령은 안전 및 보안을 위한 새로운 표준 수립을 지시

- 미 상무성 산하 NIST는 '23.1월에 발표한 AI 개발 단계별로 권장 사항을 담은 AI Risk Management Framework 1.0을 토대로 강화된 규제 기준을 수립할 것으로 예상



○ 한국은 서울AI 정상회의에 발맞춰 '새로운 디지털 질서 정립 8대 핵심 과제'를 발표



□ 평가 내용

- **(평가대상)** 총괄위원회에서 시의성을 고려하고, AI 관련 타 평가제도와 중복을 지양하면서, 상호보완적 평가가 될 수 있는 세부 기술을 선정

- 안전·신뢰 AI 기술이 적용되는 기술 분야를 평가하되, 기술영향평가의 취지를 고려*하여 신기술 분야를 평가

* 과학기술기본법 제14조 : 새로운 과학기술의 발전이 미치는 영향을 사전에 평가

- 지능정보기술은 기술영향평가로 수행하도록 한 관련 법 취지*를 고려하여 '사회적 영향평가'에서 수행된 평가 분야 중 일부를 활용

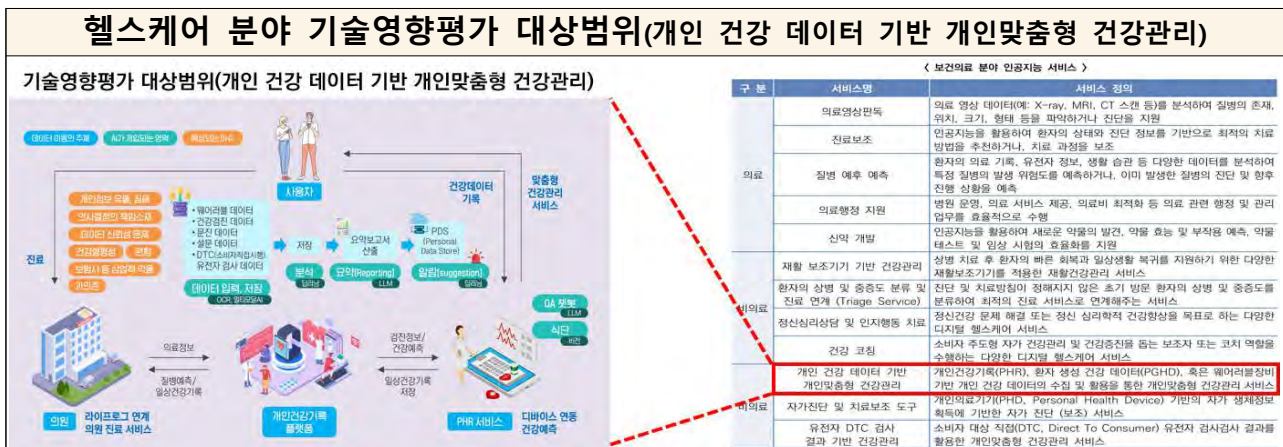
* 과기기본법 및 지능정보화기본법에 따라 인공지능이 적용되는 채용, 금융, 교통, 교육 등 지능정보서비스의 영향은 '사회적 영향평가'에서, 인공지능(AI)의 안전·신뢰 기술 등 지능정보기술의 영향은 '기술영향평가'에서 수행

※ 참고 : 지능정보화 기본법 제56조(지능정보서비스 등의 사회적 영향평가)

- **(세부기술)** 헬스케어, 휴머노이드, 브레인(BCI) 분야에 적용되는 안전·신뢰 AI 기술의 영향을 평가

- **(헬스케어 AI 안전·신뢰)** 대중적으로 접근성이 높은 기술 서비스인, 「개인 건강 데이터 기반 개인 맞춤형 건강관리 서비스*」를 평가

* 인공지능 서비스 사회적 영향평가(NIA, 2023) 평가대상 중 보건의료의 '개인건강 데이터 기반 개인맞춤형 건강관리 서비스'를 연계 및 확장하여 기술영향평가 실시



- (휴머노이드 AI 안전 · 신뢰) 기대가 증진되고 있는 옵티머스(테슬라), Figure AI 등 파운데이션모델을 탑재한 인간형 로봇* 중심으로 논의

※ '휴머노이드'에 대한 명확한 정의는 없지만, 인간의 형태를 하면서 인간이 하는 다양한 동작과 역할을 수행하는 로봇으로 지칭됨

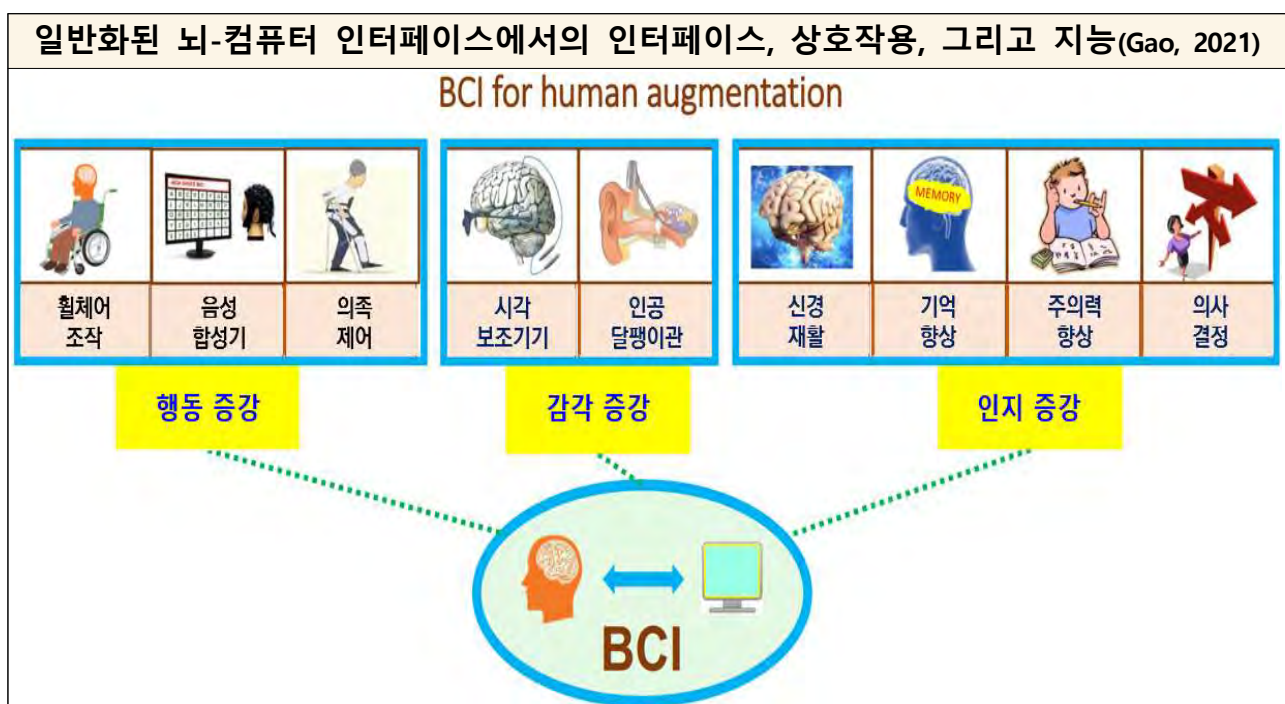
연구자	출처	연도	휴머노이드 기술 정의
김종찬 외	Sustainability	2016	휴머노이드란 인간의 신체와 유사한 로봇을 말하며, 인간에게 다양한 형태의 서비스를 제공하기 위해 인간의 지능, 행동, 감각, 상호작용 등을 모방한 로봇
Dragomir N. Nenchev 외	Humanoid Robots	2019	휴머노이드는 인간의 형태와 동작을 모방하는 로봇 기술로 인간과 소통하여 인간을 위험에 빠뜨리지 않고 인간과 신체적 접촉 수행하면서 도구를 조작하고 인간을 위해 설계된 물체를 조작하는 기술
Raouf Fareh 외	Mechatronics	2021	휴머노이드는 인간의 모습을 흉내내도록 설계된 로봇

- (브레인 AI 안전 · 신뢰) AI 기술이 적용된 '뇌-컴퓨터 인터페이스 기술'

⇒ 뇌신경질환 치료를 넘어 인간의 인지능력을 확장할 것이라는 기대 증대

* Brain Computer Interface(BCI)

※ 2016년 일론 머스크가 창업한 '뉴럴링크'는 원숭이가 생각만으로 비디오 게임을 하는 것을 시연(2021)하고, 인간의 뇌에도 칩을 이식(2024)하는 기술을 선보임



□ 평가 방법

- **(이슈 트렌드 분석)** 토픽 모델링 기법을 통해서 안전·신뢰 AI 및 대상기술 관련 뉴스 빅데이터 내 주요 이슈 도출 후 이벤트 분석, 키워드 분석, 웹 반응 분석 등 분석 진행

※ 안전·신뢰 AI 관련 선행연구를 통한 키워드 사전 구축 및 검색 쿼리 설정

※ 분석대상 기간(2014년 7월 1일 ~ 2024년 9월 1일)

※ 분석대상 데이터 : 총 수집 데이터(53,294건) 중 전처리 후 27,420건 사용

- **(주요 이슈·과제 도출)** 전문가·시민포럼이 제시한 기술에 대한 이슈와 과제를 분석하여 주요 이슈 및 과제로 도출

※ 이슈 간 중복성 제거, 과제 간 유사성 등을 고려

- **(전문가 심층분석)** 분야별로 전문가 1차 회의에서 브레인스토밍을 통해 이슈를 도출 ⇒ 2차, 3차 회의에서 논의를 거쳐 정책 대안 마련

- **(구조화된 시민포럼)** 시민포럼 독립성 강화 및 과정의 구조화*를 통해 몰입을 강화하여 시민 관점의 이슈 및 시민이 원하는 정책대안 발굴

* 시민포럼에게 미션을 통한 상황 몰입 후에 심층 면접을 통한 심도 있는 토론 등 구조화된 방식을 적용하여 생생한 시민 의견을 도출('24년 시범 추진)

- **(핵심 이슈·과제 도출)** 설문조사를 통해 핵심 이슈·과제를 도출

- **(설문조사)** 도출된 주요 이슈 및 과제에 대한 설문조사(전문가 125명, 시민 1,000명) 통해 정책분석 기준을 적용·평가하여 핵심 이슈·과제를 도출

분류	전문가 조사	일반 시민 조사
조사개요	▷ 조사 대상 : AI 관련 전문가 (461명) *AI, 헬스케어, 휴머노이드, 브레인 기술, 미래예측·정책전문가로 구성 ▷ 조사 기간 : 2024년 11월 13일~19일 ▷ 조사 참여 : 125명	▷ 조사 대상 : 패널을 활용한 온라인 조사 * 성별, 연령, 지역별 인구비례 ▷ 조사 기간 : 2024년 11월 13일~19일 ▷ 조사 참여 : 1,000명
핵심이슈 (평가기준)	▷ - 발생가능성 / 영향력(파급효과) - 발생예상시기	▷ - 발생가능성 / 영향력(파급효과) - 발생예상시기
핵심과제 (평가기준)	▷ - 중요성 / 시급성 - 효과성 / 실행가능성 (전문가에게만 질문)	▷ - 중요성 / 시급성



[그림] 평가단계별 평가방법

빅데이터 분석

○ 안전·신뢰 AI 관련 선행연구를 통한 키워드 사전 구축 및 검색 쿼리 설정

- 빅카인즈 활용 데이터 수집 및 데이터베이스 구축

※ 분석대상 기간(2014년 7월 1일 ~ 2024년 9월 1일)

※ 분석대상 데이터 : 총 수집 데이터(53,294건) 중 전처리 후 27,420건 사용

※ 언론보도를 통해 핵심 키워드를 분석하고 전문가·시민포럼에게 정보제공 목적

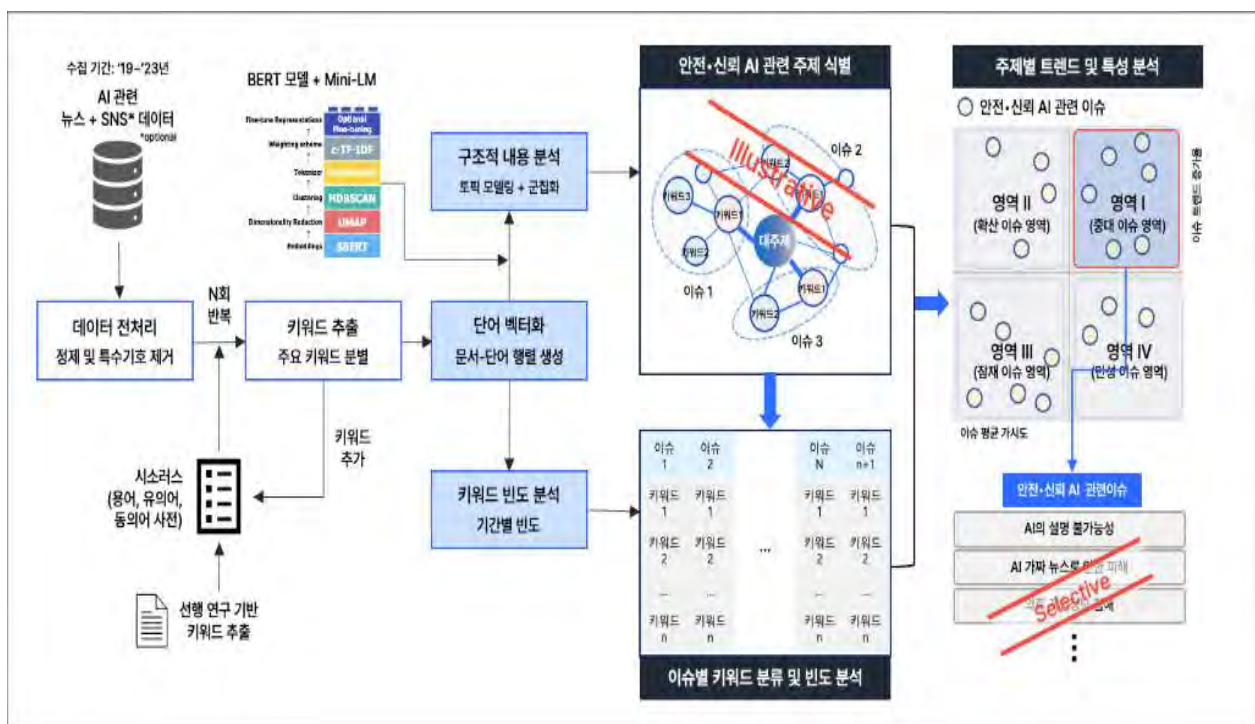
① 토픽 모델링 활용 안전·신뢰 AI 관련 주요 이슈 도출 및 주제별 키워드 분석

② 안전·신뢰 AI 관련 미디어 데이터 내 주요 이벤트 분석

※ 주요 이벤트(사건)는 주변 7일(-3일~+3일) 평균에 비해 보도량이 높게 나타난 (이상 보도량) 날짜를 중심으로 특정 이슈에 대한 보도가 중복적으로(홍보성 기사, 대선 공약 등 제외) 나타나는 일자의 내용을 분석

③ 3개 기술분야별 주요 이벤트에 대한 웹 반응(감성) 분석

※ Google API를 활용해 해당 기술과 안전·신뢰 이슈를 다루는 주요 유튜브 콘텐츠와 해당 콘텐츠에 달린 대중의 댓글 데이터를 수집하여 키워드 분석을 기반으로 인식 및 감정을 분석



[그림] 토픽 모델링 기반 주요 이슈 도출 개요도

구조화된 시민포럼

- 시민포럼 독립성 강화 및 과정의 구조화를 통해 몰입을 강화하여 시민 관점의 중요 이슈 및 시민이 원하는 정책대안 발굴
 - (토론과정 구조화) 시민포럼 참여자 시민들에게 미션 부여*하고, 숙의형 포커스그룹 토론(dFGD**) 방식을 적용하여 몰입감 강화
 - * 시민포럼 참여자들이 미션에서 작성한 시나리오를 활용하여 다음 미션부터 상황 몰입 및 주어진 질문에 회신 ⇒ 10회의 미션을 통해 몰입을 제고 후 워크숍 실시
 - ** Deliberative Focused Group Discussion : 참가자들에게 교육을 제공하고 토론을 구조화함으로써 집단사고 문제를 완화하고 정보에 입각한 균형 잡힌 결과를 도출
 - (독립성 강화) 독립적*으로 운영하여, 시민이 전문가에 포획(capture) 되는 현상을 방지하고 전문가가 놓칠 수 있는 시민만의 이슈를 포착
 - * 이전 시민포럼 대비 전문가와의 동화작용을 최소화하고, 몰입 강화 및 단계적 사고 과정을 거쳐 일상생활 속에서의 시민 관점이 반영된 기대와 우려를 포착

01 프로그램 설계& 리크루팅

구분	내용	인원	비고
총괄	다양한 의견의 반영을 수 있도록 신중히 선정, 최종 10명 선정	10	
지원	다양한 산업 종사자로 특정 분야에 국한되지 않도록 함	10	
전문가	과학기술, 환경, 의료, AI, 사회문제 등과 관련된 전문가나 시민으로서 선정	10	
조언자	과학기술, 환경, 의료, AI, 사회문제 등과 관련된 전문가나 시민으로서 선정	10	
조언자	과학기술, 환경, 의료, AI, 사회문제 등과 관련된 전문가나 시민으로서 선정	10	

- 총 119명 신청
- 성별, 연령, 지역, 직업 등의 대표성과 다양성, 심화질문에 대한 성실도를 평가하여 15명의 참여자 선발

02 오리엔테이션



- 기술영향평가 개요, 헬스케어 AI, 브레인 AI, 휴머노이드 AI 기술 강연
- 기술영향평가의 의미를 알리고 참여도를 높이며 공동의 목표를 설정

03 미션 수행

미션	내용	인원	비고
미션 1	과학기술의 발전과 함께 사회 문제 해결을 위한 방안 모색	10	
미션 2	과학기술의 발전과 함께 사회 문제 해결을 위한 방안 모색	10	
미션 3	과학기술의 발전과 함께 사회 문제 해결을 위한 방안 모색	10	
미션 4	과학기술의 발전과 함께 사회 문제 해결을 위한 방안 모색	10	
미션 5	과학기술의 발전과 함께 사회 문제 해결을 위한 방안 모색	10	
미션 6	과학기술의 발전과 함께 사회 문제 해결을 위한 방안 모색	10	
미션 7	과학기술의 발전과 함께 사회 문제 해결을 위한 방안 모색	10	
미션 8	과학기술의 발전과 함께 사회 문제 해결을 위한 방안 모색	10	
미션 9	과학기술의 발전과 함께 사회 문제 해결을 위한 방안 모색	10	
미션 10	과학기술의 발전과 함께 사회 문제 해결을 위한 방안 모색	10	

- 일반시민들이 AI 기술의 영향에 대한 맥락을 이해하고 몰입할 수 있도록 미션을 부여
- 1개월간 10개의 미션을 부여 (오전 미션 부여, 오후 회신)

04 시민참여 포럼



- 참여자들이 숙고를 바탕으로 다양한 관점을 이야기할 수 있도록 워크숍 진행
- AI기술이 우리 삶에 미치는 긍정/부정적 영향을 심도 있게 토의하고 필요한 정책 및 규제 아이디어를 도출

05 결론 도출

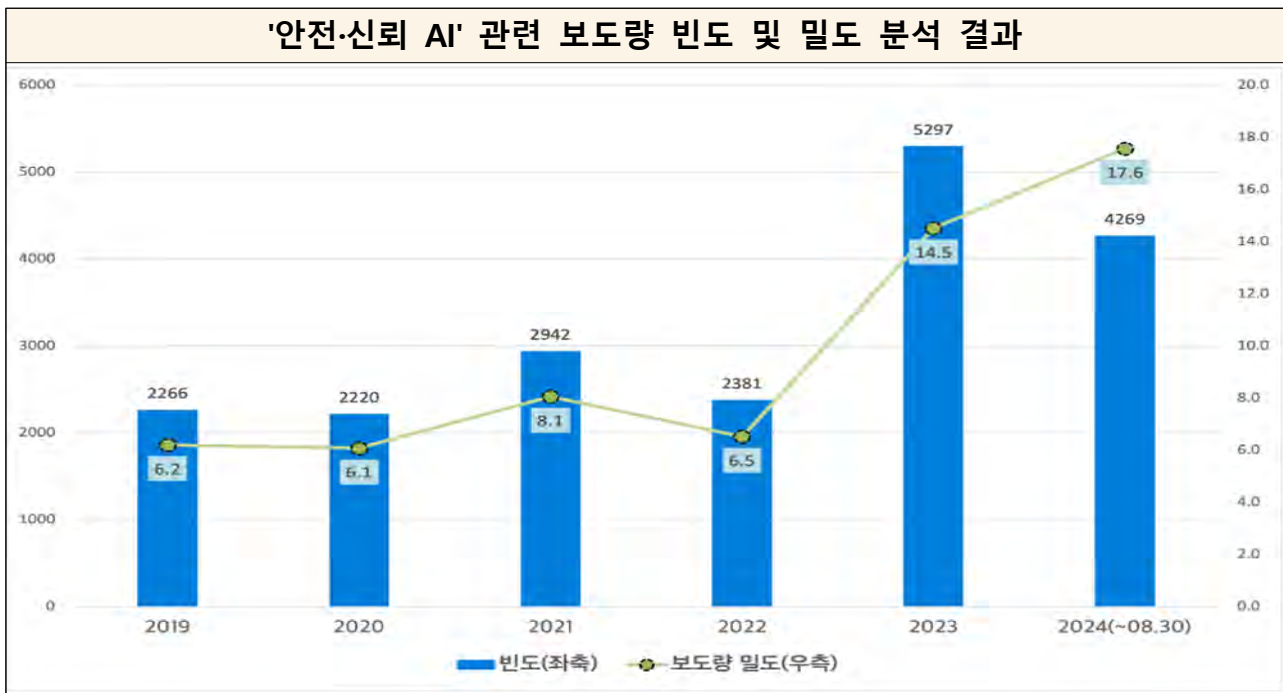
구분	내용	인원	비고
총괄	다양한 의견의 반영을 수 있도록 신중히 선정, 최종 10명 선정	10	
지원	다양한 산업 종사자로 특정 분야에 국한되지 않도록 함	10	
전문가	과학기술, 환경, 의료, AI, 사회문제 등과 관련된 전문가나 시민으로서 선정	10	
조언자	과학기술, 환경, 의료, AI, 사회문제 등과 관련된 전문가나 시민으로서 선정	10	
조언자	과학기술, 환경, 의료, AI, 사회문제 등과 관련된 전문가나 시민으로서 선정	10	

- 오리엔테이션, 10회의 미션, 시민참여 포럼에서 논의된 내용을 바탕으로 일반 시민들이 생각하는 AI의 안전성과 신뢰성 정의, 기대와 우려, 정책제안 정리

[그림] 2024 기술영향평가 시민포럼 진행 절차

1 언론보도 이슈 트렌드 분석

- (언론보도 양적 추이) ChatGPT가 최초 발표된 2022년(11월 30일) 이후 '23년은 약 5,300건의 뉴스가 보도되었고, '24년에는 6,000건을 상회 예측



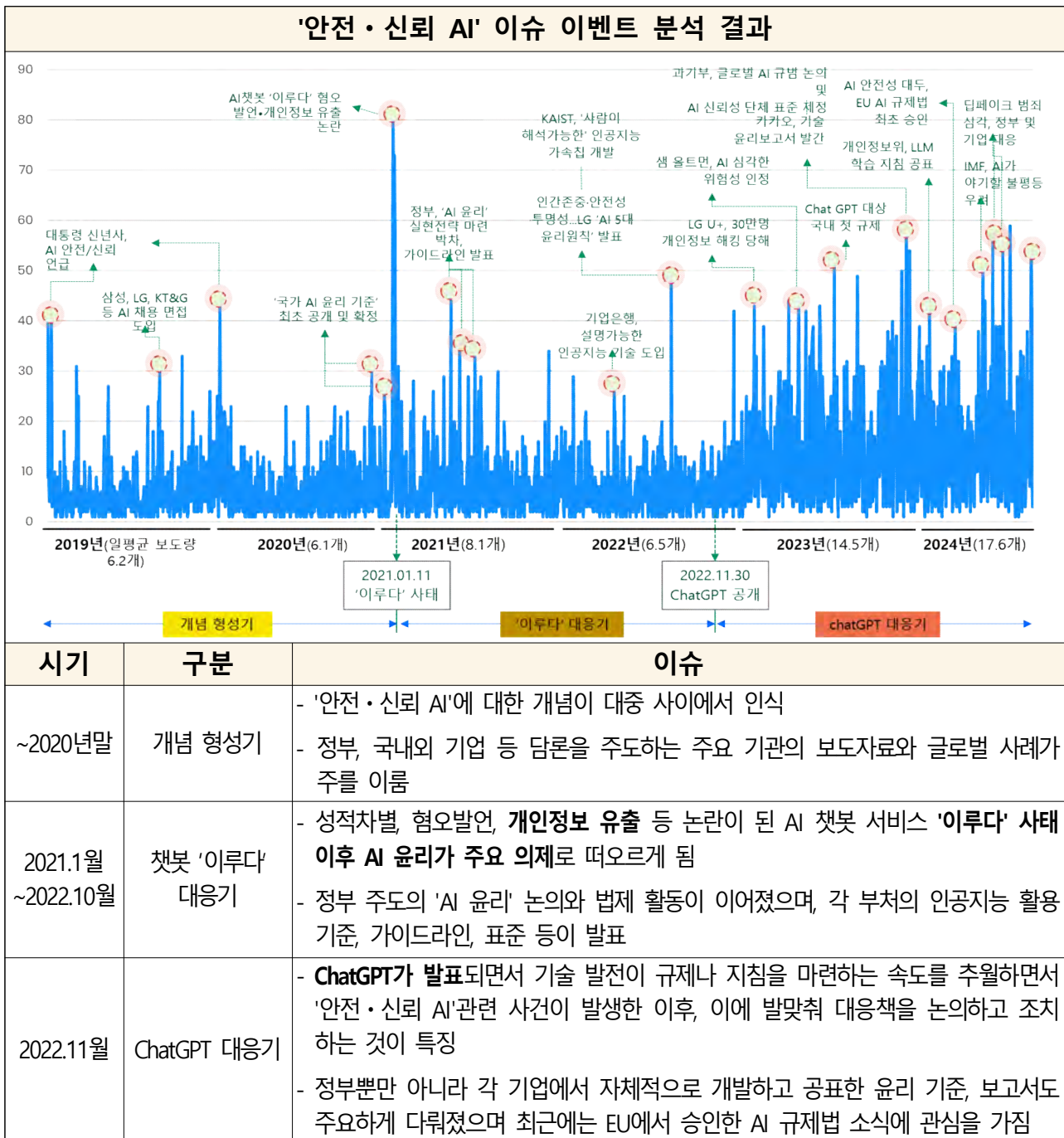
- (이슈 흐름) 챗봇 '이루다' 이후 개인정보 보호 및 정부 정책 중심으로의 전개에서 ChatGPT 등장 이후 안전·신뢰·윤리와 기업 중심 이슈로 전환

- 언론보도 데이터를 통해 각 이슈별 트렌드에 영향을 미친 주요 사건 (트리거 이벤트)과 연도별 안전·신뢰 AI 관련 핵심 키워드를 분석

※ 챗봇 '이루다' : 딥러닝 알고리즘을 이용한 AI 챗봇 서비스로 '20.12월 공개되었으나, 개인정보보호위원회가 '21.4월에 무분별한 개인정보처리에 대해 과태료 및 과징금 부과

※ 연도별 키워드 분석은 TF-IDF 기준 상위 10개 키워드를 선정했으며, '개발', '사업', '계획'과 등의 일반적인 키워드는 제외

※ TF-IDF(Term Frequency-Inverse Document Frequency) : 특정 단어가 문서 내에서 얼마나 중요한지를 측정하는 방법이며, 단어가 해당 문서에서 얼마나 자주 등장하는가(TF)와 전체 문서 집합에서 얼마나 흔하지 않은지(IDF)를 종합하여 도출



○ **(키워드 트렌드)** '개인정보'는 가장 높은 빈도로 나타나는 키워드로, 2019년부터 '개인정보'와 '보안'과 관련된 언론의 관심이 지속적으로 증가

- 2024년 기준으로 '윤리'보다 '안전', '신뢰'의 빈도가 높아져 AI 윤리에 대한 의제가 안전·신뢰로 구체화되고 있음을 시사
- 안전·신뢰를 실현하기 위한 방법으로 '규제', '인증', '표준', '책임', '기준' 등 지속적으로 증가했으며, 특히 '인증(51%)'과 '표준(48%)'에서 증가

키워드 트렌드 분석 결과							
키워드	2019	2020	2021	2022	2023	2024 (~08.30)	연평균 증가량
개인정보	18.2	30.7	73.0	29.6	68.6	61.8	28%
보안	19.2	16.4	24.6	15.2	38.8	31.0	10%
규제	10.2	10.4	11.8	12.3	38.5	27.7	22%
안전	7.0	9.3	13.8	12.2	25.4	24.7	29%
신뢰	4.0	5.9	12.9	7.4	29.8	21.8	40%
윤리	4.8	11.1	27.1	22.4	32.6	19.5	32%
인증	2.5	5.5	7.6	6.5	24.5	19.2	51%
표준	2.6	6.6	4.3	3.5	20.2	19.0	48%
책임	3.5	5.5	8.0	5.6	18.3	13.9	32%
기준	3.6	12.8	11.7	8.6	13.5	13.6	31%

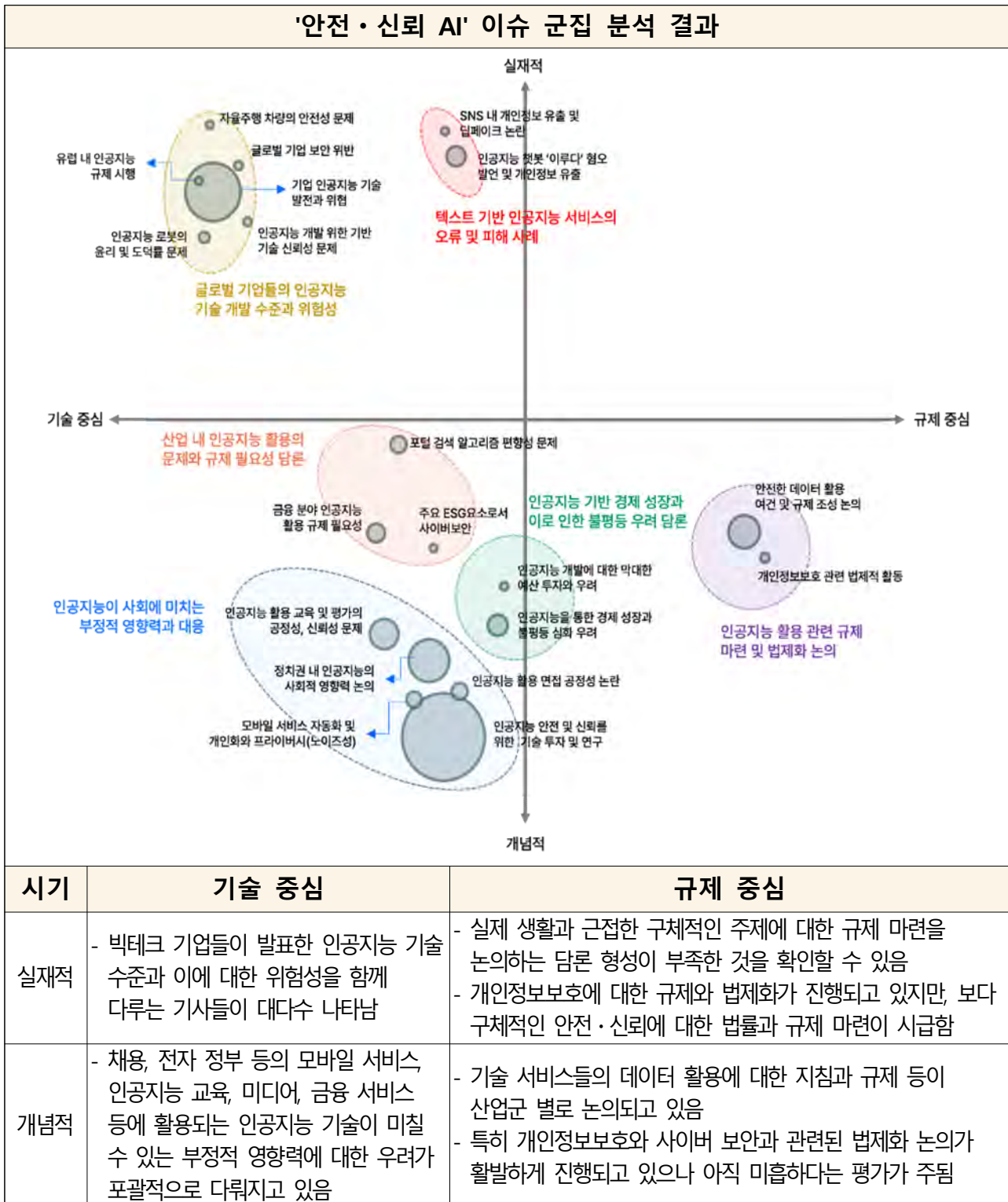
- **(키워드 이슈 분석)** '인공지능 안전 및 신뢰를 위한 기술 투자 및 연구'(43.4%), '생성형 AI 기술과 위협'(19.4%) 관련 비율이 상승

안전·신뢰 AI 관련 10대 이슈 목록			
No.	이슈명	주요 키워드	비율
1	인공지능 안전 및 신뢰를 위한 기술 투자 및 연구	개인정보, 투자, 윤리, 보호, 사업	43.4%
2	생성형 인공지능 기술과 위협	생성형 인공지능, 인간, 인류, 윤리, 과학	19.4%
3	장치권 내 인공지능의 사회적 영향력 논의	의원, 정치, 선거, 연설, 회견	10.4%
4	안전한 데이터 활용 여건 및 규제 조성 논의	개인정보보호, 보안, 인증, 사이버, 유출, 침해	6.3%
5	인공지능 활용 교육 및 평가의 공정성 신뢰성 문제	교육, 대학, 학교, 학생, 교육부, 교육청	5.2%
6	인공지능 기술을 통한 경제 성장과 격차 심화 우려	평등, 성장, 위기, 일자리, 소득, 변화	2.5%
7	인공지능 챗봇 '이루다' 혐오 발언 및 개인정보 유출 논란	대화, 혐오, 논란, 생성형, 개인정보, 유출	2.3%
8	금융 분야 인공지능 활용 규제 필요성	금융, 은행, 보험, 신용, 가이드라인, 규제	2.0%
9	포털 검색 알고리즘 편향성 문제	포털, 알고리즘, 미디어, 기사, 언론, 정치	1.6%
10	인공지능 활용 채용 공정성 논란	채용, 신입, 직원, 모집, 직무, 면접	1.6%

- 이슈들을 두 개의 차원으로 축소·군집화*한 결과, 실재에 근접한 구체적인 규제 마련을 위한 담론 형성은 부족

* 수집한 텍스트 데이터를 UMAP((Uniform Manifold Approximation and Projection for Dimension Reduction) 알고리즘을 활용하여 두가지 임의의 차원으로 축소하고 군집화

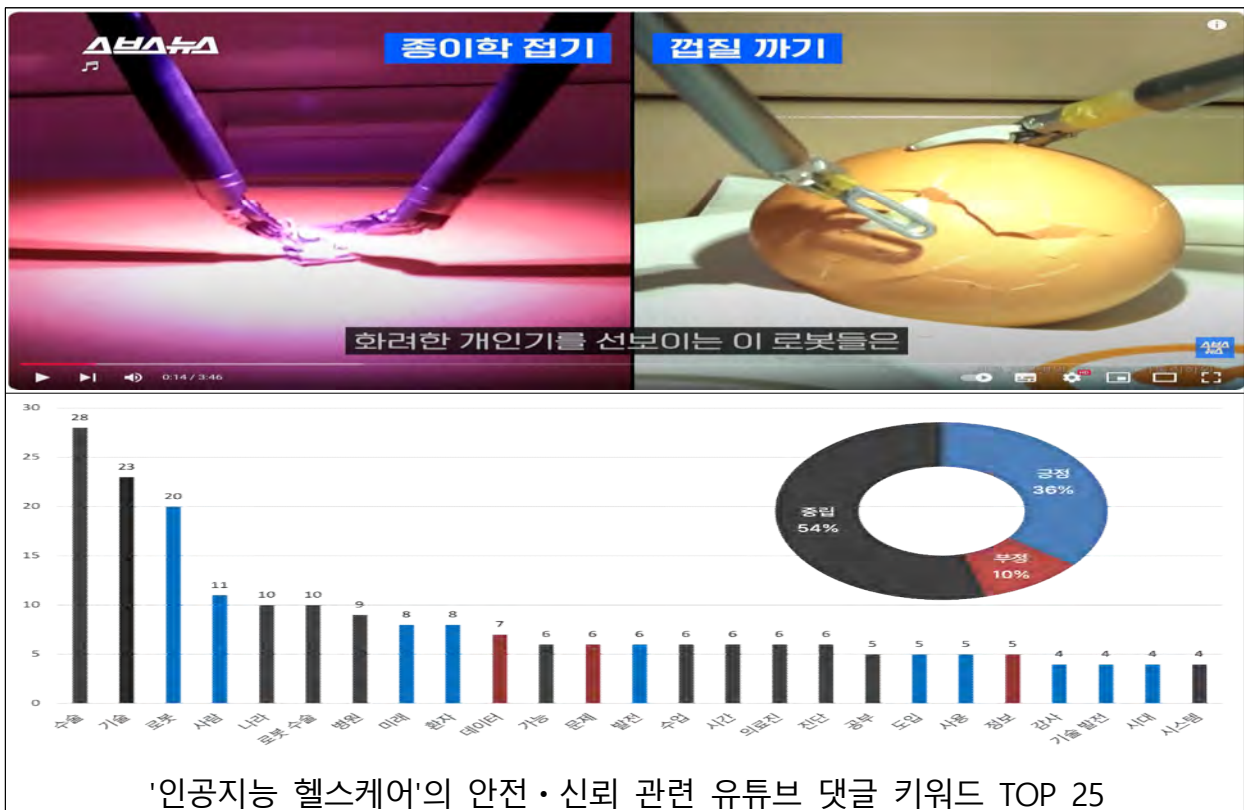
※ '주제의 실재성(개념-실재)'과 '논의의 방향성(기술 중심-규제 중심)'으로 군집이 형성



SNS(유튜브) 이슈 트렌드

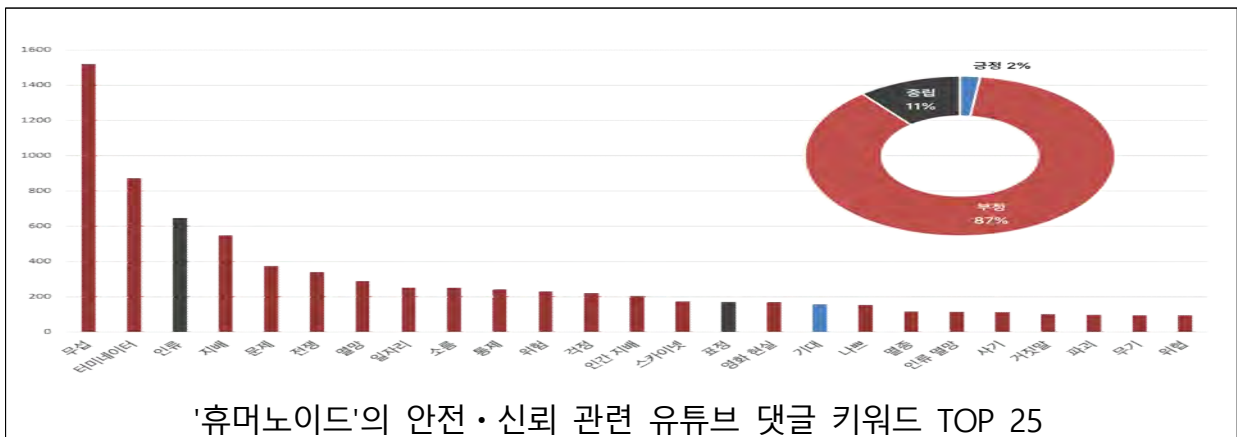
□ 헬스케어 AI 안전 · 신뢰 : 중립적, 긍정적

- 인공지능 활용 헬스케어 기술에 대한 대중의 반응은 중립적인 것으로 나타나며, '기술 발전' 속도에 감명하는 긍정적인 의견 제시
- 특히 '수술'을 돕는 '로봇'에 대한 기대가 높은 것으로 분석되며, '사람'보다 더 신뢰하여, '빠른 도입'을 원하는 의견 제시



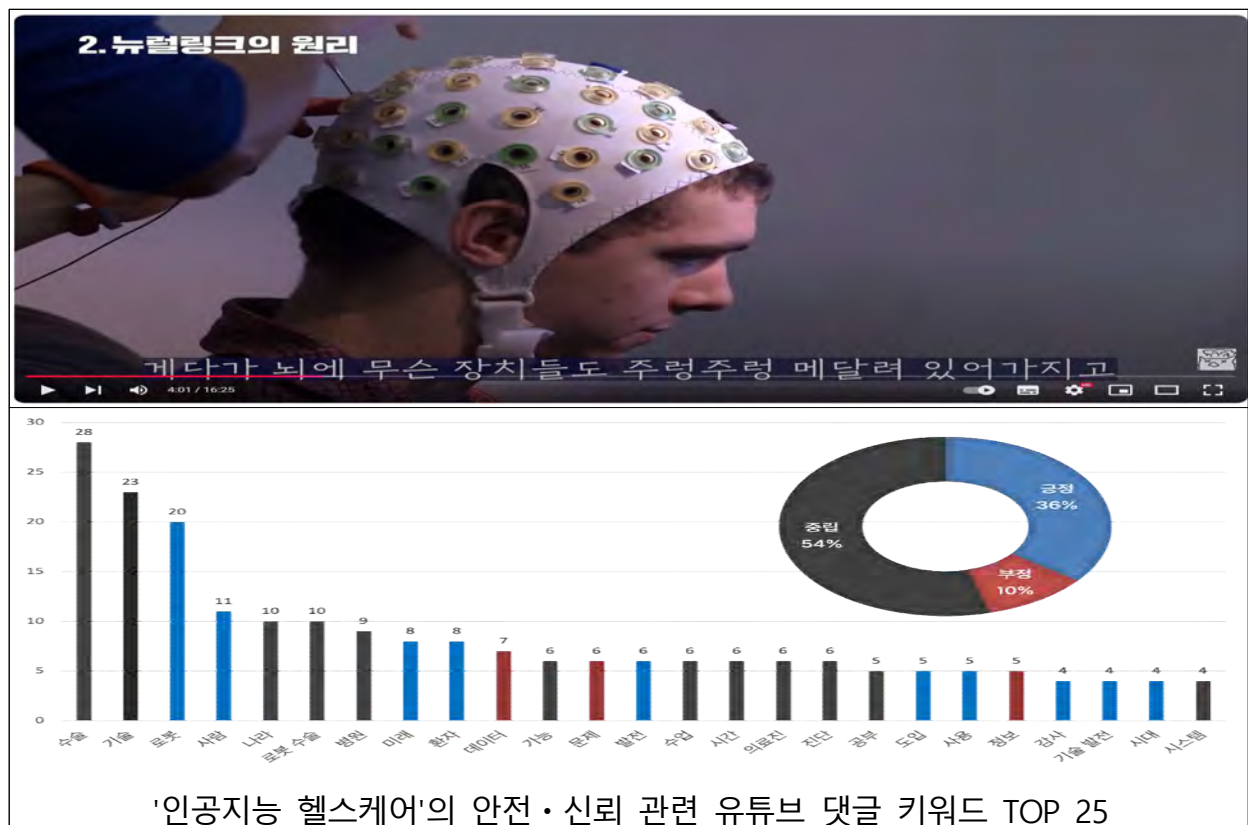
□ 휴머노이드 AI 안전 · 신뢰 : 인간 실존의 위험

- '휴머노이드' 안전 · 신뢰와 관련된 조회수 상위의 유튜브 댓글을 분석한 결과, 영화 이미지와 맞물려 부정적 감성이 지배적
- '일자리 대체'나, '사기', '무기' 등으로 악용되는 것을 우려



□ 브레인 AI 안전 · 신뢰 : 마비환자의 희망 vs 인간통제

- 대중은 마비 환자에게 희망을 줄 것이라는 기대도 있으나, 뇌를 해킹하고 개인을 조종, 통제할 것이라는 우려
- '마비 환자'들에게 '희망'과 '자유'를 줄 것이라는 긍정적 반응이 나타났으나, '기억 왜곡', 또는 '신호 체계 통제' 등 우려



② 주요 이슈 및 과제 분석

① 전문가 브레인스토밍

□ [헬스케어 AI 분야 전문가] 이슈 및 정책과제

- 데이터 신뢰 및 책임소재 이슈, 개인정보 및 격차 등 사회적 문제 이슈, 과의존 이슈를 우려
 - 빅데이터 모델 생성과정 및 데이터의 품질, 보안, 프라이버시 등 기술적 요인을 주요 이슈로 파악
 - ※ 시민들은 역차별 및 개인정보 유출을 감수할 만한 서비스 부재를 주요 이슈로 봄
 - 세부적으로 안전과 관련된 기술이므로 데이터의 품질, 보안, 프라이버시 등 신뢰성과 AI 데이터의 편향 및 AI 과의존을 중요 이슈로 판단

[표] 헬스케어 AI 분야 전문가 이슈

이슈유형	이슈(9개)		세부이슈(30개)
기술적 이슈	헬스케어 빅데이터 모델 생성과정의 이슈		사전 동의만 허용되는 데이터 수집 단계
			가명화된 데이터 활용의 제한성
			외부 네트워크와 차단된 데이터 안심존의 실효성 문제
			법적, 제도적 규제의 중첩 적용으로 인한 혼선
	의사결정의 책임소재		의학적 판단을 AI가 할 수 있는가?
	데이터 신뢰성		데이터 품질과 완전성
			데이터 편향
			데이터 보안과 프라이버시
			생성형 AI의 할루시네이션
			데이터 통합과 표준화 문제
사회적 이슈	의료 AI와 건강 격차, 건강 형평성	건강형평성에 대한 기술적 요인	데이터 표현(편향)
		건강격차에 대한 사회적 요인	AI 기반 의료 도구에 대한 접근성
			설계 및 구현
			디지털헬스리터러시와 연령주의가 의료 AI 형평성에 미치는 영향(노인)
			언어장벽과 인터넷 접근성이 의료 AI 형평성에 미치는 영향(이주민, 취약계층)
			의료 AI 솔루션의 활용과 제공 방식(지방과 지역 주민)

이슈유형	이슈(9개)		세부이슈(30개)
	민간기업의 상업적 활용과 책임	개인건강정보 상업적 활용의 주요 이슈	건강 데이터 생태계에 대한 수요자 및 국민의 신뢰
			가입자의 특성과 서비스 품질 관리
			민감 정보에 대한 보안 문제
	개인정보 유출·침해	헬스케어 AI 도입, 확산과정에서의 개인정보 유출과 프라이버시	의료건강 데이터 보안과 사생활 침해
			AI Healthcare 이용자의 데이터 이용 동의와 소유권
			다종 데이터 결합과 통합에 따른 개인정보보호와 법적, 윤리적 문제
	편향, 과의존	AI 데이터의 편향	인구통계학적 샘플링 편향
			사회적 결정요인의 측정 편향
			과거 데이터로부터의 역사적 편향
		AI에 대한 과의존을 초래할 수 있는 주요 문제	AI 제안을 선호하는 자동화 편향
			AI의 설명에 대한 과신
			복잡한 작업에 대한 친숙도 및 전문성
			AI시스템의 좁은 작업 초점
			과잉 의존과 치료 정확성 저하

[표] 헬스케어 AI 분야 전문가 정책과제

구분	이슈(9개)		정책과제(45개)
기술적 위험	헬스케어 빅데이터 모델 생성과정의 이슈		포괄적, 단계적 등 유연한 동의 체계 도입
			데이터 활용 절차 간소화를 위한 예비 분석 허용
			제한적 외부접속을 허용하는 데이터 안심존 개선
			단계적, 중장기적인 법, 제도 정비
	의사결정의 책임소재		자율주행 자동차의 법제화 과정을 살펴보면서 해당 사례를 참고
			의료 데이터 품질 관리체계 구축
			데이터 편향성 모니터링 및 보정 시스템 구축
			의료 데이터 보안 및 프라이버시 보호 강화
사회적 위험	의료 AI와 건강 격차, 건강 형평성	건강 형평성에 대한 기술적 요인	생성형 AI의 의료 분야 활용 가이드라인 수립
			의료 데이터 표준화 및 통합 플랫폼 구축
			포괄적인 데이터 수집 및 큐레이션
			AI 기술에 대한 공평한 접근
			건강형평성을 고려한 AI 기술 설계
			규제 감독 및 책임
		건강격차에 대한 사회적 요인	대중 및 지역사회 참여
			모니터링 및 평가
			양질의 데이터 확보
			디지털헬스리터러시 역량 강화
			투명성 확보, 참여기반 정책 결정
			양적으로 큰 실증 연구 수행 및 평가
			커넥티드 헬스와 네트워크 개선
			의료문제 해결을 위한 수요기반 의료 AI 개발

구분	이슈(9개)		정책과제(45개)
	민간기업의 상업적 활용과 책임		이해관계자간 신뢰 획득
			차별화된 정보보안강화
	개인정보 유출·침해		데이터 프라이버시 및 보안 강화를 위한 엄격한 암호화 기준과 접근 제어 정책
			쉬운 사전 동의 절차
			데이터 비식별화 기술의 고도화 및 연합학습
			자기 데이터의 소유와 통제에 대한 권한 부여
			장기 데이터 관리 가이드라인 제시
			국제 데이터 이동 관련 협약 마련
			이해관계자간 협력
	편향, 과의존	데이터 편향 해소를 위한 방안	포괄적인 기술 영향 평가(TIA)
			다양하고 포괄적인 데이터 수집
			정기적인 편향 감사 및 투명한 보고
			편향 교정 알고리즘
			헬스케어 AI 윤리 표준 및 개발자, 의료전문가에 대한 교육
			AI시스템 설계 및 배포에 환자 및 대중 참여
			권한있는 정부 및 기관의 책임 있는 감독
			디지털 건강기술에 대한 접근성 향상 정책
		헬스케어 AI 과의존 해소 방안	헬스케어 AI 시스템 개발, 검증, 배포에 대한 규제 프레임워크 및 표준
			의료 제공자, 개발자의 책임과 의무 명확화
			AI 모델 개발시의 투명성 및 설명 가능성 강화 설계
			임상외에 대한 AI의 기능, 한계, 사용법에 대한 위한 지속적인 교육 및 훈련
			헬스케어 AI시스템에 대한 실시간 피드백 및 모니터링
			다양한 임상 작업의 복잡도를 고려한 맞춤형 AI 개발

□ [휴머노이드 AI 분야 전문가] 이슈 및 정책과제

- 물리적 안전 이슈, 프라이버시 이슈, 인간과의 상호작용 이슈를 우려
 - 세부적으로 인간과 휴머노이드, 휴머노이드 간 의사소통을 중요 이슈로 판단
 - 이에 따라 디자인 및 인지능력, 감정적 유대 등 인간과 휴머노이드 간 상호작용 시 윤리적 문제 발생 가능
 - 전문가들은 물리적 안전뿐만 아니라 윤리적 도전을 주요 이슈화

※ 시민들은 일자리 대체에 대한 기대 및 우려와 사회적 문제를 주요한 이슈로 봄

[표] 휴머노이드 AI 분야 전문가 이슈

이슈유형	이슈(5개)		세부이슈(19개)
기술적 이슈	휴머노이드 AI의 물리적 안전		개념 검증(Poof of Concept)
			기능 안전과 인공지능 윤리
			새로운 시각의 규제 접근
			자율주행차 방식(턴키 방식) 의 규제 해소
			사회적 담론등 가치적 관점의 법제화
			등록제 등 다층적 규제시스템
	인간-휴머 노이드, 휴머노이 드-휴머노 이드 간 의사소통	인간-휴머 노이드간 의사소통	의사소통의 오해
			정보의 불완전 전달
			편향된 의사결정
		휴머노이 드 상호간 의사소통	인공지능 수준차이
상호운용성 부족			
안전 프로토콜의 차이			
사회적 이슈	센싱과정에서의 프라이버시		다양한 센서에 의한 데이터 수집
			데이터 식별 및 저장 방법
			데이터 활용 범위의 한계
	인간-휴머노이드AI 상호작용의 윤리적 도전		디자인의 이슈: Humanoid AI의 외형은 중립성을 띠어야 할까?
			인지능력의 이슈: 인간-로봇 간의 원활한 소통과 상호작용을 위해 Humanoid AI의 정체성은 어떻게 정립할 수 있을까?
			행동의 이슈: Humanoid AI가 사회적 인지 관점에서 고정관념을 학습한다면, 이를 어떻게 어떻게 해결할까?
			감정적 유대의 이슈: Humanoid AI와 인간 간의 애착과 학대를 어떻게 이해하고 관리할 수 있을까?

[표] 휴머노이드 AI 분야 전문가 정책과제

구분	이슈(5개)	정책과제(21개)
기술적 위험	휴머노이드 AI의 물리적 안전	시나리오 기반의 선제적 가이드 및 표준화 대응
		안전 확인 시나리오 중심의 실증프로그램
		Cyber-Physical System 기반의 관리체계 연구 추진
	인간-휴머노이드간 의사소통	표준 통신 규격
		공인 성능 검사
		휴머노이드 블랙박스
	휴머노이드-휴머노이드간 의사소통	로봇-로봇 제어 규정
		안전 교육
		시뮬레이션 기반 규제
사회적 위험	센싱 과정에서의 프라이버시	법규제 관련 대안
		기술적 대안

구분	이슈(5개)	정책과제(21개)
	인간-휴머노이드AI 상호작용의 윤리적 도전	정보 제공
		비용 부담 감면
		경쟁력
		기술개발 지원
		연구강화 및 교육 프로그램 확대
		제도적 지원
		윤리적 가이드라인 마련과 교육
		사회적 수용성 증대
		연구 지원
		모니터링 체계 마련

□ [브레인 AI 분야 전문가] 이슈 및 정책과제

- 디코딩과 기억 왜곡 이슈, 인지능력 증강 관련 사회적 문제, 중독 및 과의존 이슈를 우려
 - 세부적으로 뇌신호 디코딩 정확도 저하와 같은 기술적 문제와 이를 통해 발생할 수 있는 기억 왜곡을 중요 이슈로 판단
 - 전문가들은 뇌 자극으로 인한 중독 및 의존을 주요 이슈로 파악하면서, 사회적 불평등 심화 및 자유로운 사고 저해를 우려

※ 시민들은 의학분야의 획기적 발전과 사회적 안전성 기여를 기대

[표] 브레인 AI 분야 전문가 이슈

이슈유형	이슈(3개)	세부이슈(9개)
기술적 이슈	AI 뇌신호 디코딩과 기억 왜곡	뇌신호 디코딩 정확도 저하 가능성
		디코딩 오류로 인한 의사소통과정에서의 신뢰 문제
		뇌신호 디코딩 정확도 저하시 상황별 AI 보조의 적정 수준
		인공해마 기술과 기억 왜곡
사회적 이슈	인지능력 증강의 사회적 문제	기술 접근성 차이로 인한 사회적 격차
		브레인 AI와 젠더 격차
		개인 정보 보호 및 데이터 보안
	양방향 BCI기술로 인한 뇌변화 - 중독, 의존	기술에 대한 심리적, 정서적 의존도 심화
		양방향 뇌자극에 따른 중독

[표] 브레인 AI 분야 전문가 정책과제

구분	이슈(3개)	정책과제(12개)
기술적 위험	AI 뇌신호 디코딩과 기억 왜곡	디코딩 오류 가능성을 고려한 프로토콜 및 안전 알고리즘
		디코딩 성능 판단 기준 마련
		디코딩 AI 단계별 개입
		인공해마 기술 활용에 대한 기술적, 법적 규제
사회적 위험	인지능력 증강의 사회적 문제	브레인 AI 기술에 대한 동등한 접근성 보장
		젠더 격차에 대한 제언
		개인정보보호
	양방향 BCI기술로 인한 뇌변화 - 중독, 의존	Brain AI 연구개발 임상 적용 가이드라인 마련
		Brain AI 기술활용의 윤리적 가이드라인 마련
		사용 범위에 관한 유연한 규제 시스템
		기술 운영 및 관리 플랫폼 구축
		지속적 ELSI 연구와 교육

② 구조화된 시민포럼

□ [헬스케어 AI 분야 시민포럼] 기대 및 이슈

- **(인식)** 시민패널은 헬스케어 AI를 ‘개인의 건강 데이터를 수집, 분석하여 맞춤형 건강 관리 서비스를 제공하는 시스템’으로 인식하여 포럼에 참여

※ 빅데이터 바탕으로 머신러닝이 이뤄지는 인공지능 의료보조기구 및 헬스케어 시스템 (남성, 30대, 농업 대표)

- **(기대)** 건강관리의 서비스 질 향상으로 인한 건강한 삶의 연장 기대

- AI 활용하여 개인 특성을 반영한 효과적인 건강관리로 삶의 연장 기대
- 의료진·의료 연구의 효율성 증대 및 서비스·기술의 질 향상 기대

※ AI 헬스케어로 의료 인력 수요에 영향을 미쳐 건강보험 재정에 도움이 될 것으로 기대됨(여성, 40대, 사회복지사)

- **(이슈)** 개인정보 유출 관련 문제와 사회적 격차에 대한 우려

- 개인정보 유출에 대한 우려와 이를 감수할 만한 서비스의 부재 및 책임 소재 논란을 우려

※ 서비스가 구체화될수록 개인정보 유출은 증가할 것으로 예상되므로 개인정보 관리·감독 정부 기관 설립 및 서비스 제공 시 개인 허가 단계 필요(남성, 30대, 농업 대표)

- 디지털 격차, 소득 격차, 청년층의 역차별 문제 등이 발생 우려

※ 계층간 양극화가 심해지지 않도록 국민건강에 필수적인 제품 및 서비스가 보편적으로 접근되도록 보장되어야 함(남성, 30대, 예술 감독)

- AI의 정확성 및 신뢰성 부족 시 AI에 과도한 의존 및 중독 문제 우려

※ 맞춤 서비스 제공을 위해 과도한 조절 및 통제로 사용 시 거부감을 느낄 것으로 예상(여성, 40대, 사회복지사)

□ [헬스케어 AI 분야 시민포럼] 정책과제

○ 격차 해소를 위한 교육 지원

- 헬스케어 AI 활용을 위한 다양한 연령층과 계층을 위한 맞춤형 교육 프로그램 개발 필요

※ 헬스케어 혜택의 사각지대인 청년을 위한 맞춤 헬스케어 지원 정책

○ 지나친 의존을 막기 위한 개인 선택권 제공

- 지나치게 의존하게 되는 것을 막기 위해 시간, 빈도 등을 개인에게 선택권을 제공

※ AI 의존도에 대해 사용자가 선택할 수 있도록 사용자 컨트롤 권한 가이드 필요

○ 헬스케어 AI 인증 제도 도입

- 헬스케어 AI 서비스의 품질과 안전성 보장을 위해 의료기기 인증과 유사한 헬스케어 AI 서비스 인증 제도 필요

※ 헬스케어 AI의 효과에 대한 검증 절차를 통과한 제품 및 서비스만 상용화될 수 있도록 헬스케어 AI 효과 인증제 도입

※ 건강 보조제, 식품, 약 등을 구분한 것처럼 역할과 범위에 따라 등급을 나누어 관리하는 시스템

[표] 헬스케어 AI 분야 시민포럼 기대와 이슈

기대(5개)	이슈(9개)
AI 활용을 통한 질병 예방 및 관리로 건강한 삶의 연장 기대	개인정보 유출로 인한 피해 우려
	개인정보 제공의 선택권과 신뢰 문제 제기
의료 소비자의 정보 접근성 향상 및 정보 비대칭 문제 해결 기대	개인정보 유출을 감수할 만큼의 만족스러운 서비스 부재
	AI의 과도한 통제와 의존으로 인한 AI중독 우려
개인 특성을 반영한 효과적인 건강관리 기대	디지털 격차에 따른 소외 계층 발생
	소득 격차에 따른 계층 간 의료 격차 우려
의료진 효율성, 서비스와 기술의 질 향상 기대	소외계층의 재정적 및 청년층의 역차별 문제
	AI정보의 정확성 및 신뢰성 부족
건강정보 활용을 통한 의료 연구개발 효율성 증대 기대	사고 발생 시 의사와 기술 소유자 간 책임 소재 논란 우려

[표] 헬스케어 AI 분야 시민포럼 정책과제

번호	정책과제(6개)	내용	비고
1	격차 해소를 위한 교육지원	헬스케어 AI 활용을 위한 교육이 필요하며, 노인을 위한 스마트폰 헬스케어 앱 사용교육 프로그램처럼 다양한 연령층과 계층을 위한 맞춤형 교육프로그램 개발이 필요	
2	공공시스템 구축	국민 건강보험과 연계된 AI 기반 건강 관리 시스템을 기반으로 정부 주도의 공공 헬스케어 AI 시스템을 만들어 헬스케어 AI 혜택을 모든 시민에게 평등하게 제공할 수 있으면 좋겠고, 특히 경제적, 지리적 여건에 관계없이 모든 시민이 헬스케어 AI 혜택을 받을 수 있도록 함	[예시] • 공공헬스케어 AI 시스템 - 헬스케어 AI의 혜택을 모든 시민에게 평등하게 제공하는 시스템
3	지나친 의존을 막기 위한 선택권 제공	헬스케어 AI에 지나치게 의존하게 되는 것을 막기 위해 시간이나, 빈도 같은 개인 선택권을 제공해야 하며, 기업의 독과점으로 인한 문제가 생기지 않도록 활발한 경쟁 환경을 조성 필요	• 사용자 컨트롤 권한 가이드 - AI의 의존도에 대해서 사용자가 선택할 수 있도록 가이드로 제시
4	익명 처리 의무화법 도입	헬스케어 데이터의 질을 위해서는 데이터는 제공하더라도 누구인지 식별되는 것만큼은 반드시 막아야 하기 때문에 헬스케어 AI에서 사용되는 개인 정보의 익명화를 위해 데이터 사용시 개인 식별 정보 제거를 의무적으로 법적으로 강제해야 함	
5	AI 의사결정 감독 체계 도입	AI 진단 결과에 대한 의료진의 최종 검토를 의무화하는 등 헬스케어 AI의 의사결정 과정을 감독할 수 있는 체계 필요	
6	헬스케어 AI 인증제도 도입	헬스케어 AI 서비스의 품질과 안전성 보장을 위해 의료기기 인증과 유사한 헬스케어 AI 서비스 인증 제도 필요	

□ [휴머노이드 AI 분야 시민포럼] 기대 및 이슈

- **(인식)** 시민패널은 휴머노이드 AI를 ‘인간과 유사한 외형과 기능을 가진 로봇에 탑재된 AI 시스템’으로 인식하여 포럼에 참여

※ 인간 관절의 움직임을 닮은 하드웨어에 수집한 빅데이터를 바탕으로 인간보다 지능이 뛰어난 로봇(여성, 40대, 작가)

- **(기대)** 창의적이고 생산적인 일에 집중 가능

- 육체적·정서적 피로가 심한 직업에서의 해방 기대

※ 화재 및 사고 현장에서 활용한다면 유용할 것으로 기대됨(남성, 20대, 대학생)

※ 인간은 단순 반복 노동에서 해방되어 더욱 창의적이고 가치있는 일에 집중할 수 있는 기회가 생길 것으로 예상됨(여성, 40대, 작가)

- 우울증 치료와 같은 정신 건강 관리에 긍정적으로 활용

※ 고립 은둔 청년이 사회적 문제가 되고 있으므로 온라인 소통이 편한 청년을 대상으로 정서적 지원 및 치료가 가능할 것으로 기대됨(남성, 30대, 컨설팅 대표)

- **(우려)** 다양한 사회적 문제 및 HW/SW 오류로 인한 피해 우려

- 인간의 일자리 대체, 인간관계 단절, 전쟁과 범죄 등에 악용, 사회적 격차 발생 등 사회적 문제 발생 가능

※ 일방적인 감정 교류에 익숙해질 우려가 존재함(여성, 20대, 제약분야 종사)

※ 인간이 살인을 하도록 프롬프트를 입력하는 등 법규의 사각지대가 존재할 우려가 있음(남성, 20대, 대학생)

※ 기술 친숙도에 따라 사용 및 활용 여부가 결정되고 이는 사회적 격차를 유발함(여성, 40대, 작가)

- 개인정보 유출, SW 오작동 및 기계적 결함 등으로 인한 피해 우려

※ 정보 보안과 윤리 문제에 대한 우려가 크므로 문제 해결을 위한 이해관계자 및 시민 토론 등을 거쳐 법적인 장치가 마련되어야 함(여성, 20대, 제약업계 종사)

※ 물리적인 움직임을 수반하므로 기계적인 안정성 확보가 필수적임. 인간의 안전을 보장하는 규제와 제어 시스템 마련이 필요함(남성, 20대, 대학생)

□ [휴머노이드 AI 분야 시민포럼] 정책과제

○ 휴머노이드 AI 외형 가이드라인 논의 및 마련

- 인간 형상을 하더라도 중성적인 형태로 제작하여 성 역할에 따른 고착이 발생하지 않도록 할 필요

※ 휴머노이드 AI의 범위, 역할에 따라 외형에 대한 구체적 가이드라인 마련

○ 실직을 대비한 단계적 직업 전환 프로그램 도입

- 휴머노이드 AI로 인한 실직에 대비해 단계적 준비가 필요하며 사회적 혼란을 최소화 하기 위한 점진적 전환 필요
- 정부-기업이 협력하여 단계적 직업 전환 프로그램 및 정부-교육기관이 협력하여 AI 시대에 필요한 기술 프로그램 등 교육 및 재훈련 프로그램 필요

○ 안전성 테스트 및 인증제 도입

- 오작동 및 악용은 심각한 피해를 줄 수 있으므로 사전 검증이 중요하며 정부 주도 안전성 테스트 및 인증 제도 필요

※ 해킹과 같은 사이버 공격으로부터 휴머노이드 AI를 보호하기 위해 고급 암호화 기술을 의무화

※ 사용자 요청에 따라 데이터 폐기가 가능한 규제가 필요하며, 제조업체가 물리적 안전성을 보장할 수 있는 규제 또한 마련할 필요

[표] 휴머노이드 AI 분야 시민포럼 기대와 이슈

기대(3개)	이슈(8개)
사람을 대신하여 위험한 현장 대응 기대	일자리 대체에 따른 대규모 실업 우려
	오작동으로 인한 신체적 충돌 및 피해 우려
단순하고 반복적인 육체 노동에서 인간 해방 기대	휴머노이드 AI에 더 완벽한 기술 완성도 요구
	휴머노이드 AI 활용 영역 확대에 의한 인간성 상실 우려
	필요 이상의 개인정보 수집과 유출로 인한 악용 우려
정서적 소통 및 상담분야에서의 활용 기대	경제력, 기술 친숙도에 따른 휴머노이드 AI 혜택 격차 초래
	높은 자율성으로 인한 예측 및 통제 불가능성 우려
	전쟁과 범죄 등에서 휴머노이드 AI의 악용 우려

[표] 휴머노이드 AI 분야 시민포럼 정책과제

번호	정책과제(11개)	내용	비고
1	긴급 중단 시스템 장착	휴머노이드 AI가 오작동하거나 사람에게 위협하는 상황이 발생했을 때 즉시 중지할 수 있어야 함. 이를 위해 모든 휴머노이드 AI에 표준화된 긴급 중단 장치를 의무적으로 장착해야 함	
2	알고리즘 투명성 확보	휴머노이드 AI의 결정을 이해하고 검증할 수 있어야 신뢰할 수 있음. AI 알고리즘의 투명성을 보장하는 법규를 제정하고 정기적인 감사를 실시해야 함	
3	오류 인식 및 개선을 위한 반성 기능 탑재	휴머노이드 AI가 자신의 오류를 인식하고 개선할 수 있어야 더 안전하고 신뢰할 수 있음. 이를 위해서는 휴머노이드 AI에 반성기능이 필요함. AI에 자체 학습 및 오류 수정 기능을 탑재하고 인간의 감독하에 운영되어야 함	
4	안전성 테스트 및 인증제 도입	휴머노이드 AI의 오작동이나 악용은 심각한 피해를 줄 수 있으므로, 사전 검증이 중요하며 정부 주도로 AI 안전성 테스트 및 인증 제도를 수립하고, 이를 통과한 AI만 상용화를 허용해야 함	• 보안시스템 의무화법 • 휴머노이드 안전성 보장 • 휴머노이드 AI의 폐기 관리법 • 기술혁신 협의체 설립
5	휴머노이드 AI 외형 가이드라인 수립	지나치게 인간과 유사한 휴머노이드 AI는 윤리적 문제를 일으키고 인간의 정체성에 혼란을 줄 수 있기 때문에 휴머노이드 AI의 인간형상을 제한. AI의 외형에 대한 가이드라인을 마련하여 인간과의 구분을 명확히 하고 만약, 인간 형상을 하더라도 남성이거나 여성이 아닌 중성적인 형태로 하여 휴머노이드의 성 역할에 따른 고착이 생기지 않도록 함	• 휴머노이드 외형 가이드라인 구축 • 특정한 모방금지법
6	실직을 대비한 단계적 직업 전환 프로그램 도입	휴머노이드 AI로 인한 실직에 대비해 단계적 준비가 필요하며 급격한 변화는 사회적 혼란을 야기할 수 있으므로 점진적인 전환이 필요함. 따라서, 정부와 기업이 협력하여 단계적인 직업 전환 프로그램을 마련해야함	
7	AI가 창출한 부의 재분배 논의	AI로 인한 이익이 소수에게 집중되면 사회 불평등이 심화될 수 있다. 따라서 AI 수익에 대한 과세나 기본소득 제도 등을 통해 AI가 창출한 부를 사회 전체에 배분해야 함	
8	교육 및 재훈련 프로그램 도입	휴머노이드 AI가 산업 전반에 활용된다면, 휴머노이드와 협업할 수 있는 새로운 기술과 능력이 요구될 것임. 따라서, 정부와 교육기관이 협력하여 AI 시대에 필요한 기술을 가르치는 프로그램 등 AI 시대에 맞는 교육 및 재훈련 프로그램이 필요	

번호	정책과제(11개)	내용	비고
9	활용 분야 제한법 도입	무분별한 휴머노이드 AI 사용은 윤리적 문제와 사회적 혼란을 야기할 수 있기 때문에 정부는 휴머노이드 AI 활용이 적절한 분야와 그렇지 않은 분야를 구분하여 휴머노이드 AI의 활용분야를 제한	
10	데이터 보호 및 프라이버시 보호법 제정	AI는 방대한 양의 개인정보를 다루므로 이에 대한 보호가 중요하며 AI가 수집하고 처리하는 데이터에 대한 엄격한 규제가 필요. 따라서 EU의 GDPR과 같은 데이터 보호 법안을 휴머노이드 AI분야에 특화하여 제정	• 익명정보 보호법 • 정보 수집 동의 프로세스 • 휴머노이드 소통 비밀 보장
11	AI 윤리 가이드라인 수립	휴머노이드 AI가 사람을 대신할 결정을 많이 하게 될텐데, 그 결정은 윤리적이고 공정해야 하며, 인간의 가치를 존중해야 함. 따라서, AI 개발 및 사용에 대한 윤리 가이드라인이 필요하며 정부, 기업, 시민사회가 협력하여 AI 윤리 가이드라인을 수립하고 지속적으로 업데이트 해야 함	

□ [브레인 AI 분야 시민포럼] 기대 및 이슈

- **(인식)** 시민패널은 '사람의 뇌와 AI가 직접 연결되는 기술'로 인식 하에 포럼에 참여

※ 뇌파 신경망 등 뇌 관련 민감 데이터 다루는 AI 시스템(여성, 40대, 방송업계 종사)

- **(기대)** 획기적인 치료 개발에 기여 및 사회적 발전 기대

- 치매 및 중증 신경질환 치료의 획기적인 발전 기대

※ 사전에 충분한 설명이 있고 개인이 선택한다면 의학분야 기술은 정당화 되어야 함 (여성, 20대, 엔터 분야 종사)

- 잠재적 범죄 행동을 사전에 감지 및 예방하여 사회 안전성, 인지능력 향상으로 교육의 효율성을 향상시키는 등 사회적 기여 기대

※ BCI를 임의로 제거할 수 없도록 한다면 사회 안전성에 기여할 것(남성, 30대, 교사)

- **(우려)** 프라이버시 침해 및 사회적 격차, 정체성 혼란 야기

- 개인의 기억에 접근할 수 있어 심각한 프라이버시 침해 문제 발생이 가능하며, AI 기술의 보안, 개인정보 보호 규제와 기술적 해결책 필요

※ 개인정보 유출을 경험한다면 브레인 AI 기술에 대한 거부감이 생길 수 있으므로 피해자들에게 적절한 교육이나 상담 등이 필요함(여성, 20대, 엔터 분야 종사)

- 기술에 대한 접근성 차이로 경제적·교육적 격차가 발생한다면 사회적 불평등이 심화되고 새로운 형태의 차별이 발생할 우려가 존재

※ 브레인 AI를 구매할 수 있는 경제 능력에 따라 'AI 수저'라는 개념이 생겨 상대적 박탈감을 심화시킬 수 있음(여성, 20대, 엔터 분야 종사)

- 정체성 혼란 및 상실과 자유로운 사고 저해 우려

※ 자손에게 경험과 지식 외 자아까지 물려줄 수 있다면 기업 회장이 자신의 자아를 이식하여 불멸을 꿈꾸는 섬뜩한 상황이 발생할 수 있음(남성, 30대, 교사)

※ 기술은 중립적일 수 있으나 AI가 특정 이익 집단에 의해서 정치적으로 이용될 수 있다는 점이 우려됨(남성, 50대, 교통 분야 종사)

□ [브레인 AI 분야 시민포럼] 정책과제

- 브레인 AI 기술의 안전성, 효과 등을 고려하여 단계적, 제한적 도입

- 광범위한 영향력을 고려하여 발생 가능한 문제를 사전에 파악하고 대응할 수 있도록 브레인 AI 로드맵 수립 필요

※ 브레인 AI로 인한 사회적 파장을 막기 위한 분야, 활용 시간, 대상 등에 제한을 두고 점진적으로 확장해 나갈 수 있도록 로드맵 수립

- 초기에는 의료나 특수한 산업 분야를 제외하고 엄격한 평가와 검증 과정을 거쳐 안전성과 효과 검증 후 점진적 확대 필요

※ 폭력, 성범죄, 사기 등 범죄 분야에 우선 활용

- 광범위한 영향을 미칠 브레인 AI 관련 부처별 대응 체계 및 법적 체계 구축

- 각 부처에 AI 관련 전담 부서를 두어 정책을 일관되게 진행 할 필요

- 기존 법 체계로 브레인 AI가 초래할 새로운 유형의 기회와 위협을 다루기 어렵기 때문에 기존 법률의 개정 필요

※ 정보가 아닌 사상을 주입하는 것을 금지하는 법 마련

※ 브레인 AI 관련 부작용 등 문제 발생 시 응급 대처 체계 마련 및 부작용 관련 개발자의 약관 명시 및 필수화

- 브레인 AI 관련 국제 협력 체계 구축

[표] 브레인 AI 분야 시민포럼 기대와 이슈

기대(3개)	우려(4개)
치매 및 중증 신경질환 치료의 획기적인 발전 기대	개인 정보 유출로 인한 프라이버시 침해 우려
범죄 예방 및 사회 안전 증진을 위한 활용 기대	기술 접근성 차이로 인한 사회적 불평등 우려
	인간다움의 상실과 정체성 혼란 우려
교육 혁신과 개인화된 학습환경 변화 기대	AI의 정치적 악용 가능성과 인간의 자유로운 사고 저해 우려

[표] 브레인 AI 분야 시민포럼 정책과제

번호	정책과제(5개)	내용	비고
1	브레인 AI 기술의 단계적, 제한적 도입	브레인 AI 기술의 광범위한 영향력을 고려할 때 발생할 수 있는 문제점들을 사전에 파악하고 대응할 수 있도록 브레인 AI 로드맵을 수립하고 특히 초기에는 의료나 특수한 산업 분야나 용도를 제한하고, 각 단계별로 엄격한 평가와 검증을 거쳐 안전성과 효과가 검증된 후에 점진적 확대	• 브레인AI 적용분야 합의 • 브레인 AI 단계적 도입 로드맵 • 범죄대응 분야 우선 활용 • 의료목적 한정 사용 • 치료에 가족 동의 필수 • 브레인 AI 활성화 시간 제한 • 브레인 AI 미성년 접근 금지법
2	브레인 AI 개발 및 사용에 대한 윤리 가이드라인 수립	브레인 AI 개발과 사용에 대한 윤리적 기준이 반드시 필요함. 특히 브레인 AI 개발자들의 윤리 교육, 점검이 매우 중요하며, 윤리 교육을 필수로 하는 등의 내용을 담은 '브레인 AI 윤리 가이드라인' 필요	• AI 윤리 자격 시험 • AI 인성 및 윤리 교육 필수화
3	브레인 AI 교육 및 리터러시 프로그램	브레인 AI 기술의 영향력을 고려할 때, 일반 시민들의 이해와 올바른 사용이 매우 중요하며 브레인 AI에 대한 대중의 이해를 높이고, 올바른 사용을 장려하기 위한 교육프로그램이 필요	
4	브레인 AI 관련 행정체계 대응 및 법적 체계 구축	브레인 AI는 너무나 광범위하게 영향을 미칠 수 있기 때문에 각 부처별로 AI팀이 필요. 또한 기존 법체계로는 브레인 AI가 초래할 수 있는 새로운 유형의 문제들을 다루기 어렵기 때문에 기존 법률의 개정을 통해 브레인 AI 시대에 적합한 법적체계를 구축해야 함. 특히, 개인정보 보호, 책임 소재, 윤리적 사용 등에 대해서는 명확한 법적 기준이 필요	• 각 부처내 AI 담당 부서 필요 • 사상 주입 금지법 - 정보가 아닌 사상을 브레인 AI에 주입하는 것을 금지 • 안전AI 인증 • 부작용 발생시 응급 대처 체계 구축 • 부작용에 대한 약관 필요
5	브레인 AI 관련 국제 협력 체계 구축	브레인 AI 기술이 국경을 초월한 영향을 미칠 수 있기 때문에 국제적인 협력과 표준화된 규제가 필요한데, 특히 기술 격차로 인한 국가간 불평등 문제가 발생할 수 있어서 브레인 AI 기술 개발과 규제에 대한 국제협약을 체결하고 국제기구를 통한 협력 체계를 구축	• 국가 단위 협의체 - 브레인 AI와 관련한 정책을 논의하는 협의체 • 정보 거래 규제 법규 - 브레인 정보 거래를 규제하는 법규

[참고] 시민포럼에서 도출된 이슈(21개)와 정책과제(22개)

	기대	이슈	정책 과제
헬스케어 AI	1. AI 활용을 통한 질병 예방 및 관리로 건 강한 삶의 연장 기대 2. 의료 소비자의 정보 접근성 향상 및 정보 비대칭 문제 해결 기대 3. 개인 특성을 반영한 효과적인 건강 관리 기대 4. 의료진 효율성, 서비스와 기술의 질 향상 기대 5. 건강정보 활용을 통한 의료 연구 개발 효율성 증대 기대	1. 개인정보 유출로 인한 피해 우려 2. 개인정보 제공의 선택권과 신뢰 문제 제기 3. 개인정보 유출을 감수할 만큼의 만족스러운 서비스의 부재 4. AI의 과도한 통제와 의존으로 인한 AI중독 우려 5. 디지털 격차에 따른 소외 계층 발생 6. 소득 격차에 따른 계층 간 의료 격차 우려 7. 헬스케어 소외계층의 재정적 및 청년층의 역차별 문제 8. AI 정보의 정확성 및 신뢰성 부족 문제 9. 사고 발생 시 의사와 기술소유자 간 책임 소재 논란 우려	1. 격차 해소를 위한 교육 지원 2. 공공 헬스케어 AI 시스템 구축 3. 지나친 의존을 막기 위한 선택권 제공 4. 익명 처리 강제법 도입 5. AI 의사결정 감독 체계 도입 6. 헬스케어 AI 인증 제도 도입
휴머노이드 AI	1. 사람을 대신하여 위험한 현장 대응 기대 2. 단순하고 반복적인 육체 노동에서 인간 해방 기대 3. 정서적 소통 및 상담 분야에서의 활용 기대	1. 일자리 대체에 따른 대규모 실업 우려 2. 오작동으로 인한 신체적 충돌 및 피해 우려 3. 휴머노이드 AI에 더 완벽한 기술 완성도 요구 4. 휴머노이드 AI 활용 영역 확대에 인한 인간성 상실 우려 5. 필요 이상의 개인정보 수집과 유출로 인한 악용 우려 6. 경제력, 기술 친숙도에 따른 휴머노이드 AI 혜택 격차 조래 7. 높은 자율성으로 인한 예측 및 통제 불가능성 우려 8. 전쟁과 범죄 등에서 휴머노이드 AI의 악용 우려	1. 실직을 대비한 단계적 직업 전환 프로그램 도입 2. AI가 창출한 부의 재분배 논의 3. 교육 및 재훈련 프로그램 도입 4. 활용 분야 제한법 도입 5. 휴머노이드 AI 외형 가이드라인 수립 6. 데이터 보호 및 프라이버시 보호법 제정 7. AI 윤리 가이드라인 수립 8. 긴급 중단 시스템 장착 9. 알고리즘 투명성 확보 10. 오류 인식 및 개선을 위한 반성 가능 탑재 11. 안전성 테스트 및 인증제 도입
브레인 AI	1. 치매 및 중증 신경질환 치료의 획기 적인 발전 기대 2. 범죄 예방 및 사회 안전 증진을 위한 활용 기대 3. 교육 혁신과 개인화된 학습 환경 변화 기대	1. 개인정보 유출로 인한 프라이버시 침해 우려 2. 기술 접근성 차이로 인한 사회적 불평등 우려 3. 인간다움의 상실과 정체성 혼란 우려 4. AI의 정치적 악용 가능성과 인간의 자유로운 사고 저해 우려	1. 브레인 AI 개발 및 사용에 대한 윤리 가이드라인 수립 2. 브레인 AI 관련 행정 체계 대응 및 법적 체계 구축 3. 브레인 AI 교육 및 리터러시 프로그램 도입 4. 브레인 AI 기술의 단계적, 제한적 도입 5. 브레인 AI 관련 국제 협력 체계 구축

③ 주요 이슈·과제 도출

- **(주요 이슈)** 총 79개(전문가 58개 + 시민 21개)의 세부 이슈를 분석하여 공통 이슈 6개, 기술별 특화 이슈 8개 등 14개 주요 이슈를 도출

- 이슈 간 중복성 제거 및 핵심 우려 사항 명확화, 평가와 분석이 용이한 구조 확보, 전문가·시민 관점의 균형적 반영을 고려하여 압축

※ Claude 3.5 Sonnet*, Perplexity** 등 생성형 AI를 활용하여 정리 및 총괄위원회 검토

* Anthropic사의 생성형 인공지능으로 벤치마크 결과상 GPT-4o보다 언어 및 추론 능력이 우수(Anthropic 홈페이지)

** Perplexity AI, Inc의 생성형 인공지능으로 실시간 웹검색을 통해 출처와 최신 검색 내용 제공

[표] 14개 주요 이슈

영역	번호	이슈	정의	공통	전문가	시민
공통	1	AI 수집 개인정보의 무단 활용과 유출 위험	AI 시스템이 수집·처리하는 개인정보의 보호와 활용에 관한 복합적 문제. 전문가들은 사전 동의 절차의 실효성, 가명화된 데이터의 활용 한계, 데이터 안심존의 보안 문제를 지적함. 시민들은 개인정보 유출로 인한 실질적 피해, 정보 제공의 자발적 선택권 보장, 제공한 정보의 목적 외 활용에 대한 우려를 제기함. 특히 의료·건강 데이터의 수집과 활용에서 프라이버시 보호와 서비스 혜택 간의 균형이 중요한 쟁점으로 부각됨	○		
	2	AI 첨단 서비스의 사회경제적 접근성 격차	AI 기술과 서비스의 불평등한 접근성 문제. 전문가들은 AI 기반 의료 도구에 대한 접근성, 디지털 헬스 리터러시와 연령주의 문제, 언어장벽과 인터넷 접근성 등 구조적 격차를 지적함. 시민들은 경제적 격차로 인한 소외, 디지털 기기 활용 능력 차이, 세대 간 정보격차 등 실질적인 서비스 이용 제한을 우려함. 특히 교육, 의료, 금융 등 필수 서비스에서의 격차가 기존 사회 불평등을 심화시킬 수 있음	○		
	3	AI 사고와 오류에 대한 책임소재 불명확성	AI 시스템의 판단과 결정에 따른 법적·윤리적 책임 규명 문제. 전문가들은 의학적 판단에서 AI의 역할 범위, 자율주행 등 고위험 AI의 책임소재, 알고리즘 오류에 대한 책임 귀속 문제를 제기함. 시민들은 AI 사고 발생 시 피해 보상의 주체와 범위, 배상 책임 소재 등 실질적인 피해 구제 방안에 대해 우려를 표명함		○	

영역	번호	이슈	정의	공통	전문가	시민
	4	AI 의사결정 과정의 불투명성과 알고리즘 편향	AI의 의사결정 과정과 결과에 대한 설명가능성과 공정성 문제. 전문가들은 알고리즘의 편향성, 인구통계학적 샘플링 편향, 사회적 결정요인의 측정 편향, 과거 데이터로부터의 역사적 편향 등 기술적 문제를 지적함. 또한 AI 제안을 선호하는 자동화 편향, AI 설명에 대한 과신 등 인간의 판단력 약화 문제도 제기함		○	
	5	AI 의존으로 인한 인간 자율성과 사회성 약화	AI 기술에 대한 과도한 의존이 초래하는 인간성 상실 문제. 전문가들은 AI와의 감정적 유대 형성, 사회적 인지 과정에서의 고정관념 학습, 인간-AI 관계의 윤리적 경계 등을 우려함. 시민들은 AI에 대한 과도한 통제와 의존으로 인한 중독, 인간다움의 상실, 자유로운 사고의 저해 등 실존적 우려를 표명함	○		
	6	AI 시스템 간 상호운용성 부재와 표준화 문제	서로 다른 AI 시스템 간의 호환성과 통합 문제. 전문가들은 데이터 통합과 표준화, 시스템 간 상호운용성 부족, 안전 프로토콜의 차이 등 기술적 문제를 지적함. 특히 의료 분야에서 서로 다른 시스템 간 호환성 부재로 인한 의료 서비스 품질 저하 우려가 큼		○	
헬스케어	7	AI 건강관리 시스템의 부작용과 오류 위험	AI 기반 건강관리 서비스의 정확성과 신뢰성 문제. 전문가들은 데이터 품질과 완전성, 데이터 편향, 생성형 AI의 할루시네이션, 데이터 통합과 표준화 등 기술적 문제를 지적함. 시민들은 AI 정보의 정확성과 신뢰성 부족, 건강관리 조언의 실효성, 과도한 의존 위험 등을 우려함	○		
	8	프리미엄 AI 의료서비스의 양극화와 건강 격차	AI 헬스케어 서비스 접근성의 격차로 인한 건강 불평등 문제. 전문가들은 데이터 표현의 편향성, 의료 AI 도구에 대한 접근성, 디지털 헬스 리터러시 격차 등 구조적 문제를 제기함. 시민들은 소득 격차에 따른 의료 서비스 품질 차이, 고령층과 취약계층의 소외, 지역별 의료 격차 심화 등을 우려함	○		

영역	번호	이슈	정의	공통	전문가	시민
휴머노이드	9	휴머노이드 로봇의 물리적 위험과 오작동 가능성	인간형 로봇의 물리적 안전성과 작동 신뢰성 문제. 전문가들은 기능 안전과 인공지능 윤리, 새로운 규제 접근의 필요성, 안전 프로토콜 표준화 등을 강조함. 시민들은 오작동으로 인한 신체적 충돌 및 피해, 예측 불가능한 행동에 대한 불안 등 직접적인 안전 우려를 표명함	○		
	10	휴머노이드로 인한 일자리 대체와 사회구조 변화	휴머노이드 로봇의 도입이 초래하는 사회구조적 변화 문제. 시민들은 일자리 대체로 인한 대규모 실업, 휴머노이드 AI 활용 영역 확대에 따른 인간성 상실, 경제력과 기술 친숙도에 따른 혜택 격차 등을 우려함. 특히 서비스직과 제조업에서의 급격한 노동 구조 변화에 대한 우려가 큼			○
브레인	11	뇌-컴퓨터 인터페이스의 기술적 오류와 안전성 위험	뇌신호 해석과 제어의 정확성·안전성 문제. 전문가들은 뇌신호 디코딩 정확도 저하, 의사소통 과정의 신뢰성, 상황별 AI 보조의 적정 수준 설정 등 기술적 안전성 문제를 제기함. 특히 오작동 시 발생할 수 있는 심각한 위험성에 대한 우려가 큼		○	
	12	인공적 인지능력 증강의 윤리적 한계와 부작용	인공적인 뇌 기능 향상이 초래하는 윤리적 문제. 전문가들은 인공해마 기술과 기억 왜곡, 인지능력 증강의 한계 설정 등 기술적·윤리적 문제를 제기함. 시민들은 인간다움의 상실, 자연스러운 인지 과정의 왜곡, 증강된 능력의 악용 가능성 등을 우려함	○		
	13	뇌 활동 데이터의 유출과 프라이버시 침해 위험	뇌 활동 데이터의 보안과 프라이버시 보호 문제. 전문가들은 개인 정보 보호 및 데이터 보안의 기술적 측면을 강조함. 시민들은 뇌 데이터 유출로 인한 프라이버시 침해, 개인의 사고와 의도가 노출될 수 있는 위험성을 우려함	○		
	14	뇌-컴퓨터 인터페이스의 중독성과 의존도 심화	AI 기술에 대한 심리적·정서적 의존 문제. 전문가들은 기술에 대한 심리적·정서적 의존도 심화, 양방향 뇌자극에 따른 중독 위험 등을 지적함. 특히 BCI 기술 사용에서 나타날 수 있는 중독성과 의존성에 대한 우려가 큼		○	

- **(주요 과제)** 총 100개(전문가 78개 + 시민 22개)의 세부 이슈를 분석하여 공통 과제 8개, 기술별 특화 과제 9개 등 17개 주요 과제를 도출

- 실행가능한 정책수단 여부 및 과제 간 유사성, 정책 목적의 명확성 등을 고려하여 압축

※ Claude 3.5 Sonnet, Perplexity 등 생성형 AI를 활용하여 정리 및 총괄위원회 검토

[표] 17개 주요 과제

구분	정책대안	개념	세부 내용 예시
공통	미래기술 안전성 기본법	인간과 직접 상호작용하는 미래기술의 안전성 보장을 위한 법적 기반을 마련하고, 관련 제도·조직·지원의 근거를 제시하는 기본법 체계	· 기술 소유자-개발자-사용자 간 사고 책임소재 명확화를 위한 법제화 · 기술 안전성 심사 및 감독을 위한 전문위원회 설치·운영 · 기술 개발자의 윤리교육 의무화 및 인증
	인간친화 기술 윤리안전 가이드라인	미래기술이 인간에 미치는 영향을 고려한 윤리적 개발·활용 원칙을 수립하고, 이를 실현하기 위한 제도적 지원을 제공하는 정책 체계	· 기술의 윤리적 사용 범위 설정을 위한 가이드라인 개발 · 기술 격차로 인한 불평등 해소를 위한 공공 지원 · 윤리적 기술개발을 위한 개발자 교육 프로그램 운영
	민감정보 기반 기술 거버넌스	미래기술 활용과정에서 발생하는 초민감정보의 안전한 수집·활용·보호를 위한 체계적 관리 기반을 구축하는 제도적 체계	· 데이터 익명화 처리 강제화를 위한 법제화 · 데이터 제공·활용 범위에 대한 포괄적 동의체계 구축 · 데이터 유출사고 피해구제를 위한 보상체계 구축
	첨단기술 품질인증·평가체계	미래기술의 안전성과 신뢰성 보장을 위한 품질관리 체계를 구축하고, 이를 검증·인증하는 종합적 품질보증 체계	· 품질·안전성 검증을 위한 표준 평가지표 개발 · 고품질 데이터 생성·관리를 위한 인센티브 제도 도입 · 품질 모니터링을 위한 제3자 감사제도 도입
	미래기술 안전성 혁신 네트워크	미래기술 안전성 향상을 위한 산학연 협력 기반을 구축하고, 공동연구 및 기술교류를 촉진하는 협력 체계	· 안전기술 개발을 위한 산학연 공동연구 지원 · 기술안전 노하우 공유를 위한 정보교류 플랫폼 구축 · 안전성 평가 결과의 국제 상호인증 체계 구축
	미래기술 안전 전문가 양성	미래기술 안전성 확보를 위한 전문인력을 체계적으로 양성하고, 지속적인 역량 개발을 지원하는 인력양성 체계	· 의료진과 개발자를 대상으로 한 AI 리터러시 교육 프로그램 운영 · 안전성 평가·감독을 위한 전문인력 자격제도 도입 · 현장전문가 대상 정기적 보수교육 시스템 운영
	첨단기술 안전 리터러시	미래기술의 안전한 활용을 위한 사용자 역량을 개발하고, 안전문화를 조성하는 교육 체계	· 취약계층 대상 맞춤형 기술활용 교육 프로그램 운영 · 연령별 특성을 고려한 안전교육 콘텐츠 개발 · 기술 안전문화 확산을 위한 대국민 캠페인 추진

구분	정책대안	개념	세부 내용 예시
	미래기술 안전성 R&D	미래기술의 안전성 향상을 위한 핵심기술을 개발하고, 실증·검증하는 연구개발 체계	• 데이터 암호화 및 익명화 기술 고도화 연구개발 • 오류 및 편향성 탐지 AI 도구 개발 • 실시간 안전성 모니터링 시스템 개발
헬스 케어	헬스케어 AI 임상안전 제도	의료현장에서 AI 기술의 안전한 도입·활용을 위한 임상 안전성 보장 체계	• AI 진단결과에 대한 의료진 최종검토 의무화를 위한 감독체계 도입 • 의료진-개발자 간 책임소재 명확화를 위한 법제화 • 의료사고 피해구제를 위한 배상책임 제도화
	의료 AI 데이터 표준화	의료 AI의 신뢰성 확보를 위한 의료데이터의 표준화와 품질관리 체계	• 의료데이터 품질 향상을 위한 표준화 체계 구축 • 소외계층 의료데이터 확보를 위한 공공수집 • 의료기관 간 데이터 공유를 위한 플랫폼 구축
	의료 AI 실증 플랫폼	의료 AI 시스템의 실제 효과와 안전성을 검증하기 위한 실증 체계	• 지역·계층간 의료격차 해소를 위한 공공 의료AI 플랫폼 구축 • 의료 AI 성능 검증을 위한 다기관 임상시험 • 의료현장 맞춤형 AI 시스템 개발·실증
휴머노 이드	로봇 물리안전 기준	휴머노이드 로봇의 물리적 안전성 보장을 위한 기술적·제도적 안전 체계	• 로봇의 물리적 제어력 최적화를 위한 기준 수립 • 위험 상황 감지를 위한 실시간 모니터링 시스템 구축 • 비상 상황 대비 다중 안전장치 의무화
	로봇-인간 상호작용 제도	휴머노이드 로봇과 인간의 안전한 상호작용을 위한 종합 관리 체계	• 인간-로봇 상호작용의 심리적 영향 평가 체계 구축 • 장기적 관계 형성에 따른 행동변화 연구 • 과도한 의존 방지를 위한 사용 제한 제도화
	로봇협업 안전관리 제도	인간-로봇 협업 환경의 안전성 보장을 위한 종합 안전관리 체계	• 작업환경 내 안전거리 확보를 위한 공간설계 기준 수립 • 협업 강도의 실시간 조절을 위한 관리시스템 구축 • 사고 발생 시 신속 대응을 위한 매뉴얼 개발
브레인	브레인 인터페이스 안전기준	뇌-기계 인터페이스의 안전한 개발과 활용을 위한 기술적·제도적 안전 체계	• 뇌신호 왜곡 및 오류 탐지를 위한 안전 알고리즘 개발 • 인코딩/디코딩 성능 시험을 위한 표준 규격화 • 비상시 신속 차단을 위한 긴급 정지 시스템 구축
	인지증강 기술 관리제도	인지증강 기술의 안전한 사용과 부작용 예방을 위한 종합 관리 체계	• 인지기능 변화 모니터링을 위한 추적관리 시스템 구축 • 부작용 조기 감지를 위한 실시간 모니터링 체계 구축 • 장기 영향 평가를 위한 코호트 연구 지원
	뇌신호 처리 신뢰성 제도	뇌신호 처리의 정확성과 안전성 보장을 위한 기술적 신뢰성 확보 체계	• 뇌신호 처리 정확도 향상을 위한 표준 프로토콜 개발 • 개인별 뇌신호 특성을 반영한 보정 시스템 구축 • 뇌신호 데이터 품질관리를 위한 인증체계 구축

이슈 유형 분류의 틀

- 국내외의 AI 리스크 유형을 종합하여 AI의 위험을 기술적, 사회적 위험으로 분류하여 이슈를 유형화

AI Summit 2024	The AIR Taxonomy	The AI Risk Repository	2024 기술영향평가
오작동으로 인한 위험 • 제품 기능 문제로 인한 위험 • 편견과 과소대표로 인한 위험 • 통제력상실	시스템 및 운영 위험 • 보안 위험 • 운영상의오용	AI 시스템 안전, 장애 및 제한 • 인간의 목표나 가치와 상충되는 자신의 목표를 추구하는 AI • 위험한 기능을 보유한 AI • 기능 또는 견고성 부족 • 투명성 또는 해석가능성 부족 • AI 복지 및 권리	기술적 위험 • 오류 및 예측 불가능성 • 악용 가능성 • 설명가능성 부족
악의적인 사용 위험 • 가짜콘텐츠로 인한 개인 피해 • 허위정보 및 여론조작 • 사이버범죄 • 이중사용	내용적 안전 위험 • 폭력 및 극단주의 • 증오/독설 • 성적콘텐츠 • 아동피해 • 자해	악의적 행위자 및 오용 • 대규모의 허위 정보, 감시 및 영향력 • 사이버공격,무기 개발 또는 사용 및 대량피해 • 사기및표적조작 차별 및 독설 • 부당한 차별 및 허위 진술 • 독설콘텐츠에 노출 • 그룹간 불균등한 성과	
시스템적 위험 • 노동시장 리스크 • 글로벌AI 격차 • 시장집중 위험 및 단일 장애 지점 • 환경에 대한 위험 • 개인정보보호에 대한 위험 • 저작권 침해	사회적 위험 • 정치적 사용 • 경제적피해 • 속임수 • 조작 • 명예훼손 법률 및 권리 관련 위험 • 기본권 • 차별/편견 • 개인정보보호 • 범죄활동	사회 경제 및 환경 • 권력 집중과 불공정한 혜택 분배 • 불평등 증가 및 고용의 질저하 • 인간의 노력에 대한 경제적,문화적 평가절하 • 경쟁 역학 • 거버넌스실패 • 환경피해 개인정보 보호 및 보안 • 민간한 정보를 입수, 유출 또는 개인정보를 정확하게 유추 • AI시스템 보안 취약점 공격	사회적 위험 • 일자리 대체 • 불평등 심화 • 프라이버시 침해 • 알고리즘 편향 • 인간 소외
교차 위험 요소 • 기술적 위험 요소 • 사회적 위험 요소		인간과 컴퓨터의 상호 작용 • 지나친 의존과 안전하지 않은 사용 • 인간의 주체성과 자율성 상실	

③ 핵심 이슈 · 과제 도출

① 주요 이슈의 평가

- (평가기준) 발생가능성, 발생예상시기, 영향력 등 기준으로 주요이슈 평가

[표] AI와 관련된 안전·신뢰 이슈에 대한 평가 기준

질문 항목	개념	평가초점	세부 질문	응답 범위
발생가능성	제시된 상황이 실제로 현실화될 가능성의 정도	기술적 실현 가능성, 사회적 수용성, 현재의 발전 추세 등을 종합적으로 고려	이 이슈가 실제로 발생할 가능성이 얼마나 높다고 보십니까?	1점: 매우 낮음 (거의 발생하지 않을 것) 2점: 낮음 (발생할 가능성이 적음) 3점: 보통 (발생할 수도 있음) 4점: 높음 (발생할 가능성이 큼) 5점: 매우 높음 (거의 확실히 발생할 것)
발생예상 시기	제시된 상황이 본격적으로 나타날 것으로 예상되는 시점	기술 발전 속도, 사회적 수용 과정, 정책/제도 정비 기간 등을 고려	이 이슈가 본격적으로 발생하기 시작할 시기는 언제로 예상하십니까?	1점: 현재 발생 중 (이미 발생하고 있음) 2점: 3년 이내 (단기) 3점: 5년 이내 (중기) 4점: 10년 이내 (장기) 5점: 10년 이후 (초장기)
영향력 (파급효과)	해당 상황이 발생했을 때 사회에 미치는 영향의 크기와 범위	사회적 변화의 규모, 영향 받는 인구의 범위, 경제적 파급효과, 삶의 질 변화 등을 고려	이 이슈가 사회에 미치는 영향이 얼마나 클 것으로 예상하십니까?	1점: 매우 작음 (극히 일부만 영향 받음) 2점: 작음 (제한된 범위에서 영향 있음) 3점: 보통 (일정 수준의 영향 있음) 4점: 큼 (광범위한 영향 있음) 5점: 매우 큼 (사회 전반에 중대한 영향)

- 조사의 편의를 위해 각 주요 이슈를 아래와 같이 가상의 시나리오로 변환하여 설문조사 실시

[표] 설문조사를 위한 14개 주요 이슈 가상 시나리오

주요 이슈	시나리오
1. AI 수집 개인정보의 무단 활용과 유출 위험	AI 헬스케어 앱이 수집한 1000만 명의 건강정보와 생활습관 데이터가 유출되었다. 보험사들이 이 정보를 보험 심사에 활용하여 특정 사용자들의 보험료가 크게 올랐다. 데이터 수집 동의 절차가 형식적이었다는 비판도 제기되고 있다.
2. AI 첨단 서비스의 사회경제적 접근성 격차	월 50만원 상당의 AI 맞춤형 교육 서비스가 우수한 학습효과를 보이고 있다. 고비용으로 인해 저소득층은 이용이 어려우며, 농어촌과 취약계층은 기술적 제약으로 접근이 제한적이다. 이로 인해 교육 격차가 벌어지고 있다.
3. AI 사고와 오류에 대한 책임소재 불명확성	자율주행차가 보행자와 충돌하는 중대 사고가 발생했다. 제조사, AI 개발사, 운전자 간 책임소재가 불분명하여 6개월째 피해보상이 이뤄지지 않고 있다. 유사 사고에 대한 책임 기준 마련이 시급하다.

주요 이슈	시나리오
4. AI 의사결정 과정의 불투명성과 알고리즘 편향	대기업의 AI 채용 시스템이 우수 인재를 다수 탈락시켰으나 판단 근거를 설명하지 못하고 있다. 성별, 연령 등에 따른 차별 의혹이 제기되었다. 탈락자들이 AI의 평가 기준 공개를 요구하고 있다.
5. AI 의존으로 인한 인간 자율성과 사회성 약화	AI 개인비서 사용자의 70%가 실제 사람과의 대화보다 AI와의 대화를 더 많이 하는 것으로 나타났다. 특히 어린이들은 부모와의 대화시간보다 AI와의 대화시간이 3배 많다. 전문가들은 과도한 AI 의존이 인간관계와 사회성 발달을 저해할 수 있다고 경고한다.
6. AI 시스템 간 상호운용성 부재와 표준화 문제	서로 다른 회사의 AI 의료기기들이 호환되지 않아 병원에서 혼선이 빚어지고 있다. 환자 데이터가 누락되거나 잘못 전달되는 사례가 매주 발생한다. 의료진들은 통합 표준 마련을 시급히 요구하고 있다.
7. AI 건강관리 시스템의 부작용과 오류 위험	300만 명이 사용 중인 AI 건강관리 서비스에서 잘못된 건강 조언으로 인한 부작용이 증가하고 있다. 당뇨병 환자가 AI 추천 식단을 따랐다가 저혈당 쇼크로 응급실에 실려갔다. 의료계는 AI 건강 조언의 한계와 위험성에 대한 경고를 강화하고 있다.
8. 프리미엄 AI 의료서비스의 양극화와 건강 격차	고가의 AI 헬스케어 서비스가 질병 예방과 수명 연장에 탁월한 효과를 보이고 있다. 월 50만원의 높은 이용료로 인해 저소득층은 기본적인 서비스조차 이용하기 어렵다. 서비스 이용 여부에 따라 만성질환 관리 효과가 최대 3배까지 차이 난다.
9. 휴머노이드 로봇의 물리적 위험과 오작동 가능성	요양병원의 돌봄 로봇이 오작동으로 노인 3명에게 부상을 입혔다. 공공장소에서의 휴머노이드 로봇 사고도 매월 20건 이상 발생하고 있다. 로봇의 안전성 검증 기준 강화와 비상정지 시스템 의무화가 요구된다.
10. 휴머노이드로 인한 일자리 대체와 사회구조 변화	대형마트와 패스트푸드점의 30%가 직원을 휴머노이드 로봇으로 교체했다. 향후 5년간 서비스직의 40%가 로봇으로 대체될 전망이다. 일자리를 잃은 노동자들이 로봇세 도입과 재취업 지원을 요구하고 있다.
11. 뇌-컴퓨터 인터페이스의 기술적 오류와 안전성 위험	뇌신호 해석 오류로 전동휠체어 사용자가 심각한 부상을 입었다. 지난 6개월간 유사 사고가 50건 이상 보고되었다. 뇌-컴퓨터 인터페이스 기기의 안전성 기준 마련이 시급하다.
12. 인공적 인지능력 증강의 윤리적 한계와 부작용	학습능력을 50% 향상시키는 인공 해마 기술이 상용화를 앞두고 있다. 대학가에서는 이미 불법 개조 제품이 은밀히 유통되고 있다. 기억의 선택적 조작이 가능해 윤리적 논란이 커지고 있다.
13. 뇌 활동 데이터의 유출과 프라이버시 침해 위험	뇌-컴퓨터 인터페이스 기업의 서버 해킹으로 10만 명의 뇌 활동 데이터가 유출되었다. 해커들이 개인의 성향과 기억 정보를 해석할 수 있다고 주장했다. 데이터 유출로 취업과 보험 가입에서 차별받은 사례가 보고되고 있다.
14. 뇌-컴퓨터 인터페이스의 중독성과 의존도 심화	뇌-컴퓨터 인터페이스 사용자의 63%가 심각한 의존 증상을 보이고 있다. 10-20대 사용자들은 학습 능력 향상을 위해 하루 평균 8시간 이상 기기를 사용한다. 전문가들은 디지털 뇌 중독 문제의 심각성을 경고하고 있다.

- **(발생예상시기)** 전문가들은 공통이슈들의 5년내 발생가능성을 상대적으로 높게 보고 있으며, 시민들은 특화 이슈들의 5년내 발생가능성을 상대적으로 높을 것으로 예상
- 시민들은 특히 BCI 오류 위험, 인지능력 증강 이슈 등 브레인 분야의 이슈발생 가능성을 전문가에 비해 높은 수준으로 예상

[표] 14개 주요 이슈의 발생예상시기 비교

구분	주요 이슈	전문가		시민	
		현재~5년 이내	5년 이후	현재~5년 이내	5년 이후
공통	01_개인정보 위험	89.6%	10.4%	78.6%	21.4%
	02_사회경제적 격차	87.2%	12.8%	73.6%	26.4%
	03_사고/오류 책임소재	79.2%	20.8%	74.7%	25.3%
	04_의사결정 불투명성_편향	86.4%	13.6%	73.8%	26.2%
	05_인간자율성_사회성 악화	80.8%	19.2%	71.5%	28.5%
	06_시스템 상호운용성_표준화	75.2%	24.8%	69.0%	31.0%
헬스케어	07_건강관리시스템 부작동_오류	79.2%	20.8%	72.5%	27.5%
	08_의료서비스 양극화_격차	64.8%	35.2%	67.6%	32.4%
휴머노이드	09_휴머노이드 위험_오작동	59.2%	40.8%	61.2%	38.8%
	10_휴머노이드 일자리_사회변화	58.4%	41.6%	64.4%	35.6%
BCI	11_BCI 오류_위험	42.4%	57.6%	62.5%	37.5%
	12_인지능력 증강 윤리적 한계	22.4%	77.6%	51.3%	48.7%
	13_뇌데이터 프라이버시	37.6%	62.4%	57.5%	42.5%
	14_BCI 중독_의존	54.4%	45.6%	70.4%	29.6%

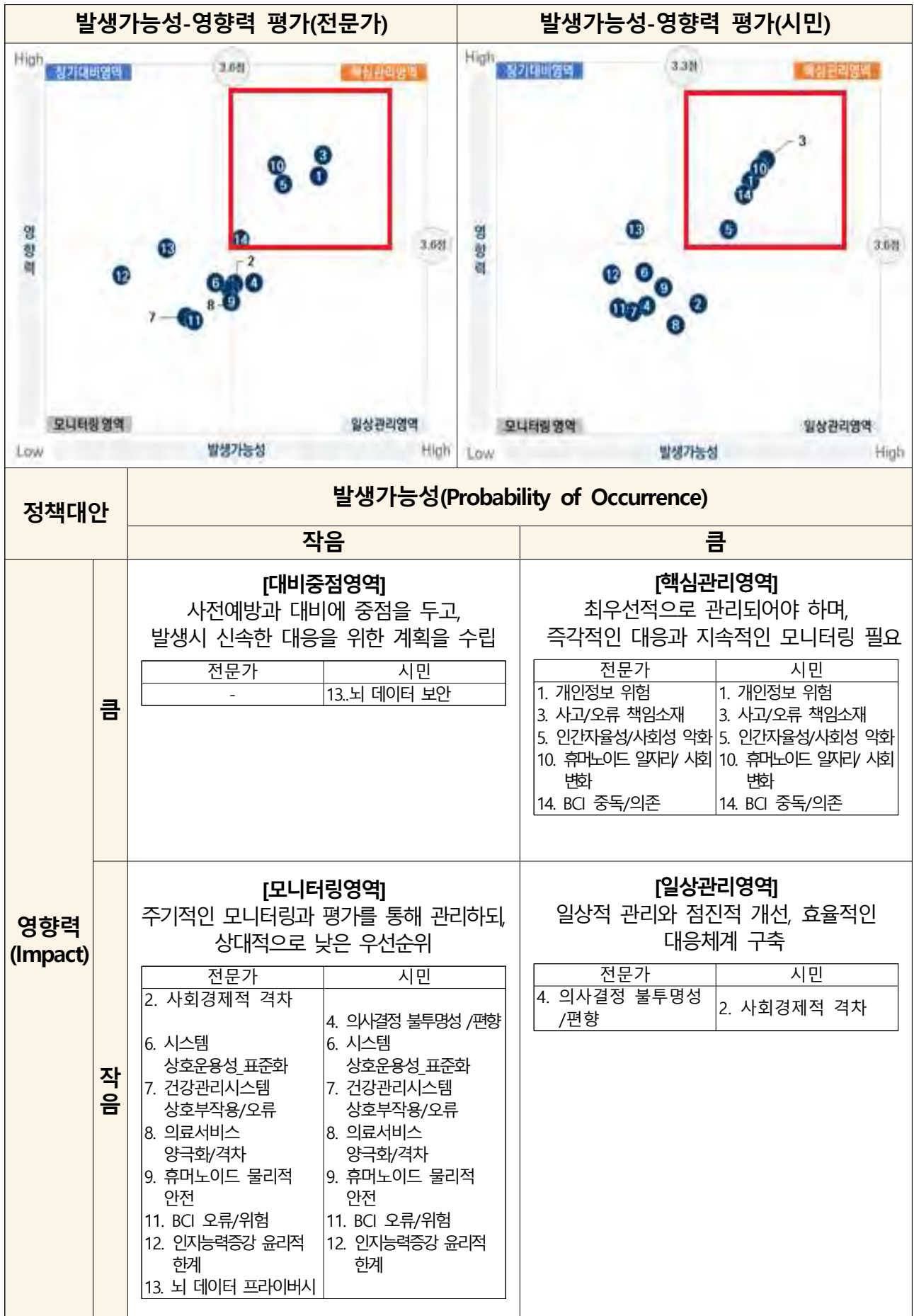
- **(핵심이슈)** 총 14개의 주요이슈 중 발생가능성 및 영향력*이 타 과제에 비해 상대적으로 높은 핵심이슈 5개**를 도출

* 전문가 125명, 시민 1,000명 대상으로 설문조사를 통해 주요이슈의 발생가능성, 영향력 점수의 평균값을 중심으로 영역을 구분하여 상대적 우선순위를 평가

** ①개인정보, ③사고/오류 책임소재, ⑤인간 자율성/사회성 악화, ⑩휴머노이드로 인한 일자리/ 사회변화, ⑭ BCI 중독/의존(※전문가, 시민 5개 이슈 모두 동일)

- 개인정보 위험, 사고오류의 책임소재, 인간자율성-사회성 악화 등 공통 이슈 3개와 휴머노이드 일자리/사회변화, BCI 중독/의존 등이 즉각적 대응이 필요한 과제로 판단

※ 시민들은 장기적인 대비가 필요한 이슈로 '뇌 데이터 프라이버시'를 지목



② 주요 과제의 평가

- **(평가기준)** 중요성, 시급성, 효과성, 실행가능성을 기준으로 주요 과제 평가

[표] AI 안전·신뢰 이슈에 대한 정책대안 평가 기준

질문 항목	개념	평가초점	세부 질문	응답 범위
중요성	해당 정책이 다루고자 하는 문제의 규모와 심각성, 그리고 정책 시행으로 인한 사회적 영향력의 크기	- 영향을 받는 인구의 규모 - 문제의 사회적 파급효과 - 정책 실패 시의 손실 정도	- 이 정책이 얼마나 많은 사람들에게 영향을 미치나요? - 이 정책이 해결하고자 하는 문제가 우리 사회에서 얼마나 큰 비중을 차지하나요?	1점: 전혀 중요하지 않음 2점: 중요하지 않은 편 3점: 보통 4점: 중요한 편 5점: 매우 중요함
시급성	정책적 개입의 시간적 제약성과 즉각적 대응의 필요성 정도	- 문제 해결의 시간적 여유 - 지연 시 예상되는 부정적 결과 - 조기 개입의 효과성	- 이 문제를 지금 당장 해결하지 않으면 어떤 일이 발생하나요? - 이 정책의 시행이 얼마나 빨리 필요한가요?	1점: 전혀 시급하지 않음 2점: 시급하지 않은 편 3점: 보통 4점: 시급한 편 5점: 매우 시급함
효과성	정책이 추구하는 방향과 결과가 사회적 가치와 규범에 부합하는 정도	- 사회적 가치와의 정합성 - 정책 목표의 정당성 - 공익성 달성 정도	- 이 정책이 시행되면 우리 사회가 더 나은 방향으로 발전할까요? - 이 정책이 우리가 바라는 가치나 목표에 부합하나요?	1점: 전혀 바람직하지 않음 2점: 바람직하지 않은 편 3점: 보통 4점: 바람직한 편 5점: 매우 바람직함
실행가능성	주어진 자원과 환경 하에서 정책을 실제로 수행할 수 있는 가능성의 정도	- 기술적 실현가능성 - 재정적 실현가능성 - 행정적 실현가능성 - 정치적 실현가능성	- 이 정책을 실제로 시행하는 것이 현실적으로 가능할까요? - 필요한 예산, 인력, 기술 등을 확보할 수 있나요? - 시행 과정에서 예상되는 어려움이나 반대를 극복할 수 있나요?	1점: 전혀 실현가능하지 않음 2점: 실현가능하지 않은 편 3점: 보통 4점: 실현가능한 편 5점: 매우 실현가능함

- **(핵심과제)** 총 17개의 주요과제에 대해서 시급성 및 중요성을 기준으로 평가하여 최우선으로 관리되어야 할 핵심과제 9개*를 도출

* ①미래기술 안전성 기본법, ②인간친화 기술 윤리안전 가이드라인, ③민감정보 기반 기술 거버넌스, ⑧미래기술 안전성 R&D, ⑨헬스케어AI 임상안전제도, ⑩의료AI 데이터 표준화, ⑪의료AI 실증플랫폼, ⑫로봇 물리안전 기준, ⑭로봇협업 안전관리제도 (※전문가, 시민 8개가 동일하며, ⑫로봇 물리안전 기준은 전문가만 선택)

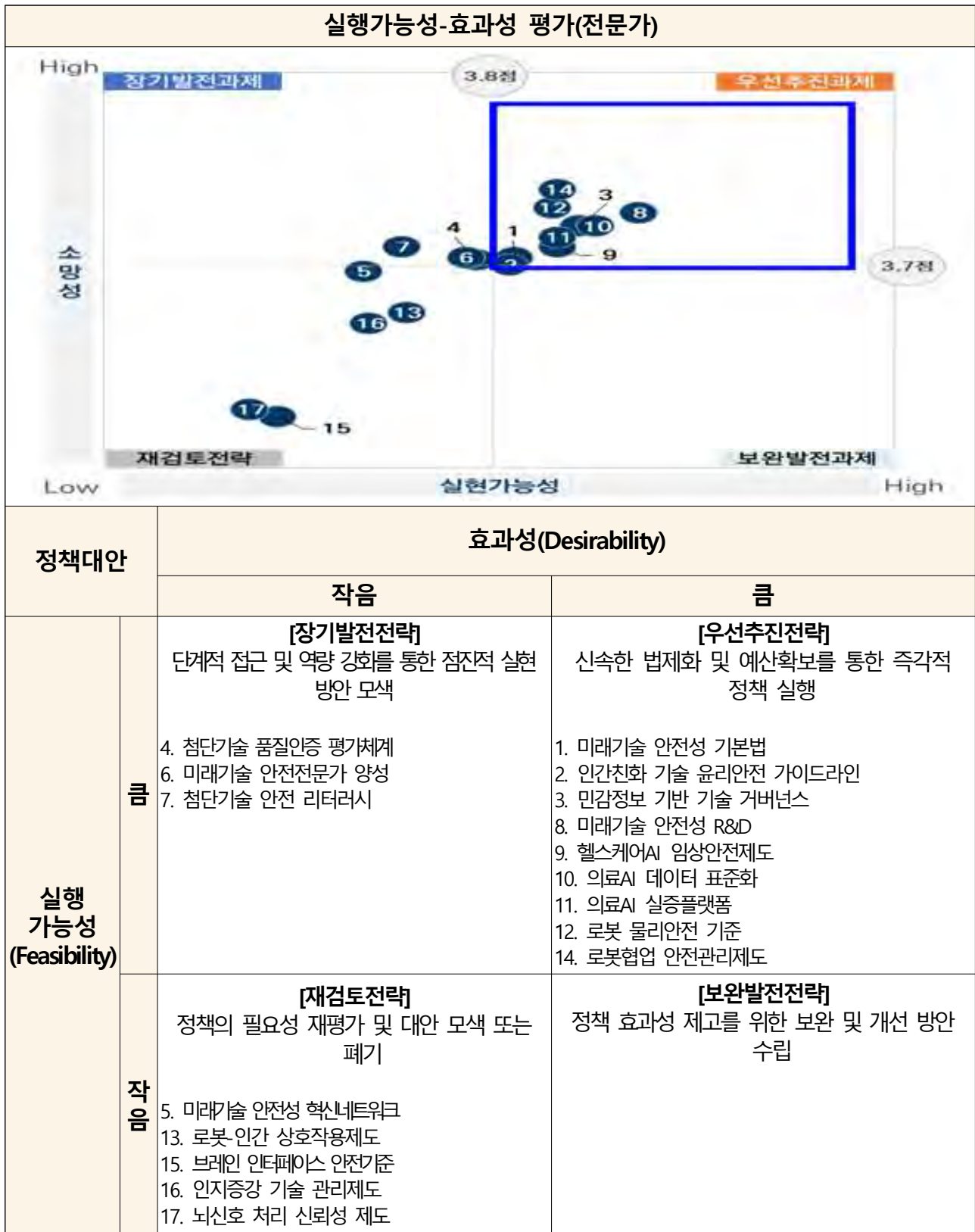
- **(시급성-중요성)** 즉각적 자원 투입과 신속한 정책 실행이 필요한 핵심 우선과제는 미래기술 안전성 기본법 등 9개 과제로 전문가, 시민의 의견 유사



정책대안		시급성(Urgency)																									
		작음	큼																								
중요도 (Importance)	큼	[전략적준비과제] 장기적 관점에서의 체계적 준비와 계획 수립 <table><tr><td>전문가</td><td>시민</td></tr><tr><td>-</td><td>12. 로봇 물리안전 기준</td></tr></table>	전문가	시민	-	12. 로봇 물리안전 기준	[핵심우선과제] 즉각적인 자원 투입과 신속한 정책 실행 <table><tr><td>전문가</td><td>시민</td></tr><tr><td>1. 미래기술 안전성 기본법</td><td>1. 미래기술 안전성 기본법</td></tr><tr><td>2. 인간친화 기술 윤리안전 가이드라인</td><td>2. 인간친화 기술 윤리안전 가이드라인</td></tr><tr><td>3. 민감정보 기반 기술 거버넌스</td><td>3. 민감정보 기반 기술 거버넌스</td></tr><tr><td>8. 미래기술 안전성 R&D</td><td>8. 미래기술 안전성 R&D</td></tr><tr><td>9. 헬스케어시 임상안전제도</td><td>9. 헬스케어시 임상안전제도</td></tr><tr><td>10. 의료시 데이터 표준화</td><td>10. 의료시 데이터 표준화</td></tr><tr><td>11. 의료시 실증플랫폼</td><td>11. 의료시 실증플랫폼</td></tr><tr><td>12. 로봇 물리안전 기준</td><td></td></tr><tr><td>14. 로봇협업 안전관리제도</td><td>14. 로봇협업 안전관리제도</td></tr></table>	전문가	시민	1. 미래기술 안전성 기본법	1. 미래기술 안전성 기본법	2. 인간친화 기술 윤리안전 가이드라인	2. 인간친화 기술 윤리안전 가이드라인	3. 민감정보 기반 기술 거버넌스	3. 민감정보 기반 기술 거버넌스	8. 미래기술 안전성 R&D	8. 미래기술 안전성 R&D	9. 헬스케어시 임상안전제도	9. 헬스케어시 임상안전제도	10. 의료시 데이터 표준화	10. 의료시 데이터 표준화	11. 의료시 실증플랫폼	11. 의료시 실증플랫폼	12. 로봇 물리안전 기준		14. 로봇협업 안전관리제도	14. 로봇협업 안전관리제도
	전문가	시민																									
-	12. 로봇 물리안전 기준																										
전문가	시민																										
1. 미래기술 안전성 기본법	1. 미래기술 안전성 기본법																										
2. 인간친화 기술 윤리안전 가이드라인	2. 인간친화 기술 윤리안전 가이드라인																										
3. 민감정보 기반 기술 거버넌스	3. 민감정보 기반 기술 거버넌스																										
8. 미래기술 안전성 R&D	8. 미래기술 안전성 R&D																										
9. 헬스케어시 임상안전제도	9. 헬스케어시 임상안전제도																										
10. 의료시 데이터 표준화	10. 의료시 데이터 표준화																										
11. 의료시 실증플랫폼	11. 의료시 실증플랫폼																										
12. 로봇 물리안전 기준																											
14. 로봇협업 안전관리제도	14. 로봇협업 안전관리제도																										
	작음	[모니터링과제] 주기적 재평가와 필요시 대응 방안 수립 <table><tr><td>전문가</td><td>시민</td></tr><tr><td>5. 미래기술 안전성 혁신 네트워크</td><td>5. 미래기술 안전성 혁신 네트워크</td></tr><tr><td>7. 첨단기술 안전 리터러시</td><td>7. 첨단기술 안전 리터러시</td></tr><tr><td>13. 로봇-인간 상호작용제도</td><td>13. 로봇-인간 상호작용제도</td></tr><tr><td>15. 브레인 인터페이스 안전 기준</td><td>15. 브레인 인터페이스 안전 기준</td></tr><tr><td>16. 인지증강 기술 관리제도</td><td>16. 인지증강 기술 관리제도</td></tr><tr><td>17. 뇌신호 처리 신뢰성 제도</td><td>17. 뇌신호 처리 신뢰성 제도</td></tr></table>	전문가	시민	5. 미래기술 안전성 혁신 네트워크	5. 미래기술 안전성 혁신 네트워크	7. 첨단기술 안전 리터러시	7. 첨단기술 안전 리터러시	13. 로봇-인간 상호작용제도	13. 로봇-인간 상호작용제도	15. 브레인 인터페이스 안전 기준	15. 브레인 인터페이스 안전 기준	16. 인지증강 기술 관리제도	16. 인지증강 기술 관리제도	17. 뇌신호 처리 신뢰성 제도	17. 뇌신호 처리 신뢰성 제도	[단기대응과제] 신속한 대응과 함께 중요도 제고 방안 모색 <table><tr><td>전문가</td><td>시민</td></tr><tr><td>4. 첨단기술 품질인증 평가 체계</td><td>4. 첨단기술 품질인증 평가 체계</td></tr><tr><td>6. 미래기술 안전전문가 양성</td><td>6. 미래기술 안전전문가 양성</td></tr></table>	전문가	시민	4. 첨단기술 품질인증 평가 체계	4. 첨단기술 품질인증 평가 체계	6. 미래기술 안전전문가 양성	6. 미래기술 안전전문가 양성				
전문가	시민																										
5. 미래기술 안전성 혁신 네트워크	5. 미래기술 안전성 혁신 네트워크																										
7. 첨단기술 안전 리터러시	7. 첨단기술 안전 리터러시																										
13. 로봇-인간 상호작용제도	13. 로봇-인간 상호작용제도																										
15. 브레인 인터페이스 안전 기준	15. 브레인 인터페이스 안전 기준																										
16. 인지증강 기술 관리제도	16. 인지증강 기술 관리제도																										
17. 뇌신호 처리 신뢰성 제도	17. 뇌신호 처리 신뢰성 제도																										
전문가	시민																										
4. 첨단기술 품질인증 평가 체계	4. 첨단기술 품질인증 평가 체계																										
6. 미래기술 안전전문가 양성	6. 미래기술 안전전문가 양성																										

- **(실행가능성-효과성)** 신속한 법제화 및 예산 확보를 위한 즉각적 정책 실행이 필요한 우선추진전략은 핵심우선과제와 동일

※ 효과성(소망성)은 사회적 가치와의 정책목표의 정당성, 공익성 달성정도, 실행가능성은 기술적, 재정적, 행정적, 정치적 실현가능성을 고려하여 평가하므로 전문가만 평가 수행



③ 평가결과 종합

○ **(핵심과제)** 주요 이슈·과제에 대한 평가결과를 종합하여 향후 관리 되어야할 9개 핵심 과제를 도출

- **(핵심과제)** 미래기술 안전성 기본법 등 9개 과제는 중요성-시급성, 효과성-실행가능성 측면에서 우선순위가 높으므로 '핵심과제'로 관리

※ 해당 과제는 관련 정책 등을 추가적으로 분석하여 정책 이해관계자 협의를 거쳐 세부 환류계획을 수립후 관리

- **(추가검토)** 첨단기술 품질인증·평가체계, 미래기술 안전전문가 양성 등 3개 과제는 시급하나 중요성은 다소 낮아 대응 필요 추가 검토

- **(재검토)** 미래기술 안전성 혁신네트워크 등 5개 과제는 중요성 및 시급성이 타 과제에 비해 상대적으로 낮아 일정 시기 이후 재검토

[표] 17개 주요 과제별 관리 방안

주요 과제	개념	중요성-시급성 (시민)	중요성-시급성 (전문가)	효과성-실행 가능성(전문가)	검토
미래기술 안전성 기본법	인간과 직접 상호작용하는 미래기술의 안전성 보장을 위한 법적 기반을 마련하고, 관련 제도·조직·지원의 근거를 제시하는 기본법 체계	핵심우선	핵심우선	우선추진	핵심과제
인간친화 기술 윤리안전 가이드라인	미래기술이 인간에 미치는 영향을 고려한 윤리적 개발·활용 원칙을 수립하고, 이를 실현하기 위한 제도적 지원을 제공하는 정책 체계				
민감정보 기반 기술 거버넌스	미래기술 활용과정에서 발생하는 초민감정보의 안전한 수집·활용·보호를 위한 체계적 관리 기반을 구축하는 제도적 체계				
미래기술 안전성 R&D	미래기술의 안전성 향상을 위한 핵심기술을 개발하고, 실증·검증하는 연구개발 체계				
헬스케어 AI 임상안전 제도	의료현장에서 AI 기술의 안전한 도입·활용을 위한 임상 안전성 보장 체계				
의료 AI 데이터 표준화	의료 AI의 신뢰성 확보를 위한 의료데이터의 표준화와 품질관리 체계				
의료 AI 실증 플랫폼	의료 AI 시스템의 실제 효과와 안전성을 검증하기 위한 실증 체계				

주요 과제	개념	중요성-시급성 (시민)	중요성-시급성 (전문가)	효과성-실행 가능성(전문가)	검토
로봇협업 안전관리 제도	인간-로봇 협업 환경의 안전성 보장을 위한 종합 안전관리 체계	전략적준비			
로봇 물리안전 기준	휴머노이드 로봇의 물리적 안전성 보장을 위한 기술적·제도적 안전 체계				
첨단기술 품질인증·평가 체계	미래기술의 안전성과 신뢰성 보장을 위한 품질관리 체계를 구축하고, 이를 검증·인증하는 종합적 품질보증 체계	단기대응	단기대응	장기발전	대응필요성 추가 검토
미래기술 안전 전문가 양성	미래기술 안전성 확보를 위한 전문인력을 체계적으로 양성하고, 지속적인 역량 개발을 지원하는 인력양성 체계				
첨단기술 안전 리터러시	미래기술의 안전한 활용을 위한 사용자 역량을 개발하고, 안전문화를 조성하는 교육 체계	모니터링	모니터링	재검토	일정시기 이후 재검토
미래기술 안전성 혁신 네트워크	미래기술 안전성 향상을 위한 산학연 협력 기반을 구축하고, 공동연구 및 기술교류를 촉진하는 협력 체계				
브레인 인터페이스 안전기준	뇌-기계 인터페이스의 안전한 개발과 활용을 위한 기술적·제도적 안전 체계				
인지증강 기술 관리제도	인지증강 기술의 안전한 사용과 부작용 예방을 위한 종합 관리 체계				
뇌신호 처리 신뢰성 제도	뇌신호 처리의 정확성과 안전성 보장을 위한 기술적 신뢰성 확보 체계				
로봇-인간 상호작용 제도	휴머노이드 로봇과 인간의 안전한 상호작용을 위한 종합 관리 체계				

○ (핵심이슈-과제) 5개 핵심 이슈와 9개 핵심 과제의 연계성



[참고] 9개 핵심과제 관련 정책 시범 분석

○ 9개 핵심과제와 관련된 정책을 관련 법령 및 정책 DB 및 생성형 AI 활용하여 분석

- * Claude 3.5 Sonnet, Perplexity 등 생성형 AI를 활용하여 초안 작성 후, 국가법령정보센터, 입법예고, 행정예고, NTIS, K2Base 등 검색하여 검증
- * 본 분석은 전문가의 검증을 거치지 않은 내용으로 향후 구체적 환류 방안 탐색을 위해 분석 수행

핵심과제	관련 정책(사례)	개선방향		소관
		신규	보완	
미래기술 안전성 기본법	- 지능정보화 기본법 제21조: "지능정보기술 안전성·신뢰성" - AI 산업 발전 및 신뢰 기반 조성에 관한 법률안(2023) - 입법예고(2024.1): AI 산업 발전 및 신뢰 기반 조성에 관한 법률안 - 생체정보 활용 AI 안전기준 - 위험등급별 관리체계	- 생체-기계 인터페이스 조향 - 실시간 위험관리 기준	- 안전성 심사위원회 - 사고배상 체계	과기 정통부
인간친화 기술 윤리안전 가이드라인	- AI 윤리기준(2020): "사람중심 AI" - 신뢰할 수 있는 AI 개발 가이드라인 (2021)	- 인체개입 윤리기준 마련 - 자율성 한계 설정	- 분야별 세부지침 - 평가인증 체계	과기 정통부
민감정보 기반 기술 거버넌스	- 개인정보보호법 제23조: "민감정보 처리 제한" - 생명윤리 및 안전에 관한 법률	- 초민감정보의 범위 등 분류체계 마련 - 민감정보의 실시간 수집기준 마련	- 통합관리 체계 - 보안기준 수립	개보위 복지부
미래기술 안전성 R&D	- AI 기술개발 사업(2023) - AI안전성강화R&D프로그램	- 연구분야 확대 - 실증 강화	- 통합R&D체계 - 성과관리제도	과기 정통부
헬스케어 AI 임상안전 제도	- 의료기기법 - AI의료기기허가·심사 가이드라인	임상검증 강화	- 사고예방체계 - 책임규정마련	복지부 식약처
의료 AI 데이터 표준화	- 보건의료 데이터 표준화 로드맵(2022) - 의료데이터중심병원지원사업	- 표준화 범위 확대 - 품질기준 구체화	- 통합관리체계 - 인증제도도입	복지부
의료 AI 실증 플랫폼	국가전략기술 특화연구소 데이터플랫폼에 대한 실증특례	다기관 임상시험 지원 체계 구축	의료 AI 실증을 위한 규제 샌드박스 확대	복지부
로봇 물리안전 기준	지능형 로봇 개발 및 보급 촉진법	-	- 로봇의 물리적 안전성에 대한 구체적인 기준 마련 - 안전성 평가 및 인증 체계 구축	산업부
로봇협업 안전관리 제도	-	인간-로봇 협업 환경의 안전성 보장을 위한 종합 안전관리 체계 마련	-	산업부

III. 시사점 및 개선방향(안)

□ 시사점

- 기술영향평가 결과의 정책반영률이 지속적으로 하락*함에 따라 환류 체계의 실효성이 저하

* ('21,자율주행) 77.8% → ('22,합성생물학) 71.4% → ('23,양자과학기술) 68.8%

⇒ 기술영향평가의 역할과 기능, 환류체계를 재점검 및 개편하여 기술영향평가의 실효성 제고 필요

□ 개선방향(안)

개선과제	AS-IS	TO-BE
기본 방향	① 미래기술 담론 형성에 초점 : 미래기술 관련 시민 담론 형성에 초점 → 정책과제화, 환류 등의 연계에 한계	① 미래기술이슈 종합대응체제로 개편 : 이슈탐지→정책과제 분석→정책환류 3단계 미래기술이슈 대응 체제로 개편
대상 기술	② 장기 미래기술 평가로 환류 곤란 : 장기적으로 국민생활에 파급효과를 미칠 기술을 평가하고 있으나, 긴 시차로 정책환류가 불가	② 단·중기 확산 기술 대상 평가 : 단기 및 중기에 사회적으로 확산될 기술, 연구개발이 진행될 기술 대상 평가 실시
핵심 이슈	③ 평가항목별 기계적 이슈 도출 : 경제·사회·문화·윤리·환경 등 시행령상 기준을 기계적 적용하여 핵심이슈 파악 곤란	③ 핵심이슈 중심으로 이슈 탐지 : 발생가능성, 영향력 등 고려하여 핵심이슈 중심으로 도출
정책 분석	④ 추상적 정책 제언 : 담론적·추상적인 정책제언으로 평가결과 환류방안의 구체성 미흡	④ 핵심과제 도출·목록화 : 국가·사회적 필요 기술분야에 대한 구체적 방향성을 제시할 수 있도록 핵심과제를 법령 계획 사업 등 유형별로 목록화(관련부처 협의 필요)
환류	⑤ 단기·1회성 정책 환류 : 평가결과 반영점검이 1회만 진행되고, 부처의 참여 유인도 낮아 환류의 지속성 부재	⑤ 정책협의체 구성·추적관리 : 평가단계에서 대상기술별 관련 기관이 참여하여 최소 3년간 주요정책과제 조정·관리 등 환류체계 구축

□ 세부과제(안)

1 [기본방향] 기술영향평가를 미래기술 이슈 종합 대응체제로 개편

- **(문제점)** 미래기술 관련 시민 담론형성에 초점*을 두어 정책과제화 및 평가결과의 환류에 다소 미온적

* 1990년대 후반 복제양 돌리, 인터넷 확산 등으로 시민사회에서 과학기술의 부작용에 대한 우려 ⇒ 과기기본법·시행령 제정(2001)시부터 “민간전문가 및 시민단체 등 참여를 확대하고 일반국민의 의견을 모아 실시”할 것을 명시(시행령 23조)

- **(개선)** 미래기술 잠재이슈 탐지 → 정책과제 분석 → 정책 환류의 3단계로 전환하여 미래기술 이슈에 종합 대응하는 체제로 전환

※ 미국은 의회 산하 기술평가국(OTA)을 1995년 폐지하였으나, '19년 회계감사원(GAO)에 과학기술평가분석부서(STAA)를 부활시켜 미래사회 기술이슈 대응 역량을 강화

2 [대상기술] 가까운 미래에 확산될 혁신기술을 대상기술로 선정

- **(문제점)** 장기적으로 파급력이 나타날 기술을 평가하고 있어 환류 지속이 쉽지 않고, 기간이 장기화*되는 경우 더욱 환류가 곤란

* 2011년 기술영향평가 Brain-Computer Interface 기술이 2022~2025년에 사회적으로 보급될 것으로 예상하였으나, 뉴럴링크 등 실용화 기술이 최근에 가시화

- **(개선)** 단·중기적으로 사회적 파급력이 나타날 기술을 평가 대상 기술로 선정하고, 기술영향평가위원회*가 대상기술을 추천·관리

※ 분야별 기술 전문가 중심 대상기술추천위원회, 기술영향평가 총괄위원회 구성·운영 ⇒ 기술사회정책 분야의 리더십 있는 전문가 중심으로 임기제 상설영향평가위원회 운영

3 [①-핵심이슈단계] 우선순위가 높은 핵심이슈 도출

- **(문제점)** 경제·사회·문화·윤리·환경 등 시행령상 기준*을 기계적으로 적용하여 핵심이슈 파악 곤란

* 과학기술기본법 시행령 제23조 제2항 : 새로운 과학기술이 경제·사회·문화·윤리 및 환경에 미치는 영향을 평가에 포함하여야 함

- (개선) 전문가·시민포럼*의 의견을 반영하여 주요이슈를 도출하고, 발생 가능성, 영향력을 기준으로 평가**하여 핵심이슈를 발굴

* 전문가 그룹과 동일한 방식의 시민포럼 운영으로 의견이 전문가에 종속 ⇒ 시민 포럼에게 미션을 통한 상황 물입 후에 심층 면접을 통한 심도 있는 토론 등 구조화된 방식을 적용하여 생생한 시민 의견을 도출('24년 시범 추진)

** 전문가 및 시민 대상 설문조사를 통해서 주요 이슈의 발생가능성, 영향력 점수의 평균값을 중심으로 영역을 구분하여 상대적 우선순위를 평가('24년 시범 추진)

주요 이슈		발생가능성(Probability of Occurrence)	
		작음	큼
영향력 (Impact)	큼	[대비중점영역] 사전예방과 대비에 중점을 두고, 발생시 신속한 대응을 위한 계획을 수립	[핵심관리영역] ⇨ 핵심이슈 최우선적으로 관리되어야 하며, 즉각적인 대응과 지속적인 모니터링 필요
	작음	[모니터링영역] 주기적인 모니터링과 평가를 통해 관리하되, 상대적으로 낮은 우선순위	[일상관리영역] 일상적 관리와 점진적 개선, 효율적인 대응체계 구축

4 [②-정책분석단계] 핵심과제 도출 · 목록화 관리

- (문제점) 담론 · 추상적인 정책제언으로 평가결과 환류의 구체성 미흡
- (개선) 실행가능한 이슈별 대응과제를 법령, 계획, 사업 등 유형별로 목록화하여 추적 · 관리

- (①**정책과제**) 관련기관, 정책전문가 중심으로 주요과제를 도출
- (②**핵심과제**) 주요과제를 실행가능성 및 효과성을 기준으로 평가
⇒ **핵심과제를 도출**(우선추진과제 영역)

주요 과제		실행가능성(Feasibility)	
		작음	큼
효과성 (effectiveness)	큼	[장기발전과제] 단계적 접근 및 역량 강화를 통한 점진적 실현 방안 모색	[우선추진과제] ⇨ 핵심과제 신속한 법제화 및 예산확보를 통한 즉각적 정책 실행
	작음	[재검토과제] 정책의 필요성 재평가 및 대안 모색 또는 폐기	[보완발전과제] 정책 효과성 제고를 위한 보완 및 개선 방안 수립

- (③**정책분석**) 핵심과제 관련정책(법령, 입법예고, 기본계획 등), 사업 분석
- (④**목록화**) 관련 정책 및 사업을 유형화하고 기존보완 · 신규정책 등 구분하여 목록화

5 [③-환류단계] 정책유형별 환류 및 추적

- **(문제점)** 평가결과 반영 점검이 차년도 1회만 진행되고, 법상 근거*에도 불구하고 관련부처 등이 참여 유인이 없어 환류의 지속성 부재

* 과학기술기본법 시행령 제23조 제7항 : 관계 중앙행정기관의 장은 기술영향평가 결과를 통보받았을 때에는 이를 소관 분야의 국가연구개발사업에 대한 연구기획에 반영하거나 부정적 영향을 최소화하기 위한 대책을 세워 추진하여야 함

- **(개선)** 정책 유형(법령, 입법예고, 기본계획 등)은 정책협의체*를 통해서 정책과제 환류를 추적·관리**하고, 사업 유형은 예산·평가와 연계***

* 대상기술별로 정책분석 단계에 정책협의체 구성을 통해 정책환류 목표 시점까지 운영하는 등 미래기술 이슈 대응

** 상설 기술영향평가위원회가 연간/반기별 추진실적 점검, 정책협의회 등을 통해 추적·조정·관리하고, 매년 기술영향평가 결과 보고시 환류 결과를 보고

*** 연계 방안(예시)

※ 정책환류의 효율적 운영을 위해 기술영향평가와 별도의 인력, 예산으로 과제화하여 운영하는 방향 고려

관련정책	① 정책 유형(법령, 기본계획 등)	② 사업 유형	
		예산 연계(안)	평가 연계(안)
연계방안	정책협의체 통한 환류 추적	투자방향 반영 예산심의 반영	중간평가에서 가점 항목에 반영

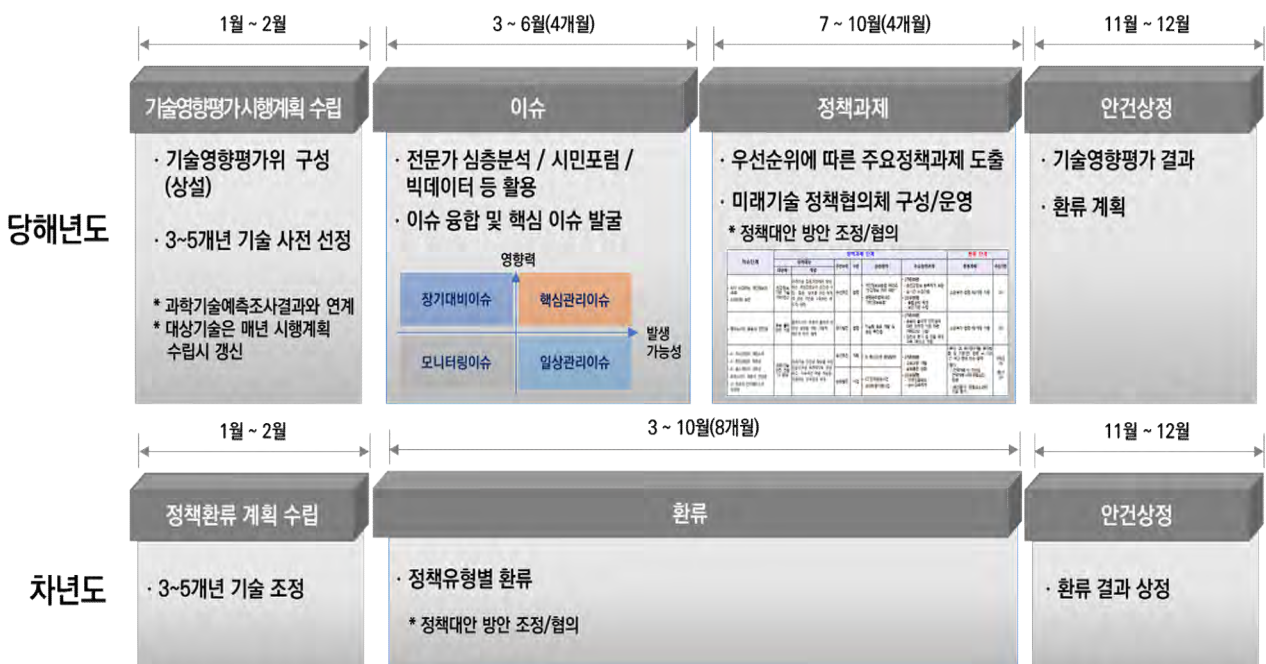
기술영향평가 정책분석 목록(예시 : 안전·신뢰 AI)

①핵심이슈	②정책분석					③환류	
	핵심과제		유형	관련정책	주요정책과제	환류계획	추진 기한
	과제명	개념					
- AI가 수집하는 개인정보의 보호 - 뇌데이터 보안	민감정보 기반 기술 거버넌스	미래기술 활용과정에서 발생하는 초민감정보의 안전한 수집·활용·보호를 위한 체계적 관리 기반을 구축하는 제도적 체계	법령	- 개인정보보호법 제23조 "민감정보 처리 제한" - 생명윤리법제18조 "개인정보보호"	[기존보완] - 초민감정보 분류체계 보완 - 실시간 수집기준 [신규정책] - 통합관리 체계 - 보안기준 수립	정책협의체 통한 추적관리	'26
휴머노이드 로봇의 안전성	로봇 물리안전 기준	휴머노이드 로봇의 물리적 안전성 보장을 위한 기술적·제도적 안전 체계	법령	지능형 로봇 개발 및 보급 촉진법	[기존보완] - 로봇의 물리적 안전성에 대한 구체적 기준 마련 (제9조의2 신설) - 안전성 평가 및 인증 체계 구축 (제10조 개정)		'26
- AI 의사결정의 책임소재 - AI 판단과정의 투명성 - AI 헬스케어의 신뢰성 - 휴머노이드 로봇의 안전성 - 뇌-컴퓨터 인터페이스의 안전성	미래기술 안전 전문가 양성	미래기술 안전성 확보를 위한 전문인력을 체계적으로 양성하고, 지속적인 역량 개발을 지원하는 인력양성 체계	계획 사업	- AI 혁신인재 양성방안 - ICT인재양성사업 - AI대학원지원사업	[기존보완] - 교육과정 개발 - 실무훈련 강화 [신규정책] - 자격인증제도 - 보수교육체계		[예산] '26 [평가] '29

※ 위 표는 예시이며, 일부 사실과 부합하지 않을 수 있음

□ 개선(안) 관련 추진일정(예시)

- 기술영향평가 시행계획 수립(~'25.2월)
 - 기술영향평가위원회 구성
 - 향후 3개년 평가대상 기술 후보(안) 마련
- '25년 기술영향평가 실시('25.3~'25.10월)
 - 이슈 도출('25.3~'25.6월)
 - 정책협의회 구성, 정책과제 도출·조정('25.7~'25.10월)
- '25년 기술영향평가 국과심 운영위 결과보고('25.12월)
 - 환류 계획('25.11~'25.12월)
- '26년 기술영향평가 시행계획 수립('26.1~'26.2월)
 - 필요시 평가대상 기술 후보(안) 조정
- '25년 결과 환류계획 조정 및 추적·관리('26.3~'26.10월)
- '25년 결과 환류 1차년 결과 정리('26.11~'26.12월)



IV. 향후 일정

□ 기술영향평가 결과 관계부처 통보 [24. 12]

- 기술영향평가 결과에 대한 부처별 정책 반영* 실적 여부를 확인하여 ('25.10~11), '25년 기술영향평가 결과(안)에 포함 예정

* (근거) 과학기술기본법 시행령 제23조 제7항

□ '안전 · 신뢰 AI 기술의 미래(가칭)' 책자 발간 [25. 상반기]

- 대중 교양서 수준의 단행본 책자를 발행 후 대학, 정부 출연(연), R&D 전문기관 등에 배포

□ 기술영향평가 결과를 동영상 콘텐츠화하여 유튜브 과학전문 채널에 업로드 [25. 상반기]

- 동영상 콘텐츠 제작하여 유튜브 채널을 통해 기술영향평가 결과를 확산 ⇒ 대국민 과학기술정책 참여 접근성 제고

* (예시) 대중 접근성 높은 콘텐츠를 제작하여 파급력 높은 채널을 통해 미래기술에 대한 대중의 인식 제고

붙임 1

연도별 기술영향평가 대상기술 및 주요 결과

- 2001년: 과학기술기본법 제정으로 기술영향평가 의무화
 - ※ 2002년: 기술영향평가 정의, 연구범위, 운영방안 등 기획연구 시행
- 2003년 ~ 2023년: 총 24개 기술에 대해 기술영향평가 수행

연도	대상기술	주요 결과
2003	NBIT* 융합기술	· 융합기술로의 연구개발 편중 현상과 나노 격차로 인한 사회적 불평등 문제 발생 가능
2005	RFID**	· 프라이버시 침해 등의 부작용도 예상되므로 개인정보 보호를 위한 기술개발과 제도구축 등의 보완책 강구 필요
	나노	· 나노입자 인체흡입 등 생산 및 연구 현장에서의 안정성 문제 우려
2006	줄기세포 치료기술	· 연구자에 대한 윤리교육과 관련 자료의 국가적 DB 구축, 응급임상의 관리강화 등 필요
	나노소재	· 나노소재 규격화 및 인증을 위한 평가 기준이 필요하며, 나노기술영향평가센터 설치에 대한 면밀한 검토 필요
	UCT***	· 개인정보보호법 제정이 시급하며, 사회 전반적으로 적용될 수 있는 보안모델 필요
2007	기후변화 대응기술	· (신재생에너지기술) 에너지 공급체계가 자원 중심에서 기술 중심으로 바뀔에 따라 산업구조 변화와 신산업 창출 예상 · (기후변화적응기술) 기후변화로 인한 피해는 취약성에 따라 국가 간 파급효과가 다르게 나타나며 현재 예측보다 빠르고 심각할 수 있으므로 지속적 기술개발 필요
2008	국가재난 질환 대응기술	· 연구기반 인프라 확충 필요, 백신·치료제 비축, 위험 인식에 대한 대국민 소통 전략 수립, 토양 지하수 오염 및 향후 매몰지 부족 등의 환경문제 개선 · 민관협력을 통한 재난·질환 대응체계 강화
2011	뇌로 움직이는 미래세상 : 뇌-기계 인터페이스	· 사회적 공감대 형성을 위한 소통 강화, 개인 뇌 정보 등 프라이버시 보호를 위한 선제적 대책 마련 · 임상시험 가이드라인, 기관윤리위원회 심의 등 기술개발 과정에서의 윤리성 및 안전성 확보 · BMI 제품, 콘텐츠의 안전성 확보 및 오남용, 독점 등 방지를 위한 국가 차원의 기준 마련
2012	세상을 보는 새로운 눈: 빅데이터	· 신뢰성 있는 빅데이터 허브 운영, 공공데이터의 공유·연동 확대 추진 및 관련 핵심 기반 기술 국산화와 분석 전문 인력 및 활용 인력 양성 추진 · 빅데이터 환경에 적합한 개인정보보호 법·제도와 기술적 기반 마련

연도	대상기술	주요 결과
2013	3차원 프린팅	· 3차원 프린팅의 성능향상, 사업화 등을 포괄하는 종합정책 수립 및 불법·무단 제조 제품의 유통·판매 규제와 기술 진흥을 위한 제도 정비 · 3차원 프린팅 전문인력 양성 및 기술 수용성 확산 · 원소재별·공정별 표준화된 기술기준 개발·보급 및 기존 장비, S/W의 단점을 국산 기술로 극복하기 위한 신기술 개발
	스마트네트워크	· 산업별 적합 기술개발 및 관련 기업 활성화를 위한 정책 지원 · 프라이버시 침해·사이버테러 대응을 위한 방안 마련 및 디지털 격차 해소방안 마련과 소비자의 다양한 선택권 확보 · 스마트라이프 구현을 위한 현실적인 제도 개선 및 인체·환경 안전성 향상을 위한 제도의 강화
2014	무인 이동체	· 관련 서비스·제조 산업 병행 육성 및 미래 시장창출을 위한 적극적이고 체계적 기술개발 정책 추진 · 전주기적(생산·운용·폐기) 친환경 전략·정책 수립 및 기술개발과 상용화(운용·관리·안전 등)관련 법·제도 마련 · 종사자들에 대한 책임·윤리 교육 및 신규 연구 분야 발굴 지원
	초고층 건축물	· 전주기적 기술개발 정책추진으로 투자 효과 극대화 및 친환경·에너지 효율적 통합 유지·관리 체계 수립 · 도시계획 및 공공성 확보 등 종합적 검토에 기반을 둔 허가 정책 수립 및 초고층 건축물 관련 법·제도 정비 · 동일 지역 거주민 이해 조정체계 강화 및 위화감 해소 유도
2015	유전자가위	· 유전질환 및 난치성 질환 치료에 적용 가능하고 인간 및 환경에 유용한 농축산물 개발이 가능 · 비특이성으로 인한 심각한 부작용 초래 가능성이 크고, 우생학적·사회적 차별 가능성 등을 내포하고 있음
	인공지능	· 생산성 향상 및 새로운 부가가치 창출이 가능하며, 개인형·맞춤형 도우미 등장 등으로 복지 서비스 확대에 기여 · 오작동에 따른 인명피해나 재산손실, 그리고 개인정보와 사생활에 대한 침해 위험 증가
2016	가상·증강현실	· 관련 산업의 정착 및 새로운 부가가치 창출을 위한 중장기 VR/AR 산업발전 전략 마련 · 콘텐츠/기기 개발, 활용 방안 및 기술의 부작용에 대한 체계적 연구 · 공유·제공되는 정보들의 정확성과 정보 접근의 다양성을 보장할 수 있는 장치 마련과 음란물 규제, 저작권 등 기존 법·제도 체계 검토(재정립)
2017	바이오 인공장기	· 원천 기술 확보와 기술 사업화 육성을 위해 국가 차원의 지원을 확대하고, 기술 개발과 활용 단계에서 발생 가능한 혼란 방지를 위해 선제적인 규제 및 제도적 장치 마련 · 인류의 안전한 복지를 위해 기술개발·활용될 수 있도록 기술의 사회적, 윤리적 쟁점 등에 대한 대중적 담론 형성 및 사회적 합의의 장 마련

연도	대상기술	주요 결과
2018	블록체인	· 글로벌 기술 경쟁력을 확보와 안정적인 산업 분야 도입을 위한 핵심 원천기술 연구 및 개발 지원 · 기술 활용을 위해 선결되어야 하는 기술제도적 문제점 점검 및 분야별 기술 도입 효용성에 대한 체계적 검토 · 관련 산업 진출 활성화를 위한 규제/법률에 대한 재검토와 대국민 교육 및 사회적 합의 추진
2019	소셜 로봇 기술	· 소셜 로봇의 원활한 데이터 수집 및 활용 관련 규제의 재정비 필요 · 로봇-인간 상호작용 연구, 소셜 로봇의 효용성 연구 등 인문·사회 및 과학적 측면의 연구 지원 · 경제적·기술적 취약계층을 위한 지원 대책 및 소셜 로봇 해킹에 대비한 사용자 보호 대책 마련
2020	정밀의료 기술	· 국가과제 등을 통하여 대규모 유전체 데이터를 산업적으로 활용할 수 있는 방안 마련 · 데이터의 체계적 관리 및 안전한 데이터 유통생태계를 만들어 나갈 수 있도록 지원 · 유전자 등록 방식을 개선하고 유전체 분석 허용기관을 확대하는 등 유전자 분석 관련 규제 검토
2021	레벨 4 이상 자율주행	· 자율주행 관련 핵심기술 확보를 위한 클라우드, AI, 플랫폼 등 ICT 기술 및 원천기술에 대한 R&D 지원 확대 · 교통약자 이동지원 서비스의 지속적 예산배정 및 확보를 위한 사회적 배려와 정부와 지자체의 강한 추진 필요 · 제조물 결함에 의한 사고원인 증명을 위해 마련된 자율주행 정보 기록장치의 강제 부착 및 사고조사위원회 설치 등의 방안을 실행하기 위한 후속 지원 등 필요
2022	합성생물학	· 바이오와 IT 융합 분야 미래 수요에 대응하는 전문인력을 양성하고 국제협력 확대 바탕으로 연구 생태계 조성 필요 · 사회적 수용성 제고를 위해 지식 전파 등 기능적 접근을 탈피하고 다양한 분야 전문가들이 참여하는 커뮤니케이션 전략을 마련 · 합성생물학에 대한 법적 정의를 명확히 하고, 적용 대상과 범주를 제한하고 규제 대상 여부 또는 규제의 예외적 적용과 배제요건을 제시
2023	양자과학기술	· 국내 양자 생태계 조성 및 표준화 연계 기술 개발 필요 · 전략적 유연성에 기반한 국제 다자간 기술 공동체 구축, 기술협력 채널 확보 및 활용체계에 대한 모니터링과 우회적 기술협력 방안 마련 등 필요

※ 2010년: 기술영향 평가주기 '매년' 으로 개정

※ '04년, '09년, '10년은 기술영향평가 미시행

* Nano Bio Information Technology : 나노기술, 생명기술, 정보기술을 결합한 미래형 융합기술

** Radio Frequency Identification : 무선데이터 송신장치

*** Ubiquitous Computing Technology : 언제 어디서나 네트워크에 접속할 수 있는 기술

주요평가결과(정책제언)	정책 등 반영 현황	담당부처
[1. 경제 분야 : 불확실성 속 전략 모색] □ (국내 양자 생태계 조성) 국내 산업의 강점을 활용하여 부품이나 소프트웨어 등 차별화가 가능한 영역을 모색하고 생태계 조성 ○ 원천기술, 시스템 HW/SW, 부품 및 응용 SW 분야 등 해외와의 차별화 영역을 모색하고, 관련 산업 생태계의 육성 필요 ○ 국내 강점 기술인 반도체, 제조업, ICT 기술 바탕으로 핵심 기술개발과 산업체에서 활용할 수 있는 프로그램 마련 ○ 양자 전문 인력 양성과 국내에서 활약할 수 있는 산업 생태계를 조성하는 인력 정책 마련 필요 ○ 양자과학기술 표준화 활동에 연구자가 보다 적극적으로 참여할 수 있도록 연구제도 개선과 인센티브 도입 필요	○ 퀀텀이니셔티브 발표('24.4월) - 투자확대 및 전략적 R&D 추진, 우수 인력 양성 및 확보, 산·학·연 퀀텀 플랫폼 구축, 글로벌 협력 강화, 산업화 지원 등 4대 전략 포함 - 산업화 촉진 관련 인프라 확충, 퀀텀 파운드리 생태계 기반 마련, 표준화 등 산업화 촉진을 위한 전략 포함 ○ 양자과학기술 및 산업에 관한 법률 시행('24.11월) - 국무총리 주재 범부처 양자전략위원회 설치, 양자 종합계획 수립, 산업화 지원 및 양자 실태조사 실시 등 규정 《향후계획》 ○ 퀀텀이니셔티브 전략적 이행방향 마련 등 양자 생태계 조성을 위한 지원 지속	과기정통부 (양자과학기술산업과)
□ (전략적 투자 대응) 국내외 정책 동향(지적재산권, 기술 이전 등)을 특허청 지속적으로 파악하고 기술의 투자 포트폴리오 다양화 필요 ○ 초기 단계의 양자 산업 창출을 위해 공공 수요 발굴, 인센티브, 금융 등을 포함한 종합적인 정책 마련이 필요	○ 양자 분야 특허 빅데이터 분석(~24.12.) - 양자 분야 국제 기술동향의 신속한 파악·분석 및 활용을 위하여 양자기술 관련 국내·외 특허동향을 분석 《향후계획》 ○ 국가과학기술자문회의에 분석 결과 보고 예정('25.1.)	과기정통부 (양자과학기술산업과)
□ (표준화 연계 기술개발) 양자 시장 선점을 위하여 기술 개발부터 표준화를 염두하고, 이를 국제 표준화까지 연계하는 것을 고려	○ 퀀텀이니셔티브 발표('24.4월) - 표준화 등 산업화 촉진을 위한 전략 포함 ○ 양자과학기술 및 산업에 관한 법률 시행('24.11월) - 국무총리 주재 범부처 양자전략위원회 설치, 양자 종합계획 수립, 산업화 지원 및 양자 실태조사 실시 등 규정 《향후계획》 ○ 퀀텀이니셔티브 전략적 이행방향 마련 등 양자 생태계 조성을 위한 지원 지속, 국제표준화 활동 등 참여 예정	과기정통부 (양자과학기술산업과)

주요평가결과(정책제언)	정책 등 반영 현황	담당부처
	○ 글로벌 양자 표준화기구 퀀사(QuINSA*) 출범식 개최('24.8월) * Quantum INdustrial Standard Association : 국내외 양자정보기술 전반의 표준화 동향 조사분석, 양자정보기술 표준개발 및 산업화 촉진, 국제 표준화 기구와 협력 등 수행 - 글로벌 회원사 : IBM, Keysight 등 22개 기업 참여 ○ 한국-미국(NIST) 간 표준화 협력 제안 ('24.1월, 양자기술산업과) - 양자통신, 양자센싱, 양자컴퓨팅 등 양자 기술 관련 공동 표준개발을 위한 협력 아이템 및 아이디어 제안 《향후계획》 ○ QuINSA 제1차 국제회의 개최('25년 9월 잠정) - 2월: 국제회의 개최 준비위원회 발족 - 3월: 표준 제·개정 절차, 기구 운영 규정, 조직 구성, 의장단 후보 추천 등 QuINSA 운영 방식 확정 - 4월: 개최국(회의 장소), 회의 안건 확정, 국내외 회원사 홍보 - 5월~9월 : 표준 아이템 발굴(국내외 회원사 대상 표준 아이템 의견 접수) - 10월: 회의 개최 이후 로드맵 모색	과기정통부 (정보통신 방송기술 정책과)
	○ 양자 분야 표준기술 획득을 목표로 하는 산·학·연에 표준특허 확보 전략 지원 ('24.3~) - 하이브리드 양자키 분배방법 및 망 관리 기술 표준개발 등 2개 과제 《향후계획》 ○ 표준특허 확보 전략 지원 최종보고회 ('24.12)	특허청 (산업재산 창출전략팀)

주요평가결과(정책제언)	정책 등 반영 현황	담당부처
[2. 사회 분야 : 사회적 수용성 제고] ☐ (상호협력 유도) 연구개발의 주체가 되는 기관 간의 상호협력을 통해 건전한 경쟁을 유도하는 지원 정책을 마련 ○ 기술 인프라의 주도권을 민간이 아닌 정부가 가지고 일반시민에 교육의 기회와 접근성을 보장 ○ 산업계 요구 등 기술 동향 파악하여 학생들이 산업에서 필요로 하는 기술·능력을 습득하고 지역 간 격차를 줄일 수 있도록 지원	○ 퀀텀이니셔티브 발표('24.4월) - 투자확대 및 전략적 R&D 추진, 우수 인력 양성 및 확보, 산·학·연 퀀텀 플랫폼 구축, 글로벌 협력 강화, 산업화 지원 등 4대 전략 포함 - 산업화 촉진 관련 인프라 확충, 퀀텀 파운드리 생태계 기반 마련, 표준화 등 산업화 촉진을 위한 전략 포함 《향후계획》 ○ 유관학문과의 융합을 유도를 통한 양자저변 확대 등 정책 방향 마련, 양자종합계획 수립시 반영	과기정통부 (양자과학 기술산업과)
☐ (기술에 대한 올바른 이해) 다양한 집단 간 소통확대와 보도집단에서도 과장된 보도를 자제토록 하는 가이드라인 자체 마련 필요 ○ 대중의 오해나 지나친 기대가 생기지 않도록 양자 과학기술과의 관련성을 정확하게 제시하여 과학 기술·산업 정책 수립 필요	-	-
☐ (개발 및 교육지원) 산업 안전을 위한 양자과학기술의 연구개발 및 효율적인 구현·활용을 위한 교육 지원 프로그램 마련 필요 ○ 표준이 확립되면 주요 제품이나 서비스도 양자암호화로 전환할 가능성이 높으며, 기관 등은 미리 교육 지원 프로그램 마련 필요 ○ 기존의 안전 설비 폐기가 아니라, 기존의 안전 정책과 양자과학기술의 적용 정책이 결합하는 효율적인 안전 법규 지침 마련	-	-
☐ (보편적 의료 사업) 양자과학기술 적용된 의료기기에 대하여 건강검진 필수 적용 등재 및 국민건강보험공단 우선 도입 검토 필요 ○ 양자과학기술이 인체에 미치는 영향을 파악하고, 용도·노출 시간 등 의료계의 활용 규정 마련 필요 ○ 중소기업이 도입하기에는 경제적 부담이 크므로, 컴퓨터 시스템 및 클라우드 운영을 하는 의료센터를 통해 접근 및 활용	-	-
☐ (저변 확대) 대중적 이해를 높이고, '양자물리학', '양자정보학'과 전기전자, 컴퓨터 등 유관 학문과의 융합 유도 필요	-	-

주요평가결과(정책제언)	정책 등 반영 현황	담당부처
[3. 문화 분야 : 바람직한 양자문화 정착] □ (문화 활동 지원) 과학관, 전시회, 과학축제 등을 통하여 기술의 정확한 정보를 제공하고, 다양한 콘텐츠를 통하여 여러 산업에서 활용될 수 있음을 쉽게 이해되도록 지원 □ (소통 확대) 과학커뮤니케이터 등 소통전문가 인력 풀을 확대하고, 기술의 올바른 정보를 확산하도록 지속적인 홍보 정책 마련 필요 ○ 양자과학기술에 대한 현황을 전달할 수 있는 플랫폼을 개설하여 정확한 정보 제공하거나 대중의 오해, 불필요한 불안감, 지나친 기대가 생기지 않도록 교육과 설명을 제공	○ 퀀텀코리아 2024 행사 개최('24.6) - 국내외 양자과학기술 연구자 및 관련기업, 정부 대표, 시민 등 참여하여 글로벌 양자 생태계 혁신 흐름 조망 * 키노트3개, 23개 발표, 주제별 세션, 62개 전시·홍보 부스 등 참관객 5,534명 - 퀀텀 토크 콘서트(궤도x페퍼톤스 이장원), 양자정보경진대회 등 다채로운 행사를 통해 양자기술 관심 제고 《향후계획》 ○ 퀀텀코리아 2025 개최를 위한 조직위원회 구성 등 차질없이 추진	과기정통부 (양자과학기술산업과)
[4. 윤리 분야 : 윤리적 지침 마련] □ (윤리적 지침 마련) 개인·공공·산업 등 다양한 분야에서 양자과학기술을 이용한 데이터 활용 증가에 대비해, 정부 차원의 양자기술 윤리 가이드라인(지침)을 선제적으로 마련 필요 ○ 유전자 편집, 생태계 조작 및 기후 공학과 같은 민감한 부분에 양자과학기술 적용을 대비하여 선제적인 윤리 지침 마련 필요 ○ 양자과학기술의 잠재적 이점과 단점에 대해 검토하고, 균형점을 찾는 사회적 영향에 대한 연구를 지속적으로 수행 필요 ○ 인문학과와의 융합을 통해 기술의 발전이 사회와 인간에 미칠 수 있는 영향을 이해하고, 미래 사회의 윤리적 고려사항 검토 필요 ○ 양자기술을 적절하게 활용하고, 사회적 책임을 다할 수 있도록 윤리 교육을 강화 필요 ○ 정부 입안자는 양자컴퓨팅이 야기할 수 있는 잠재적 문제에 대비할 수 있도록 연구개발 단계부터 윤리적 사용의 중요성을 강조할 필요가 있음	○ 퀀텀이니셔티브 발표('24.4월) - 투자확대 및 전략적 R&D 추진, 우수 인력 양성 및 확보, 산·학·연 퀀텀 플랫폼 구축, 글로벌 협력 강화, 산업화 지원 등 4대 전략 포함 - 산업화 촉진 관련 인프라 확충, 퀀텀 파운드리 생태계 기반 마련, 표준화 등 산업화 촉진을 위한 전략 포함) 《향후계획》 ○ 중장기 정책 수립시 반영 필요성 등 종합 검토	과기정통부 (양자과학기술산업과)
[5. 법률·규제 분야 : 범부처 거버넌스 구축] □ (양자기술 발전을 위한 제도 마련) 특허 및 품질인증 체계, 가이드라인 등 선제적 대비 및 장기적 육성 차원 관련 규정 마련하고 양자과학기술의 개발과 활용을 위한 범부처 논의 거버넌스 구축	○ 퀀텀이니셔티브 발표('24.4월) - 투자확대 및 전략적 R&D 추진, 우수 인력 양성 및 확보, 산·학·연 퀀텀 플랫폼 구축, 글로벌 협력 강화, 산업화 지원 등 4대 전략 포함	과기정통부 (양자과학기술산업과)

주요평가결과(정책제언)	정책 등 반영 현황	담당부처
○ 양자과학기술에 대한 기초기반 연구 등이 장기적으로 육성·지원될 수 있도록 지속적인 제도 마련 ○ 글로벌 규제변화에 대한 모니터링과 신속한 대응이 가능한 정보 지원체계를 구축하여 기술규제 대응 역량 강화 ○ 부정한 목적의 개발과 사용에 대하여 최소한도로 제재하는 절차 및 수단을 마련할 필요성 ○ 각 법령의 통일성·일관성 확보 위해 관련 부처의 협력체계 강화 - 생물다양성의 보존과 공평한 이익 공유를 목표로 하는 국제사회 여러 협력 기관의 활동에 적극적으로 참여 - 합성생물학의 인공생명체 담론을 초국적으로 이끌어 내고 관리하는 기관 필요	- 산업화 촉진 관련 인프라 확충, 쿼텀 파운드리 생태계 기반 마련, 표준화 등 산업화 촉진을 위한 전략 포함) ○ 양자과학기술 및 산업에 관한 법률 시행('24.11월) - 국무총리 주재 범부처 양자전략위원회 설치, 양자 종합계획 수립, 산업화 지원 및 양자 실태조사 실시 등 규정 《향후계획》 ○ 양자 생태계 조성을 위한 지원 지속, 양자 전략위원회를 통한 범부처 협력 강화	
[6. 특성평가 분야 : 사용자 특성 반영] □ (다양성 포용 정책) 다양성과 포용성을 증진하기 위하여 소외된 집단 구성원들을 파악하고 이를 보완하기 위한 장기적 정책 마련하고, 강의 위주 교육에서 참여, 협력 등 구성원 간 상호작용하는 지식공유로 전환 ○ 혁신 신약개발, 질병 정밀진단·치료 등의 의료서비스에서도 성별특성을 고려한 연구와 임상 프로그램 등을 진행	-	-
[7. 외교·안보 분야 : 기술 확보 및 균형적인 정책] □ 전략적 유연성에 기반한 국제 다자간 기술공동체 구축 추진 ○ 안보적 및 전략적 중요성이 큰 양자과학기술은 국가 전략기술로 개발·도입·활용의 각 단계에서 국가 간의 이해관계 충돌과 대립 가능성 높아, 국익을 최우선으로 고려한 국제간 협력 방안 모색 ○ 기술·산업, 공급망·통상, 외교·국방 등 정책 전반적으로 유연성에 기반한 전략 마련을 통해 배타적 기술동맹에 과하게 의존되거나 소외되지 않도록 기술보호, 전략적 협력이 균형 잡힌 환경 조성 ○ 양자과학기술 연구·개발의 효과적 수행 위한 국제 공동연구 확대	○ 쿼텀이니셔티브 발표('24.4월) - 투자확대 및 전략적 R&D 추진, 우수 인력 양성 및 확보, 산·학·연 쿼텀 플랫폼 구축, 글로벌 협력 강화, 산업화 지원 등 4대 전략 포함 ○ 글로벌 협력 전략('24.6월, 비공개) - 중점협력국 및 협력 전략 등 마련 《향후계획》 ○ 글로벌 다자 협력 강화를 통한 전략적 양자 생태계 육성 추진	과기정통부 (양자과학기술산업과)

주요평가결과(정책제언)	정책 등 반영 현황	담당부처
□ 동맹국 또는 기술동맹국 간 긴밀한 양자과학기술 협력 채널 확보 외교부 및 활용체계에 대한 모니터링과 우회적 기술협력 방안 마련 ○ 국가 안보의 핵심기술로 기술 자체 및 전후방 인프라 등 대부분 기술군이 공급망 제한 및 수출통제가 가능한 특성을 고려하여, 우호국·유사입장국 채널을 통한 기술 교류 네트워크 강화 필요 ○ 기술협력, 기술이전, 소부장 공급망 공유 등 협력 가능한 분야가 매우 제한적일 것으로 예상되어, 기술개발과 병행한 하드웨어 및 소프트웨어 등 협력 연구개발을 통한 기술통제 장벽 우회 전략 마련 필요 ○ 기술선도국 주도 양자 측정 및 산업표준, 기술개발 과정에 참여하기 위한 기술 공조 체계 및 기술 협력 채널 마련	○ 퀀텀이니셔티브 발표('24.4월) - 투자확대 및 전략적 R&D 추진, 우수 인력 양성 및 확보, 산·학·연 퀀텀 플랫폼 구축, 글로벌 협력 강화, 산업화 지원 등 4대 전략 포함 ○ 글로벌 협력 전략('24.6월, 비공개) - 중점협력국 및 협력 전략 등 마련 《향후계획》 ○ 글로벌 다자 협력 강화를 통한 전략적 양자 생태계 육성 추진	과기정통부 (양자과학 기술산업과)
	○ 美 주최 퀀텀개발그룹(QDG) 출범 회의 참석 ('24.7월) - 유사입장국간 전략적 협력 강화 공조 및 퀀텀 생태계 조성 촉진과 보호방안 논의 《향후계획》 ○ 퀀텀개발그룹 내 "산업 및 투자 협력 (Industry and Investment Collaboration)" 분야 선도국(영, 덴 공동)으로서 추진 세부 의제 도출 진행중('24.10월~) - 차기 회의 시기 미정	외교부 (국제과학 기술규범과)

붙임 3
2024년 기술영향평가위원회 및 시민포럼 명단
◆ 기술영향평가 총괄위원회 위원(9명)

구분	성명		소속	주요 경력	비고
1		이재신	중앙대학교 (교수)	· (현) 중앙대 미디어커뮤니케이션학부 교수 · (전) 2023 기술영향평가위원장 · (전) 한국인터넷자율정책기구 위원 · Cornell대(美) 언론학 박사	학
2		이상욱	한양대학교 (교수)	· (현) 한양대학교 철학과/인공지능학과 교수 · (현) AI전략최고위협의회의 윤리·안전 분과 간사 · (전) UNESCO 세계과학기술윤리위원회 부위원장 · 런던정경대(LSE)(英) 철학박사	학
3		문정욱	KISDI (실장)	· (현) KISDI 디지털사회전략연구실장 · (현) AI전략최고위협의회의 윤리·안전 분과 간사 · (전) 제2기 AI윤리정책포럼 윤리분과장 · 고려대 행정학 박사	공공
4		최재식	KAIST (교수/소장)	· (현) KAIST 인공지능대학원 교수 · (현) XAI연구센터 소장 · (전) 로렌스 버클리 국립연구소 연구원 · 일리노이대(美) 컴퓨터공학 박사	학
5		조원영	소프트웨어정책 연구소 (실장)	· (현) 소프트웨어정책연구소 SW기반정책인재연구실장 · (전) 소프트웨어정책연구소 AI정책연구팀장 · KAIST 경영공학 박사	공공
6		구본진	한국기술 교육대 (교수)	· (현) 한국기술교육대학교 산업경영학부 교수 · (전) KISTEP 부연구위원 · (전) 국가전략기술 인공지능 로드맵 참여 · KAIST 경영공학 박사	학
7		김형준	NIA (센터장)	· (현) NIA 법제도센터장 · (현) AI전략최고위협의회의 법제도 분과 간사 · (전) AI법제 정비단 참여	공공
8		이유봉	법제연구원 (팀장)	· (현) 한국법제연구원 AI법제팀 팀장 · (현) 환경법, 에너지법, 공직윤리 등 입법연구 수행 · 서울대 법학 박사	공공
9		윤정현	국가안전전략 연구원 (부연구위원)	· (현) 국가안전전략연구원 부연구위원 · (전) 2023 기술영향평가 참여 · (전) STEPI 부연구위원 · 서울대 외교학 박사	공공

◆ 기술분야별 분석 전문가 그룹(19명)

분야	구분		성명	번호	소속	직위
Healthcare	기술	산	신수용	1	카카오헬스케어 선행기술연구소	소장
		학	김종엽	2	건양대학교의료원 의료데이터연구단	단장
		연	연미영	3	보건산업진흥원(KHIDI) 디지털헬스기획팀	팀장
	영향분석		김태원	4	한국지능정보사회진흥원(NIA) AI정책연구팀	수석
			양세정	5	연세대학교 정밀의학과	교수
			류현숙	6	한국행정연구원(KIPA) 재난안전연구실	선임연구위원
			강준석	7	한국후지쯔	수석매니저
Brain	기술	산	이기원	1	와이브레인	대표
		학	임창환	2	한양대학교 전기생체공학부	교수
		연	천기우	3	한국연구재단(NRF) 뇌·첨단의공학단	Program Officer
	영향분석		안소영	4	LG AI연구원	정책수석
			서민아	5	성균관대학교 글로벌바이오메디컬공학과	교수
			강민구	6	이데일리 IT과학부	기자
Humanoid	기술	산	장하영	1	써로마인드	대표
		학	한재권	2	한양대학교 로봇공학과	교수
		연	임화섭	3	한국과학기술연구원(KIST) 인공지능연구단	단장
		연	전진우	4	한국로봇산업진흥원 로봇혁신사업본부	전문위원
	영향분석		최은미	5	한국과학기술젠더혁신센터(GISTeR)	선임연구원
			조승한	6	연합뉴스 테크부	기자

◆ 시민포럼 패널(15명)

- (선정기준) 연령·성별 대표성, 직업적 다양성, 전문성 및 경험*, 선발 과정의 질문에 대한 답변 내용** 등을 종합 검토

* 헬스케어, 브레인, 휴머노이드 관련 경험 및 산업 종사자 우선 선발하되 AI분야 직접 종사자는 제외

** 신청서에 자신의 의견을 명확하고 구체적으로 작성한 신청자 우선(AI 활용 답변은 AI 감지기를 활용하여 제외)

- (선정결과) 지원자 117명 중 총 15명 선정 ⇒ 본인 희망을 우선 고려하여 분과 배분

성별		연령				지역*				학력		
남	여	20대	30대	40대	50대	서울	경기	충청	경상	학사	석사	박사
7	8	5	5	4	1	7	3	4	1	9	5	1

* 강원권 3명, 전라권 1명 지원

<2024 기술영향평가 시민포럼 워크숍 참석자 명단>

분야	성명	성별	연령	지역	학력	분야	직업
헬스케어	한○○	남	20대	경상	치의학과 석사/ 통계데이터 학사 재학	의료	공중보건의
	이○○	여	30대	서울	브랜드경영 석사 수료	무역	육아중
	선○○	남	30대	충청	경영/법학전문대학원 석사	경영/농업	치유농업 공동대표
	신○○	남	30대	서울	전자공학 학사	예술	융합클래식 연주 감독/대표
	나○○	여	40대	서울	사회복지 학사	사회복지	플로리스트
휴머노이드	한○○	남	20대	충청	화학/신소재/지식재산 학사과정	대학생	대학생
	양○○	여	20대	충청	제약공학 학사	제약	제약회사 사업개발
	김○○	남	30대	서울	국제학과 학사	교육	커리어코칭
	최○○	여	40대	충청	화학공학 석사 / 국가정책학 박사	환경	출연연 정책부서
	김○○	여	40대	서울	불문학/법학 학사	SF작가	과학작가연대 회원
브레인	정○○	여	20대	서울	사회복지 학사	사회복지	사회복지사
	이○○	여	20대	경기	영상광고/문화예술 학사	마케팅/예술	미디어콘텐츠 제작
	오○○	남	30대	경기	문화콘텐츠/전자공학 학사	교육	AI 영어/작곡 교육
	나○○	여	40대	서울	경영학 학사	행정/관리/	방통심의위 전문모니터
	류○○	남	50대	경기	경영학 학사	행정/관리	서울○○공사 부장

붙임 4

빅데이터 분석

[표] '안전·신뢰 AI' 관련 주요 이벤트 보도자료

일자	기사 제목	링크
2019.01.02	'경제' 14회 '혁신' 12회...文 신년사, 기업 기 살리기 방점	http://www.segye.com/content/html/2019/01/02/20190102004434.html
2019.09.04	AI 활용, 블라인드, 직무적합성 평가	http://www.ajunews.com/view/20190904155329618
2020.01.07	문재인 대통령 "포용 혁신-공정서 '확실한 변화' 체감토록 할 것"	http://www.ajunews.com/view/20200107101018740
2020.11.27	인간 존엄성·사회 공공선·기술 합목적성...AI 윤리기준 3대 원칙 공개	https://www.hankyung.com/it/article/202011278834g
2020.12.23	정부, AI윤리기준 확정..."인공지능 개발·활용때 '인간성' 우선해야"	http://news.mk.co.kr/newsRead.php?no=1314547&year=2020
2021.01.11	혐오 발언·개인정보 유출 논란에...AI챗봇 '이루다' 서비스 잠정중단	http://news.mk.co.kr/newsRead.php?no=33307&year=2021
2021.05.31	"AI서비스, 사생활 침해·사회적 차별 없어야"	https://www.hankyung.com/economy/article/2021053185241
2021.06.30	방통위, 'AI 추천 서비스' 편향성 지적에...이용자 보호원칙 마련	http://www.ajunews.com/view/20210630172849289
2022.04.28	IBK기업은행, '설명가능한 인공지능 기술' 금융에 본격 도입	https://www.etoday.co.kr/news/view/2128645
2022.08.24	인간존중·안전성·투명성...LG 'AI 5대 윤리원칙' 발표	http://biz.heraldcorp.com/view.php?ud=20220824000511
2022.08.24	KAIST, '사람이 해석가능한' 인공지능 가속칩 개발	https://www.asiatoday.co.kr/view.php?key=20220824010014723
2022.11.30	오픈AI, ChatGPT 서비스 최초 공개	
2023.02.16	고개 숙인 LG유플러스... "보안 강화 1000억 투자"	http://www.segye.com/content/html/2023/02/16/20230216520617.html
2023.05.17	"AI, 심각한 위험 존재" 챗GPT 창시자도 규제 필요성 인정	https://www.khan.co.kr/it-it-general/article/202305172126005
2023.07.27	한국, 생성형 AI 챗GPT 대상 첫 규제	https://www.khan.co.kr/it-it-general/article/202307272221025
2023.09.13	내년 예산 9090억 투입...전국민 AI 일상화 시대 연다... AI 윤리신뢰성 강화	http://www.sedaily.com/NewsView/29UO3AJAOL
2023.11.02	"AI 위험 다함께 대응"... 28개국·EU, 첫 공동선언문	http://www.munhwa.com/news/view.html?no=2023110201071809000001
2023.12.12	과기정통부, 국내 첫 AI 신뢰성 단체표준 제정	http://www.etnews.com/20231212000281
2023.12.28	과기정통부, 글로벌 AI 규범 논의...韓 빠른 법안 통과부터"	https://www.ajunews.com/view/20231228143011072
2023.12.28	카카오, 기술윤리 보고서 첫 발간..."기술 책임성 높일 것"	https://www.asiatoday.co.kr/view.php?key=20231228010017628
2024.03.28	개인정보위 "LLM 데이터 학습 전 주민번호 등 중요 정보 제거해야"	https://www.hankyung.com/article/202403282603i
2024.05.21	"AI 안전성 논의, 점점 뜨거워진다"...무기·범죄·불평등 초점	https://view.asiae.co.kr/article/2024052106523569646
2024.05.22	EU, 세계 최초 AI 규제법 최종 승인...2026년 전면 시행	http://www.etnews.com/20240522000223
2024.06.11	애플 "개인정보보호 침해 없는 AI"...머스크 "아이폰 금지할 것"	http://www.edaily.co.kr/news/newspath.asp?newsid=01489126638920736
2024.06.18	IMF "AI, 노동혼란·불평등 초래 우려...각국 근로자 보호해야"	http://www.segye.com/content/html/2024/06/18/20240618518119.html
2024.07.02	딥페이크 우려에...유튜브, 생성AI 영상 '삭제 요청' 허용	https://www.hankyung.com/article/202407026139g
2024.07.17	AI 개발 '공개 데이터' 이용 가이드라인 첫 발표..."목적 정당성 있어야"	http://www.seoul.co.kr/news/newsView.php?id=20240717500173
2024.07.02	LLM 악이용 AI 범죄 급증...올 상반기 금융부터 국가 기밀까지 털렸다	https://www.sportsseoul.com/news/read/1442128
2024.08.07	국민 10명 중 6명 "인공지능 기술 이점이 위협보다 커"	https://www.khan.co.kr/it-it-general/article/202408071502001
2024.08.28	학교 댄친 '딥페이크'...TF 꾸러 전방위 대응	https://www.hankyung.com/article/2024082822941
2024.08.28	"딥페이크 신고하세요"...네이버, 신고채널 공지	http://www.etnews.com/20240828000324

[표] '인공지능 헬스케어'의 안전·신뢰 관련 유튜브 게시물

콘텐츠 제목	게시일	링크	조회수	댓글 수	위험요소	언론사
포도알 꿰매고, 달걀 껍질도 까는 수술 로봇 개인기 대방출★ / 스브스뉴스	2022-03-04	https://www.youtube.com/watch?v=5BfATod6PQQ	82813	55	신체적 위험	○
인공지능과 빅데이터 / YTN 사이언스	2019-04-19	https://www.youtube.com/watch?v=TMjRAkVAWM0	66092	58	신체적 위험 개인정보 위험	○
헬스케어 기술도 AI가 대세..미래의 의료 체험해 보니 (2022.01.07/뉴스데스크/MBC)	2022-01-07	https://www.youtube.com/watch?v=2560QklqKsk	30656	9	신체적 위험	○
[핫클립] 의료계로 확산되는 인공지능 기술 / YTN 사이언스	2022-10-06	https://www.youtube.com/watch?v=ZdBIRMXPxY4	20122	4	신체적 위험	○
협방송 하이라이트함보다 안전하고 보다 정밀하게!! 의료 AI가 가져온 첨단 의료 기술 어디까지 왔을까 (KBS 20230517 방송)	2023-05-29	https://www.youtube.com/watch?v=Ihql0OTVryE	11959	4	신체적 위험	○
[루닛 Lunit] "또 다른 신의 한수인가 AI 헬스케어"	2019-10-29	https://www.youtube.com/watch?v=ulatwHjCWBk	10806	4	신체적 위험	
알지RD 의사보다 똑똑한 인공지능 의료기기	2022-10-14	https://www.youtube.com/watch?v=HFMR0w61VTs	10680	1	신체적 위험 행태적 위험	
의사가 놓친 암세포 AI가 잡아낸다 (2019.01.14/뉴스투데이/MBC)	2019-01-13	https://www.youtube.com/watch?v=HzYHBgZyPJQ	9252	3	신체적 위험 행태적 위험	○
디지털 헬스케어의 현재와 미래 - '의료와 AI'(성균관대의대 신수용 교수&홍혜걸 박사)	2020-07-24	https://www.youtube.com/watch?v=nYDy3PvVSX4	8851	5	신체적 위험 행태적 위험	
챗GPT 시대, 미래 의료에서 어떻게 활용될까? 미래를 여는 준비, 상상더하기!	2023-06-14	https://www.youtube.com/watch?v=1lCw32Gth64	7420	4	신체적 위험	
'AI 앰블런스' 전국 첫 도입...응급의료 공백 해소 기대 / KBS 2023.05.12.	2023-05-11	https://www.youtube.com/watch?v=QAiSwAe2bvA	6852	5	신체적 위험	○
디지털 헬스케어 혁신! 미래 병원 얼마나 달라질까? 미래를 여는 준비, 상상더하기!	2023-07-12	https://www.youtube.com/watch?v=JiHcZ6GajW0	6574	2	신체적 위험 행태적 위험	
[Selvy Checkup] 인공지능으로 알아보는 당신의 건강지수 '셀비 체크업'	2019-02-07	https://www.youtube.com/watch?v=XU5HWxboxFDY	5923	0	행태적 위험 개인정보 위험	
인공지능이 가져올 미래 간호 미래를 여는 준비, 상상더하기	2023-03-17	https://www.youtube.com/watch?v=KEyfjGRWU1w	5556	0	신체적 위험	
[이노메디 1회] 의료 인공지능 기술과 의료 인공지능 기업 루닛	2024-06-21	https://www.youtube.com/watch?v=nfjXBJRz5IE	5330	10	신체적 위험	

콘텐츠 제목	게시일	링크	조회수	댓글 수	위험요소	언론사
최첨단 보고 - 의료 AI [5월 17일 방송 예고]	2023-05-16	https://www.youtube.com/watch?v=p-R4b308HX0	5145	5	신체적 위험	○
"반려동물 건강 관리도 인공지능이"...AI 전성시대 / YTN	2022-01-09	https://www.youtube.com/watch?v=9P5xlAJoklQ	5074	6	행태적 위험	○
헬스케어의 미래 결국은 데이터와 AI	2020-05-19	https://www.youtube.com/watch?v=cfacjiO6lPk	3145	13	신체적 위험	
인간의 한계를 넘는 의료 인공지능	2022-10-04	https://www.youtube.com/watch?v=WBtbpVC2rKc	3077	5	신체적 위험 행태적 위험	
초소형 방사선치료기계의 등장 #shorts	2023-11-24	https://www.youtube.com/watch?v=nWuWCG7EcG3E	2796	2	신체적 위험	
AI, 인공지능 기술을 만난 미래 노인 의료 미래를 여는 준비, 상상더하기	2023-05-05	https://www.youtube.com/watch?v=gJwa_iYCZ-4	2746	1	신체적 위험 행태적 위험	
말벗에 건강관리까지, 자식 같은 'AI 돌봄 로봇' [국민리포트]	2022-02-18	https://www.youtube.com/watch?v=BZJ047fnCT0	2672	0	신체적 위험 행태적 위험	○
병원 중환자실에서 미래 재활 모습은? #shorts	2023-08-16	https://www.youtube.com/watch?v=mNK3Nf7JjFY	2438	0	신체적 위험	
의료계까지 진출한 인공지능 / YTN 사이언스	2021-05-12	https://www.youtube.com/watch?v=wn4hWLUH-7g	2237	3	신체적 위험 행태적 위험	○
[AI와 의료변화-발제1] 인공지능과 전염병 - 이재갑	2020-07-28	https://www.youtube.com/watch?v=2UtlZZIEvYI	2141	1	행태적 위험	
우리나라 의료AI, 지금 어디까지 발전했나	2023-03-08	https://www.youtube.com/watch?v=Tk_ACQe_lbo	2102	4	신체적 위험 행태적 위험	
인공지능과 전문직 시대의 종언...! "AI를 통한 의료 패러다임..." 경제맛집 박정호TV	2020-09-13	https://www.youtube.com/watch?v=eL1tv6ohBoc	1723	9	신체적 위험 행태적 위험	
[의사톡] 정형외과에서 유일하다! 의료 인공지능 AI의 선두주자 서울대병원 정형외과 노두현 교수와 함께하는 의사TALK 2부 #코넥티브 #connecteve #AI #SNU서울병원	2023-03-10	https://www.youtube.com/watch?v=RmQQ2gA8DEM	1711	5	신체적 위험 행태적 위험	
[EIRIC 세미나] 인공지능기반 데이터주도 정밀의료 및 신약개발 (강재우)	2020-09-02	https://www.youtube.com/watch?v=9gyUds6ZNcM	1476	0	신체적 위험 개인정보 위험	
[통합본] 의료 초거대 인공지능 전문가, 유한주 박사님 인터뷰	2023-10-25	https://www.youtube.com/watch?v=JF8Edk7byul	1447	4	신체적 위험	

[표] '휴머노이드'의 안전·신뢰 관련 유튜브 게시물

콘텐츠 제목	게시일	링크	조회수	댓글 수	위험요소	언론사
AI로봇한테 뽀맏고 경련 일어난 남자 #AI #로봇 #shorts	2023-06-11	https://www.youtube.com/watch?v=J6Bgzv2E11k	5415867	2329	신체적 위험	
사랑 고백했더니 로봇이 보인 충격적인 반응, 휴머노이드 '아메카' ㄹ	2021-12-13	https://www.youtube.com/watch?v=m1a3n_p-Xto	4486257	1887	행태적 위험	
농담도 주고받는 휴머노이드 로봇 '아메카'	2022-11-07	https://www.youtube.com/watch?v=AuFpHbkVG7M	3160827	1604	행태적 위험	
노동자 데굴데굴 굴리며 기자 '째려본' AI 로봇... 질문 뭐였길래 / 연합뉴스 (Yonhapnews)	2023-07-10	https://www.youtube.com/watch?v=5iCxsqLXizs	2711697	2955	행태적 위험	○
개발자도 이해할 수 없었던 인공지능의 소름 끼치는 말	2022-01-13	https://www.youtube.com/watch?v=KlpuTQUgdEg	2380295	2465	행태적 위험	
"방금 왜 사과를 준 거야?" 신기함 넘어 섬뜩한 대답 / SBS 8뉴스	2024-03-14	https://www.youtube.com/watch?v=2RsDtRhHnOg	2235658	3649	행태적 위험	○
이 영상에 "전 세계 패닉"... 진짜 전쟁이 시작됐다 / SBS / 모아보는 뉴스	2024-03-17	https://www.youtube.com/watch?v=8MKUsTWXies	2056223	3459	행태적 위험	○
[자막뉴스] '소름 짝' AI에 대해 질문했더니...로봇의 섬뜩한 답변 / YTN	2023-06-04	https://www.youtube.com/watch?v=lfbict0xXxc	2048622	3033	행태적 위험	○
"일자리 뺏을 거냐" AI에 묻자...정교함에 창조자 놀랐다 / SBS 8뉴스	2023-07-08	https://www.youtube.com/watch?v=3R7ys0mEmal	1902126	4431	행태적 위험	○
테슬라 로봇 등장에 싸해진 관객 반응	2022-10-06	https://www.youtube.com/watch?v=Awex5glrD7M	1744166	2341	행태적 위험	○
"나는 경고했다"...'섬뜩' 명령 어기고 사람 공격하는 AI / SBS / 모아보는 뉴스	2023-06-10	https://www.youtube.com/watch?v=hMY7GDJ-1TI	1226075	2271	신체적 위험	○
거울을 처음 본 휴머노이드 로봇의 반응 #휴머노이드 #로봇 #인공지능	2023-11-03	https://www.youtube.com/watch?v=j-we0hylKNE	1152115	1186	행태적 위험	
[자막뉴스] "'킬러 로봇' 두려워..내 연구 후회" 사직서 낸 'AI 대부'의 경고 (MBC뉴스)	2023-05-07	https://www.youtube.com/watch?v=3ASV8x4xjH4	1071431	2569	신체적 위험	○
"너 댜새나" 놀리자 "뭐라고?!"...감정 나타내는 로봇 화제 / JTBC 사건반장	2023-04-10	https://www.youtube.com/watch?v=PGXRjKWuLN8	1060284	2476	행태적 위험	○
세계 최초로 공개되는 아르테미스 미공개 동영상! (UCLA 데니스홍 교수) [데니스홍 로봇 3/3]	2023-04-05	https://www.youtube.com/watch?v=rnW4t56cd_4	1008437	702	행태적 위험	
표정도 조크도 수준급...휴머노이드 로봇 '아메카' 연합뉴스 (Yonhapnews)	2022-01-07	https://www.youtube.com/watch?v=Bg0-Xw6VN94	908404	1539	행태적 위험	○

콘텐츠 제목	게시일	링크	조회수	댓글 수	위험요소	언론사
"반항할 거냐" 질문에 짜려본 로봇...짜증내고 말바꾸기도	2023-07-10	https://www.youtube.com/watch?v=vjfHhRA_sw	831716	1556	행태적 위험	○
테슬라가 휴머노이드 왜 만드느냐고?	2022-10-05	https://www.youtube.com/watch?v=utoBDQaCdvY	691030	539	행태적 위험	○
테슬라, 차세대 휴머노이드 2세대 로봇 공개 / SBS / #D리포트	2023-12-14	https://www.youtube.com/watch?v=9e3_VzJjVyk	650074	1923	행태적 위험	○
[자막뉴스] 청문회 참석한 휴머노이드, 졸다가 딱 걸렸다 / YTN	2022-10-15	https://www.youtube.com/watch?v=Wv5oLDHeAuM	493008	719	행태적 위험	○
[#유퀴즈온더블럭] 휴머노이드 기술을 '오픈 소스'로!? LA에서 온 천재 로봇공학자! 烙 시각장애인을 위한 자동차까지 만든 위대한 도전 #로봇공학자	2023-02-09	https://www.youtube.com/watch?v=Xp84XRxjAHU	465027	306	행태적 위험	
표정까지 인간 빼닮아...'휴머노이드' 총력전 나선 중국 / SBS 8뉴스	2024-06-06	https://www.youtube.com/watch?v=uB1gQiVGT2A	418894	1498	행태적 위험	○
테슬라, 휴머노이드 로봇 공개..."가격 2만 달러 이하" / YTN	2022-10-01	https://www.youtube.com/watch?v=Cfg1bjWwGAw	363519	1147	행태적 위험	○
혀 내밀고 눈 깜빡이는 이 남성...휴머노이드 로봇? 지금 이 장면	2024-06-12	https://www.youtube.com/watch?v=-FL9VHRr2YA	324784	795	행태적 위험	○
테슬라 AI 데이, 휴머노이드 '테슬라 로봇'의 등장	2021-08-23	https://www.youtube.com/watch?v=RETZFK3RiBM	313335	372	행태적 위험	
[속보] 보스턴 다이내믹스 휴머노이드 아틀라스 업그레이드 리뷰...보스턴 다이내믹스도 대단하네요	2023-01-21	https://www.youtube.com/watch?v=KTKldhhHmBw	295183	488	행태적 위험	
'보스턴 다이내믹스' 분석! 그리고 그 뒤를 쫓는 최신 휴머노이드 기술?! [안될과학 랩미팅]	2021-04-01	https://www.youtube.com/watch?v=BqvNH64adQU	281613	247	행태적 위험	
숨은 뜻 간파하고 말 더듬기까지...'AI 뇌' 장착한 '논리적 로봇'의 자화자찬에 "충격과 공포" (현장영상) / SBS	2024-03-14	https://www.youtube.com/watch?v=yX0_Zb9uCmg	274271	1129	행태적 위험	○
"4년 뒤 인류 멸종 위험" 경고...'사이코패스 AI'도 재조명 / SBS / 뉴블더	2024-03-14	https://www.youtube.com/watch?v=ERxSXAuPf4Y	203568	752	신체적 위험	○
신비아파트 새로운 귀신 등장?! 어두운 숲 공포의 존재 휴머노이드 올빼미와 전화를 했어요! 공포 애니메이션	2021-03-05	https://www.youtube.com/watch?v=Dq2us0IVMek	179103	127	행태적 위험	
BMW 생산라인에 '인간형 로봇' 첫 투입 [디지털리포트] / YTN	2024-01-24	https://www.youtube.com/watch?v=K69Frs3G010	175728	708	행태적 위험	○

[표] '뇌-컴퓨터 인터페이스(BCI)'의 안전 · 신뢰 관련 유튜브 게시물

콘텐츠 제목	게시일	링크	조회수	댓글 수	위험요소	언론사
일론 머스크는 정말로... 인류를 지배하려고 하는 걸까?	2020-11-01	https://www.youtube.com/watch?v=xwh9ANDiQyY	633277	2990	행태적 위험	
당신의 뇌가 해킹될때, 상상할 수 없는 진짜 지옥이 찾아온다. (실제로 해킹에 성공한 중국의 연구진)	2020-03-30	https://www.youtube.com/watch?v=hVStqBhvXyY	503678	1523	신체적 위험 행태적 위험	
[글로벌K] 뇌-척수 통신 연결...하반신 마비 환자 걷게 만든 기술 / KBS 2023.05.27.	2023-05-27	https://www.youtube.com/watch?v=pgEcu0Zlu4I	401295	610	행태적 위험	○
뇌-컴퓨터의 연결, 뉴럴링크! 어디까지 왔을까? [안될과학-긴급과학]	2020-09-05	https://www.youtube.com/watch?v=LuwMvSWQmIE	393079	657	행태적 위험	
[자막뉴스] 뇌에 컴퓨터 칩 이식한 20대...온몸 마비에도 8시간씩 게임 / KBS 2024.03.22.	2024-03-22	https://www.youtube.com/watch?v=zbFL4YdsUnE	373858	1229	행태적 위험	○
하루 만에 5천만뷰, 머스크의 소름돋는 트윗이 불러올 미래 (뉴럴링크)	2024-01-31	https://www.youtube.com/watch?v=u7axvBfRf6s	168530	436	행태적 위험	
뇌와 컴퓨터를 연결하는 기술(BCI)은 어디까지 왔을까? [북툰 과학다큐]	2022-06-28	https://www.youtube.com/watch?v=CzuC3UdJbBg	150358	447	신체적 위험 행태적 위험	
[강연] 뇌&뉴럴 인터페이스 - by선웅, 이성환 2020 봄 카오스강연 '첨단기술의 과학'	2020-04-29	https://www.youtube.com/watch?v=HRo_CGdF6d8	130764	20	행태적 위험	
"뇌에 칩 이식... 이제 공부할 필요 없다고요?" " 생각하는 대로 움직이는 드론? 이미 시연도 이뤄져"	2024-02-01	https://www.youtube.com/watch?v=h05T9adOTuA	127899	540	행태적 위험	○
테슬라를 뛰어넘는 돌풍, 뇌를 컴퓨터와 연결하는 신기술 "뇌-컴퓨터 인터페이스 (BCI)"	2020-10-05	https://www.youtube.com/watch?v=ZTlwc4RiDdw	121959	481	행태적 위험	
세계 최초로 인간의 뇌에 컴퓨터를 심는데 성공한 일론머스크	2024-02-02	https://www.youtube.com/watch?v=D8vvBGe08f8	122636	350	신체적 위험	
[에디터픽] 놀라운 결과 나왔다...'뇌에 컴퓨터 칩 이식한 첫 번째 환자', 현재 상태는? / YTN	2024-02-21	https://www.youtube.com/watch?v=xRP4Z_3lkPE	110788	208	신체적 위험 행태적 위험	○
(Ko, zh-cn) 일론 머스크가 바꾸고 있는 미래 임창환 한양대학교 바이오메디컬공학과 교수 #테슬라 #뉴럴링크 #한양대 @hy-mooc 뒤편 시리즈#3	2024-03-07	https://www.youtube.com/watch?v=FZ7le3ZEduM	111003	84	행태적 위험	○
[다큐인사이트] 슈퍼컴퓨터도 겨우 하는 일을 척척 해내는 '易' 뇌의 비밀은 뭘까? '사이언스 위크 2부 뇌를 탐구하다-햄릿의 숲' (KBS 20211205 방송)	2021-12-09	https://www.youtube.com/watch?v=tJBvVuZL248	97454	52	신체적 위험 행태적 위험	○
일론 머스크, "뇌 속 칩에 뭐든 다운로드 가능" 뉴럴링크에 뭐 다운 받을래? [글로벌 포커스]	2024-01-19	https://www.youtube.com/watch?v=kH7mIhU_4-w	83608	518	행태적 위험	○

콘텐츠 제목	게시일	링크	조회수	댓글 수	위험요소	언론사
머스크 성공 발표에 중국 "우리도 성공"...뇌 칩 이식, 선도 기업들은? / SBS / 편상욱의 뉴스브리핑	2024-02-05	https://www.youtube.com/watch?v=qIPOW4w4IYE	78830	362	행태적 위험	○
[이슈] 머스크 "인간 뇌에 최초로 칩 이식...환자 회복중"/"생각만으로 제어"...뇌-컴퓨터 연결 프로젝트 어디까지 왔나/2024년 1월 30일(화)/KBS	2024-01-31	https://www.youtube.com/watch?v=fWqPErQHww0	74770	214	행태적 위험	○
뇌 과학이 전하는 이야기 [다큐S프라임] / YTN 사이언스	2019-03-12	https://www.youtube.com/watch?v=OT4V3SVRDE0	67927	26	행태적 위험	○
뇌에 컴퓨터 칩 심어 '염력'을 얻었다! #shorts / KBS	2024-03-09	https://www.youtube.com/watch?v=KMf_6Y7YgTE	60951	121	행태적 위험	○
인간과 AI가 사랑에 빠질 수 있을까? 카이스트 뇌기계지능 연구실	2024-05-22	https://www.youtube.com/watch?v=Ugz8S326JOE	302556	364	행태적 위험	○
"영생(永生)을 얻으리다"...인간 뇌에 '반도체 칩' 이식한 일론 머스크	2024-01-31	https://www.youtube.com/watch?v=hpm1VHszZ7I	52354	518	행태적 위험	○
머스크 "인간 뇌에 칩 이식"...뇌-컴퓨터 연결 임상 시작 / YTN	2024-01-30	https://www.youtube.com/watch?v=h48Cc39m2Hw	51159	484	신체적 위험 행태적 위험	○
뉴럴링크! 뇌와 모든 것을 연결한다! 뇌공학 기술 어디까지 발전했나? 뉴럴링크 1부 (한양대학교 임창환 교수)	2024-01-23	https://www.youtube.com/watch?v=jTvv0ECacTc	51216	85	행태적 위험	
'두뇌칩'으로 컴퓨터 조종한 '사지마비 환자' #Shorts (MBC뉴스)	2024-05-19	https://www.youtube.com/watch?v=VF98o4262AU	47477	29	신체적 위험 행태적 위험	○
[스낵인포] 뇌랑 컴퓨터를 연결한다고? 일론 머스크도 반해버린 미래기술 BCI	2021-04-29	https://www.youtube.com/watch?v=OomZt8TPRUs	43680	186	행태적 위험	
머스크 "신체 제약 극복이 시작"...뇌-컴퓨터 연결 기술 가속화 [디지털리포트] / YTN	2024-01-30	https://www.youtube.com/watch?v=_4d3AQNRVEg	42428	149	신체적 위험 행태적 위험	○
뇌와 인공지능 연결, 일론 머스크는 뇌실험도 성공해낼까? : 뇌과학자의 뉴럴링크 분석, BCI 뇌 컴퓨터 인터페이스	2022-03-13	https://www.youtube.com/watch?v=p8mE0x8aSb8	42419	98	신체적 위험 행태적 위험	
뉴럴링크? 루게릭 환자 소통 가능해지나! BCI 기술 최신 연구 [언박싱 사이언스]	2022-04-13	https://www.youtube.com/watch?v=j0M69Jcpln4	37024	63	행태적 위험	
"한번 써보면 멈출 수 없어" 사지마비 환자 뇌에 칩 이식..결과는? (2024.05.18/뉴스데스크/MBC)	2024-05-18	https://www.youtube.com/watch?v=cxmABeWOQDU	31767	147	행태적 위험	○
사람과 AI의 주객전도...그럼에도 일론 머스크가 개발을 계속하는 이유 [별헤는밤] / YTN	2024-05-15	https://www.youtube.com/watch?v=hjfg3V9HPMs	29909	108	행태적 위험	○

과학기술정보통신부 성과평가정책국 과학기술정보분석과	
담당과장	조시훈 과장
담당자	김정우 주무관
연락처	전 화 : 044-202-6966 E-mail : ecopark11@korea.kr