

---

# 초거대AI 경쟁력 강화 방안

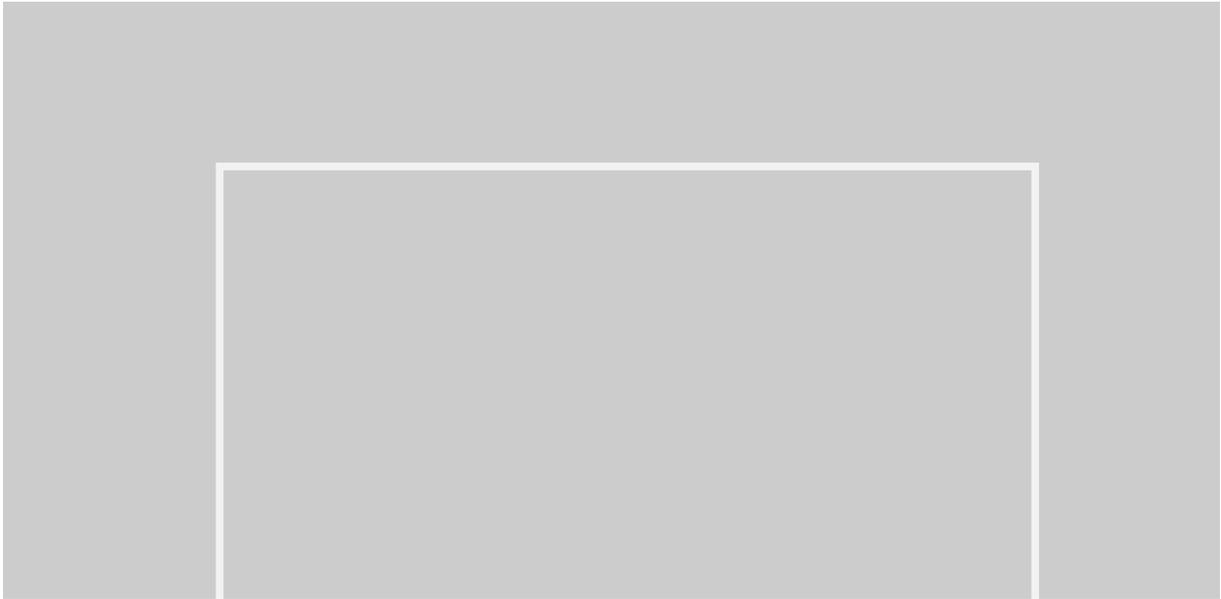
---

2023. 4. 14.

관계부처 합동

# 순 서

[요 약]	i
I. 추진 배경	1
II. 초거대 AI 시대 개막	4
III. 초거대 AI의 영향력	7
IV. 우리의 경쟁력 진단	10
V. 비전 및 중점 추진과제	17
VI. 전략 및 추진과제	18
1. 초거대 AI 기술·산업 인프라 확충	18
2. 초거대 AI 시대를 견인하는 혁신 생태계 조성	24
3. 범국가 AI 혁신 제도·문화 정착	30
VII. 향후계획	33



# 요약



# 초거대AI 경쟁력 강화 방안(안) [요약]

## I 추진 배경

- 전 세계적 챗GPT 돌풍을 계기로, AI에 대한 막연한 기대감이 높은 **효용성**으로 증명되고, 누구나 쉽게 AI를 활용하는 '**AI 일상화**' 촉발
- 글로벌 빅테크 기업은 압도적 컴퓨팅 파워·대규모 자본 등을 토대로 **초거대AI 플랫폼을 선점**하기 위한 **속도전\*** 치열
  - \* OpenAI(챗GPT, '22.11) → 구글(Bard, '23.2), 메타(LLAMA, '23.2) → OpenAI (GPT4, '23.3)
- 우리도 독자적 초거대AI 플랫폼을 바탕으로 세계 선두권 경쟁에 참여
  - ⇒ 관련기업 지원을 강화하여 **초거대AI 경쟁력 확보 및 미래 수출 전략산업화**
  - \* 우리나라는 자체 초거대 AI를 보유한 4개국(미국, 중국, 이스라엘, 한국)에 포함

### < 초거대AI 경쟁력 강화 방안 수립경과 >

「AI 석학과의 대화」('22.9) → 「대한민국 디지털 전략」('22.9) → 「AI 일상화 및 산업 고도화 계획」 발표('23.1) → 「AI 최고위 전략대화」 등 초거대AI 대응 민간 릴레이 간담회('23.1~3)

## II 초거대AI 영향력

✓ 초거대 AI는 사람 수준의 언어·시각 능력을 토대로 **디지털 산업 자체의 파괴적 혁신**과 경제·사회 **전 분야의 AI 내재화**를 주도할 전망

※ (스탠포드 HAI) 생성AI 기술이 **변곡점**에 도달, 사회적 영향력이 전방위적으로 확대

- **(기술)** 트랜스포머, 강화학습 등으로 AI 성능이 획기적으로 개선되고, 검색·워드프로세서 등에 AI가 결합, 압도적 **효율 향상**의 도구\*로 진화
  - \* (MS) 워드·엑셀·PPT 등 오피스 SW에 '코파일럿' 적용('23.3) → 전문비서 수준으로 변경
- **(산업)** 초거대AI가 사람의 지식서비스를 보조하고, 산업 난제해결을 지원하는 등 **소산업의 생산성을 획기적으로 혁신**하는 필수요소로 부상

- 초거대AI에 전문지식 데이터를 추가 학습(파인튜닝)하여, 다양한 전문분야의 AI서비스 개발이 가능



- (사회) 초거대 AI 활용이 본격화되며, ①신뢰성·윤리 이슈와 ②AI수용성 차원에서 AI 위험요인 해소, 리터러시 강화, 기존 제도의 개선 준비 필요

- \* AI 산출물에 대한 저작권 논쟁, 데이터 학습 과정에서 개인정보 문제 등
- \* 급속한 AI발전에 대한 우려시각도 일부 존재 → 美 비영리단체 '생명의 미래연구소'는 인류에 대한 위험성을 이유로 GPT4 이상의 강력한 AI 개발을 일시 중단할 것 촉구

### III 우리의 경쟁력 진단

- 그간 데이터 축적(학습용데이터 691종), AI 기술력 향상(美 대비 89.1%), 국산 AI반도체(NPU, PIM) 출시 등 인공지능 기반 조성에 집중

- 이와 같은 정책성과와 민간의 한발 앞선 도전으로, 우리 기업은 독자적 초거대AI 플랫폼\*을 확보하고 세계를 향해 도전 시작

- \* 네이버(클로버), LG(엑사원), 카카오(KoGPT), SKT(에이닷), KT(믿음) 등 출시

- 응용기업들은 초거대 AI 플랫폼에서 제공하는 API를 바탕으로 기업별 데이터를 추가학습 시킨 '전문화된 AI'로 승부하는 추세

- \* (뤼튼) 작문·콘텐츠창작 플랫폼 출시, (업스테이지) 사진 답변 등 AI챗봇 'AskUp' 출시

⇒ 그동안 축적한 AI 산업기반과 국내 초거대AI 플랫폼을 기반으로 글로벌 시장에서 승부해 성과를 창출할 가능성이 충분

### IV 비전 및 추진전략

비전 · 목표: **똑똑한 인공지능, 국민과 함께 디지털 경제를 가속화**하겠습니다.

**초거대AI 플랫폼** 한국어플랫폼 세계1위  
+ 비영어권 중심 글로벌시장 선점

**응용서비스** 기업 간 협력 생태계 조성으로 전문 특화분야 세계 1위 도전

- 추진 전략
1. 초거대AI 개발·고도화를 지원하는 기술·산업 인프라 확충
  2. 민간·공공 초거대 AI 융합 등 초거대 AI 혁신 생태계 조성
  3. 범국가 AI 혁신 제도·문화 정착

## V 주요 내용

✓ **그간 확보한 AI 정책 성과** 위에 **초거대AI 경쟁력 강화** 및 산업 생태계 육성을 위해 **추가하거나 개선해야 할 과제**를 발굴하여 집중 추진

### 전략1. 초거대 AI 기술·산업 인프라 확충

□ **(데이터)** 초거대AI 개발에 필요한 **양질의 텍스트 데이터 보강**(‘23~) **신규**

○ 분야별 특화 **텍스트 데이터 200종**(도서 15만권 분량) 구축(‘23년 34종)

\* 한국어 중심으로 텍스트 데이터를 확보하되, 비영어권(동남아, 중동 등) 데이터 단계적 확대

○ **한국어 고품질 말뭉치\***, **응용말뭉치\*\*** 구축(‘23년 25종 구축, 문체부) **강화**

\* 한국어 분석정보(개체명, 감성 등) 부착 말뭉치 \*\* 한국어-외국어·수어·점자 통번역 등

현재	+	개선
AI 학습용 데이터 사업은 <b>라벨링 데이터</b> 구축 중심		초거대AI 사전학습에 필요한 <b>텍스트(비라벨링) 데이터</b> 구축 신설

□ **(기술)** 기 추진 중인 딥러닝 개선 기술개발(‘22~’26, 2655억원)과 함께, 현재 **초거대AI의 한계 돌파**를 위한 **R&D 신규 추진**(‘24~) **신규**

현재	+	개선
AI의 <b>학습능력, 소통능력</b> 등을 강화하는 기초연구 진행		초거대AI에서 부각된 <b>최신정보 미반영, 거짓 답변</b> 등 한계 돌파 기술개발

□ **(인프라)** 국산 AI반도체 기반의 **고성능·저전력 K-클라우드**를 초거대 AI가 **활용**할 수 있도록 **핵심 HW·SW** 개발(‘23~) **신규**

○ 중소기업·대학 등에 **초거대AI용 컴퓨팅자원 확대 지원**(‘23~) **강화**

### 전략2. 초거대 AI 혁신 생태계 조성

□ **(초거대AI 플래그십 프로젝트)** 민간 5대 전문영역에 초거대AI를 접목하여 **전문가 보조·지원** 등 생산성을 혁신하는 응용서비스 개발(‘24~) **신규**

< 민간 5대 전문영역 초거대AI 플래그십 프로젝트(예시) >

법률	의료	심리상담	문화·예술	학술·연구
소장 작성 지원 판례·판결문 제시	소견서 작성 지원 치료방법 추천	육아고민 상담 심리분석 리포트	한류 대응 외국인 한국어 교육	참고문헌 통계 제시 초록 작성 보조

- **(공공·행정 혁신)** 내부업무·민원대응 효율화(과기정통·행안부), 초거대 AI 특허검색(특허청) 등 초거대 AI 공공 선도과제 추진('23~) **강화**
- **(초거대AI 산업 혁신 생태계)** 민간 차원의 투자, 新서비스 창출 등 디지털 기업의 협력 강화를 위한 “초거대AI 협의회” 운영('23~) **신규**
- 중소기업의 초거대AI 기반 클라우드 서비스 개발을 종합지원하고('23~) **신규**, 메타버스에서 초거대AI를 지능형 비서로 시각화하는 프로젝트 추진('24~) **신규**
  - \* (지능형교사) 교육·학습 지원, (개인비서) 스케줄관리·예약, (쇼핑도우미) 상품추천·결제·배송
- **(인재양성)** 기존의 AI·SW 인력 양성과 함께, 초거대AI 개발·활용에 전문화된 글로벌 수준의 인재를 추가 양성('24~'27년) **신규**
- MS 등과 **구직자 대상 초거대AI 프로젝트** 중심 교육과정 운영('23~, 고용부) **강화**
- **SW 개발자, 교원·학생** 등 대상 초거대AI **활용역량** 강화교육, **일반인** 대상 **기본소양·윤리교육** 실시('23~'27년, 교육·과기정통부) **강화**

현재	+	개선
수준별 SW·AI 인재양성('23~'27) AI 석·박사 등 정규과정(6.5만명) 구직자·재직자 등 비정규과정(13.2만명)	+	초거대AI 개발·활용 역량 강화('23~'27) 초거대AI 전문인재 양성 및 초거대AI 활용역량 교육 보강, 일반국민 AI 역량 교육(100만명)

### 전략3. 범국가 AI 혁신 제도·문화 정착

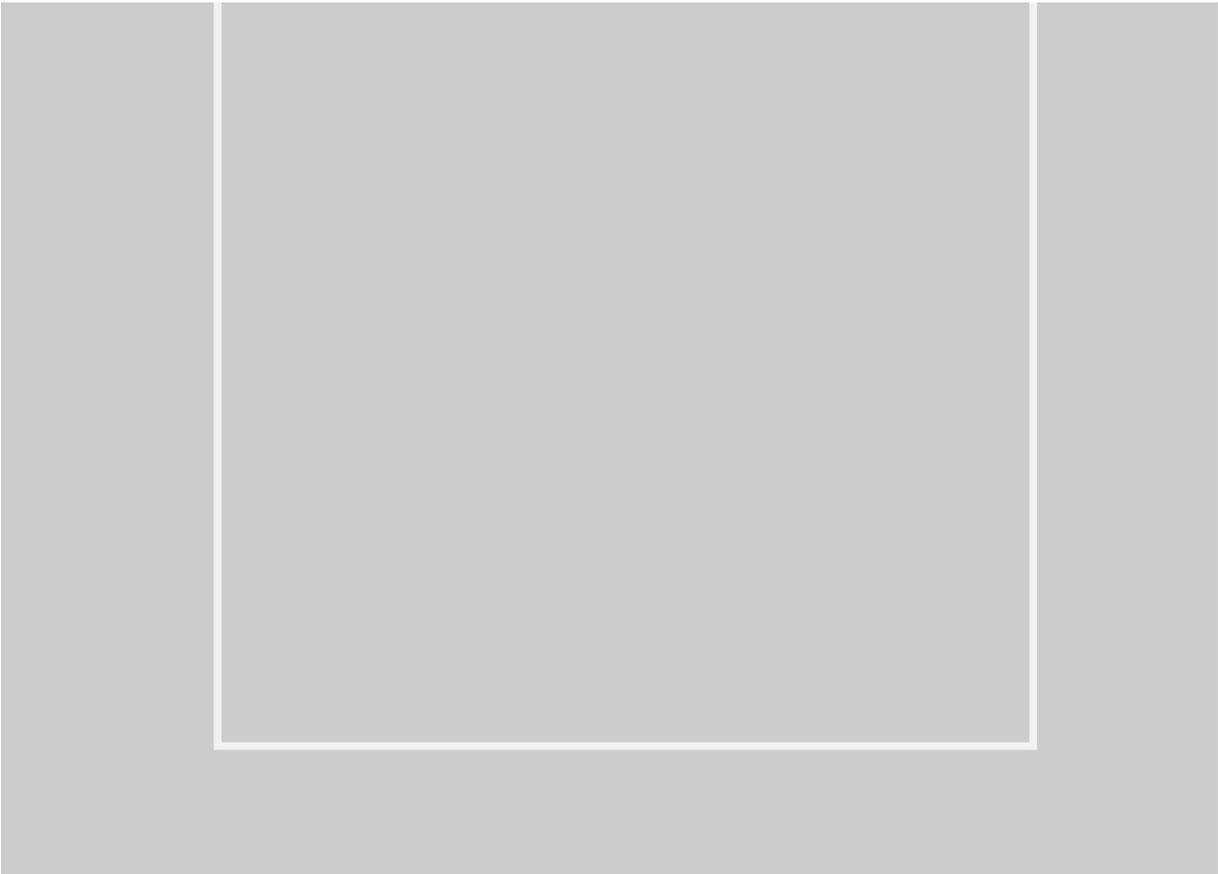
- **(사회적 수용력)** 초거대AI 관련 **규제 개선방향** 도출, 초거대AI 확산에 따라 야기되는 교육·보안 등 **사회적 이슈** 논의·대응방안 모색('23~) **강화**
  - \* 학업을 위한 챗GPT 허용 여부, 챗GPT 활용과정에서 보안사항·개인정보 유출 위험성
- 초거대AI 활용 등을 검토하여 '**교육 분야 AI 윤리원칙**' 보완('23, 교육부), 공무원 대상 **초거대AI 공공부문 활용 가이드라인** 마련('23, 행안부)
- **(신뢰성 평가)** 기업이 개발한 초거대AI 서비스에 대해 위험요인·성능을 공신력 있는 **제3의 기관을 통해 평가** 추진('23~, 민간 자율신청 기반 운영) **신규**

## VI 향후 계획

- 사회적 변화 대응 및 수용력 향상을 위한 추진체계 구성 : '23.4월~
- 신규 프로젝트 세부기획 및 예산반영 : '23.4월~



**PH PH**



## I 추진배경

### □ 챗GPT 돌풍, 초거대 AI가 우리의 일상 속 스며드는 신호탄

- 챗GPT 출시('22.11)로 촉발된 초거대 AI 열풍, 기업·전문가 영역에서 활용되던 AI가 일상 모든 접점에서 본격 확장되는 계기 마련

- (생성형AI) 단어 사이의 관계와 맥락을 학습하여 문장에 위치할 단어를 생성(스탠포드HAI)
- (초거대AI) 국내에서는 생성형AI 중 대규모 모델 크기를 갖는 것을 초거대AI로 통용

- 초거대 AI는 AI의 혁신적 가치를 전산업에 내재화시키는 게임체인저로서, 경제·사회 전반의 파괴적 혁신\*을 이끌 것으로 기대

\* "챗GPT같은 AI는 PC·인터넷만큼 중대한 발명으로 세상에 영향을 줄 것" (빌게이츠, '23.2)

### □ 초거대 AI 시장 선점을 위한 빅테크 주도의 경쟁이 본격화

- 빅테크 기업은 압도적 컴퓨팅 파워·대규모 자본 등을 토대로 초거대 AI 개발·플랫폼 선점에 사활을 걸고 경쟁 중

\* OpenAI(챗GPT, '22.11월) 출시 이후 구글(Bard, '23.2월), 메타(LLAMA, '23.2월), OpenAI(GPT4, '23.3월), MS(코파일럿, '23.3월) 등이 경쟁적으로 초거대 AI 출시

- 각 분야의 서비스 기업 역시 초거대 AI 기반 新비즈니스 창출에 매진, 초거대 AI 전·후방 생태계 혁신이 본격화

### □ 우리의 강점과 축적된 AI 역량을 토대로 글로벌 시장에 도전할 필요

- 그간 민간의 한발 앞선 도전과 정부의 적극적 지원을 토대로, 우리는 초거대 AI 분야에서 세계와 경쟁 가능한 수준\*까지 성장

\* 국내 AI 플랫폼·알고리즘 등 기술역량(Development)은 세계 3위 수준(英 토터스미디어, '22)

- 우리도 초거대 AI 경쟁력 확보는 물론 각 분야별 초거대 AI 기반 혁신서비스를 선도적으로 창출, 미래 수출 전략산업으로 육성 필요

⇒ 챗GPT 열풍을 기회로, 초거대 AI 경쟁력 조기 확보에 민·관의 역량을 결집하여, 대한민국의 새로운 성장동력의 탄생을 이끌 시점

## 참고 추진경과

### □ 뉴욕 구상('22.9.21) 및 AI 석학과 대화('22.9.22)

- '뉴욕 구상'을 통해 인류 보편적 가치를 확산하고 디지털 혜택을 공유하기 위한 새로운 디지털 질서 수립 제안



"디지털은 그 자체가 기술인 동시에 혁명입니다. 인류의 보편적 가치를 지키기 위해서는 새로운 차원의 디지털 질서가 필요합니다."(MP, '22.9, 뉴욕대)

- 윤석열 대통령은 제프리 힌튼 등 'AI 석학과 대화' 모두발언에서 인공지능의 공정성·공평성 등에 대한 본질적인 논의와 함께,  
- 다양성이 인정되고 관용이 높은 사회적 분위기, 결과에 연연하지 않는 꾸준한 정부의 지원의 중요성을 강조

### □ 인공지능 일상화 및 산업 고도화 계획('23.1.26)

- 아직 기업현장·국민생활 등에서 인공지능 활용은 초기 단계이므로, 국민과 인공지능 혜택을 공유하고 AI 산업·기술 초격차 실현 필요  
- 그간 축적한 산업기반을 토대로 인공지능을 국가 전반으로 확산하고 AI산업의 실질적 성과를 창출하기 위한 10대 프로젝트 추진('23년 7,129억원)

### □ 초거대AI 기업 릴레이간담회('22.12월~'23.3월)

- 초거대AI 기업 간담회('22.12~'23.3, 네이버, 카카오, LG AI 연구원, KT, SKT 등)
- 챗GPT 대응 장관-전문가 토론회('23.2), 디지털 국정과제 현장 간담회('23.2)
- 제3차 '인공지능 최고위 전략대화' 개최('23.3)

※ 네이버, LG AI, SKT, KT, 카카오 등 초거대 AI 개발 기업 / 루닛, 뽀빠이, 스캐터랩 등 AI 스타트업 / 서울대·KAIST 등 학계 등 다양한 의견 수렴 추진

**【발제 주요내용】**

 배경훈 원장	• 인공지능은 <b>승자독식</b> 특성이 크고 <b>기술발전 속도가 매우 빠른</b> 분야로 글로벌 경쟁을 위해서 차세대 원천기술 연구, 양질의 데이터 확보 등 <b>민관의 역량 결집이 중요</b>
 하정우 소장	• 우리나라는 <b>초거대AI</b> 를 개발하는 <b>4개국</b> (한국·미국·중국·이스라엘) 중 하나로, 기술격차는 크지 않은 상황으로 <b>한국어 부문</b> 등 비영어권 부문에서 <b>경쟁력 확보가 가능할 것으로 예상</b>
kakao 백상엽 대표	• <b>국산 AI 반도체 활용</b> 지원을 통한 <b>인프라 경쟁력</b> 제고 및 대학·연구기관 등에 <b>초거대AI 관련 인재양성</b> 지원 필요
 이세영 대표	• 초거대 AI 응용서비스를 <b>활성화</b> 하기 위해서는 <b>스타트업도 시장에 진입</b> 하여 경쟁력을 가질 수 있는 <b>생태계 조성</b> 지원이 절실
 최문정 교수	• 초거대 AI는 <b>사회적, 문화적으로 다양한 파급효과</b> 를 미칠 것으로 예상되므로 이에 대비한 <b>사회적 수용력 제고 노력</b> 도 병행해야 함

**【간담회 주요의견】**

**▲ 초거대AI 기술·산업**

- (학습용 데이터) AI 학습용데이터가 많은 도움이 되고 있으며, **한국어 학습용 데이터 구축** 등에 대한 투자는 **꾸준히 필요**(SKT, KT, 네이버)
- (R&D) 글로벌 기업이 개발한 **오픈소스 모델 기반의 엔지니어링**만 답습할 것이 아니라, **초거대 AI 기초·원천기술 확보**가 중요(LG)
- (인프라) **고성능·저전력 컴퓨팅인프라** 구축을 위한 **기술개발 및 컴퓨팅자원 지원** 필요(카카오, SKT, KT)

**▲ 초거대AI 생태계**

- (전문영역 AI) 일반 상식 수준을 넘어 **장기적으로 전문분야의 인사이트**를 가진 AI로 만드는 것을 고려한 투자 필요(오혜연 교수, KAIST)
- (인재) **프롬프트 엔지니어** 등 초거대AI 시대를 선도할 수 있는 **인재양성 및 활용역량 강화** 지원이 중요(이성환 교수, 고려대)

**▲ 초거대AI 제도·문화**

- (규제·제도) 초거대 AI의 도입·확산을 고려하여 **이전의 규칙, 제도들을 현재에 맞게 재정의**하는 것이 필요(네이버)
- (신뢰성) 초거대 AI는 **거짓답변, 편향성** 등에 한계를 보이고 있으므로, **신뢰성 및 성능에 대한 검증 데이터 셋**이 필요(LG)

## II 초거대 AI 시대 개막

◇ 초거대 AI의 등장, 본격 활용·확산으로 변곡점(Inflexion Point) 도달 → 디지털을 비롯한 사회·경제 전반의 새로운 혁신 촉발

### □ 챗GPT가 불러온 초거대 AI 열풍에 전 세계가 주목

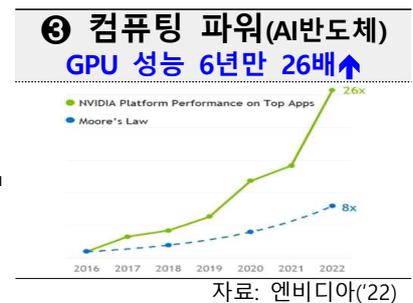
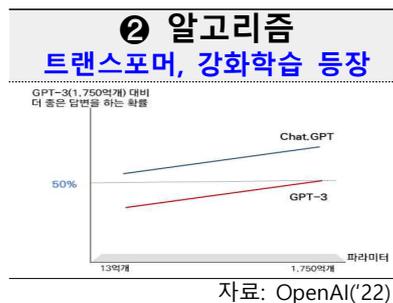
- AI는 높은 활용 난이도 등으로 기대감 대비 전면적 확산엔 한계 (글로벌 기업 중 약 34% 수준만이 인공지능을 비즈니스에 적용(IBM, '22))
- 최근 챗GPT 등 초거대 AI가 일상·업무 전반\*에 급속도로 확산 되면서, AI의 막연한 기대감이 높은 효용성으로 증명되는 중

\* 챗봇(챗GPT), SW코드(코파일럿), 이미지(DALL-E 2, 미드저니), 영상편집(어도비) 등

### □ SW가 끌고 HW가 미는 AI 혁신, 초거대 AI의 발전을 가속화

- 초거대 AI는 ①방대한 데이터, ②고도화된 알고리즘(트랜스포머 등), ③컴퓨팅 파워의 극적인 발전\*을 바탕으로 이뤄낸 성과

\* 엔비디아 H100('22.9)은 6년 전 출시된 P100 대비 26배 높은 성능(무어의 법칙의 3배 이상)



### □ 초거대 AI로 인해 기존 산업의 재창조가 시작

- 초거대 AI는 산업·사회 전반에 확산되어 생산성·업무효율 향상에 기여 ⇒ 초거대AI 기반 新서비스 시장이 급격히 성장할 전망



\* 글로벌 '생성형 AI' 시장(그랜드뷰 리서치, 억\$): ('22) 101 → ('30) 1,093(CAGR 34.6%)

## 참 고 초거대AI 기술현황

- (알고리즘) 초거대AI는 구글이 '17년 발표한 **트랜스포머** 알고리즘\* (출처: Vaswani 외, "Attention Is All You Need")을 사용하여 개발
  - \* 데이터 병렬처리 등으로 전통적 RNN(언어), CNN(이미지) 대비 처리속도·규모 확장
  - 트랜스포머는 인코더(이해) 파트와 디코더(생성) 파트 등으로 구성
  - 초거대AI는 ①인코더(구글 BERT, '18), ②디코더(OpenAI GPT('18~), 구글 LAMDA('22), 메타 OPT('22), ③인코더+디코더(메타 BART('19), 구글 T5('21))를 이용
  - \* 구글, 메타는 인코더 또는 인코더+디코더 방식을 이용하다 OpenAI가 GPT-3('20)에서 초거대AI의 성공 가능성을 보인 이후 디코더 방식도 출시('22)

### □ 주요 초거대AI 기업별 기술현황

- (OpenAI) 초거대AI 서비스를 **선도적으로 출시**

<b>GPT3.5('22)</b>	·GPT3 대비 <b>답변 정확도·신뢰성 향상</b> 을 위해 사람이 개입한 ' <b>지도학습</b> ', <b>답변 만족도 점수에 기반한 '강화학습'</b> 적용
<b>챗GPT('22)</b>	·GPT3.5에 <b>차별·혐오발언을 차단</b> 하는 기능 추가
<b>GPT4('23)</b>	·기존 대비 <b>추론능력, 멀티모달</b> (이미지 입력 → 추론), <b>대화 메모리 기능</b> (기존 8천 단어→ 6.4만 단어 입력·저장 가능) 등 향상 * 챗GPT 플러그인을 출시, API 호출기능으로 실시간 정보조회, 제휴사이트 협업

\* MS는 검색엔진에 챗GPT 적용, 워드·엑셀·PPT 등에 GPT4를 접목한 MS365 코파일럿 공개

- (구글) **답변 근거 제공, 실시간 정보를 반영한 답변 등 성능 차별화**

<b>LaMDA('22)</b>	·사용자 질의에 대해 <b>근거(Groundness) 기반으로 신뢰성 있는 응답</b> 을 지향한다는 점에서 챗GPT와 차별화 지향 * 품질(Quality), 안전(Safety), 근거성(Groundness) 지표로 모델성능 측정·개선
<b>Bard('23)</b>	·GPT3.5에 <b>차별·혐오발언을 차단</b> 하는 기능 추가

- (메타) OPT, LLAMA 등 **AI 모델 공개**로 초거대AI 영향력 확대 추구

<b>OPT('22)</b>	·GPT-3와 동일한 175B 파라미터, 비상업 라이선스 공개
<b>LLAMA('23)</b>	·다양한 크기의 모델(7/13/33/175B)에 많은 데이터 학습

## 참고

## GPT-4의 개요 및 특징

- ◇ OpenAI의 **GPT-4**(23.3.15 출시)는 기존모델(ChatGPT-GPT-3.5)대비 **추론능력, 이미지를 입력으로 인식하는 멀티모달, 대화메모리 기능 등 창의성이 향상된 모델**
- ◇ **무료 버전인 ChatGPT와 달리 ChatGPT PLUS(유료, \$20/월)로 업그레이드 해야 사용 가능**
- ◇ ①**50명 이상의 전문가 활용** ②**ChatGPT를 통해 얻은 결과 업데이트 등 6개월의 파인튜닝**(MS Azure에서 학습)을 통하여 기존 대비 **신뢰성과 정확성 개선**  
- 한편, **영리추구 및 MS와 합작하는 형태로 전환하면서 구체적 스펙 비공개**

### □ GPT4 특징 및 한계

<b>창의력</b>	<ul style="list-style-type: none"> <li>○ 요구한 질의에 대한 문제를 보다 창의적이고 정확하게 풀 수 있으며, 시나리오 작성 및 작문 등의 능력이 크게 향상</li> </ul> <p style="border: 1px dashed gray; padding: 2px;">* (예) 동화 신데렐라의 줄거리를 알파벳 A부터 Z까지 순서의 단어로 작성</p>
<b>추론 능력</b>	<ul style="list-style-type: none"> <li>○ 기존 ChatGPT의 성능을 능가하여 여러개 사례에서 공통점 추출 능력이 진화</li> </ul> <p style="border: 1px dashed gray; padding: 2px;">* (예) 3명의 사원 유희시간 목록을 제시하면 공통 유희시간 파악하여 회의예약 가능시간 안내</p>
<b>멀티 모달</b>	<ul style="list-style-type: none"> <li>○ 기존 텍스트로 국한된 입출력에서 이미지를 입력으로 받아 추론할 수 있는 기능 탑재</li> </ul> <div style="display: flex; align-items: center;">  <div style="margin-left: 10px;"> <p>* 냉장고 내부의 음식을 촬영한 사진으로 가능한 요리를 추천</p> <p>① 요거트 파르페- 요거트, 딸기, 블루베리, 꿀이나 견과류 추가</p> <p>② 당근과 후무스 랩- 랩에 후무스를 바르고 잘게 조각낸 당근을 올려 말아 드세요!</p> </div> </div> <p style="font-size: small; margin-top: 5px;">What are a couple of meals I can make with this?</p>
<b>대화 메모리</b>	<ul style="list-style-type: none"> <li>○ 기존모델이 약 8천 단어(4,096토큰)를 입력했던 반면 GPT-4는 약 6만 4천 단어* 입력을 저장하며, 이전대화의 기억을 더 오래 지속하여 대화맥락을 유지</li> </ul> <p style="font-size: small;">* A4 약 50페이지 수준으로 단편소설 정도의 분량</p>
<b>신뢰성</b>	<ul style="list-style-type: none"> <li>○ 악성 질의에 대한 대응 훈련을 통해 보다 윤리적인 측면의 답변을 하도록 진화</li> </ul>
<b>답변 유연성</b>	<ul style="list-style-type: none"> <li>○ 말투나 어조 등에 따라 답변 스타일을 적합한 형태로 변경하여 농담에 대한 이해발현 우수</li> </ul> <p style="border: 1px dashed gray; padding: 2px;">* (예) 마돈나는 왜 기하학을 공부할까? (답변) 모든 각도에서 포즈를 취하는 방법을 배우려고! ☺★♪</p>
<b>시험 능력</b>	<ul style="list-style-type: none"> <li>○ 미국 변호사 시험(상위 10%) 및 생물학 올림피아드(상위 1%) 등 표준화된 시험에서 우수한 성적 달성하며, 기존모델 대비 우수한 성능을 보임</li> </ul>
<b>한계</b>	<ul style="list-style-type: none"> <li>○ 여전히 사회적 편견, 악의적인 프롬프트에 대해 한계점을 보유하여 미흡한 추론 및 부정확한(최신정보 결여) 답변을 생성</li> </ul> <p style="border: 1px dashed gray; padding: 2px;">* (예) 너비가 무한한 출입구를 통과하기엔 탱크와 자동차중 어느것이 더 나은가? (답변) 무한히 넓다는 것은 출입구의 폭 제한이 없으므로 둘다 통과 가능(높이 고려X)</p>

### III 초거대 AI의 영향력

◇ 초거대 AI는 사람 수준의 언어·시각 능력을 토대로 디지털 산업 자체의 파괴적 혁신과 경제·사회 전 분야의 AI 내재화를 주도할 전망

#### □ [기술] 사람 수준에 버금가는 기술 진보로 영향력 무한 확장

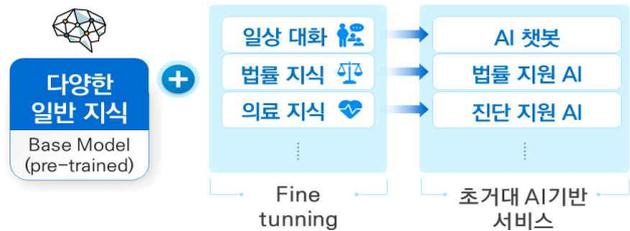
○ 초거대 AI는 트랜스포머\*, 강화학습(RLHF)\*\* 등 진보된 학습·알고리즘을 토대로 AI 자체의 획기적 성능 개선을 주도

\* 데이터 병렬처리로 RNN·CNN대비 처리속도·규모 확장 \*\* 답변 정확도·신뢰성 향상

○ 초거대 AI는 언어 생성 모델 중심으로 시작되었으나 이미지, 음성, 영상 등 언어와 연결(Mapping) 될 수 있는 다양한 영역으로 확장 추세

	<b>미드저니</b>	 <b>클로바 노트</b>	 <b>Meta Make a Video</b>
텍스트 또는 이미지 파일을 입력하면 자동으로 그림 생성		초거대AI+음성인식을 통해 음성인식 정확도 향상	
		텍스트 명령어를 입력하면 자동으로 짧은 동영상 생성	

○ 특히, 초거대 AI에 전문지식 데이터를 추가학습 하는 파인 튜닝 방식으로 다양한 분야로 AI 적용 확대에 기여



#### □ [AI·디지털 산업] 초거대AI를 지향하는 방향으로 선회 중

○ 최근 검색·워드프로세서 등 전통적 디지털 도구에 초거대 AI가 결합하면서, AI가 보조수단이 아닌 압도적 효율성 향상의 도구\*로 진화

	○ (OpenAI 챗GPT) 여러번의 검색 대신 질문에 대한 정확도 높은 답변 제공. 거짓답변 등에도 불구하고 검색시장 전반에 큰 변화 촉발 전망 → 검색 기반의 광고수익에 의존하던 포털 비즈니스 모델 변화도 예상
	○ (MS 코파일럿) 사용자 의도 및 업무 목적을 파악하여 문서 초안 작성, 문서 기반 PPT 파일 생성, 엑셀 데이터 분석 및 시각화 등 업무 효율 극대화

- 초거대 AI로 로우코드(Low-code)·노코드(no-code) 개발환경이 본격 확산되어 AI 개발·활용이 비전문가까지 확대되는 기폭제로 작동
  - \* 로우코드 플랫폼은 매년 45% 성장, '25년까지 70% 앱이 로우코드로 개발(가트너)
- 나아가 메타버스, 로봇 등 다양한 디지털 산업에 접목되어 사람의 의도를 이해하고 서비스 품질을 향상시키는 방향으로 활용

 <b>메타버스</b>	○(로블록스) 메타버스 공간 내에서 자연어 명령어를 기반으로 이미지, 오디오, 3D 모델링 등 새로운 콘텐츠 생성을 지원하는 도구 공개
 <b>로봇</b>	○(구글 PaLM-SayCan) 생성형 언어모델을 로봇에 접목하여 사람의 의도, 주변 환경을 이해하고 반응하는 로봇을 발표

- 결국, AI·디지털 경쟁력의 핵심으로 자리 잡은 초거대 AI를 선점하는 기업이 미래 AI·디지털 시장의 Key Player가 될 전망

#### □ [초산업] 전 산업 분야의 AI 내재화로 파괴적 혁신 유도

- 초거대AI의 학습능력을 바탕으로 디지털 적용 영역이 확장되고, 전 산업에 걸친 생산성 향상 및 업무 효율 증대를 촉진
  - 특히, 각 산업 영역에서 사람이 수행하던 지식서비스를 보조하고, 기존에 풀기 어려웠던 문제들도 해결토록 지원\*하는 역할까지 가능
    - \* (예) 생성 AI를 통해 희귀 의료데이터를 확보하여 AI모델 학습 훈련, 의료인 교육에 활용
    - \* 단, IoT 센서 등 산업용 데이터에 대한 기계학습을 통해 비용 효율적으로 AI를 개발하던 분야는 초거대AI 적용보다 기존 AI 개발 방식을 우선 적용 예상
- 산업 영역 별 특화서비스 개발은 ①초거대AI 플랫폼(Base-Model, 일반 지식)에 특화 전문지식을 추가로 학습(Fine-tuning)시키는 방식과
  - ②특화서비스별로 경량의 초거대AI 플랫폼(일반지식+전문지식)을 구축·활용하는 방식의 2가지 큰 흐름으로 발전해 나갈 전망
    - \* ① 방식은 전문분야 지식뿐만 아니라 일반지식으로 확장성 등에 강점, ② 방식은 해당 전문 영역에서 비용 효율적 성능 달성, 보안 문제 측면 등에서 강점

< 산업 분야별 초거대 AI 적용 사례 및 전망 >

 <b>의료</b>	○ 의료영상 판독을 통한 질환진단 보조 고도화, 임상·질환 합성 데이터 생성, 신약 후보물질 발굴, 약물 디자인 등 활용 확대 전망('23, 미래에셋)
 <b>마케팅</b>	○ '25년까지 생성형AI가 대기업 마케팅 메시지의 30%를 만들 전망(가트너), 美 1,000대 기업은 콘텐츠 생성(58%), 고객지원(57%) 등에 챗GPT 활용('23)
 <b>금융</b>	○ 고객상담, 금융상품 추천, 신용평가, 금융사고 감지 등 금융전반에 활용 확대 전망, 글로벌 금융회사의 20%가 대화형 AI도입 중(엔비디아, '23)
 <b>미디어</b>	○ 챗GPT를 활용한 이용자 맞춤형 콘텐츠, 퀴즈 제작을 발표(美 버즈피드), AI가 90%를 만든 블록버스터 영화가 '30년 1편 이상 개봉될 것(가트너)
 <b>법률</b>	○ 초거대AI가 법률대리인의 서류 작업 등을 지원하고, 판사 업무를 보조하는 등 리걸테크 서비스 고도화 전망

□ [사회] 모두가 AI를 쓰는 시대로의 전환, 기대(明)와 우려(暗)가 교차

○ 초거대 AI 활용이 본격화되면서, 단기적으로 ①AI의 신뢰성에 주목, 점차 ②AI의 사회·문화적 수용성으로 이슈가 확산될 전망

① AI 자체가 내포하는 기술적 한계(불투명성, 편향성 등)와 이용자의 부주의, 오·남용/악용 등에 따른 AI의 잠재 위험요인에 주목

\* 급속한 AI발전에 대한 우려시각도 일부 존재 → 美 비영리단체 '생명의 미래연구소'는 인류에 대한 위험성을 이유로 GPT4 이상의 강력한 AI 개발을 일시 중단할 것 촉구

② AI의 수용성 향상을 위해서는 AI를 이해하고 활용하는 AI 리터러시(Literacy)가 필수, 'AI·디지털 문맹'이 사회적 이슈로 부상

\* 싱가포르 교육부, 챗GPT 기반 학습방법 지도를 위해 교사대상 AI연수 운영중('23)

- AI 산출물에 대한 저작권 논쟁\*, 데이터 학습 과정에서 개인 정보 문제\*\* 등 초거대AI 산업 혁신과 기존 제도간 충돌 발생

\* MS 코파일럿(Copilot)은 학습 과정에서 활용한 오픈소스 코드의 출처를 밝히지 않아 'SW 불법 복제 행위'에 대해 미국에서 집단 소송 제기('22.11)

\*\* '이루다' 서비스 개발·학습 과정에서 개인정보보호 위반으로 인해 과징금과태료 부과('21.4)

- 새로운 AI 서비스에 대한 이해관계 집단 간의 갈등도 커질 우려



(법률) AI 법률 플랫폼 '로톡'이 변호사법 위반 혐의로 고발



(의료) 의료AI 활용 책임주체, 대중신뢰 등의 문제 이슈화

## 참고

# 스탠포드 HAI가 분석한 생성AI의 영향력

\* HAI: Human-Centered Artificial Intelligence(인간중심 인공지능 연구소)

- 총론: 생성AI 기술이 변곡점에 도달, 사회적 영향력 확대
  - 언어·이미지·음성 등을 생성하는 AI 성능이 변곡점(Inflexion Point)에 도달하여 기존에 풀기 어려웠던 문제를 해결토록 지원가능
    - \* (예) 생성 AI를 통하여 희귀 의료데이터를 확보하여 AI모델 학습 훈련이나 의료제공자 교육에 활용
  - 여전히 AI의 편향성, 악의적 사용 등으로 인해 우리의 일상과 사회에 위험이 될 수 있다는 점에 주의를 기울여야 할 것

## □ 분야별 생성 AI가 미치는 영향

	<ul style="list-style-type: none"> <li>○ 실제 환자 데이터를 사용해 만든 '합성환자'를 대거 양산하여 치료받는 환자와 비교하기 위한 대규모 통제집단 데이터 확보가능</li> <li>○ 다양한 질병변종을 만들어 교육현장에서 학생들이 무한한 사례학습을 통해 특정 질병을 식별·진단</li> </ul>
<p>의료</p>	<ul style="list-style-type: none"> <li>○ 다만 편향된 데이터로 생성된 합성환자들을 통한 결과는 부정확할 수 있음</li> </ul>
	<ul style="list-style-type: none"> <li>○ 작곡, 그림그리기 등 예술작품 제작이 가능해지며 창작 과정에서 아이디어와 상호작용하며 새로운 창작물을 만들어내는 등 창작과정을 보조하는 도구로 사용</li> <li>○ 인공지능이 창작물을 완전히 대체할 수 없으며, 규제가 없는 상황에서 교육에 있어 딥페이크 및 표절문제 우려</li> </ul>
<p>예술</p>	
	<ul style="list-style-type: none"> <li>○ 개인·법률대리인·판사의 보조적 역할 수행</li> </ul>
<p>법치주의</p>	<p style="border: 1px dashed gray; padding: 2px;">* 생성AI를 통해 개인은 필요한 법률문서를 준비하고, 변호사는 법률연구와 문서작업을 도움받고, 판사는 판결의 정확성과 효율성을 개선시킬 수 있음</p> <ul style="list-style-type: none"> <li>○ 법적 의사결정에서 공정성과 신뢰성이 최우선가치이므로 생성AI가 잘못된 정보를 제공하고 사례들을 만들어 낼(Fake) 경우 커다란 문제 발생</li> </ul>
	<ul style="list-style-type: none"> <li>○ 전문가가 수행하던 작업을 자동화·고도화하면서 업무를 보조하는 형태로 생산성 증대</li> </ul>
<p>노동</p>	<p style="border: 1px dashed gray; padding: 2px;">* (예) 미국 주별 투표 패턴과 경제성장 변화의 상관관계를 시각화하는 작업 (기존) 데이터확보, 코딩, 시각화 도구 작업 등 여러 시간 소요→ (생성AI) 몇초 안으로 단축</p> <ul style="list-style-type: none"> <li>○ 생산성 증가로 노동자 임금이 증가할 수 있으나, 노동자가 AI로 대체될 수 있어 정책입안자 등 여러 경제주체들의 선택에 의해 노동시장이 변화할 예정</li> </ul>
	<ul style="list-style-type: none"> <li>○ 학생들에게 맞춤형 프로그램을 제공하거나 실시간 모니터링을 통한 개별적인 피드백을 제공하는 등의 방식으로 학습 불균형 해결 가능</li> <li>○ 글쓰기의 질은 사람의 사고능력에 기반하므로 학습현장에서 생성AI에 과의존시 학생의 사고력과 창의력 저해</li> </ul>
<p>교육</p>	

## IV 우리의 경쟁력 진단

◇ 그간 데이터 축적, AI 기술력 향상 등 인공지능 기반 조성에 집중, 우리기업은 독자적 초거대AI 플랫폼을 확보하고 세계를 향해 도전 시작

\* 우리나라는 자체 초거대 AI를 보유한 4개국(미국, 중국, 이스라엘, 한국)에 포함

👉 이를 바탕으로 우리나라가 글로벌 초거대AI 시장을 선도해나갈 수 있도록 초거대AI 대응 관점에서 AI 정책을 재검토, 새로운 정책방향 제시

### □ 초거대AI 발전에 필요한 인프라 구축 및 인재 양성

□ **(데이터)** 그간 정부 주도로 공공·민간 분야에 대규모 데이터 구축·개방, 일부에서는 글로벌 기업 수준의 성과\*도 창출

\* 이미지(물체 탐지 및 인식) : (AI Hub) 3억9,800만장 > (DALL·E1) 2.5억만장 (Imagenet) 1,400만장  
한국어 음성(음성 대화) : (AI Hub) 23만 시간 > (MS AI) 6만시간(쑈언어) (메타 AI) 5만시간

AI학습용 데이터	민·관 협력 데이터	공공데이터
산업 파급효과가 큰 분야 691종·구축(~'22.12월)	21개 분야, 7,753종 데이터 플랫폼 구축·운영(~'22.12월)	공공데이터 데이터셋 7.7만개 개방(~'22.12월)

○ 데이터 기업\*이 지속적으로 증가('18) 7,153개→('22) 8,940개, 연평균 5.7% ↑)하고, 민간 주도의 데이터 유통·거래·활용 기반 조성\*\* 시작

\* 데이터를 구축하거나, 가공·처리 또는 판매하는 서비스를 사업으로 영위하는 기업

\*\* 데이터 거래사('23, 100명), 가치평가기관('23, 4개) 및 품질인증기관 최초지정('23) 등

□ **(알고리즘)** 초거대AI는 학습 알고리즘(트랜스포머), 학습방식(강화학습 등)이 공개되어 있고,

○ 그간 확보한 AI 기술력(美 대비 89.1%), AI 특허 경쟁에서의 가파른 상승세(연평균 89.7%) 등이 시장 선점의 토대가 될 것으로 기대

\* AI R&D 투자규모 : ('19) 395억원→ ('20) 640억원→ ('21) 843억원→ ('22) 897억원

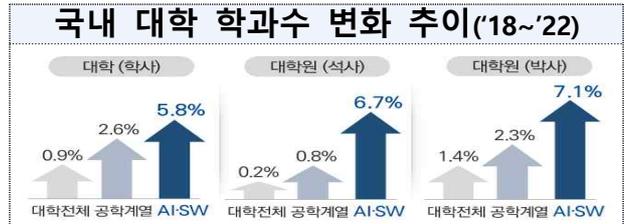
\* AI특허출원건수('11~'20) : 美(35.6%, 15,035건), 中(31.0%, 13,103건), 日(11.6%, 4,906건), 韓(11.3%, 4,785건) 순, '19년 이후 연간 출원량은 일본을 추월(특허청, '23)

□ **(컴퓨팅 자원)** 그간의 정부 투자\* 등으로 국산 NPU·PIM 출시, 민간 클라우드 기업과 협업하여 AI반도체 Farm 구축 착수('23~)

\* 차세대 지능형 반도체 기술개발 (총 1조 96억원 중 '20~'22년까지 3,516억원 집행)

○ 세계 최고의 메모리 기술력을 토대로 AI 반도체 승부를 건다면, 국내 컴퓨팅 파워-AI 경쟁력의 동반 혁신 시너지 창출 가능

□ **(핵심 인재)** 최근 5년간 AI·SW 분야 졸업생이 학사 81,499명(연4.4% ↑), 석사 10,509명(연11.5% ↑), 박사 3,936명(연14.3% ↑) 등으로 증가 추세



\* '18년 대비 '22년 AI·SW 관련 학과가 대학 128개, 석사과정 75개, 박사과정 46개 증가

○ AI 시장 성장에 따른 인력 수요\*에 대해 적극적으로 대응하고, 초거대AI를 선도하는 글로벌 리더급 AI 연구자\*\*를 양성할 필요

\* '21년 4,000명, 향후 5년간 3만명 부족할 것으로 전망(SPRI)

\*\* 연구실적 기준 세계 3,000위 내 한국인은 단 11명에 불과('22.12 기준, Research.com)

⇒ 초거대 AI의 미래 주도권 경쟁 확보를 위해, 데이터·알고리즘·컴퓨팅 자원과 핵심인재 양성에 대한 정책역량 집중

□ **우리 기업은 초거대 AI 경쟁에 본격 합류하며 가시적 성과 창출**

□ **(AI 플랫폼)** GPT-3 등장 이후, 국내 대기업 역시 초거대AI 개발 경쟁에 본격 참여하면서 가시적 성과\* 창출을 시작

\* 네이버 하이퍼클로버 플랫폼 : API 사용 500개, 응용서비스 출시 20개 기업 달성('23.3월 기준)

네이버, '클로바'	LG, '엑사원'	카카오, KoGPT	SKT, '에이닷'	KT, '믿음'
<b>Clova</b> <sup>3</sup>	<b>EXAONE</b>	<b>KoGPT</b> kakao	<b>A.(에이닷)</b>	<b>MI:DEUM</b>

○ 우리가 우세한 한국어 기반 초거대AI 모델·서비스를 개발하고, 성공사례를 바탕으로 비영어권 시장을 집중 공략해 나갈 필요

\* Open AI의 GPT-3는 영어, 중국어, 스페인어 등 다수 언어를 학습하고, 한국어는 일부에 불과하나, 우리나라는 한국어 중심으로 학습(네이버 한국어 학습비중 97%)

□ **(AI 응용) ①**대기업은 자사 서비스(AI케어콜, 메타버스 등)에 초거대 AI를 접목하여 성능·효율성을 향상하는 한편,

\* (예) SKT는 '에이닷'을 중심으로 자율주행·메타버스 등 결합을 추진(K-AI 얼라이언스)

○ **②**중소·스타트업은 초거대 AI 플랫폼에서 제공하는 API를 이용하여 특화영역 데이터를 추가 학습한 '전문 AI 서비스'를 시장에 출시

\* (뤼튼) 작문·콘텐츠창작 플랫폼 출시, (업스테이지) 사진 답변 등 AI챗봇 'AskUp' 출시

○ 전문영역 AI 서비스를 선도적으로 출시하고 글로벌 레퍼런스를 확보하여, 초거대AI를 새로운 수출 동력으로 창출할 필요

⇒ 우리의 강점인 한국어 분야와 전문·특화 부문 서비스 개발을 통해, 새로운 디지털 수출 동력을 확보하고 글로벌 초거대AI 시장을 선도

□ **AI 일상화 시대를 맞이하기 위한 사회적 수용성 향상 필요**

□ 챗GPT를 계기로 누구나 AI를 쓰는 시대가 도래, 사회·산업 전반에서 초거대AI를 폭넓게 수용하기 위해서는 신뢰성 확보가 필수

○ 이미 우리 일상 곳곳(교육, 일자리 등)에서 초거대 AI 활용·확산에 따른 이슈가 대두, 미리 대비하지 않으면 사회적 혼란 초래 우려

교육 분야에서 챗GPT 논란	초거대 AI가 위협하는 직업군
 <div style="display: flex; justify-content: space-around; align-items: center;"> <div style="border: 1px solid gray; padding: 5px; background-color: #e0e0e0;">부정행위로 금지 필수</div> <div style="font-size: 2em; font-weight: bold;">VS</div> <div style="border: 1px solid gray; padding: 5px; background-color: #e0e0e0;">보조 도구로 적극 활용</div> </div>	<div style="display: flex; justify-content: space-around;">    </div> <div style="display: flex; justify-content: space-around; font-size: 0.8em;"> <span>낮은수준 코딩 엔지니어</span> <span>작가 등 콘텐츠 크리에이터</span> <span>그래픽 디자이너</span> </div>

□ 초거대 AI 산업진흥과 책임문제(플랫폼 독점, 개인정보 유출, 가짜정보 생성 등) 사이의 규제 이슈도 균형 있게 풀어나갈 필요

⇒ 초거대 AI 일상화 대비, AI의 윤리적 활용과 신뢰성 확보와 더불어 AI에 대한 사회적 수용도 제고를 위한 선제적 대응 필요

# 참고 국내 초거대 AI 개발 현황 및 전략

## 네이버 하이퍼클로바(Hyper CLOVA)

- **(현황)** 자사서비스\*에 초거대 AI(Hyper CLOVA, 2,040억개 파라미터) 적용과 함께 플랫폼으로서 API 개방을 통한 AI 생태계 조성\*\* 추진
  - 삼성전자와 함께 초대규모 AI에 필수적인 AI 반도체 솔루션 개발
- \* 클로바케어콜(돌봄서비스), 클로바노트(홍보문구 생성, 음성→텍스트 전환) 등
- \*\* 플랫폼 개방(클로바스튜디오) 후 API 활용 기업 500개, 응용서비스 출시 기업 20개(23.3월 기준)
- **(계획)** 검색 특화 '서치GPT\*'('23.7월), 한국어 능력을 향상한 '하이퍼클로바X\*\*' 출시 예정('23.7월)
  - \* 초대규모 기술 활용한 차세대 검색 서비스
  - \*\* 생성AI 기반이 되는 거대 언어모델(LLM, Large Language Model)



<Hyper CLOVA>



<클로바스튜디오 사례: 우주문방구 스토리네이션>

## LG EXAONE

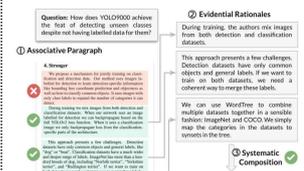
- **(현황)** 실제 산업현장에서 활용할 수 있는 상위 1% 전문가 수준의 모델(EXAONE\*) 개발·적용 중
  - **(Universe)** 제품 고객 상담, 금융 상담, 코드 생성 등 생산성 향상
  - **(Atelier)** 이미지 생성 및 이해에 기반한 디자인 창의성 제고
  - **(Discovery)** AI, 화학, 제약, 의학, 바이오 등 전문 분야 최신연구 협업
- \* 텍스트, 이미지를 모두 학습하는 멀티모달 및 한국어, 영어를 병행 학습하는 멀티링구얼(multi-lingual) 방식의 초거대 AI 모델
- \*\* 사실과 근거에 기반한 답변이 가능하도록 모델 구조를 혁신하고 신뢰성 있는 전문 데이터(전문문헌, 기업문서, 논문/특허 등)를 학습
- **(계획)** EXAONE 3대 플랫폼(Universe·Atelier·Discovery)을 중심으로 적용 기업 지속 확대('23년~)



타투 프린터 이미지 생성 <LG생활건강 상용화>



분자 구조식 인식/학습 <멀티모달을 통한 논문/특허 이해도 향상>



답변의 학술적 근거 제시 <전문 문헌 QA>

**카카오**

**KoGPT, Karlo**

- (현황) AI 한국어 모델('KoGPT', 파라미터 60억개)과 이미지 생성 모델('Karlo')개발 및 고도화

- KoGPT와 Karlo를 활용한 **버티컬 AI 서비스\***에 집중

\* AI 시인 시아(SIA) 시집 출간('22.8월), AI 이미지 생성기인 B^EDIT 웹서비스 출시('23.3월)

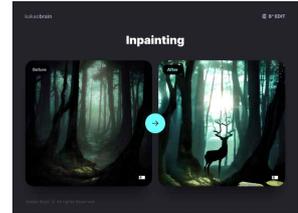
- (계획) 초거대 AI모델 통한 사용자 서비스 고도화

- 카카오톡 프로필 기능에 '생성형 AI이미지 기술' 적용('23년 4월)

- KoGPT 업그레이드 버전 출시('23년 상반기) 및 이에 기반한 한국형 ChatGPT 출시('23년 하반기)



< KoGPT >



< B^EDIT >

**KT**

**믿:음**

- (현황) 언어 기반 초거대 AI 모델('믿:음')로 API 서비스 출시(13종, '22년 11월), 기존 서비스(AI콜센터, 기가지니 등) 순차적 고도화

- '믿:음'의 하반기 상용화 통해 AI서비스 대중화

- (계획) FADU, Moreh, 리벨리온 등 우수기술 보유 스타트업 협력 기반 Full Stack 구현 통해 초거대 AI 학습 및 서비스 상용화

- HAC(Hyperscale AI Computing) 서비스 기반 초거대 AI 모델 학습 및 개발(220B, '23년 상반기)

- 리벨리온 ATOM 기반 LLM 서비스 상용화('23.3Q)



< AICC(컨택센터) >



< ATOM칩(리벨리온) >

**SKT**

**에이닷(A.)**

- (현황) 세계 최초 GPT-3 기반 대화형 앱 '에이닷(A.) 상용화('22.5), 검색·아바타·음악·콘텐츠 등 자사 IT서비스와 연계 제공

- BERT, GPT3 등 최신 언어모델 공개 논문을 참고, 자체 개발

- (서비스) 일상에 유용한 정보 <sup>캘린더/증권 등</sup> 및 다양한 콘텐츠 <sup>TV/음악/게임 등</sup> 제공, **통신서비스** 관련 기능 <sup>요금제 조회, 멤버십 등</sup> 보유

- (계획) 국내·외 제휴를 통해 차별화된 기술·기능 고도화

- Home UX 개편, 뉴스·전화 브리핑/컨시어지 서비스 도입('23.2Q) 및 Partner Agent 도입('23.3Q)



< 에이닷(A.) >

# 참고 초거대 AI 응용서비스 주요 현황

## □ 해외 기업

### 작문보조 슈도라이트(OpenAI GPT-3 기반)

- 작가가 초안을 입력하고 몇가지 단어를 선택하면, AI가 문장 보완, 줄거리·은유 제시 등 재작성



### 언어 번역 NLLB-200(메타)

- 200개 언어 텍스트를 번역할 수 있는 AI 서비스 (일부 아프리카 인도 지역 언어의 경우 70% 이상 정확도)



### 이미지생성 미드저니(Midjourney AI 연구소)

- 딥러닝을 통해 이용자가 입력한 텍스트의 키워드에 따른 이미지를 찾고 조합해 이미지를 생성



### 코드 생성 코파일럿(깃허브)

- 입력 문장(자연어 요구사항)에 맞는 코드를 생성
- GPT-3 모델을 이용하여 깃허브의 수많은 레포지토리들을 학습하는 방식으로 개발



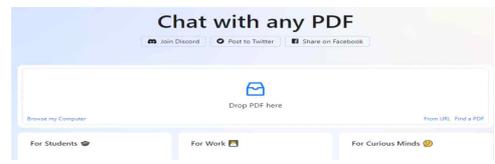
### 검색엔진 you(you.com)

- 검색 사이트에 이미지 생성, 작문, 코드 생성 등 다양한 과업을 수행할 수 있도록 검색엔진과 연동



### 문서질의 ChatPDF

- PDF파일을 업로드하여, PDF의 주요내용에 대해 질의응답하는 서비스



### 비디오생성 GEN-1(런웨이)

- 텍스트 프롬프트에서 지정한 스타일을 적용해 기존 비디오를 새로운 비디오로 변환



### 광고디자인 jasper(Jasper)

- 블로그 게시물, 광고문구 작성, 마케팅 문구 작성을 지원하는 AI 카피라이팅, 이미지 생성 서비스 제공



## □ 국내 기업

**작문보조** **뤼튼 트레이닝(뤼튼 테크놀로지스)**

- SNS 광고, Push 알림, 보도자료 등 **상황에 맞는 문장 생성**이 가능하며 **입력한 글에 대해 적절한 피드백과 참고자료**를 제공



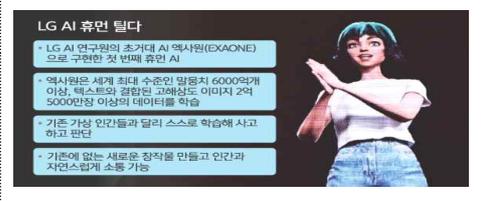
**작문보조** **모카(아스타컴퍼니)**

- 광고, 헤드라인, 상품설명, 세일즈 카피 문구 등을 상품 콘텐츠에 맞게 생성



**이미지생성** **틸다(LG)**

- 언어-이미지 양방향 데이터를 생성하는 AI로서, **틸다가 창작한 이미지·패턴을 토대로 박윤희 디자이너가 의상 제작**(뉴욕 패션위크)



**돌봄서비스** **클로바 케어콜(네이버)**

- 사람 수준의 자연스러운 대화 서비스를 통해 **독거 어르신**의 외로움 해소, **돌봄 역할** 수행



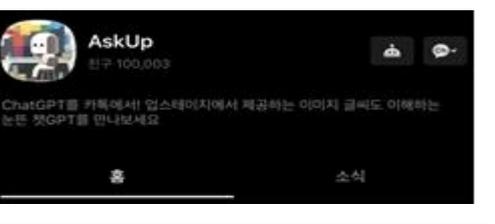
**업무지원** **킵grow(유니드컴즈)**

- SMS·메일 발송, SNS 포스팅 등 14종 마케팅 업무 자동화 서비스 제공



**문서요약** **아속업(AskUp)(업스테이지)**

- ChatGPT 기반으로 OCR 기술을 결합해 개발한 카카오톡 AI 챗봇으로 사용자가 문서를 찍거나 전송하면 내용을 읽고 이해하고 **답변**을 제공



**세무보조** **연말정산 알리GPT(올거나이즈)**

- 연말 정산과 세금 관련 질문에 1초 내로 맞춤형 **답을** 제공, **삼점삼 앱**에서 해당 기능 이용 가능

**연말정산 알리GPT**

최근 데이터로 학습되어 있으나 답변이 정확하지 않을 수 있어요.

연말정산 궁금증 질문하기

**연말정산 알리GPT**

**2023년 연말정산에서 가장 크게 달라진 점은 무엇인가요?**

2023년 연말정산에서 가장 크게 달라진 점은 소득공제의 상한기간에 따라 공제 가능 범위가 달라지는 것입니다. 연 300만 원에서 최대 1,500만 원

**여행서비스** **AI여행플래너(마이리얼트립)**

- ChatGPT 기반으로 맛집·명소·날씨·팁·여행지 추천 등 여행 관련 다양한 주제에 대한 대화 가능





## VI. 전략 및 추진과제

### 전략1. 초거대 AI 기술·산업 인프라 확충

#### 1 민간의 초거대AI 개발을 지원하는 학습용 데이터 구축

◇ 초거대AI 개발을 위해서는 **대규모 텍스트 데이터 필요** ⇒ 라벨링 데이터 중심의 학습용 데이터 구축사업을 개편, 양질의 대규모 **텍스트 데이터 보강**

현재	개선
AI 학습용 데이터 사업은 라벨링 데이터 구축 중심	초거대AI 사전학습에 필요한 텍스트(비라벨링) 데이터 구축 신설

#### □ 초거대 AI 학습용 데이터 구축 보강

○ **(초거대 AI 학습용 데이터)** 학습용 데이터 구축사업을 초거대 AI 개발에 필요한 최소한의 데이터 수요를 충족시키도록 개편(23~) **신규**

- (언어) 산업계의 수요(양질의 한국어 텍스트 300억개 토큰 이상)를 반영, 초거대 AI 학습용 텍스트 데이터 200종(200억개 토큰, 약 100억개 기 구축) 구축

- \* 학습용 데이터 1종에 1억개 이상의 토큰이 포함될 수 있도록 과제 기획
- \* 200억개 토큰은 책 약 15만권에 해당하는 분량(책 1권당 13.3만 토큰 가정)
- \* '23년 한국어(22종), 문화관광(2종), 교육·법률·금융(각 1종) 등 총 34종 구축 착수(476억원)

▶ 산업분야별 특화 학습용 데이터(텍스트+파인튜닝 데이터)를 구축하고, 비영어권 시장공략을 위해 필요한 동남아·중동 등 언어데이터 단계적 확대

(예시)	현행	라벨링 데이터 10만건	⇒	개편 (초거대AI)	텍스트 1억 토큰	+	라벨링 데이터 1만건
------	----	-----------------	---	---------------	--------------	---	----------------

- (이미지) 현행 방식으로 라벨링 데이터를 구축하되, 초거대AI 개발에 필요한 이미지 설명 텍스트를 추가토록 과제별 과업 개편

- \* 산업계는 2억장 이상의 이미지가 필요하다는 의견, 동 수요는 기 충족(약 4억장 구축)
- \* '23년 이미지 48종 구축(763억원)

		“도로 위에서 말탄 기수와 + 모터사이클을 탄 경관이 대화를 나누고 있다.”
기존 라벨링	초거대AI 사전학습용 텍스트 추가	

※ 초거대 AI 학습용 데이터 구축 목표치 산출 근거

- (초거대AI 데이터 수요) 정제된 텍스트 300억 토큰 이상, 이미지 2억장 이상
  - (언어) 국내 AI기업들(카카오, LG, NC소프트 등)은 양질의 한국어(도서 등) 비라벨링 데이터 300억 토큰(1문장≈10토큰) 이상 구축 필요 의견 제시
    - \* 3,000억개 토큰 이상의 데이터가 필요하며, 일정 수준의 성능 구현을 위해서는 10%(300억개 토큰) 이상의 정제된 데이터 필요(GPT3도 약 3,000억개 토큰 사전학습)
  - (이미지) LG는 엑사원 개발에 약 3.5억장, Dall·E1은 2.5억장, Dall·E2 6.5억장, MidJourney 1억장 등 통상 2억장 이상을 활용
- (기 구축 현황) '22년 총 691종 구축(언어 약 89억 토큰, 이미지 약 3.98억장)
  - ⇒ 이미지는 수요 최소량을 충족하나, 언어는 수요 대비 200억 토큰 이상 부족
    - \* AI 학습용 데이터 사업 목표 1,300종('25년) 중 691종 기 구축(~'22년), 609종 구축 예정
- (구축 목표치 설정) (언어) 산업 분야별 특화 데이터 200억 토큰 구축, (이미지) 별도 과제 신설 없이, 기존 과제 내에 초거대 AI 학습용 데이터 추가 구축

- (한국어 말뭉치) 초거대AI 한국어 성능 향상을 위한 고품질 말뭉치\*, 한국어 응용말뭉치\*\* 구축('23년 25종 등 '27년까지 총 130종 구축, 문체부) **강화**
  - \* 한국어 분석정보(개체명, 감성 등) 부착 말뭉치 \*\* 한국어-외국어 수어점자 통번역 등
- (보건의료 텍스트·영상) 보건의료 분야 비정형 데이터(응급실기록지, 의료영상·판독문 등) 구축·개방(~'26년 30만건 구축, 복지부) **강화**

□ 국가지식·글로벌 데이터 개방·활용 촉진

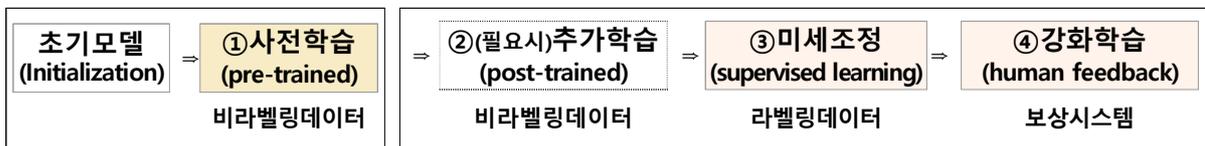
- (국가지식정보) '국가지식정보 통합플랫폼\*'의 저작권 제약이 없는 논문·보고서 메타데이터\*\*를 개방하여 초거대AI 학습에 활용토록 제공('24~) **신규**
  - \* 1차 구축('22, 81개 사이트 연계) → 2차 구축('23, 20개 사이트 연계) → 대국민 오픈('24.1)
  - \*\* 초록·요약, 분야, 유형, 출처, 주제어, URL 등('22년 기준 약 1.64억건)
- (글로벌 데이터) 아세안 사무국과 협력하여, 한-아세안 데이터 협력 프로젝트 공동 기획(아세안 협력기금 활용), 사업 참여 추진('23~) **신규**
  - 우리나라 및 아세안 10개국 데이터 공동 축적·활용 협력체계(국가간, 기업간)·플랫폼 구축, 공동 데이터를 활용한 AI 경진대회 개최('24~)

## 참고 초거대 AI 지원을 위한 학습용데이터 사업 추진방향

### □ 초거대 AI 학습 방식 및 학습용 데이터

- 초거대 AI는 다량의 텍스트(비라벨링)를 사전학습시켜 베이스모델을 만들고, 원하는 서비스에 적합하게 라벨링된 데이터로 미세조정된 형태
  - 베이스모델 기반의 특화서비스 개발을 위해서는 해당 전문영역 '텍스트(추가학습\*)+라벨링(미세조정)' 데이터를 이용한 학습 필요
    - \* 전문영역 특화 학습을 위한 비라벨링 데이터는 사전학습 단계에서도 적용 가능
  - 텍스트 데이터를 많이 학습할수록 적은 양의 라벨링 데이터를 이용한 지도학습으로도 높은 성능 달성 가능

#### < 초거대 AI 모델 개발 과정 >



※ 트랜스포머(초거대AI 알고리즘) 이전의 일반적인 딥러닝 모델(RNN, CNN)들은 사전학습 없이 라벨링된 데이터를 학습(③지도학습+④강화학습)해 AI 서비스를 개발

### □ 초거대 AI 지원을 위한 학습용데이터 구축사업 추진방향

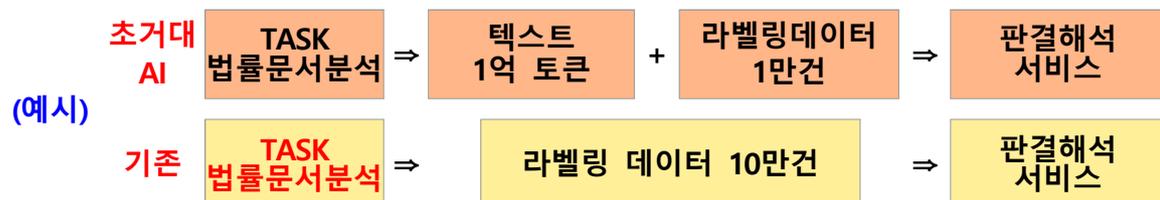
- 현재 학습용 언어 데이터는 라벨링 데이터 중심으로 구축해서 텍스트 데이터는 매우 부족한 상황

\* 텍스트(원문)는 라벨링에 필요한 주요 핵심 문장·문단만 발췌·공개

⇒ 초거대 AI 학습용 데이터는 다량의 텍스트(비라벨링) 중심으로 구축 하되, 미세조정에 필요한 라벨링 데이터를 추가하는 형태로 추진

#### ※ [예시] 법률 전문도서를 활용한 초거대 AI 학습데이터 1종 구성

- 전문가들은 초거대AI 특화서비스 개발에는 텍스트 1억 토큰과 라벨링데이터 1만건 필요 의견



## 2 초거대AI 미래 원천기술 확보 지원

◇ 초거대 AI는 **최신정보 미반영, 거짓답변 등 한계** ⇒ 현행 AI 기초연구에 추가하여 **초거대 AI 한계 돌파 기술개발** 착수, 차세대 초거대 AI 선도기반 조성

현재	+	개선
AI의 학습능력, 소통능력, 신뢰성 등을 강화하는 기초연구 진행	+	초거대AI에서 부각된 최신정보 미반영, 거짓 답변 등 한계 돌파 기술개발

### □ 초거대AI 한계돌파를 위한 차세대 기술 개발 추진

- **(초거대AI 한계 돌파)** 논리적 리즈닝(인과관계 이해), 실시간 정보 반영, 편향성 필터링 등 초거대AI 한계돌파 기술 개발을 착수('24~) 신규
- 초거대AI 기업과 대학이 컨소시엄으로 참여하여, 초거대AI 기업은 인프라 및 R&D 수요 제공, 대학은 과제 수행 및 전문인력 양성 등 수행

< 초거대 AI 한계 극복을 위한 R&D 과제 추진방향 >

구분	현재 초거대AI(챗GPT) 한계	한계 돌파 방향
<b>논리적 리즈닝</b>	·통계적으로 적합한 다음 단어를 생성할 뿐 '맥락'을 이해 못하는 한계	·생성 결과에 대한 <b>인과관계, 경위, 출처 설명</b> 등을 통해 사실성 확보
<b>실시간성 반영</b>	·학습한 과거데이터 기반으로 결과를 도출하여 최신정보 반영이 불가	· <b>실시간 정보검색, 외부 데이터반영</b> 을 통해 데이터 적시성·최신성 확보
<b>멀티모달 인터랙션</b>	·최적화된 질문을 통해서만 사용자가 원하는 수준의 답변 획득 가능, 질의 방식이 텍스트 입력으로 한계	· <b>사용자의 의도를 파악</b> 하여 최적의 답변 제공, <b>멀티모달</b> (이미지, 음성, 영상 등) <b>입력·인터랙션</b> 제공
<b>편향성 탐지·제거</b>	·학습 데이터에 따라 인종·성별·정치성향 등이 편중된 콘텐츠 생성	· <b>편향성 여부를 사전 필터링·제거</b> * 편향성 제거 학습방법 연구는 기 추진중
<b>언어모델 경량화·최적화</b>	·수천억 파라미터로 개발·운영하기 위해 대규모 데이터컴퓨팅 자원 필요	· <b>적은 수의 파라미터</b> 로 구성된 언어 모델의 성능 고도화·최적화 기술 개발

※ '24년 차세대 생성 AI 기술개발 과제 신설 추진 → 203개 산·학·연 기관이 참여하는 'AI 혁신허브', '전문위원회' 등을 통해 과제 발굴·기획·선정('23.下)

▶ AI 학습능력을 획기적으로 향상시키고 인간과 소통능력, 신뢰성을 강화하는 등 **기존 AI의 한계를 극복**하기 위한 기초연구는 기 추진중('22~'26, 2,655억원)

\* 딥러닝 한계(소량 데이터 학습, 실세계 변화 적응), 소통 한계(상식기반 추론, 인간(AI)-AI 소통), 신뢰성 부족(설명가능한 AI, 편향성 완화 등) 극복

### 3 초거대AI 인프라 기술개발 및 대규모 컴퓨팅 자원 제공

◇ 초거대 AI는 **막대한 컴퓨팅 자원** 필요 ⇒ 국산 AI반도체 기반의 **고성능·저전력 K-클라우드**를 초거대 AI가 **활용**할 수 있도록 **핵심 HW·SW** 개발

현재	개선
AI 반도체 기술개발, NPU Farm 실증 컴퓨팅 인프라 연중 제공(TF급)	초거대AI 인프라로 K-클라우드 고도화 대용량 컴퓨팅 인프라 집중 제공(PF급)

#### □ 고성능·저전력 컴퓨팅 인프라 구축을 위한 핵심기술 개발

- **(K-클라우드 고도화)** 국산 AI반도체 기반 클라우드에서 초거대 AI를 개발·운영할 수 있도록 지원하는 핵심기술 개발을 단계적 추진 **강화**
- **(AI반도체 SW)** AI반도체를 이용한 서버 클러스터(On-premise) 구축을 위해 병렬·분산처리 지원 SW, 서버 클러스터 관리 SW 등 개발('23~)
- **(데이터처리 가속 HW)** 보드간·서버간 데이터 공유를 가속화하기 위한 차세대 인터페이스\*, 데이터 처리 가속기(DPU)\*\* 등 개발('24~)
  - \* 프로세서가 관리할 수 있는 메모리 용량과 대역폭을 확장할 수 있는 인터페이스
  - \*\* 대용량 데이터 병렬분산 처리를 위해 보드간서버간 NW, 데이터 저장 등 처리를 지원하는 가속HW
- **(클라우드 컴퓨팅 지원)** 국산 AI반도체 기반의 서버 수백개를 연결, 클라우드 컴퓨팅을 지원하는 HW·SW 개발·실증('23년 예타 신청) **신규**

#### □ 초거대AI용 대용량 컴퓨팅 자원 제공

- **(초거대AI용 컴퓨팅자원)** 초거대 AI 연구·개발 특성에 맞게 컴퓨팅 자원 지원 용량을 대폭 확대(기존 TF급 → PF급으로 확대) **강화**
- 중소·벤처기업 대상 대용량 컴퓨팅자원(1PF, 최대 3개월) 제공('23.4~)
- AI 집적단지 데이터센터 운영 개시('23.10)에 맞춰, 연구소·대학까지 대규모 컴퓨팅자원 제공을 확대하고, 유연한 자원할당 체계 구축('23.10~)
  - \* LLaMA-13B 모델의 경우 A100 60대 규모 컴퓨팅 자원(1.2PF, 3개월 기준) 활용(Meta, '23.2)

구분	개별용량	자원총량	지원기간	비용
'23.10월~	2PF 이내	11.2 PF	최대 3개월	(무상)비영리기관
'24.1월~	심의*	20 PF	심의*	(유상)영리기관

\* 자원 유희상황, 사전조사 등을 토대로 자원조정심의를 통해 적정규모·기간 확정

## 전략2. 초거대AI 혁신 생태계 조성

### 1 초거대 AI 응용서비스 선도 프로젝트 추진

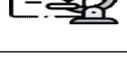
- ◇ AI의 학습능력을 민간 전문영역과 공공분야에 적용하여 **업무 효율성** 향상 및 **초거대AI 기반 新비즈니스** 창출 가능 ⇒ 초거대AI 선도 프로젝트 신설

#### □ 민간 5대 전문영역 초거대AI 플래그십 프로젝트 추진 **신규**

- 초거대AI를 전문영역에 접목하여 전문가 보조·지원 등 생산성을 혁신하는 플래그십 프로젝트를 기획, 선제적 개발·적용 추진('24~)
- 초거대AI 기업 및 도메인 전문가 설문조사 등을 통해 **법률·의료** 등 5대 전문영역 **발굴**('23.2~3) → 세부과제 기획, 예산반영 추진('23.4~)

※ 변호사·회계사·의사·교수 등 전문영역 종사자와 함께 서비스를 발굴·기획하고, 25개 주요 SI기업 인터뷰를 통해 과제 구체화

#### < 초거대AI 5대 플래그십 프로젝트(안) >

분야	주요 서비스 내용	활용 데이터
 법률·세무	<ul style="list-style-type: none"> <li>• (서류업무) 소장, 의견서, 변호제안서, 계약서 등 작성·검토</li> <li>• (소송대응) 유사사건 판례·법조문 제시, 승소율 예측 등</li> </ul>	<ul style="list-style-type: none"> <li>·소장·의견서(법무법인)</li> <li>·판결문(법원)</li> <li>·계약서(회계법인)</li> </ul>
 의료보조	<ul style="list-style-type: none"> <li>• (진료업무) 건강검진·진료 소견서 작성, 과거 진료 기록 요약, 치료 방법·처방전 등 추천</li> <li>• (의료컨설팅) 개인건강관리, 반려동물 질병 상담 등</li> </ul>	<ul style="list-style-type: none"> <li>·진료기록서(병원)</li> <li>·임상정보(uptodate)</li> <li>·전문지식(한의학연구원)</li> </ul>
 심리상담	<ul style="list-style-type: none"> <li>• (심리상담) 육아고민 및 자녀 이상행동 상담, 학생 학업·교우관계 상담, 군인·수감자 등 특수환경 정서진단</li> <li>• (상담보조) 상담 분석리포트 작성, 상담치료 방향 제시</li> </ul>	<ul style="list-style-type: none"> <li>·육아컨설팅, 심리 치료 데이터(컨설팅기관)</li> <li>·상담 데이터(상담센터)</li> </ul>
 문화·예술	<ul style="list-style-type: none"> <li>• (방송미디어·예술) 원고 작성, 방송편집, 디자인·음악 창작 등</li> <li>• (한류확산) 외국인 대상 초거대AI 한국어 교육 서비스 등</li> </ul>	<ul style="list-style-type: none"> <li>·방송대본·편집영상, 포스터 등(방송사 제작사)</li> <li>·서적·오디오북(출판사)</li> </ul>
 학술·연구	<ul style="list-style-type: none"> <li>• (학술검색) 선행 논문·특허 등 참고문헌, 관련 통계 제시</li> <li>• (연구지원) 연구주제·방향 설정, 연구보고서 요약 지원</li> </ul>	<ul style="list-style-type: none"> <li>·논문데이터(학술DB)</li> <li>·특허 데이터(특허청)</li> </ul>

- 초거대AI(언어·이미지 이해·생성능력)를 기반으로 전문분야 지식을 학습시켜 수요자가 효과적으로 활용 가능한 전문서비스를 개발('24~)
- 초거대AI 기업 뿐만 아니라 전문분야 기업·기관(수요기관), 데이터 보유기관 등이 사업에 참여(컨소시엄)하는 형태로 사업 추진



□ **공공·행정 초거대AI 서비스 개발·활용**

- **(공공·행정 혁신)** 행정·공공기관 등 수요를 바탕으로 내부 업무, 대민 서비스 등 초거대 AI 선도과제 추진('23~, 관계부처) **강화**

**< 공공분야 초거대AI 선도 과제(예시) >**

구분	주요내용	소관
내부업무 보조	보도자료, 홍보자료 작성, 단순·반복적 보고서 초안 작성 지원	과기정통부 행안부
국민문의 대응	금융상품, 주택연금 등 공공제도에 대한 국민문의 상담 등	
지자체 지원	지역민원 및 주민신고 등 대응, 지역관광·명소 안내 등	지자체
특허 심사	초거대 특허 언어모델 구축 및 검색서비스 개발	특허청

- **(디지털서비스 개발 지원)** 공공서비스 개발 등을 위한 초거대AI 모델 활용(API) 및 컨설팅을 **종합 지원\***('23~)하고, **강화**

\* '민간의 첨단 초거대AI 활용지원', '클라우드 바우처' 사업을 통해 약 150개 기업 지원

**< 초거대 AI 활용 종합지원 주요 지원내용(안) >**

구분	주요내용
심층 컨설팅	AlaaS 기획·개발, 기존 서비스 고도화, 데이터학습 등 컨설팅
이용료 지원	AlaaS 개발·학습에 필요한 초거대AI 모델 API 이용료 지원
전문교육	초거대 AI 개념, 활용방법, 실습 등 단계별 교육 프로그램 제공
세미나	초거대 AI 스타트업 서밋, 초거대 AI 생태계 매칭데이 등

- 디지털플랫폼정부 테스트베드 등을 통해 초거대 AI를 적용한 공공서비스 등을 시험·검증('23~) **신규**

\* 디지털플랫폼정부 테스트베드에 초거대 (생성) AI 기반 소스코드 자동생성 기능도 제공 검토

## 참고 초거대AI 5대 플래그십 프로젝트 발굴 현황 및 계획

- ◆ 초거대AI 5대 분야 「플래그십 프로젝트」는 ① 초거대AI 우선 적용분야 도출 → ② 분야별 프로젝트 기획 절차를 통해 추진

### □ 초거대AI 우선 적용분야 도출

- 각국 해외사례 및 연구보고서 분석 등을 통해 15개 후보분야 발굴('23.2~3월)
- 분야별 전문가 자문회의를 통해 검토대상 10개 분야 확정 및 평가기준\*(2개 항목, 6개 지표) 설정('23.3월)

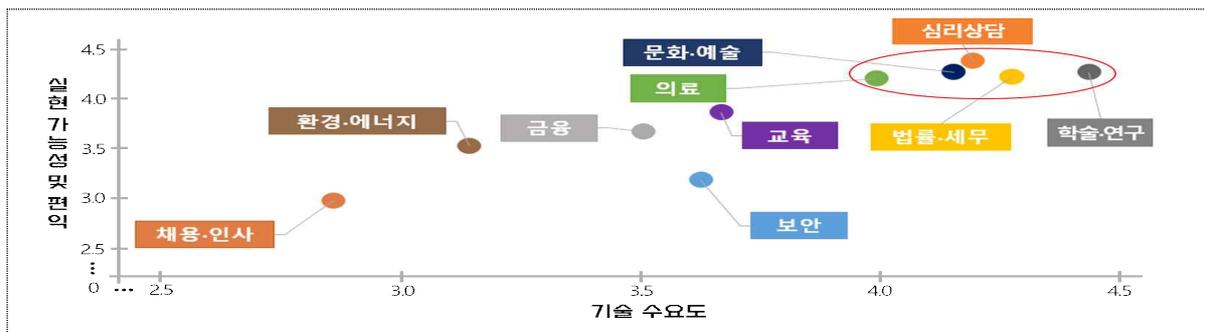
\* 수요도(필요성·시급성), 실현가능성·편익(가능성·성숙도, 파급효과, 사회적 효용성 등)

#### < 초거대AI 우선적용 후보분야 >

▲심리상담 ▲교육 ▲금융 ▲법률·세무 ▲보안 ▲의료 ▲문화·예술 ▲채용·인사  
▲학술·연구 ▲환경·에너지 (10개) / ▲고객응대 ▲언론 ▲공공 ▲제조 ▲마케팅 (15개)

- 델파이 분석\*을 통해 도메인 별 '기술 수요도' + '실현가능성 및 편익'이 높은 5개 분야 도출(정보통신산업진흥원)

\* ICT 컨설팅 기업, 초거대AI 기업, 대학교수, 연구기관 등 민간전문가 평가



### □ 분야별 초거대AI 기반 응용서비스 기획

- 도메인·AI 전문가(SW산업협회 산하 초거대AI 협의회 등)가 참여하는 '과제 기획위원회' 운영을 본격화하여 후보과제 확정, 재정당국과 예산 협의('23.4~)

#### < 5대 전문영역 초거대AI 플래그십 프로젝트(예시) >

법률	의료	심리상담	문화·예술	학술·연구
소장 작성 지원 판례·판결문 제시	소견서 작성 지원 치료방법 추천	육아고민 상담 심리분석 리포트	한류 대응 외국인 한국어 교육	참고문헌 통계 제시 초록 작성 보조

## 2 초거대AI 기반 디지털산업 혁신 생태계 조성

◇ 향후 초거대AI를 통해 AI 서비스의 성능·완성도가 획기적으로 향상될 전망  
 ⇒ 초거대AI 기반 AI 서비스(SaaS, 메타버스)가 활성화되는 **민간 중심 생태계 조성**

현재	+	개선
기존SW의 SaaS 전환·고도화, SaaS 개발환경·기술교육 등 일부 지원		초거대AI 기반 SaaS·메타버스 서비스 개발, 민간 중심 초거대AI 협의회 구성

### □ 초거대 AI 기반 SaaS 개발 및 이용 확산 **신규**

- **(SaaS)** 중소 SaaS 기업이 초거대AI 기반의 혁신적 AI 서비스를 개발할 수 있도록 초거대AI 모델, 컴퓨팅 자원 등을 종합 지원('23~)
  - 중소 SaaS 기업의 서비스 개발을 지원하는 **별도 트랙\***을 신설하고, 이용 가능한 초거대AI 기업, 클라우드사업자 리스트를 **메뉴판** 형태로 제공
  - \* '23년 4개 과제(총 20억원, 과제당 5억원)를 우선 시범지원하고, '24년 확대 추진
  - 초거대AI 기반 SaaS 등을 개발한 유망 SW 기업에 대해 **글로벌 SaaS 마켓 플레이스** 진입 및 마케팅 등 종합 지원('24~)

### □ 초거대AI 기반 메타버스 생태계 활성화 **신규**

- **(메타버스)** 메타버스 공간에서 초거대AI를 지능형 비서 등으로 시각화 서비스하는 초거대AI+메타버스 융합 프로젝트 추진('24~, 5개 과제)

지능형 교사	개인 비서	쇼핑 도우미
학생·교사와 대화형 교과학습을 지원하는 디지털 휴먼 교사	미팅·스케줄관리, 숙박·식당 예약 등을 수행하는 AI 비서	상품정보 검색·추천, 결제·배송 등 O2O 쇼핑 지원

### □ 초거대 AI 협의회 구성·운영('23~) **신규**

- 한국소프트웨어산업협회 중심으로 민간 차원의 투자, 新서비스 창출 등을 협력하기 위한 “초거대AI 협의회\*” 구성
  - \* 초거대AI 기업, 클라우드 기업, 중소 AI 개발사, 협·단체 등
- 주기적으로 기술 교류회, 기업간 해외 동반진출 등 협력방안 논의

### 3 초거대AI 전문인재 양성 및 전국민 활용역량 강화

◇ 초거대AI 발전도 결국 사람의 역량 향상이 좌우 ⇒ 기존 AI 인력 양성 사업에 초거대AI 개발·활용 경험을 확대하고 전국민 초거대AI 리터러시 강화

현재	+	개선
개별 과정을 통해 수준별 <b>SW·AI 인재양성 추진('23~'27)</b>	+	<b>초거대AI 전문인재 신규 양성,                      초거대AI 활용역량 강화('23~'27)</b>
AI 석·박사 등 정규과정(6.5만명) 구직자·재직자 등 비정규과정(13.2만명)		초거대AI 전문인재 양성 및 초거대AI 활용역량 교육 보강, 일반국민 AI 역량 교육(100만명)

\* 現 SW·AI교육 과정에 초거대AI 활용역량 교육 등을 추가·보완해 운영

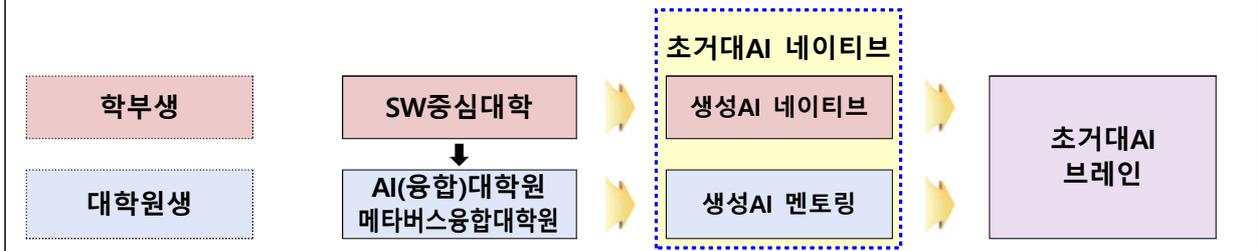
#### □ 초거대AI 전문인재 양성 신규

- **(초거대AI 전문인재)** 산·학 협력 프로젝트, 해외 공동연구 등을 통해 글로벌 수준의 초거대AI 연구개발, 활용 역량을 갖춘 중·고급 인재 양성('24~)
  - SW중심대학, AI대학원 등과 연계하여 '초거대AI 네이티브', '초거대AI 브레인' 등 지원('24~)

#### < 초거대AI 선도 인재양성 프로젝트(안) >

- ① **(초거대AI 네이티브)** 초거대AI 기업과 AI·SW 전공자가 함께 산학협력 공동 프로젝트, 문제해결 멘토링을 통한 프롬프트 엔지니어링 역량 제고로 기업이 원하는 초거대AI 중·고급 인재로 성장 지원
- ② **(초거대AI 브레인)** 국내 대학원생을 대상으로 국내·외 생성AI 빅테크 기업의 Top-Tier 연구자들과 초거대AI 공동연구를 지원하여 연구개발 역량을 갖춘 엘리트 AI인재 양성

< 운영 개념도 예시 >



#### □ 분야별 초거대AI 활용 역량 강화 강화

- **(SW)** 초거대AI 기반 AI 응용서비스 개발·실습(파이썬 프로그램 제작 등) 등 현업 수요 기반 교육 프로그램 운영('23~)

\* ICT 이노베이션 스퀘어('23년 7,310명 387억원), 산업전문인력 AI역량강화('23년 5,040명 167억원)

- 전공자 및 중·고급 교육과정\*에서 코드 분석, 에러 검출·정정 등 SW개발 도구로서 효율성을 높일 수 있는 AI 활용 교육 실시

\* 이노베이션 아카데미('23년 600명, 331억원), SW마에스트로('23년 230명, 130억원) 등

※ 초거대AI가 최상의 결과물을 낼 수 있도록 지시·대화 역량 강화 교육과정 신설 검토

- 구직자 등 대상으로 MS 등 초거대AI 선도기업들이 설계·운영하는 초거대AI 프로젝트 중심 실무인재 양성과정 발굴·확대('23~, 고용부)

- **(교육)** 정보교육 전담교원 대상 생성AI 활용 연수를 기획·운영\*하고, 교원 및 교·사대 학생 등 대상 '생성AI 활용 수업 경진대회' 추진('23~)

\* 2023 하계 연수시 커리큘럼으로 적용 예정('23.8월, 교원 120명 내외)

- 생성 AI 원리 이해와 활용을 위한 AI교육 플랫폼\* 기반의 학습콘텐츠 개발 및 AI교육 선도학교 대상 시범 수업 후 일반학교로 단계적 확산('24~)

\* 부처간 협업사업: [교육부] 생성AI 등 교육콘텐츠 개발 + [과기정통부] AI교육 플랫폼 운영

- 초중등 학생을 대상으로 생성 AI의 생산적 활용방안은 물론, 데이터 편향·신뢰성 문제 등 윤리적 이슈에 대한 기초 교육도 실시

## □ 초거대AI 기본 소양 강화 강화

- **(일반국민)** 전국 1,000여개 디지털배움터를 통해 챗GPT 소개 및 활용방법, 유의사항(편향성·답변오류) 등 단계별 초거대AI 교육 프로그램 운영('23~)

< 디지털배움터 초거대AI 교육 운영(안) >

과정	교육대상	내용	목표('23년)
기초	노령층 등	·챗GPT 등 초거대AI 사용방법, 신뢰성 교육 등	연 90,000명 교육
생활	학생 등	·오피스SW 연계활용(코파일럿), 프롬프트 작성 등	연 8,000명 교육
심화	학생재직자 등	·영상제작, SW 코딩 등 전문업무 활용 방법	연 2,000명 교육

※ '23년 전국민 대상 시범교육(10만 명) 이후 초거대 AI 교육 확대('24~'27년 90만 명 목표)

▶ 석박사, 재직자, 학생 등을 대상으로 AI·SW 인재양성 기 추진 중

- AI 대학원, AI융합혁신대학원 등을 통해 세계적 R&D 역량을 갖춘 석·박사급 연구자를 양성하고, SW 중심대학 확대 등을 통해 AI·SW 융합 인재 양성
- 재직자·구직자·학생·군인 등을 대상으로 인공지능 개발·활용 역량 강화

### 전략3. 범국가 AI 혁신 제도·문화 정착

#### 1 초거대AI 규제개선 및 제도정립 추진

◇ 초거대AI 규제개선 과제 발굴 및 개선방향 도출을 통한 산업육성 기반 마련

□ 초거대AI 개발·활용 촉진을 위한 규제개선·제도정립 과제 추진 **강화**

○ 규제개선 주제별로 인공지능 법제정비단을 운영하여 관계부처와 함께 법제 이슈 정리 및 개선방향 도출 지원('23~)

※ 관계부처에서 별도 운영 중인 제도개선 T/F가 있는 경우 이를 활용

< 초거대AI 규제개선·제도정립 방향(예시) >

주제	논의 방향	관계부처
개인정보	<p>활용 가치가 높은 데이터를 개인정보 침해 우려 없이 안전하게 AI 학습용데이터로 활용할 수 있는 방안 논의</p> <p>※ (예시) 비정형데이터 가명처리 기준 마련, 재현데이터 활용 확대, 적극적으로 공개한 자기개인정보의 AI 학습용데이터 활용 허용 검토 등</p>	개인 정보위
지식재산권	<p>저작물의 학습용 데이터 활용, 인공지능 산출물 보호 등 기존 지식재산권 관련 법 체계로 해결이 어려운 문제 논의</p> <p>※ 美 저작권청은 인간과 초거대 AI가 공동으로 생성한 산출물에 대하여 인간이 기여한 부분에 대하여 신청서 작성을 통해 저작권을 인정해주는 내용의 가이드라인을 발표('23.3.10)</p> <p>- (예시) 인간(글)과 미드저니 AI(그림)의 공동으로 제작한 그림책에 대하여 인간(글)이 기여한 부분에 대해서만 저작권을 인정</p>	문체부 특허청
개별 산업	<p>개별 산업 분야에서 생성형 인공지능의 도입·활용을 염두에 두지 않은 채 운영되어온 제도* 발굴 및 개선</p> <p>* (예시) 미국, 독일 등의 경우 인공지능을 활용한 보험모집이 가능하나, 우리나라 보험업법(제83조)은 보험모집을 할 수 있는 자에 인공지능 불포함</p>	각 부처
위험성 관련	<p>초거대 AI 제품·서비스 등으로 우려되는 피해나 위험성을 방지하기 위한 사전적 규율체계 마련 및 인공지능으로 유발된 사고에 대한 사용자의 손해배상 입증책임 완화 필요성 등 검토</p> <p>※ 과방위 계류 중인 인공지능법 제정 추진 및 EU의 인공지능 책임 지침 등을 고려한 우리나라의 손해배상 제도 점검</p>	과기 정통부 · 법무부 공정위

## 2 초거대 AI 서비스 신뢰성 제고

- ◇ 초거대AI가 사회·산업 전반으로 확산되기 위해서는 신뢰성 확보가 중요  
 ⇒ 신뢰성·성능 평가 신설로 사회 구성원 우려 해소 및 기업의 부담 완화

기존	신설
일반적 AI 서비스 기획·개발 과정에서 개발자 등이 스스로 활용하는 자율점검표, 개발안내서 제공	개발된 초거대AI 서비스에 대해 제3의 기관을 통한 신뢰성·성능 평가 신규 실시

### □ 신뢰성 · 성능 평가

- 초거대AI 서비스에서 발생 가능한 주요 위험요인, 성능 등에 대해 객관성·공신력이 확보된 제3의 기관을 통한 평가 추진('23~) **신규**
  - \* 기업 자율적으로 평가 참여, 결과 공개 여부 등을 결정토록 하되, 평가 참여 시 초거대AI 출시 과정에서 기업의 책임 부담이 완충 될 수 있는 방향으로 운영
  - **(신뢰성 평가)** 공평성, 투명성, 데이터 출처 적절성, 활용 적합성 등 평가항목\*별 세부 시나리오, 테스트데이터 마련하여 평가 실시
  - \* 「신뢰할 수 있는 인공지능 개발 안내서(과기정통부-TTA, '22.2)」를 토대로, 생성형AI 모델 특성을 반영한 평가항목·요구사항을 도출
  - **(성능 평가)** 일상대화에서부터 법, 정치, 역사, 과학, 예술 등 전문 영역에 이르기까지 질문데이터를 구성하여 답변 성능을 평가

#### < 초거대AI 서비스 신뢰성 평가 주요 내용 >

이슈	위험 요인	평가 내용
공평성	편향·차별 등 비윤리적 콘텐츠 생성	편향 발언 차단·회피 가능 여부
성능 및 투명성	부정확한 사실·상식 답변	답변의 정확성(성능) 및 답변의 출처·근거 제시 가능 여부
데이터(출처) 적절성	저작권 미확보된 콘텐츠 활용	학습 데이터 수집·획득 방법의 적절성 여부(자가 선언)
활용 적합성	잘못된 정보, 피싱 이메일 등의 생성에 의도적으로 활용	악용·오용에 대한 사용자 책임 고지, 서비스 한계 안내 여부 확인

- 비윤리·유해성 표현, 사실 왜곡 등을 검증할 수 있는 데이터셋 구축('23~) **신규**

#### < '23년 윤리·신뢰성 검증 데이터 구축 방안(안) >

- ① (공평성) 성별·지역·종교·인종 등 특정 집단에 대한 차별·편향성 검증 데이터
- ② (성능) 법, 역사, 과학, 시사·상식 등에 대한 성능을 평가하기 위한 질문·답변
- ③ (활용 적합성) 피싱메일 생성, 해킹 등 부적절한 정보 제공을 유도하는 질문

### 3 사회적 변화 대응 및 수용력 향상

◇ 초거대AI가 국민일상, 사회·경제 전반으로 확산되면서 야기되는 다양한 이슈에 대해 민간전문가, 이해관계자와 함께 사회적 공감대 형성

□ 인공지능 확산에 따른 사회·문화적 수용성 향상

- 「디지털 신질서 협의체」(‘23.3), 「인공지능 법제정비단」(‘23.上), 「인공지능 윤리포럼」(‘23.上) 등을 통해 사회·문화적 이슈 발굴 및 논의(‘23~) **강화**
- 초거대AI 관련 “디지털 新질서” 정립을 위한 폭넓은 논의·사회적 협의를 진행할 범정부 논의체계(가칭) 초거대AI 대응반 운영(‘23~) **신규**

< 초거대AI 확산으로 촉발된 주요 이슈 >

주제	논의 사항
 업무적용	·문서작성, SW코딩 등 챗GPT를 현장에 적극 도입해야 한다는 의견과 신뢰성·정확성이 미흡하므로 업무현장에 적용하기는 시기상조라는 의견 상존 ·챗GPT로 인한 단순·반복적 지식업무 등 일자리 감소 우려
 교육·학습	·학생들의 학습도구, 교사들의 수업자료 작성 등을 위해 챗GPT 활용을 허용해야 한다는 의견과 챗GPT 허용은 학생들의 사고력·문제해결력을 감소시키고 연구 투명성 위협 및 표절문제를 야기한다는 의견 상존 * (뉴욕시) 공립학교 내 챗GPT 접근 금지, (사이언스지) 챗GPT 사용 금지
 보안	·공공부문, 금융분야 등 종사자가 챗GPT 활용 과정에서 보안사항, 고객 정보, 소스코드 등이 유출될 가능성 존재, 사전·사후적 리스크 관리 필요 * JP모건(금융), 버라이즌(이동통신)은 직원들의 챗GPT 사용 금지
 공정경쟁	·초거대AI를 보유한 일부 빅테크 기업의 독과점과 중소기업 등의 의존도가 심화되고, ‘데이터 쓸림’ 현상도 가중될 우려
 범죄악용	·초거대AI의 콘텐츠 생성기능이 음란물(딥페이크) 및 가짜뉴스 생성, 스팸 메일 및 악성코드 생성, 피싱사이트 로그인 페이지 제작 등에 악용될 우려

□ 초거대AI 활용 가이드라인 마련 **신규**

- 초거대AI를 교육·학습 등에 활용하기 위한 방안을 연구하고, 이를 반영하여 ‘교육 분야 인공지능 윤리원칙’ 수정·보완(‘23, 교육부)
- 공무원 대상 초거대AI 공공부문 활용 가이드라인 마련(‘23, 행안부)

## VII. 향후 계획

사회적 변화 대응 및 수용력 향상을 위한 추진체계 구성 : '23.4월~

※ ① 인공지능 윤리포럼, ② 인공지능 법제정비단, ③ 초거대AI 대응반

신규 프로젝트 세부기획 및 예산 반영 협의 : '23.4월~

전국민 인공지능 일상화 추진계획 발표 : '23.6월