

생성형 AI와 경쟁

GENERATIVE AI AND COMPETITION

2024. 12.



공정거래위원회



본 저작물은 공공누리 제2유형으로 개방한 것으로 이용자는 공공저작물을 이용하는 경우 출처 또는 저작권자를 표시하여야 하며, 상업적 이용이 금지되어 있으므로 영리행위와 직접 또는 간접으로 관련된 행위를 위하여 이용될 수 없습니다. 구체적인 저작물 이용 조건은 공공누리 홈페이지 (<https://www.kogl.or.kr/info/license.do>)에서 내용을 확인할 수 있습니다.

본 보고서는 생성형 AI 시장구조 및 경쟁 상황 등에 대한 이해를 돕는 목적 등으로 작성되었으며, 향후 공정거래위원회가 소관 법령에 따라 의결하는 사건 또는 관련 정책의 수립 등 공정거래위원회의 공식적 의견에 영향을 미치지 않습니다.

<목 차>

발 간 사	1
1 도입	3
1.1 발간의 목적	3
1.2 작성 범위 및 한계	4
1.3 국내외 동향	6
2 생성형 AI 연관 산업 현황	12
2.1 AI와 생성형 AI	12
2.1.1 AI의 개념	12
2.1.2 AI의 발전 단계	13
2.1.3 생성형 AI의 개념과 의의	15
2.1.4 생성형 AI의 유형	17
2.2 생성형 AI의 구축과 실행	19
2.2.1 구축·실행 제반 과정	19
2.2.2 생성형 AI의 가치사슬 현황	20
2.2.2.1 AI 반도체	20
2.2.2.2 클라우드 컴퓨팅	24
2.2.2.3 데이터	27

2.2.2.4	전문인력	31
2.2.2.5	기반모델	33
2.2.2.6	AI 서비스	38
2.3	국내 생성형 AI의 시장구조 및 현황	39
2.3.1	가치사슬 관련 시장 구조	39
2.3.2	시장 경쟁상황 분석	42
2.3.2.1	AI 반도체	42
2.3.2.2	클라우드 컴퓨팅	49
2.3.2.3	기반모델	56
2.3.2.4	AI 서비스	61
3	생성형 AI 산업에서의 경쟁·소비자 쟁점 및 과제	63
3.1	시장경쟁에 영향을 미치는 구조적 요인	63
3.1.1	자본·기술 집약적 산업에 따른 진입장벽	63
3.1.2	규모의 경제, 범위의 경제 등에 따른 대규모 사업자의 우위	65
3.1.3	시장 선점 효과에 따른 선도적 사업자의 우위	67
3.2	경쟁 저해 및 소비자이익 저해 우려	68
3.2.1	단독행위	68
3.2.1.1	필수 요소에 대한 접근 제한	68
3.2.1.2	결합 판매	71

3.2.1.3	고객 유인 및 고객의 이탈 방지	74
3.2.1.4	기술 부당 이용	76
3.2.1.5	인력의 유인·채용	77
3.2.2	기업결합	78
3.2.3	소비자 이익 저해	80
4	결론	84
	<참고 문헌>	86

발 간 사

공정하고 자유로운 경쟁은 창의적인 기업활동을 조성하고 소비자를 보호하며, 나아가 국민경제의 균형있는 발전을 가능하게 합니다. 빠른 속도로 발전하는 인공지능(AI) 시장에서도 건전한 경쟁질서 마련이 시장에 참여하는 사업자의 혁신 역량을 고취시키고, 소비자의 편익을 제고할 것이라는 점에는 이견이 없을 것입니다.

공정위는 역동적인 AI 시장의 경쟁상황을 분석하고 경쟁법적 위험성을 미연에 방지하는 한편, 시장참여자들의 예측가능성을 제고하기 위해 「생성형 AI와 경쟁」이라는 정책보고서를 발간하게 되었습니다. 이번 정책보고서를 발간하기 위하여 「AI 시장 실태조사 연구」 용역(책임연구원: 황태희 교수)을 추진한 한국경쟁법학회와 긴밀하게 논의하는 한편, 국내외 사업자들과 수차례의 간담회를 실시하는 등 이해관계자와도 적극 소통하였습니다.

이 보고서는 오픈AI의 챗GPT(Chat GPT) 등장과 함께 폭발적인 관심을 받고 있는 생성형 AI(Generative AI)를 중심으로, AI 반도체·클라우드 컴퓨팅 등 핵심 인프라와 기반모델, AI 서비스 등 가치사슬 내 세부 시장의 경쟁 상황을 분석하고, 경쟁·소비자 이슈 등을 점검하였습니다.

현재 AI 생태계 내 각 분야별 시장에서는 글로벌 빅테크뿐만 아니라 국내 대기업, 다수의 스타트업이 AI 시장에서 경쟁우위를 차지하기 위해 치열한 각축전을 벌이고 있으며, 이는 기술 경쟁, 자본·인재 유치, 인접 시장 수직통합, 사업자 간 파트너십 등 다양한 형태를 통해 나타나고 있습니다.

그러나, 활발한 경쟁 속에서 이미 일부 시장에서는 지배력을 가진 소수의 사업자가 등장하고 있는 것도 사실이며, 복수의 해외 경쟁당국은 시장지배력을 가진 사업자의 남용 행위 등 경쟁제한 행위가 발생할 가능성에 대한 우려를 표명하고 있는 상황입니다.

시장지배적 사업자의 등장 자체가 문제라고 보지는 않으나, 규모의 경제, 네트워크 효과 등이 나타나는 AI 시장의 특성상 지배력을 가진 사업자의 행위는 시장에 미치는 영향이 더욱 클 수 있습니다. 따라서, 저는 지배력을 인접 시장으로 전이하거나, 필수적 요소에 대한 접근을 제한하는 등 경쟁제한행위가 나타나는지 여부에 대해서는 면밀히 모니터링할 필요가 있다고 생각합니다.

또한, 공정위는 급격하게 변화하고 있는 AI 시장에서 나타나는 행위를 사전적으로 혹은 사후적으로 점검하는 제도개선 필요성도 검토할 계획입니다. 예를 들어, 최근 AI 시장 내 파트너십이 경쟁을 저해할 수 있다는 우려에 대해 현행 기업결합 심사제도의 개선을 검토하고, 시장지배적 지위를 가진 AI 개발사들의 데이터 수집·활용 행위가 소비자의 이익을 저해할 우려가 있다는 점에 대해서는 관련 규정의 정비 필요성 등을 검토할 예정입니다.

시장의 경쟁질서를 해치는 반칙행위에 대한 모니터링을 강화하고, 경쟁제한적 우려가 있는 행위를 점검하는 제도를 개선함으로써 AI 생태계 안에 경쟁질서가 확립될 수 있도록 노력하겠습니다. 이는 AI 시장 내 품질에 의한 활발한 경쟁을 촉진시켜 기업의 혁신 정신을 고취하고, 소비자의 후생을 제고할 것입니다. 이 혜택이 결국 국민 전체에게 돌아갈 것이라고 저는 믿습니다.

아울러, 이번 보고서가 국민들께서 혁신성장의 동력으로 새롭게 떠오르고 있는 생성형 AI 시장구조 등을 이해하는 데 많은 도움이 되기를 희망합니다. 감사합니다.

2024년 12월

공정거래위원회 위원장 **관 기 정**

1 도입

1.1 발간의 목적

1. 공정거래위원회(이하 ‘공정위’)는 신기술을 중심으로 빠르게 변화하는 시장구조에 적시 대응하고, 경쟁·소비자 이슈를 선제적으로 분석하여 시장참여자들의 예측가능성 등을 도모하기 위하여 정책보고서 주제로 인공지능(Artificial Intelligence, 이하 ‘AI’)을 선정하였다.

2. 특히, AI의 다양한 범위 중에서 향후 AI 관련 시장 및 산업 재편 등에 핵심적인 역할을 수행할 것으로 기대되는 ‘생성형 AI(Generative AI)’를 중점적으로 검토·분석하였다. 최근 전세계적으로 생성형 AI에 대한 관심과 수요가 높아지고 있고, 생성형 AI 관련 시장도 형성되어 가고 있기 때문에 시장연구(Market Study) 등을 통한 분석 필요성이 매우 높아졌다.

3. 또한, 생성형 AI 시장에서는 수많은 사업자들이 시장경쟁에서 우위를 차지하기 위하여 노력 중이나, 경쟁 과정에서 발생할 수 있는 진입장벽 구축·거래거절·경쟁사업자 배제 등 경쟁저해 우려도 나타나고 있다. 이에 대해 이해관계자는 물론 해외 경쟁당국의 관심도 높아지고 있으며, 일부 해외 경쟁당국은 이미 생성형 AI 관련 시장조사 또는 경쟁제한 행위 여부에 대한 조사 등을 진행하고 있다. 국내 생성형 AI 관련시장도 글로벌 사업자 및 국내 대기업, 스타트업 등 많은 사업자들이 진출하여 경쟁하고 있다는 점 등을 고려할 때 국내 생성형 AI 시장의 경쟁 상황 등을 살펴볼 필요가 있다.

4. 이에 따라 이 정책보고서에서 국내 생성형 AI 관련 시장구조, 경쟁 상황 및 경쟁·소비자 쟁점 등을 검토·분석함으로써, 시장에 대한 이해도를 높이고 공정한 경쟁질서 확립을 위해 공정위가 집중해야 할 역할을 제시하고자 한다.

1.2 작성 범위 및 한계

5. 이번 정책보고서를 작성하기 위하여 AI 시장에 참여하고 있는 회사를 대상으로 한 실태조사(Market Study), AI 회사를 비롯한 학계·전문가 의견 청취, 문헌조사 및 검토 등을 진행하였고, 그 결과를 종합·분석하여 본 보고서에 반영하였다.

6. 실태조사는 「독점규제 및 공정거래에 관한 법률」(이하 ‘공정거래법’) 제87조에 근거하여 서면으로 실시하였다.¹ 조사 대상 사업자는 이해관계자 등과 실무 면담을 통해 국내·외에서 AI 사업을 영위하는 주요 회사 및 AI 관련 협회·민관 협의체 소속 회사 등을 포함하여 50개로 선정하였으며, 주요 조사내용은 생성형 AI 관련 사업 현황, 관련시장의 경쟁 상황에 대한 견해 및 경쟁 저해·불공정거래 경험 여부 등에 관한 것이었다.

7. 실태조사 외에도 국내·외 AI 사업을 영위하고 있는 주요 회사와 인터뷰를 진행하여 산업계 의견을 청취하였으며, 학계·법조계·연구기관 등 외부 전문가 간담회와 자문을 실시하는 등 이해관계자 의견을 충분하고 객관적으로 반영하고자 하였다.

8. 또한, 국내·외 사업자 및 공공·민간기관 등의 AI 산업 보고서, 주요 해외 경쟁당국의 AI 시장 연구(Market Study) 보고서 및 조사·조치 관련 보도자료 등 다양한 문헌 자료를 참고하여 보고서의 완결성 등을 제고하고자 하였다.

9. 이하 정책보고서 2. 목차에서는 생성형 AI 관련 가치사슬² 및 시장 현황을 분석하고, 3. 목차에서는 관련 산업 내 경쟁·소비자 관련 쟁점 등을 검토한 후 4. 목차에서 결론을 기술하였다.

1 이하 본 보고서상 ‘AI 실태조사’ 라 함

2 가치사슬은 마이클 포터(Michael Porter)가 1985년 정립한 개념으로서, 제품의 생산과 판매, 홍보와 사후서비스 등 사슬처럼 연결된 모든 단계에서 부가가치가 만들어지는 것을 말함

10. 구체적으로, 다음 1.3 목차에서 국내외 동향을 살펴보고, 2. 목차에서 AI와 생성형 AI의 비교를 통해 생성형 AI가 구축·구현되기 위하여 필요한 가치사슬을 정의하고 각 가치사슬 분야별 경쟁상황을 분석하였으며, 3. 목차에서 관련시장에서 발생 가능한 구조적·행태적 경쟁 저해 요인 및 소비자 권익 침해 등을 살펴봄으로써 4. 목차에서 생성형 AI 생태계에서의 지속적인 혁신과 공정한 거래질서 확립을 위한 경쟁정책 방향, 경쟁당국의 역할 등을 제시하였다.

11. 다만, 이번 정책보고서는 다음과 같은 한계가 존재함을 밝힌다. 먼저, 생성형 AI 관련 시장은 2022년 11월 챗GPT의 공개 이후 기반모델 출시 기준으로 크게 성장하고는 있지만 현재 초기 단계에 머무르고 있으며, 각 분야별로 유·무료 제품 또는 서비스가 혼재되어 있어 시장규모 및 거래규모를 명확히 파악하는 데 한계가 있다.

12. 또한, AI 실태조사 대상 회사들이 참여하고 있는 분야 관련 매출액이 아직 미발생하였거나 미미하여 생성형 AI 관련 산업에서 주요 회사들의 시장점유율을 추정하기 어렵고, 역동적인 변화가 발생하는 시장 특성상 관련시장 획정 자체도 어렵다는 점에서 계량적 분석에도 한계가 존재한다.

13. 따라서, 이번 보고서에서는 생성형 AI 인프라 구축, 개발, 그리고 서비스 구현으로 이어지는 가치사슬 전반을 살펴 주요 사업자 위주로 시장 경쟁 상황을 서술하였으며, 상대적으로 관련 시장이 더 형성되었거나 발전한 분야를 중점적으로 검토·분석하였다.

14. 아울러, 본 보고서는 생성형 AI의 구축과 구현, 시장 경쟁상황 및 경쟁·소비자 쟁점 등을 검토·분석함으로써 향후 경쟁정책 방향을 제시하기 위한 것으로서, 어느 한 사업자의 공정거래법 위반 여부 등을 사전에 판단하지 않고 향후 공정위의 조사 개시 여부 및 위법성 판단 등에 대한 구속력도 없다는 점을 명확하게 밝혀둔다.

1.3 국내외 동향

15. 우리나라에서는 생성형 AI에 관한 주요 이슈에 대하여 공정위는 물론, 과학기술정보통신부, 산업통상자원부, 문화체육관광부, 개인정보보호위원회, 방송통신위원회 등 여러 부처가 담당하고 있다.

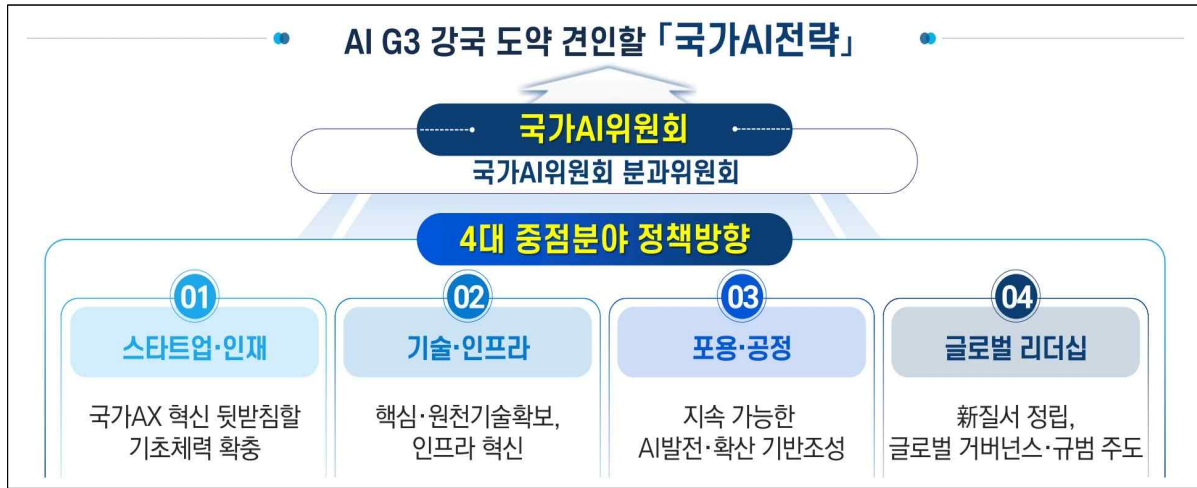
16. 최근 AI 기본법 제정 요구가 지속되고 있고, 2019년부터 관련 법안이 지속 발의되고 있으나 2024년 11월 현재 제정된 법률은 없는 상황이다. 현재까지 국회에 발의된 법안들은 주로 인공지능 산업의 육성과 신뢰·책임 확보에 관한 내용 위주인 것으로 확인된다.

17. 다만, 그간 발의된 법안을 병합하여 대안으로 발의된 「인공지능 발전과 신뢰 기반 조성 등에 관한 기본법안」이 2024년 11월 국회 과학기술정보방송통신위원회에서 통과되었다. 해당 법안은 인공지능 기본계획의 수립, 국가 인공지능위원회 설치, 인공지능 관련 시책 추진 등 정부의 역할을 규정하고 있고, 민간에 대해서는 인공지능 사업자의 인공지능 이용 사실 고지·표시 의무, 안전성·신뢰성 확보 조치 의무 등을 규정하고 있다.

18. AI 기술이 빠르게 진화하고, 전 세계적으로 AI 혁신경쟁이 심화되는 상황에서 인프라 등에 대한 정부의 정책적 지원을 위해 과학기술정보통신부를 중심으로 「AI-반도체 이니셔티브», 「클라우드 컴퓨팅 기본계획」 등이 수립되었으며, 민·관 AI 역량을 결집하기 위하여 2024년 9월 「국가인공지능위원회(위원장 : 대통령)」가 출범되고 「국가 AI전략 정책방향」이 발표되었다.

19. 특히, 「국가 AI전략 정책방향」 4대 분야 정책 방향 중 ③포용·공정에 AI 시장의 공정경쟁 질서 확립 내용이 반영되었다. 이는 AI 혁신을 뒷받침하기 위해 경쟁·소비자 친화적 AI 시장 생태계 조성 필요성 및 중요성에 대한 정부의 의지가 반영된 것으로 볼 수 있다.

표 1. 국가 AI 전략 정책방향



* 출처: 과학기술정보통신부 보도자료³

20. AI 기술 발전에 따라 발생할 수 있는 위협 요소에 대응하여 AI 안전 확보를 위한 글로벌 거버넌스 구축이 강조되고 있으며, 2023년 영국이 개최한 제1차 「AI 안전성 정상회의(AI Safety Summit)」에 이어 2024년 5월 우리나라에서 제2차 「AI 서울 정상회의(AI Seoul Summit)」가 개최되기도 하였다.

21. 이번 회의에서는 지난 제1차 안전성 정상회의에서 발표된 장관급 합의 ‘블레츨리 선언(Bletchley declaration)’을 격상하여, 미국·영국·유럽연합(EU) 등 정상 간 합의로 AI 안전과 더불어 혁신·포용의 내용을 포함한 ‘서울 선언’을 채택하였다. 또한, 국내·외 14개 주요 AI 기업들이 AI 책임·발전·혜택 공유와 관련된 자발적 약속인 ‘AI 서울 기업 서약’을 발표하였다.

22. 향후 AI 정상회의는 글로벌 거버넌스를 보다 강화할 수 있는 플랫폼이 될 것으로 기대되며, 제3차 AI 정상회의는 오는 2025년 2월 프랑스에서 개최될 예정이다.

3 과학기술정보통신부(2024), 인공지능 3대 강국(AI G3) 도약을 위한 대한민국 AI 혁신의 청사진 제시 보도자료

23. 미국에서는 연방 차원의 포괄적인 AI에 관한 법률은 제정되지 않았으나, 2023년 10월 조 바이든 대통령은 ‘안전하고 보안이 보장되며 신뢰할 수 있는 AI의 개발 및 사용에 대한 행정명령’⁴을 발표하였다.

24. 행정명령은 AI 안전 및 보안을 위한 기준을 제시하고, 프라이버시 보호 등 분야별 조치에 대한 내용을 담고 있으며, 특히 ‘혁신과 경쟁’에서는 공정하고 개방적이며 경쟁적인 AI 생태계 촉진을 위해 FTC의 권한 행사를 장려하는 내용이 반영되었다.

25. 다만, 최근 치러진 미국 대통령 선거에서 승리한 도널드 트럼프 당선인이 바이든 정부의 AI 행정명령 폐지를 공약으로 발표하여 향후 트럼프 정부 출범에 따른 미국 AI 정책의 변화가 예상된다.

26. 한편, 유럽연합(European Union, 이하 ‘EU’)은 지난 2021년 발의되었던 세계 최초의 인공지능 규제법 ‘인공지능법(AI Act)’⁵을 2024년 5월에 최종 승인하였으며, 이 법은 2024년 8월부터 발효되어 단계적으로 시행될 예정이다. EU 지역에서 AI 시스템·서비스를 제공·사용하는 사업자를 대상으로 적용되며, AI를 위험 수준별로 4가지 등급(허용 불가 위험, 고위험, 투명성 위험, 최소 위험)으로 구분하여 차등 규제할 것을 명시하고 있다.

27. EU 인공지능법 중 중 경쟁법 관련 내용으로는 각 회원국당 하나 이상의 통고당국(Notifying Authority) 및 시장감시당국(Market Surveillance Authority)을 지정하거나 신설할 것을 규정하고 있으며, 시장감시당국이 잠재적 경쟁법 적용 관련 정보를 EU 집행위와 경쟁당국에 보고할 것을 규정하고 있다.

4 The White House(2023), Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence

5 <http://eur-lex.europa.eu/eli/reg/2024/1689/oj>

28. 또한, 주요 경쟁당국들은 생성형 AI 시장상황 등에 대하여 공동성명서를 발표하거나 시장조사(Market Study)를 진행하기도 하는 등 생성형 AI가 초래할 수 있는 경쟁 문제 등에 대하여 높은 관심을 나타내고 있다.

29. 2024년 7월 미국 법무부(Department Of Justice, 이하 ‘DOJ’)와 연방거래위원회(Federal Trade Commission, 이하 ‘FTC’), EU 집행위원회(European Commission, 이하 ‘EC’), 영국 경쟁당국(Competition & Markets Authority, 이하 ‘CMA’)은 AI 시장의 잠재적 경쟁 저해 요인, 경쟁 강화 요인, 경쟁당국 역할 등이 포함된 공동성명⁶을 발표하였다. 구체적으로, 생성형 AI 개발 등에 필요한 요소의 공급, 기존 대규모 사업자들의 상당한 시장지배력 및 시장참여자 간 투자·파트너십 등과 관련된 경쟁 우려를 표명하고, AI 사업자의 부당한 이용자 데이터 사용에 따라 나타날 수 있는 소비자 이익 저해 우려도 표명하였다.

30. 2024년 10월 이탈리아 로마에서 열린 G7(미국, 일본, 영국, 독일, 프랑스, 이탈리아, 캐나다) 경쟁당국 정상회의에서는 2023년 11월에 발표한 「디지털 경쟁에 관한 공동성명」에 포함된 내용 중 AI 관련 내용만을 상세히 발표⁷하였다. 구체적으로, 해당 성명에는 투입요소에 대한 진입장벽, 빅테크의 시장지배적 지위 남용 등 AI 관련 경쟁 이슈와 함께 AI 시장의 지도 원칙과 경쟁당국의 역할이 포함되어 있다. 이러한 주요 경쟁당국 간 공동성명은 각 당국이 AI 시장의 반독점 행위를 주시하고 있다는 신호를 시장에 전달·강조하기 위한 것으로 해석할 수 있다.

31. 생성형 AI 시장조사와 관련해서는, 경제협력개발기구(Organization for Economic ooperation and Development, 이하 ‘OECD’)를 비롯하여 영국, 프랑스, 캐나다 등 주요 경쟁당국이 아래와 같이 보고서를 발간하였다. 각 당국은 보고서를 통해 AI 시장 상황과 시장 내 경쟁 위협 요인을 분석하는 한편, AI 시장의 공정한 질서 확립 필요성 등을 강조하였다.

6 DOJ·FTC·EC·CMA(2024), Joint Statement on Competition in Generative AI Foundation Models and AI Products

7 G7 Competition Authorities and Policymakers’ Summit(2024), Digital Competition Communiqué

표 2. 주요 경쟁당국별 보고서 발간 현황

구분	발간 시기	보고서 명
OECD	2024년 5월	Artificial Intelligence, Data and Competition
영국	2023년 9월	AI Foundation Models Initial Report
	2024년 4월	AI Foundation Models Update paper
프랑스	2024년 6월	On the competitive functioning of the generative artificial intelligence sector
캐나다	2024년 3월	Artificial intelligence and competition
포르투갈	2023년 11월	Competition and Generative Artificial Intelligence

32. 주요 경쟁당국의 보고서 세부 내용을 살펴보면, OECD는 생성형 AI 시장에서 구조적으로 발생할 수 있는 특성 등을 분석하고, 생성형 AI 개발에 필요한 주요 요소에 대한 접근과 관련된 경쟁 우려 등에 대하여 분석하였다. 구체적으로, 생성형 AI 시장 특성상 규모의 경제, 네트워크 효과, 선두 주자 이점 등이 나타날 수 있으며, 경쟁 우려에 관해서는 생성형 AI 개발에 필요한 데이터에 접근하는 것이 곤란해질 경우 시장경쟁에 위협을 초래할 수 있고, 다른 요소인 컴퓨팅 자원을 제공하는 사업자들의 기반모델 등 연관 서비스 결합 판매에 따른 경쟁 우려 등도 있을 수 있다고 분석하였다.

33. 다음으로, CMA는 2023년 발간한 보고서에서 기반모델을 중심으로 이와 연관된 AI 인프라, AI 개발, AI 배포 단계를 분석하였고, 향후 시장 상황이 데이터·컴퓨팅 자원에의 접근성, 경쟁력 있는 오픈소스 모델의 존재 여부 등에 달려 있다고 전망하였다. 즉, 자체 데이터 및 컴퓨팅 자원 등에 대한 접근이 기반모델 개발과 성능 향상 등 경쟁에 주요한 영향을 미칠 수 있다고 보았으며, 경쟁력 있는 기반모델이 오픈소스로 계속 제공될 경우 사업자들의 시장 진입 및 혁신을 촉진할 뿐만 아니라 수요자의 기반모델 선택권도 제공된다고 보았다.

34. CMA는 2024년 보고서에서는 2023년 보고서 발간 이후 생성형 AI 시장 변화양상 등을 분석하면서, 한편으로는 시장경쟁에 부정적 영향을 미칠 수 있는 위험 요인으로서 중요한 투입 요소에 대한 접근 제한, 시장지배력의 전이, 주요 시장참여자 간 파트너십 등을 제기하기도 하였다.

35. 프랑스 경쟁당국은 생성형 AI 시장상황 등을 살펴본 후 컴퓨팅 자원への 접근, 클라우드 서비스 종속, 데이터 접근, 시장참여자 간 파트너십 등의 측면에서 경쟁제한 우려를 표명하였다. 각 내용을 살펴보면, 주요 AI 반도체인 엔비디아의 GPU 공급 등에 관한 우려, 크레딧(Credit)과 같은 이익 제공 등을 통한 대규모 클라우드 서비스에의 종속 우려를 제기하였다. 데이터에 관해서는 접근 제한 등의 우려 외에도 AI 사업자가 데이터 소유자 동의·승인 없이 데이터를 사용하는 것에 대한 우려, 파트너십을 통해 나타날 수 있는 주요 인력 채용, 기반모델의 독점적 배포 등에 따른 경쟁 우려를 제기하기도 하였다.

36. 캐나다 경쟁당국은 AI가 경쟁에 미칠 수 있는 영향을 캐나다 경쟁법 맥락에서 탐구하여 AI 시장의 경쟁 촉진 방안을 보고서에 담았다. AI 시장을 AI 인프라-AI 개발-AI 배포 단계로 구분하고, 시장에서 발생할 수 있는 경쟁제한 우려를 진입장벽, 경쟁자 배제·차별적 행위, 합병 등을 통한 시장 지배력 강화, AI를 활용한 공동행위, 소비자를 기만하는 마케팅 행위 등을 중심으로 분석하였다. 한편, AI 시장에서 경쟁을 촉진하기 위한 방안으로 정책 입안 시 경쟁정책 평가 도구(Competition Policy Assessment Toolkit)의 활용을 언급하였다.

37. 포르투갈 경쟁당국은 보고서에서 생성형 AI를 디지털 시장에서 제기되는 경쟁 이슈의 대표적인 사례로 인식하였다. 생성형 AI 시장에서는 데이터, 클라우드 컴퓨팅 등 필수요소 시장에서 경쟁자 배제 전략이 사용될 위험이 있으며, 디지털 시장의 특성으로 인하여 고도로 집중화될 수 있는 경향이 있어 일부 시장 참여자들이 경쟁 과정을 왜곡할 수 있다고 보았다. 동 보고서에서는 이러한 경쟁 위험성을 경고하면서, 만약 경쟁 위험이 실제로 발생한다면 소비자 혜택과 경쟁법의 촉진·적용을 보장하기 위해 경쟁당국이 적극 개입할 것임을 밝히기도 하였다.

38. 해외 경쟁당국이 발간한 초기 보고서가 생성형 AI 시장 구조 분석에 초점을 맞추었다면, 최근 발간한 보고서는 다양한 경쟁법적 이슈들을 보다 상세히 다루고 있다. 공정위는 국내 시장을 포함한 생성형 AI 시장의 구조와 현황을 살펴보고, 각 세부 시장에서 발생할 수 있는 경쟁·소비자이익 저해 우려를 공정거래법 관점에서 분석하여 제도개선 등 향후 과제를 제시하고자 하였다.

2 생성형 AI 연관 산업 현황

2.1 AI와 생성형 AI

2.1.1 AI의 개념

39. 현재 다양한 유형·기능의 AI가 출시되고 있으며, 기술의 진화속도가 매우 빠르고 끊임없이 발전하고 있어 AI 개념을 확고하게 정의하기는 곤란하다. OECD 등 주요 기관도 아래와 같이 AI를 다양하게 정의하고 있다.⁸ 우리나라에서도 AI 관련 법안이 다수 발의되고 있으나 2024년 11월 현재 관련 법이 제정되지는 않았으며⁹, 법안마다 AI 개념을 다르게 정의하고 있다.

표 3. 주요 기관별 AI 정의

구분	AI 정의
OECD	· 명시적·암묵적 목표를 위해 받은 입력으로부터 물리적·가상 환경에 영향을 미칠 수 있는 예측·콘텐츠·추천·결정과 같은 결과물의 생성 방식을 추론하는 기계 기반 시스템
유럽연합	· 다양한 수준의 자율성을 가지고 작동하도록 설계되고 배포 후 적응력을 발휘할 수 있으며 명시적 또는 암묵적 목적을 위해 수신한 입력으로부터 물리적 또는 가상 환경에 영향을 미칠 수 있는 예측, 콘텐츠, 추천 또는 결정과 같은 산출물을 생성하는 방법을 추론하는 기계 기반 시스템
미국	· 인간이 정한 일련의 목표를 위해 실제 또는 가상 환경에 영향을 미치는 예측, 권고 또는 결정을 할 수 있는 기계 기반의 시스템

40. AI 개념은 기술 발전 등에 따라 변경될 수 있으나, 그 개념에는 목표 설정에 따른 작동, 자율적 작동과 적응력, 예측·추천·결정 등 기능 등이 포함된다고 볼 수 있다.

8 한국지능정보사회진흥원(2024), OECD AI 시스템 정의 개정의 주요 내용 및 시사점
소프트웨어정책연구소(2024), 유럽연합 인공지능법(EU AI Act)의 주요 내용 및 시사점
법제처(2024), 인공지능(AI) 관련 국내외 법제 동향

9 그간 발의된 법안들을 병합하여 대안으로 발의된 「인공지능 발전과 신뢰 기반 조성에 관한 기본법안」이 2024. 11. 26. 과학기술정보방송통신위원회에서 통과되었고, 해당 법안은 AI 개념을 “학습, 추론, 지각, 판단, 언어의 이해 등 인간이 가진 지적 능력을 전자적 방법으로 구현한 것”으로 정의

2.1.2 AI의 발전 단계

41. AI는 1950년대 최초 언급된 이래로 머신러닝(Machine Learning, 기계 학습)과 딥러닝(Deep Learning, 심층 학습) 등 발전 과정을 거쳐 현재 생성형 AI에 이르렀으며, 인간의 지능 혹은 그 이상을 달성하는 범용 인공지능(Artificial General Intelligence, AGI) 출현 등 그 수준은 지속 발전할 것으로 예상된다.

표 4. 인공지능 세부 분류



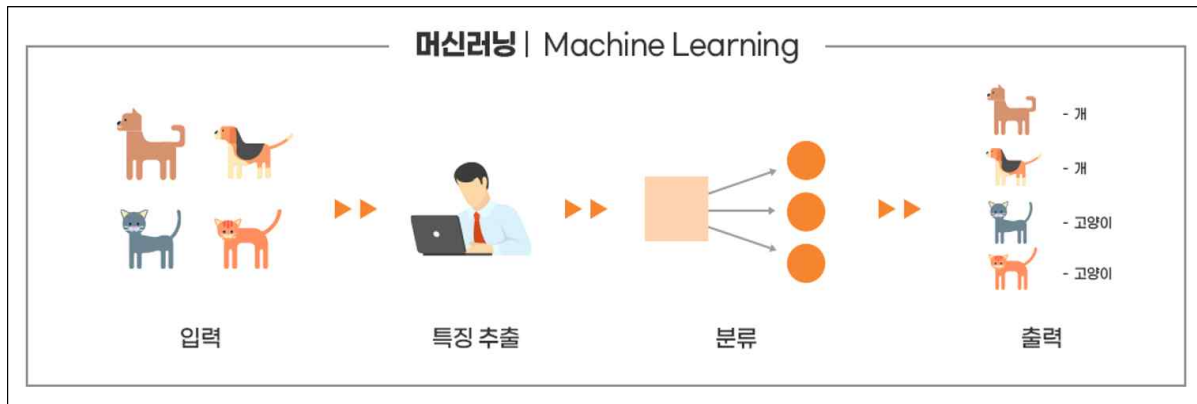
주) LLM(Large Language Model), GAN(Generative Adversarial Networks), VAE(Variational Autoencoder)

* 출처: 한국방송통신전파진흥원¹⁰

42. AI는 1956년 존 매카시(John McCarthy) 등이 참석한 다트머스 회의에서 처음으로 언급되었고, 당시에는 ‘기계를 인간 행동의 지식에서와 같이 행동하게 만드는 것’으로 이해되었다. 그 이후 AI는 데이터에서 특징과 패턴 등을 추출하여 학습시키고, 학습된 데이터를 기반으로 예측 및 추론 등을 수행할 수 있도록 하는 머신러닝 알고리즘으로 발전하였다.

10 한국방송통신전파진흥원(2023), ChatGPT를 넘어 생성형(Generative) AI 시대로: 미디어·콘텐츠 생성형 AI 서비스 사례와 경쟁력 확보 방안

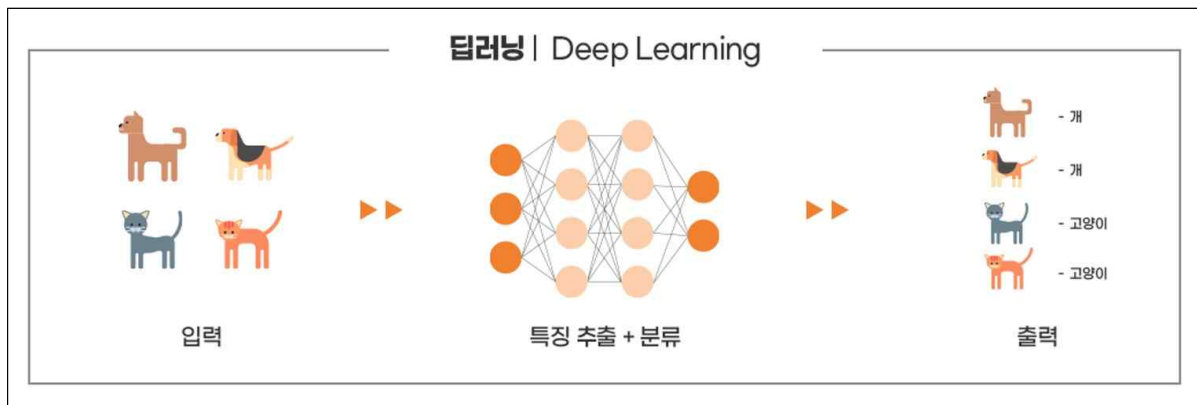
표 5. 머신러닝 방식 구조도



* 출처: SK C&C¹¹

43. 또한, 제프리 힌튼(Geoffrey Hinton)¹²에 의해 인공신경망 연구가 더욱 발전하면서, 딥러닝 알고리즘도 나타났다. 딥러닝은 인간의 뇌 신경망을 모방한 심층 신경망을 사용하여 복잡한 데이터 특성을 스스로 추출·학습하여 예측 또는 추론 등을 수행할 수 있도록 하는 기술을 말하며, 생성형 AI도 이러한 딥러닝 등에 기반하여 나타나게 되었다.

표 6. 딥러닝 방식 구조도



* 출처: SK C&C

11 SK C&C(2023), 머신러닝? 딥러닝? 알쏭달쏭한 인공지능 용어 파헤치기

12 제프리 힌튼은 존 홉필드(John Hopfield)와 함께 인공신경망으로 머신러닝의 기초를 세운 공로로 2024 노벨 물리학상을 수상

2.1.3 생성형 AI의 개념과 의의

44. 생성형 AI는 수많은 데이터를 사전에 학습하고, 이를 기반으로 사용자 요구에 대한 새로운 결과물을 생성할 수 있는 AI 기술 등을 말한다. 기존까지의 AI가 학습 데이터에 기반한 예측·분류 등의 기능을 수행하였던 것에 비해, 생성형 AI는 학습 데이터를 통해 새로운 콘텐츠 등을 생성할 수 있다는 것이 주요한 차이점이라고 할 수 있다.

45. 생성형 AI는 2014년 적대적 생성 신경망(Generative Adversarial Networks, GAN) 개념이 등장한 이래¹³ 아래와 같이 꾸준히 발전하였고, 2022년 말 오픈AI의 챗GPT(Chat GPT) 출시 이후 많은 사업자들이 생성형 AI를 개발·출시하거나 연관산업에 활용하여 다양한 서비스를 제공하고 있다.

표 7. AI 분야 핵심 사건

모델	합성곱 신경망(CNN) 2012년 이미지넷 챌린지에서 CNN 기반 딥러닝 알고리즘 AlexNet이 우승을 차지하며 딥러닝 부흥의 계기 이미지 처리 딥러닝 본격화	텐서플로(Tensor Flow) 구글에서 인공지능 개발을 위한 프레임워크를 오픈소스로 공개 오픈소스 문화에 기여	트랜스포머(Transformer) 구글이 개발한 자연어 처리 모델 기존 RNN구조의 단점을 극복하여 여러 모델파생(ELMo, BERT, GPT의 기반) 언어모델의 혁신적 돌파구 마련	GPT-2언어모델 BERT언어모델 언어모델의 새로운 표준 역할 기술 성숙의 시대	AI반도체 전성시대 GPU가 AI반도체의 중심으로 부상 엔비디아의 폭발적 성장 DALL-E 출시 이후, 글로벌 빅테크 들은 멀티모달 모델 개발 본격화 멀티모달 모델의 서막	생성AI 유니콘의 탄생 ChatGPT
모델	적대적 생성 신경망(GAN) 생성자와 식별자가 서로 대립하며 데이터를 생성하는 모델 (기존 생성AI 대비 성능 우수) (DGGAN, starGAN 등 파생) 생성 AI의 새로운 혁신	알파고(AlphaGo) 구글 딥마인드가 개발한 인공지능 바둑 프로그램 인공지능 대중화	OpenAI의 GPT Paper(6월) 기술 성숙의 시대	GPT-3 초거대 언어모델의 시작 (1,750억개 매개변수) 알파폴드2(Alphafold)2 단백질의 구조 예측 분야에서 뛰어난 성과 달성 난제 해결을 위해 AI 응용 본격화	DALL-E2 Midjourney Stable diffusion Pati Imagen 멀티모달의 전성시대	GPT4

* 출처: 한국저작권위원회

46. 생성형 AI 관련시장에는 구글, 마이크로소프트와 같은 주요 해외 빅테크 뿐만 아니라 국내 대기업 및 스타트업 등도 적극 진출하여 참여하고 있는 것으로 나타난다. 아울러, 새로운 결과물을 생성하기 위해서는 높은 연산·추론 능력이 요구되므로, 컴퓨팅 자원과 데이터 같은 AI 인프라 구축 과정을 비롯한 AI 개발 과정에서 기술과 자본 등에 대한 상당한 투자도 요구되고 있다.

13 한국저작권위원회(2023), 생성형 인공지능(Generative AI) 산업 현황 보고서

47. 한편, 생성형 AI는 높은 효용성 등으로 인하여 스마트폰이나 PC 같은 온디바이스(On-device), 자율주행, 보안, 금융 등 전 산업 응용처로 확대되고 있는 상황이며, 민간에서도 그 효용성을 상당히 인식하고 있는 것으로 나타난다.

48. Pwc Survey에 따르면, 60% 이상의 CEO들이 생성형 AI를 통해 작업 효율도가 5% 이상 개선될 것으로 보았으며, Deloitte CEO Survey에 따르면 생성형 AI 도입으로 CEO들은 기업운영 효율성 및 운영 자동화 등 도움을 받을 것으로 기대한 것으로 나타났다.¹⁴ 아울러, McKinsey에 따르면, 63개 생성형 AI 사용 사례를 분석할 때 연간 2.6조~4.4조 달러 가치를 창출할 것으로 추정되었다¹⁵.

49. 생성형 AI 개발·응용 분야의 지속적 확장으로 인하여 정확한 시장규모를 측정하기는 곤란하지만, 주요 시장조사 기관들은 글로벌 생성형 AI 시장 규모가 향후 크게 성장할 것으로 전망하였다. 시장 조사기관 Statista는 2023년 약 449억 달러에서 2030년 약 2,070억 달러까지 성장할 것으로¹⁶, Precedence Research는 2023년 약 176억 달러에서 2030년 약 2,556억 달러까지 성장할 것으로 보았다¹⁷.

50. 생성형 AI는 높은 생산성 등으로 산업혁신을 촉진하며 급속하게 발전하고 있으나, 한편으로는 생성형 AI 관련 저작권 분쟁이 발생¹⁸하거나 생성형 AI 관련 산업에서 경쟁저해 우려 등도 제기되는 상황이다. 전 세계적으로 생성형 AI 수요의 폭발적 증가와 시장 확대 등에 따라 많은 사업자들이 시장 우위를 점하기 위해 경쟁하는 과정에서 필수 요소에 대한 접근 제한, 결합 판매 등에 따른 경쟁저해 우려 등이 있는 것으로 나타난다.

14 삼일 PwC 경영연구원(2024), 생성형 AI를 활용한 비즈니스의 현주소

15 McKinsey(2023), OECD(2024), Artificial Intelligence, Data and Competition에서 재인용

16 Statista(2023), 삼일 PwC 경영연구원(2023), 미리보는 CES 2024에서 재인용

17 Precedence Research(2024), Generative AI market size 2023 to 2033

18 해외 사례를 예시로 들면, Getty Images는 Stability AI의 Stable Diffusion 학습에 자사 이미지가 무단 이용되었음을 주장하며 2023년 2월 미국에서 저작권 침해 및 상표권 침해 소송을 제기. 다만, 본 보고서에서는 생성형 AI에 따른 저작권 분쟁보다는 경쟁법적 쟁점 등에 대하여 보다 중점적으로 검토하였음

2.1.4 생성형 AI의 유형

51. 생성형 AI는 상당한 기술과 자본력 등을 통해 구축한 ‘AI 모델 (Model)’에 기반하여 작동하며, 이러한 AI 모델은 산출물·규모·작동 방식 등과 같이 다양한 기준에 따라 구분할 수 있다.

52. 우선, 최종 산출물의 성격에 따라 언어·이미지·음성 등 생성 모델로 구분할 수 있으며¹⁹, 텍스트·이미지 등 여러 유형의 데이터를 생성할 수 있는 다중(Multimodal, 멀티모달) 모델 등도 존재한다.

표 8. 최종 산출물 성격에 따른 분류 예시

구분	설명	주요 모델
언어 생성 모델	· 문장 생성, 기계 번역, 질문 답변 및 요약문 생성 등에 활용되는 모델	· GPT, LaMDA, 엑사원 등
이미지 생성 모델	· 이미지 증강, 스타일 변환, 캐릭터 생성 등에 활용되는 모델	· Imagen, Stable Diffusion 등
음성 생성 모델	· 음성 합성, 음성 변환, 음악 생성 등에 활용되는 모델	· VALL-E 등

53. 다음으로, 언어 생성 모델의 경우 매개변수(Parameter)²⁰의 개수에 따라 대규모 언어 모델(Large Language Model, 이하 ‘LLM’)과 소규모 언어 모델(Small Language Model, 이하 ‘SLM’)로 구분할 수도 있다²¹. LLM과 SLM의 분류를 명확히 정의하는 기준은 확인되지 않지만, SLM은 통상적으로 수십만에서 수십억 개의 매개변수를 보유하고, LLM은 수천억 개 이상의 매개변수를 보유하는 것으로 볼 수 있다.

19 한국저작권위원회(2023), 생성형 인공지능(Generative AI) 산업 현황 보고서

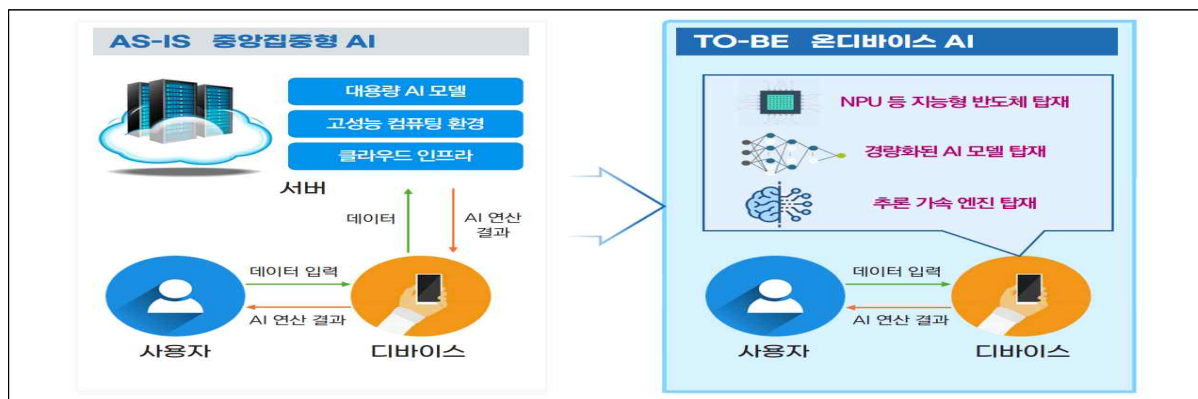
20 매개변수는 AI 모델 내부에서 결정되는 변수로 AI 훈련 과정에서 입력이 결과값으로 변환될 수 있도록 조정하는 가중치를 말하며, 매개변수가 많을수록 더 많은 정보를 저장하고 처리할 수 있음

21 한국전자통신연구원(2024), 온디바이스 소형언어모델 기술개발 동향

54. 예를 들면, LLM 중 GPT-3.5 모델이 1,750억 개, 하이퍼클로바 모델이 2,040억 개를 보유하고²², SLM 중 LLaMa3 8B 모델은 80억 개, Mistral 7B 모델이 70억 개의 매개변수 보유한 것으로 확인된다. 또한, SLM은 LLM과 비교하여 상대적으로 훈련을 위한 데이터 규모가 작고 컴퓨팅 자원 요구사항도 낮아 메모리·컴퓨팅 자원이 제한적인 모바일 기기 등에서 구동될 수도 있다.

55. 다른 구분 기준으로서, AI가 어떠한 인프라로 운영되는지 등에 따라 클라우드 기반 AI와 온디바이스(On-device) AI로 구분할 수도 있다²³. 클라우드 기반 AI는 사용자의 질문 및 요구를 받아 이를 외부 서버에 전송하고, 그 연산 결과를 전달받아 사용자에게 답변을 제공한다. 다만, 이는 개인 또는 회사 데이터 등을 외부 서버로 전송함에 따라 정보 유출 우려가 발생할 수 있고, 클라우드 등 외부 서버와의 네트워크 환경이 강제됨에 따른 데이터 전송 및 결과 수신 과정상 지연이 발생할 가능성이 존재한다. 온디바이스 AI는 데이터를 외부 서버에 전달하지 않고 별도의 인터넷 연결 없이 기기에 탑재된 컴퓨팅 자원 등을 활용하여 자체적으로 AI 연산을 수행하여 사용자에게 결과를 제공하는 AI를 말하며, 클라우드 기반 AI의 정보유출 우려, 네트워크 환경에 따른 지연 우려 등에 대한 대안으로 그 필요성이 대두되었다고 볼 수 있다.

표 9. 클라우드 기반 AI 및 온디바이스 AI 비교



* 출처: 한국전자정보통신산업진흥회

22 소프트웨어 정책연구소(2023), 초거대언어모델의 부상과 주요 이슈

23 한국전자정보통신산업진흥회(2024), 온디바이스 AI 기술동향 및 발전방향

2.2 생성형 AI의 구축과 실행

2.2.1 구축·실행 제반 과정

표 10. 생성형 AI 주요 가치사슬



* 영국 CMA 보고서 등 참고하여 작성

56. 생성형 AI는 위와 같이 컴퓨팅 자원·데이터·전문인력 등의 AI 인프라 구축 단계, 이러한 인프라를 활용하여 AI 기반모델을 구축하는 개발 단계, AI 기반모델에 기초한 서비스 제공 단계를 통해 구현된다.

57. AI 인프라 단계는 AI 개발·구현 등을 위하여 갖추어야 하는 AI 반도체·클라우드 등 컴퓨팅 자원, 데이터, 전문인력 등 필수 요소를 구축하는 단계를 말한다. 이는 초기 개발 과정뿐만 아니라 그 이후의 AI 기능 개선 등을 위하여 계속 요구되고 상당한 투자도 필요한 단계로 볼 수 있다.

58. AI 개발 단계는 특정 AI 기능을 구현하는데 기초가 되는 기본모델을 직접 개발하거나, 자사 AI 서비스를 개발·제공하기 위하여 다른 사업자의 기본모델을 미세조정하는 단계 등을 말한다.

59. AI 구현 단계는 기본모델에 근거하여 텍스트·이미지 생성 등 AI 기능을 구현하고, 이를 활용하여 수요자에게 AI 서비스를 제공하는 단계를 말한다.

2.2.2 생성형 AI의 가치사슬 현황

2.2.2.1 AI 반도체

60. AI 반도체는 데이터 학습과 추론, 결과물 생성 등 AI 서비스 구현에 필요한 대규모 연산을 효율적으로 수행할 수 있도록 하는 반도체 등을 말하며²⁴, 대표적으로 GPU(그래픽 처리 장치, Graphic Processing Unit), NPU(신경망 처리 장치, Neural Processing Unit)²⁵ 등 여러 유형이 존재한다.

61. 그 중 GPU는 주로 게임 산업에서 3D 그래픽 등을 처리하기 위하여 활용되었으나, AI를 위한 대규모 데이터 연산에 효율적이라는 점이 드러나면서 AI 반도체로 주목받게 되었다²⁶. 2012년 이미지 인식을 겨루는 이미지넷 대회(Imagenet Large Scale Visual Recognition Challenge, ILSVRC)에서 GPU 기반의 딥러닝 모델 ‘알렉스넷’이 약 84.7%의 높은 이미지 인식률²⁷로 우승하였고, 이로 인해 AI 분야에서 GPU의 비약적인 발전 가능성이 대두되었다.

62. 최근에는 인공신경망 연산에 특화된 저전력 AI 반도체인 NPU뿐만 아니라 프로세서와 메모리를 융합하여 저전력 AI 연산을 지원하는 PIM(Processing in Memory), 인간 뇌의 동작 방식을 모방하여 디자인 되고 에너지 효율성을 높이는 뉴로모픽(Neuromorphic) 등 새로운 AI 반도체 기술도 등장하고 있다.

63. 주요 AI 반도체 회사들은 GPU·NPU 등 AI 반도체를 개발하고, 이에 HBM 같은 메모리 반도체²⁸ 등 부품을 결합한 제품인 AI 가속기²⁹를 제조하여 서버 제조사 또는 AI 개발사 등 수요자에게 공급하고 있다.

24 한국전자통신연구원(2024), 주요국 반도체 정책과 AI반도체 정책에의 시사점

25 NPU는 특정 AI 알고리즘에 특화된 반도체로서 전력을 상대적으로 적게 소모하는 등의 장점을 가지고 있으나, GPU 대비 범용성이 낮아 다른 AI 알고리즘을 사용할 경우 연산 자체가 불가능하거나 속도가 크게 떨어지는 등의 단점도 가지고 있음

26 한국과학기술평가원(2019), 기술동향 브리프 인공지능

27 그 이전에는 이미지 인식률이 75%를 넘지 못하였음

64. AI 반도체 회사는 AI 반도체 설계부터 제조까지 전반적 과정을 직접 수행하는 종합 반도체 회사(Integrated Device Manufacturer, IDM) 및 AI 반도체 설계만 담당하고 제조 과정을 다른 회사에 위탁하는 팹리스(Fabless) 회사 등으로 구분된다. 예를 들면, 대표적인 AI 반도체 기업 엔비디아는 팹리스 회사로서 GPU를 설계하여 TSMC 등 파운드리(Foundry) 회사에 제조를 위탁하며, 파운드리 회사는 타 사업자로부터 HBM 등 메모리 반도체를 공급받아 GPU와 HBM 등을 결합하여 AI 가속기를 제조하는 것으로 알려져있다.

표 11. AI 반도체 및 AI 가속기 예시



* 출처: CNET, 사피온코리아, HPE 등³⁰

28 HBM(고대역폭 메모리, High Bandwidth Memory)은 메모리 반도체의 일종으로서, GPU가 연산해야 하는 데이터를 전달하고 그 결과를 받아 저장하는 기능을 수행함

29 AI 가속기, AI 카드 등 다양한 명칭으로 불리고 있으며, 이하 ‘AI 가속기’ 라 함

30 CNET(2022), See Nvidia’s H100 Hopper Chip Up Close
사피온코리아(2024), X330 제품 개요 / HPE 사이트 등

65. 한편, 반도체 회사가 AI 가속기를 서버³¹ 제조사에 공급하면, 서버 제조사는 AI 가속기와 전원 장치 등 다른 부품을 함께 결합하여 서버를 제작하고, 이를 AI 개발사 등 수요자에게 제공한다.

66. 위와 같은 과정을 거쳐 AI 개발사는 AI 개발을 위한 컴퓨팅 자원으로 AI 반도체를 활용할 수 있으며³², 생성형 AI 개발 및 구현, 성능 고도화를 위해서는 고도의 연산능력이 필요하므로, AI 개발사에게 AI 반도체 및 가속기를 다수 확보하는 것이 매우 중요하다.

67. AI 반도체·가속기 등 컴퓨팅 자원의 필요성과 중요성은 아래와 같이 AI 실태조사 결과에서 반도체 회사 및 AI 개발사 응답을 통해서도 확인되었다.

표 12. AI 실태조사 결과 중 AI 반도체 중요성 응답

구분	주요 의견
반도체 회사 응답	<ul style="list-style-type: none"> • 고성능 연산 자원을 필요로 하는 생성형 AI 시장에서 AI 반도체는 AI 서비스의 근간을 이루는 인프라 • AI 모델이 방대한 데이터를 빠르게 처리하기 위해 고성능 컴퓨팅 자원이 필요함
AI 개발사 응답	<ul style="list-style-type: none"> • 방대한 데이터를 활용하여 AI 기반모델을 학습시키기 위해 GPU 등 AI 반도체 성능 및 수량이 절대적으로 중요 • AI 모델의 학습 및 추론과 관련하여 GPU 등 AI 반도체의 영향력이 매우 큼

68. 현재 전세계적인 생성형 AI 수요 증가로 많은 사업자들이 AI 개발에 착수함에 따라 이를 위한 AI 반도체·가속기 수요도 함께 증가하고 있다. 다음과 같이, 주요 AI 모델을 개발하고 있는 사업자들은 AI 모델 개발 등을 위하여 상당한 수량의 AI 가속기를 활용하였던 것으로 확인되었으며, 대표적인 AI 가속기는 엔비디아 社의 제품 A100, H100 등으로 알려져 있다.

31 서버는 네트워크를 통해 정보나 서비스 등을 제공하는 역할을 수행하는 장치 등을 말함

32 AI 개발사의 AI 반도체 활용 방식은 이에 한정되지 않으며, 다른 활용 방식이 더 존재할 수 있음

표 13. 주요 AI 모델별 AI 가속기 활용 개수

회사	AI 모델	AI 가속기	활용 개수
메타	LLaMa	A100	2,048개
구글	LaMDA	TPU v3	1,024개
오픈AI	GPT-3	A100	1,024개(추정)
마이크로소프트, 엔비디아	MT-NLG	A100	4,480개

* 출처: CMA³³

69. AI 가속기는 공급 물량 부족 등에 따라 품귀 현상이 발생하여 확보에 상당한 시일이 소요되었던 사례도 있으며, 일부 AI 가속기의 경우 높은 가격을 형성하고 있는 것으로 확인되기 때문에³⁴ AI 가속기를 대량으로 확보하기 위해서는 상당한 투자 비용이 필요하다고 볼 수 있다.

33 CMA(2023), AI Foundation Models Initial Report

34 엔비디아 주요 제품인 H100의 경우 2023년 당시 납품 주기가 8~11개월, 가격이 3~4만 달러 수준을 형성하였던 것으로 나타남(출처: 한국수출입은행, AI 반도체 시장 현황 및 전망)

2.2.2.2 클라우드 컴퓨팅

70. 클라우드 컴퓨팅은 AI 개발사가 AI 개발·구현 등에 필요한 컴퓨팅 자원을 클라우드 서비스 제공 사업자(Cloud Service Provider, 이하 ‘CSP’)로부터 클라우드 서비스 방식으로 제공받는 것을 말한다.

71. 앞서 살펴본 바와 같이 생성형 AI 수요는 증가하고 있으나, AI 개발사가 직접 AI 반도체·가속기를 다수 확보하여 자체 컴퓨팅 능력을 갖추는 것이 쉽지 않고, 그 기간과 비용이 상당히 소요될 수 있기 때문에 클라우드 서비스를 통해 컴퓨팅 자원을 활용하는 사례가 나타나고 있는 상황이다.

72. 아울러, 기술 발전으로 인하여 기존보다 효율적인 AI 반도체가 개발되는 경우 이미 상당한 자본으로 구매하였던 AI 반도체의 효용성 등이 떨어질 수도 있기 때문에 AI 개발사 입장에서는 직접 투자보다 클라우드 서비스를 통해 컴퓨팅 자원을 활용할 유인도 가지게 된다고 볼 수 있다.

73. CSP가 AI 개발사 등에게 클라우드 컴퓨팅을 제공하는 경우, CSP가 AI 반도체·가속기의 수요자로서 이를 대량으로 확보하고 부지·설비 등에 대한 상당한 투자를 진행하여 데이터센터 등을 구축할 필요가 있다.

표 14. 데이터센터 내부 사진(예시)



* 출처: 하나금융경제연구소³⁵

35 하나금융경영연구소(2024), 국내 데이터센터 산업, AI와 ESG 사이에서 길을 찾다

74. CSP는 AI 가속기 등을 확보하기 위하여 제조사 등으로부터 이를 공급받을 수도 있으나, 최근 CSP가 자체적으로 AI 반도체를 개발하여 자신의 클라우드 서비스 인프라로 활용하는 사례도 나타나고 있다. 다만, CSP의 자체 AI 반도체 개발·활용 방식은 상당한 자본과 기술 투자 또는 제조사와의 협업 등이 요구되기 때문에 현재 이러한 방식으로 클라우드 컴퓨팅을 제공할 수 있는 사업자는 일부에 한정된 것으로 보인다.

75. 클라우드 컴퓨팅 또한 AI 반도체와 마찬가지로 AI 개발·구축을 위한 중요한 요소에 해당하며, 특히 AI 가속기 등 컴퓨팅 자원을 대량으로 조달할 여력이 충분하지 않은 스타트업 등 소규모 회사 입장에서는 그 중요성이 더욱 크게 작용한다고 볼 수 있다.

76. AI 실태조사 결과에서도 CSP 및 AI 개발사 등 사업자들이 클라우드 컴퓨팅의 중요성과 장점 등에 대해 아래와 같이 의견을 제출하기도 하였다.

표 15. AI 실태조사 결과 중 클라우드 컴퓨팅 중요성 의견

구분	주요 의견
CSP 응답	<ul style="list-style-type: none"> · 대용량 자원이 필요할 때마다 적시에 공급받기 위해서는 클라우드 서비스가 중요함 · 대규모 데이터 처리, 복잡한 학습 등을 위해서는 컴퓨팅 자원 및 유연한 확장성을 제공하는 클라우드가 필수적
AI 개발사 응답	<ul style="list-style-type: none"> · 컴퓨팅 자원 인프라를 자체 구축하고자 할 경우 운영·유지보수 관련 인력 확보가 어려워, 클라우드 활용이 효율적 · AI 발전 시 많은 컴퓨팅 자원이 필요한데, 클라우드는 그 발전에 따라 유연하게 자원을 확장할 수 있는 장점 보유

77. 아울러, 현재 생성형 AI 수요 증가 등으로 인하여 일부 CSP가 AI 개발사 등에게 클라우드 컴퓨팅 자원을 제공하는 것 외에도 자사 클라우드 플랫폼에서 AI 기반모델에 대한 접근·사용 기능도 제공하는 등 클라우드 서비스 제공 범위를 확장하고 있는 것으로 나타난다.

78. CSP의 서비스 제공 범위 확장은 CSP와 AI 기반모델 개발사 간 사업 제휴·파트너십 사례 등을 통해서 나타날 수 있다. 일부 CSP가 AI 기반모델 개발사에 대해 컴퓨팅 자원을 제공하거나 대규모 자본 투자를 하고, AI 기반모델 개발사로부터 기반모델 라이선스를 제공받은 경우도 확인된 바 있다.

79. 위와 같은 CSP의 서비스 확장은 AI 기반모델 개발사뿐만 아니라, 기반모델을 활용하여 AI 서비스를 개발하는 회사들에게도 이점을 제공할 수 있다. 즉, AI 서비스 개발사들은 CSP의 클라우드 컴퓨팅 인프라를 직접 이용하기도 하지만, 플랫폼에서 특정 AI 기반모델을 사용할 수 있게 되어 AI 서비스를 개발·구현하는 것이 보다 용이할 가능성도 있다³⁶.

80. 다만, CSP의 서비스 확장은 생성형 AI 가치사슬 내 수직통합 등으로 이어질 수 있고, 이로 인해 경쟁법적 쟁점을 야기할 수 있다는 우려도 존재한다. 이는 이하 3. 경쟁·소비자 쟁점 등에서 다루기로 한다.

36 Autorité de la concurrence(2024), On the competitive functioning of the generative artificial intelligence sector

2.2.2.3 데이터

81. 데이터는 AI 모델의 학습 등을 위하여 활용되는 데이터를 말하며, 말뭉치 데이터 외에도 이미지, 음성 등 다양한 유형의 데이터를 포함한다.

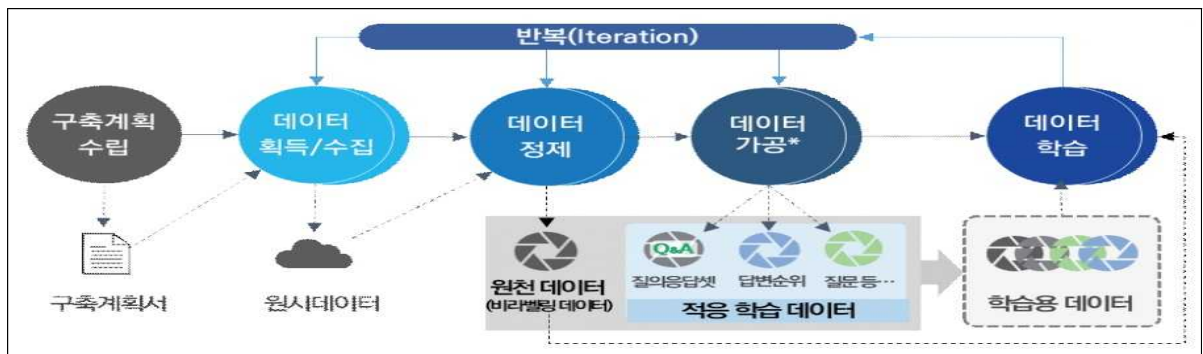
82. AI 학습에 활용할 데이터를 구축하는 방법은 AI 기반모델의 종류 등에 따라 다양하지만, 기본적으로 ‘구축계획 수립 → 데이터 획득·수집 → 데이터 정제 → 데이터 가공 → 학습’ 단계를 반복하여 구축된다³⁷.

표 16. AI 학습 데이터 구축 단계별 설명

단계	설명
구축계획 수립	· 데이터 구축 목적을 명확하게 정의하고, 그 목적에 맞는 AI 모델 임무, 데이터 구축 수량 및 방법, 절차 등 설계
데이터 획득·수집	· 학습에 필요한 데이터를 직접 획득하거나, 이미 보유하고 있는 조직 등으로부터 수집하는 등 원시데이터 ^{주)} 확보
데이터 정제	· 원시데이터를 학습에 필요한 형식 및 크기로 맞추고, 개인정보 비식별화, 민감정보 처리 등을 통해 원천데이터 확보
데이터 가공	· AI 모델의 적절한 답변을 위한 질의응답 셋 구축, 응답결과 우선순위 설정 등 미세조정 학습을 위한 다양한 형태의 데이터 구축
데이터 학습	· 학습 데이터 셋을 이용하여 AI 모델을 학습시키고, 사전학습된 모델의 성능을 향상시키거나 보정하는 활동을 수행

주) 원시데이터는 아직 가공되지 않은 최초의 데이터 등을 말함

표 17. AI 학습 데이터 구축 단계 구조도



37 과학기술정보통신부 등(2024), 인공지능 학습용 데이터 품질관리 가이드라인 v3.1 제2권 초거대 AI 데이터 품질관리 가이드라인

83. 사업자가 데이터를 수집하는 방식으로는 자신이 제공하는 서비스 등을 통한 데이터 수집, 공공·개방 데이터 수집, 데이터 보유 사업자와의 거래를 통한 수집, AI 생성 데이터(합성 데이터) 활용 등 여러 방식이 존재한다. AI 실태조사 결과에 따르면, 데이터 보유 사업자와의 거래를 통한 데이터 수집 방식은 ‘저작권법 등에 따라 보호받는 데이터, 정제되거나 고품질의 데이터 확보 목적’ 등에 따라 이루어지는 것으로 나타났다.

84. 데이터는 AI 기반모델의 학습과 훈련 등을 위한 필수적 요소이고, 데이터의 다양성과 수량 등 품질이 기반모델의 성능과 가치 및 경쟁력³⁸ 등으로 이어질 수 있으므로, 그 중요성은 상당히 크다고 볼 수 있다. 실제로 주요 AI 기반모델 개발사들이 자사 AI 모델을 학습시키기 위하여 아래 표 18과 같이 대량의 데이터를 활용한 것으로 확인되며³⁹, AI 실태조사 결과에서도 표 19와 같이 데이터가 중요하다는 응답이 다수 제출되었다.

표 18. 주요 AI 모델별 학습 데이터

(단위: 토큰, B: 10억)

회사	AI 모델	훈련 데이터
메타	LLaMa	1,400B
구글	LaMDA	2,810B
오픈AI	GPT-3	300B
마이크로소프트, 엔비디아	MT-NLG	270B

* 출처: CMA

표 19. AI 실태조사 결과 중 데이터 중요성 응답

주요 의견
<ul style="list-style-type: none"> 동일한 AI 모델이라고 하더라도, 데이터의 품질·다양성·정교함 등에 따라 모델 성능 차이가 크게 나타남 데이터는 AI 모델의 지속적 업데이트와 품질 개선에 중요한 역할을 하며, 최신 정보와 트렌드를 반영하여 성능을 향상시키는 데 기여함

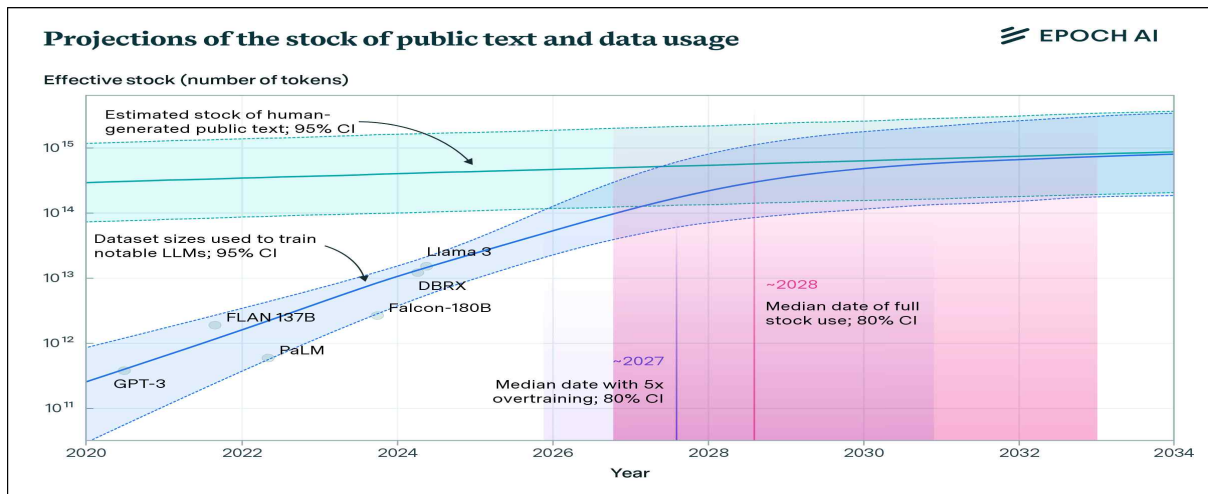
38 OECD(2024), Artificial Intelligence, Data and Competition

39 CMA(2023), AI Foundation Models Initial Report

85. 또한, 불충분한 학습 등으로 인하여 데이터 편향이 나타나는 ‘AI 왜곡(Hallucination)’ 현상, AI 생성 데이터만으로 반복 학습할 때 그 답변이 부정확해지는 ‘AI 모델 붕괴(Collapse)’ 우려⁴⁰ 등을 고려할 때, AI 학습 및 훈련을 위한 원시데이터의 중요성이 더욱 더 강조되는 상황이다. 이러한 AI 왜곡 및 붕괴 우려 등을 방지하기 위해 ‘인간 피드백을 통한 강화 학습(Reinforcement Learning from Human Feedback, RLHF)⁴¹’ 등이 활용되기도 한다.

86. 한편, 사업자가 수집할 수 있는 데이터 중 공공·개방 데이터의 경우, 일부에서는 생성형 AI 수요 증가 등에 따라 향후 수년 이내에 고갈될 수 있다는 의견도 나타나고 있다. 예를 들어, 미국 민간 연구단체 EPOCH AI는 아래와 같이 지금의 거대 AI 모델 개발 추세라면 2026~2032년 AI 학습에 적합한 고품질 공개 텍스트 데이터가 소진될 것(파란색 곡선이 초록색 선에 수렴)으로 예측하기도 하였다.⁴²

표 20. 데이터 고갈에 대한 EPOCH AI 분석



* 출처: EPOCH AI

40 Nature(2024), AI models collapse when trained on recursively generated data

41 이는 인간의 피드백을 기반으로 보상받으며 학습하는 방법 등으로서, 모델이 생성한 결과에 대한 인간의 반응이나 평가를 통해 보상을 조정하고, 이를 통해 모델은 인간에게 맞춰진 반응을 할 수 있도록 학습됨(SK하이닉스(2024), All around AI 3편 머신러닝의 이해)

42 EPOCH AI(2024), Will We Run Out of Data? Limits of LLM Scaling Based on Human-Generated Data

87. 공공·개방 데이터의 고갈로 인하여 향후 특정 사업자가 보유하고 있는 자체 데이터 등에 대한 의존도가 높아질 수 있으며, 이는 데이터 비용 인상과 AI 학습·훈련 비용의 상승 등으로 이어질 가능성도 존재한다. 영국 경쟁당국은 데이터를 보유한 사업자가 수요자에게 자사 데이터에 대한 접근·사용 등의 권한을 제공하면서, 이러한 권한 제공을 주요 수익원으로 전환하는 등 수익화 추세에 있다고 분석하기도 하였다.⁴³

88. 다만, AI 개발의 필수 요소인 데이터의 중요성이 커질수록 가격 등에 대한 거래조건 협상과 시장 경쟁 상황에 대한 사업적 판단에 따라 데이터 접근제한이 발생할 수 있다는 우려도 나타나고 있다. 아울러, 기반모델의 개발과 추론에 있어 필수적인 데이터의 수집과 관련한 경쟁법적 이슈도 제기되고 있어 이는 이하 3. 경쟁·소비자 쟁점에서 다루기로 한다.

43 CMA(2023), Trends in Digital Markets

2.2.2.4 전문인력

89. 전문인력은 컴퓨팅 자원의 유지·활용, 데이터 정제·가공, AI 기반모델 개발, AI 서비스 구축 등 AI 개발·구현 업무를 전문적으로 처리할 수 있는 인력을 말한다.

90. AI 사업을 위해서는 관련 전문인력이 반드시 필요하고, 능력 있는 전문인력을 어느 정도 확보하는지가 각 사업자의 제품·서비스 기술 수준 및 경쟁력 등에 영향을 미칠 수 있기 때문에 AI 가치사슬에서 다른 요소 못지않게 전문인력도 중요한 요소에 해당한다.

91. 최근 생성형 AI 수요 증가에 따라 가치사슬 전 분야에서 전문인력에 대한 수요도 증가하고 있어 전문인력 확보를 위해 사업자들이 상당한 비용을 지출하고 있는 상황이다. 실제로 상당한 자본력 등을 갖추고 있는 AI 사업자가 특정 AI 개발사에 있던 직원을 대거 채용하는 사례도 나타나는 등 전문인력 확보를 위한 경쟁도 심화하고 있다.

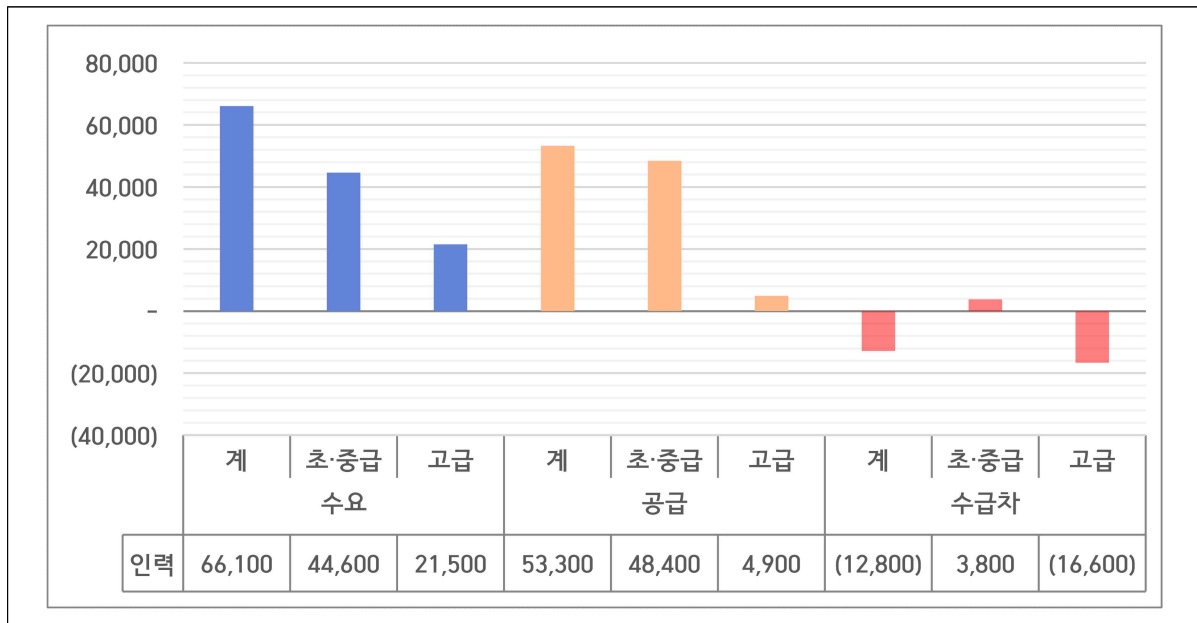
92. AI 실태조사 결과에서도 다수 사업자들이 전문인력 수요에 비해 실무 경험과 전문 지식을 갖춘 인재 공급이 제한적이고, 전문인력의 해외 유출, 대기업들의 공격적 인재 확보, 이로 인한 상당한 수준의 급여 제공 필요 등에 따라 전문인력 확보가 어렵다는 의견을 제시하였다.

표 21. AI 실태조사 결과 중 전문인력 채용의 어려움 응답

주요 의견
<ul style="list-style-type: none"> · AI 인재에 대한 기업들의 수요는 계속 늘어나고 있으나, 국내 인력 풀이 크지 않고 우수 인력들의 해외 진출이 많이 나타나고 있음 · AI 전문인력에 대한 높은 수요로 인하여 상당한 수준의 급여가 요구되는 등 회사 입장에서 부담으로 작용

93. 고용노동부에서 발표한 인력수급 전망 결과에서도 2023년~2027년 AI 분야에서 12,800명이 부족한 것으로 나타났으며⁴⁴, “고급인력 해외유출 우려 등 고급 수준 인력난 심화 전망”으로 분석하였다.

표 22. 고용노동부의 AI 분야 인력수급 전망 결과



94. 또한, 스탠퍼드 인간중심 인공지능 연구소(Human-Centered Artificial Intelligence, 이하 ‘HAI’) 분석⁴⁵에 따르면, 아래와 같이 2023년 국내에서 전문인력이 해외로 유출되는 현상이 발생한 것으로 나타났다.

표 23. 2023년 주요 국가별 AI 전문인력 유입·유출 지수

(단위: 명)

국가	룩셈부르크	스위스	영국	미국	한국
지수	3.67	1.6	0.41	0.4	-0.3

44 고용노동부(2023), 27년까지 인공지능(AI) 12,800명, 클라우드 18,800명 신규인력 부족 전망 보도자료

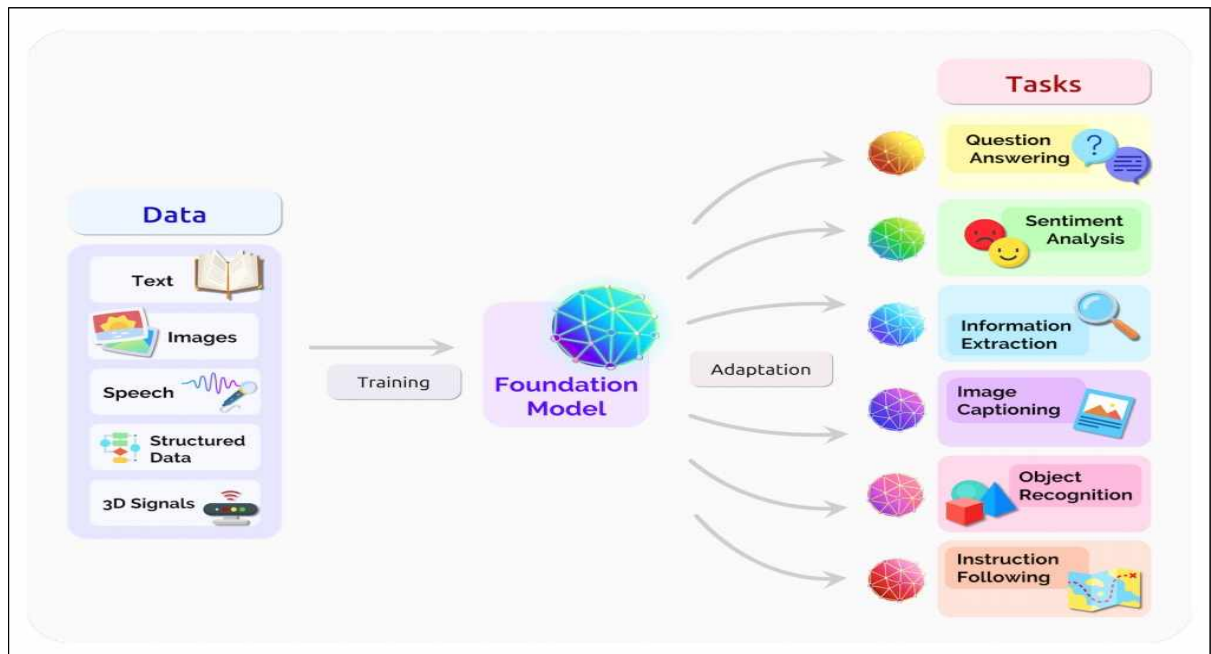
45 HAI는 글로벌 채용 플랫폼인 링크드인(LinkedIn)을 기준으로, 10,000명당 AI 전문인력 유입·유출 등을 분석(출처: HAI(2024), Artificial Intelligence Index Report 2024)

2.2.2.5 기반모델

95. 기반모델(Foundation Model, 이하 ‘FM’ 또는 ‘기반모델’ 이라 함)은 방대한 양의 데이터를 학습하여 텍스트·이미지 생성 등 다양한 작업을 수행할 수 있는 생성형 AI의 기초가 되는 모델 등을 말한다.

96. 기반모델 용어는 스탠퍼드 대학교 파운데이션 모델 연구 센터(Center for Research on Foundation Models, 이하 ‘CRFM’)가 처음 제안하였으며, 그 용어를 사용한 이유에 대하여 ‘특유의 완성되지 않은 채 배포되는 특성’ 을 언급하였다⁴⁶. 즉, 기반모델은 아직 명확하게 수행해야 할 작업이 정해지지 않은 상태로 배포되어 사용자가 원하는 목적에 맞게 미세조정(Fine-Tuning) 등과 같은 과정을 거침으로써 완성되는 모델이라고 할 수 있다.

표 24. 기반모델 도식화



* 출처: CRFM⁴⁷

46 한국지능정보사회진흥원(2023), 파운데이션 모델의 이해와 미래 전망

47 CRFM(2021), On the opportunities and risks of foundation models

97. 기반모델은 생성형 AI 기능을 구현할 수 있도록 하는 틀·기초가 되기 때문에 생성형 AI 개발·구현을 위하여 필수적인 요소로 볼 수 있고, AI 실태조사 결과에서도 많은 사업자들이 아래와 같이 기반모델의 중요성에 대하여 의견을 제시하였다.

표 25. AI 실태조사 결과 중 기반모델의 중요성 응답

주요 의견
<ul style="list-style-type: none"> • 기반모델은 다양한 AI 서비스의 기반이 되고, 기반모델의 성능이 AI 서비스 품질을 결정짓는 중요 요인 중 하나에 해당함 • 기반모델이 소버린(Sovereign) AI⁴⁸ 관점의 대응 및 자체 기술력 등에 가장 큰 영향을 미침

98. 사업자는 기반모델을 직접 구축할 수도 있고, 한편으로는 이를 직접 구축하지 않고 제3자가 구축한 기반모델을 사용할 수도 있으며, 어떠한 방식을 선택할지는 각 사업자의 역량이나 사업전략⁴⁹ 등에 달려 있다.

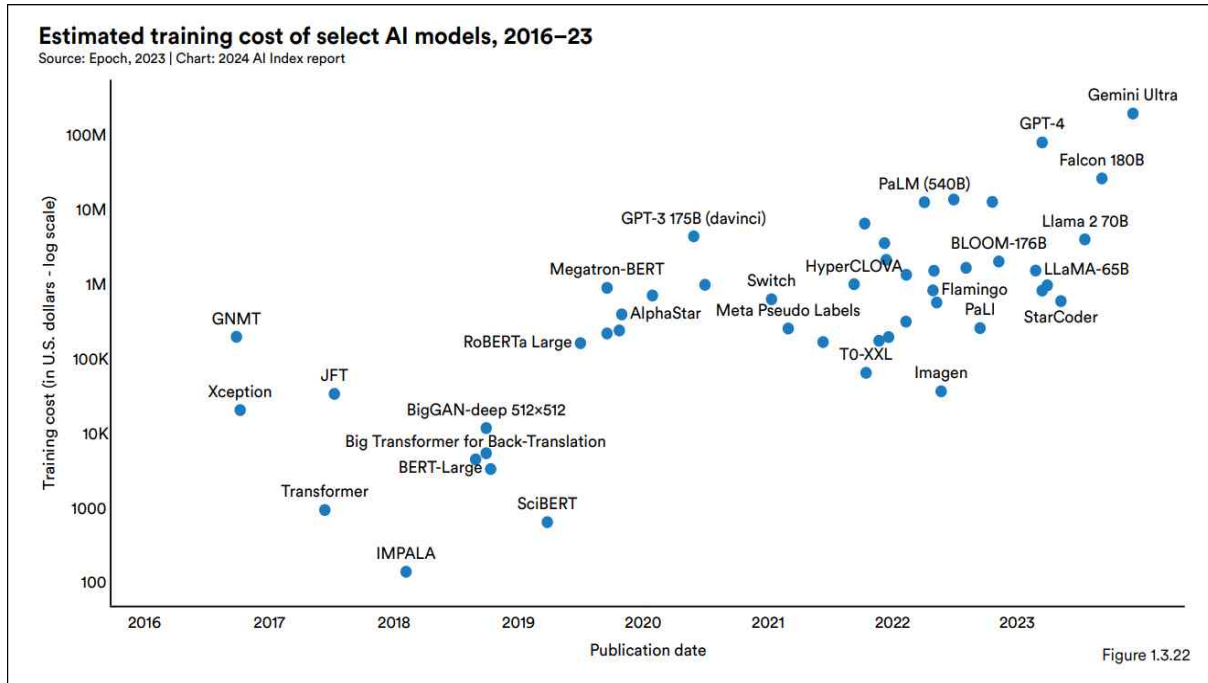
99. 기반모델을 직접 구축하기 위해서는 앞서 살펴본 컴퓨팅 자원, 데이터 및 전문인력 등 AI 인프라를 충분하게 갖추어야 하고 그에 따른 투자 비용도 상당히 지출하여야 하므로, 자금력 등이 풍부한 사업자가 기반모델을 직접 구축하는 것이 상대적으로 수월할 수 있다.

100. HAI 분석에 따르면, 다음과 같이 2016년부터 2023년 동안 기반모델 훈련비용이 급격하게 증가하였고, Gemini Ultra 및 GPT-4 모델 훈련 비용이 각각 약 1억 9천만 달러, 약 7천 8백만 달러가 소요된 것으로 추정되는 등 기반모델 훈련 등 구축 과정에는 상당한 비용이 투입되는 것을 알 수 있다.

48 국가·기업이 자체 인프라 등을 활용하여 독립적인 AI 역량을 구축하는 것을 말함

49 자체 기반모델 개발을 진행하던 도중 타사 기반모델 활용으로 전환하는 사례도 있음

표 26. 주요 기반모델 훈련비용(추정치)



* 출처: HAI⁵⁰

101. 기반모델을 구축한 사업자(기반모델 개발사)는 기반모델을 내부적으로만 사용할 수도 있으나, 다른 개발사 등 수요자에게 자신의 기반모델을 배포하여 이를 사용하게 할 수도 있으며, 배포 방식(유·무료 여부, 개방 여부 등)은 사업상 자율적 판단에 따라 결정할 수 있다.

102. 기반모델을 무료로 배포하는 경우에는 수익 실현은 어려울 수 있으나 개발자를 포함한 많은 이용자를 확보하여 이를 토대로 기반모델 성능을 보다 향상시킬 수도 있다. 한편, 유료 배포의 경우에는 그 수익을 통해 AI 인프라 등에 대한 투자 비용 일부를 회수할 수 있을 것이나, 이용자를 확보하기 위해서는 기반모델의 성능을 높이고 상당한 인지도를 확보할 필요가 있다.

103. 또한, 기반모델 개발사는 기반모델을 개방형(오픈소스, Open Source) 또는 폐쇄형(클로즈드 소스, Closed Source)으로 배포할 수도 있다.

50 HAI(2024), Artificial Intelligence Index Report 2024

104. 개방형은 기반모델의 코드와 매개변수 등을 공개하는 방식, 폐쇄형은 기반모델 또는 세부 정보 및 훈련 방법 등을 공개하지 않은 방식이며, AI 실태조사에서 사업자들은 개방형·폐쇄형 기반모델 사용에 따른 장단점을 아래와 같이 응답하였다.

표 27. 개방형·폐쇄형 기반모델 사용에 따른 장점·단점

구분	장점	단점
개방형	<ul style="list-style-type: none"> · 무료 또는 저렴한 가격 · 기반모델 조정이 상대적 용이 	<ul style="list-style-type: none"> · 상대적으로 낮은 성능 · 공식적 기술 지원에 한계
폐쇄형	<ul style="list-style-type: none"> · 상대적으로 더 좋은 성능 · 기반모델의 안정성·신뢰성 	<ul style="list-style-type: none"> · 비용적 부담 · 조정 및 성능 최적화 어려움

105. 아울러, 기반모델 개발사는 앞서 살펴본 CSP와의 파트너십 등을 통해 해당 클라우드 서비스 플랫폼에서 자신의 기반모델을 배포할 수 있으며, 이는 이용자를 보다 용이하게 확보할 수 있는 방법이 될 수 있다.

106. 한편, 사업자들은 기반모델 구축에 관한 기술적·재정적 문제 또는 사업상 판단 등에 따라 기반모델을 직접 구축하지 않고 다른 사업자의 기반모델을 사용할 수도 있으며, 그 방식에는 미세조정, API⁵¹ 호출 등 다양한 방식이 있을 수 있다.

107. 미세조정(Fine-Tuning)은 일정 분야에 특화된 데이터셋을 구축하고, 이를 활용하여 기반모델을 학습시킴으로써, 기반모델을 특정한 목적 등에 맞게 조정하는 것을 말한다. 예를 들면, 텍스트 생성 기반모델에 컴퓨터 프로그래밍 코드 관련 데이터 등을 학습시켜 해당 기반모델이 프로그래밍 코드에 대한 답변 내용을 생성할 수 있도록 조정하여 새로운 모델을 구축할 수 있다.

51 Application Programming Interface의 약자로서 소프트웨어 어플리케이션 간 통신하여 데이터 등을 교환할 수 있도록 하는 기술 등을 말하며, 예를 들면 스마트폰 내 날씨 어플리케이션이 API를 통해 기상청의 날씨 정보를 받아와 어플리케이션 화면에 표기하는 것 등이 있음

108. 이러한 미세조정은 기반모델을 직접 구축하는 것보다 비용이 적게 소요될 수는 있으나 기반모델의 가중치를 조정할 수 있는 접근 권한을 가지고 있어야 하며, 특정 작업을 수행할 수 있도록 조정하기 위해서는 추가로 학습시켜야 하는 데이터의 품질과 고유성이 더욱 중요해질 수 있다⁵².

109. API 호출 방식은 사업자가 특정한 기반모델에 데이터 등을 보내고 그 응답을 받아 사용하는 방식을 의미하며, 이 또한 기반모델을 직접 구축하는 것보다 비용이 적게 들 수 있다. 다만, 기반모델의 결과값 생성·출력 등에 대한 API 접근만 제한적으로 제공되기 때문에 그 유연성이 제한될 수 있고, 해당 기반모델 개발사에 의존하게 될 수도 있다는 우려도 존재한다⁵³.

52 OECD(2024), Artificial Intelligence, Data and Competition

53 CMA(2023), AI Foundation Models Initial Report

2.2.2.6 AI 서비스

110. AI 서비스는 기반모델을 활용하여 AI 기능을 구현하고 이를 개인·기업 등 수요자에게 제공하는 것을 말하며, 별도의 AI 어플리케이션(Application) 구축 또는 다른 어플리케이션 내 AI 기능 추가 등 여러 형태로 나타날 수 있다.

111. AI 서비스를 구분하는 일률적인 기준은 없지만 ‘AI 에이전트’와 같이 기반모델을 활용한 광범위한 작업 등을 수행하는 수평적 AI(Horizontal AI) 서비스 및 특정 산업에서 문제를 해결하기 위하여 맞춤형으로 제공되는 수직적 AI(Vertical AI) 서비스 등으로 구분하는 견해도 있다⁵⁴.

112. AI 서비스는 인프라 구축·개발 단계를 통해 구축되는 결과물로서 최종 수요자와 직접 맞닿아 있고 상당한 부가가치와 효용 등을 창출할 수 있기 때문에 가치사슬에서 중요한 의미를 지니고 있다.

113. AI 실태조사 결과에서도 사업자들은 AI 서비스의 중요성에 대하여 아래와 같이 응답하였다.

표 28. AI 실태조사 결과 중 AI 서비스의 중요성 의견

주요 의견
<ul style="list-style-type: none">· 고성능 AI 기술을 특정 작업에 적용할 수 있도록 하는 등 사용자에게 가치를 전달하는 핵심 역할· 광범위한 영역에서 활용할 수 있는 생성형 AI 기술 기반의 AI 서비스 등은 각 산업 분야의 경쟁력 증대를 위하여 필수적

114. 한편, AI 서비스 제공사는 주로 자신이 마련한 약관 등 계약 내용에 근거하여 서비스를 제공하며, 해당 약관에는 데이터 수집·학습에 관한 사항을 포함하는 사례도 나타나고 있다.

54 한국지능정보사회진흥원(2024), 버티컬 AI로의 변화 및 과제

2.3 국내 생성형 AI의 시장구조 및 현황⁵⁵

2.3.1 가치사슬 관련 시장 구조

115. AI 반도체, 클라우드 컴퓨팅 및 기반모델 등 가치사슬 주요 분야별로 국내 시장에는 글로벌 빅테크, 대기업, 스타트업 등 다양한 사업자들이 참여하고 있으며, 상이한 경쟁구도를 형성하고 있다.

116. 가치사슬은 AI 반도체·클라우드 컴퓨팅·데이터 등 핵심 인프라와 기반모델(조정모델 포함) 개발 등을 포함한 ‘상류 부문(Upstream part)’ 과 기반모델을 활용한 AI 서비스 제공을 포함하는 ‘하류 부문(Downstream part)’ 으로 구분할 수 있으며, 부문별로 주된 시장참여자가 다르게 나타나고 있다.

117. 가치사슬 상류 부문의 사업을 영위하기 위해서는 상당한 자본력과 기술력 등이 요구되기 때문에 이를 갖추고 있는 글로벌 빅테크 및 국내 대기업 등을 중심으로 관련시장에 참여하고 있는 것으로 나타난다⁵⁶.

118. 가치사슬 하류 부문에서는 타 사업자의 기반 모델을 활용하여 AI 서비스를 개발할 수도 있으므로, 글로벌 빅테크 및 국내 대기업 외 상대적으로 규모가 작은 스타트업 사업자들도 다수 참여하고 있는 것으로 나타난다.

119. 생성형 AI 시장에서 주목할 만한 구조적 특징은 각 사업자들이 일부 가치사슬 분야에만 진출하여 거래관계를 형성하고 있는 경우도 있지만, 주요 빅테크 사업자들이 전 분야에 진출하여 수직통합을 구축하고 있다는 것이다.

55 앞서 살펴보았던 6개의 가치사슬(AI 반도체, 클라우드 컴퓨팅, 데이터, 전문인력, 기반모델, AI 서비스) 중 관련시장이 상대적으로 더 형성되어 있는 4개 가치사슬(AI 반도체, 클라우드 컴퓨팅, 기반모델, AI 서비스) 위주로 기술하였음

56 AI 반도체, 기반모델 등 가치사슬 상류 부문에 스타트업 같은 소규모 회사들이 전혀 참여하고 있지 않은 것은 아니며, 이하 후술하는 바와 같이 가치사슬 상류 부문에서는 상대적으로 대규모 사업자들 위주로 경쟁하고 있는 상황

120. 이는 글로벌 빅테크 등 대규모 사업자들이 이미 클라우드 등 일부 분야에 진출해 있었고 상당한 자본력과 기술도 갖추고 있어, 생성형 AI 개발과 구현 과정에 요구되는 필수요소를 확보하거나, 관련 투자를 진행하기에 상대적으로 수월하기 때문인 것으로 볼 수 있다.

121. 아울러, 생성형 AI 생태계가 기반모델을 중심으로 인접시장에서 연계성이 높은 서비스가 제공된다는 점을 고려할 때, 수직통합을 구축한 사업자는 어느 하나의 제품을 제공하면서 다른 제품 등과 결합하고자 하는 유인을 갖게 될 뿐만 아니라 그 제공도 용이하다고 할 수 있다.

122. 그에 따라 생성형 AI 시장에서 수직적으로 통합되어 전체 가치사슬에 걸쳐 사업을 영위하고 있는 사업자는 구글, 마이크로소프트, 아마존, 메타 등 주요 빅테크이며, 일부 회사들은 ‘다른 AI 사업자에 AI 인프라 등 필수요소를 제공하는 공급자’ 지위에 있으면서 동시에 ‘AI 서비스 등을 개발·제공하는 사업자에 대한 경쟁자’ 지위도 구축하고 있다고 볼 수 있다.

표 29. 주요 사업자별 생성형 AI 가치사슬 진출 현황

	Google	Microsoft	amazon	Meta	NAVER	kakao
AI 반도체	TPU	Maia	Inferentia	Meta MTIA		
클라우드 컴퓨팅	Google Cloud	Azure	amazon webservices		NAVER CLOUD PLATFORM	kakaoicloud
기반모델	Gemini	Phi-3	Amazon Titan	Meta LLaMA	HyperCLOVAX	kanana
AI 서비스	Gemini Nano/Pro	Copilot	alexa	Meta AI	CLOVAX	kanana

123. 주요 수직통합 사업자를 구체적으로 살펴보면, 구글은 AI 반도체 ‘TPU(Tensor Processing Unit)’를 개발하고 이를 자사 클라우드 인프라로 활용하여 AI 개발사 등에 컴퓨팅 자원을 제공하며, 한편으로는 AI 개발사로서 자사 기반모델 Gemini 등을 활용한 AI 서비스를 제공하고 있다.

124. 마이크로소프트도 ‘Maia 100’ 등 AI 반도체를 개발하고 자사 클라우드 인프라로 활용하여 AI 개발사 등에 컴퓨팅 자원을 제공하면서, 스스로도 AI 개발사로서 Phi-3 등 기반모델을 구축하기도 하고, AI 서비스 Copilot도 제공하고 있다.

125. 이와 유사하게, 아마존도 ‘Inferentia’ 등 AI 반도체를 자체 개발하고 자사 클라우드 인프라로 활용하여 AI 개발사 등에게 컴퓨팅 자원을 제공하면서, 자체 기반모델 Amazon titan을 구축하거나 AI 서비스로 Alexa를 출시하려고 하는 등 AI 개발사에 대한 경쟁사 지위에도 있다고 볼 수 있다.

126. 국내에서 대표적인 생성형 AI 사업자 네이버는 자체 AI 반도체 사업을 영위하고 있지 않아⁵⁷ AI 가치사슬 전체를 수직 통합한 것은 아니나, 클라우드 사업(네이버클라우드)에 진출하여 컴퓨팅 인프라를 구축하고, 자사 기반모델인 ‘하이퍼클로바 X’를 개발하여 제3자에게 배포하는 한편, 이를 활용한 AI 서비스 ‘클로바 X’도 제공하고 있다.

127. 카카오도 자체 AI 반도체 사업을 영위하고 있지는 않으나, 클라우드 컴퓨팅 사업을 영위하는 계열회사 카카오엔터프라이즈社를 통해 AI 개발사 등에게 컴퓨팅 자원을 제공하고 있으며, ‘카나나’라는 통합 AI 기반모델⁵⁸ 및 서비스 등 출시를 준비하고 있는 것으로 확인된다.

57 다만, 네이버는 삼성전자와 협력하여 AI 반도체 개발을 추진하였던 것으로 확인됨

58 카카오는 KoGPT, Karlo 모델도 구축하였고, 이를 활용한 AI 서비스(Karlo AI 프로필 특채널 등)를 출시하였으나 2024년 10월 모든 서비스를 종료

2.3.2 시장 경쟁상황 분석

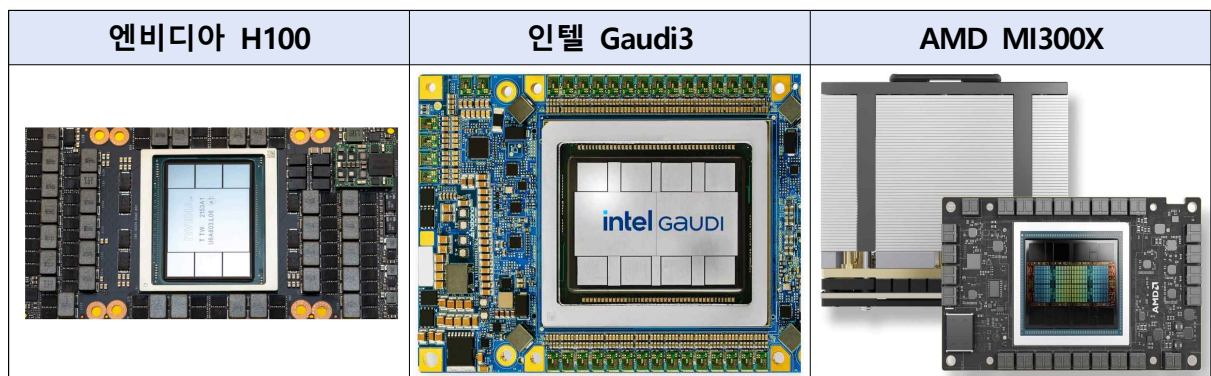
2.3.2.1 AI 반도체

2.3.2.1.1 시장 상황

128. 국내·외 많은 회사들이 GPU·NPU 등 AI 반도체를 개발하여 AI 가속기 등의 제품을 공급하고 있으며, 엔비디아·인텔·AMD 등 글로벌 사업자와 사피온코리아·리벨리온·퓨리오사AI 등 국내 사업자가 주요 사업자로써 시장에 참여하고 있다.

129. 엔비디아, 인텔, AMD는 각각 아래와 같이 AI 반도체가 탑재된 AI 가속기 등 제품을 출시·공급하고 있으며, AI 실태조사 결과에 따르면, 엔비디아 등의 AI 가속기 제품⁵⁹은 HPE, Dell 등 서버 제조사의 서버 제품에 탑재되어 국내 AI 개발사 등 수요자에게 공급되는 것으로 확인된다.

표 30. 엔비디아, 인텔, AMD 각 사 제품



* 출처: CNET, AMD, 인텔 사이트 참조

59 소프트웨어 정책연구소의 2023년 인공지능산업 실태조사 결과보고서에 따르면, 국내 사업자들이 엔비디아의 AI 가속기 제품 중 H100을 약 1,961개 보유하고 있는 것으로 조사되었으며, 해당 조사는 인공지능산업 생태계 조성 및 산업 활성화 지원 정책 수립을 위한 기초자료 확보를 위해 2023년 9월~11월 인공지능산업 관련 2,354개 사업자를 대상으로 실시되었다.

130. 한편, 엔비디아, 인텔, AMD의 전세계 데이터센터 부문⁶⁰ 매출은 아래와 같으며, 엔비디아의 매출이 2023년에 전년 대비 크게 증가하였고, 다른 사업자에 비해서도 높게 나타나고 있다.

표 31. 글로벌 3개 회사의 전세계 데이터센터 부문 매출액

(단위: 백만 달러)

구분	엔비디아**	인텔	AMD
2022년	15,005	19,445	6,043
2023년	47,525	15,521	6,496

* 출처: 각 사 사업보고서

** 엔비디아는 1월 결산법인으로서, 2022.2월~2023.1월 매출 및 2023.2월~2024.1월 매출을 2022년 및 2023년 매출로 기재함

131. 국내 사업자인 사피온코리아, 리벨리온, 퓨리오사AI도 아래와 같이 AI 반도체가 탑재된 AI 가속기 제품을 출시하여 수요자에게 공급하고 있는 것으로 나타났다.⁶¹

표 32. 사피온코리아, 리벨리온, 퓨리오사AI 각 사 제품



* 출처: 사피온코리아, 리벨리온, 퓨리오사AI 각 사 사이트

60 AI 반도체 및 가속기 등 제품이 CSP 등의 데이터센터 인프라로써 활용되기 때문에 데이터센터 부문 매출액을 기재하였으며, 해당 매출액에는 AI 반도체 및 가속기 외 다른 제품 매출도 포함되어 있을 수 있음

61 소프트웨어정책연구소(2023), 생성 AI 산업 생태계 현황과 과제

132. 구체적으로 살펴보면, 사피온코리아는 X330 제품에 대하여 글로벌 서버 제조사인 슈퍼마이크로컴퓨터의 서버 적격성 평가를 통과하여 해당 서버에 장착될 수 있는 제품으로 검증받았으며⁶², 현재 리벨리온과 합병⁶³도 진행하고 있는 것으로 확인된다.

133. 리벨리온은 국내 CSP인 KT클라우드와 협력하여 ATOM을 KT클라우드 인프라에 활용하고, 삼성전자와 협력하여 새로운 AI 반도체 ‘REBEL’ 개발을 추진하고 있는 것으로 확인된다.

134. 퓨리오사AI는 카카오엔터프라이즈와 협력하여 워보이 제품을 카카오 클라우드 인프라로 활용하고, LG AI연구원의 기반모델(EXAONE)의 기술 검증 과정에 자사 2세대 반도체인 ‘레니게이드’를 활용했던 것으로 확인된다.

135. 위의 국내 3개 사업자들의 제품 매출액은 아래와 같으며, 2023년 기준 사업자 간 매출액 규모는 큰 차이가 없으나, 향후 사피온코리아와 리벨리온 간 합병에 따라 매출액 규모 차이 발생 등 시장 상황의 변동 가능성이 예측된다.

표 33. 국내 3개 회사의 제품 매출액

(단위: 백만 원)

구분	사피온코리아	리벨리온	퓨리오사AI
2022년	2,472	-	-
2023년	2,447	2,526	1,954

* 출처: 나이스비즈라인, 각 사 감사보고서

62 사피온코리아(2024), 사피온 AI 반도체 ‘X330’, 슈퍼마이크로 서버 적격성 평가 완료

63 사피온코리아와 리벨리온은 AI 반도체 역량 결집을 통한 경쟁력 강화 및 향후 AI 반도체 개발 성공 가능성 제고 등을 목적으로 2024년 8월 합병 계약을 체결하였으며, 사피온코리아가 리벨리온을 흡수합병하고 그 상호를 리벨리온으로 변경할 예정인 것으로 확인됨

136. 앞서 서술한 바와 같이, 최근에는 AI 반도체 수요자인 구글·마이크로소프트·아마존웹서비스 등 일부 회사들이 최근 자체 AI 반도체를 개발하여 자사 클라우드의 인프라로 활용하는 등 제조사에 대한 의존을 줄이려고 노력하고 있다. AI 반도체 수요 사업자들이 자체 AI 반도체 개발에 착수한 것은 전세계적인 생성형 AI 수요 증가에 따라 AI 반도체 공급이 부족해지고 품귀 현상이 발생하였기 때문이라는 의견도 존재한다⁶⁴.

표 34. 구글, 마이크로소프트, 아마존웹서비스 자체 AI 반도체

구글 TPU	마이크로소프트 Maia	아마존웹서비스 Inferentia
		

* 출처: 한국과학기술평가원⁶⁵

137. 주요 글로벌 사업자들의 AI 가속기가 국내 시장에서 얼마나 판매되었는지 명확하게 확인되지 않아⁶⁶ 국내 시장에서 주요 AI 반도체 사업자별 시장점유율 등을 구체적으로 산정하기는 곤란하다. 하지만, AI 실태조사 결과 사업자의 의견·해외 경쟁당국의 평가·시장조사기관의 분석 등에 따르면 엔비디아의 AI 반도체·가속기 제품이 관련시장에서 상당히 높은 선호도와 점유율을 가지고 있는 것으로 볼 수 있다.

64 Autorité de la concurrence(2024), On the competitive functioning of the generative artificial intelligence sector

65 한국과학기술평가원(2024), AI 반도체 기술개발 현황과 시사점: 클라우드 산업을 중심으로

66 엔비디아의 AI 가속기 등 제품은 주로 서버 제조사 및 중간 유통사 등을 통해 국내 수요자에게 공급되기 때문에 글로벌 사업자들의 데이터센터 부문 매출액 중 국내로 유통된 제품 매출액을 명확하게 확인하기 곤란함

AI 실태조사 결과, 엔비디아는 자신의 AI 가속기 등 제품이 주로 서버 제조사 및 중간 유통사 등을 통해 국내 최종 수요자에게 공급되기 때문에 국내 최종 수요자에 유통된 자신의 제품 매출액 등을 알지 못한다고 설명하였음

138. 먼저, AI 실태조사 결과에 따르면 AI 반도체 회사 및 AI 개발사 등 많은 사업자들이 아래와 같이 엔비디아가 시장에서 상당히 높은 점유율을 가지고 있다는 의견을 제시한 바 있다.

표 35. AI 실태조사 결과 중 엔비디아 점유율 관련 의견

구분	주요 의견
반도체 회사 의견	<ul style="list-style-type: none"> · 국내 시장에서 엔비디아가 독점적 지위를 유지하고 있고, 이에 대항하여 국내 업체들이 노력하고 있음 · 엔비디아가 절대적인 시장점유율, 업계 인지도를 보유하고 있고, 국내 업체는 낮은 인지도 등 약점 극복 필요
AI 개발사 등 의견	<ul style="list-style-type: none"> · 엔비디아 GPU의 점유율이 95% 이상으로 높게 나타나며, 특히 학습 환경의 경우 엔비디아 쓸림 현상이 더욱 심함 · 시장이 엔비디아 독점 형태로 수렴하고 있으며, 엔비디아 GPU에 대한 글로벌 대기업 간 경쟁적 수주 상황

139. 한편, 영국·프랑스 등 해외 경쟁당국도 AI 시장 보고서에서 아래와 같이 엔비디아가 관련시장에서 상당한 지위를 가지고 있는 것으로 검토하기도 하였다.

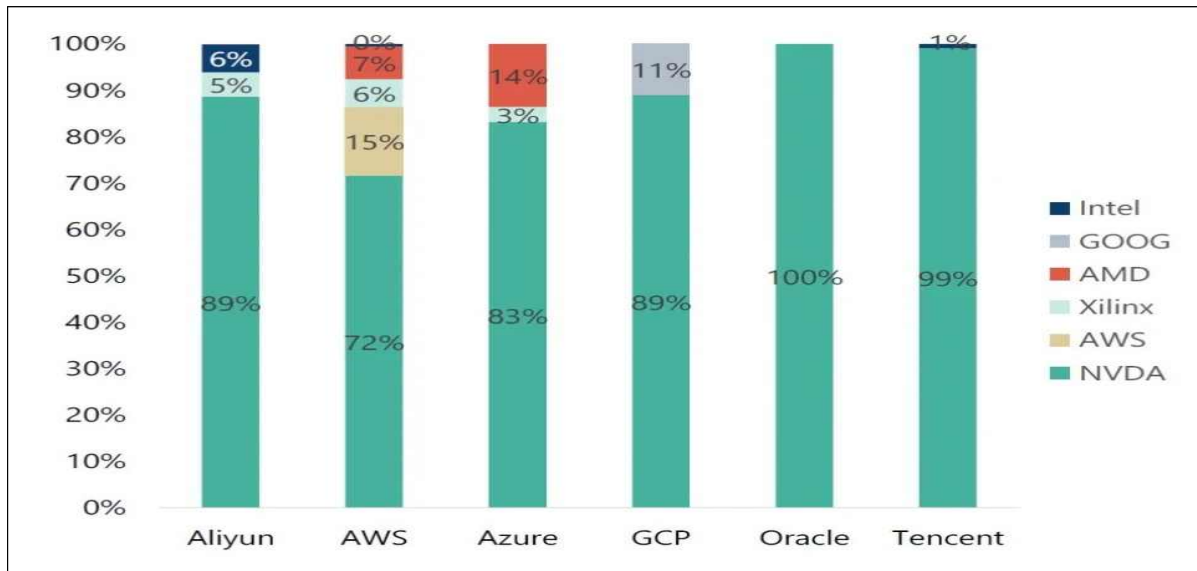
표 36. 해외 경쟁당국 보고서 내용(발췌)

구분	보고서 내용
영국	A few businesses produce the necessary inputs, creating dependencies for startups and other companies developing and deploying AI tools. For example, NVIDIA currently leads the market in AI accelerator chips , and Google provides its own TPUs only on GCP.
프랑스	It is clear from the above developments (see paragraphs 124 et seq.) that a single operator, Nvidia , appears to have a dominant position in the sector for the IT components needed to train foundation models.

* 출처: CMA, Autorité de la concurrence

140. 또한, 시장조사기관 리프터 인사이트(Liftr Insights)에 따르면, AI 가속기의 주요 수요자인 아마존웹서비스·마이크로소프트 등 주요 CSP가 자사 클라우드 인프라로 엔비디아 GPU를 다수 활용한 것으로 확인된다.

표 37. 2022년 6월 기준 주요 CSP별 클라우드 인프라 비중



* 출처: 한국과학기술평가원⁶⁷

141. 엔비디아가 높은 시장점유율 및 선호도 등을 갖게 된 배경에 대해서는 여러 의견이 있으나, 주요 의견으로는 엔비디아가 지난 2006년 무료로 배포한 AI 프로그래밍 지원 도구(CUDA⁶⁸) 등을 통해 관련시장에 먼저 진입함에 따른 것이라는 의견이 존재한다. 구체적으로, 많은 개발자들이 엔비디아 GPU 및 CUDA를 활용하여 AI 개발 등을 진행하였고, 이 과정에서 경험과 사용의 편리성이 축적되면서 엔비디아 GPU와 CUDA에 대한 선호도가 높아지고 엔비디아가 시장에서 상당한 입지를 구축할 수 있었다는 의견으로 확인된다.

67 Liftr Insights, 한국과학기술평가원(2024), AI 반도체 기술개발 현황과 시사점: 클라우드 산업을 중심으로 재인용

68 CUDA는 Compute Unified Device Architecture의 약자이며, 엔비디아 사업보고서에 따르면 전세계적으로 약 470만 명 이상의 개발자들이 이를 사용하는 것으로 나타남

2.3.2.1.2 경쟁상황 평가

142. 현재 국내 AI 반도체 시장에서 여러 사업자들이 AI 반도체를 개발·출시하고 있으나, 엔비디아 제품에 대한 높은 수요·선호도와 관련시장 내 엔비디아 지위 등 제반 사정을 고려할 때 엔비디아의 독주 체제 안에서 글로벌 및 국내 사업자가 돌파구를 모색 중인 상황으로 평가할 수 있다.

143. AI 실태조사 결과에 따르면, 엔비디아의 상당한 지위 등으로 인하여 경쟁상황을 비경쟁적으로 인식하는 의견이 제시되었고, 한편으로는 성능 향상 등을 위하여 사업자들이 치열하게 경쟁하고 있다는 의견도 제시되었다.

표 38. AI 실태조사 결과 중 경쟁상황 등 주요 의견

구분		주요 의견
경쟁상황	경쟁적	· 국내 기업도 기술 발전을 통한 경쟁 참여 · 효율적 반도체 구축을 위하여 치열한 경쟁
	비경쟁적	· 엔비디아가 상당한 지위를 차지하여 비경쟁적 · 제조·공급 가능 업체가 한정적
경쟁사 고객 유인	용이함	· 범용성, 전력 효율성 등 차별화 통해 고객 유인 · 충분한 기술 완성도 갖출 경우 경쟁사 대체 용이
	어려움	· 경쟁사보다 높은 성능·품질 및 가격경쟁력 필요 · 엔비디아 HW 및 CUDA에 높은 종속성

144. 현재 아마존웹서비스, 마이크로소프트 등 일부 CSP가 자체 AI 반도체를 개발하여 클라우드 인프라에 활용하고 있고 AI 반도체 기술도 계속 발전하고 있는 상황을 고려할 때, 향후 엔비디아 집중 현상이 완화되어 시장 경쟁상황이 변동될 수 있는지에 대하여 주의 깊게 살펴볼 필요가 있다. 컴퓨팅 자원을 AI 반도체 제조사인 엔비디아가 아닌 CSP를 통해 공급받게 된다면 엔비디아 집중이 완화될 여지가 있으며, 향후 기술 발전 등을 통해 수요자의 요구사항에 특화되어 보다 효율적인 AI 반도체를 개발·공급하게 되는 경우에도 시장 경쟁상황이 변동될 수도 있을 것이다.

2.3.2.2 클라우드 컴퓨팅

2.3.2.2.1 시장 상황

145. 클라우드 컴퓨팅 서비스를 제공하는 사업자는 국내·외 다수 사업자가 있으며, 아마존웹서비스·마이크로소프트·구글 등 글로벌 CSP와 네이버 클라우드·KT클라우드·NHN클라우드 등 국내 CSP가 시장에 참여하여 경쟁하고 있는 것으로 나타난다.

146. 먼저, 각 CSP별 사업 내용 등에 대하여 간략히 살펴보면, 아마존 웹서비스, 마이크로소프트, 구글은 클라우드 컴퓨팅을 제공함과 아울러 자사 클라우드 플랫폼에서 수요자들에게 특정 기반모델에 대한 접근·사용 기능 등을 아래와 같이 제공⁶⁹하고 있는 것으로 확인된다.

표 39. 글로벌 3개 사업자들의 주요 서비스

회사	주요 서비스 명칭	제공 기반모델
아마존웹서비스 (Amazon Web Services)	베드록 (Bedrock)	Amazon Titan, LLaMa2 등
마이크로소프트 (MS Azure)	애저 AI 스튜디오 (Azure AI Studio)	Phi-3, GPT-4o 등
구글 (Google Cloud Platform)	버텍스 AI (Vertex AI)	Gemini, Imagen 등

* 출처: 각 사 사이트

147. 아울러, 위 3개 CSP는 엔비디아와 계약을 체결하여 각 사 클라우드에서 엔비디아 GPU 기반의 DGX 클라우드 서비스도 제공할 수 있는 등 생성형 AI 관련 클라우드 서비스 사업을 확장하고 있는 것으로 나타났다. 예를 들어, 아마존웹서비스는 아래와 같이 엔비디아와의 다양한 협력을 추진하고 있다.

69 이를 MaaS(Model as a Service)라고 칭하기도 함

표 40. 아마존웹서비스의 DGX 클라우드 제공 관련 설명

AWS and NVIDIA Extend Collaboration to Advance Generative AI Innovation

- **AWS** to offer NVIDIA **Grace Blackwell GPU-based Amazon EC2 instances and NVIDIA DGX Cloud** to accelerate performance of building and running inference on multi-trillion-parameter LLMs
- Project Ceiba — an AI supercomputer built exclusively on **AWS with DGX Cloud** — to feature 20,736 GB200 Superchips capable of 414 exaflops for NVIDIA's own AI R&D

* 출처: 엔비디아 사이트

148. CSP의 이러한 서비스 범위 확대는 각 CSP의 시장 경쟁력 제고 노력의 일환으로 클라우드 컴퓨팅 시장에서 CSP 간 경쟁을 더욱 촉발시키는 요인으로 작용할 수도 있다. 또 다른 한편으로는, 이와 같은 소수의 빅테크 사업자 간 파트너십 등 쏠림 현상에 대하여 경쟁법적 우려 등을 제기하는 시각도 있다⁷⁰.

149. 위 3개 CSP별 클라우드 서비스 매출액⁷¹은 아래와 같다.

표 41. 글로벌 CSP의 전세계 클라우드 서비스 매출액

(단위: 백만 달러)

구분	아마존웹서비스	마이크로소프트**	구글
2022년	80,096	74,965	26,280
2023년	90,757	87,907	33,088

* 출처: 각 사 사업보고서

** 마이크로소프트는 6월 결산법인으로서, 2021.7월~2022.6월 매출 및 2022.7월~2023.6월 매출을 2022년 및 2023년 매출로 기재함

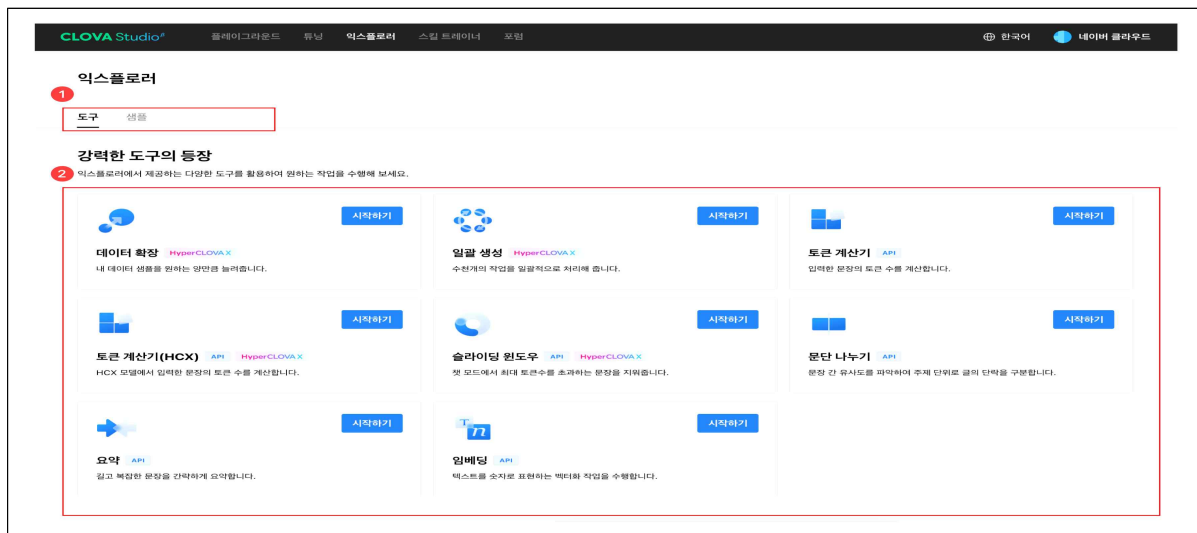
150. 한편, 국내 CSP도 엔비디아 GPU, 국내 AI 반도체 회사의 NPU 등에 기반한 컴퓨팅 자원을 제공하고 있으며, 글로벌 CSP와 같이 AI 기반모델에 대한 접근·사용 권한을 제공하는 경우는 많지 않은 것으로 확인된다.

70 대규모 수직통합 사업자의 경쟁 우위 등 구조적 경쟁 저해에 대해서는 이하 3. 목차에서 후술

71 각 CSP의 클라우드 사업 전체 매출액 중 AI 컴퓨팅에 한정된 매출액을 따로 확인할 수 없었으며, 그에 따라 이하 매출액은 각 CSP의 클라우드 사업 전체 매출액임

151. 실태조사 대상 국내 사업자 중 네이버클라우드가 서버 인프라를 비롯하여 자사 AI 기반모델인 하이퍼클로바X 기반의 클로바 스튜디오 기능⁷²을 함께 제공할 수 있는 Neurocloud(뉴로클라우드) for HyperCLOVA X 서비스 등을 제공하고 있는 것으로 나타난다.

표 42. 클로바 스튜디오 화면(예시)



* 출처: 클로바 스튜디오 사이트

152. KT클라우드의 대규모의 GPU Cluster를 가상화하여 사용자가 연산을 실행할 때 Cluster 내의 GPU 자원을 동적으로 할당하는 서비스(Hyperscale AI Computing, HAC) 등을 제공하고 있으며, NHN클라우드의 자신의 데이터센터에 엔비디아 H100 제품을 비롯하여 1,000개 이상의 AI 반도체를 갖춘 ‘멀티 AI GPU 팜(Farm)’ 을 구축하였다고 공표하기도 하였다⁷³.

153. 위 국내 3개 CSP의 매출액은 다음과 같으며, 네이버클라우드가 타 CSP에 비하여 상대적으로 높은 매출액 규모를 보이고 있다.

72 클로바 스튜디오는 기업이 보유한 대량의 데이터 셋을 하이퍼클로바 X 언어모델과 결합하여 AI 기반의 특화된 서비스를 만들 수 있는 개발도구를 말함

73 NHN클라우드, AI 인프라 시장 선두 목표로 ‘NHN 클라우드 2.0’ 전략 발표

표 43. 국내 CSP의 클라우드 서비스 매출액

(단위: 백만 원)

구분	네이버클라우드	KT클라우드	NHN클라우드
2022년	990,928	411,462	80,787
2023년	1,048,758	636,318	98,791

* 출처: 각 사 사업보고서

** 네이버클라우드 감사보고서상 IT 서비스 매출, KT클라우드 감사보고서상 서비스 제공 수익, NHN클라우드 감사보고서상 클라우드 매출 금액을 기재

154. 국내 클라우드 서비스 시장에서 AI 개발을 위하여 제공되는 클라우드 컴퓨팅만을 별도로 명확히 측정하기는 곤란하지만⁷⁴, AI 실태조사 결과를 비롯한 클라우드 시장 관련 조사 결과 등에 따르면 아마존웹서비스·마이크로소프트 등 주요 글로벌 CSP가 상당히 높은 점유율을 가지고 있다고 볼 수 있다.

155. 먼저, AI 실태조사 결과에 따르면, 아래와 같이 국내 클라우드 시장에서 글로벌 CSP가 상당한 지위를 가지고 있다는 의견이 다수 확인되었다.

표 44. AI 실태조사 결과 중 CSP 시장점유율 관련 의견

구분	주요 의견
CSP 의견	<ul style="list-style-type: none"> 국내 클라우드 서비스 시장은 해외 CSP의 높은 점유율로 고착화되어 있음
AI 개발사 등 의견	<ul style="list-style-type: none"> AI 관련 클라우드 사업은 기존 클라우드 사업 구조에서의 해외 CSP 등 소수 기업 중심으로 진행 중 해외 업체인 구글, 아마존, MS 3자 경쟁을 하고 있으며, 국내 업체가 해외 업체를 대체하기 쉽지 않은 상황

74 클라우드 서비스 유형은 서버 등 인프라를 제공하는 IaaS(Infrastructure as a Service), 소프트웨어 개발 및 테스트 등을 위한 컴퓨팅 플랫폼을 제공하는 PaaS(Platform as a Service), 각종 소프트웨어를 제공하는 SaaS(Software as a Service) 등으로 구분할 수 있으며, 현재 CSP가 컴퓨팅 자원 외에도 AI 개발도구 등을 제공함에 따라 인프라·플랫폼·소프트웨어 등 어느 범위까지 포함하는지에 따라 주요 CSP의 매출액 등이 달라질 수 있음

156. 아울러, 공정위가 지난 2022년에 실시한 클라우드 실태조사 결과에 따르면, 2021년 기준 주요 CSP별 점유율은 아래와 같이 아마존웹서비스가 상당히 높게 나타나는 것으로 확인되기도 하였다.

표 45. 주요 CSP별 점유율 (IaaS·PaaS 기준)

아마존웹서비스	마이크로소프트	구글	네이버클라우드	소계
62.1%	12.0%	6.5%	7.0%	87.6%

* 출처: 공정위

157. 또한, 과학기술정보통신부의 2023년 부가통신사업 실태조사 결과⁷⁵에서도 아마존웹서비스, 마이크로소프트 등 해외 CSP의 클라우드 서비스를 이용하는 비중이 더 높게 나타나는 것으로 조사되었다.

표 46. 주요 클라우드 서비스 이용 비중

(클라우드 이용 사업자 171개 회사 대상, 복수 응답)

아마존웹서비스	마이크로소프트	구글	네이버클라우드	KT클라우드	NHN클라우드
60.2%	24.0%	19.9%	20.5%	8.2%	7.0%

* 출처: 과학기술정보통신부

158. 영국 경쟁당국도 클라우드 서비스 시장에 대한 조사 결과에서 아마존 웹서비스·마이크로소프트가 상당한 시장지배력을 가진 것으로 분석하였다⁷⁶.

표 47. 영국 경쟁당국 보고서 내용(발췌)

보고서 내용
Based on the evidence we have seen to date, our emerging view is that there are indicators of significant market power being held by the largest two providers, AWS and Microsoft . (후략)

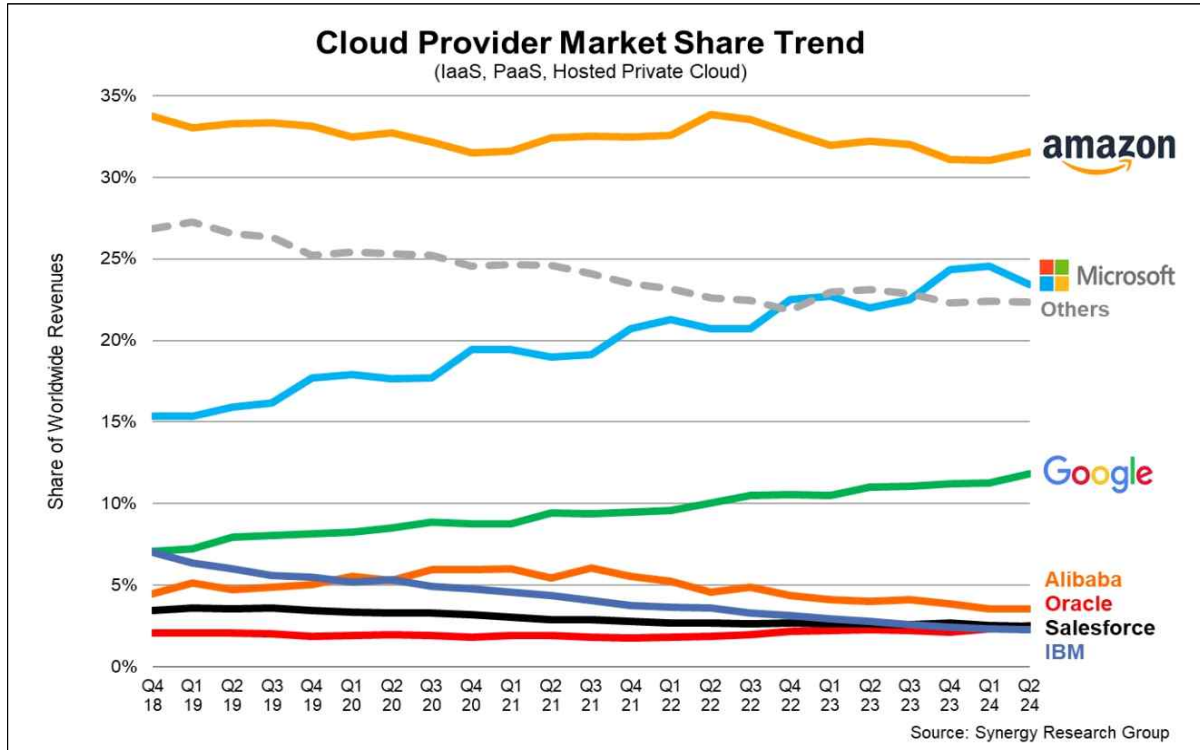
* 출처: CMA

75 2023년 부가통신사업 실태조사는 「전기통신사업법」을 근거로 2023년 10월~12월 실시되었으며, 부가통신사업이란 웹사이트 및 어플리케이션 등을 통해 온라인쇼핑·콘텐츠·SNS·게임·인터넷금융 등을 제공하는 것을 말함

76 CMA(2024), Cloud services market investigation, Competitive landscape working paper

159. 시장조사기관 시너지 리서치 그룹(Synergy research group)의 조사 결과에서도 전세계적으로 아마존웹서비스, 마이크로소프트 등 글로벌 CSP의 점유율이 높게 나타나는 것으로 확인된다.

표 48. 클라우드 서비스 시장점유율 추이



* 출처: Synergy research group⁷⁷

160. 이처럼 아마존웹서비스가 상위 사업자가 될 수 있었던 배경에 대해서는 아마존웹서비스의 클라우드 서비스 시장 선점 효과, 규모의 경제 등에 따른 것이라는 의견⁷⁸이 존재한다. 구체적으로 아마존웹서비스가 2006년 클라우드 서비스를 최초 개시함에 따라 개발자 생태계를 초기에 장악할 수 있었고, 대규모 데이터센터 구축을 통해 낮은 가격에 경쟁력 높은 서비스를 제공하는 등 선순환에 따라 후발주자들과 격차를 벌릴 수 있었다는 것이다.

77 Synergy research group(2024), Cloud Market Growth Stays Strong in Q2 While Amazon, Google and Oracle Nudge Higher

78 삼성증권(2019), 클라우드 전쟁 1조 달러 클럽을 향한 왕좌의 게임

2.3.2.2.2 경쟁상황 평가

161. 현재 국내 AI 시장에서도 글로벌 CSP의 점유율이 높고, 그 경쟁에 필요한 AI 반도체 등 인프라 구축과 AI 기능 확장 등에서 글로벌 CSP가 국내 CSP보다 상대적 우위⁷⁹에 있는 점을 고려할 때, 글로벌 CSP 간 경쟁이 활발한 가운데, 국내 CSP도 그 경쟁에 같이 참여하고 있다고 평가할 수 있다.

162. AI 실태조사 결과에서는 아래와 같이 경쟁상황에 대하여 경쟁적 또는 비경쟁적으로 바라보는 의견이 모두 나타났으며, 비경쟁적이라고 보는 회사는 그 이유에 대하여 글로벌 CSP 위주의 시장 경쟁상황 등을 언급하였다.

표 49. AI 실태조사 결과 중 경쟁상황 등 주요 의견

구분	주요 의견
경쟁적	<ul style="list-style-type: none"> · 온프레미스^{주)} 솔루션 회사, CSP 등의 치열한 경쟁 · 시장 선점을 위한 CSP 간 경쟁적 사업 확대 중
비경쟁적	<ul style="list-style-type: none"> · 해외 주요 클라우드 사업자 위주의 시장 · 규모의 경제를 달성한 빅테크들의 시장경쟁 완료

주) 온프레미스(On-premise)는 회사의 자체 시설에 AI 반도체·가속기 및 서버 등 설비를 갖추는 것 등을 말함

163. 향후 수요자의 다중 클라우드 활성화, AI 반도체 등 인프라 확충 등을 통해 국내 CSP의 경쟁 여건이 개선되거나 글로벌 CSP에 대한 집중도가 완화되는 등 경쟁 상황이 변동될 수 있는지 예의주시할 필요가 있다.

164. 즉, AI 개발사 등이 안정적인 서비스 제공 등을 고려하여 2개 이상의 클라우드 서비스를 병렬적으로 이용할 수도 있으며⁸⁰, 이는 CSP 간 경쟁을 촉발시키는 요인으로 작용할 여지가 있다고 볼 수 있을 것이다.

79 아마존웹서비스·마이크로소프트·구글의 자체적인 AI 반도체 개발, 엔비디아 협력을 통한 인프라 확충 및 클라우드 서비스 내 AI 기반모델 접근·사용 기능 제공 등

80 2023년 부가통신사업 실태조사 결과에 따르면 주요 사업자들의 다중 클라우드 이용 비율이 44.7%로 나타났으며, 과학기술정보통신부가 2024년 7월 클라우드 장애 문제 등을 계기로 일부 사업자들에게 다중화를 권고한 사례도 존재

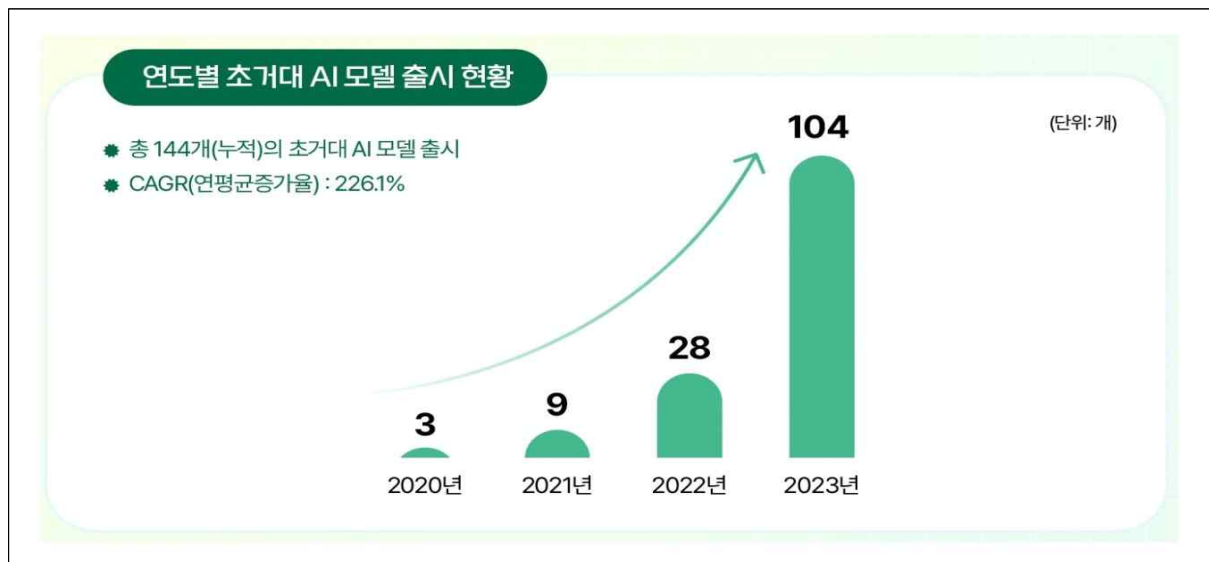
2.3.2.3 기반모델

2.3.2.3.1 시장 상황

165. 기반모델은 앞서 살펴본 바와 같이 AI 인프라를 활용하여 방대한 데이터를 학습한 생성형 AI의 기초가 되는 모델로서, 이를 직접 개발하여 수요자에게 배포·제공하는 기업이 있는 한편, 타 사업자가 개발한 기반모델을 미세조정하여 활용하는 기업도 존재한다.

166. 2020년대에 들어서 기반모델의 대형화 및 초거대 AI 기반모델 개발 시도가 급격하게 증가하고 있으며, 모델 성능이 규모, 학습 데이터 등의 양에 따라 향상된다는 ‘스케일링 법칙(*scaling law*)’도 등장하였다⁸¹. 소프트웨어 정책연구소 분석에 따르면, 초거대 AI 기반모델 출시 개수는 2022~2023년에 급격하게 증가하였던 것으로 확인된다.

표 50. 2020년~2023년 연도별 초거대 AI 모델 출시 현황



* 출처: 소프트웨어 정책연구소(2024), 글로벌 초거대 AI 모델 현황 분석

81 소프트웨어정책연구소(2024), 글로벌 초거대 AI 모델 현황 분석(2020~2023년). 해당 자료에서 초거대 AI 모델은 대용량 연산 인프라를 바탕으로 방대한 데이터를 학습해 인간처럼 종합적인 인지·판단·추론이 가능해진 ‘큰 규모’의 AI 모델로 정의함

167. 현재 기반모델 시장의 주요 사업자를 살펴보면, 글로벌 사업자에게는 구글, 메타, 오픈AI, 마이크로소프트 등이 있고, 국내 사업자에게는 네이버, 카카오, LG, KT 등이 있으며, 각 사업자별 주요 기반모델은 아래와 같다.

표 51. 주요 국내·외 기업별 주요 기반모델 현황

구분	사업자명	주요 모델('24.11월 기준)
해외	메타	LLaMa 3.2
	구글	Gemini 1.5
	오픈AI	GPT-4o
	마이크로소프트	Florence-2
	엔비디아	Nemotron-4
	Anthropic	Claude 3.5 Sonnet
국내	네이버	HyperCLOVA X
	카카오	KoGPT
	LG	EXAONE 3.0
	KT	Mi:dm
	NC소프트	VARCO LLM
	업스테이지	Solar
	솔트룩스	LUXIA GPT
	코난테크놀로지	코난 LLM

168. 대표적 글로벌 사업자인 구글은 자사 클라우드 플랫폼 Vertex AI를 통해 다양한 기반모델을 개발하여 출시하고 있으며, 2024년 2월 영역별로 고도의 추론 및 이해 능력을 갖춘 Gemini 1.5를 출시하고, 온디바이스용 AI Gemini Nano 등 다양한 모델을 출시하였다.

169. 오픈AI는 GPT-1을 시작으로 2024년 5월 최신 언어모델인 GPT-4o를 출시하였고, 2023년 9월 출시한 text-to-image 모델 Dall-E3를 통합하여 챗GPT를 통해 AI 서비스를 제공하고 있다.

170. 국내 사업자인 네이버는 자회사 네이버클라우드를 통해 AI 인프라를 구축하였고, 2021년 자체 개발한 기반모델 HyperCLOVA를 처음 발표하였으며, 2023년에는 HyperCLOVA X를 출시하였다.

171. 카카오도 언어 모델로 KoGPT, 이미지 모델로 Karlo를 구축하여 자사 AI 서비스의 개발 및 구현 용도로 활용하였고, 현재 ‘카나나’라는 명칭의 기반모델 및 AI 서비스 출시를 준비하고 있다.

172. 위와 같이 AI 반도체, 클라우드, 데이터 등 대규모 자원이 필요한 기반모델을 직접 개발·배포하는 사업자 중에는 글로벌 빅테크 및 국내 대기업 등이 있다. 해당 사업자들은 관련 산업에 이미 진출해 있었거나 자본력, 기술 등 역력이 충분하여 진입이 상대적으로 용이하였기 때문으로 보인다.

173. 업스테이지, 솔트룩스, 코난테크놀로지 등 상대적으로 소규모 사업자들도 시장에 참여하고 있으며, 이들은 투자 유치, 코스닥 시장 상장 등을 통해 상당한 자금을 확보하여⁸² AI 기반모델 개발 등을 추진할 수 있었던 것으로 보인다.

174. 한편, 기반모델 개발사의 매출은 기반모델 API 사용에 따른 이용대가 수취, AI 서비스를 도입하고자 하는 사업자의 AI 서비스 구축 요청⁸³에 따른 대가 수취 등을 통해 발생할 수 있으며, 모델 성능 향상 등을 목적으로 대가 수취 없이 기반모델을 제공⁸⁴하기도 한다.

82 솔트룩스는 2020년 7월 코스닥 시장 상장 시 ‘AI 관련 연구개발비, GPU 등 데이터센터 설비 투자 자금’ 등에 사용할 목적으로 약 187억 원 조달, 코난테크놀로지도 2022년 7월 코스닥 시장 상장 시 ‘AI 관련 연구개발비, GPU farm 확충 자금’ 등에 사용할 목적으로 약 300억 원 조달, 업스테이지는 현재 상장하지 않았으나 유상증자 등을 통해 자본 확보

83 기반모델 사업자 코난테크놀로지는 2024년 6월 한국남부발전(주)와 생성형 AI 구축 및 학습 용역 계약을 체결하여 코난테크놀로지의 AI 기반모델 등을 활용한 생성형 AI 서비스를 한국남부발전에 도입하기로 함(출처: 코난테크놀로지, 국내 첫 대형언어모델 공공 도입... 남부발전과 41억대 계약 보도자료)

84 Meta는 자사 기반모델인 LLaMa를 오픈소스로 개방하였는데 이는 많은 AI 개발자와의 협력을 통해 모델 성능을 개선시켜 혁신의 기회를 확대하려는 의도에 따른 것이라는 의견도 존재

175. 또한, 앞서 살펴본 것처럼, 최근에는 기반모델 개발사가 상당한 컴퓨팅 자원과 성능 향상·수익 실현을 위한 이용자 확보를 위해 컴퓨팅 자원을 제공할 수 있고 자사 기반모델을 수요자에게 배포할 수 있도록 도와주는 CSP와 파트너십을 맺는 사례도 나타나고 있다. 대표 사례에는 마이크로소프트(CSP)와 오픈AI(FM) 간 파트너십이 있으며, 이외에도 구글(CSP)과 Mistral AI(FM), 아마존웹서비스(CSP)와 Cohere(FM) 간 파트너십 등이 있는 것으로 확인된다⁸⁵.

176. 기반모델 개발사는 파트너십 등을 통해 자신의 기반모델을 용이하게 배포할 수 있고, 다수의 이용자를 확보할 수 있게 됨에 따라 다른 기반모델 개발사에 비하여 상대적으로 경쟁상 우위를 차지할 가능성도 존재한다.

177. 다만, 기반모델은 과업유형, 성능, 안정성, 매개변수, 학습 연산량, 데이터 규모 등 여러 기준이 있어 모델 간 단순 비교가 어렵고, 생성형 AI가 본격적으로 주목받은 2022년 챗GPT 등장 이후 아직까지 많은 기반모델 기반 서비스가 무료로 제공되고 있다는 점 등을 고려하면 매출액 비교 등을 통해 시장 점유율을 명확히 분석하기는 어려운 상황이다.

178. AI 실태조사 결과에서도 기반모델 시장에 대하여 일률적으로 판단하기 어렵다는 취지의 답변이 아래와 같이 확인되기도 하였다.

표 52. AI 실태조사 결과 중 기반모델 개발사 의견

주요 의견
<ul style="list-style-type: none"> · 기반모델에 대하여는 일원화된 정의가 없고, 이용자·개발자 등은 용례, 상황에 따라 AI모델을 선택...사업자의 통계 리스트나 순위는 빠른 속도로 달라질 수 있음 · 기술 개발이 매우 빠른 속도로 전개되며 시장의 형태도 일률적으로 정의하기 어려운 단계이기 때문에, 안정화된 시장에 적용하는 기준을 적용하기 어려움

85 CMA(2024), Cloud services market investigation

2.3.2.3.2 경쟁상황 평가

179. 현재 국내 AI 기반모델 시장에서 다양한 모델이 출시되고 있고, 사업자들의 모델 인지도 제고, 이용자 확보 등을 위한 성능 향상 노력 등을 고려하면, 기반모델 사업자 간 경쟁은 활성화되어 있다고 평가할 수 있다.

180. 그러나, 해외뿐만 아니라 국내에서도 필수 인프라를 이미 확보하고 있거나 확보가 용이한 빅테크 위주로 시장을 선점하고 있으며, 기반모델 개발과 성능 제고를 위한 인프라 구축, 투자 규모⁸⁶ 등 측면에서는 글로벌 사업자들이 국내 사업자보다 상대적으로 경쟁 우위에 있다고 볼 수 있다.

181. AI 실태조사에서는 초기 시장 특성상 다양한 의견이 제시되었으나, 경쟁상황을 비경쟁적으로 인식하는 사업자는 아래와 같이 글로벌 사업자들의 시장 선점 등을 언급하기도 하였다.

표 53. AI 실태조사 결과 중 경쟁상황 등 의견

구분		주요 의견
경쟁상황	경쟁적	· 다수의 사업자들이 다양한 기반모델을 출시하여 경쟁 심화
	비경쟁적	· 시장 선점한 소수 글로벌 사업자 대비 국내 사업자는 경쟁 열위
진입장벽	높음	· 모델 구축 과정상 기술적 어려움, 자본·인력 등에 대한 많은 투자 필요
	낮음	· 클라우드 인프라 접근성, 기술혁신에 따른 비용 감소, 역동적 자금 투자

182. 한편, 최근 기반모델의 구독 서비스화, CSP 등 컴퓨팅 인프라 사업자와의 파트너십, 모델 개발사의 사업전략 변화⁸⁷ 등이 나타나고 있어 시장 내 경쟁상황은 지속적으로 변할 것으로 예상된다.

86 한국지능정보사회진흥원(2024), 「글로벌 정부·민간 분야 AI 투자 동향 분석」에 따르면 2023년 기준 미국 정부와 민간의 AI 투자 규모는 874.1억 달러로, 전 세계 AI 투자의 62%를 차지

87 국내 주요 초거대 AI 모델을 개발한 KT는 최근 기업들의 맞춤형 sLLM(smaller LLM) 공급 사업으로 선회하는 한편, 거대모델은 마이크로소프트와 협력하여 공동투자를 통해 한국형 LLM을 개발하겠다고 발표

2.3.2.4 AI 서비스

2.3.2.4.1 시장상황

183. AI 서비스는 생성형 AI가 텍스트 이해, 대화, 이미지·영상 분석과 새로운 콘텐츠 생성까지 가능한 수준으로 발전하면서 이러한 기능을 접목하여 일정한 작업을 수행하는 등 다양한 기능을 수요자에게 제공하고 있다.

184. 앞서 언급한 구분에 따르면, 수평적(Horizontal) AI 서비스에는 챗GPT와 같이 거대언어모델 기반의 대화 서비스를 제공하는 것뿐만 아니라 AI 에이전트와 같이 인식, 처리, 행동 등의 과정을 통해 주어진 환경에 대응하는 시스템⁸⁸ 등이 포함된다. 한편, 수직적(Vertical) AI 서비스는 특정 산업 분야의 요구사항과 데이터를 기반으로 학습하여 해당 분야의 전문성과 높은 정확성, 신뢰성을 갖춘 솔루션을 제공하는 형태로 나타난다. 이는 거대 기반모델보다 경량화된 모델을 활용하는 특징이 있으며, 의료·금융·제조 등 다양한 분야에서 구현될 수 있다.

185. AI 서비스를 개발하여 제공하는 유형도 기반모델 개발사가 기반모델로 자사 AI 서비스를 직접 개발·제공하는 유형, 수요자의 요구사항에 따라 AI 서비스를 개발하여 그 소유권을 이전하는 유형, 파트너십을 통하여 AI 공동 개발을 추진하는 유형 등 다양하게 존재한다.

186. 또한, 기존에 제공되던 서비스에 AI 기능을 접목하는 경우도 있으며, 분야별로 수요자의 요구와 특화된 데이터를 기반으로 차별화된 AI 기반 서비스들이 개발되고 있어 일정한 기준으로 AI 서비스 시장을 획정하거나 시장 점유율을 추정하는 것은 한계가 있다.

187. 또한, 시장에서 경쟁력이 있다고 인정되는 AI 서비스이더라도 무료로 제공되는 경우도 있어 사업자의 매출액 산정을 통한 단순 비교, 시장규모 추정에도 현실적인 어려움이 존재한다.

⁸⁸ 정보통신기획평가원(2024), ICT Brief 2024-38호

2.3.2.4.2 경쟁상황 평가

188. AI 서비스 시장은 고비용의 AI 인프라 사업을 직접 영위하지 않더라도 진입할 수 있고, AI 기술을 통해 이익을 창출할 수 있으며, 거대언어 모델을 보유한 해외 사업자 외에도 국내 대기업 및 스타트업 등이 다양한 AI 서비스를 출시하고 있다는 점에서 시장 경쟁이 활발한 상황으로 평가할 수 있다.

189. AI 실태조사 결과에서도 AI 서비스 시장의 전반적 경쟁상황에 대해 응답한 사업자의 90% 이상이 “경쟁적(다소 경쟁적 포함)” 이라고 답변하기도 하였다.

190. 수평적 AI 서비스 대표적 사례인 거대언어모델 기반 대화 서비스에서는 오픈AI를 비롯한 구글, 아마존, 마이크로소프트, 메타 등 해외 사업자와 더불어 네이버, 카카오 등 기반모델을 보유한 국내 사업자도 같이 경쟁하고 있다.

191. 기반모델을 직접 개발하지 않더라도, 타 개발사 모델 활용 또는 협업을 통해 AI 서비스를 개발할 수 있으며, 최근 SK텔레콤이 미국 AI 스타트업 퍼플렉시티(Perplexity)와 파트너십을 통해 AI 검색 서비스를 공동 개발하기로 한 사례, 뤼튼테크놀로지스가 국내 스타트업 사이오닉AI와 협력하여 AI 검색 서비스 고도화를 발표한 사례⁸⁹가 대표적이라고 할 수 있다.

192. 이외에도 리걸테크(Legal Tech) 등과 같이 특정 분야별로 스타트업들도 진출 중이며, 국내 생성형 AI 서비스 시장은 국내 AI 산업이 기반모델 자체 개발에도 불구하고 인프라와 자본력, 인력 등을 다수 보유한 글로벌 기업과의 현실적 격차로 인해 차별화된 AI 서비스로 경쟁력을 확보하려는 전략을 채택하고 있다고 평가할 수 있다.

193. 또한, 최근 주목받는 온디바이스 AI와 같이 기반모델을 활용하는 AI 서비스는 모델의 다양화에 따라 서비스 개발 및 사업 전략도 다양해질 수 있어 향후 시장 경쟁상황은 더욱 더 치열해질 것으로 예상된다.

89 정보통신기획평가원(2024), ICT Brief 2024-32호

3 생성형 AI 산업에서의 경쟁·소비자 쟁점 및 과제

◆ 본 목차에서는 생성형 AI 시장에서의 경쟁에 영향을 미칠 수 있는 구조적 요인을 살펴보고, 시장참여자들의 어떠한 행태가 경쟁저해 및 소비자이익 저해 우려를 불러일으킬 수 있는지 분석한 후, 제도 개선 및 모니터링이 필요한 사안을 제시함

3.1 시장경쟁에 영향을 미치는 구조적 요인⁹⁰

◆ AI 시장의 특성으로 인하여 구조적으로 발생할 수 있는 요인들을 이론적인 틀로 분석하였으며, 각 내용이 법 위반을 반드시 의미하지는 않는다는 점을 명확히 함

3.1.1 자본·기술 집약적 산업에 따른 진입장벽

194. 생성형 AI 개발·구현을 위해서는 컴퓨팅 자원, 데이터, 전문인력, AI 기반모델 등 여러 요소가 필수적으로 요구되며, 각 필수 요소를 갖추기 위해서는 자본 및 기술 등에 대한 상당한 투자가 필요하다.

195. 각 요소별로 살펴보면, 컴퓨팅 자원으로서 AI 반도체의 경우 엔비디아 제품이 시장에서 상당한 수요를 가지고 있어 공급받기까지 시간이 소요되고 그 가격도 고가에 달하는 등 수급이 쉽지 않으며, 이를 대체할 AI 반도체를 직접 개발하는 것 역시 대규모 자본과 기술력이 필요하다고 볼 수 있다.

196. 데이터도 AI 개발의 성패 등을 좌우하는 주요한 요소로, 공공·개방 데이터 학습만으로는 성능 향상 등에 한계가 있을 수 있어 자체 데이터 등이 필요하며, 최근 데이터 수요 증가 및 수익화 추세 등에 따라 그 데이터 확보를 위해서는 상당한 비용을 들일 수밖에 없는 상황이다⁹¹.

90 이는 OECD 보고서(Artificial Intelligence, Data and Competition)에서 논의된 내용을 중심으로 분석

91 한편으로는, AI 모델 및 훈련 방식 개선 등 기술 혁신을 통해 AI 개발을 위한 데이터 의존도가 감소할 수 있다는 의견도 존재(Competition Bureau Canada(2024), Artificial Intelligence and competition, Discussion Paper)

197. 전문인력에 대한 수요도 AI 산업 수요 증가에 따라 함께 높아지고 있고, 우수한 인재 유치를 위해 높은 수준의 급여를 제시해야 하는 등 상황에 따라 전문인력 확보에도 상당한 투자가 필요하다.

198. 기반모델을 구축하는 과정에서는 위 컴퓨팅 자원, 데이터, 전문인력 등이 기본적으로 요구될 뿐만 아니라 모델의 성능 개선과 타 기반모델과의 차별화 등을 위해서는 상당한 기술과 전문성 등이 추가로 요구될 수 있다.⁹²

199. AI 서비스의 경우 시장진입 자체가 다른 요소만큼 어렵지 않을 수는 있으나, 다른 AI 서비스와 차별화되거나 고도화된 서비스 개발을 위해서는 높은 기술력과 전문인력 등이 필요할 수 있으며, 특히, 컴퓨팅 인프라나 기반모델을 직접 보유한 기업과 그렇지 않은 기업 간 투자 비용의 차이 등도 진입장벽으로 작용할 가능성이 있다.

200. 위와 같이 주요 가치사슬에서 많은 자본과 고도의 기술력 등이 요구됨에 따라 생성형 AI 시장에 새로운 사업자의 진입이 용이하지 않을 수 있으며, 이는 경쟁 압력을 저하시키는 구조적 요인에 해당할 수 있다.

201. 다만, 클라우드를 통한 컴퓨팅 자원 제공 또는 기반모델의 오픈소스 배포 등은 생성형 AI 시장에서 진입장벽을 낮추는 요인으로 작용할 수도 있다. 즉, 일부 한계에 대한 우려⁹³도 존재하지만, AI 인프라를 직접 구축하지 않고 CSP로부터 클라우드 서비스 형태로 컴퓨팅 자원을 제공받거나, 기반모델을 자체 구축하지 않고 오픈소스로 배포된 모델을 활용함으로써 비용을 절감하여 AI 서비스를 구현하는 것도 가능하다.

92 OECD(2024), Artificial Intelligence, Data and Competition

93 클라우드 서비스 형태의 컴퓨팅 자원 활용은 특정 CSP에 대한 의존도를 불러일으키거나, 오픈소스 기반모델 활용은 차별화 및 성능 향상의 한계 등으로 이어질 우려도 있음

3.1.2 규모의 경제, 범위의 경제 등에 따른 대규모 사업자의 우위

202. 생성형 AI 생태계에서는 대규모 사업자 및 수직통합을 구축한 사업자가 규모의 경제, 범위의 경제, 네트워크 효과 등을 실현할 가능성이 크고, 그에 따라 해당 사업자들이 효율성 및 경쟁력 등을 더욱 제고할 가능성이 존재한다.

203. 먼저, 규모의 경제(economies of scale)는 생산량·투입량 등이 증가할수록 평균비용 등이 절감되는 것 등을 말하며, 생성형 AI 가치사슬 중 대규모로 자원이 투입되는 분야에서 규모의 경제가 발생할 수 있다. 예를 들면, 클라우드 컴퓨팅 자원을 제공하기 위해서는 AI 반도체, 데이터센터 등에 대한 대규모 초기 투자가 필요하며, 이후 운영비용 등 절감을 통해 규모의 경제를 달성할 수 있다. 또한, 기반모델 개발 과정에도 컴퓨팅 자원과 데이터 등이 대량으로 투입되며, 이와 같이 대규모로 개발된 기반모델을 활용하면 추후 다른 모델을 효율적으로 개발·훈련할 수도 있다⁹⁴.

204. 따라서, 규모의 경제는 컴퓨팅 자원, 데이터 등 AI 인프라에 상당한 투자 비용을 지출할 수 있는 대규모 사업자들에게 나타날 가능성이 상대적으로 더욱 클 수 있다고 볼 수 있다.

205. 다음으로, 범위의 경제(economies of scope)는 한 사업자가 복수의 제품 및 서비스 등을 같이 생산할 때 그 비용을 절감할 수 있는 것 등을 말하며, 생성형 AI 시장은 가치사슬 각 분야가 연계되어 있고 이미 수직통합을 구축한 회사도 존재하는바, 범위의 경제가 발생할 수 있다. 예를 들어, 클라우드 컴퓨팅 자원을 제공하는 사업자는 데이터센터에 필요한 AI 반도체를 직접 개발함으로써, AI 반도체 공급에 소요되는 기간 단축 또는 구매 비용 절감 등을 통해 범위의 경제를 실현할 수 있다. 또한, 기반모델을 개발하는 사업자가 자사 기반모델을 활용하여 최종 수요자에게 AI 서비스도 제공할 경우, 이용자 데이터를 직접 수집할 수 있고 그 데이터로 기반모델 학습을 추가로 진행할 수 있게 되는 등 데이터 구매·투자 비용을 줄이면서 범위의 경제를 실현할 수 있다.

94 CMA(2023), AI Foundation Models Initial Report

206. 이러한 범위의 경제도 생성형 AI 가치사슬 여러 분야에 진출할 여력이 충분한 대규모 수직통합 사업자에게 나타날 가능성이 상대적으로 더 크다고 볼 수 있다.

207. 네트워크 효과(network effect)는 특정 이용자의 제품 등에 대한 수요가 다른 이용자 수요에도 영향을 미치고, 이용자가 많아질수록 그 가치가 커지는 효과 등을 말하며, 생성형 AI 생태계에서도 사업자들이 이용자를 많이 확보하여 성능을 개선하는 네트워크 효과가 발생할 수 있다.

208. 예를 들면, 엔비디아의 GPU 및 CUDA의 경우, 많은 개발자들이 상당한 기간동안 이를 활용하여 AI 개발을 진행함으로써 다양한 경험과 사례 등이 축적되었고, 이것이 다른 개발자들의 엔비디아 GPU 수요를 이끌어내는 등 네트워크 효과가 실현된 것으로 해석할 수도 있다. 또한, AI 서비스 개발사도 많은 이용자를 확보하여 데이터를 축적하고 그 데이터를 통해 AI 서비스 성능과 사용자 경험을 보다 개선함으로써, 결과적으로 더 많은 이용자를 확보할 수 있게 되는 등⁹⁵ 네트워크 효과를 실현할 수 있다.

209. 위와 같은 경제적 효과 등으로 인하여 대규모 사업자, 수직통합 사업자 등이 스타트업과 같은 소규모·신규 사업자 등에 비하여 경쟁상 우위를 점할 수 있으며, 이는 경쟁 과정에서 기존 또는 신규 시장참여자들의 어려움을 초래하는 요인 등으로 작용할 우려가 존재한다.

95 Competition Bureau Canada(2024), Artificial Intelligence and competition, Discussion Paper

3.1.3 시장 선점 효과에 따른 선도적 사업자의 우위

210. 생성형 AI 가치사슬 각 분야별 시장에서는 선제적으로 진입한 사업자가 해당 시장을 선점하여 선두주자의 이점 등을 가지게 될 가능성도 존재한다.

211. 구체적으로, 선도적 사업자는 초기 시장에서 자신의 제품 및 서비스 인지도 등을 제고하여 이용자를 상대적으로 용이하게 확보할 수 있으며, 해당 이용자로부터 수집한 데이터로 자사 제품 및 서비스의 성능 등을 추가적으로 개선·향상시킴으로써 후발주자와의 격차를 확대할 수 있다.

212. 또한, 선도적 사업자는 제품 및 서비스를 공급하기 위하여 갖춰야 하는 필수 요소 등에 대한 수요가 상당히 높아지기 전에 선제적 투자를 진행함으로써, 후발주자보다 상대적으로 적은 투자비용을 지출하는 등 효율성을 높일 수 있고, 대규모 투자를 통해 규모의 경제도 달성할 수 있다.

213. 선도적 사업자는 위와 같은 경쟁력 및 효율성 등을 기반으로 자신의 제품 등을 중심으로 한 생태계를 구축하고 강화할 수 있으며, 이는 잠금효과(Lock-in effect) 등을 통해 시장지배력 강화로 이어질 가능성 또한 존재한다. 예를 들면, AI 반도체 시장에서는 엔비디아가 GPU 및 CUDA 등을 통해 해당 시장에 선제적으로 진입하여 시장을 선점할 수 있었고, 클라우드 컴퓨팅 시장에서는 아마존웹서비스가 해당 시장을 개척함으로써 선두 주자로서의 이점을 확보할 수 있었다고 볼 수 있다.

214. 결과적으로, 관련시장에 상대적으로 늦게 진입한 사업자는 위와 같이 이미 상당한 입지를 구축하여 경쟁상 유리한 위치에 있는 회사와 경쟁하여야 하므로, 이는 경쟁 과정에서 어려움을 겪을 수 있는 요인 또는 새로운 시장 진입을 저해하는 요인으로 작용할 수 있다.

3.2 경쟁 저해 및 소비자이익 저해 우려

- ◆ 본 목차에서는 가정적 상황 또는 이론적 가능성, 그리고 업계의 우려 등을 고려하여 생성형 AI 시장에서 행태적으로 발생할 수 있는 경쟁 저해 가능성, 소비자이익 저해 우려에 대하여 기술하였음
- ◆ 다만, 각 우려 등에 대한 기술 내용이 국내에서 발생한 경쟁 저해 등 혐의를 포착하였음을 의미하거나 공정위의 즉각적 조사가 필요함을 함의하지는 않으며, 법 위반을 예단하지도 않음을 명확히 함

3.2.1 단독행위

3.2.1.1 필수 요소에 대한 접근 제한

215. 생성형 AI의 개발 및 구현을 위해서는 컴퓨팅 자원, 데이터 등 필수적 요소들이 상당히 필요한데, 이러한 필수 요소에 대한 접근 측면에서 경쟁 저해 등이 발생할 우려가 있을 수 있다.

216. 사업자는 원칙적으로 자신이 제공하는 상품·용역의 가격 및 수량 등 거래조건을 자유롭게 설정하여 거래할 수 있다. 다만, 특정한 상품·용역이 필수 요소에 해당하고 관련시장에서 지배적 지위에 있는 사업자가 해당 상품의 가격, 물량 등 거래조건에 관하여 그 지배력을 남용하는 경우에는 이론상 경쟁 저해 등 우려가 나타날 수도 있다.

217. 앞서 언급한 바와 같이, 생성형 AI 생태계에서 필수 요소인 AI 반도체 시장에서는 엔비디아의 GPU에 대한 초과 수요가 지속되고 있다. 엔비디아의 수요자에 대한 GPU 물량 공급·분배 기준은 확인되지 않지만, AI 실태조사 결과에 따르면, 일부 회사가 엔비디아 파트너십을 통해 GPU 물량을 원활히 공급받을 수 있다는 취지의 의견도 일부 존재하였다.

표 54. AI 실태조사 결과 중 엔비디아 제품 공급 관련 의견

주요 의견
<ul style="list-style-type: none"> • 향후 GPU 등 수요 급증에 따라 현 글로벌 독과점 상황에서 수요물량 규모에 따른 물량 배정 순위 조정 등에 기인한 납품 지연 발생 가능성 존재 • 엔비디아와 파트너십을 통하여 우선순위로 배정받는 등 일부 긍정적 영향이 나타날 수 있음

218. 한편, 프랑스 경쟁당국이 엔비디아에 대한 경쟁법 위반 여부 조사에 착수한 것으로 확인되었고, 구체적인 조사 내용 및 결과는 공개되어 있지 않다. 다만, 프랑스 경쟁당국은 2024년 6월 발간한 AI 시장 보고서에서 ‘엔비디아 공급 제한, 불공정 계약 조건’ 등을 언급하면서 경쟁 우려를 기술하였고, 엔비디아 조사를 개시하였다고 밝힌 바 있다.

219. 엔비디아의 2023년 사업보고서를 보면, 프랑스 경쟁당국 외 유럽연합·영국 등 경쟁당국도 엔비디아의 GPU 판매·공급 할당 등에 대하여 살펴보고 있는 것으로 확인되며, 해외 경쟁당국의 이와 같은 조사 경과 등에 대하여는 주의 깊게 지켜볼 필요가 있다.

표 55. 엔비디아 2023년 사업보고서(발췌)

<p>For example, the French Competition Authority collected information from us regarding our business and competition in the graphics card and cloud service provider market as part of an ongoing inquiry into competition in those markets.</p> <p>We have also received requests for information from regulators in the European Union, the United Kingdom, and China regarding our sales of GPUs, our efforts to allocate supply, foundation models and our investments, partnerships and other agreements with companies developing foundation models, and we expect to receive additional requests for information in the future.</p>
--

* 출처: 엔비디아 사업보고서

220. 생성형 AI 생태계에서 양질의 데이터 확보 여부도 AI 서비스 시장 내 경쟁에 주요한 영향을 미칠 가능성이 있고, 공공·개방 데이터 활용의 한계점 등으로 인하여 사업자 간 데이터 거래의 필요성이 대두되고 있다.

221. SNS, 브라우저 등 주요 서비스를 통해 다수 이용자를 확보한 사업자는 그 서비스를 통해 데이터를 용이하게 수집할 수 있는 반면, 서비스 이용자 수가 적은 사업자는 자체 데이터 수집이 상대적으로 곤란하여 데이터 거래에 의존할 수밖에 없다고 할 수 있다. 다만, 이러한 데이터 거래 과정에서 개인정보보호법 등 관련 법령 등에 따라 데이터 거래가 제한될 가능성도 존재한다.

222. 원칙적으로 사업자는 자신이 보유한 데이터의 거래 또는 제공 여부에 대하여 관련 법령의 제약 하에 데이터의 가치에 대한 사업적 판단 등에 따라 자유롭게 결정할 수 있다. 「불공정거래행위 심사지침」⁹⁶에서도 생산 또는 판매 정책상 합리적 기준을 설정하여 이에 맞지 않는 불특정 다수 사업자와의 거래를 거절하는 행위는 원칙적으로 공정거래법 적용 대상이 아님을 명시하고 있다.

223. 따라서, 어느 한 사업자의 데이터 거래 요청 거절에 대하여 반드시 경쟁법적 관점에서 검토해야 하는 것은 아니나, 공정한 경쟁을 저해하거나 불공정 거래에 해당하는 경우에 한하여 경쟁법적 관점에서 검토해볼 수 있다.

224. AI 실태조사 결과, 데이터 거래를 거절당한 경험이 있는 사업자는 그 사유에 대해 ‘데이터 소유자가 데이터를 경쟁력으로 보거나, 데이터 제공 시 잠재적 경쟁사업자의 시장 진입 등을 우려하기 때문’ 이라고 응답한 경우도 존재하였다. 만약 거절 대상 데이터가 생성형 AI 사업 영위에 필수적이고, 대체거래처를 찾는 것도 쉽지 않아 사업활동이 곤란해질 수 있는 경우에는 이론적으로 경쟁 저해 우려가 나타날 가능성도 있다.

96 공정거래법 제45조 제1항에 규정된 불공정거래행위의 해석 및 적용과 관련된 내부 지침으로서, 이하 ‘심사지침’이라 함

3.2.1.2 결합 판매

225. 결합 판매는 사업자의 주된 상품·용역에 종된 상품·용역을 결합하여 판매하는 것을 말하며, 생성형 AI 시장에서는 실제로 다양한 결합 판매 사례가 나타나고 있다.

226. 마이크로소프트는 워드·엑셀 등 365 제품과 윈도우 OS에 자신의 AI 서비스 코파일럿(Copilot)을 결합하고 있고, 구글은 Gmail 및 Docs 등에 자사 AI인 Gemini를 결합하고 있는 것으로 나타난다. 또한, 구글, 아마존웹서비스, 마이크로소프트 등 주요 글로벌 CSP는 클라우드 서비스를 통한 컴퓨팅 자원 제공 외에도 자사 또는 타사의 기반모델에 대한 접근·사용 기능을 제공하는 등 클라우드 서비스와 기반모델 배포 서비스 등을 결합하고 있다.

227. 생성형 AI 분야의 여러 가치사슬에 진출하여 수직통합을 구축하고 다양한 서비스를 제공하고 있는 사업자는 각 서비스를 결합하여 제공하고자 하는 유인을 가지며, 이러한 결합 판매 방식은 시장 경쟁에서 자연스럽게 나타날 수 있는 사업 방식에 해당한다고 볼 수 있다.

228. 즉, 사업자가 자신이 제공하는 주된 상품·용역의 경쟁력을 높이기 위한 수단으로 종된 상품·용역을 함께 공급할 수 있고, 이러한 행위가 주된 상품·용역 시장에서 성능 향상 경쟁으로 이어지는 경우에는 친경쟁적 효과가 나타난 것으로 볼 수 있다. 예를 들어, 한 사업자가 문서 작성 프로그램에 생성형 AI 기능을 도입하여 서비스를 개선한다면, 다른 사업자도 경쟁에서 뒤처지지 않기 위하여 AI 기능을 추가함으로써 문서 작성 프로그램의 전반적인 성능이 향상될 수 있고, 이는 소비자 후생의 증가로 이어질 수 있다.

229. 다만, 주된 상품·용역 시장에서 상당한 지배력을 가지고 있는 사업자가 거래상대방으로 하여금 종된 상품의 구매를 강요하는 등 결합 판매가 강제될 경우에는 종된 상품·용역시장에서 경쟁 저해 등이 발생할 우려가 있다.

230. 이는 주된 시장에서 지배력을 가지는 사업자가 그 지위를 이용하여 종된 시장에서 고객을 다수 확보하려고 하는 것으로서, 결과적으로 시장지배력의 전이를 도모하고 수요자의 자율적 선택을 제약할 우려가 있으며, 종된 시장의 경쟁사업자가 공급기회 등을 상실하여 경쟁상 열위에 처할 수도 있다.

231. 현재 생성형 AI 시장에서도 일부 결합 판매가 나타나고 있으며, 주요 해외 경쟁당국이 조사를 진행하고 있는 사례도 존재한다.

232. EU 경쟁당국은 삼성전자 및 구글 간 계약 등에 따라 삼성전자 스마트폰에 구글의 AI 서비스 Gemini nano가 기본적으로 먼저 탑재되어 출시되는 것에 관하여 모니터링 중이다. 조사에 착수한 배경은 빅테크 사업자가 다른 기반모델 개발사의 최종 사용자에게 서비스 제공 등을 어렵게 할 수 있다는 우려 때문인 것으로 확인되었다.⁹⁷

233. 한편, AI 실태조사 결과 응답에서도 구글 AI 서비스 선택제에 대한 의견이 아래와 같이 제시되기도 하였다.

표 56. AI 실태조사 결과 중 AI 모델 선택제 관련 의견

주요 의견
<ul style="list-style-type: none"> • 앱 개발자 및 이용자가 스마트폰에 선택제된 AI에 의존하게 되고, 다른 경쟁 AI의 앱 개발자 및 이용자에 대한 접근이 제한될 우려 • 스마트폰 내 영구적·임시적 메모리의 한계 및 특정 AI 선택제 등으로 인하여 경쟁 AI의 성능 저하 및 경쟁 배제 우려

234. 한편, 독일 경쟁당국은 구글이 차량 운영체제(Operating System, 이하 ‘OS’)인 안드로이드 오토모티브 OS 라이선스를 차량 제조사에 제공할 때, 구글의 음성인식 AI 등이 기본 탑재될 수 있도록 한 것에 대하여 조사를 진행하고, 구글의 자진시정 절차 등을 진행한 바 있다.

97 European Commission(2024), Speech by EVP Margrethe Vestager at the European Commission workshop on “Competition in Virtual Worlds and Generative AI”

235. 구체적으로, 독일 경쟁당국은 2023년 6월 아래와 같은 구글의 결합 판매 행위 등이 경쟁을 저해할 수 있다는 예비적 견해를 밝히고, 이에 대한 구글의 시정안에 대하여 2023년 12월 시장 테스트를 진행하였다.

표 57. 안드로이드 오토모티브 OS 및 음성인식 AI 결합 판매 관련 내용

주요 결합 판매	<ul style="list-style-type: none"> 안드로이드 오토모티브 OS에 구글 맵, 구글 플레이, 구글 어시스턴트 서비스를 결합하여 번들로 제공 구글 어시스턴트 외 다른 음성 지원 기능을 설치하지 않는 조건으로 광고 수익의 일부를 차량 제조사에 제공 등
구글 주요 시정안	<ul style="list-style-type: none"> 번들 외 구글 맵 OEM 소프트웨어 개발도구, 구글 플레이 스토어, 차량 내 AI 음성지원 솔루션을 추가 제공 구글 어시스턴트 독점적 사전 설치 조건의 수익 공유 관련 계약 조항 삭제 등

* 출처: 독일 경쟁당국⁹⁸

236. 독일 경쟁당국은 예비적 견해 이후 최종적 판단을 내리지는 않았으며, 예비적 견해의 주요 내용은 기본값 설정 행위로 인하여 다른 대체적 서비스가 거의 사용되지 않을 위험이 초래되고, 시장 접근에도 방해가 될 수 있다는 것 등으로 확인되었다.

237. 이상과 같이, AI 시장에서 결합 판매는 경쟁제한 효과와 효율성 증대 효과 및 소비자 후생 증대 효과가 동시에 나타날 수 있다. 따라서 특정 결합 판매가 국내 생성형 AI 시장에서 경쟁을 제한하는지 여부에 대해서는 지속적 모니터링이 필요할 것이다.

98 Bundeskartellamt(2023), Statement of objections issued against various of Google's practices in connection with Google Automotive Services and Google Maps Platform(Press Releases), Google Automotive Services proceeding - Market test regarding commitments proposed by Google(Press Releases)

3.2.1.3 고객 유인 및 고객의 이탈 방지

238. 생성형 AI 가치사슬에서 사업자들은 고객을 확보·유지하기 위하여 다양한 사업전략·수단을 실행할 수 있으며, 그 수단에는 이익 제공·제의를 통하여 고객을 유인하는 방법 또는 고객이 자신의 경쟁사와 거래하지 않도록 하는 방법 등이 있을 수 있다.

239. 먼저, 이익을 제공하는 것에 관하여 살펴보면, 일부 CSP가 자사 고객을 확보하기 위하여 클라우드 서비스 크레딧(Credit)⁹⁹을 제공하고 있는 것으로 알려졌으며, 클라우드 서비스 외 다른 가치사슬 분야에서도 이익 제공을 통한 고객 확보 방법이 활용될 가능성이 존재한다.

240. 사업자가 자신과 거래하도록 하기 위하여 상품 등 판매량에 따른 가격 할인 또는 일정 이익을 제공하는 것은 통상의 경제원리에서 흔하게 나타날 수 있는 것이며, 이는 소비자 후생 등을 증대하는 긍정적 효과도 나타날 수도 있다. 특히, 생성형 AI 시장과 같은 초기 시장에서는 주요 사업자들이 자사 제품의 인지도를 제고하고, 고객 또는 이용자 확보 등을 위하여 이익을 제공할 유인을 가지게 된다고 볼 수 있다.

241. 다만, 그 이익을 제공하는 정도가 부당하거나 과다할 경우에는 관련시장에서 가격과 품질 등에 의한 바람직한 경쟁질서를 저해할 우려가 나타날 수 있다.

242. 만약 상당한 자금력을 가진 사업자가 관련시장에서 통상 제공되는 이익의 수준을 넘는 과도한 이익 등을 제공하여 고객을 유인하는 경우, 상대적으로 이러한 여력이 부족한 사업자는 고객을 확보하기 어렵고, 결국 경쟁상 열위에 놓이거나 시장경쟁에서 배제될 우려도 있다.

99 이는 클라우드 서비스를 일정 기간 무료로 이용할 수 있도록 일정 규모의 금전적 혜택을 제공하는 것 등을 말함

243. 다음으로, 사업자는 자신의 고객이 다른 경쟁사업자와 거래하게 되는 등 거래관계에서 이탈하는 것을 방지하고자, 거래조건에 경쟁사업자와의 거래제한 등의 내용을 명시하거나 간접적으로 거래상대방 전환 시 불리한 조건을 적용하는 등 다양한 형태로 배타조건부 거래를 실행할 가능성이 있다.

244. 생성형 AI 생태계에서도 기반모델 배포 및 데이터 거래 등에서 이러한 배타조건부 거래가 나타날 가능성이 존재한다. 예를 들면, 특정 CSP가 기반모델 계약사와의 계약 등을 통해 자신의 클라우드 서비스에서 해당 기반모델에 대한 접근·배포 기능을 제공하면서, 다른 CSP의 클라우드에서는 제공할 수 없도록 제한하는 사례가 있을 수 있다.

245. AI 실태조사 결과에 따르면, 일부 회사는 ‘마이크로소프트가 자사 Azure 클라우드 서비스를 통해서만 배타적으로 오픈AI의 기반모델(GPT)을 배포하고 있다’는 의견이 제시되기도 하였다. 다른 가정적 예시로는, 특정 데이터를 제공받는 사업자가 데이터 제공사로 하여금 경쟁사업자에게 해당 데이터를 제공하는 것을 제약하는 조건으로 계약하는 사례 등이 있을 수 있다.

246. 사업자가 이윤 추구 등을 위하여 자사 고객이 경쟁사업자와 거래하는 것을 선호하지 않을 수 있고, 배타조건부 거래 등을 실행하는 것에 합리성이 인정될 수도 있기 때문에 그 거래에 대하여 곧바로 경쟁법적 관점에서 검토해야 하는 것은 아니다. 다만, 배타조건부 거래가 강제성을 띠고 그로 인하여 관련 시장에서 경쟁제한 우려를 불러일으킬 수 있는 경우에는 경쟁법적 관점에서 검토해 볼 여지가 있으며, 그 경쟁제한성 여부는 심사지침에 따라 배타조건부 거래가 물품 구입처 또는 유통경로 차단, 경쟁수단 제한 등을 통해 경쟁사업자를 관련시장에서 배제하거나 배제할 우려가 있는지 여부를 위주로 판단한다.

3.2.1.4 기술 부당 이용

247. 생성형 AI 가치사슬에서 사업자의 경쟁력을 결정하는 주요 요인에는 기술력 수준도 있으며, 최근 기술 수준 향상과 AI 서비스의 개발, 품질 제고 등을 위하여 사업자 간 기술 교류·협력 등이 나타나고 있다.

248. 사업자 간 교류·협력 방식은 제휴(Partnership), MOU(Memorandum Of Understanding), 공동 연구개발, 협의체 참여 등 명칭을 불문하고 다양한 형태로 존재하며, 이를 통해 기술 수준의 발전, 새로운 서비스 개발 및 품질 개선 등이 이루어질 경우 소비자 후생도 증대되는 등 친경쟁적 효과도 발생할 수 있다.

249. 다만, 사업자 간 협력 과정에서 법적 보호대상 등에 해당하는 기술을 부당하게 이용할 경우, 그 기술을 보유한 사업자의 권익이 침해되고 거래상대방 확보 및 수익 실현 등 장래의 사업활동이 곤란해질 수 있으며, 공정거래법은 이를 ‘기술의 부당 이용¹⁰⁰’으로 규율하고 있다. 예를 들어, AI 기술 보유사와 미보유사 간 협력 과정에서 법적 보호 대상에 해당하는 기술 정보를 공유한 후, 기술 미보유사가 해당 기술을 부당하게 활용하여 서비스를 출시할 경우, 기술 보유사의 미래 사업활동이 상당히 곤란해질 가능성¹⁰¹이 존재한다.

250. AI 관련 산업이 빠른 속도로 변화·발전함에 따라 법적으로 보호받을 수 있는 기술 범위를 명확하게 특정하기 어려운 문제 등이 나타날 수는 있으나, 관련 시장에서 기술 부당이용 사례 발생에 대해서는 예의주시할 필요가 있다.

100 공정거래법 시행령 [별표 2] 불공정거래행위의 유형 또는 기준에서 규정한 사업활동 방해 중 기술의 부당 이용

101 앞서 살펴보았던 시장 선점 효과 등에 따라 기술을 부당 이용한 사업자는 자사 서비스 인지도를 높이고 경쟁에서 유리한 위치를 차지하는 반면, 당초 기술을 보유하고 있던 사업자는 자사 고객을 확보하지 못하고 관련 수익 창출도 어려워질 수 있음

3.2.1.5 인력의 유인·채용

251. 현재 생성형 AI의 개발과 구현에서는 상당한 전문인력이 필요하므로 주요 사업자들이 전문인력 확보를 위해 경쟁하고 있으며, 관련 시장에서 전문인력들의 이직도 꾸준히 발생하고 있는 것으로 나타나고 있다.

252. 전문인력 개개인의 직업 선택 자유를 고려할 때, 그 인력이 이직하는 것 자체를 경쟁법으로 규율할 수는 없으나, 특정 사업자가 다른 사업자 소속의 전문인력을 자사 인력으로 부당하게 유인·채용하는 행위에 대해서는 경쟁법 관점에서 검토해볼 수 있다¹⁰².

253. 즉, 다른 사업자의 인력을 부당하게 유인·채용하여 해당 사업자의 사업 활동을 곤란하게 하는 행위는 바람직한 경쟁 질서를 저해하는 불공정한 경쟁 수단에 해당할 수 있다. 다만, 심사지침에 따라 해당 인력의 유출과 그로 인한 사업자의 상당한 매출 감소 등의 인과관계를 입증할 필요가 있다.

254. 한편, 인력의 유인·채용 행위는 특정 사업자의 사업활동을 곤란하게 할 수도 있지만, 사업자 간 경쟁에도 상당한 영향을 미칠 수 있으므로 경쟁제한 효과를 야기할 수 있는 인력의 유인·채용에 대한 공정위의 다른 수단도 고려해볼 필요가 있다. 예를 들어, 지난 2024년 3월 공정위가 공무원 학원 사업자 간 결합을 불승인하면서, 기업결합 시 특정 회사에 인기 강사 등 인력이 집중됨으로써 경쟁 저해 우려 등이 나타날 수 있다고 보았는바¹⁰³, AI 관련 사업자 간 기업결합 심사 시에도 인력 집중의 효과를 검토하는 방안도 고려해볼 수 있다.

102 공정거래법 시행령 [별표 2] 불공정거래행위의 유형 또는 기준에서 규정한 사업활동 방해 중 인력의 부당 유인·채용

103 공정위(2024), 공정위, 메가스터디교육의 공단기 인수 불허 보도자료

3.2.2 기업결합

255. 사업자는 자신의 경쟁력 등을 제고하기 위하여 자체적인 연구·개발을 진행할 수도 있지만, 전문성이 있거나 유력한 기업을 인수하는 전략 또는 해당 기업에 투자하는 전략 등을 통해 기술력을 향상시킬 수 있다.

256. 이러한 사업자의 인수·투자는 시장경제에서 자연스레 나타날 수 있는 전략이고, 기술 발전과 혁신을 촉진하여 시장규모의 성장과 경쟁 활성화 등 긍정적 효과를 야기할 수도 있다.

257. 현재 생성형 AI 생태계에서도 자금력이 풍부한 수직통합 사업자 등이 높은 기술력을 보유하고 있으나 자본이 부족한 스타트업 등에 대하여 상당한 규모로 투자하거나 파트너십을 체결하는 사례들이 나타나고 있다. 이러한 파트너십은 전통적인 자본 투자 및 지분 취득 외에도 인력 채용, 기반모델 라이선스 확보 등 여러 형태로 나타나고 있으며, 주요 AI 사업자 간 파트너십 내용은 아래와 같이 확인된다¹⁰⁴.

표 58. 주요 사업자 간 파트너십 내용

구분	내용
마이크로소프트-오픈AI (2019년~)	· MS의 오픈AI에 대한 약 130억 달러 투자 및 AI 모델 라이선스, 오픈AI 지분 49% 확보 등
구글-엔트로픽 (2023년)	· 구글의 엔트로픽에 대한 약 20억 달러 투자 계획 및 클라우드 인프라 제공 등
마이크로소프트-인플렉션AI (2024년)	· MS의 인플렉션AI 핵심 인력 채용, 주요 지적재산 라이선스 확보(약 6.5억 달러 지불) 등
아마존-어댑트 (2024년)	· 아마존의 어댑트 공동 창업자 및 주요 인력 채용 및 어댑트 AI 모델 라이선스 확보 등

104 공정위(2024), AI 시장 실태조사 연구 / 어댑트 사이트 등 참고

258. 다만, 최근 대규모 AI 사업자의 투자·인수 등이 관련시장 내 시장지배력 강화 등으로 이어질 수 있다는 우려가 제기되고 있어 AI 시장에서의 투자·인수 등이 시장경쟁에 미치는 영향을 면밀히 살펴볼 필요성이 높아지고 있다. 특히, 현행 공정거래법은 기업결합 유형을 5개 유형으로 규정¹⁰⁵하고 있으나, 최근 생성형 AI 생태계에서 나타나는 파트너십은 전통적인 기업결합 유형에 해당하지 않으면서도, 실질적으로는 기업결합의 효과를 가질 수 있다.

259. 한편, 영국 경쟁당국은 마이크로소프트의 인플렉션 AI 인력 채용 등 파트너십을 기업결합 심사 대상으로 보아 심사를 거쳐 이를 승인하는 등 기업결합 심사 대상 범위를 폭넓게 인정하고 있다.

표 59. 영국 경쟁당국의 기업결합 심사 배경(발췌)

- | |
|--|
| <p>9. The definition of a relevant merger situation under the Enterprise Act 2002 (the Act) covers various kinds of transactions and arrangements, subject to the following three criteria being met: (i) two or more enterprises cease to be distinct. (후략)</p> <p>11. There is no particular combination of assets that constitutes an enterprise. As set out in the CMA's guidance, it may include a group of employees and their know-how where this enables a particular business activity to be continued.</p> <p>13. In addition to hiring the core former Inflection team, Microsoft also acquired additional assets, including access to Inflection IP. (후략)</p> <p>14. On this basis, the CMA believes that Microsoft has substantively acquired Inflection's pre-Transaction FM and chatbot development capabilities. (중략) and, as a result, that two enterprises have ceased to be distinct such that the Transaction falls within the CMA's merger control jurisdiction for review.</p> |
|--|

* 출처: 영국 경쟁당국¹⁰⁶

260. 위와 같은 파트너십이 기존의 기업결합과 사실상 동일·유사한 효과를 가질 경우 기업결합 심사 필요성이 있을 수 있는바, 이러한 신유형 기업결합에 대한 정책적 대응 방안을 모색할 필요도 있다.

105 주식 취득 또는 소유, 임원 겸임, 합병, 영업 양수, 회사 설립 참여

106 CMA(2024), Microsoft Corporation's hiring of certain former employees of Inflection and its entry into associated arrangements with Inflection(Summary)

3.2.3 소비자 이익 저해

261. 사업자가 AI 개발 및 기능 향상 등을 위하여 자신의 서비스를 이용하는 이용자로부터 데이터를 수집하여 이용할 수 있으며, 이러한 데이터 수집·이용에 관해서는 공정거래법 또는 약관법¹⁰⁷ 등의 관점에서 검토해볼 수 있다.

262. 먼저, 공정거래법상 사업자와 소비자 간 거래를 규율하는 규정은 많지 않으나 ‘시장지배적 사업자가 부당하게 소비자의 이익을 현저히 해칠 우려가 있는 행위’를 하는 것을 금지하고 있으며, 데이터 수집·이용 행태에 대하여 해당 규정을 적용·검토할 수 있는지에 대한 분석과 논의가 필요하다.

263. 현재 국내에서 AI 서비스를 제공하는 사업자들이 이용자의 데이터를 수집·이용하기 위해서는 관련 법령에 따라 이용자 동의를 받아야 하므로¹⁰⁸, 각 사업자들은 제각각 다른 방식으로 이용자 동의를 받고 있다. 동의 방식에는 최초 가입 당시 약관에 다른 내용과 함께 포괄적 동의받는 방식, 서비스 이용 시 개별 고지 등을 통해 동의받는 방식, 이용자가 동의 여부를 선택하도록 하는 방식 등 다양한 방식이 존재한다.

264. 한편, 독일 경쟁당국은 페이스북의 ‘서비스 이용 약관 등을 근거로 이용자 데이터를 수집·이용한 행위’를 독일 경쟁법에 따라 시장지배적지위 남용으로 검토하여 시정조치를 부과한 바 있다.

265. 해당 사건에 관하여 독일 경쟁당국은 이용자들의 실질적 동의가 없었다는 점 등을 근거로 페이스북의 행위가 유럽연합 일반 데이터 보호 규정¹⁰⁹에 위반되고, 부당한 데이터 수집·이용 등을 통해 시장지배력이 강화되는 등 이용자(소비자)에 대한 거래조건 관련 착취 남용이라고 판단하였다¹¹⁰.

107 「약관의 규제에 관한 법률」의 약칭

108 참고로, 개인정보보호위원회는 메타가 이용자 민감정보를 수집하고 활용하면서 별도동의를 받지 않은 행위 등에 대하여 2024년 11월 과징금 약 216억 원 등을 부과

109 General Data Protection Regulation, 이하 ‘GDPR’

표 60. 페이스북 사건 관련 독일 경쟁당국 등 판단 내용

경쟁당국 판단(2019.2.)
<ul style="list-style-type: none"> · (대상 행위) 페이스북은 서비스 이용 약관 등을 근거로 제3자*로부터 이용자 데이터를 수집하고, 이를 해당 이용자의 페이스북 계정에 연계·통합한 행위 * 왓츠앱(WhatsApp), 인스타그램(Instagram), 페이스북 API 연동 웹사이트 등 · (적용 법조) 독일 경쟁법 제19조 제1항(시장지배적지위 남용)*을 적용하였고, 유럽연합 조약상 시장지배적지위 남용 규정은 적용하지 않음 * '1개 이상의 기업이 시장지배적 지위를 남용하는 것을 금지한다.' · (주요 판단) ①페이스북은 독일 소셜네트워크 시장지배적 사업자, ②이용자가 충분히 인지하지 못하고 실질적 동의 없는 데이터 수집·이용은 GDPR 위반, ③이를 통한 시장지배력 형성·강화 등 · (시정조치) 행위 종료 및 이행계획 제출 명령(2019.2.)
고등법원(2019.8.) 및 연방대법원(2020.6.) 판단
<ul style="list-style-type: none"> · (고등법원: 효력 정지) 이용자에게 경제적 피해·데이터 통제권 손실이 미발생, 자발적 동의가 없었다고 보기 어려움 등 · (연방대법원: 원심 파기) 이용자가 페이스북 서비스에 대한 선택권* 등을 갖지 못하고, 이러한 선택권의 상실은 이용자에 대한 착취에 해당, 페이스북이 더 많은 데이터로 이용자의 고착 현상 강화 및 수익성 개선 등 * 광범위한 데이터 수집·이용에 따라 보다 개별화된 경험이 제공되는 서비스, 이용자가 데이터 제공을 선택하되 덜 개별화된 경험을 제공하는 서비스 중 선택할 수 있는 권리

266. 사업자들이 AI 모델 학습 등을 위한 데이터 수집·이용과 관련하여 이용자 동의를 받기는 하지만, 위 사례와 같이 실질적 동의를 받지 않는 경우가 있을 수 있다. 그 경우 독일과 유사하게 착취 남용으로 규율할 수 있는지에 대해서는 심층적인 분석과 많은 논의가 필요하다.

267. 우선, 현행 법령상 소비자 이익 저해 행위에 대하여 소비자 이익의 개념과 범위 및 이를 저해하는 행위, 현저성의 판단기준 등을 구체화할 필요가 있다. 또한, 개인정보보호법 등을 통해 보호하는 데 한계가 있을 수 있는 데이터 관련 경쟁법적 쟁점, 소비자 이익 저해행위 등에 대하여 공정거래법 등을 통해 보호할 수 있는지 등에 대한 후속 연구도 필요할 것이다.

110 이상운(2020), 디지털 플랫폼 사업자의 소비자 착취 행위에 대한 경쟁법의 적용: 독일 페이스북 사건 / 유진희·심재한(2021), 온라인 플랫폼 사업자의 이용자 데이터 보호 위반행위와 경쟁법의 적용 등 참고

268. 한편으로는, EU 경쟁법 체계¹¹¹와 같이 착취남용 유형으로 ‘불공정한 거래 조건 부과’를 규정하는 방식도 고려할 필요가 있다는 의견도 존재한다¹¹².

269. 다음으로, 약관법 관점에서는 사업자들이 데이터 수집 등에 관해 이용자 동의를 받기 위하여 마련한 약관 내용 등의 불공정성 여부를 검토해 볼 수 있다.

270. 약관은 계약의 한쪽 당사자, 통상적으로 사업자가 여러 명의 상대방과 계약을 체결하기 위하여 일정한 형식으로 미리 마련한 계약의 내용을 말하며, 데이터 수집·이용 관련 약관 내용은 사업자별로 다르게 나타난다.

271. 사업자가 AI 서비스 제공 및 기능 향상 등을 위하여 이용자 데이터를 수집·이용할 수 있도록 하는 등의 약관 내용이 이용자에게 부당하게 불리한지, 즉 불공정 약관에 해당하는지 여부에 대해서는 신중하게 검토할 필요가 있다. 즉, 사업자의 데이터 수집으로 인하여 이용자 권리가 침해될 가능성이 있으나, 한편으로는 AI 기능 향상 등을 통해 궁극적으로 이용자의 이익이 증대될 가능성도 배제할 수는 없으므로 사업자와 이용자 간 이익의 비교형량 등을 통해 ‘부당하게 불리한지’ 여부를 면밀하게 검토하여야 한다.

272. 이를 위해서는 저작권법 및 개인정보보호법 등 관계법령, AI 서비스 업종의 특성 및 국내외 거래관행 등을 종합적으로 고려하여 판단하는 것도 필요할 것이다.

111 EU 경쟁법 「Treaty on the Functioning of the European Union(TFEU)」 제102조는 시장지배적지위 남용 행위 유형 중 하나로 ‘부당하게 높은 가격 또는 낮은 가격을 설정하는 행위, 불공정한 거래조건의 부과’ 등을 규정하고 있으며, 독일 경쟁법 「Gesetz gegen Wettbewerbsbeschränkungen(GWB)」 제19조는 시장지배적지위 남용 행위 유형으로 ‘실질적인 경쟁 시 발생한 개연성이 높은 것과 다른 대가 또는 그 밖의 거래조건을 요구하는 경우’ 및 ‘유사한 시장에서 시장지배적기업이 동종 구매자에게 요구하는 것보다 불리한 대가 또는 그 밖의 거래조건을 객관적 이유 없이 요구하는 경우’ 등을 규정하고 있음

112 최난설현(2018), 디지털 시장에서의 독과점 규제 적용 가능성에 대한 검토 - 독일의 Facebook 사례를 중심으로 -

273. 이 외에도 데이터 수집·이용 약관 외 AI 서비스 결과값, 품질 등에 대하여 사업자를 면책하는 약관 내용 등에 대해서도 AI 기술의 특징을 고려하더라도 과도한 측면이 없는지 검토해 볼 수도 있다.

274. 한편, 앞서 살펴본 바와 같이 사업자의 데이터 수집·이용 행태 등에 대하여 사후적으로 접근하는 방식도 있으나, 당사자 간 분쟁을 사전에 예방하고 혁신을 촉진하기 위하여 ‘데이터 수집·이용 시 동의를 받는 방식’ 등에 관한 구체적인 방안도 고려해 볼 수 있다.

275. 즉, 대면 계약과 달리 사업자의 약관상 주요 내용 설명¹¹³ 등이 곤란할 수 있는 비대면 계약에서 사업자가 이용자 데이터 수집·이용에 관한 사항을 어떻게 고지하고 동의를 받는 것이 바람직한지 논의해 볼 수 있다. 예를 들면, 사업자가 이용자에게 데이터 수집·이용 사실을 고지하고 이용자가 본인의 자율적인 판단으로 옵트인 또는 옵트아웃 등¹¹⁴으로 동의 여부를 선택할 수 있도록 하는 등 방안 등이 존재하며, 이 경우 옵트아웃은 이용자가 용이하게 선택할 수 있는 절차가 마련되어야 할 것이다.

276. 참고로, 영국 정부는 AI 사업자들이 옵트아웃 방식으로 이용자 데이터를 수집하는 것을 허용하는 방안에 대하여 이해관계자 등과 협의할 계획인 것으로 확인되었다¹¹⁵.

113 약관법 제3조 제3항은 “사업자는 약관에 정하여져 있는 중요한 내용을 고객이 이해할 수 있도록 설명하여야 한다. 다만, 계약의 성질상 설명하는 것이 현저하게 곤란한 경우에는 그러하지 아니하다” 라고 규정

114 옵트인(Opt-in) 방식은 사전에 이용자 동의를 받은 후 데이터를 수집·처리하는 방식을 말하며, 이에 상반되는 옵트아웃(Opt-out) 방식은 먼저 데이터를 수집·처리하고 이용자가 거절 의사를 나타내면 이를 중단하는 방식을 말함

115 한국저작권위원회(2024), 영국, AI 콘텐츠 수집 옵트아웃 모델 검토에 콘텐츠 제공자와 창작자 반발

4 결론

277. 생성형 AI 시장은 주요 시장 참여자들 간 활발한 경쟁이 이루어지고 있는 시장일 뿐만 아니라 기술 혁신 등을 통해 매우 빠르게 변화·발전하고 있는 역동적인 시장에 해당한다. 대화형 AI 서비스인 챗GPT 출시 이래 현재는 스스로 판단하여 업무를 처리하기까지 하는 AI 에이전트 서비스도 개발되며, 초거대 AI 모델도 짧은 기간에 급격하게 출시되는 등 생성형 AI 시장에서 활발한 경쟁이 이루어지고 있다.

278. 앞서 살펴보았던 바와 같이, 국내 생성형 AI 생태계 내 가치사슬 전 분야에서도 국내·외 사업자들이 매우 활발하게 경쟁하고 있으나, 세부적인 시장별로 시장지배력을 가진 사업자 혹은 수직통합 사업자가 등장한 것도 사실이며, 그에 따른 시장참여자의 우려도 나타나고 있다.

279. AI 반도체 시장에서 엔비디아가 AI 생태계의 필수요소인 GPU 제품의 막강한 성능을 바탕으로 지배력을 가지고 있다고 시장참여자들이 평가하고 있고, 구글·아마존·마이크로소프트 등 글로벌 사업자와 네이버와 같은 국내 사업자는 생성형 AI 생태계 내 주요 가치사슬별 시장에 진출하여 수직통합 사업자로써 AI 서비스 등을 개발·출시하고 있다.

280. 다만, 어느 한 사업자가 시장지배력을 가진다는 사실 자체가 곧바로 공정거래법 위반을 의미하지는 않으며, 그 지배력을 통해 별도의 남용행위 또는 경쟁제한 행위를 하는 경우 공정거래법 등 위반 여부를 검토해볼 수 있다. 이론적으로 생성형 AI 생태계는 대규모 투자의 필요성, 시장선도자의 우위 등 구조적 특징으로 인하여 경쟁법적 우려가 제기될 가능성이 존재한다. 자본·기술 집약적인 산업의 특성상 진입장벽이 발생할 수 있으며, 규모의 경제, 범위의 경제 등이 작동하기 때문에 대규모 사업자가 우위를 점할 수 있고, 아울러 네트워크 효과 등으로 인해 시장을 선점한 선도적인 사업자가 경쟁에 있어서 유리한 위치를 점할 수도 있다.

281. 경쟁당국은 구조적으로 쏠림 현상이 나타날 수 있는 생성형 AI 시장에서 시장참여자들의 행태가 경쟁 및 소비자에 미치는 영향을 분석하고, 경쟁제한적 우려가 있는 특정 시장에 대한 모니터링을 강화하는 한편, 제도개선을 통해 경쟁질서 마련을 위해 노력할 필요가 있다.

282. 많은 행위들이 경쟁제한적인 효과와 친경쟁적인 효과가 동시에 나타낼 수 있기 때문에 해당 행위에 대해서는 면밀하게 살펴볼 필요가 있다. 다만, 각 국가별로 처한 시장 경쟁상황이 다르기 때문에 해외 경쟁당국의 조사 및 제재가 반드시 공정위 조사 필요성을 의미하지는 않는다.

283. 역동적인 AI의 생태계의 특성상 기존의 제도의 개선을 통해 경쟁을 제한하는 행위를 사전적(ex-ante) 혹은 사후적(ex-post)으로 점검할 필요성이 있는지 검토해볼 필요가 있다. 기존의 기업결합 심사 제도를 통해 점검하는 데 한계가 있을 수 있는 새로운 유형의 파트너십에 대한 사전적인 심사 필요성을 검토하고, AI 생태계의 필수 요소인 ‘데이터’의 수집·이용과 관련한 남용 행위에 대하여 소비자이익 저해 관점에서 평가할 수 있는지 제도 개선을 살펴볼 필요가 있다.

284. 이번 「생성형 AI와 경쟁」 정책보고서는 시장의 초기 단계이자 급격하게 변화하는 국내 생성형 AI 시장의 경쟁상황 등을 각 가치사슬 전반에 걸쳐 조망했다는 점에서 의의가 있다.

285. 생성형 AI 세부 시장별로 경쟁상황의 분석, 경쟁·소비자 이슈의 점검 등을 통해 모니터링, 제도개선 등 향후 추진 과제를 제시하였으며, 앞으로도 국내외 시장참여자와 지속적으로 소통하여 역동적으로 발전하고 있는 AI 시장의 경쟁상황 등에 유연하게 대응하며 필요 시 추가적인 시장분석도 추진할 예정이다.

<참고 문헌>

- 공정거래위원회(2024), AI 시장 실태조사 연구
- 공정거래위원회(2022), 클라우드 서비스 분야 실태조사 연구
- 공정거래위원회(2024), 메가스터디교육의 공단기 인수 불허 보도자료
- 고용노동부(2023), '27년까지 인공지능(AI) 12,800명, 클라우드 18,800명
신규인력 부족 전망 보도자료
- 과학기술정보통신부(2024), 2023년 부가통신사업 실태조사 결과
- 과학기술정보통신부(2024), 인공지능 3대 강국(AI G3) 도약을 위한 대한민국 AI
혁신의 청사진 제시 보도자료
- 과학기술정보통신부 등(2024), 인공지능 학습용 데이터 품질관리 가이드라인 v3.1
제2권 초거대 AI 데이터 품질관리 가이드라인
- 법제처(2024), 인공지능(AI) 관련 국내외 법제 동향
- 소프트웨어정책연구소(2024), 유럽연합 인공지능법(EU AI Act)의 주요 내용 및
시사점
- 소프트웨어 정책연구소(2024), 글로벌 초거대 AI 모델 현황 분석(2020~2023년)
- 소프트웨어 정책연구소(2023), 생성 AI 산업 생태계 현황과 과제
- 소프트웨어 정책연구소(2023), 초거대언어모델의 부상과 주요 이슈
- 소프트웨어 정책연구소(2024), 2023년 인공지능산업 실태조사 결과보고서
- 정보통신기획평가원(2024), ICT Brief 2024-38호
- 정보통신기획평가원(2024), ICT Brief 2024-32호
- 한국과학기술평가원(2019), 기술동향 브리프 인공지능
- 한국과학기술평가원(2024), AI 반도체 기술개발 현황과 시사점: 클라우드 산업을
중심으로
- 한국수출입은행(2024), AI 반도체 시장 현황 및 전망
- 한국저작권위원회(2023), 생성형 인공지능(Generative AI) 산업 현황 보고서
- 한국저작권위원회(2024), 영국, AI 콘텐츠 수집 옵트아웃 모델 검토에 콘텐츠
제공자와 창작자 반발
- 한국전자통신연구원(2024), 온디바이스 소형언어모델 기술개발 동향

한국전자통신연구원(2024), 주요국 반도체 정책과 AI반도체 정책에의 시사점
한국지능정보사회진흥원(2024), 버티컬 AI로의 변화 및 과제
한국지능정보사회진흥원(2023), 파운데이션 모델의 이해와 미래 전망
한국지능정보사회진흥원(2024), OECD AI 시스템 정의 개정의 주요 내용 및
시사점

이상윤(2020), 디지털 플랫폼 사업자의 소비자 착취 행위에 대한 경쟁법의 적용:
독일 페이스북 사건
유진희·심재한(2021), 온라인 플랫폼 사업자의 이용자 데이터 보호 위반행위와
경쟁법의 적용

사피온코리아(2024), 사피온 AI 반도체 'X330', 슈퍼마이크로 서버 적격성 평가 완료
사피온코리아(2024), X330 제품 개요 등
삼성증권(2019), 클라우드 전쟁 1조 달러 클럽을 향한 왕좌의 게임
삼일 PwC 경영연구원(2023), 미리보는 CES 2024
삼일 PwC 경영연구원(2024), 생성형 AI를 활용한 비즈니스의 현주소
코난테크놀로지(2024), 국내 첫 대형언어모델 공공 도입... 남부발전과 41억대
계약 보도자료
하나금융경영연구소(2024), 국내 데이터센터 산업, AI와 ESG 사이에서 길을 찾다
한국전자정보통신산업진흥회(2024), 온디바이스 AI 기술동향 및 발전방향
NHN클라우드(2024), AI 인프라 시장 선두 목표로 'NHN 클라우드 2.0' 전략 발표
SK하이닉스(2024), All around AI 3편 머신러닝의 이해
SK C&C(2023), 머신러닝? 딥러닝? 알쏭달쏭한 인공지능 용어 파헤치기

Autoridade da Concorrência(2023), Competition and generative artificial intelligence
Autorité de la concurrence(2024), On the competitive functioning of the generative
artificial intelligence sector
Bundeskartellamt(2023), Statement of objections issued against various of Google's
practices in connection with Google Automotive Services and
Google Maps Platform(Press Releases)

Bundeskartellamt(2023), Google Automotive Services proceeding - Market test regarding commitments proposed by Google(Press Releases)

Competition Bureau Canada(2024), Artificial Intelligence and competition, Discussion Paper

Competition & Markets Authority(2023), AI Foundation Models Initial Report

Competition & Markets Authority(2024), AI Foundation Models Update paper

Competition & Markets Authority(2024), Cloud services market investigation, Competitive landscape working paper

Competition & Markets Authority(2024), Microsoft Corporation’s hiring of certain former employees of Inflection and its entry into associated arrangements with Inflection(Summary)

Competition & Markets Authority(2023), Trends in Digital Markets

DOJ·FTC·EC·CMA(2024), Joint Statement on Competition in Generative AI Foundation Models and AI Products

European Commission(2024), Speech by EVP Margrethe Vestager at the European Commission workshop on “Competition in Virtual Worlds and Genertive AI”

G7 Competition Authorities and Policymakers’ Summit(2024), Digital Competition Communiqué

OECD(2024), Artificial Intelligence, Data and Competition

The White House(2023), Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence

CRFM(2021), On the opportunities and risks of foundation models

EPOCH AI(2024), Will We Run Out of Data? Limits of LLM Scaling Based on Hunman-Generated Data

HAI(2024), Artificial Intelligence Index Report 2024

Nature(2024), AI models collapse when trained on rescursively generated data

CNET(2022), See Nvidia's H100 Hopper Chip Up Close

Precedence Research(2024), Generative AI market size 2023 to 2033

Synergy research group(2024), Cloud Market Growth Stays Strong in Q2 While
Amazon, Google and Oracle Nudge Higher

생성형 AI와 경쟁 (GENERATIVE AI AND COMPETITION)

발행처 공정거래위원회

발행인 위원장 한기정

발행일 2024년 12월

편집장 경쟁정책국장 신동열

시장감시정책과장 이준헌

시장감시정책과 사무관 전해준, 조사관 임연수

본 「생성형 AI와 경쟁」의 내용에 관하여 문의사항이 있으면 아래로 연락하여 주시기 바랍니다.

(우) 30108 세종특별자치시 다솜3로 95 공정거래위원회(2동)

Tel. 044-200-4330, 4333 / Fax. 044-200-4344

발간등록번호 11-1130000-000571-01

인쇄처 문중인쇄(주)

Tel. 02-503-7764

