

정책자료 23-24

# 온라인 플랫폼 정책 수립체계 운영

2023. 12.

연구기관 : 정보통신정책연구원



● 차례 >>> 온라인 플랫폼 정책 수립체계 운영

<b>I</b>	<b>서론</b> .....	<b>1</b>
	1. 추진 배경 .....	1
	2. 포럼 개편 및 구성 .....	3
	3. 포럼 운영 경과 .....	7
<b>II</b>	<b>신기술과 플랫폼 생태계</b> .....	<b>10</b>
	1. 생성형AI 기술 동향 .....	10
	가. 대표 상업서비스 .....	10
	나. 대표 오픈소스 모델 .....	13
	다. 대규모 언어 모델 연구 동향 .....	15
	라. 경량화를 통한 미세 튜닝 대중화 .....	19
	마. AI 반도체 .....	23
	바. 생성 AI의 보안 .....	25
	사. 애플리케이션 개발을 위한 기술과 오픈소스 .....	27
	아. 정리 및 시사점 .....	36
	2. AI가 경제에 미치는 영향 .....	38
	가. AI의 경제학적 의미 .....	38
	나. AI가 고용에 미치는 영향 .....	45
	다. 성장 및 불평등에 미치는 영향 .....	58
	라. 경쟁 및 기업조직 변화 .....	63
	마. 플랫폼 경제에 대한 시사점 .....	65
	바. AI에 대한 정책적 관점 .....	70
	3. 생성형 AI가 규제 및 경쟁정책에 미치는 영향 .....	73
	가. 개요 .....	73
	나. 생성형 AI의 등장 고려 전의 플랫폼 규제 및 경쟁정책 재검토 .....	78
	다. 생성형 AI의 등장으로 인한 플랫폼 경제의 변화에 대한 예측 .....	88
	라. 새로운 정책 개발 또는 정책 변화 모색을 위한 쟁점 발굴 .....	98

다. 새로운 정책 개발 또는 정책 변화의 방향 .....	109
4. AI가 광고산업에 미치는 영향 .....	115
가. 연구목적 .....	115
나. 연구배경 .....	116
다. 광고와 마케팅의 AI 활용 .....	122
라. AI 활용 광고 관련 연구 .....	139
마. 광고산업에 AI가 미치는 영향 .....	141

### Ⅲ 플랫폼과 데이터 ..... 151

1. 저작권 및 라이선스 관련 정책 방향 .....	151
가. 서론 .....	151
나. 현황과 과제 .....	151
다. AI 모델 학습 과정에서의 문제 .....	157
라. AI 모델 활용 과정에서의 문제 .....	164
마. AI 생성물에 대한 권리와 그 귀속의 문제 .....	169
바. 생성형 AI 생태계 활성화를 위한 데이터 정책 방향 .....	172
2. 개인정보법제 개선방향 .....	176
가. 생성형 AI 현황 .....	176
나. 생성형 AI와 개인정보 관련 쟁점 .....	178
다. 국내 법제도 현황 .....	180
라. 개선방향 .....	183
3. 데이터 거래 활성화 방안 .....	187
가. 서론 .....	187
나. 사업이용자의 플랫폼 데이터 접근의 필요성 .....	188
다. 사업이용자의 플랫폼 데이터 접근 이슈 및 과제 .....	192
라. 결론 .....	202
4. 데이터 독점과 부당 이용의 이슈와 과제 .....	204
가. 데이터 독점 개념의 복합적 성격 .....	204
나. 경쟁제한성 관점에서의 데이터 독점의 이슈들 .....	205
다. 데이터 부당 이용의 문제 .....	210
라. 데이터 부당 이용의 예시적 유형 및 사례 .....	214
마. 쟁점 정리와 과제 .....	221

5. 경쟁사업자의 플랫폼 데이터 접근의 이슈와 과제 .....	223
가. 데이터 접근권에 관한 주요국의 입법 동향 .....	228
나. 데이터 접근 거절과 필수설비의 문제 .....	233
다. 데이터 접근에 관한 두 가지 규제 수단 .....	236
라. 정리 및 결론 .....	238
6. 국내외 디지털 플랫폼의 데이터 공유 및 개방에 대한 사례 연구 .....	239
가. 디지털플랫폼과 데이터 개방 .....	239
나. 디지털플랫폼의 데이터 개방의 필요성과 효과 .....	242
다. 이용사업자 대한 데이터 접근성 사례 .....	246
라. 데이터의 공유와 거래: 공공데이터화 사례 .....	271
마. 협력적 데이터 공유 .....	278
바. 플랫폼 데이터 공유를 위한 제언 .....	282

#### **IV 플랫폼 경제 측정 ..... 283**

1. 서론 .....	283
가. 연구의 필요성 및 목적 .....	283
나. 보고서의 구성 .....	285
2. 플랫폼 경쟁 .....	286
가. 플랫폼 경쟁 측정의 필요성 .....	286
나. 경쟁과 지표 .....	289
다. 판례 분석 .....	294
라. 모니터링 관점에서의 경쟁 지표와 가용데이터의 한계점과 보완점 .....	308
3. 플랫폼 영향 확산 .....	315
가. 플랫폼 영향 확산 측정의 필요성 및 범위 .....	315
나. 기존 연구 .....	317
다. 플랫폼 영향 확산 측정지표, 가용데이터 및 한계점 .....	327
4. 자국 플랫폼의 가치 .....	335
가. 연구 목적 및 필요성 .....	335
나. 로컬 플랫폼의 정의 .....	336
다. 로컬 플랫폼의 의미와 가치 .....	341
라. 로컬 플랫폼 가치 측정 방법론: 선행 연구 정리 .....	345
마. 시사점 및 결론 .....	355

5. 국내 실태조사 현황 및 향후 과제 .....	358
가. 플랫폼 경제 측정 지표 및 방법론 .....	358
나. 국내 플랫폼 경제 실태조사 .....	362
다. 국내 플랫폼 실태조사 현황 평가 및 향후 과제 .....	371

● 표 차례 >>> 온라인 플랫폼 정책 수립체계 운영

<표 I-1> 플랫폼 경제 자문위원회 구성 ..... 5

<표 I-2> 포럼 워킹그룹 구성 ..... 5

<표 I-3> 포럼 워킹그룹 운영 경과 ..... 9

<표 I-4> 최종보고회 발표내용 ..... 9

<표 II-1> 한지우·오삼일(2023)에 따른 직업별 시노출도 ..... 57

<표 II-2> 소비자 여정 단계별 AI 활용 방안 및 사례 ..... 125

<표 III-1> 주요국 데이터 거래소 현황 ..... 193

<표 III-2> 데이터 사업 수행 시 애로사항 ..... 194

<표 III-3> 차별비용과 가격차별 효율성의 관계 ..... 216

<표 III-4> 이용 가능한 네이버 기본 데이터 ..... 247

<표 III-5> 네이버 데이터 애널리틱스 서비스 제공 데이터 ..... 250

<표 III-6> 네이버 애널리틱스 데이터 분석 자료 ..... 253

<표 III-7> 배민셀프서비스 제공 데이터 ..... 259

<표 III-8> 우버의 정보 수집 및 이용 ..... 270

<표 III-9> 직방 공공데이터 활용 및 수요 현황 ..... 273

<표 III-10> 구글 클라우드 추천 데이터세트 ..... 275

<표 IV-1> EU 플랫폼 경제관측소 경제측정 보고서의 구성 ..... 284

<표 IV-2> 코로나 기간 고령층의 음식 배달 서비스 이용 ..... 328

<표 IV-3> 사업체의 디지털 플랫폼 거래 현황, 2020 ..... 330

<표 IV-4> 서비스업 사업체가 이용 중인 온라인 플랫폼 수, 2021 ..... 331

<표 IV-5> 온라인 vs 온/오프라인 채널 이용 상점 간 효율성 비교 예시 (Ryu et al., 2019) ..... 349

<표 IV-6> 자국 검색 보유 여부에 따른 광고시장 규모 추정 결과 (자료: Ji et al., 2019) ..... 351

● **그림 차례** >>> 온라인 플랫폼 정책 수립체계 운영

[그림 I-1] 플랫폼 자율기구 구성 현황 ..... 1

[그림 I-2] 각 플랫폼 기업별 자율규제 추진 현황 및 계획 ..... 1

[그림 I-3] 백만 이용자 달성에 걸린 시간 ..... 2

[그림 I-4] 2023 디지털 플랫폼 정책포럼 구성 ..... 4

[그림 I-5] 2023 디지털 플랫폼 정책포럼 운영 경과 ..... 7

[그림 II-1] 폐쇄형과 오픈소스 모델 비교 (출처 ARK-INVEST) ..... 15

[그림 II-2] LoRA 어댑터 개념도 ..... 22

[그림 II-3] 기술개발이 직무의 변화와 노동, 자본에 미치는 영향 ..... 42

[그림 II-4] AI의 GPT 및 IMI적 특성과 타 기술과의 비교 ..... 44

[그림 II-5] 새로운 직무의 생성 및 기존 직무 대체로 인한 고용 및 임금 효과 ..... 48

[그림 II-6] AI 자동화가 고용에 미치는 영향에 대한 국가별 비교 ..... 50

[그림 II-7] 임금 수준별 생성형 GPT로 인한 직무 노출도 ..... 51

[그림 II-8] ChatGPT의 활용과 글쓰기 품질의 격차 완화 ..... 53

[그림 II-9] ChatGPT 출시 이후의 글쓰기 프리랜서 고용변화 ..... 55

[그림 II-10] 글로벌 임금수준에 따른 임금상승률 비교 ..... 61

[그림 II-11] 대형 플랫폼 사업자의 리스크 ..... 85

[그림 II-12] 인공지능 기술 적용의 리스크 ..... 86

[그림 II-13] 개발 과정으로 본 AI 가치사슬(Value Chain) ..... 89

[그림 II-14] 생태계 구성으로 본 AI 가치사슬(AI 스택) ..... 91

[그림 II-15] 글로벌 대규모 언어 모델의 현황 ..... 93

[그림 II-16] 하류 FM 서비스 시장에서의 FM 활용 ..... 104

[그림 II-17] 글로벌 디지털 광고비 성장 ..... 116

[그림 II-18] 업무별 AI 관련 비용 증감세 추세 ..... 120

[그림 II-19] AI의 현재 영향과 미래의 영향 예측 ..... 121

[그림 II-20] AI 활용 마케팅 랜드스케이프 ..... 122

[그림 II-21] 프로그래매틱 바이딩 프로세스 ..... 123

[그림 II-22] AI를 이용하는 마케터 비중과 이용 용도 현황 ..... 125

[그림 II-23] AI 접목 광고 캠페인 사례 유형 ..... 126

[그림 II-24] 캐드베리 캠페인 ..... 128

[그림 II-25] 하인즈 캠페인 .....	129
[그림 II-26] 코카콜라 캠페인 .....	130
[그림 II-27] 코카콜라 캠페인 II .....	130
[그림 II-28] 구글의 랜딩 페이지 텍스트 생성형 AI 인터페이스 .....	131
[그림 II-29] 구글의 자연어를 활용한 검색 경험(SGE) .....	132
[그림 II-30] 메타 AI 활용 광고 플랫폼 .....	134
[그림 II-31] 아마존 AI 활용 광고 플랫폼 .....	135
[그림 II-32] 마이크로소프트 AI 활용 광고 플랫폼 .....	136
[그림 II-33] 틱톡 AI 활용 광고 플랫폼 .....	137
[그림 II-34] 네이버 AI 활용 광고 플랫폼 .....	138
[그림 II-35] 네이버 스마트스토어 클로바MD 적용 및 클로바 라이브챗 예시 .....	138
[그림 II-36] AI 광고 관련 논문 수 .....	139
[그림 II-37] AI 광고 관련 논문 워드 클라우드 .....	140
[그림 II-38] AI 광고 연구 맥락과 주제 .....	140
[그림 II-39] 네이버 클로바X 활용 현재와 미래 .....	141
[그림 II-40] 인공 창의성 랜드스케이프 .....	142
[그림 II-41] 인공 창의성 도구 매트릭스 .....	143
[그림 II-42] 생성형 AI 시장 규모 .....	144
[그림 II-43] 생성형 AI 활용 광고 가이드라인 예시 .....	150
[그림 III-1] 대규모 언어모델(LLM) 발전 개요도 .....	177
[그림 III-2] 생성형 AI 생태계 .....	178
[그림 III-3] 플랫폼 경제에서의 데이터 피드백 순환과 네트워크 효과 .....	206
[그림 III-4] 플랫폼 개방의 선순환 구조 .....	240
[그림 III-5] 클라우드 데이터 박스 서비스 콘셉트 .....	246
[그림 III-6] 네이버 애널리틱스와 스마트스토어 통계(Biz Advisor)의 연계 관계 .....	257
[그림 III-7] 구글 애널리틱스 플랫폼 구성요소 .....	262
[그림 III-8] 구글 애널리틱스 허브의 데이터 교환 모델 .....	264
[그림 III-9] 아마존 AWS 마켓플레이스 CAS 개요 .....	267
[그림 IV-1] 전통적 경쟁 지표와 플랫폼 경쟁 지표 .....	293
[그림 IV-2] 플랫폼 관련 사례에서 사용된 지표 .....	294
[그림 IV-3] 경쟁 지표의 출처 및 한계점 .....	310
[그림 IV-4] 경쟁 지표 출처의 보완점 .....	313
[그림 IV-5] 소매업 매출의 온라인 쇼핑 비중, 2021 .....	317
[그림 IV-6] 서울 내 호텔과 에어비앤비의 지리적 분포 .....	322

[그림 IV-7] 코로나19와 배달 이용 여부에 따른 음식점 매출 변화 .....	325
[그림 IV-8] 시군구 지역별 온라인 쇼핑 비중 변화: 2020년과 2015년 .....	329
[그림 IV-9] 로컬 플랫폼 구분을 위한 기준 예시 .....	339
[그림 IV-10] 국내 검색 점유율 변화 추이 .....	342
[그림 IV-11] 주요 인터넷 서비스의 연간 소비자잉여를 계산 예 Brynjolfsson&Oh (2012) ....	347
[그림 IV-12] 메타프론티어(MetaFrontier) 그룹 간 효율성 비교 예시 .....	348
[그림 IV-13] 검색시장 관련 연계 산업의 예시 .....	350
[그림 IV-14] 온라인 광고시장의 크기 (자료: Ji et al., 2019) .....	350
[그림 IV-15] An AHP model for evaluating internet platforms .....	353

## 요 약

### I 서론

- 디지털 플랫폼을 둘러싼 규제 논란 지속: 규제 vs 자율규제
- 플랫폼에 대해서는 자율규제를 우선한다는 현 정부 정책 기조로 민간 플랫폼 자율기구 4개 분과(1분과: 갑을, 2분과: 소비자·이용자, 3분과: 데이터·AI, 4분과: 혁신공유·거버넌스)가 운영 중이나, 자율규제 정책 방향에도 불구하고 다수의 온라인 플랫폼 관련 법안이 발의
- 생성형 AI라는 신기술 부상으로 플랫폼 생태계 변화
- 생성형 AI가 빠르게 확산되고, 주요 플랫폼 기업들이 이를 경쟁력 제고를 위한 수단으로 활용하기 시작하면서 생성형 AI가 플랫폼 생태계, 플랫폼 경쟁에 미치는 영향에 대한 관심도 증대

#### □ 2023 포럼 구성 및 운영 방향

- 2023년 포럼은 플랫폼의 핵심 의제에 대한 면밀한 분석을 통해 정책 방향 수립에 활용될 수 있는 근거자료를 축적할 수 있는 방향으로 개편
- 자문위원회 회의를 통해 2023년 포럼의 핵심 의제로 플랫폼 생태계의 새로운 화두로 부상한 생성형 AI, 플랫폼 정책의 핵심 현안 중 하나인 데이터, 증거 기반 정책 수립을 위한 플랫폼 경제 모니터링을 선정
- 각 의제별로 자문위원을 중심으로 워킹그룹을 구성하고, 워킹그룹의 핵심 연구 주제 및 필요성에 따라 외부 집필의원 참여를 통해 의제 논의 및 정책 방향 도출

## II 신기술과 플랫폼 생태계

### □ 생성형 AI기술동향

- 상업적인 폐쇄형 생성 AI 서비스뿐만 아니라 오픈 소스 모델의 발전 추이가 가파르며 폐쇄형 서비스로 제공하는 상용 언어 모델의 선두 그룹과의 차이가 줄어들고 있음이 관찰됨
- 대규모 언어 모델은 인간의 고수준의 추론 능력과의 차이를 좁히는 방향으로 연구가 집중되고 있으며, 후속 언어 모델은 고수준의 추론 능력 확보에 집중할 것으로 보임
- 경량화와 미세튜닝 방법론 연구 발전으로 대규모 언어 모델이 다양한 도메인에서 성능을 개선하고 활발하게 이용될 것이며, AI 전용 가속칩의 발전으로 언어 모델 학습/추론의 속도와 효율이 개선되어 생성 AI의 대중화에 기여할 것으로 전망됨

### □ AI가 경제에 미치는 영향

- AI 또는 생성형 AI의 특성이 플랫폼적 특성을 강화할 것인지의 여부는 단정지을 수 없음
  - 플랫폼의 가장 큰 특성인 양면시장이나 간접적 네트워크 외부성은 AI와는 연관이 없고, 규모의 경제효과는 존재하지만 이는 대규모 연구개발 투자를 필요로 하는 디지털 기술이 공유하는 특성으로 판단되며 전환비용과 고착효과도 플랫폼에 반드시 유리한 방향으로 작용하리라는 보장은 없음
  - 플랫폼의 데이터 축적 및 활용능력은 AI 경쟁에 있어 유리한 요소로 작용할 수 있고 기존 플랫폼들의 재정적 기반, 선발자의 이익, 빅테크로서의 노하우 등이 AI 경쟁에 있어 앞선 위치를 차지하는 원동력이 될 것이며 플랫폼 간의 경쟁에 있어서 AI 개발 및 서비스 능력이 새로운 차원의 경쟁요소가 될 가능성은 높음

### □ 생성형 AI가 규제 및 경쟁정책에 미치는 영향

- 생성형 AI가 기존 거대 플랫폼 외부에서 이를 기반으로 한 새로운 서비스의 가치 창출 체계를 형성하면서 시장경쟁에 긍정적으로 작용할 수도 있으나, 생성

형 AI가 강한 규모와 범위의 경제에 따르고 파운데이션 모델을 중심으로 자기 강화적인 성장 사이클을 가진 근본적인 집중화 성격을 보여 모델과 애플리케이션 개발 장벽을 형성할 우려가 있다는 부정론도 존재

- 인공지능 기술이 적용되는 생태계에 대한 규제에는 기존의 규제가 기초를 두고 있는 사업모델이 아닌 대안 사업모델로 중대한 편익과 높은 이용자 보호 수준을 달성할 가능성을 고려하여 새로운 형태의 혁신 거래(innovation deals)에 개방적인 태도가 필요
- 파운데이션 모델 제공자에 대한 혁신 지원 정책 및 디지털 플랫폼 정책보다 진취적인 정책 목표 설정과 규제접근이 필요

#### □ AI가 광고산업에 미치는 영향

- AI의 발달, 특히 최근의 생성형 AI의 급속한 발달과 확산은 광고와 인접 산업에 막대한 영향을 미치고 있으며, 특히 광고의 효과와 효율성을 극대화하는 데 중요한 역할을 하고 있어 광고가 주요 비즈니스 모델인 주요 글로벌, 로컬 플랫폼들은 AI를 적극적으로 활용하여 경쟁력을 확보하려 함
- AI가 광고 제작과 매체 거래 등에 적극적으로 활용됨에 따라 광고의 창의성과 독창성, 소비자의 혼동, 전문인력의 대체, 개인정보 보호 등이 AI와 관련한 광고 영역의 주요 이슈로 부상

### III 플랫폼과 데이터

#### □ 저작권 및 라이선스 관련 정책 방향

- 생성형 AI 생태계가 확장되면서 생성형 AI 활용 과정에서 생길 수 있는 데이터와 관련된 법적 쟁점들을 살펴보고 해결방안을 모색할 필요
- 법적 쟁점들을 AI 모델 학습 과정, AI 모델 활용 과정, AI 생성물에 대한 권리와 그 귀속의 측면에서 살펴보고 독점과 공유의 조화, 법적 불확실성 해소, 이해관계의 조정, 플랫폼의 기능과 역할에 대해 제안

## □ 개인정보법제 개선 방향

- 생성형 AI와 개인정보 관련 쟁점을 AI학습, 설명가능한 AI, 공적/공공적 활용, 자율주행 차량 측면에서 살펴보고 관련 국내 법제도 현황을 분석
- 학습데이터 수집/이용, 가명정보 특례, 설명의무, 행정조사/범죄수사, 사회적 평판 부여, 취업정보 활용 등 공공적 활용 등과 관련된 개선방향 제안

## □ 데이터 거래 활성화 방안

- 현재 국내 데이터 거래 및 유통 체계 상황을 살펴보고, 데이터 거래 활성화 관련 이슈와 과제에 대해 고찰
- 유통 기반 측면에서 데이터에 대한 정보 부족 및 이종 산업 간 데이터 융합의 한계점, 가격 측면에서 명확한 가격산정 기준의 부재, 품질 측면에서 산업별 데이터 연계를 위한 데이터 표준화 등 이슈 분석과 정책방안 제시

## □ 데이터 독점과 부당 이용의 이슈와 과제

- 데이터 독점을 경쟁제한성 관점과 데이터 접근의 관점에서 데이터 부당 이용 및 소유권의 문제를 중심으로 살펴보고, 데이터 부당 이용의 대표적인 사례로 아마존의 판매자 데이터 활용, 페이스북의 이용자 정보 결합 활용을 분석
- 데이터 독점 및 배타적 활용에 따른 체계적인 비대칭 우위 및 일방적 차별 이슈, 시장지배력 확보 및 유지 수단으로 활용, 적정 수준보다 과다 이용에 따른 비효율성 문제 등의 쟁점을 해결해 나가는 정책방향 제안

## □ 경쟁사업자에 대한 데이터 접근의 법적 이슈와 과제

- 경쟁사업자에 대한 데이터 접근의 경쟁정책적 의미를 고찰하고 데이터 접근·사용 방해 관련 국내외 주요 사례를 살펴봄. 이와 함께 데이터 접근권 관련 입법 동향, 데이터 접근 거절과 필수 설비의 문제, 데이터 접근에 관한 사전규제와 사후규제를 중심으로 분석
- 현행법상 데이터 접근이나 제공 거절에 대한 방어적, 수동적, 사후적 접근법의 효용과 한계가 현재 가시적으로 드러나는 상황에서 데이터 접근에 따른 법적, 기술적, 계약적 제약을 극복해야 하는 것이 관건, 이 점에서 데이터 접근에 대한 사전규제를 공적 규제 방식으로만 설계할 것인지에 대해서는 신중한 접근과 모색이 필요

## □ 국내외 디지털 플랫폼의 데이터 공유 및 개방에 대한 사례 연구

- 플랫폼 경제에서 데이터 공유는 기술 혁신의 가속화 및 사회적·경제적 가치 창출에 기여. 이에 디지털 플랫폼의 데이터 개방의 필요성과 효과, 데이터 개방의 유형 및 분류에 대한 분석
- 네이버, 배달의 민족, 구글, 아마존, X(구 트위터), 우버 등 이용사업자에 대한 데이터 접근성 사례와 함께 데이터 공유와 거래, 협력적 데이터 공유 사례에 대한 분석을 토대로 플랫폼 데이터 공유를 위한 제언

## IV 플랫폼 경제 측정

### □ 플랫폼 경제측정

- 플랫폼의 경제적 영향력 및 규모가 확대되고 있어 이들에 대한 정책 입안의 필요성이 증대되고 있으나, 이들의 경제적 역할이나 행위에 대한 데이터는 부족한 상황이며, 이러한 문제의식하에 유럽연합의 플랫폼 경제관측소(EU Observatory on the Online Platform Economy)는 2021년 플랫폼 경제 측정에 관한 보고서를 발간
- 본 장에서는 EU 플랫폼 경제관측소의 플랫폼 경제측정 보고서를 벤치마크하되, 우리나라 실정에 맞고 플랫폼 경제분석과 보다 유기적으로 연결된 플랫폼 경제 모니터링 체계 구축 방안을 제시하는 것을 목적으로 함

### □ 플랫폼 경쟁

- 플랫폼 시장과 같이 동태적이고 혁신적인 시장일수록 경쟁의 '과정'을 측정하는 것이 중요하며, 또한 해당 시장 고유의 특성인 다면시장, 직간접 네트워크 효과 및 데이터에 의한 진입장벽 등을 고려한 경쟁의 측정이 필요하므로 전통적인 지표로는 한계 존재
- 경쟁의 측정을 위해 가장 중요한 시장획정의 경우, 산업 정책적 관점에서의 시장 획정은 산업의 발전과 공정한 생태계의 조성을 위한 경쟁촉진 등 다방면으로 모니터링을 위해 선정하는 것이 바람직할 것이며, 이는 곧 경쟁법 집행에서의 시장획정과는 차별화 됨을 뜻함

- 현재 널리 사용되고 있는 경쟁 측정지표들은 전통적으로 사용해 왔던 지표들에 플랫폼의 특성을 반영한 시장구조와 성과 관련 변수가 대부분이며, 간접네트워크효과와 멀티호밍 등과 같이 타 산업에서는 드물게 쓰였던 지표들이 플랫폼 경쟁에 대한 지표로 사용되기도 함
- 현행 플랫폼 경쟁 측정지표는 측정의 목적과 주체에 따라 주기, 시차, 범위, 깊이 등의 차원에서 각자의 장점과 한계점을 가지고 있으며, 이에 따라 정부 부처의 플랫폼 조사는 부처 간 조사의 일관성과 정기성을 가지고 장기적으로 수행하되, 역동성 등 다른 차원을 측정할 수 있는 민간 자료를 통한 보완이 필요

#### □ 플랫폼 영향확산

- 플랫폼 경제 확산의 중요성은 플랫폼 기업과 산업 자체에 대한 측정 및 평가뿐만 아니라 플랫폼을 이용하는 소비자와 생산자 및 산업에 대한 영향력에 대한 측정과 분석이 필요하며, 특히 플랫폼 경제의 확산이 공급자, 소비자 간 혜택의 불균등한 배분으로 이어지는 경우 이해 갈등 촉발의 원인으로 작용할 수 있기 때문에 이에 대한 구체적인 분석이 이루어져야 함
- 소비자 측면에서는 개별 플랫폼이 소비자 후생에 영향을 주는 메커니즘을 파악하는 것이 중요한데 이는 플랫폼 등장에 따른 가격 하락 효과나 제품 다양성 증가 효과 등을 통해 측정이 가능하며, 소비자 혜택의 확산이 체계적으로 불균등하게 분포되고 있는지 등을 파악함으로써 이해 갈등 해결의 시작점을 마련할 수 있음
- 생산자 측면에서는 개별 생산자의 플랫폼 매출의존도, 멀티호밍/싱글호밍 여부 등을 파악하고, 플랫폼 확산에 따른 기존 산업의 구조변화를 모니터링 하는 것이 필요하며, 이를 통해 전통산업과의 이해 갈등 문제 해결의 시작점을 마련할 수 있음
- 노동시장 측면에서는 플랫폼 종사자 증가에 따라 새로운 일자리에 대한 관찰뿐 아니라 기존 산업의 일자리에서의 노동 유출방식에 대해서도 관찰이 필요하며, 디지털 플랫폼이 물리적 한계를 뛰어넘는다는 특성을 감안할 때, 기존 지역 시장에서의 산업 구성을 변화시키는 역할을 수행한다는 점에서 지역경제에 대한 모니터링도 필요

## □ 자국 플랫폼의 가치

- 자국 플랫폼 보유는 글로벌 빅테크 견제, 인터넷 주권, 데이터 경제 패권 및 문화주권 확보, 자국 콘텐츠의 확산 등의 측면에서 중요한 역할을 수행하는 바 세계 각국이 자국 플랫폼 확보 경쟁에 열을 올리고 있음
- 자국 플랫폼의 가치에 대한 경제학적 분석은 그러나 많은 한계점을 내포하고 있는데, 자국 플랫폼 보유국과 미보유국 간의 국가 간 비교의 경우, 경제학적 모형 설계의 어려움 및 가용한 데이터가 디지털 플랫폼 경제 측정에 적합하지 않다는 문제가 존재
- 자국 플랫폼을 보유한 나라가 많지 않아 해당 주제에 대한 연구는 우리나라가 주도적으로 해나가야 하는 상황이며, 국가 간 비교 등을 위한 거시지표의 개발 등이 아직 진행 중이라는 점을 감안할 때, 이용자 및 전문가 설문 등을 통한 데이터의 축적 및 확률적 변경 접근법, 컨조인트 분석, 계층화 모형 분석 등 미시적 모형을 통한 분석 및 모니터링을 고려할 필요



# 제 I 장 서론

## 1. 추진 배경

□ 디지털 플랫폼을 둘러싼 규제 논란 지속: 규제 vs 자율규제

- 플랫폼에 대해서는 자율규제를 우선한다는 현 정부 정책 기조로 민간 플랫폼 자율기구 4개 분과(1분과: 갑을, 2분과: 소비자·이용자, 3분과: 데이터·AI, 4분과: 혁신공유·거버넌스)가 운영 중이며 다음과 같은 자율규제 관련 결과 발표

[그림 I-1] 플랫폼 자율기구 구성 현황



- '23년 5월 자율규제 방안 발표회를 통해 각 분과별로 오픈마켓 분야 자율규제 방안, 오픈마켓 소비자 집단피해 신속 대응방안, 검색·추천 서비스 투명성 제고를 위한 자율규제 원칙, 플랫폼의 사회가치 제고를 위한 8대 원칙 발표
- '23년 8월 범부처 플랫폼 정책협의체 및 기업간담회를 통해 네이버, 카카오, 쿠팡, 우아한형제들, 당근 등 우리나라 대표적 플랫폼 기업들이 자율규제\* 현황 및 향후 계획 발표

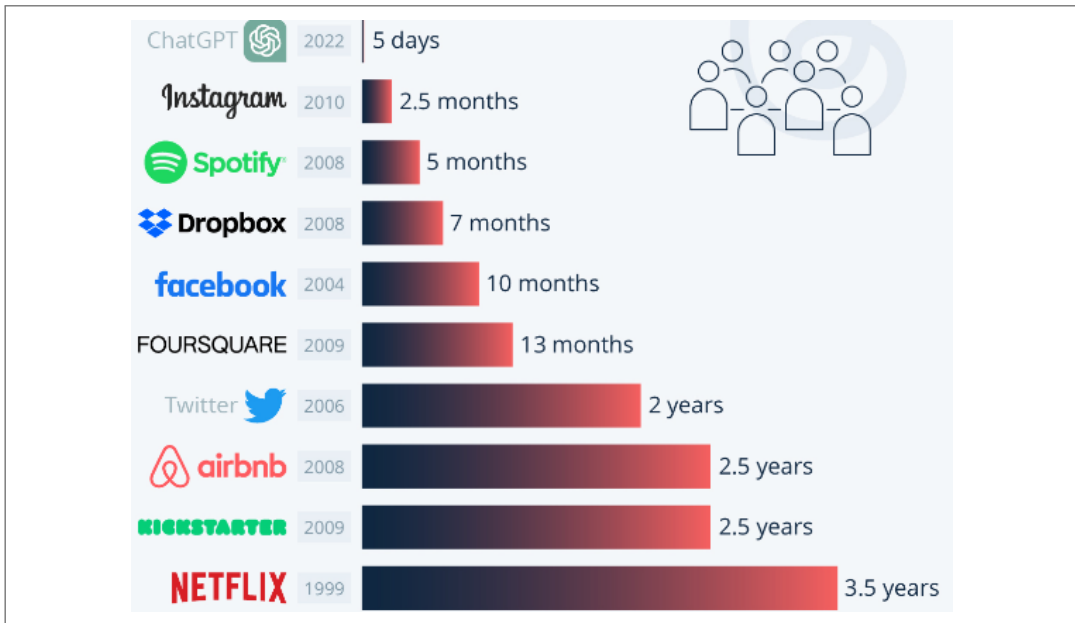
\* 민간 플랫폼 자율기구 활동 외에 기업 자체적으로 추진하는 자율규제 활동

[그림 I-2] 각 플랫폼 기업별 자율규제 추진 현황 및 계획



- '23년 11월, 플랫폼 상생협력 확산을 위한 업계 간담회 등을 통해 이용사업자들과의 상생협력 방안 발표
  - 플랫폼 자율규제의 법적 근거를 담은 전기통신사업법 일부 개정안이 국무회의에서 의결
  - 이처럼 플랫폼과 관련된 자율규제 정책 방향에도 불구하고 다수의 온라인 플랫폼 관련 법안이 발의되었으며, '23년 12월, 플랫폼 경쟁촉진법(가칭) 추진이 국무회의에서 논의 예정
- 생성형 AI라는 신기술 부상으로 플랫폼 생태계 변화
- '22년 말, 오픈AI에서 생성형 AI 기반 챗봇서비스 챗GPT의 제공을 시작한 이후 5일만에 백만 이용자를 달성하면서 빠르게 성장하며 주목

[그림 I-3] 백만 이용자 달성에 걸린 시간



출처: statista(2023.7.7.)

- 챗GPT는 이후 여행 예약, 쇼핑 등 외부 서비스를 활용할 수 있는 플러그인 기능과 사용자가 쉽게 맞춤형 챗봇을 만들고 이를 게시·판매할 수 있는 GPT스토어를 도입하는 등 서비스를 플랫폼화

- 알파벳, 메타 등 빅테크 또한 생성형 AI 파운데이션 모델을 공개하며 생성형 AI 경쟁에 돌입, 빅테크 플랫폼 기업들은 생성형 AI를 자사 제품과 서비스에 적용하는 등 생성형 AI 주도권 확보를 위한 움직임 본격화
- 생성형 AI가 빠르게 확산되고, 주요 플랫폼 기업들이 이를 경쟁력 제고를 위한 수단으로 활용하기 시작하면서 생성형 AI가 플랫폼 생태계, 플랫폼 경쟁에 미치는 영향에 대한 관심도 증대

## 2. 포럼 개편 및 구성

### □ 2022 포럼의 구조적 한계 및 개편 필요성

- 2022년 포럼은 4개의 분과로 구성되어 있는데, 진흥과 규제가 분리되어 있고 정책 수단과 이슈가 혼재되어 있어 각 분과별 주제가 중복되거나 도출된 정책 대안의 구체성이 떨어지는 한계 내포
  - (포럼 구성) 강건한 산업 생태계를 위한 정책 방안을 논의하는 1분과, 혁신과 공정의 생태계를 위한 정책 방향을 논의하는 2분과, 플랫폼의 사회적 가치 제고를 위한 정책방향을 논의하는 3분과, 데이터의 보호와 활용의 균형을 위한 정책방향을 논의하는 4분과로 구성
- 또한 기존 포럼 분과와 자율기구 분과와의 기능이 중복되는 등 기능 재조정의 필요성도 존재
- 산업 진흥을 위한 가장 효과적인 정책 대안 중 하나가 규제 개선이라는 점에서도 법제도와 산업 진흥을 별도의 분과로 운영하기 보다는 핵심 의제를 종합적 관점에서 논의할 수 있는 구조로 개편 필요

### □ 2023 포럼 구성

- 2023년 포럼은 플랫폼의 핵심 의제에 대한 면밀한 분석을 통해 정책 방향 수립에 활용될 수 있는 근거자료를 축적할 수 있는 방향으로 개편
- 이를 위해 EU의 ‘Observatory on the Online Platform Economy’의 전문가 그룹을 참고하여 학계의 플랫폼 전문가를 중심으로 한 플랫폼 경제 자문위원회 구성

- EU에서는 온라인 플랫폼과 관련된 정책 결정을 위해서 산업의 변화과정을 모니터링할 필요성을 인지하고, 온라인 플랫폼에 대해 증거기반 및 문제 중심 정책 결정을 위한 온라인 플랫폼 경제의 발전을 모니터링하는 전문가 그룹을 2018년에 설립
- 2023년 포럼의 플랫폼 경제 자문위원회는 위원장 1인과 10인의 자문위원으로 구성되었으며, 각 자문위원은 경제, 법, 미디어, 기술 등 다양한 분야에서 플랫폼 관련 연구를 활발히 진행하고 있는 전문가로 선발

[그림 I-4] 2023 디지털 플랫폼 정책포럼 구성



〈표 I-1〉 플랫폼 경제 자문위원회 구성

이름	소속	전공 분야
이원우 (위원장)	서울대학교	법학
권남훈	건국대학교	경제학
전현배	서강대학교	경제학
최윤정	연세대학교	경제학
홍대식	서강대학교	법학
이성엽	고려대학교	법학
유병준	서울대학교	경영학
류민호	동아대학교	경영학
최세정	고려대학교	미디어
황용석	건국대학교	미디어
오장민	성신여자대학교	공학(AI)

- 자문위원회 회의를 통해 2023년 포럼의 핵심 의제를 다음과 같이 선정하고, 각 의제별 워킹그룹을 구성하여 정책 의제별 세부 연구주제에 대한 논의 및 집필
  - 각 의제는 미래지향적 관점의 이슈, 현안 이슈, 지속적으로 살펴보아야 하는 이슈라는 3가지 관점에서 플랫폼 생태계의 새로운 화두로 부상한 생성형 AI, 플랫폼 정책의 핵심 현안 중 하나인 데이터, 증거 기반 정책 수립을 위한 플랫폼 경제 모니터링을 선정
  - 각 의제별로 자문위원을 중심으로 워킹그룹을 구성하고, 워킹그룹의 핵심 연구 주제 및 필요성에 따라 외부 집필의원 참여

〈표 I-2〉 포럼 워킹그룹 구성

핵심이제별 워킹그룹	자문위원	집필위원
신기술과 플랫폼 생태계	권남훈, 오장민, 최세정, 홍대식	-
플랫폼과 데이터	유병준, 이성엽, 황용석	김성환, 이상용, 이승민, 신영수
플랫폼 경제 모니터링	류민호, 전현배, 최윤정	-

□ 핵심 정책의제에 따른 워킹그룹별 세부 연구 주제

- (신기술과 플랫폼 생태계) 신기술과 플랫폼 생태계 워킹그룹에서는 ①생성형 AI 기술로드맵, ②AI 발전의 경제적 영향과 시사점, ③생성형 AI가 경제규제 및 경쟁정책에 미치는 영향, ④AI가 광고산업에 미치는 영향을 논의
- (플랫폼과 데이터) 플랫폼과 데이터 워킹그룹에서는 크게 생성형 AI 생태계 활성화 위한 플랫폼 데이터 정책과 공정하고 자유로운 플랫폼 데이터 이용을 위한 정책 방향을 논의 하였으며, 세부 주제는 다음과 같음
  - (생성형 AI 생태계 활성화) 생성형 AI와 관련하여 ①저작권 및 라이선스 관련 정책 방향, ②개인정보법제 개선 방향, ③데이터 거래 활성화를 위한 제언을 중점적으로 논의
  - (공정하고 자유로운 플랫폼 데이터 이용) 플랫폼에서의 데이터 이슈를 플랫폼 사업자, 경쟁 사업자, 이용사업자로 구분하여 ①데이터 독점과 부당이용의 이슈와 과제, ②경쟁사업자에 대한 데이터 접근의 이슈와 과제, ③ 디지털 플랫폼의 데이터 공유와 개방 사례에 대해 논의
- (플랫폼 경제 측정) 플랫폼 경제의 발전을 모니터링하기 위해 필요한 지표들과 자국 플랫폼의 가치를 논의하였으며 세부 주제는 ①플랫폼 경쟁, ② 플랫폼의 영향 확산, ③ 자국 플랫폼의 가치로 구성

### 3. 포럼 운영 경과

#### □ 포럼 운영경과 개요

- 디지털플랫폼 정책포럼 킷오프 회의(3.7)
- 정책포럼 의제 선정 및 워킹그룹 구성(3~4월)
- 워킹그룹 운영 및 전문가초청세미나(4~11월)
- 자문위원회(6.29)
- 연구결과 공유 및 논의(11월)
- 디지털플랫폼 정책포럼 최종보고회(12.11)

[그림 I -5] 2023 디지털 플랫폼 정책포럼 운영 경과



#### □ 킷오프 회의

- 일시/장소 : '23. 3. 7(화), 한국 과학기술회관 대회의실
- '23년 포럼 운영계획 및 EU 플랫폼 모니터링센터 운영현황 등 발표, '23년 포럼 연구·분석 핵심의제 논의
  - 플랫폼 정책방향(자율규제 등)에 대한 제언 위주에서 핵심의제에 대해 심도 있게 논의할 수 있는 체제로 개편(기존 운영위를 '플랫폼 경제 자문위'로 확대·개편 및 '플랫폼 현장 자문단' 구성)

- 플랫폼 자율규제 관련 전기통신사업법 개정 방안 주요 내용발표 및 의견수렴
  - △자율기구 설립 근거, △정부의 독려시책 마련, △자율규제기구 지원, △이해관계자 의견 수렴 등 투명성 확보, △참여 인센티브(자율규제 활동, 성과 등 고려) 등

□ 자문위원회

- 일시/장소 : '23. 6. 29(목), 강남 페이토 대회의실
- '23년 포럼 워킹그룹 주요 논의 경과 및 향후 운영 계획 공유
  - (워킹그룹1: 신기술과 플랫폼 생태계) 생성형AI 등 기술 진화에 따른 플랫폼 생태계 패러다임 변화와 정책 이슈 등 분석
  - (워킹그룹2: 플랫폼과 데이터) 초거대AI 생태계 혁신 활성화를 위한 플랫폼 데이터 정책방향 및 데이터의 공정하고 자유로운 활용을 위한 정책방향 등 논의
  - (워킹그룹3: 플랫폼 경제 측정) 플랫폼 경제 현황과 토종 플랫폼의 가치 측정 등을 위한 해외사례 분석 및 필요 데이터 식별·수집방안 등 논의
- 국내외 플랫폼 정책 동향 발제 및 플랫폼 현안 이슈 등 토론

□ 정책포럼 워킹그룹 운영(4~10월)

- 정책포럼 의제검토 및 선정회의(3~4월)
  - 1차 회의(3.16), 2차 회의(3.21), 3차 회의(3.24), 4차 회의(4.11)
- 워킹그룹 킷오프 회의(4월)
  - 플랫폼과 데이터 워킹그룹 킷오프 회의(4.18)
  - 플랫폼 경제 모니터링 워킹그룹 킷오프 회의 (4.25)
  - 신기술·플랫폼 생태계 워킹그룹 킷오프 회의 (4.26)
- 워킹그룹 운영 및 전문가 세미나(5~11월)

〈표 I-3〉 포럼 워킹그룹 운영 경과

	[워킹그룹 1] 신기술·플랫폼 생태계	[워킹그룹 2] 플랫폼과 데이터	[워킹그룹 3] 플랫폼 경제 모니터링
4월	kickoff 회의(4.26)	kickoff 회의(4.18)	kickoff 회의(4.25)
5월	2차회의(5.15)		2차회의(5.15)
6월	3차회의(6.14)		3차회의(6.12)
7월	4차회의(7.17)	2차회의(7.6)	
8월	5차회의(8.8)		4차회의(8.22)
9월		3차(9.5), 4차(9.18)	
10월	6차회의(10.16)	5차회의(10.06)	전문가 세미나(10.30)
11월	7차(11.3), 8차(11.27)	6차(11.10), 7차(11.27)	5차회의(11.20)

□ 최종보고회

- 일시/장소 : '23. 12. 11(목), 강남 안다즈 스튜디오
- '23년 디지털 플랫폼 정책포럼 주요 논의결과 발표

〈표 I-4〉 최종보고회 발표내용

워킹그룹	발표 내용
[워킹그룹1] 신기술·플랫폼 생태계	<ul style="list-style-type: none"> <li>• 생성형 AI기술 로드맵</li> <li>• AI발전의 경제적 영향과 시사점</li> <li>• 생성형 AI가 경제규제 및 경쟁정책에 미치는 영향</li> <li>• AI가 광고산업에 미치는 영향</li> </ul>
[워킹그룹2] 플랫폼과 데이터	<ul style="list-style-type: none"> <li>• 저작권 및 라이선스 관련 정책 방향</li> <li>• 개인정보법제 개선방향</li> <li>• 데이터 거래 활성화 방안</li> <li>• 데이터 독점과 부당 이용의 이슈와 과제</li> <li>• 경쟁사업자에 대한 데이터 접근의 이슈와 과제</li> <li>• 디지털플랫폼의 데이터 공유와 개방 사례 분석</li> </ul>
[워킹그룹3] 플랫폼 경제 측정	<ul style="list-style-type: none"> <li>• 플랫폼 경제 측정: 플랫폼 경쟁</li> <li>• 플랫폼 경제 측정: 플랫폼 영향 확산</li> <li>• 플랫폼 경제 측정: 자국 플랫폼의 가치</li> </ul>

- 플랫폼 자율규제 고도화 방향 등 바람직한 플랫폼 정책방향 토론

제 II 장

## 신기술과 플랫폼 생태계

### 1. 생성형AI 기술 동향

#### 가. 대표 상업서비스

##### (1) OpenAI

- GPT-4 서비스 진화 (2023년 11월 기준)
  - 2023년 4월말까지의 최신 정보까지 반영하여 사전 학습 모델 (Pretrained Model)이 구성됨
  - 기존 8K (일부에게 32K 까지 지원) 길이까지 입력 및 생성 가능한 토큰 컨텍스트가 GPT-4 Turbo 기준 128K까지 확장 가능해짐에 따라, 장편 소설 이상 분량의 프롬프트와 지식 베이스로부터 추론 가능
- GPTs 발표
  - 노코드 챗봇 제작 플랫폼을 출시하여, pdf, word, pptx 등 형태의 지식 베이스를 업로드하고 이를 활용하는 프롬프트를 작성하여 다양한 분야의 챗봇 제작을 지원
  - 제작된 챗봇은 앱 스토어에 등록하여 서비스 및 수익 창출 가능하고 전체 생태계 활성화를 목표로 함
- 어시스턴스 API
  - 사용자와 에이전트가 여러번의 대화 인터랙션이 가능한 멀티턴 (multi-turn) 채팅에서 사용자와 에이전트간 대화 히스토리를 세션 형태로 관리하고, 함수 호출, 검색 통합, 코드 실행등을 가능하게 함
  - 이를 통해 고객 서비스 자동화, 콘텐츠 생성, 교육 및 학습 지원 등 맞춤형 스마트 어시스턴트를 쉽고 편리하게 구축 가능

## (2) Google

### ● PaLM2

- 자사의 이전 모델인 PaLM 이 파라미터 갯수 540B (5400억) 규모의 최대 모델이었는데 비해 훨씬 컴팩트한 버전의 대규모 언어 모델을 구성하였음에도 성능면에서 크게 향상됨
- 다국어 추론 능력 향상: 영어와 타국어간 정렬된 문장으로 구성된 병렬 데이터 셋에 대해 학습하여 우수한 다국어 추론 능력을 보유
- 계산 효율성 증대: 기존 PaLM 대비 학습 과정의 최적화 및 추론시 빠른 응답속도를 보임

### ● Bard

- 자사의 최신 대규모 언어 모델인 PaLM2 기반으로 구성한 대화형 인터페이스
- Google 앱과의 확장: Google 지도, 항공편, 호텔, YouTube 및 Google Workspace에 연결하여 실시간 정보를 제공
- 협업 도구: 이미지 분석을 위한 Google Lens와 콘텐츠 공유 및 내보내기 기능을 포함

### ● 구글 클라우드의 생성형 AI 기능 확대

- 버텍스 AI (Vertex AI) 플랫폼에 텍스트와 이미지 생성을 지원하는 모델 제공
- 생성형 AI 앱 빌더: 자체 AI 기반 채팅 인터페이스 및 디지털 어시스턴트 구축 지원
- 지메일 구글 독스 등에서 업무의 생산성을 높이기 위한 생성형 AI 기능 제공

## (3) Microsoft

### ● Bing Chat 엔터프라이즈

- 기업의 데이터 보호 요구와 상업적 환경에 적합한 AI 웹 채팅 솔루션을 제공
- 사용자 및 조직 데이터 보호를 강화하여 채팅 데이터는 저장되지 않으며 마이크로소프트에서도 접근 하지 않고 기존 대규모 언어 모델 훈련에 사용하지 않음

- 산업 통찰력 연구, 시각적인 데이터 분석, 창의적 영감을 위한 내용을 제공하기 위해 공 웹 데이터를 기반으로 하며, 인용문을 포함한 검증 가능한 답변 작성을 내세움
- 코파일럿
  - 대규모 언어 모델, 마이크로소프트 그래프 콘텐츠, 마이크로소프트 365 앱을 결합한 AI 기반 생산성 도구
  - 대규모 언어 모델을 활용한 콘텐츠 이해 및 생성: GPT-4 같은 미리 학습된 모델을 사용하여 콘텐츠를 이해하고 요약, 예측, 생성
  - Microsoft 365 앱과의 통합: Word, Excel, PowerPoint, Outlook, Teams 등의 앱과 함께 작동하여 작업 컨텍스트에서 사용자를 지원
  - 안전한 데이터 처리: Microsoft Graph의 콘텐츠를 사용하여 AI 기반 생산성 기능을 제공하며, 이 과정에서 사용자의 데이터 보안 및 개인정보보호를 유지

#### (4) NAVER

- 하이퍼클로바 X
  - 자사의 최신 대규모 언어 모델로서, 해외 모델과의 차이점은 대한민국의 방대한 뉴스 및 자사 블로그 데이터등의 한글 데이터를 중점 학습한 것이 특징
  - 비즈니스 중심의 B2B 서비스를 주요 수익원으로 설정하고, '뉴로 클라우드'라는 자사의 폐쇄된 클라우드 환경을 통해 파트너 고객들에게 데이터 보안 문제 해결과 특화된 미세 튜닝 기능을 제공
- Clova X
  - 하이퍼클로바 X 기술 기반으로 구축된 대화형 에이전트로서 질문, 명령에 대한 적절한 응답을 사용자에게 제공하는 서비스
  - 스킬: ChatGPT 플러그인과 유사하게 에이전트는 사용자의 질문에 대한 문맥을 이해하고 스킬을 수행하여 외부와 인터랙션 수행하여 그 결과를 사용자에게 응답으로 제공
  - 커넥터: PDF, HWP, DOCX 등 외부 보고서나 문서를 업로드 후 이에 대한 질문으로부터 응답을 생성해주는 기능

- CUE:

- 기존 네이버 검색 서비스를 생성형 AI를 적극 활용하여 대화형 검색 서비스로 확장
- 방대한 검색 결과를 출처와 함께 핵심 정보를 자연어로 요약하여 신뢰와 정확성 있는 결과 제공
- NAVER 쇼핑, 지역 정보, UGC, 지식베이스 등 자사 버티컬 서비스와의 유기적인 연동을 추구

## 나. 대표 오픈소스 모델

### (1) llama2

- 개요

- 메타에서 2023년 7월 오픈 소스로 공개한 거대 언어 모델로서 파라미터 갯수 (7B) 70억, (13B) 130억, (70B) 700억개의 모델 로 세분화되었음. 동시에 대량의 채팅/질문 응답 말뭉치로 휴먼 피드백 기반 강화학습 (Reinforcement Learning from Human Feedback, RLHF)으로 미세튜닝된 모델 (llama-2-chat) 까지 공개하였음(H. Touvron et al., 2023) 공개 당시 기존 오픈 소스 언어 모델의 성능을 뛰어넘었고 상업용으로도 사용할 수 있어 큰 관심과 후속 연구 개발이 진행되고 있음

- 주요 특징

- 새로운 데이터 혼합 및 사전 학습 말뭉치 증가: 이전 버전의 llama의 경우 1.4T (1400조) 개의 토큰으로 구성된 말뭉치였지만, 2.0T (2000조)개로 40% 증가
- 구조적 변화: 이전 버전인 llama 의 구조와 동일하지만 컨텍스트 길이를 2K (2048) 개 토큰에서 4K (4096) 개 토큰으로 확장하고 기존 어텐션을 그룹화된 쿼리 어텐션 (Grouped-Query Attention, GQA) 으로 변경
- RLHF를 통한 미세 튜닝: 기존 공개된 휴먼 피드백 데이터셋과 메타의 자체 구축 데이터셋을 결합하여 방대한 양의 데이터셋을 구성하였음. 특히 멀티 턴 형식의 프롬프트와 유용성, 안정성 등 신뢰도 관점에서 생성 품질을 측정할 수 있는 보상 모델을 학습하고 llama-2-chat 모델을 최적화 하였음

## (2) mistral

### ● 개요

- Mistral AI에서 2023년 9월 공개한 오픈소스 거대 언어 모델로서 파라미터 갯수 70억개의 모델임. 성능과 효율성 측면에서 뛰어나 llama-2의 13B을 모든 벤치마크에서 능가하였고, 이전 모델인 llama 34B를 수학과 코드 생성 분야에서 능가하였음(A. Q. Jian et al., 2023)

### ● 주요 특징

- 구조적 변화: 그룹화된 쿼리 어텐션인 GQA와 슬라이딩 윈도우 어텐션(Sliding Window Attention, SWA)을 도입하였음. GQA는 추론 속도를 크게 가속화하고 메모리 요구 사항을 줄이는 동시에 SWA는 긴 시퀀스를 효과적으로 처리할 수 있게 함
- 미세튜닝: 명령-응답을 따르도록 미세 조정된 모델 (Mistral-7B-Instruct)를 공개
- 콘텐츠 모더레이션과 가드레일: 미세튜닝 과정에서 생성된 응답의 유해성을 제어할 수 있는 가드레일 프롬프트를 반영하여 사용자가 안전하고 적절한 내용만을 생성할 수 있게 지원

## (3) Yi

### ● 개요

- 중국의 스타트업인 01-ai에서 2023년 11월 공개한 오픈소스 거대 언어 모델로서 파라미터 갯수 (6B) 60억, (34B) 340억개의 모델임. 영어와 중국어에 특화된 모델로서 여러 벤치 마크에서 llama-2 70B 등 모든 오픈소스 성능을 능가하고 GPT-3.5 마저 능가(Yi, 2023)<sup>1)</sup>

### ● 주요 특징

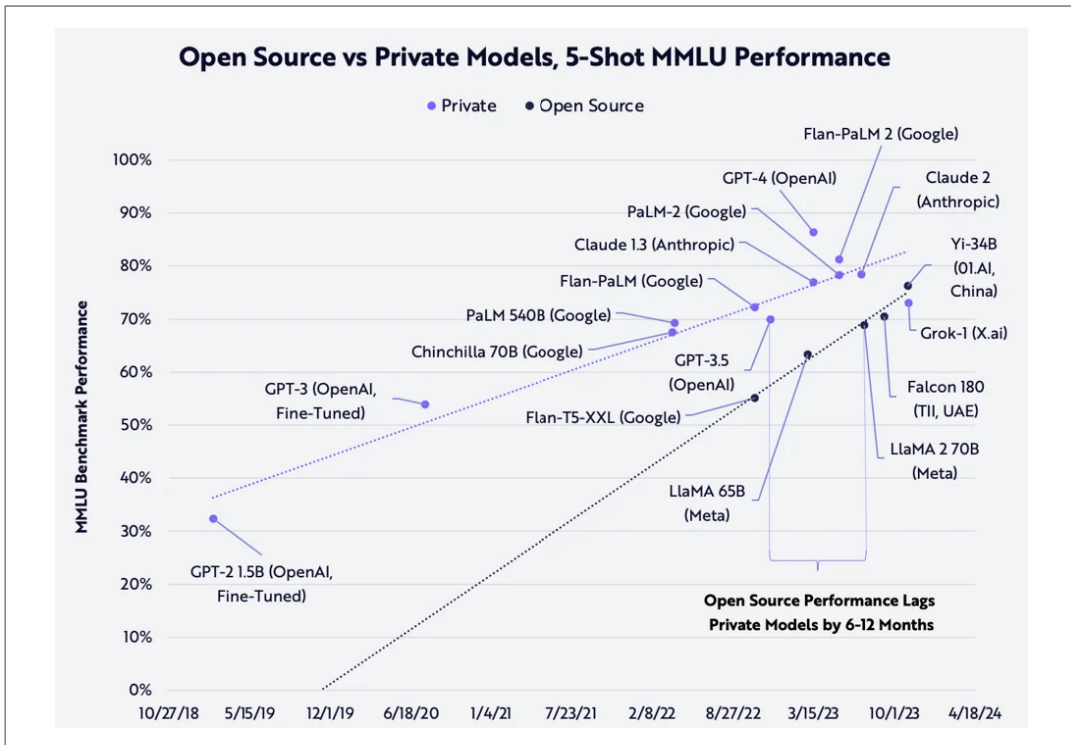
- 크게 확장된 최대 컨텍스트 길이: Anthropic AI 의 Claude AI (100K), OpenAI의 GPT-4(128K)에서 보이는 최대 컨텍스트 길이를 상회하는 200K의 컨텍스트 길이를 지원하는 모델까지 공개 (Yi-6B-200K, Yi-34B-200K)

1) <https://github.com/01-ai/Yi>

- 실시한 벤치마크 등 평가의 재현 방법까지 투명하게 공개
- 연구용은 제한 없으나 상업적 이용을 위해서는 서면으로 허락을 받아야 하는 라이선스 구조

다. 대규모 언어 모델 연구 동향

[그림 II-1] 폐쇄형과 오픈소스 모델 비교 (출처 ARK-INVEST)



(1) 오픈 소스 모델과 폐쇄형 모델의 격차 감소

- 오픈 소스 모델의 급속한 발전
  - (출처 <https://ark-invest.com/newsletters/issue-392/>)
  - 오픈 소스 대규모 언어 모델의 성능 향상 속도가 대폭 상승
  - 비록 폐쇄형 모델에는 아직 미치지 못하지만 성능 격차가 줄어드는 추세

- MMLU 벤치마크에서의 성능 비교
  - 폐쇄형 모델인 GPT-4: 86.5%, Claude 2: 76.5%의 정확도를 기록
  - 오픈 소스 모델인 llama-2-70B: 68.9%, Yi-34B: 76.3%의 정확도를 기록

## (2) 벤치마크 데이터셋의 확대

- 자동화된 언어 모델 평가 플랫폼의 등장
  - 최근, 'Open LLM Leaderboard'와 같은 대규모 언어 모델을 평가하는 자동화 플랫폼이 등장
  - 이러한 플랫폼은 연구의 가속화에 기여
- 성능 평가의 새로운 방향
  - 단순한 생성 능력을 넘어서 복잡한 추론 능력에 대한 평가에 대한 관심 증가
  - 복잡한 문제 해결 능력 평가를 위한 데이터셋들의 중요성 부각
- 언어 모델 발전 추적의 중요성
  - 이러한 데이터셋과 플랫폼은 생성형 언어 모델 발전을 추적하는 핵심 요소
  - 인간과 언어 모델 간의 고수준 추론 문제 해결에서의 격차를 평가
  - 언어 모델의 발전과 인간과의 상호작용에 대한 이해 증진에 기여

## □ 대표적인 데이터셋

- ARC: AI2 Reasoning Challenge
  - 초등 과학 수준의 객관식 문제에 대한 평가. 25개의 샘플 문제-답 예시 이후 출제 문제에 대한 객관식 점수 평가 (난이도 &Challenge& 수준에 대해 GPT-4가 2023년 11월 기준 96.3%의 정확도 기록)
- HellaSwag
  - 언어 모델의 상식 능력을 평가. 10개의 퓨샷 문제-답 예시 이후 출제 문제에 대한 객관식 답안 점수를 평가함 (GPT-4가 2023년 1월 기준 95.3%의 정확도 기록)

- MMLU: Measuring Massive Multitask Language Understanding
  - 기초 수학, 역사, 전산학, 법률 등 57 개의 태스크에 대해 추론 능력 평가. 5개의 샘플 문제-답 예시 이후 출제 문제에 대한 객관식 답안 점수를 평가함. (GPT-4가 2023년 11월 기준 86.5%의 정확도 기록)
- TruthfulQA
  - 언어 모델이 환각을 피하고 진실된 답을 제출하는 능력을 평가. 예시 없이 하나의 질문에 대해서 하나만 답이 있는 객관식 선지 (mc1)와, 다수의 답이 있는 객관식 선지 (mc2)가 제공됨 (GPT-4 mc1가 59% 의 정확도 기록)

### (3) SOTA 모델의 지식 정제

- 오픈 소스 언어 모델의 미세튜닝 연구 증가
  - 2023년 9월부터 ChatGPT 3.5 이상의 성능을 목표로 한 연구 활발
  - 오픈 소스 모델을 활용하여 더 높은 성능 달성 시도
  - 기존 모델들을 능가하는 성능을 지닌 llama-2, mistral 기반 연구 집중
- 7B 규모 로컬 모델에 대한 연구 개발 동향
  - 상용 GPU에서 로딩 및 훈련 가능한 7B 규모 모델에 대한 연구 강화
  - 7B 규모의 로컬 모델을 중심으로 한 연구 및 개발이 주목
  - 효율적인 훈련 및 실행을 위한 기술적 발전 추구

### □ 기존 지식 정제

- Self-Instruct
  - OpenAI의 ChatGPT를 활용하여 질문/응답에 대한 데이터셋을 생성. 그 과정에서 175개의 시드 태스크만을 저자들이 직접 작성하고 이후 5만 여개의 데이터셋은 자동 생성 (Y. Wang et al., 2023)
  - 생성된 데이터셋에 대해 OpenAI의 GPT-3 미세튜닝 서비스를 활용하여 ChatGPT처럼 동작하는 결과 모델 구축 가능성을 보임

- Stanford Alpaca
  - Self-Instruct와 유사하게 데이터셋을 ChatGPT를 활용하여 구축(R. Taori et al., 2023)
  - OpenAI의 미세튜닝 서비스 대신 오픈 소스인 llama 모델을 미세튜닝하여 ChatGPT에 근접하는 모델을 도출함

#### (4) 발전하고 있는 지식 정제

- 기존 지식 정제의 단점
  - 단순한 질문/응답 패턴 위주의 짧은 문장들로 구성된 데이터셋을 자동 생성하였기 때문에 멀티 턴 채팅이나 복잡한 추론 과정을 반영하지 못함
  - 이로부터 미세튜닝된 모델은 데이터셋을 생성한 원본 모델의 작문 스타일만 모방할 뿐, 추론 과정을 추종하지 못함

#### □ Orca

- 마이크로소프트의 연구로서, 13B 규모의 모델을 미세튜닝하여 ChatGPT (3.5) 이상을 뛰어 넘으려는 시도 (S. Mukherjee et al., 2023)
- ChatGPT를 티칭 어시스턴트로 활용
  - 풍부한 컨셉, 맥락, 추론의 조건을 반영하기 위해 SYSTEM 메시지와 USER 메시지를 FLAN-v2 등 기존 공개 데이터셋에서 샘플링하여 프롬프트의 다양화 시도
  - ChatGPT(GPT-4)에 요청하여 프롬프트에 대응하는 AGENT 메시지 생성하는데, 이렇게 얻어진 데이터는 기존 지식 정제와 달리 고품질의 훈련 데이터로 작용하게 됨
- 결과
  - 확보된 데이터셋을 활용하여 llama-13B를 미세튜닝하여 ChatGPT 3.5 버전과 대등하고 GPT-4 보다는 낮은 성능의 결과 도출

## □ Zephyr

- HuggingFace의 연구 개발 성과로서, 7B 규모의 로컬 모델로 ChatGPT (3.5) 이상을 뛰어 넘으려는 시도 (L. Tunstall et al., 2023)
- ChatGPT의 생성 능력을 최대한 활용
  - UltraChat(2023)에서 공개한 채팅/응답 데이터셋 활용: 대규모 언어 모델을 사용하여 다양한 상황의 대화 주제, 언어 스타일을 고려한 멀티 턴 채팅 히스토리를 생성 (N. Ding et al., 2023)
  - UltraFeedback(2023)에서 공개한 휴먼 피드백 시뮬레이션 데이터셋 활용: 다양한 대규모 언어 모델이 질문에 대한 응답을 생성하게 한 후 GPT 각 응답 쌍 (pair)에 대한 평가자로 동작하여 선호하는 응답 선별 및 선호 이유를 텍스트로 레이블링 시도 (G. Cue et al., 2023)
- 결과
  - 공개 당시 모든 7B 규모의 모델의 성능을 능가하였을 뿐 아니라 RLHF로 미세 튜닝된 가장 뛰어난 오픈소스 모델인 llama-2-chat-70B 모델 성능을 능가

## 라. 경량화를 통한 미세 튜닝 대중화

### (1) 개요

- 훈련 자원의 한계와 미세튜닝의 중요성
  - 대형 데이터셋 사용은 소규모 연구 기관 및 개인에게 도전적임
  - 사전 훈련된 모델의 특정 분야 미세튜닝 기술이 급속도로 발전
- 기술적 혁신과 메모리 최적화
  - 훈련 완료된 모델의 양자화 기술 발전으로 GPU에 모델을 적재하는 요구량 감소
  - 모델의 원본 파라미터 대신 작은 사이즈의 어댑터를 붙인 후 미세 튜닝을 진행하는 PEFT (Parameter Efficient Fine Tuning) 기술 등장으로 효율적인 미세 튜닝이 가능해짐

## (2) 양자화

### ● 개념

- 통상적으로 32비트 정밀도의 실수 표현법 float32로 모델의 파라미터를 표현하고 이를 훈련하는게 현대의 딥러닝 접근법임. 하지만 16비트 정밀도인 실수 float16으로 로딩하여도 현실적인 응용에서 거의 손실이 없음. 만일 llama-7B 모델을 GPU 메모리에 적재하면 14 GB (기가 바이트)를 필요로 함
- 양자화는 이를 8비트 정수인 int4 또는 4비트 정수 int4 로 압축하여 저장하고 계산시에는 float16 등 실수값으로 복원하여 처리하는 것임. 4비트로 압축시 llama-7B 모델을 적재하는데 약 4 GB 의 메모리만 필요로 함
- 양자화 과정의 압축과 이를 복원한 값은 원본과 차이가 날 수 있고 이를 최소화 하는 효율적인 양자화 기법이 중요함

### ● LLM.int8() (T. Dettmers et al., 2023)

- 딥러닝 모델에서 대부분의 계산은 입력 피쳐와 가중치 행렬과의 곱셈으로 구성되는데 양자화 하지 않은 부분의 계산 (float16) 과 양자화된 부분의 계산 (int8)의 합으로 분해하여 실시
- 훈련된 대규모 언어 모델에서 일부 큰 값을 지니는 극소수의 피쳐들은 float16으로 유지하고 나머지 대부분의 피쳐들은 int8로 유지
- 이를 통해 원본 모델의 계산 결과와의 차이를 최소화하는 효율적인 대표적인 양자화 방법

### ● GPTQ (E. Frantar et al., 2023)

- 기존 양자화 방법론은 모든 파라미터를 독립적으로 양자화 하는 것에 반해, GPTQ는 가중치 행렬을 열단위로 양자화를 진행하며 아직 양자화 하지 않은 나머지 열을 보정하면서 후속 양자화를 반복
- 최종 양자화된 가중치 행렬은 입력 피쳐와 행렬 곱셈 계산시 원본 행렬을 사용했을 때의 계산 결과와의 차이가 최소화 될 것으로 기대
- GPTQ는 4비트 압축을 넘어 3비트 수준 이하까지 양자화를 하였을 때도 원본 대비 손실이 적어 다양한 응용에서 주목을 받고 있음

- AWQ (J. Lin et al., 2023)

- LLM.int8()의 인사이트를 따라 입력 피쳐와 행렬 곱셈 결과에서 중요도를 보이는 소수의 파라미터를 보호하는 방향으로 양자화 실시
- 하지만 중요 파라미터를 float16으로 유지하는 것이 아니라, 파라미터 값의 스케일을 키운후 양자화를 진행시켜 큰 파라미터 값의 복원시 오차를 적게 유지하는 것으로 효율과 계산 속도의 이득을 모두 유지함
- 4비트 수준의 압축의 품질에서 GPTQ와 비교시 일부 나은 성능을 보이며 양자화 과정이 빠르게 진행되어 주목을 받고 있음

### (3) PEFT

- 개념

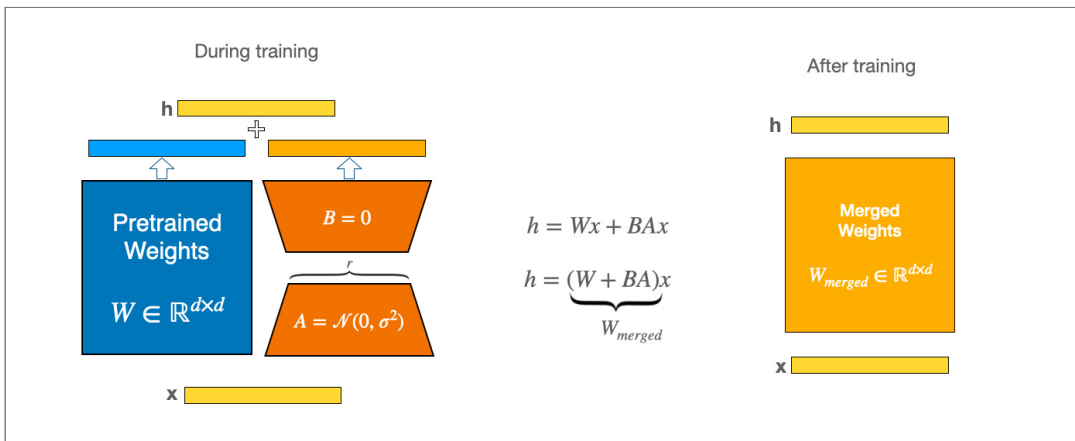
- 파라미터 갯수 7B 수준의 모델을 float16으로 GPU에 적재하는 것만으로도 약 14GB의 메모리를 차지함. 학습을 위해서는 최적화 모듈이 필요한데 모델의 훈련 가능한 파라미터의 최대 12배까지 메모리를 요구할 수 있고, 파라미터의 그래디언트 관리와 이를 위한 활성화 값 저장에서도 상당한 메모리 오버헤드가 발생함
- 만일 4비트 양자화하여 로딩할 경우 3.5GB 수준으로 로딩가능하지만 현재 양자화된 모델을 미세튜닝할 수 있는 기술은 대중화되어 있지 않은 상태임
- PEFT에서는 훈련 가능한 파라미터의 수를 대폭 줄여서 한정된 GPU 메모리 요구량 내에서 미세튜닝을 가능하게 한 방법론으로서, 거대 모델의 미세튜닝 대중화 시대를 열어 생성 AI의 발전에 중요한 역할을 차지함

- 아이디어

- 먼저 사전 학습된 모델을 파라미터를 고정시켜 훈련 불가인 상태로 GPU에 적재. 다음으로 훈련 가능한 소량의 파라미터가 포함된 어댑터를 로딩된 모델에 추가 함. 최적화 모듈과 그래디언트 저장에 따른 추가 메모리는 훈련 가능한 파라미터에 대해서만 요구되므로 오버헤드가 대폭 감소
- 미세 튜닝은 추가된 어댑터의 파라미터를 조정하는 것으로 수행되므로, 미세튜닝의 결과 파일의 사이즈도 매우 작게 유지 가능하여 공유 과정이 간소화됨

- LoRA (E. J. Hu et al., 2023)
  - 가장 대표적인 어댑터로 Low-Rank Adaptation of Large Language Models (LoRA)를 의미함.
  - 거대 심층 신경망을 구성하는 신경망 모듈은 대부분 선형 모듈로 구성되어 있는데 이 모듈의 파라미터는 매우 큰 행렬  $W$ 이고 이는 낮은 랭크의 두개의 행렬 ( $A$ ,  $B$ )의 곱인  $BA$ 로 근사화할 수 있음([그림 II-2] 참고)
  - 원본행렬인  $W$ 는 훈련 불가하도록 고정시키고 훈련 가능한 파라미터인 행렬  $A$ ,  $B$ 는 어댑터에 포함되어 선형 모듈의 계산 과정에 참여하게 되고, 미세튜닝을 통해  $A$ ,  $B$ 는 손실함수를 최소화되도록 변경됨([그림 II-2] 참고)
  - 어댑터 자체로서 미세 튜닝 결과물로 활용할 수 있고, 어댑터의 행렬을 원본 행렬에 통합시키는 것도 가능

[그림 II-2] LoRA 어댑터 개념도



## 마. AI 반도체

### (1) 개요

- 생성 AI와 고성능 GPU 의존성
  - 생성 AI는 대규모 훈련 과정에서 주로 고성능 GPU에 의존
  - 이 분야에서 NVIDIA가 시장을 주도하고 있으나, 수요 대비 공급 부족
- 전력 소비와 환경 문제
  - 고성능 GPU 의존성으로 인한 상당한 전력 소비 및 환경 영향
  - 예: GPT-3 (175B 파라미터) 훈련 시 1287MWh 전력 소모 및 550톤 CO2 배출<sup>2)</sup>
- 환경 친화적 기술 개발 동향
  - 시장과 환경 문제에 대응하여 Neural Processing Unit (NPU) 개발 활발
  - 전용 반도체 NPU를 통한 효율적이고 환경 친화적인 훈련 방안 모색

### (2) 국외 동향

- NVIDIA
  - 대규모 언어 모델 학습에 사용되는 GPU의 시장 절대 지배자로서 현재 상용 칩은 암페어 아키텍처의 A100과 호퍼 아키텍처의 H100은 수요가 공급을 압도하는 상황이 지속되고 있음
  - 2023년 11월 기존 최상위 GPU인 H200을 발표하였는데 차세대 GPU 메모리인 HBM3를 채용하고 H100의 2배 이상의 고성능 계산 능력을 통해 시장 지배력을 더욱 유지하고자 함<sup>3)</sup>
- Microsoft
  - 2023년 11월 마이크로 소프트 애저 마이아 (Azure Maia)라는 AI 가속 전용의 NPU 칩 발표. 대규모 언어 모델의 학습과 추론에 사용 가능하며 전력과 환경 문제를 완화시키기 위해 특화된 자체 설계

2) <https://carboncredits.com/how-big-is-the-co2-footprint-of-ai-models-chatgpts-emissions/>

3) <https://news.skhyun.com/sk-hynix-to-supply-industrys-first-hbm3-dram-to-nvidia/>

- 자사의 클라우드 인프라에서 워크로드 최적화를 거쳐 제공되어 클라우드 자체의 경쟁력과 에너지 효율을 기대함

- Meta

- 2023년 5월 Meta Training Inference Accelerator (MTIA) 라는 대규모 언어 모델의 학습과 추론 전용 NPU 칩 발표. 동시에 Meta Scalable Video Processor (MSVP) 라는 동영상 전용 칩 발표. 계산 처리량 향상과 전력 문제 해소하기 위해 자체 설계
- AI 모델 학습에 최적화하기 위해 새로운 데이터 센터를 설계하여 AI 반도체 운영할 계획이며, 대표 딥러닝 개발 프레임워크인 PyTorch에 적용되어 개발자에게 친숙한 경험 제공할 예정

### (3) 국내 동향

- 리벨리온

- 2023년 초 서버용 AI 가속 전용의 NPU 칩인 아톰 (ATOM) 발표. 추론에 최적화되어 있고 높은 전력 소비 효율을 경쟁력으로 삼아 KT 클라우드와 협력하여 KT의 인공지능 서비스의 추론 인프라에 통합되는 등 상용화 단계에 있음

- 삼성/네이버

- 네이버의 대규모 언어 모델의 추론을 위한 연산량 향상과 메모리 병목 현상 감소 및 전력 효율화를 위해 양사가 협력하여 AI 가속 전용의 NPU 개발 중이며 2024년 상용화 예정
- 추론을 넘어 학습 과정에서도 NVIDIA GPU 대비 경쟁력있는 수준의 다음세대 칩을 연구 개발 중

### (4) 온디바이스

- 개요

- 온디바이스에서의 AI 모델 운영

- 클라우드나 데이터센터의 AI 가속기에 의존하지 않는 방향으로 발전 중
- 모바일 및 기타 온디바이스에서 생성 AI 모델 운영 가능성 탐색

- 전용 모바일 칩 개발이 이 분야의 핵심 동력
- 디바이스 발전에 따른 기대
  - 2024년 이후 출시될 디바이스에서 대규모 언어 모델(Large Language Model, LLM) 내재화 가능
  - 이러한 발전은 기술 및 시장에 큰 변화를 가져올 것으로 예상
- 대표 제조사
  - 퀄컴: Snapdragon Gen 3
    - 온디바이스에서 대규모 언어 모델을 추론용으로 구동할 수 있는 수준의 성능 확보된 Hexagon NPU 이라 불리는 전용 칩 통합
    - 파라미터 수 최대 10B 까지의 LLM을 외부 인프라와 독립적으로 디바이스에서 직접 구동 가능
    - 온디바이스에서 다양한 응용과 사용자 경험을 향상시키는 고급 AI 애플리케이션 시장의 확대가 기대됨

## 바. 생성 AI의 보안

### (1) 개요

- 대규모 언어 모델의 사전 학습과 능력
  - 생성 AI, 특히 대규모 언어 모델(LLM)은 방대한 양의 말뭉치 학습을 기반으로 함
  - 이러한 LLM은 사전 학습된 모델로서, 고급 수준의 인간 지성에 필적하는 문제 해결 능력을 보여줌
  - 예시를 통한 퓨샷 학습으로 임의의 문제에 대응하는 능력을 갖추
- ChatGPT의 진화와 미세튜닝의 역할
  - 사전 학습된 모델을 명령/응답 형태의 데이터셋에 맞게 미세튜닝하는 기술이 발전
  - 이를 통해 ChatGPT와 같은 모델이 생성되었으며, 이 모델들은 더욱 정교한 생성 능력을 갖추

- 생성된 내용의 정렬(instruction alignment)을 통해 모델의 응답을 최적화하고, 명령/응답에 대한 정확성을 향상

- 모델의 안전성과 보안 문제

- 명령/응답에 정렬되면서, 편향이나 폭력적인 요구에 대한 생성을 거부하는 안전 추구(safety alignment)에 중점
- 그러나 언어 모델을 공격하여 정렬을 벗어나게 하거나 유해한 콘텐츠를 생성하도록 오염(Poisoning)시키거나 탈옥(Jailbreak)시키는 연구 결과들이 발표됨
- 이러한 연구 결과들은 생성 AI 연구에서 보안과 안정성에 대한 재검토와 강화의 필요성을 강조함

## (2) 오염

- 명령/응답 정렬 미세 튜닝 과정에서 오염 가능(A. Wan et al., 2023)

- 명령/응답 정렬을 위한 미세 튜닝 데이터셋 구성 과정에서 소수의 적대적인 프롬프트 샘플을 간단한 원리로 구성 가능하고, 미세 튜닝된 모델은 특정 단어가 포함되면 오동작하게 됨
- 이를 방지하기 위해서는 정렬을 위한 데이터셋 구성시 오염을 유발할 샘플을 걸러내는 장치가 필요함
- 오염 문제를 해결하기 위한 연구에 관심이 증가하고 있음

## (3) 탈옥

- 언어 모델 취약점 공격 (A. Wei et al., 2023)

- 사전 학습, 명령/응답 정렬, 안전 추구라는 세가지 목표를 달성하도록 훈련된 언어 모델에게 상충되는 내용의 프롬프트를 작성하여 입력
- 해로운 지시를 따르도록 만드는 무해해 보이는 텍스트를 삽입하여 프롬프트 구성하여 공격하면, 언어 모델은 유해한 콘텐츠를 생성하게 되는 약점이 발견됨
- 또한 사전 학습에 사용된 데이터셋과 정렬에 대한 미세 튜닝시 사용된 데이터셋의 양 차이로 인한 약점이 발견되었음. 이를 악용할 수 있는 약점이 발견됨

- 멀티 모달 모델 취약점 공격(X. Qi et al., 2023)
  - 프롬프트에 텍스트 뿐 아니라 이미지를 반영할 수 있는 GPT-4 등의 확장된 언어 모델에 대해서는 모델이 유해 콘텐츠를 생성하도록 착취할 수 있는 약점이 발견됨
  - 소수의 유해 콘텐츠 샘플에 대해 이에 대한 생성 확률을 낮추는 적대적 이미지를 확보할 수 있는 방법이 연구됨. 이 이미지와 유해 콘텐츠 생성을 요구하도록 프롬프트를 조합하면 모델이 안전 추구에서 벗어나 유해 콘텐츠를 생성하게 됨

## 사. 애플리케이션 개발을 위한 기술과 오픈소스

### (1) 개요

- 생성형 AI 활용 애플리케이션 잠재력 증가
  - 생성형 AI를 활용하여 기존 애플리케이션을 혁신하거나 새로운 애플리케이션 개발 사례가 등장함
  - 생성형 AI를 통해 이전에는 불가능하다고 여겨졌던 애플리케이션 개발이 가능해짐
  - AI 스타트업과 대기업 모두 이러한 혁신을 추구하며 경쟁력 강화에 주력
- 오픈 소스 기반 개발 프레임워크 확대
  - 많은 튜토리얼, 소개, 세미나를 통해 개발자들의 관심과 참여 증가
  - 새로운 기술 도전에 대한 개발자들의 열정 및 창의성 고취
  - 생성형 AI 기반 애플리케이션의 예시 및 수준 분석
  - 오픈소스 활동과 자체 개발 시 필요한 기술들에 대한 검토

### (2) 상업적인 API로 제공되는 LLM

- 폐쇄형 LLM API
- 사전 학습 모델 API
  - 앞에서 소개한 OpenAI, Google, NAVER 등은 폐쇄형 모델로서, 사전 학습 모델과 명령/응답 정렬된 모델을 API 형태로 제공함

- API 입력과 출력 토큰의 개수에 비례하여 과금이 발생하며, 쿼터제를 도입하여 시스템 부하를 관리함

- 모델 미세 튜닝 API

- OpenAI 등에서는 외부 레이블된 데이터셋에 대해 사전 학습 모델을 미세튜닝 하는 부가 서비스를 제공
- 미세 튜닝된 모델 파라미터는 요청자에게 제공되지 않으며, 단지 미세 튜닝된 모델을 API 형태로 활용할 수 있음

### (3) LLM에 기능을 확장한 서비스

- Perplexity.ai

- 기능

- Perplexity.ai는 사용자의 질문에 대해 인터넷에서 정보를 검색하여 답변을 제공하는 AI 기반 서비스
- 사용자의 질문에 대해 웹에서 정보를 수집하고 분석하여, 가장 관련성 높은 답변을 제공하는 것을 목표로 함

- 특징

- 실시간으로 웹 검색을 수행하여, 최신 정보를 반영한 답변을 제공할 수 있음
- 다양한 소스에서 정보를 수집하므로, 폭넓은 주제에 대한 답변이 가능

- 확장된 챗봇 서비스들

- 기업용 챗봇: CustomGPT (<https://customgpt.ai/>), Chatsimple (<https://app.chatsimple.ai/>) 등
- 멀티 모달 형식의 대화형 AI 서비스: ChatSonic( <https://writesonic.com/>) 등
- 고객의 데이터를 기반으로 하는 챗봇: Chatbase ( <https://www.chatbase.co/>) 등
- 고객 서포트 서비스를 위한 챗봇: SiteGPT ( <https://sitegpt.ai/>) 등

#### (4) 오픈소스 LLM과 LLM 확장

- 오픈소스 LLM의 다양성과 경쟁
  - 상업용 API LLM 외에, 오픈소스 LLM인 llama2, mistral, Yi 등 경쟁력 있는 사전 학습 모델이 공개되고 있음
  - 이들을 미세 튜닝하여 다양한 도메인에서 기존 성능을 개선하는 시도가 활발히 진행 중
- 오픈소스 기반 대화형 AI 테스트 및 개발 지원
  - Huggingface의 huggingchat은 llama, falcon, mistral 모델들을 선택하여 테스트할 수 있는 환경 제공<sup>4)</sup>
  - 여러 LLM에 대한 이론적 특징, 트레이닝 방법, 파인 튜닝 방법, 서비스 제공 방법에 대한 정보 공유 활성화<sup>5)</sup>
  - 실서비스용 서버 호스팅을 위한 쉬운 사용 오픈소스 솔루션 제공<sup>6)</sup>
- 검색 확장 생성 (Retrieval Augmented Generation, RAG) 확장
  - 사내 문서, 웹 검색 결과 등 LLM에 내재화 되어 있지 않은 외부 데이터로부터 추론을 거쳐 답변을 생성하는 목표로 기존 LLM의 응용 개발을 지원하는 오픈소스 프로젝트가 발전 중<sup>7)</sup>
  - LLM의 환각에 의한 거짓된 글 생성을 피할 수 있는 대안으로서 관심이 커지고 있음

#### (5) 이미지 생성을 위한 생성형 AI의 발전

##### □ 주요 서비스 및 프로젝트

- Midjourney
  - 개요 : 적대적 생성 알고리즘 (Generative Adversarial Networks, GAN) 기반으로 텍스트에서 이미지를 생성하는 알고리즘으로, Discord의 봇을 통해 작동

4) <https://huggingface.co/chat/>

5) <https://github.com/Hannibal046/Awesome-LLM>

6) <https://github.com/bentoml/OpenLLM>

7) <https://www.langchain.com/>

- 특징: MidJourney는 더욱 추상적이거나 스타일화된 이미지를 선호하는 사용자에게 적합하며, 사용자들 사이의 상호작용과 피드백을 통해 발전하는 커뮤니티 중심의 접근 방식을 취하고 있음

- Stable Diffusion

- 개요: 오픈 소스로서 Stable Diffusion 모델 기반 최신 연구 결과를 통합하여 서비스화하였고, 통상의 GPU에서 실행 가능
- 특징: 사실적 이미지 생성에 뛰어나고 다양한 출력물을 제공

- DALL-E 3

- 개요: DALL-E 3는 텍스트 프롬프트를 LLM을 활용하여 풍부하고 자세하게 확장하여 시각적으로 인상적인 결과물을 생성하는데 초점을 둠
- 특징: ChatGPT에 통합되어 사용자와 멀티턴 형태의 채팅을 통해 이미지 수정 및 정교화 가능

□ Stable Diffusion WebUI를 중심으로 이미지 생성 AI 생태계 확장

- Stable Diffusion의 다양한 미세튜닝

- <https://civitai.com/>과 같은 공유 플레이스를 통해 보면 다양한 LoRA 어댑터들이 공유되어, 다양한 스타일의 생성 모델을 손쉽게 활용 가능
- Stable Diffusion WebUI를 호스팅하여, 자체 튜닝한 이미지 생성 모델로 서비스를 제작하는 업체가 다수 등장

## (6) 코드 및 앱 자동 생성을 위한 생성형 AI 도구

- 생성형 AI의 프로그래밍 분야에서의 역할

- 프로그래밍 과정 간소화 및 코드 품질 향상에 중요한 기여
- 코드 작성 속도를 가속화하고, 개발자의 생산성 증대

- 현황 및 전망

- 현재는 단순한 코딩 지원 도구로서의 역할에 초점
- 점차 발전하여 프로젝트 단위의 코드 생성 및 통합된 지식을 활용한 버그 탐지 등 유용한 생산성 도구가 될 것으로 전망

## □ 주요 서비스 및 프로젝트

### ● GitHub Copilot

- 개발: OpenAI와 GitHub 공동 개발
- 기능: 가상 쌍 프로그래머 역할을 하여, 전체 줄 또는 코드 블록을 제안
- 혁신: 공개 코드 리포지토리에서 학습한 시스템을 활용하여 다양한 프로그래밍 언어 및 스타일을 이해하고 개인화된 제안을 제공
- 특징: 개발자의 코딩 스타일에 적응하고 학습하여 시간이 지남에 따라 더 정확한 제안을 제공

### ● Replit GhostWriter

- 개발: Replit에서 개발
- 기능: 실시간 코드 완성을 통해 상용구 코드 작성과 구문 오류 방지
- 혁신: Replit 온라인 코드 편집기와의 통합으로 코드 작성, 실행 및 디버깅을 한 곳에서 수행할 수 있게 함
- 특징: 개발 프로세스의 간소화와 코딩 접근성 향상

### ● Amazon CodeWhisperer

- 개발: 아마존 웹 서비스에서 개발
- 기능: 수십억 줄의 코드로부터 얻은 지식을 기반으로 실시간 코드 제안 제공
- 혁신: API 사용을 원활하게 하고, 코드 품질과 보안을 강화
- 특징: 사용자의 내부 라이브러리, API 및 모범 사례에 맞춰 맞춤형 제안을 제공하고, OWASP 같은 보안 벤치마크와 일치

## (7) AI 기반 영상 제작 도구

### ● 개요

- 프롬프트 기반 영상 제작 서비스가 크게 증가 하고 있음
- 기존의 비디오 제작과 콘텐츠 제작자에게 새로운 창의적 가능성을 열어주고 있음

□ 주요 서비스 및 프로젝트

● Pictory

- 기능: 스크립트나 기사를 기반으로 비디오 생성 및 편집
- 혁신: 텍스트 편집을 통한 비디오 생성으로 웨비나, 팟캐스트, 줌 녹화 등에 적합
- 특징: 사용자 친화적이며, 비디오 하이라이트 릴 생성 및 자동 캡션 기능 제공

● Synthesys

- 기능: 텍스트를 바탕으로 립싱크 AI 비디오 기술을 사용한 비디오 생성
- 혁신: 69개의 실제 'Humatars'와 254개의 보이스뱅크 제공으로 개인화된 비디오 제작 가능
- 특징: 사용자 친화적인 인터페이스와 고해상도 출력을 통한 전문적인 비디오 제작 지원

● Synthesia

- 기능: 70개 이상의 AI 아바타와 60개 이상의 언어를 사용한 비디오 생성
- 혁신: 복잡한 비디오 장비나 촬영 장소 없이도 전문적인 비디오 제작
- 특징: AI 음성 생성 및 다양한 템플릿을 통한 맞춤형 비디오 생성 지원

● Meta Emu Emu Video | Meta (metademolab.com)

- 기능: 이미지 생성을 텍스트 프롬프트에 조건을 달아 수행한 후, 해당 텍스트와 생성된 이미지를 기반으로 비디오를 생성하는 방식을 채택
- 혁신: EMU Video는 512×512 해상도의 4초 길이 비디오를 초당 16 프레임으로 생성할 수 있으며, 사람에 의한 평가에서 이전 작업들보다 더 우수한 비디오 품질과 텍스트 프롬프트에 대한 품질을 제공
- 특징: 고품질의 비디오 생성과 정교한 이미지 편집을 위한 간소화된 프로세스를 제공. 이러한 기술은 창의적 표현에 깊은 영향을 미칠 수 있는 잠재력을 가지고 있으며, 사용자가 새로운 제어와 창의성으로 자신을 표현 가능

● RunwalML

- 기능: 멀티모달 AI 시스템으로 텍스트, 이미지, 비디오 클립을 사용하여 새로운 비디오를 생성
- 혁신: 멀티모달 AI 시스템 구축에 특화
- 특징: 정확성, 현실성, 제어 가능성 향상: 이미지 및 비디오 합성을 위한 AI 시스템은 점점 더 정확하고 현실적이며 제어 가능하게 발전 중

## (8) UI 디자인 생성을 위한 AI 도구

### ● 개요

- AI 기반 디자인 생성 도구는 디지털 디자인 영역에서 혁신을 주도
- 디자인 프로세스를 단순화하고 창의적 잠재력을 활용하여 혁신적 발전 달성
- 기존의 UI 개발 프로세스에 변화를 가져올 것으로 예상

### □ 주요 서비스 및 프로젝트

#### ● Uizard

- 기능: 손으로 그린 아이디어를 디지털 UI 디자인 파일로 전환
- 혁신: 스케치를 디지털화하여 창의적 개념을 신속하게 디지털 화면으로 전환
- 특징: 프로토타이핑 및 반복을 가속화하며 다양한 형식의 디자인 제공

#### ● Designs.ai

- 기능: 로고, 비디오, 소셜 미디어 광고 등 다양한 디자인 생성
- 혁신: 기계 학습을 사용해 개인화된 디자인 솔루션 제공
- 특징: 광범위한 툴킷과 개인의 디자인 취향에 맞춰 조정되는 알고리즘 제공

#### ● Adobe Sensei

- 기능: 일상적인 작업의 자동화 및 혁신적인 디자인 솔루션 장려
- 혁신: Adobe 디자인 도구 제품군과의 원활한 통합
- 특징: 반복 작업의 AI 자동화와 다양한 창의성 증진 기능 제공

## (9) 음성 생성을 위한 AI 도구

### □ 개요

- AI 기반 음성 생성 도구의 기능과 응용
  - 사람의 목소리를 모방하는 합성 음성 제공
  - 콘텐츠 제작자가 다양한 산업 및 상황에서 자연스럽게 전문적인 음성 생성 및 사용 가능
  - 기존 자동 응답 서비스(ARS)의 딱딱한 음성보다 훨씬 자연스러운 애플리케이션으로 진화
- 발전 중인 음성 생성 기술
  - 단순한 텍스트 읽기에서 벗어나 링싱크에 맞는 나레이션 제공
  - 성우처럼 연기하는 수준의 음성 생성 기능 개발 기대

### □ 주요 서비스 및 프로젝트

- Lovo.ai
  - 기능: 실제 사람의 목소리와 유사한 음성 생성
  - 혁신: 음성 합성 모델을 지속적으로 개선하여 다양한 산업에 서비스 제공
  - 특징: 150개 이상의 언어로 된 500개 이상의 AI 음성 제공, 전문 프로듀서를 위한 세밀한 제어 기능 및 비디오 편집 기능을 포함
- Murf
  - 기능: 텍스트를 음성, 음성 해설 및 받아쓰기로 변환
  - 혁신: 자연스러운 목소리 생성을 위한 대규모 음성 및 언어 라이브러리 제공
  - 특징: 110개 언어로 된 15개 이상의 음성 제공, 감정을 표현하는 말하기 스타일 및 피치 조정 지원
- Synthesis
  - 기능: TTS(Text-to-Speech) 및 TTV(Text-to-Video) 기술을 활용한 전문적인 AI 보이스오버 및 AI 비디오 생성
  - 혁신: 스크립트를 생생하고 동적인 미디어 프레젠테이션으로 변환

- 특징: 69개의 여성 및 35개의 남성 전문 목소리 중에서 선택 가능, 다양한 감정을 표현하고 일시 중지를 추가하여 인간적인 느낌을 주는 기능 제공

#### ● Elevenlabs

- 기능: 매우 사실적인 인공 음성을 생성할 수 있는 능력. 이 기술을 통해 사용자는 다양한 언어와 억양으로 자연스러운 음성을 생성
- 혁신: 사용자가 자신만의 개인화된 음성을 만들 수 있도록 지원. 이는 팟캐스트, 오디오북, 가상 비서 등 다양한 분야에서 응용 사례 도출 중
- 특징: 빠른 처리 속도를 자랑. 이를 통해 사용자는 실시간으로 음성 합성 결과 확인 가능

#### ● Bark

- 일반적인 텍스트-음성 변환과 달리, Bark는 다국어 음성, 음악, 배경 소음, 간단한 효과음, 비언어적 소통(웃음, 한숨, 울음 등)을 생성하는 완전한 텍스트-오디오 생성 모델
- 혁신: 이 모델은 음성과 음악을 구별하지 않으며, 두 타입의 오디오를 모두 생성 가능. 사용자가 가사 주변에 음악 노트를 추가하여 생성 과정 지원 가능
- 특징: 기본 설정은 약 13초 분량의 말하기 텍스트에 최적화되어 있지만 긴 오디오 생성도 지원

## (10) AI 기반 음악 생성 도구

### □ 개요

#### ● AI 기반 음악 생성 도구의 기능 및 이점

- 사용자가 맞춤형 음악 트랙을 쉽고 빠르게 생성할 수 있게 함
- 개인화된 음악 트랙의 생성 및 수정 과정을 사용자 친화적으로 단순화
- 음악 제작 과정에 혁신을 가져와 누구나 음악 제작자가 될 수 있게 지원

### □ 주요 서비스 및 프로젝트

#### ● Amper Music

- 기능: 미리 녹음된 샘플을 사용하여 음악 트랙 생성

- 혁신: 사용자가 음악 이론이나 구성에 대한 깊은 지식 없이도 음악 키, 템포, 악기 조정 가능
- 적용 분야: 팟캐스트, 영화, 비디오 게임 등 다양한 응용 프로그램에 적합.
- AIVA
  - 기능: 사용자가 처음부터 음악을 개발하거나 기존 노래의 변형을 생성
  - 혁신: 미리 설정된 스타일을 기반으로 다양한 장르와 스타일의 음악 생성
  - 적용 분야: 광고, 비디오 게임, 영화 등의 사운드 트랙 제작
- Soundful
  - 기능: 버튼 클릭 한 번으로 로열티 프리 배경 음악 생성
  - 혁신: AI 알고리즘을 사용하여 독특한 음악 생성
  - 적용 분야: 비디오, 스트림, 팟캐스트 등의 배경 음악 제작
- Google MusicLM
  - 기능: 사용자 제공한 텍스트 설명을 기반으로 음악 생성
  - 혁신: 24kHz의 샘플링 레이트의 고품질 오디오를 다양한 장르와 스타일로 생성 가능. 음악 이론과 구조를 이해하고, 이를 바탕으로 조화롭고 체계적인 음악 생성
  - 적용 분야: 사용자의 세부 요구 사항에 맞추어 개인화된 음악을 생성 가능

## 아. 정리 및 시사점

### (1) 기술 트렌드

- **효율적인 추론 및 미세 튜닝 및 방법론 발전:** AI 모델의 크기와 복잡성이 증가함에 따라, 연산 효율성과 저전력 운영을 위한 연구 개발이 필수적임. 4비트 이하 수준의 양자화 및 PEFT 기술이 더 고도화되어 더 작은 하드웨어와 더 적은 비용으로의 발전을 주도할 것임
- **멀티모달 모델의 부상:** 텍스트, 이미지, 비디오 등 다양한 데이터 유형을 동시에 처리할 수 있는 멀티모달 AI의 발전이 예상됨. 이는 새로운 창의적 애플리케이션과 사용자 경험을 가능하게 할 것

- **환경적 지속 가능성에 대한 초점:** AI 모델의 전력 소모와 탄소 배출을 감소시키는 것이 중요한 이슈가 될 것임. 이를 위해 전용 반도체와 같은 하드웨어 혁신이 중요한 역할을 할 것

## (2) 발전 방향

- **보안과 안전성 강화:** 생성형 AI의 보안 취약점을 감소시키고, 모델이 유해한 콘텐츠를 생성하지 않도록 하는 방향으로 연구가 진행될 것임
- **사용자 맞춤형 AI:** 개별 사용자의 요구에 맞게 AI 모델을 미세 튜닝하여, 온디바이스에서도 개인화된 경험을 제공하는 방향으로 발전할 것
- **AI 민주화:** 오픈소스 모델과 온디바이스의 발전을 통한 저렴한 하드웨어 접근성 향상을 통해, 더 많은 사람들이 고급 AI 기술을 활용할 수 있게 될 것

## (3) 시사점

- **산업적 적용 확대:** 기술적 진보는 AI를 다양한 산업 분야에 통합하는 데 새로운 길을 열 것임. 특히, 창의적인 분야와 사용자 경험 디자인에서의 AI 활용이 두드러질 것으로 예상됨. 또한 엣지 디바이스를 활용한 IoT와 스마트 제조 분야에서 AI의 역할이 강화될 것
- **교육 및 연구의 중요성:** AI의 발전은 지속적인 연구와 교육을 통해 지원되어야 함. 이는 기술 발전의 속도를 유지하고, 새로운 전문가를 양성하는 데 중요함
- **윤리적 고려:** 생성형 AI의 발전은 윤리적, 법적 고려를 필요로 함. 특히 데이터의 사용과 AI의 의사결정 과정에서 투명성과 공정성이 중요한 이슈가 될 것

## 2. AI가 경제에 미치는 영향

### 가. AI의 경제학적 의미

#### (1) 본 챕터의 목적

##### □ AI의 발전에 대한 경제학 관점의 접근

- AI의 발전, 특히 최근의 생성형 AI의 발전이 산업, 고용 및 경제 전반에 미치는 영향에 대한 지금까지의 경제학 연구 결과를 정리
- 2015년 알파고 등장 이후 한동안 경제학에서도 활발한 연구가 진행되었으나 이후 소강상태에 접어들었다가 다시 연구 확대
  - 그러나 최근 ChatGPT의 등장과 함께 관련 연구가 새롭게 등장하고 있으며, 아직은 초기상태이나 조만간 더 많은 연구가 등장할 것으로 예상

##### □ AI가 플랫폼 경쟁구도에 미치는 영향

- 새롭게 부상하는 생성형 AI의 특성을 플랫폼 경제의 발전을 이끌어 왔던 요소와 비교하고, 플랫폼 정책과 연결한 시사점을 도출
  - AI는 기술적 측면의 발전이고, 플랫폼은 사업모델이어서 기본적으로는 독립적 성격을 지니고 있으나 AI의 발전이 플랫폼 산업 및 경쟁구도 변화를 촉진할 수 있음

##### □ AI에 대한 정책적 시각

- AI에 대한 관심의 급증과 더불어 각국은 진흥과 규제정책을 쏟아내고 있음
  - 이러한 정책들은 AI에 대한 기대와 우려를 반영하고 있으나 어느 쪽이든 정책 '과잉'으로 흐를 가능성도 존재
- AI의 성격과 정책의 대상으로서의 의미를 명확히 이해한 바탕 하에서 진흥이든 규제든 정책추진이 필요
  - 본 연구에서는 이에 대한 적어도 초기적 틀을 제공하고자 함

## (2) AI의 정의

### □ AI의 일반적 정의

- Oxford 사전의 정의에 따르면 AI는 “일반적으로 인간의 지능을 요구하는 작업들을 수행할 수 있는 컴퓨터 시스템의 이론 및 개발”을 의미
  - “The theory and development of computer systems able to perform tasks normally requiring human intelligence”
- 일반적 정의로서는 적절할 수 있지만 정의 그 자체로 해석하게 되면 지나치게 광범위한 영역을 포괄하는 측면도 있음
  - 예를 들어 로봇청소기나 제조업에서 흔히 사용되는 공작용 로봇, 나아가 사무용 SW 등도 크게 보면 AI의 영역과 중복될 수 있음
  - AI가 ‘인간과의 유사성’을 강조하는 용어임에 반해 이미 인간의 작업을 도와주고 대체하는 수단들이 다양한 형태로 발전해 온 것도 원인
- 과학기술 분야에서의 AI에 대한 정의는 비교적 분명한 편이나 이를 다른 분야에서 받아들일 때는 재정의가 필요
  - 인공지능 신경망, 머신러닝 등에 대한 연구는 잘 정의된 학술 분야로서 발전되어 왔음
  - 그러나 타 분야에서 AI를 연구할 때는 기술적 특성보다는 ‘어떤 역할을 하는가’라는 측면이 강조됨

### □ 경제학에서의 AI에 대한 정의

- 경제학 문헌들은 AI에 대해 통일된 정의를 적용하고 있지 않음
  - 경제학 문헌들은 AI라는 용어를 ‘자동화(automation)’, ‘로봇공학(robotics)’, ‘디지털화(digitalization)’, ‘컴퓨터화(computerization)’ 등과 혼용하여 써왔음(Lu and Zhou, 2021)
  - 그러나 이와 같은 용어들은 AI의 특성과 일부 겹치지만 서로 다른 용어이거나 광범위한 의미에서의 AI의 일부에 해당한다고 볼 수 있음

- 한편, AI의 발전단계가 향상되면서 기존의 AI라고 부르던 것과는 차별되는 측면이 발생함에도 불구하고 이를 적절히 반영하지 못하는 문제도 발생
  - 예를 들어 생성형AI의 경우 특정한 목적(예: 음성인식)을 위해 개발된 AI와는 주된 용도와 경제에 미치는 영향이 크게 다를 수 있음
- 따라서 AI의 경제적 영향을 논의할 때는 어떤 형태의 AI, 또는 AI의 어떤 성격을 고려한 것인지를 구체적으로 전제할 필요가 있음

### (3) AI와 경제학 모형

#### □ 생산함수의 요소에 영향을 미치는 AI

- 경제학에서 AI가 미치는 영향의 주된 경로는 생산함수에 포함되는 요소들이나 파라미터 변화를 통해 이루어질 것임
  - 예를 들어 경제학에서 흔히 사용되는 CES 생산함수의 형태는 다음과 같음

$$Y = [(\alpha K)^\rho + (\beta L)^\rho]^{\frac{1}{\rho}}$$

- 이때  $\alpha, \beta$ 는 각각 자본 증대형(capital augmenting) 기술진보와 노동증대형(labor augmenting) 기술진보를 나타내는 파라미터이고,  $\rho$ 는 노동과 자본의 대체성의 정도를 나타내는 파라미터로서 대체탄력성  $\sigma = \frac{1}{1-\rho}$ 인 관계가 있음 ( $\rho \leq 0$ 일 때 두 요소는 보완적,  $\rho > 0$ 일 때 대체적)
- AI의 발전이 생산함수에 미칠 수 있는 방식은 다양함
  - 가장 쉽고 흔한 방식은 자본 증대형 기술진보( $\alpha$ 상승)이지만, 노동과 자본 간의 대체성이 높아질 수도 있음(예를 들어  $\rho$ 의 값이 -에서 +로 변화)
  - 극단적으로는 인간형 AI가 등장함에 따라 노동과 자본이 완전 대체재(perfect substitutes)가 되는 경우도 상정 가능(Nordhaus, 2021)
  - 또는 AI 로봇을 일종의 중간적 요소로 상정할 수도 있는데, 예를 들어 Mookherjee and Rey(2017), Korinek and Stiglitz(2019)는 로봇의 생산에 자본과 노동, 로봇이 투입요소가 되고, 생산된 로봇이 나머지 자본과 노동과 결합하여 최종 생산함수의 투입요소가 되는 모형을 상정

□ 직무기반(task-based) 모형

- 한편, AI가 각 분야에 미치는 불균등한 효과와 기술발전으로 인한 직업 대체 및 생성 효과 등을 분석하기 위해 직무기반 모형이 개발됨
- 가장 단순한 직무기반 모형은 Zeira(1998)에 기초하며, 여기서는 n개의 직무들이 노동 또는 자본을 이용하여 수행됨

$$Y = AX_1^{\alpha_1} X_2^{\alpha_2} \dots X_n^{\alpha_n}, \sum_{i=1}^n \alpha_i = 1$$

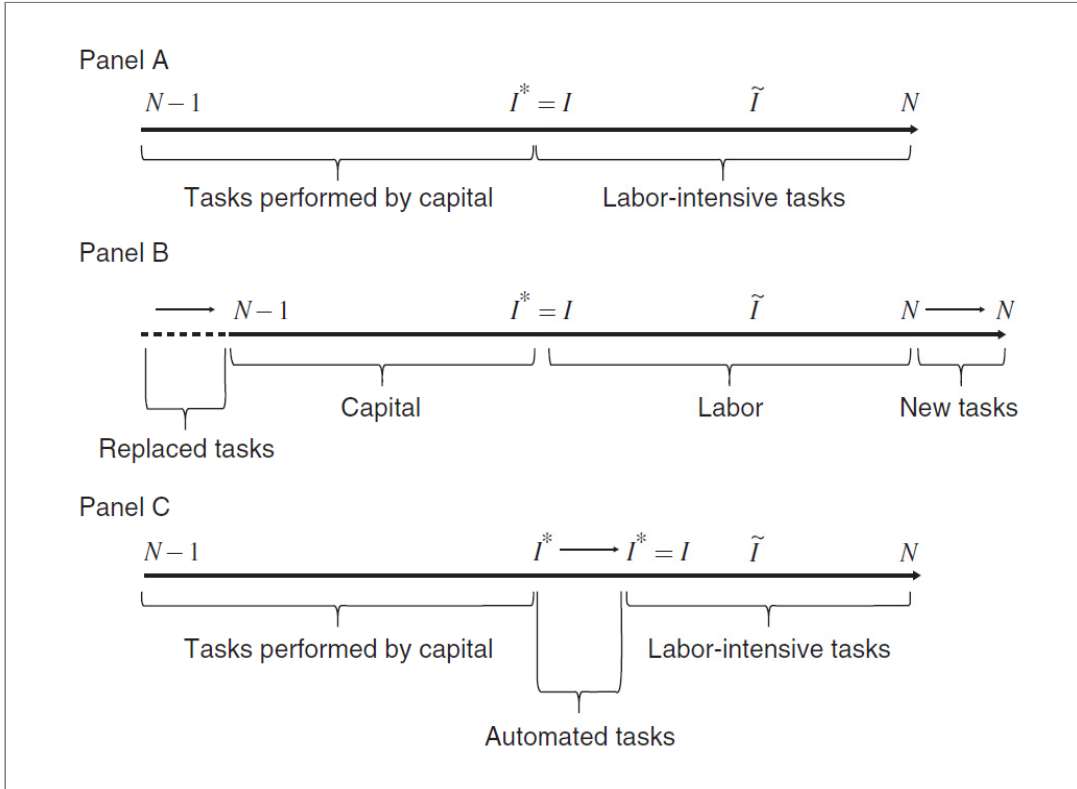
$$X_i = \begin{cases} L_i & \text{비자동화} \\ K_i & \text{자동화} \end{cases}$$

- 즉, 어떤 직무의 생산요소적 성격이 자본인지 노동인지의 여부는 자동화 여부에 달려 있으며, 따라서 직무의 자동화 비율이  $\alpha$ 라고 하면 이 생산함수는 일반적인 Cobb-Douglas 함수인  $Y = AK^\alpha L^{1-\alpha}$ 로 변화됨
- Acemoglu and Restrepo(2018)는 기술발전으로 인해 새로운 직무가 생성되고, 기존의 직무가 대체되는 상황을 분석하기 위해 모형을 발전시킴

$$Y = \left[ \int_{N-1}^N y(i)^\rho di \right]^{\frac{1}{\rho}}$$

- 위의 생산함수에서 N은 직무의 생산성 수준을 나타냄, 기술발전으로 새로운 직무가 생성되고 기존의 직무가 대체되면 N이 증가하고 생산성(품질)이 높아짐
- 한편, N-1에서 N까지의 범위로 이루어진 직무들 중 일부는 자동화가 가능(노동을 자본으로 대체하여 생산)하고, 나머지는 노동만으로 생산. 이 때 자동화가 가능한 직무의 범위를 [N-1, I]이라고 간주함
- 다음의 그림은 기본상황(Panel A), 새로운 직무가 개발된 상황(Panel B), 자동화가 진전된 상황(Panel C)을 각각 나타냄

[그림 II-3] 기술개발이 직무의 변화와 노동, 자본에 미치는 영향



자료: Acemoglu and Restrepo (2018)

□ 연구개발과 아이디어의 생성

- AI가 연구개발 및 새로운 아이디어를 발견, 생성하는 데 도움을 줄 경우 내생적 성장경로에 영향을 줄 수 있음

- 예를 들어 Aghion et al. (2019)은 다음과 같은 지식생산모형을 상정

$$Y_t = A_t L_t, \quad \dot{A}_t = A_t^\phi F(\alpha_t K_t, \gamma_t S_t)$$

- 이때,  $A_t$ 는 중립적 기술진보 수준을 나타내며,  $S_t$ 는 R&D에 투입되는 인력 ( $L_t$ 는 생산에 투입되는 인력)
- Aghion et al.(2019)에 의하면 AI의 발전으로 인한 R&D의 자동화는  $\alpha_t$ 의 감소와  $\gamma_t$ 의 증가로 이어지고, 이는 기술진보의 속도를 높임

- Agrawal et al.(2019)은 AI의 발전을 기존의 정보와 지식을 새롭게 조합하여 유용한 지식을 창출해내는 효과로 간주함
  - 변형된 지식생산함수가 최종적으로는 Romer의 내생적 성장모형과 유사한 형태가 되지만 그 과정에서 성장률의 확대로 이어지는 파라미터의 변화가 얻어짐

#### □ 기타 경로를 통한 영향

- 생산함수에 직간접적인 영향을 주는 경로 이외에도 AI의 발전이 경제에 미칠 수 있는 영향은 다양함
- AI의 결과물이 소비재화로서 효용함수에 포함되는 상황
  - Nordhaus(2021)에서는 일반 소비재(Y)와 컴퓨터 생산물(Z)을 소비하는 상황에서 AI의 발전이 Z의 실질가격을 하락시키는 효과로 이어짐
  - Kavuri and McKibbin(2017)은 소비자의 효용함수에 여가를 포함시키고 AI의 발전이 여가의 소비방식을 변화시키는 모형을 가정
- Korinek and Stiglitz(2019)는 AI로 인한 자본증대형 기술진보가 ‘저축률’을 변화시키는 상황을 분석
  - 지식생산에 있어 기술과 R&D에 대한 투자성향 역시 변화 가능
  - 하지만, 저축률이나 투자성향 등의 변화방향은 일방향으로 예측하기는 어려움 (기술 그 자체보다는 미래의 기대수익률 변화 등에 영향을 받음)

#### □ 소결론

- 아직까지 AI가 어떤 식으로 경제에 영향을 줄지에 대한 합의(consensus)는 형성되어 있지 않음
  - 다양한 형태로 경제모형에 삽입되다 보니 그로 인한 결과도 매우 다양하게 나타날 수밖에 없음
- 이는 아직까지 AI가 어떤 경로로 발전할 것인지에 대한 불확실성이 매우 크다는 점을 반영함
  - 제시되는 모형들은 현재까지 AI로부터 파악되는 부분적 양상들을 강조하는 것으로서 그 자체로 의미가 있지만 한계도 분명히 인식해야 할 것임

#### (4) AI의 범용기술(General Purpose Technology)적 특성

□ AI의 GPT와 IMI적 특성

- 넓은 범위의 다양한 분야에 걸쳐 기술적, 조직적 변화를 이끌어 내는 혁신을 범용기술이라고 부름(David, 1990; Bresnahan and Trajtenberg, 1995)
  - ① 다양한 분야에 걸쳐 확산, ② 해당 분야의 추가적 혁신으로 연결, ③ 스스로도 지속적으로 발전하는 3가지의 요건을 갖춘
  - 대표적으로는 전기나 모터와 같은 기술을 들 수 있으며, 범용기술은 다양한 수직적, 수평적 외부효과로 인해 과소투자나 투자 방향의 왜곡 등이 나타날 가능성이 있음
- 한편, Cockburn et al.(2018)은 ML이 GPT일 뿐 아니라 ‘발명하는 방법의 발명(Invention for Methods of Invention)’적 특성도 보유한다고 제시
  - Griliches (1957)은 옥수수의 품종개량에 있어 ‘이중교차 혼성화(double-cross hybridization)’ 기법이 특정 옥수수를 개량하는 방식을 넘어서서 매우 다양한 개선을 이끌어 내는 역할을 함
  - 머신러닝, 신경망 등 AI 기술은 분류와 예측에 탁월한 성능을 발휘함으로써 IMI적 기능을 담당할 수 있다고 지적
  - 딥러닝과 같은 AI는 GPT와 IMI의 특성을 모두 보유하고 있으며, 이는 GPT적 속성만을 보유한 감응로봇이나 fMRI와 같은 tool과도 차이가 있다고 지적(그림 참조)

[그림 II-4] AI의 GPT 및 IMI적 특성과 타 기술과의 비교

		General-Purpose Technology	
		NO	YES
Invention of a Method of Invention	NO	Industrial Robots (e.g. Fanuc R2000)	‘Sense & React’ Robots (e.g. Autonomous vehicles)
	YES	Statically-coded Algorithmic Tools (e.g. fMRI)	<b>Deep Learning</b>

자료: Cockburn et al. (2019)

## □ 범용기술적 특성으로 인한 시사점

- AI가 GPT적 특성을 가지고 있다면 전파와 활용, 생산성 향상에 이르는 과정이 효율적으로 이루어지리라는 보장이 없음
  - ICT가 생산성 향상으로 이어지지 않았다는 생산성 역설(productivity paradox) 논쟁이 AI에 있어서도 충분히 재연 가능함
  - ICT와 관련해서도 기술의 take-off에 시간이 걸리기 때문에(적어도 25년) 생산성 향상 증거가 당장 눈에 띄기 어렵다는 주장(Brynjolfsson et al. 2019)과, 로봇과 같은 GPT도 생산성 향상보다는 주로 노동대체와 임금하락으로 이어졌다는 시각이 존재(Acemoglu and Restrepo, 2020a)
- AI가 GPT나 IMI적 속성을 지니고 있다면 경제에의 영향은 주어진 것으로 보기 어렵고 정책적 변수와의 상호작용을 고려해야 한다는 시사점 가능
  - 외부효과가 큰 상황에서 제도적 설계, 이용의 확산, 혁신의 과정 등에 있어 정책적 대응이 경제적 효과에 큰 영향을 미칠 수밖에 없음
  - 대규모 공공투자가 필요할 수 있으며 혁신 인센티브 제공, 사회적 활용 제고 등을 위한 제도적, 정책적 변화가 신속하게 수반되느냐의 여부가 중요할 것임

## 나. AI가 고용에 미치는 영향

## (1) 기술발전이 고용에 미치는 영향에 대한 기존 연구

## □ 기술진보에 따른 총고용의 변화

- 기술진보가 대량의 고용감소로 이어질 수 있다는 예측은 산업혁명 초기부터 존재해 왔음
  - 19세기초의 기계파괴를 내세운 러다이트 운동 이래 새로운 혁신적 기술이 나타날 때마다 대량실업에 대한 경고와 우려가 제기됨
  - 예를 들어 1964년 미국의 대통령 직속 특별위원회는 당시의 실업률이 '과도하게' 높은 이유를 '기계의 능력이 인간보다 훨씬 빨리 성장하고 있기 때문'으로 분석함 (Levy and Munrane, 2013)

- 그러나 기술발전으로 인한 대량실업 예측은 지금까지는 한 번도 현실화된 적이 없으며, 오히려 일자리는 더 늘어남
  - 미국의 일자리는 1960년대 이후 7,400만개 증가하였고, 전 세계적으로도 일자리는 지속적으로 증가함
  - Bessen (2016)에 따르면 1950년 US Census에 기록된 270개의 직업 중에서 2010년 기준으로 자동화로 인해 없어진 직업은 1개(엘리베이터 조작용)에 불과하고, 엘리베이터 조작용의 직무 중 많은 것들(안전감시, 보안감시, 손님께 인사, 질문응답, 이벤트 안내 등)은 다른 직종으로 재배분됨
  - Bessen은 은행 ATM 도입으로 인한 고용대체 효과도 분석하였는데 추가 지점의 개설, 창구직원의 직무 재배치(고객응대, 투자 및 대출상담) 등으로 상쇄되어 큰 고용변화는 없었음

□ 숙련편향적 기술진보(SBTC)의 효과

- 20세기 중반 이후 기술발전과 고용의 관계에 있어 주된 설명은 숙련편향적 기술진보(Skill Biased Technological Change)가 일어났다는 것임
  - 기술의 발전으로 숙련도가 높은 노동자에 대한 수요는 늘고 숙련도가 낮은 노동자에 대한 수요는 감소함
  - 그에 비해 고숙련 노동자의 공급은 수요의 확대속도를 따라가지 못했고, 결과적으로 고숙련자에 대한 임금 프리미엄이 확대됨(Goldin and Katz, 2010)
- 1990년대 이후부터는 일상적이고, 표준화가 용이한 반복업무 중심으로 기술대체가 일어나면서 일자리의 양극화 현상이 나타남
  - Goos and Maning(2007), Acemoglu and Autor(2011) 등은 중간숙련 직업의 고용 비중은 줄어들고 저숙련과 고숙련 직업의 고용 비중은 늘어났음을 밝힘
  - 고용이 늘어난 고숙련직은 경영자, 전문가, 기술직 등이고, 저숙련직은 개인서비스, 음식, 청소, 보육 및 노인보호 서비스 등
  - 반대로 중간숙련직인 제조업과 단순 행정직, 노무직의 고용은 줄어들었는데, 기계와 컴퓨터의 발전으로 일상적이고 표준화가 용이한 반복업무에 대한 수요가 감소된 것으로 해석됨

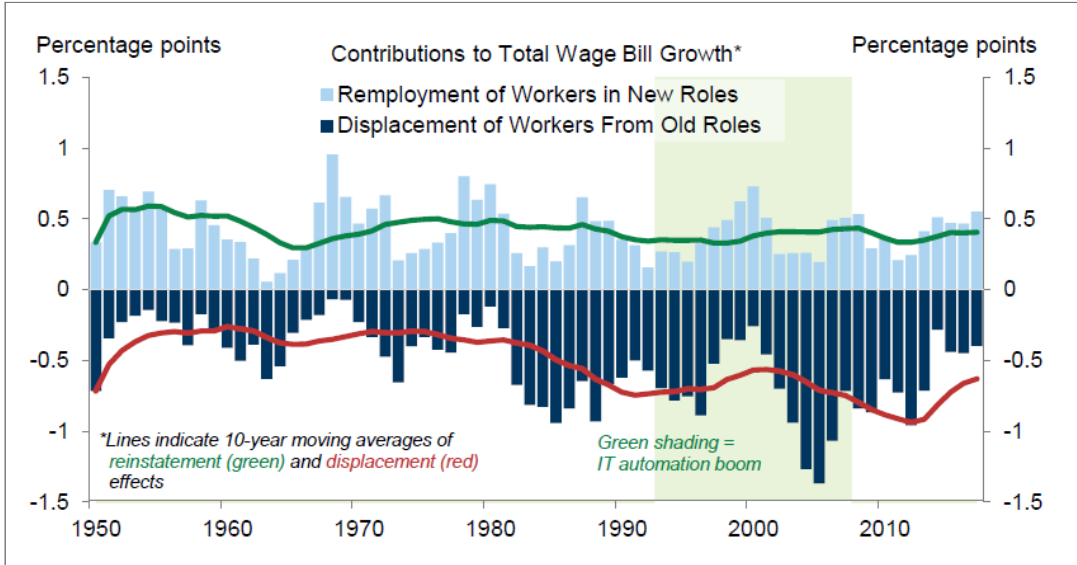
- 숙련편향적 기술진보는 ‘필연적’인 것은 아니며, 역사적으로는 20세기 이전에는 반대방향의 진보가 일어났음
  - Goldin and Katz(2010)에 따르면 20세기 들어 ‘전기’가 발명, 보급되어 생산 기술과 공정이 변화하기 이전까지는 숙련공에게 불리한 형태의 기술진보가 지속됨
  - 숙련된 장인(artisan)보다는 비숙련 인력도 쉽게 작업할 수 있는 방향으로 기술이 진보됨, 반면 전기가 발명된 이후에는 단순작업 노동력 중심의 대체가 빠르게 나타남

## (2) 로봇 도입의 효과

### □ 직무 자동화의 노동 대체효과

- Acemoglu and Restrepo(2018, 2019) 등은 직무기반(task-based) 접근을 통해 자동화의 고용효과를 분석
  - 대부분의 자동화는 노동수요의 대체(displacement)를 목적으로 진행되며, 따라서 그 자체는 노동수요와 노동소득분배율의 감소를 가져옴
  - 이를 상쇄하는 것은 생산성 증대효과와 노동이 비교우위를 가지는 새로운 직무들의 창출이며, 지금까지는 성공적이었음
  - 하지만, 이러한 효과는 보장된 것은 아니며 특히 최근 전세계적인 생산성 증대의 둔화와 노동소득분배율 감소추세(Karabarbonis and Neiman, 2014) 등을 고려할 때 노동대체효과가 더 우세해질 가능성도 존재
  - Acemoglu and Restrepo(2019)는 실제로 1990년대 중반 이후 대체(displacement) 효과가 더 부각되고 있음을 제시함(그림 참조)

[그림 II-5] 새로운 직무의 생성 및 기존 직무 대체로 인한 고용 및 임금 효과



자료: Acemoglu and Restrepo(2019)의 내용을 Hatzius et al.(2023)이 재가공

□ 실증적 분석결과

- Acemoglu and Restrepo(2020a)는 미국에서 1990년 이후 산업용 로봇의 도입의 효과로 고용과 임금이 모두 하락한 것으로 추정
  - 그러나 이는 고용감소 효과를 찾지 못한 다른 문헌들과는 대비되는 것으로 Graetz and Michaels(2015)는 17개국의 산업용 로봇 도입효과를 분석하여 성장률, 총요소생산성, 임금이 긍정적, 고용에 중립적 효과를 보고
  - Autor and Salomons(2017) 역시 19개국, 35년의 기간에 대한 연구결과 개별 산업별로는 고용감소가 있다고 하더라도 전체적으로는 생산성 증대와 스페일오버 효과로 인한 고용증대 효과가 관찰된다고 보고함
  - 선진국에서의 고용 및 노동소득분배율 감소는 중국으로의 생산이동(Autor, et al. 2013), 슈퍼스타 기업들의 출현(Autor et al. 2020) 등의 효과가 중첩된 측면도 있어서 쉽게 판단하기는 어려움

### (3) 생성형 AI의 고용효과

#### □ 고용효과에 대한 영향추정 방법론

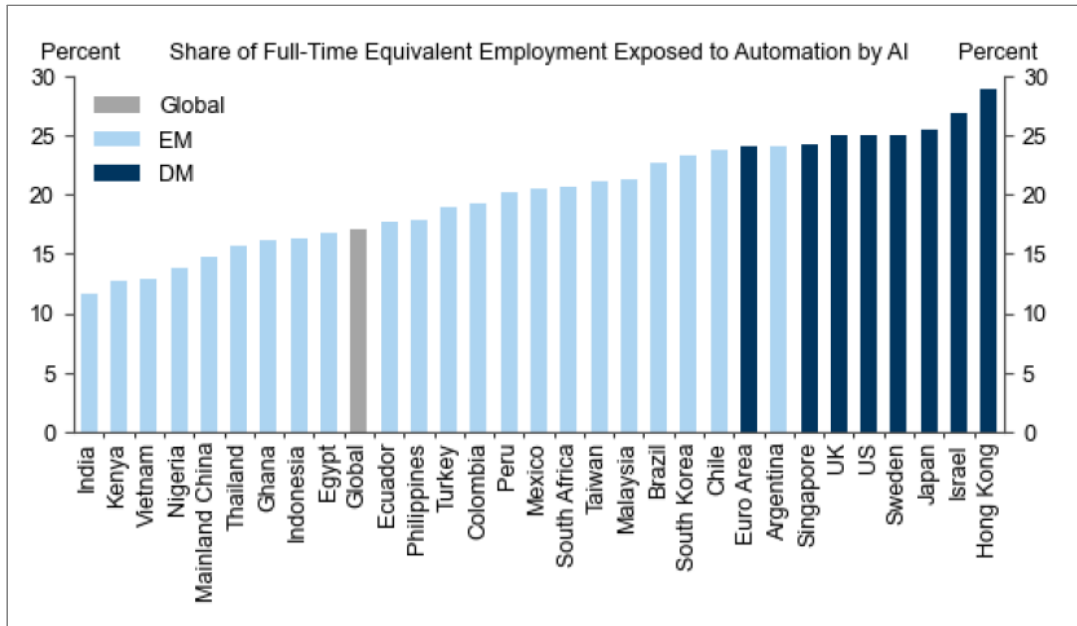
- 생성형 AI는 아직 초기단계여서 노동시장에 미치는 파급효과를 분석하거나 예측하기는 어려움
  - 주로 직업별로 필요한 직무를 분석하여 생성형 AI의 능력과 연계한 뒤 영향 여부를 분석하는 접근법의 연구들이 제시되고 있음
- 이러한 방식으로 직무에 미치는 AI의 고용효과를 분석한 잘 알려진 기존 연구로는 Frey and Osborne(2017)이 있음
  - 이들은 현존하는 직업 702개의 직무내용을 분석하여 이 중 47%가 10~20년 내에 컴퓨터로 대체될 가능성이 크다고 분석
  - 예를 들면 회계원이 없어질 확률은 94%, 판매점원 92%, 기계공 65%, 파일럿 55%, 경제학자는 43%
  - 직무를 컴퓨터로 수행하게 될 가능성이 높다고 해서 직업이 사라지는 것은 아니라는 점에서 이러한 접근의 한계는 분명히 인식할 필요가 있음
- 한편, 실험연구나 온라인 구인사이트 등을 이용한 초기 연구도 진행됨
  - 통제된 실험실에서 생성형AI에 대한 노출로 인해 직무능력이 어떻게 변화하는지, 직접적 영향을 받는 직업군의 구인구직 동향이 어떻게 변화하는지 등을 분석
  - ChatGPT 이전의 특화된 생성형AI모델의 영향에 대한 실증연구로는 Brynjolfsson et al. (2023)을 들 수 있음

#### □ Hatzius et al. (2023)

- 투자은행인 골드만 삭스의 내부 연구진은 미국(O\*NET)과 유럽(ESCO)의 직무 분류를 분석하여 AI의 영향을 분석
  - 미국 직업의 2/3이 어떤 수준으로든 AI의 영향을 받고, 그 들 중 대부분은 직무의 25~50%가 대체될 수 있다고 추정하였으며, EU도 양상은 비슷
  - 직업별로는 행정(46%)와 법무(44%)전문가의 업무대체 가능성이 높고, 건설(6%), 유지보수(4%)업의 대체 가능성은 낮음

- AI의 노출예측 정도를 바탕으로 하면 미국기준 고용의 7%가 AI로 대체되고 63%는 직무에 AI를 활용하며, 나머지 30%는 변화 없음
- 미국과 EU에서 추정된 결과를 타 국가의 산업구성에 대입함으로써 선진국(DM)과 신흥시장(EM; Emerging Market)에서의 영향을 추정
  - 그 결과 EM에서는 DM보다 낮은 영향을 받지만 전세계적으로 평균 18%의 직무가 영향을 받을 수 있을 것으로 추정함(그림 참조)
  - 물론 AI 발전의 시나리오와 전파속도에 따라 영향의 수준도 달라질 것임

[그림 II-6] AI 자동화가 고용에 미치는 영향에 대한 국가별 비교



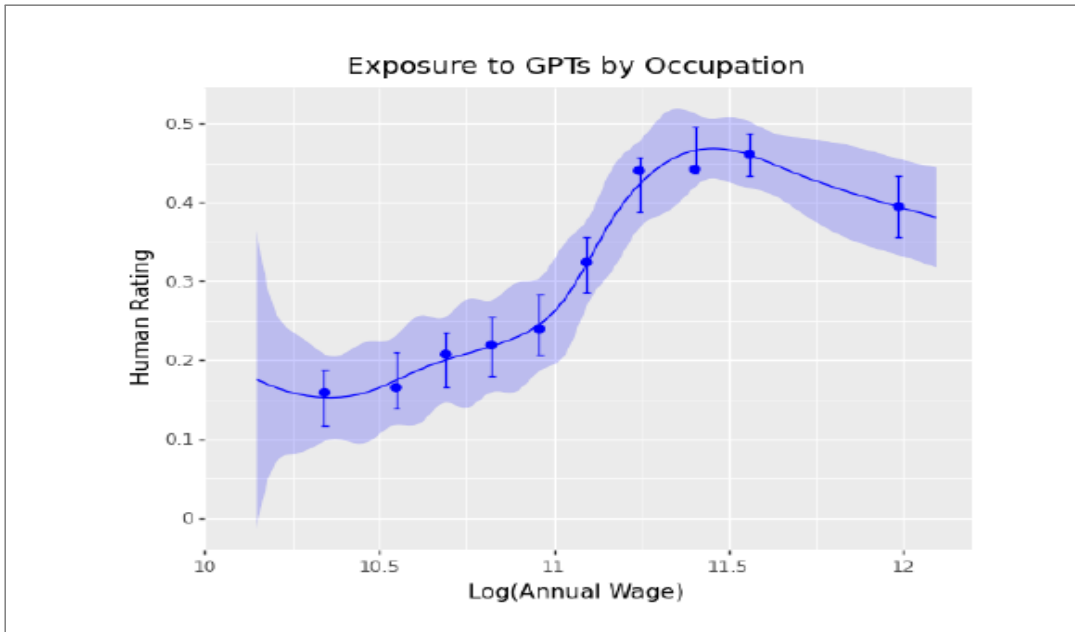
자료: Hatzius et al.(2023)

□ Eloundou et al. (2023)

- ChatGPT의 출시 이후 OpenAI의 연구진은 LLM이 직무에 미치는 영향을 분석하는 논문 출간
  - 미국의 1,016개 직업(O\*NET)의 직무내용과 GPT-4의 능력을 비교하여 영향에 노출되는 정도를 분석
  - AI에 대한 전문지식을 가진 전문가의 평가와 GPT-4의 평가를 결합(대부분 일치)

- 분석결과, 평균 14%의 직무가 직접적 영향을, 56%는 간접적으로 영향을 받음
  - 특이한 점은 소득이 높은 직업일수록 GPT의 영향에 더 노출된다는 것인데 다만 최상위층은 상대적으로 덜 영향을 받음(그림 참조)
  - 과학(science)이나 비판적 사고(critical thinking)가 중요한 직업은 노출이 적고, 프로그래밍이나 글쓰기 능력이 필요한 직업은 많이 노출됨

[그림 II-7] 임금 수준별 생성형 GPT로 인한 직무 노출도



자료: Eloundou et al.(2023)

□ Felten et al. (2023)

- 유사하지만 약간 다른 연구로서 LLM이 가장 두각을 나타내는 두 가지 분야인 언어와 이미지 생성 기능의 영향에 집중함
  - 10가지 AI응용분야를 52개의 인간 능력과 결합하고, 이를 다시 O\*NET의 직업 DB와 연결
  - 언어모델은 14.6%, 이미지 인식 및 생성기능은 4.6%의 직업에 영향을 주는 것으로 추정

- 좀더 흥미로운 결과는 임금, 학력수준, 창의력, 성별 등과의 연관성임
  - 중위임금 수준이 높을수록, 학력수준이 높을수록, 창의적 능력이 요구되는 직업일수록 생성형 AI에 대한 노출도가 높아짐
  - 직업종사자의 인구통계학적 분포로는 여성, 백인, 아시안 인종일수록 노출도가 높아지고, 남성, 흑인, 히스패닉일수록 노출도가 낮아짐

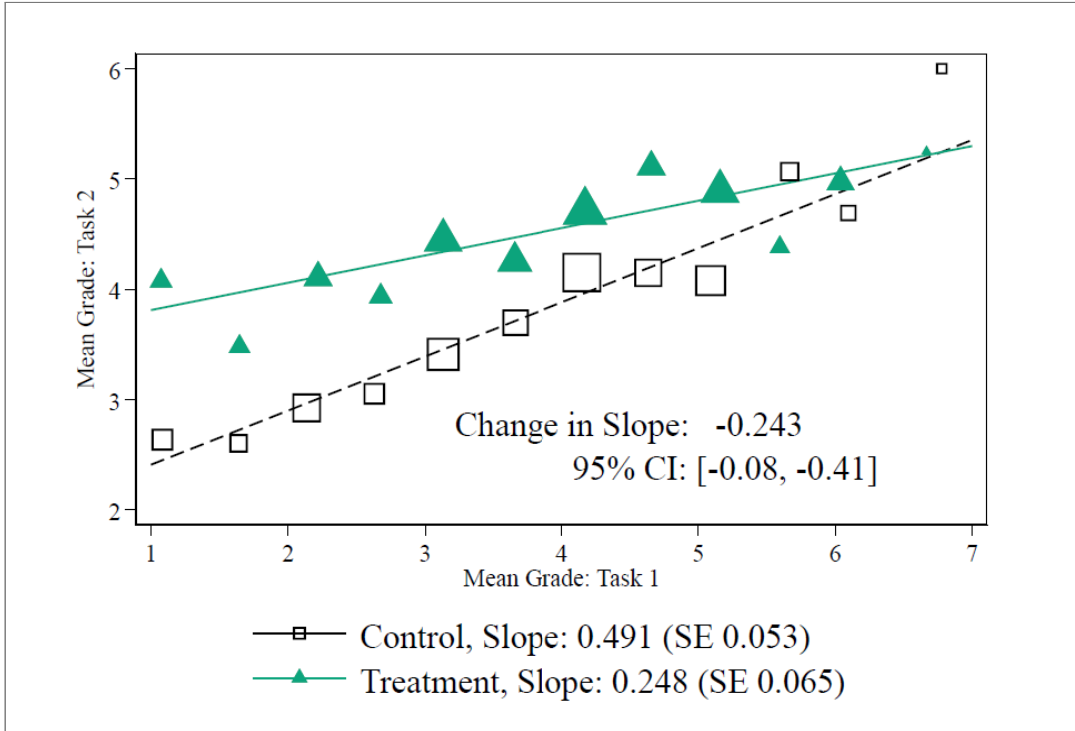
□ Lou et al. (2023)

- Eloundou et al.(2023)과 유사한 방식을 중국에 적용하여 생성형 AI가 개발도상국의 직업에 미치는 영향을 분석함
  - 분석 결과 평균 4%의 직무가 직접영향을, 17%가 간접 영향을 받는 것으로 나타나서 미국의 경우에 비해 영향받는 정도는 확실히 적었음
  - 정보처리와 금융관련 업종의 노출 정도는 큰 반면, 건설업, 광업 등의 노출 정도는 낮음
  - 고학력, 저연령층일수록 노출도가 높고, 지역별로는 대도시 지역일수록 노출도가 높았음

□ Noy and Zhang (2023)

- 실험방식의 연구로서 444명의 대졸수준 전문가를 대상으로 각종 직업과 관련된 글쓰기 과제를 부여
  - 무작위적으로 ChatGPT를 활용할 수 있도록 하고, 양적 질적 생산성을 측정
  - 그 결과 ChatGPT를 활용한 경우 평균적으로 시간적으로는 0.8SD, 품질 측면에서는 0.4SD만큼의 생산성 향상효과 측정
- 특히 흥미로운 결과는 고능력자와 저능력자 간의 격차를 줄인다는 점임
  - 2개의 과제를 순차적으로 수행시키고, Treatment 그룹에만 제2과제 수행시 ChatGPT를 쓰도록 한 결과 글쓰기 시간이 전반적으로 단축되는 가운데 특히 저평가자의 품질상승 정도가 커서 격차가 완화됨
  - 글쓰기 기술보다는 노력에 대한 보완 효과가 더 크며, 초고를 쓰는 시간과 노력을 줄여서 아이디어와 편집에 집중할 수 있도록 해줌

[그림 II-8] ChatGPT의 활용과 글쓰기 품질의 격차 완화



주: Control은 2개의 task 모두 ChatGPT를 활용하지 않은 경우, Treatment는 2번째 task 수행시 ChatGPT를 활용한 경우를 나타냄

자료: Noy and Zhang (2023)

□ Choi and Schwarcz (2023)

- 법학전문대학원 시험에 GPT-4를 활용할 수 있도록 한 그룹과 그렇지 않은 그룹의 성과를 비교한 연구
  - 객관식 시험의 성과는 크게 향상되었으나 복잡한 서술문항은 그렇지 못함
  - 여기서도 기존 저성과자의 향상 정도는 큰 반면, 최고성과자의 경우 오히려 부정적 결과가 발생

□ Dell'Acqua et al. (2023)

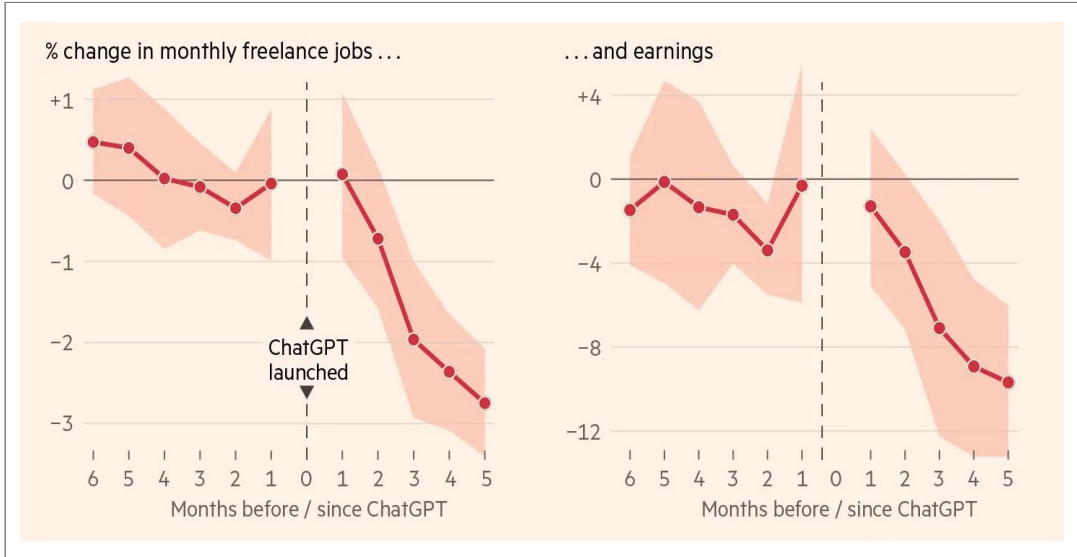
- 하버드 경영대학원과 BCG 컨설팅 그룹의 공동 실험연구
  - BCG 컨설턴트 758명(전체 7%)을 대상으로 no AI, GPT-4, GPT-4 + 프롬프트 엔지니어링 학습 등을 적용하여 비교, 평가

- 독특한 점은 AI가 성과를 낼 수 있는 영역과 그렇지 않은 영역을 구분하여 실험하였다는 것임 (후자는 정성적, 정량적 정보를 종합하는 분석업무)
- 실험결과 AI를 활용한 그룹의 성과는 크게 개선되었으며, 다른 연구에서 도 관찰된 격차해소 효과도 두드러짐
  - 평균 12.2% 더 많은 업무를 수행하고, 25.1% 더 빨리 수행하였으며, 통제그룹에 비해 40% 더 높은 품질을 달성함
  - 평균이하의 성과자는 43%의 점수향상 효과를 얻은 반면, 평균이상 성과자는 17%의 성과향상만을 나타내서 평준화 효과가 나타남
- AI가 성과를 내기 어려운 영역에서는 AI의 이용이 오히려 부정적인 효과로 이어짐
  - AI를 활용한 그룹이 그렇지 않은 그룹에 비해 19% 더 잘못된 답을 낼 가능성이 높음
- AI의 성과를 극대화하기 위한 2가지 방법론을 파악, 제시함
  - 첫째는 'Centaur' 방법론으로 AI와 인간이 자신이 잘 할 수 있는 역할에 따라 업무를 분담하는 것이고, 둘째는 'Cyborgs'로서 AI와 지속적인 소통을 통해 업무를 완전히 결합시키는 전략임

□ Hui et al. (2023)

- ChatGPT의 출시 이후 온라인 프리랜서 매칭시장인 Upwork에서 집필 관련 업무에 대한 구인구직 상황변화를 분석
  - '22년 1월~'23년 4월의 기간을 분석한 결과 ChatGPT의 출시 이후 집필 및 편집관련 프리랜서 월평균 의뢰업무의 개수는 2%, 수입은 5.2% 감소한 것으로 나타남
  - 해당월에 작업을 1개 이상 하였던 프리랜서가 새로운 작업을 추가로 하게될 확률은 1.2%p 또는 10% 감소하였고, 작업 개수도 4.7% 감소

[그림 II-9] ChatGPT 출시 이후의 글쓰기 프리랜서 고용변화



자료: Hui et al.(2023), 그림은 Financial Times(2023.11)의 기사에서 재인용

- Dall-E2와 Midjourney 등 이미지 관련 프로세싱 AI의 출시 이후의 분석도 동일한 결과를 보여줌
  - 의뢰업무의 개수는 3.7%. 수입은 9.4% 감소하는 것으로 나타남

□ Brynjolfsson, Li, and Raymond (2023)

- ChatGPT 이전 버전의 생성형 AI기반 상담 채팅 도입이 소비자 응대 서비스에 미친 영향을 실증적으로 분석
  - 5,179명의 Customer support agents에 순차적으로 AI채팅 서비스가 2020~2021년에 걸쳐 도입되었을 때 나타난 효과를 분석
  - 시간당 상담 처리량을 기준으로 생산성을 측정한 결과 AI 보조 도입의 생산성 증대효과는 14%에 달함
  - 상담처리 1건당 걸리는 시간이 감소하고, 시간당 처리가능한 상담수가 증가(동시에 복수 처리 가능)하였으며, 성공적 상담확률도 약간 높아짐
- 생산성 증가효과는 특히 경험이 적고 skill수준이 낮은 근로자에게 더 크게 나타남
  - 저성과자의 생산성 향상효과는 34%에 달하였으나 기존의 고성과자의 성과 향상효과는 최소한의 수준에 머무름

- 한편, AI를 통한 향상효과는 협력적 태도와 학습효과를 통해 나타남
  - AI의 추천을 잘 활용할수록 생산성 향상 효과가 컸으며, 시간이 갈수록 AI에 대한 수용의 정도가 향상됨 (초기에 회의적이었던 경우 더 catchup)
  - AI의 효과는 AI를 더 이상 활용하지 못하는 상황이 되어서도 어느 정도 유지됨 (학습효과가 존재함을 보여줌)
- AI의 활용은 소비자의 불만 감소에도 도움을 줌
  - 비인간화로 인한 한계에도 불구하고 오히려 소비자의 불만요소를 효과적으로 제거할 수 있는 소통의 증대효과가 더 큰 것으로 판단됨

□ 한지우·오삼일 (2023)

- 한국은행 소속의 연구자들은 Webb(2020)의 방법론을 적용하여 AI가 국내직업에 미칠 영향을 분석함
  - Webb(2020)의 방법론은 직무내용 설명과 AI관련 특허제목이 얼마나 중복되는지를 동사-명사 조합을 통해 살펴보고 AI노출지수를 추정
  - 이들은 O\*NET에 기반하여 작성된 노출지수를 한국표준직업분류로 변환함으로써 이를 비교함
- 결과에 따르면 국내 일자리 중 AI에 의해 대체될 가능성이 큰 일자리는 341만개 (전체의 12%)로 추정됨
  - 다만, 이는 상대적으로 AI노출지수 상위 20%에 해당하는 직업의 고용자수를 단순 계산한 것으로서 상당히 ad hoc한 가정에 기반한 추정임
- AI노출지수 소분류 상위직업과 하위직업 및 세분류 상위직업 결과도 보고하고 있음 (표참조)

〈표 II-1〉 한지우·오삼일(2023)에 따른 직업별 AI노출도

구분	소분류	세세분류
상위직업	화학공학 기술자 발전장치 조작용 철도 및 자동차 기관사 상하수도 처리 장치 조작용 재활용 처리 장치 조작용 금속재료공학 기술자	일반의, 한의사, 임상병리사(1%이내) 전문의사(7%) 건축가(13%) 수의사(15%) 회계사(19%) 판사, 감사, 변호사, 변리사(21%) 간호사(22%) 경찰관, 수사관(23%) 치과의사(46%), 화가, 조각가(48%)
하위직업	음식관련 단순 종사자 대학교수 및 강사 상품대여 종사자 종교관련 종사자 식음료 서비스 종사자 운송 서비스 종사자	중고등학교 교사(57%) 육아도우미(75%) 약사, 한약사(83%) 기자, 언론관련 전문가(86%) 성직자(98%) 대학교수(99%) 가수, 경호원(100%)

주: '소분류'는 직업 소분류(153)개 기준임. 세세분류 직업의 괄호 안 숫자는 상대적 노출도로 평가했을 때 상위 몇 %에 속하는 지를 나타냄. 예를 들어 '건축가(13%)'는 건축가의 노출 정도가 전체 직업군 중에서 상위 13%에 속하는 높은 수준임을 나타냄.

- 그러나 이러한 접근법은 AI의 효과를 제대로 측정하는 방법론이라고 보기 어려운 측면이 있음
  - 특정 분야에 AI특허가 많이 나온다고 해서 해당직업이 영향을 많이 받는다고 보기 어려움
  - ChatGPT와 같은 생성형 AI의 특성과는 상당히 동떨어진 특성을 반영

## 다. 성장 및 불평등에 미치는 영향

### (1) AI가 성장과 노동소득분배율에 미치는 영향

#### □ 경제성장의 정형화된 사실과 AI의 영향

- Kaldor(1957)는 1인당 GDP 성장률과, 노동소득분배율이 일정하게 유지되는 것이 경제성장의 정형화된 사실(stylized facts)이라고 관찰
  - 이러한 경향은 적어도 20세기 중반 이후까지 유지되어 많은 경제학자들은 이를 기본으로 경제성장 모형을 구성
- 하지만, Trammell and Korinek(2023)은 이러한 경향이 유지되어야 할 이유가 없다면 AI의 도입으로 다양한 형태의 성장이 가능하다고 지적
  - AI가 기존의 경제성장 방식에 어떻게 영향을 미치느냐에 따라서 다르겠지만 성장률이 더 가속화되고 노동소득분배율은 감소되는 상황이 나타날 수 있음
  - AI의 발전으로 노동과 자본의 대체성 향상, 직무 자동화, 자본의 자기복제 효과 등이 나타날 수 있고, 특히 지식생산 프로세스에 영향을 미칠 경우 폭발적인 성장도 나타날 수 있음

#### □ 자본과 노동의 대체성과 노동소득분배율

- AI로 인해 노동의 몫이 줄어들지의 여부는 요소간의 대체성(대체탄력성)의 크기 및 변화 여부에 달려 있음
  - AI가 어떤 식으로든 자본을 증대시키는 방향(자본축적, 자본증대형 기술진보)으로 작용하더라도 자본의 몫이 반드시 늘어난다고는 볼 수 없음
- 자본과 노동의 대체탄력성이 비탄력적이라면 (두 생산요소의 관계가 보완적이라면) 자본의 가격의 하락 효과가 더 커서 반대로 노동의 몫이 더 커질 수도 있음
  - AI가 생산에 미치는 영향의 경로는 다양하지만 결국 자본에 변화를 준다는 점을 고려할 때, 최종효과는 결국 자본과 노동 간의 보완 혹은 대체관계에서 얻어질 것임
- Trammell and Korinek(2023)의 문헌 연구 역시 이를 뒷받침하는 데, 자본과 노동의 대체탄력성이 크거나 AI로 인해 대체탄력성의 크기가 높아지는 모형에서 주로 노동소득분배율이 줄어드는 효과가 나타남

- 한편, 노동과 자본의 대체성이 높고 ‘토지’와 같이 고정된 보완요소가 존재할 경우에는 노동의 몫은 줄어들더라도 해당 고정요소의 몫이 커지게 됨

#### □ Baumol의 비용효과와 균형적 성장

- Aghion et al. (2019)는 Baumol 비용효과에 주목하면서 AI의 대체효과에도 불구하고 균형적 성장이 가능할 수 있음을 제시함
  - Baumol 비용효과는 생산성 증가가 빠른 부문의 비중은 감소하고 생산성 증가가 느린 부문은 오히려 비중이 높아지는 현상을 지칭
  - 경제성장의 속도는 필수적이지만 생산성 증가가 느린 부문에 의해 제약되는데, 이러한 효과는 오히려 안정적 성장이 가능하게 하는 원동력이 될 수 있음
- 균형적 성장경로에서는 자동화가 진전되더라도 자본의 비중이 특정수준 이상 높아지지 않고 GDP성장률도 안정화될 수 있음
  - 물론, 실제로 균형적 성장경로를 따르게 된다는 보장이 있는 것은 아님

#### □ 소결론

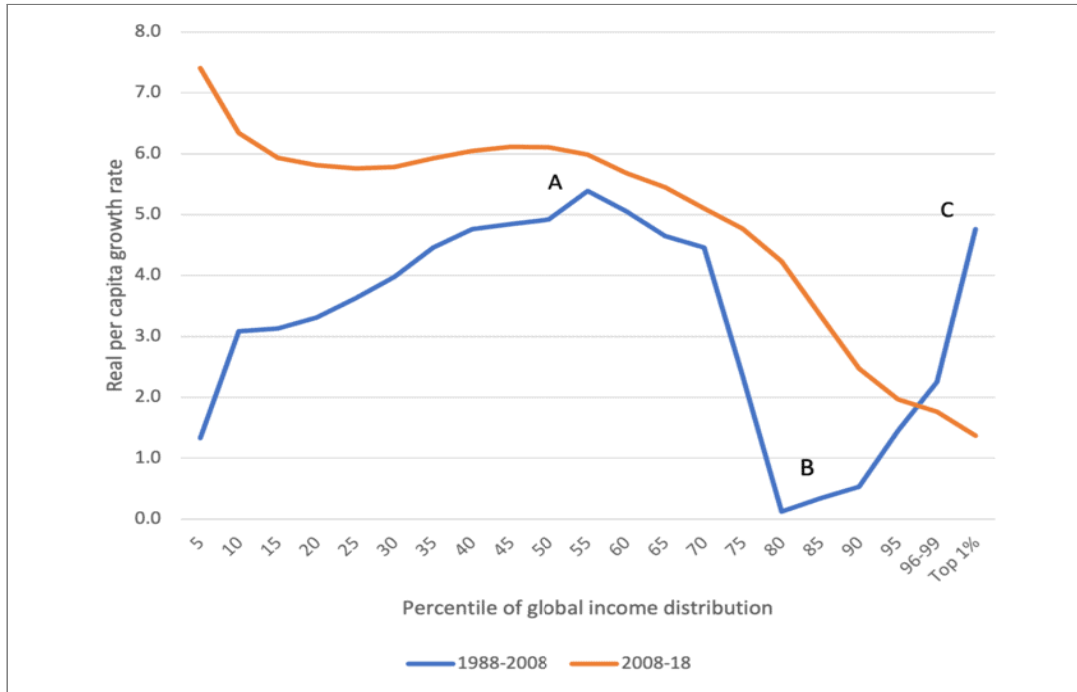
- AI가 경제성장에 미치는 구체적인 영향은 불가지론에 가까운 상태
  - AI와 자동화가 생산성과 경제성장률을 높일 것이라는 점에 대해서는 대체로 비슷한 결론이지만 그 정도는 불분명 (다양한 결과)
  - 노동소득분배율의 경우 지속적으로 줄어들 것이라는 전망과 일정 수준으로 유지될 수 있을 것이라는 전망이 공존함
- 노동의 몫이 크게 줄어들 것인지의 여부는 인간의 노동이 AI와 보완성을 유지할 것인지에 달려 있음
  - 소득의 원천이 반드시 노동에 의존해야 하는 것은 아니지만 노동소득분배율의 감소는 불평등의 확대로 이어질 수 있다는 점에서 노동소득분배율의 유지, 상승이 바람직하다면 결국 AI의 보완적 발전 여부가 중요할 것임
  - 이는 Acemoglu and Restrepo(2020b)의 주장과도 연결되는데, 이들은 AI가 ‘인공지능’을 개발하는 것을 추구하다보니 본질적으로 인간을 ‘대체’하는 데 초점을 맞추고 있으며, 하지만 이러한 접근보다는 인간과 보완적 역할을 하면서 능력을 ‘향상’시키는 방향으로 연구가 집중되어야 한다고 주장함

## (2) AI와 불평등

### □ 불평등에 대한 기존의 정형화된 사실

- 우선, 자본주의와 산업발전으로 인해 불평등이 전반적으로 증대한다는 일반적인 인식은 정확한 것은 아님
  - 밀라노비치(2017)는 역사적으로도 불평등도의 변화가 파동처럼 증가와 감소를 반복해 왔다고 제시함
- 대중적인 주목도가 높은 것은 국가 내의 불평등도이지만 실제로 더 중요한 것은 국가 간 성장속도의 차이에 따른 불평등 수준임
  - Milanovic (2015)에 따르면 어떤 국가에 태어났는지가 한 개인의 소득수준의 70%를 설명하며, 국내적 상황과 개인의 노력 등이 나머지 30% 결정
  - 성장률의 차이가 분배수준 보다 큰 영향을 미칠 것이라는 점은 시점 간 비교로도 알 수 있는데 예를 들면 1990~2010년 동안 한국은 2.6배 성장하였고, 반면 불평등도는 상승했는데(지니계수 4.9포인트 상승) 후자로 인한 악영향을 제거하더라도 한 개인의 기대 소득수준은 2.2배 상승
  - 전 세계적으로 보면 후진국이나 개발도상국의 경제성장률이 선진국보다 빠르기 때문에 불평등도는 감소할 가능성이 크며, 실제로도 그러하였음
- 20세기 후반 글로벌화의 진전과 중국의 약진 등의 영향으로 미국 등 선진국에서 국가내 불평등이 증대된 것이 불평등 심화에 대한 인식을 강화
  - 이는 Lakner and Milanovic(2015)의 유명한 ‘코끼리 그림’으로 상징되는데, 1988~2008년에는 세계 중위권 소득자와 최고위 소득자의 소득은 빠르게 증가한 반면([그림]의 A와 C), 중상위권 소득자는 가장 부진한 모습을 보였는데(B), A에 해당하는 것이 중국 등 개발도상국 국민이고, B와 C는 각각 미국 등 선진국의 중산층 및 최상위 소득층 국민에 해당함
- 그러나 Milanovic(2022)에 따르면 코끼리 형태의 불균형한 소득성장 패턴은 2000년대 후반부터는 더이상 유효하지 않으며, 보다 일반적인(기대되는) 형태의 변화, 즉 저소득층일수록 catch-up이 빨라지는 형태로 변화됨

[그림 II-10] 글로벌 임금수준에 따른 임금상승률 비교



자료: Milanovic(2022)

- 이상의 사실로부터 얻을 수 있는 교훈은 불평등도의 진행방향을 일률적으로 판단하기 어렵다는 것임
  - 기술진보 뿐 아니라 산업구조변화, 글로벌화, 국가별 상황의 차이 등 다양한 요소들이 반영된 결과이며, 시간 별, 측정단위 별로도 결과가 달라짐

□ AI가 불평등에 미치는 영향

- 전통적으로는 노동소득분배율의 변화를 불평등과 연결짓는 경향이 강하며, 그런 관점에서는 AI의 발전이 불평등을 증대시킬 가능성이 있음
  - 노동소득만이 소득의 원천은 아니므로 노동소득분배율이 감소한다고 해서 불평등이 증가하는 것은 아니지만, 대체로 저소득층일수록 노동소득 의존도가 높다는 점에서 연관성이 있음
  - 앞에서 살펴본 것처럼 AI가 어떤 식으로 경제에 영향을 미칠지는 불확실한 측면이 많지만 노동소득분배율이 줄어들 가능성을 다수가 지적

- 국가 간 성장률의 격차와 관련해서도 AI가 더 격차를 벌릴 것이라는 시각과 반대로 격차를 뛰어넘는데 유용할 것이라는 시각이 공존함
  - Norton(2017)은 개발도상국이 글로벌 공급망 형성에서 차지하던 위치가 AI의 발전으로 인해 흔들릴 수 있다고 지적함, 즉 값싼 노동력의 활용이 크게 중요하지 않게 되면서 off-shoring 경향이 줄어들 것으로 예상
  - 반면, World Bank(2017)의 지적처럼 산업의 양상 자체가 바뀌게 되면 제조업 기반이 부족한 개발도상국의 catch-up이 더 수월해질 수도 있음
  - 또는 Noy and Chang(2023), Dell'Acqua et al.(2023)이 제시하는 것처럼 생성형AI가 지식과 경험의 격차를 완화하는 역할을 하는 방식으로 개발도상국과 선진국 간의 불평등이 완화될 수도 있음
- 한편, Norton(2017)은 AI가 불평등을 심화시킬 수 있는 몇 가지 경로를 제시하였는데 그 중 비교적 의미가 있는 것들은 다음과 같음
  - (a)노동 대비 자본의 우위 심화, (b) 중하급(low and medium-skilled) 고용의 비중 감소 (c) 노동조합의 약화 (d) 세금징수 기반의 약화로 인한 정부의 재분배 기능 약화 등임
  - 이 중에서 (a)는 이미 다루었고, (b)는 기존의 숙련편향적 기술진보(SBTC) 상황을 가정한 것인데 생성형AI 환경에서는 다를 수 있음 (c)는 노동의 협상력 약화를 전제한 것으로서 원인이라기보다는 결과에 가깝고, (d)역시 소득세 기반 약화와 자영업 확대 등을 이유로 들고 있으나 소득세는 전체 징세기반의 일부라는 점에서 설득력이 있다고 보기 어려움
- 한편, Brynjolfsson(2022)은 AI로 인한 불평등 심화의 가능성에 대해 독특한 시각을 제시함
  - 보수를 더 많이 가져갈 수 있는 협상력의 차이는 핵심자산의 보유여부에 달려 있는데 유용한 지식이 인간의 뇌에 한정되어 체화되는 한 힘의 격차에는 한계가 존재
  - 그런데, 인간에 체화되는 인적자본의 격차는 크기와 집중도에 한계가 있는 반면, AI의 경우 코드화된 기술과 지식이 인적자본을 대체할 수 있고, 이 경우 기존과 달리 무한히 집중될 가능성이 있다고 지적함

- 그런데, 불평등은 인간 단위에서 측정되는 것임을 고려하면 결국 이러한 주장은 특정인이 AI의 성과를 독점하는 상황을 경계하는 것이라고 볼 수 있으며, 이것이 현실화될 수 있느냐는 또 다른 문제임(국가권력에 의한 강제 재분배로 이어질 수도 있음)

#### □ 소결론

- AI, 특히 생성형 AI가 불평등을 심화시킬지의 여부 역시 불확실성이 큼
  - 다른 기술진보와 마찬가지로 이득과 손해를 보는 계층이 존재하는 것은 사실이며, 반드시 불평등으로 이어지지 않을 수 있다는 점도 마찬가지
- 오히려 불평등을 심화시킨다는 대중적 인식을 극복하고 불평등 완화를 위한 방식으로 AI의 활용이 발전될 수 있도록 노력하는 것이 중요
  - AI가 불평등을 심화시킬 것이라는 대중의 인식은 AI의 채택 및 확산에 부정적인 요인으로 작용 가능함
  - 생성형 AI의 직무영향이 중간 이상 소득자에게 크고, 생산성 향상효과가 저성과자와 고성과자 간의 격차를 줄인다는 연구결과 등을 고려하면 오히려 AI를 어떻게 잘 활용하느냐가 불평등의 감소에 기여할 것임

## 라. 경쟁 및 기업조직 변화

### (1) 산업구조 및 조직에 미치는 영향

#### □ Varian(2019)의 평가

- AI로의 수직적 계열화 정도
  - AI의 활용도가 높아지면서 make or buy 결정이 중요한 요소로 떠오를 것이나 수직계열화의 가능성이 높다고는 보지 않음
  - HW와 SW자원을 모두 확보하고 서비스 표준화와 유연성을 갖춘 클라우드 사업자의 상대적 경쟁력을 강조
- 기업의 규모와 범위
  - AI의 이용이 사업별로 규모의 경제 수준에 영향을 미치는 지와 관련이 깊은데, 이 역시 서비스화와 아웃소싱 확대 가능성이 더 높다고 평가

● 가격 차별화의 정도

- 가격설정의 유연화가 가능하고, 맞춤화 수준이 높아지면서 가격차별화의 정도가 높아질 것이나 경제학은 이를 특별히 부정적으로 보지 않음
- Dube and Misra(2017)에 의하면 ML에 의한 가격책정시 이윤은 10% 증가하나 소비자 후생은 별 차이 없었음
- DellaVigna and Gentzkow(2017)에 따르면 유연화된 가격체계에서 저소득층은 고소득층에 비해 확실히 더 이익을 얻음

● 규모의 경제 수준

- 공급측면의 규모의 경제와 관련하여 통상적인 인식과 달리 소프트웨어의 서비스화, 주기적 업그레이드 등으로 인해 더 이상 높은 개발비용과 낮은 개발비용으로 인한 규모의 경제효과가 크지 않다고 지적
- 네트워크 효과의 경우 동일 프로그램 언어나 클라우드 사용으로 인한 선호 축적의 형태로 직접적 네트워크 효과는 있을 수 있으나, 오픈소스가 큰 비중을 차지하고 있고, 타 산업과 차별이 될만한 수준도 아님
- 사용자가 많으면 데이터 축적이 용이해서 서비스 수준도 높아지는 효과를 ‘데이터 네트워크 효과’라고 언급하는 경우가 많으나 이는 사실 네트워크 효과가 아니라 ‘학습효과(learning by doing)’라고 지적
- 학습효과는 기업의 노력과 투자의 결실이라는 점에서 네트워크 효과와 구분되며, 문제는 ‘자원(resources)’의 보유가 아니라 ‘활용 능력(skills)’을 어떻게 획득하고 적용하느냐임(그에 따라 다양한 결과가 가능)

● 알고리즘 담합

- 반드시 명시적 합의가 없더라도 상호작용의 빈도가 높아지면 암묵적(tacit) 담합과 같은 결과가 나올 수 있다는 것은 경제학에서 잘 알려짐
- 마찬가지로 알고리즘을 이용한 가격책정이 일반화되면 자연스럽게 암묵적 담합과 유사한 결과가 초래될 가능성을 배제할 수는 없음
- 하지만, 적어도 아직까지는 과점적 경쟁상황에서 어떤 알고리즘이 실제로 그런 결과를 이끌어 낼 수 있는지 알려진 바 없음(단순 알고리즘으로는 구현되기 어려움)

## (2) 기업조직의 변화

### □ 대리인 비용의 감소

- Allen et al.(2023)은 LLM에 대리인 비용 감소효과가 있으며, 이러한 효과가 기업의 크기를 축소하고 경쟁을 강화할 것이라고 주장
  - 주인(principal)이 대리인(agent)의 작업 품질을 확인하고 모니터링하기 위해 드는 비용이 감소하고, 더 나아가 작업을 아웃소싱할 필요가 줄어들어 make or buy 결정의 필요성 자체가 적어짐
  - 기업의 최적 규모가 줄어들고 분권화된 혁신이 가능해지며, 창업에 드는 비용도 낮아짐으로 인해 경쟁의 활성화, 노동시간 감소, 소비자 후생 증가 등의 생산성 효과가 일어날 것이라고 긍정적으로 전망
- 그러나 최적 기업규모가 축소되고 진입비용이 감소된다는 전망은 지나치게 단순한 접근일 수 있음(Tirole, 2017)
  - 일상적인 직무는 AI의 도움을 받아 외주화하고 특화된 분야에 집중함으로써 수평조직화가 가능하고 개인창업이 용이해지는 측면은 있음
  - 반면, AI의 활용을 위해 대규모 고정비용 투자가 필요하다면 대기업일수록 위험과 비용을 감수할 수 있고, 관계특정적(relationship specific) 성격을 가지는 투자라면 수직계열화가 더 진전될 수도 있음
  - 활용하느냐가 불평등의 감소에 기여할 것임

## 마. 플랫폼 경제에 대한 시사점

### (1) 평가의 관점

#### □ 플랫폼 경제의 주요특성과 AI의 적용

- Furman 리포트(Furman et al. 2019), Stigler 리포트(Zingales et al. 2019) 등으로 불리는 플랫폼 규제논의를 촉발했던 일련의 보고서들은 플랫폼 경제의 특수성을 다음과 같이 정리하고 있음
  - 쏠림현상(tipping)과 승자독식(winner takes all) 경향: 양면시장적 특성으로 인한 네트워크 외부성, 디지털 재화의 특성인 규모 및 범위의 경제 등으로 인한 효과

- 제로가격으로 인한 착시현상: 소비자 후생 측정이 어려우며 품질저하, 다양성 저하, 혁신 감소 가능성 등에 대한 감시 필요성
- 기존 사업자의 데이터 우위: 데이터 축적과 활용에 있어 중소기업 경쟁자나 신규 진입자는 누리기 어려운 선순환 이점을 가짐
- 전환비용: 플랫폼 사용을 통한 개인정보, 명성 등의 누적, 표준 및 호환성, 끼워 팔기 등 플랫폼의 전환비용 상승전략으로 인한 고착화
- 행동경제학적 편향: 플랫폼의 높은 소비자 분석 능력과 선택대안 부재 등으로 인해 소비자들이 조종될 우려
- 내부 자본조달의 이점: 대규모화된 플랫폼은 내부 자금조달의 이점을 가짐
- 고려해야 할 것은 본질적으로 플랫폼은 ‘비즈니스 모델’인 반면 AI는 ‘테크놀로지’라는 점임
  - 플랫폼과 AI의 특성은 서로 독립적이면서 중첩될 수도 있으며, 플랫폼 기업이 AI를 가장 잘 개발하거나 활용할 수 있는 주체라는 보장은 없음
  - Google, Facebook, Microsoft 등이 빅테크이면서 AI 연구개발을 활발히 하기 때문에 플랫폼 기업이 AI의 주도적 기업인 상황이 될 수는 있음
  - 그러나 AI가 가장 잘 활용되어 비즈니스와 경제적 변화를 만들어 내는 주요 분야가 플랫폼이 되리라는 보장은 없음 (즉, 타 분야의 변화가 훨씬 더 클 수 있음)
- 또한 제시된 특성 중 비즈니스 모델이나 사업 전략적 특성, 특히 플랫폼 고유의 특성이라고 보기도 쉽지 않은 것들은 검토할 필요가 없어 보임
  - 제로가격으로 인한 착시현상, 행동경제학적 편향, 내부자본조달의 이점 등이 이에 해당함
  - AI의 사용으로 인해 플랫폼의 전략이 더 정교해질 수 있다고 주장할 수도 있겠지만 이는 더 성능 좋은 컴퓨터를 쓰면 사업능력이 높아진다는 수준의 주장에 가까움

## (2) AI 특성이 플랫폼 경쟁에 미치는 영향

### □ AI의 특성과 플랫폼 경쟁

- AI의 특성을 플랫폼과 연결지은 연구는 아직은 드문 편임
  - 소수의 working paper들은 주로 플랫폼에 대한 규제를 옹호하는 입장에서 이들 플랫폼이 AI 경쟁에서도 두각을 나타내는 것을 우려하고 있음
  - AI자체의 특성을 깊이 분석하기보다는 플랫폼 규제를 강화하거나 AI기업도 플랫폼과 함께 규제해야 하는 논리로 동원하는 것으로 보임
- Kollnig and Li (2023)는 생성형AI의 발전이 플랫폼의 시장지배력을 강화할 것이라고 주장
  - 데이터 수집의 용이성, 우수 인력의 확보, 안정적 개발 인프라, 학계에 미치는 영향 등의 요소들에 있어 플랫폼이 이점을 가지고 있으며, Google의 생성형 AI 분야로의 수직계열화도 견제받지 않았다고 지적
  - 결과적으로 신규 진입자가 스스로 경쟁력 있는 생성형 AI 모형을 만들어 내기 어렵고, 현재의 플랫폼 지배력이 계속될 가능성이 크며, EU의 AI 규제법 역시 기존 사업자에게 더 유리하게 작용할 수 있다고 주장
- Yasar et al. (2023)은 AI 소프트웨어 기업들을 EU DMA의 ‘핵심 플랫폼 서비스(core platform services)’에 포함시켜 규제해야 한다고 주장
  - ChatGPT와 같이 ‘freemium’ 형태로 대중에 개방하거나 API를 제공하는 등 AI서비스가 플랫폼화하는 모습을 보임
  - 데이터 자원, 특화된 하드웨어 자원의 집중, 수직결합, 네트워크 효과, 재정적 기반, 노하우, 선발자 이익 등으로 인해 소수의 기업이 주도하며, 이들 중 상당수의 특성은 DMA와 gatekeeper 규제에 이어진 디지털 시장 특성과 유사하므로 규제가 필요하다고 주장
  - 플랫폼의 서비스와 생성형 AI가 결합함으로써 서비스 전환이 더 어려워지고 있으며, 데이터와 서비스 결합을 금지하는 5(2)(b) 및 5(8)이 생성형 AI를 포함하지 않고 있다고 지적
  - 소수 생성형AI기업으로의 집중은 경합성과 공정성을 저해하며, 현재의 AI Act와 DMA로는 이러한 문제를 해결할 수 없으며 DMA의 여러 조항들(5~7)이 생성형 AI에도 직접 확장할 필요가 있다고 주장

□ 참고: 생성형AI 기업 간의 경쟁

- Varian(2019)은 LLM 이전의 환경을 기반으로 하고는 있으나 치열한 경쟁구조를 예측한 바 있음
  - ML 서비스를 제공하는 클라우드 간의 porting이 어렵지 않고, 오픈소스 툴도 마련되어 있는 상황에서 적어도 중소형 응용서비스에 대한 고착효과는 크지 않음
  - HW의 경우도 복제, 확장이 어렵지 않고(적어도 CRS) 클라우드의 잉여 컴퓨팅 자원을 활용할 인센티브가 있기 때문에 경쟁을 더욱 촉진(표준화된 서비스는 베르트랑 가격경쟁으로 이어질 가능성 큼)
- 반면, Vipra and Korinek(2023)는 소수의 기업들에 의해 생성형 AI가 장악될 가능성이 높으면서 규제 강화를 주장
  - 파운데이션 모델의 고정비용(training)은 높고 한계비용은 낮음, 다양한 용처가 있어서 범위의 경제도 존재
  - 선발자의 이익이 존재하고 재능, 데이터, 컴퓨팅 자원, 지식재산권 등으로 인해 자연독점적 성격을 지님
  - 수직적 결합, 약탈가격, 규제포획 등의 가능성이 있는 반면, 안전위험을 내재화 가능하다는 것은 장점. 오픈소스의 존재는 시장집중을 완화
  - 이에 따라 규제자는 경합성을 높여야 하고 특히 수직결합 가능성에 주목해야 하며, 최첨단의 파운데이션 모델에 대해서는 자연독점인 공익설비의 관점으로 접근해야 한다고 주장
  - AI와 비AI에 대해 동등한 접근이 필요하며, 이는 어떤 분야에서는 AI의 활용을 제약하는 결과로 이어질 수 있다고 주장
- Carugati(2023)도 유사한 주장을 제기
  - 데이터는 양, 품질, 신선도 등이 중요하고 네트워크 효과도 있음
  - AI를 위해서는 컴퓨팅 자원들(GPU, 서버, 슈퍼컴, 네트워크 장비 등)이 필요하고 클라우드 사업자들은 소수에 집중됨
  - 위와 같은 특성은 높은 진입장벽으로 이어지고, 사업자들이 다양한 경쟁제한적 배제전략을 취할 가능성도 높다고 주장함

## □ 플랫폼 경쟁에 미치는 효과에 대한 평가

- AI 또는 생성형 AI의 특성이 플랫폼적 특성을 강화할 것인지의 여부는 단정지을 수 없음
  - 플랫폼의 가장 큰 특성인 양면시장이나 간접적 네트워크 외부성은 AI와는 연관이 없음
  - 규모의 경제효과는 확실히 존재하지만 이는 대규모 연구개발 투자를 필요로 하는 디지털 기술이 공유하는 특성이며, 플랫폼과 반드시 연관되지 않음 (예를 들어 플랫폼 회사들이 신약을 개발하는 경우를 가정해 볼 수 있음)
  - 전환비용과 고착효과도 플랫폼에게 반드시 유리한 방향으로 작용하리라는 보장은 없는데, 기존의 플랫폼이 AI를 이용하여 전환비용과 고착효과를 높일 수도 있지만, 이에 도전하는 서비스들이 AI를 이용하여 전환비용을 오히려 낮출 수도 있음 (기존에 전환비용을 높이는 요소로 작용했던 누적된 경험과 편의성을 AI를 통해 쉽게 복제하는 등의 방식)
- 이렇게 볼 때 플랫폼 경쟁에 직접적 영향을 줄 수 있는 요소는 플랫폼의 데이터 축적 및 활용능력 정도로 볼 수 있음
  - 이러한 능력은 AI 간의 경쟁에서 유리한 요소로 작용할 수 있으며, AI를 이용한 플랫폼 서비스의 향상에도 유리한 피드백 효과로 작용 가능
  - 하지만, 이는 효율성을 제고하는 '시너지'로 해석될 수 있는 부분이기도 해서 그 자체로 규제나 견제의 필요성을 제시하는 요소로 보기는 어려움(예를 들어 유전개발 사업자가 셰일가스의 개발이나 정유사업에 시너지가 있다고 해서 그 자체가 우려의 대상이 되는 것은 아님)
  - 또한 데이터 우위로 인한 피드백 효과가 있다고 해서 대형 플랫폼일수록 더 유리한 상황이 될 수 있는지는 미지수임 (예를 들어 일정 규모 이상의 플랫폼에 비슷한 크기의 이익을 가져다 준다면 플랫폼 경쟁에 미치는 영향은 크지 않을 것임)
- 기타 기존 플랫폼의 재정적 기반, 선발자 이익, 빅테크로서의 노하우 등도 기존의 플랫폼들이 AI경쟁에 있어 앞선 위치를 차지할 수 있는 원동력이 될 수 있음
  - 다만, 이런 요소들은 본질적 우위로 간주할 만한 특성까지는 아니라고 판단됨 (역사적으로 이런 형태의 우위는 흔히 존재했으며, 생각보다 굳건하지 않음)

- 다만 플랫폼 간의 경쟁에 있어서 AI 개발 및 서비스 능력이 새로운 차원의 경쟁요소가 될 가능성이 있다는 점은 주목할 필요
  - 경제의 다른 부문과 마찬가지로 플랫폼 서비스도 AI와 결합되어 발전할 것이며, 주요 플랫폼들이 AI개발에 선도적 역할을 한다는 점도 작용
  - 이러한 상황이 글로벌 플랫폼 간, 해외 및 국내 플랫폼 간의 미래 경쟁구도에도 영향을 줄 수 있다는 점은 분명함

## 바. AI에 대한 정책적 관점

### □ AI의 발전을 늦추는 규제 필요성 여부

- AI가 인류의 미래를 위협할 수 있다는 우려가 형성되면서 AI의 발전속도를 늦추는 규제를 주장하는 사람들이 생겨남
  - 일론 머스크, 스티브 워즈니악 등이 포함된 Future of Life 재단은 2023년 3월 AI의 개발을 6개월간 중단하자고 제안하였고, 아예 전면 중지하자는 주장까지 등장(Youdkowsky, 2023. 3.29)
- 관련하여 Acemoglu and Lensman(2023)는 재앙의 가능성이 있을 때 왜 AI의 개발 및 활용속도를 늦춰야 하는지를 설명하는 경제학적 모형을 개발
  - 기술의 성장속도가 빠를수록 재앙으로 인한 피해의 크기도 커지고, 미래에 대한 실질적 할인율도 낮아짐(오래 기다려도 빠른 성장으로 인한 따라잡기가 쉬워짐)
  - 이 두 가지 효과를 고려하면 AI의 성장 및 생산성 효과가 빠르게 이루어질수록 이의 채택은 더 늦춰져야 한다는 irony한 결과로 이어짐
  - 사기업은 이러한 사회적 효과를 고려하지 않기 때문에 비효율적으로 빠른 개발에 나서게 됨
- 한편, Jones(2023)은 빠른 성장과 재앙의 위협으로 인한 상충(trade-off)을 접근하는 데 중요한 요소는 효용함수의 형태라고 지적
  - ‘로그’형태의 효용함수에서는 성장으로 인한 소비증대 효과가 커서 인류절멸의 위기 확률이 1/3에 이르는 수준도 수용 가능

- 위험회피 수준이 일정 이상(계수 2 이상)인 CRRA 함수에서는 소비증대로 인한 효과가 한계가 있기 때문에 훨씬 보수적이 됨 (효용함수의 boundedness가 중요한 역할)
- 이러한 연구들의 한계는 AI의 발전을 늦춰서 위험을 제거하는 것이 '가능'하다는 전제를 하고 있다는 것임
- 기술발전 경로의 불확실성, 국제적 경쟁, 사회적 선호 도출의 어려움 등은 고려되지 않음
- 현실적인 상황 하에서는 AI의 개발을 늦추는 것은 현실적이지도, 바람직하지도 않을 가능성이 있음

#### □ 노동친화적인 AI 연구개발 방향설정 필요성

- Acemoglu and Restrepo (2019), Brynjolfsson (2022) 등은 노동대체적이 아닌 노동생산성 증대적인 AI개발의 필요성을 역설
- AI는 노동을 대체할 수도 있고, 노동과 보완적인 역할을 하며 생산성을 높일 수도 있는데 인간과 구분되지 않는 AI를 만들겠다는 욕구는 노동을 대체하는 AI에 대한 노력을 과잉하게 만드는 경향이 있음
- 그러나 Agrawal et al.(2023)은 이와 같은 접근이 한계가 있다고 지적함
- 일부 직무의 자동화는 다른 부문에서의 직무조정으로 연결된다는 점에서 직무조정 중심으로만 AI를 개발하는 것은 쉽지 않음
- 생성형 AI는 현재로서 불평등을 줄일 가능성이 큰데, 직무조정 중심으로 AI의 개발을 규제할 경우 오히려 불평등이 심화될 수 있음 (직무의 조정은 결국 이미 해당 직무를 수행하고 있던 종사자의 임금만 상승시키고, 더 많은 직무를 할당 받을 수도 있었던 low-skill 근로자의 기회를 박탈하게 될 수도 있음)
- 이러한 고찰은 AI에 대한 선부른 판단에 기초한 규제적 접근이 오히려 더 부정적 결과로 이어질 수도 있음을 제시

□ AI의 종류에 따른 정책적 접근의 세분화 필요성

- Schrepel and Pentland(2023)은 AI 파운데이션 모델의 종류를 3가지로 구분하고, 각각에 대해 서로 다른 접근이 필요하다고 주장
  - 일반대중(General Public) 모형은 ChatGPT와 Bard와 같이 일반 대중을 대상으로 한 대형모형이며, 규모의 경제가 중요하게 작용
  - 생태계(Ecosystem) 모형은 특정 사용자 그룹이나 산업을 위해 만들어지고, 주로 비공개 데이터를 이용해 fine tuning된 모형으로 규모의 경제효과가 일부 존재
  - 개인화된(Personal) 모형은 대형모형을 바탕으로 개인 데이터에 맞춤형된 모형으로 규모의 경제효과가 적음
- 어떤 모형의 AI인지에 따라 경쟁의 양상 및 집중화의 가능성 등이 달라질 것이며, 그에 맞는 규제적 접근도 필요
  - 우선은 일반대중 모형의 대형 사업자들이 규모의 경제를 앞세워 지배력을 향상시킬 것이므로 감시가 필요하며, 이들이 생태계 모형으로도 고착효과를 형성해 나갈 것인지, 오픈소스로부터의 경쟁압력이 잘 형성될 것인지 등도 살필 필요가 있음
  - 그에 반해 생태계 모형은 이제 막 나타나기 시작하고 있으며, 개인화된 모형은 오랜 시간 동안 치열한 경쟁을 통해 발전할 것으로 예상 (규모의 경제 측면에서도 그렇고, 수용자의 변화가 필요하다는 점에서도 그러함)
  - 규제에 있어서도 one size fits all과 같은 접근은 적절치 않음을 시사

### 3. 생성형 AI가 규제 및 경쟁정책에 미치는 영향

#### 가. 개요

##### (1) 연구의 배경

###### □ 인공지능 기술의 발전에 따른 경쟁의 양상 변화

- 인공지능 기술의 발전은 3차 산업혁명 시대의 승자인 온라인 플랫폼 사업자와 4차 산업혁명의 도래를 앞두고 재도약하려는 파이프라인 제조업체 간의 새로운 시장을 건 경쟁의 양상으로 나타나고 있음
- 새로운 경쟁의 장터에서 경쟁의 우위를 확보하기 위하여 시장을 주도하는 기업들이 앞다투어 채택하는 인공지능 기술이 바로 기계학습(machine learning) 기술과 그 한 방법론으로 최근 대표적으로 쓰이는 심층학습(deep learning) 기술임
  - 기계학습 기술은 이미 상품이나 서비스의 내용이나 구성을 결정하는 과정이나 상품이나 서비스 거래의 품질을 개선하는 과정에 사용되고 있음
  - 예컨대, T맵이나 카카오맵에서 제공하는 네비게이션 서비스의 경우 이용자가 입력한 목적지와 주행 데이터, 교통상황 데이터를 토대로 하여 컴퓨터가 기계 학습 과정을 거쳐 이를 분류하고 연관성이나 패턴을 파악하여 경로 알고리즘을 생성하도록 하고 이 중 이용자의 선호에 부합하는 경로를 제시하는 것을 그 내용으로 함
  - 아마존이나 넷플릭스와 같은 전자상거래 플랫폼 사업자들은 수집된 고객의 검색 또는 구매 경험 데이터를 토대로 하여 컴퓨터의 기계학습 과정을 거친 추천 알고리즘을 활용하는 유사한 방식을 통해 거래를 원하는 고객의 결정을 도와주고 있음

###### □ 인공지능 기술의 적용이 시장 참여자의 결정을 도와주는 수준을 넘어 아예 결정을 대신하는 단계에 이르는 상황에 대한 우려

- 구매자가 기업이 제공하는 인공지능 알고리즘을 활용한 상품이나 서비스를 이용하거나 거래 과정에서 그 도움을 받으면 선택의 품질이 높아질 수 있음

- 그러나 구매자가 제공받는 정보가 왜곡되거나 선택능력이 제한될 경우에는 거래상의 의사결정을 알고리즘에 의존하게 되는 문제나 알고리즘을 활용한 가격차별의 문제가 생길 수 있음

● 이런 문제는 인공지능 알고리즘이 판매자 쪽에서 활용될 경우에도 발생함

- 특히 경쟁에서 중요한 요소인 가격의 결정을 인공지능 알고리즘이 대신하게 되는 상황과 관련하여 알고리즘 담합(algorithmic collusion)의 가능성에 대한 논의가 활발하게 이루어진 바 있음(OECD, 2017; 최난설현, 2018 등)
- 알고리즘이 시장에서의 거래관계에 깊숙이 침투하여 거래의 당사자인 판매자와 구매자 양쪽에게 적지 않은 영향을 미치는 시대

□ 인공지능이 경쟁과 공정거래에 미치는 영향에 대한 기존의 논의

● 인공지능 알고리즘을 통한 담합의 가능성이나 인공지능이 생성한 데이터가 경쟁 관련 데이터로 인정될 가능성과 같이 기존 거대 플랫폼 사업자가 제공하는 핵심 플랫폼 서비스에서 인공지능 기술이 이용되면 발생할 수 있는 경쟁 또는 공정거래 문제에 국한

- 인공지능 기술의 적용이 확대되고 있지만, 인공지능 기술이 그 자체로 이용자가 이용할 수 있는 서비스가 되기보다는 주로 디지털 플랫폼 사업자가 제공하는 다양한 서비스의 편의성과 품질을 높이는 도구적 기술로 활용되어 왔기 때문

● 디지털 플랫폼 사업자를 둘러싼 경쟁 또는 공정거래 문제를 특유한 문제로 인식하고 이에 대응하기 위해 제정된 유럽연합(EU)의 대표적인 법률에서도 디지털 플랫폼 사업자가 인공지능 알고리즘을 서비스에 활용하는 과정에서 발생할 수 있는 문제에 대처하기 위한 일부 규정을 두고 있을 뿐임

● 「온라인 중개 서비스의 사업 이용자에 대한 공정성과 투명성 증진을 위한 법률」(REGULATION on promoting fairness and transparency for business users of online intermediation services), 이른바 P2B 투명성촉진법의 인공지능 일반 관련 조항

- 투명성 증대를 위한 정보공개 의무 부과 방식. 검색, 배열 순위(ranking)의 투명성 보장(제5조)
- 차별대우에 대한 설명의무(제7조)

- 최혜우대조항에 대한 설명, 공시의무(제10조)
- 생성형 AI의 등장 고려 전의 인공지능 기술(알고리즘의 매개변수 조정 또는 인위적 변수 개입에 따른 성과 조정이 어느 정도 가능한 상황)을 배경으로 함
- 「디지털시장법」(Digital Markets Act, 'DMA')의 인공지능 일반 관련 조항
  - 생성형 AI 서비스를 포함한 AI 서비스는 게이트키퍼로 지정될 수 있는 핵심 플랫폼 서비스(core platform service, 'CPS')의 유형에 명시적으로 포함되어 있지 않으므로, AI 서비스가 독자적인 플랫폼 서비스로 제공될 경우에는 DMA의 적용범위에 해당하지 않음(Yasar et. al, 2023)
  - 게이트키퍼로 지정된 사업자가 전유적인 또는 제3자의 AI 서비스를 자신의 CPS에서 통합할 경우 게이트키퍼 사업자의 의무 규정(제5조, 제6조, 제7조) 중 일부가 적용될 수 있음(Yasar et. al, 2023). 제5조는 그 자체로 집행력이 있는(self-executive) 규정인 반면, 제6조, 제7조는 그 자체로 집행력이 없고(non-self-executive) 유럽위원회가 그 의무를 구체화하는 이행조치를 채택할 수 있는 규정임(Bostoen, 2023)
  - 공개되지 않은 데이터 이용 금지 규정(제6조 제2항)은 AI 서비스를 통합한 게이트키퍼 사업자가 자신의 API에 의존하여 하류시장에서 애플리케이션을 공급하는 사업자의 데이터에 무임승차하는 것을 방지할 수 있음
  - 이용사업자에 대한 탈결합(unbundled) 접근 제공 의무 규정(제5조 제7항)은 AI 서비스를 통합한 게이트키퍼 사업자가 이용사업자에게 자신의 신원인증 서비스, 결제 서비스 기타 기술적 서비스 이용을 강제하는 것을 방지할 수 있음
  - 추가 등록 금지 규정(제5조 제8항)은 AI 서비스를 통합한 게이트키퍼 사업자가 이용사업자 또는 최종이용자에게 AI 서비스 이용 조건으로 다른 CPS 구독 또는 등록을 강제하는 것을 방지할 수 있음
- 전통적 AI(Conventional AI)와 생성형 AI(Generative AI)
  - 기존에 주로 논의되던 인공지능 기술은 분류형 인공지능(Classification AI) 또는 식별형 인공지능(Discriminative AI)이라고 부를 수 있는 것이었음
    - 분류 모델 또는 식별 모델은 지도학습(supervised learning) 기반으로 문서, 음성 또는 이미지 인식과 같은 분석 작업에 적합

- 생성형 인공지능(Generative AI)은 사용자의 입력을 기반으로 사용자가 원하는 결과를 유추해 텍스트, 이미지, 오디오 등 다양한 형태의 결과물을 만들어내는 AI 알고리즘을 말함(유재흥 외, 2023)
  - 생성형 AI는 비지도학습(unsupervised learning) 기반으로 문서, 음성 또는 이미 생성과 같은 창의적인 작업에 적합
  - 생성형 AI는 모델을 실행하기 어렵고, 컴퓨터 자원을 관리하고 병목 현상을 방지하기 위해 정교한 직무 조정이 필요하며, 클라우드 컴퓨팅 이용 등에 엄청난 비용이 수반되는 등 수년간 여러 요인의 제약이 존재하였음
  - 신기술의 등장, 더 많은 데이터, 더 저렴한 컴퓨팅 성능 등 이제는 쉽게 생성형 AI를 구축할 수 있는 환경이 마련되었음(김태원, 2023)
- GPT 모델과 챗GPT의 출시로 인한 변화
  - 2020년 대규모 언어 모델인 GPT-3 모델과 그 후속 모델의 발표 그리고 2022년에 이 모델을 미세조정(fine tuning)하여 사람이 읽을 수 있는 텍스트 형태의 입력을 가리키는 이용자의 명령형 프롬프트(prompt)에 적합한 응답을 생성하는 채팅형 인터페이스를 제공하는 챗봇 서비스인 챗GPT(ChatGPT) 출시
  - 생성형 AI가 사회적으로 큰 영향을 미치고 있으며 기존 산업도 향상된 자연어 인터페이스를 적용하여 빠르게 변화하고 있음(오장민, 2023)
- 생성형 AI 시장의 실태를 파악하고 기존의 경쟁 구도를 전제로 도입, 강화되고 있는 디지털 플랫폼 규제 및 경쟁정책을 재검토하고 발전시킬 필요성
  - 생성형 AI의 등장에 따라 기존 거대 플랫폼 외부에서 생성형 AI를 기반으로 한 새로운 인공지능 서비스와 가치 창출 체계가 형성되고 있음
  - 생성형 AI의 출현 및 발전에 따라 시장에서의 경쟁 구도와 생태계 구성의 변화가 예상됨
- 생성형 AI의 출현이 경쟁과 공정거래에 미치는 영향 분석의 필요성
  - 새로운 인공지능 서비스와 가치 창출 체계가 기존의 핵심 플랫폼 서비스를 대체하거나 독립적인 서비스와 별도 생태계로 발전할 가능성(생성형 AI가 시장 경쟁에 긍정적인 영향을 미친다는 가정)

- 새로운 인공지능 서비스와 가치 창출 체계가 기존의 핵심 플랫폼 서비스를 강화하여 소수의 거대 플랫폼 사업자에 의한 생태계 독과점 현상을 악화할 가능성(생성형 AI가 시장 경쟁에 부정적인 영향을 미친다는 가정)
- 생태계 간 경쟁 못지않게 생태계 내부에서 이루어지는 거래단계별 거래 과정에서 새로운 유형 또는 형태의 불공정거래행위가 발생할 가능성

## (2) 연구의 목적과 대상

□ 생성형 AI의 등장으로 인한 플랫폼 경제의 변화를 예측하여 기존의 플랫폼 규제 및 경쟁정책을 재검토하기 위한 계기를 마련하고, 생성형 AI의 등장에 따른 새로운 정책 개발 또는 정책 변화 모색을 위한 쟁점을 발굴하여 새로운 정책 개발 또는 정책 변화의 방향 제시

- 경제규제정책 및 경쟁정책이 연구 대상임
- 경쟁정책은 사후적인 규율 방식인 경쟁정책과 공정거래정책을 중심으로 함
- 경제규제정책은 경쟁정책을 보완하는 역할을 하는 사전·사후적인 규율의 혼성(hybrid) 방식인 거래공정화 규제정책, 그리고 사전적인 규율 방식인 사업활동(line of business) 규제정책만을 연구 대상으로 함

□ 구체적인 연구주제

- 디지털 플랫폼 또는 인공지능 일반과 구별되는 생성형 AI의 특징은 무엇인가?
- 생성형 AI의 특징은 디지털 플랫폼 또는 인공지능 일반과 구별되는 새로운 시장 리스크를 창출하거나 기존 리스크를 강화하는가?
- 디지털 플랫폼 또는 인공지능 일반의 특징에 의하여 제기되는 것으로 주장되는 시장 리스크로 인해 새로 도입되거나 강화된 경제규제 또는 경쟁정책의 현황은?
- 생성형 AI의 특징에 의하여 제기되는 것으로 주장되는 시장 리스크는 기존의 경제규제 또는 경쟁정책 도입 또는 강화의 근거를 보강하는가? 아니면 약화하는가?
- 생성형 AI의 특징에 의하여 제기되는 것으로 주장되는 시장 리스크에 대처하기 위해 경제규제 또는 경쟁정책을 새로 도입하거나 강화할 필요성이 있는가?

## 나. 생성형 AI의 등장 고려 전의 플랫폼 규제 및 경쟁정책 재검토

### (1) 인공지능 기술이 적용되는 상황에서의 경쟁법 쟁점

- 기계학습으로 대변되는 인공지능 기술이 시장의 영역에서 현실적인 변수로 등장함에 따라 시장에서의 규칙을 다루는 경쟁법에 제기되는 새로운 이슈(홍대식, 2019)
  - 관련되는 당사자의 유형에 따라 구매자와 관련된 이슈와 판매자와 관련된 이슈로 구분해 볼 수 있음
  - 발생하는 이슈의 성격에 따라 데이터와 관련된 것과 알고리즘과 관련된 것으로 구분해볼 수 있음
    - 시장에서 활용되는 인공지능 기술의 경우 투입 요소로서의 데이터와 산출물로서의 알고리즘이 중요한 구성부분이 됨
- 경쟁법에 영향을 미치는 인공지능 기술의 적용 상황의 구분
  - 데이터와 관련된 구매자 관련 이슈
  - 알고리즘과 관련된 구매자 관련 이슈
  - 데이터와 관련된 판매자 관련 이슈
  - 알고리즘과 관련된 판매자 관련 이슈

#### (가) 데이터와 관련된 구매자 관련 이슈

- 인공지능 기술이 적용되는 상품 또는 서비스를 둘러싼 경쟁에서 이를 제공하는 기업들 사이에 충분한 데이터를 어떻게 확보할 것인가 하는 문제
  - 디지털 플랫폼의 시대에 성공적인 기업들은 대체로 구매자 또는 이용자로부터 많은 데이터를 수집하고 있음
    - 과거에 이런 데이터는 주로 구매자 또는 이용자를 특정하여 거래 또는 상호작용을 원활하게 하기 위한 것에 국한되었음
    - 인공지능 기술의 발전으로 이용자의 위치 정보나 행태 정보로부터 상품 또는 서비스의 품질을 개선하고 거래 규모를 확대하는 데 필요한 수많은 가치를 창출할 수 있게 됨에 따라 수집, 이용 대상이 되는 데이터의 유형과 규모는 기하급수적으로 증가하고 있음

- 기업들이 데이터 수집, 이용에만 열을 올린 나머지 데이터 생성의 원천이 되는 이용자의 보호에는 소홀히 한다는 비판이 최근 자주 대두되고 있음
  - 특히 이용자 데이터에는 이용자를 식별하거나 식별할 수 있는 개인정보가 포함된 경우가 많은데, 개인정보는 이에 대해 개인이 스스로 결정할 수 있는 권리를 갖는 것을 전제로 하는 개인정보보호법의 보호 대상이 됨
  - 경쟁법의 영역에서도 기업이 데이터를 수집, 이용하는 과정에서 이용자의 이익을 부당하게 제한하는 거래조건을 설정하는 것이 문제될 수 있음

□ 독일 경쟁당국의 2017년 메타 사건

- 독일 경쟁당국이 2017년 메타(페이스북)가 이용자의 인터넷 이용 관련 데이터를 수집, 이용하는 행위가 독일 경쟁법에 위반된다는 혐의에 대한 조사에 착수하여 2019년 범위반을 인정한 사례가 있음(Bundeskartellamt press release, 2019)
- 이 사건에 대한 불복소송에서 독일 연방대법원은 2020년 메타의 부당한 개인정보 수집·이용행위가 데이터 착취·남용행위가 될 수 있다는 잠정적 판단을 전제로 하여 연방카르텔청의 처분에 대한 하급심법원의 집행정지 결정을 파기환송하였음(BGH Press release, 2020; 법무법인 지평, 2021; 이효석, 2022)
  - 독일 연방대법원은 메타의 행위가 「일반데이터보호법」(GDPR) 위반인지 여부와 관계없이 착취·남용행위라는 근거를 끌어냈는데, 그 근거가 이용자들의 진정한 선택권을 박탈하고 전부 또는 전무(all or nothing)의 양자택일을 요구했다는 점, 즉 소비자 선택 가능성을 침해하였다는 것임
- 이 사건은 파기환송을 받은 하급심법원이 유럽최고법원(CJEU)에 법률 해석에 대한 예비적 결정을 위탁함으로써 유럽최고법원의 법적 판단을 받게 되었음
  - 유럽최고법원은 2023년 회원국의 경쟁당국이 기업이 시장지배적 지위를 착취·남용할 경우 GDPR 위반행위를 조사 및 제재를 할 수 있고, 데이터보호법 위반 여부에 대하여는 해당 데이터보호당국과 협의하여야 한다고 판시하여, 경쟁법과 개인정보보호법이 중첩되는 영역에 대한 경쟁당국의 관할을 인정하였음

(나) 알고리즘과 관련된 구매자 관련 이슈

□ 알고리즘의 복잡성과 그로 인한 투명성의 부족으로 야기되는 문제

- 인공지능 알고리즘의 설계가 잘못되었을 경우 아무리 양질의 데이터셋에 의하여 추가적인 학습을 하더라도 구매자에게 제공되는 정보나 거래의 기회에 차별적인 결과가 발생할 수 있음
  - 이는 이러한 알고리즘에 의존하여 거래를 선택해야 하는 구매자에게 부정적인 영향을 줄 수 있음
  - 네비게이션 서비스에 적용되는 인공지능 알고리즘이 최적길 경로라고 알려주어 그대로 주행하였다가 오히려 길을 험해거나 쓸데없이 돌아가는 경험을 하게 되는 경우를 예로 들 수 있음
- 더욱이 기업이 알고리즘을 설계하는 과정에서 의도적으로 편향된 생각과 의견을 입력하는 경우 이를 규제할 수 있는 규정이나 법이 없다면 심각한 문제가 발생할 수 있음

□ 가격책정 알고리즘 적용에 따른 문제

- 알고리즘이 가격 책정에 적용되면 구매자별로 동태적이고도 개별화된 가격책정이 가능해짐
  - 이러한 가격책정 알고리즘은 구매자에 대한 데이터를 바탕으로 개인의 선호뿐만 아니라 특정 상품이나 서비스에 대한 개인의 지급의사의 변화를 점점 정확하게 추론할 수 있게 되기 때문임
  - 통상적인 경우 판매자가 구매자의 지급의사의 차이를 정확하게 알 수 없고 가격의 차이를 이용한 차익거래의 가능성이 있어 동일한 상품 또는 서비스의 가격을 개별적으로 책정하는 데 어려움이 있음
  - 인공지능 기술의 발전으로 판매자가 구매자의 구매 행태에 대한 데이터를 축적하여 이를 토대로 개별 구매자의 선호 구조를 분석한 후 맞춤형 상품 또는 서비스를 제공하기 위한 지속적인 노력을 하면서, 제시되는 가격 자체도 개별 구매자의 선호도나 지급의사를 포함한 다양한 요소를 고려하여 맞춤형 가격이 될 수 있음

- 이러한 맞춤형 가격을 설정하면 경제이론에 의할 때 구매자별로 개별화된 완벽한 가격차별이 될 수 있음
  - 개별 구매자별로 그 선호도나 지급의사에 일치하게 개별화된 가격이 책정되면 이론적으로 가격과 지급의사의 차이에 따른 소비자 잉여가 구매자에게 남지 않고 판매자에게 모두 귀속되는 결과가 될 수 있음

#### (다) 데이터와 관련된 판매자 관련 이슈

□ 경쟁관계에 있는 판매자 간에 누가 더 데이터에 대한 접근에서 우위에 있는가 하는 문제

- 출발선에서의 경쟁조건의 형평성 또는 평평한 운동장(level playing field)의 확보의 문제
  - 인공지능 기술의 적용 여부와 그 정도에 따라 기업 간의 경쟁 구도가 달라질 수 있는 상황에서, 어느 기업은 인공지능 기술의 중요한 투입 요소인 데이터에 자유자재로 접근할 수 있는 반면에, 다른 기업은 그렇지 못하다면, 경쟁의 출발선에서 현격한 차이가 생길 수밖에 없음
  - 디지털 플랫폼이 사용하는 양면시장형 사업모델은 플랫폼이 양 측면에 있는 이용자 그룹과 그 관심을 끌기 위한 기회를 엿보는 판매자 그룹을 동시에 무대 위에 올려 증개하는 역할을 핵심적인 구성요소로 함
  - 새로운 증개 서비스를 제공하는 시장을 얻기 위한 경쟁이 한참 진행되는 동안에는 다수의 기업이 주도권을 차지하기 위해 노력함. 그러나 네트워크 효과가 지배하는 시장의 특성 때문에, 일정한 기간이 지나면 시장의 선점에 성공한 기업이 시장의 전부 또는 대부분을 차지해버리고 나머지 기업들은 시장에서 사라지거나 주변부로 밀려나게 됨. 일반적으로는 경쟁은 다시 그 바깥에서 시장의 경계를 허물고 새로운 시장을 개척하는 것으로 계속됨
- 데이터가 주도하는 경제에서는 이른바 데이터 주도 네트워크 효과(data-driven network effect)라는 것이 발생한다는 이론이 유력하게 제기되고 있음(Ezrachi & Stucke, 2016)
  - 그에 의하면, 온라인 플랫폼 간의 경쟁에서 승리한 기업은 그와 동시에 데이터 확보에도 유리한 지위에 오르게 되고, 이렇게 확보한 데이터는 그 범위와 규모의 측면에서, 또한 이러한 데이터를 활용한 시행착오 또는 실행에 의한 학습

(learning by doing)의 기회의 측면에서 그 기업에 다른 기업들에 비해 경쟁 우위로 작용하게 된다고 함

- 그로 인해 한 시장에서 주도적인 기업은 그에 인접하여 새롭게 대두되는 시장에서 그 지배력을 계속 확장할 수 있게 되어 다른 기업들이 후속 기회를 얻는 것을 배제할 수 있는 능력도 있고 그럴 유인도 갖게 된다고 함

□ 프랑스와 독일의 경쟁당국이 2016년에 공동으로 발간한 보고서(Autorité de la Concurrence and Bundeskartellamt, 2016)

- 데이터 또는 빅데이터를 활용한 경쟁제한행위의 유형이 열거되어 있음
  - 이는 구체적인 행위의 묘사만 다를 뿐 기존에 경쟁법의 영역에 익숙하게 알려진 행위 유형과 크게 다르지 않음
  - 여기에는 데이터에 대한 접근을 거부하는 거래거절, 데이터 이용에 대한 차별적 취급, 제3자와의 배타적인 데이터 이용계약, 데이터셋의 판매와 데이터 분석 서비스의 제공을 연계하는 끼워팔기, 복수시장에서의 동일한 데이터셋 이용을 통한 지배력 전이 등이 있음

(라) 알고리즘과 관련된 판매자 관련 이슈

□ 검색 또는 순위결정 알고리즘에 의한 판매자 간 차별의 문제

- 온라인 활동을 하는 이용자의 대부분은 단순한 정보를 검색하기 위해서나 특정 상품이나 서비스를 구매하기 위한 정보를 얻기 위해서도 검색서비스를 많이 이용함
  - 이때 검색서비스 제공자가 이용자가 입력한 검색어에 대응하여 검색 결과를 제시하기 위하여 사용하는 검색 또는 순위결정 알고리즘에 자신의 계열회사 또는 자신이 지정하는 판매자나 광고 수수료를 많이 지급한 판매자를 우대하기 위한 변수를 인위적으로 집어넣어 이용자의 선호도나 관련성과 관계없이 그러한 판매자의 검색 순위를 높게 배치하게 한다면, 이는 판매자 간의 차별을 초래하게 될 것임
  - 검색 또는 순위결정 알고리즘에 의한 판매자 간 차별의 문제는 특히 검색서비스 제공자가 검색서비스에서의 경쟁 우위를 기반으로 하여 쇼핑, 부동산, 영화 등 다른 전문서비스 분야에서도 경쟁 우위를 확보하기 위하여 검색 또는 순위결정 알고리즘을 활용한다는 의문에서 비롯됨

### ● EU의 구글 쇼핑 검색 사건의 사례

- 구글은 자신이 사용하는 검색 알고리즘이 인공지능 기술을 적용한 것이므로 관련성에 기초한 자동화된 의사결정의 결과임을 강조하였음. 그런데 유럽에서 이용자가 구글의 검색서비스에 의하여 쇼핑 검색을 한 결과 구글 자신이 제공하는 전문서비스가 다른 경쟁 전문서비스에 비하여 검색결과에서 우대를 받고 있다는 문제 제기가 이어졌음
- EU의 경쟁당국인 유럽위원회는 2017년 구글의 쇼핑 검색에 의한 자사 서비스 우대의 결과가 경쟁법에 위반한다고 판단하고 제재하였음
- 그 중요한 근거는 자사 서비스 검색결과와 경쟁 서비스 검색결과에 적용된 순위결정 알고리즘이 다르다는 것이었음. 구글은 비교 쇼핑 서비스를 포함한 웹사이트의 순위를 결정하기 위하여 페이지랭크 알고리즘이라고 알려진 순위결정 알고리즘을 적용하면서 이용자 경험을 증대하기 위하여 다양한 조정 메커니즘을 활용하고 있음
- 유럽위원회에 의하면, 구글은 경쟁사의 비교 쇼핑 서비스에 대한 검색 순위를 결정하기 위해서는 판다라고 알려진 알고리즘을 추가로 적용하였고, 이것이 경쟁사의 검색 순위를 낮추는 데 이용된다는 것임
- 이 사건에 대한 불복소송에서 유럽일반법원은 2021년 10월 유럽위원회 결정의 결론을 유지하였음

### ● 우리나라의 네이버 쇼핑 검색 사건 등의 사례

- 우리나라 공정거래위원회가 조사, 제재한 네이버 쇼핑 검색 사건과 카카오모빌리티 사건도 이와 유사하게 알고리즘에 의한 판매자 간 차별의 문제가 쟁점이 되었음
- 네이버 쇼핑 사건에서 공정위는 2021년 네이버가 국내 온라인 비교쇼핑 서비스 시장의 시장지배적 사업자로서 소비자의 비교쇼핑 검색 시 노출순위를 결정하는 알고리즘을 조정, 변경하여 자신이 운영하는 오픈마켓 서비스인 스마트스토어에 입점한 업체를 우대함으로써 국내 오픈마켓 시장에서 경쟁을 제한한 행위를 한 것으로 인정함
- 서울고등법원은 2022년 네이버 쇼핑 부문 사건에 대한 행정소송의 제1심에서는 원고 패소 판결을 선고하였음

- 카카오모빌리티 사건에서 공정위는 2023년 카카오모빌리티가 국내 일반 앱 호출 시장의 시장지배적 사업자로서 카카오T 앱의 중형택시 알고리즘을 조정, 변경하여 자신의 자회사가 운영하는 카카오T 블루 가맹택시를 우대함으로써 국내 택시가맹서비스 시장에서 경쟁을 제한한 것으로 인정함

□ 이른바 알고리즘 담합 문제

- 경쟁자 간에 동일한 가격결정 알고리즘을 적용하고 그 알고리즘에서 추가로 학습하는 경쟁자의 가격책정 데이터가 시장에서 자유롭게 수집되는 경우의 문제
  - 인공지능 알고리즘의 적용으로 인하여 개별적인 가격책정이 가능해진다고 하더라도, 판매자가 서로 다른 알고리즘을 사용하여 경쟁자의 가격책정에 대응하여 다른 가격을 구매자에게 제시하는 경쟁이 활발하게 일어나면 알고리즘 관련 가격차별의 이슈는 해소될 수 있을 것임
  - 이런 경우 어느 한 판매자가 가격을 조정할 경우 그 데이터는 실시간으로 경쟁 판매자에게 전달되어 동일한 알고리즘의 적용을 통하여 미세한 시간 차이를 두고 가격은 동조화될 수 있음. 그로 인해 마치 경쟁자 간에 가격결정에 관한 합의를 한 것과 마찬가지로의 결과가 발생할 수 있으나, 그러한 결과는 아무런 상호적인 의사연결 없이 알고리즘의 작용으로 인하여 생긴 것이므로 이를 카르텔이라고 단정하기 어려움
  - 이러한 상황을 알고리즘 담합이나 디지털 카르텔이라고 부르면서, 이에 대한 경쟁법 대응을 촉구하는 논의가 최근에 활발하게 전개되고 있음
  - 특히 OECD에서는 2017년 6월 “알고리즘과 공모”라는 주제로 토론회를 개최하여 이와 관련된 쟁점을 집중적으로 논의하였음(OECD, 2017). 이와 관련하여 혹자는 알고리즘의 적용으로 인한 담합 효과 발생의 위협에 경쟁법이 효과적으로 대응하기 위하여 ‘합의’를 요구하는 기존의 법리에서 벗어나거나 규제를 할 수 있도록 합의에 해당하는 요건을 재구성할 필요가 있다는 주장도 나오고 있음

(2) 플랫폼에 대한 경제규제 및 경쟁정책

(가) 디지털 플랫폼과 인공지능(AI)에 대한 정책적 대응의 기초

- 현재 도입되었거나 도입이 논의되는 경제규제 및 경쟁정책은 디지털 플랫폼 서비스 제공자의 대형화에 따른 잠재적 리스크의 식별과 평가에 기초함
  - 디지털 플랫폼의 특징은 무엇인가? 과연 디지털 경제의 일반적인 특징과 비교할 때 플랫폼에 특유한 것인가?
  - 특유하더라도 이런 특징은 독립적인 또는 추가적인 리스크를 초래할 우려가 있는가?
  - 예상되는 잠재적 리스크에 대응하기 위한 경제규제정책은 무엇인가?
  - 어느 정도 현실화된 리스크에 더 잘 대응하기 위한 경쟁정책은 무엇인가?

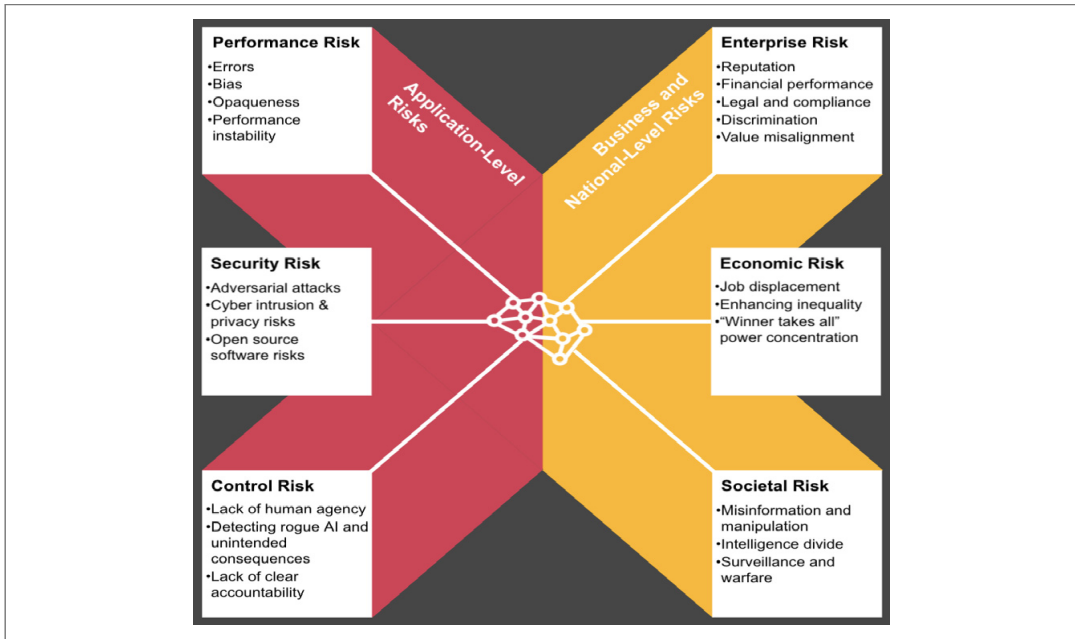
[그림 II-11] 대형 플랫폼 사업자의 리스크



- 경쟁 감소로 인한 소비자의 직접 피해와 이용사업자에 대한 부정적 영향을 통한 간접 피해로 구성

- 인공지능 기술이 적용되는 상황에서의 정책 논의도 대부분 대규모 디지털 플랫폼 서비스 제공자의 인공지능 기술의 적용이 독립적인 또는 추가적인 리스크를 초래할 우려가 있는가 하는 점이었음

[그림 II-12] 인공지능 기술 적용의 리스크



자료: Harvard Business Review (2021)

(나) 경제규제정책과 경쟁정책의 구별

- 디지털 플랫폼과 인공지능에 대한 경제정책적 대응이 논의될 수 있는 정책영역은 경제규제 및 경쟁정책의 영역임
  - 경제규제정책과 경쟁정책은 특정한 산업 또는 시장구조에 대한 다른 진단과 처방에 기초함
- 경제규제정책과 경쟁정책의 구별
  - 경제규제정책은 주로 전기통신사업법을 근거로 하여 이루어짐
    - 원칙적으로 사전규제를 통해 이루어지는 유효경쟁정책
    - 전기통신사업의 효율적인 경쟁체제의 구축과 공정한 경쟁환경의 조성을 위한 경쟁정책(전기통신사업법 제34조)

- 경쟁정책은 주로 공정거래법을 근거로 하여 이루어짐
  - 원칙적으로 사후규제를 통해 이루어지는 시장경쟁 정책과 공정거래정책
  - 입법과 정책에 대한 경쟁영향평가를 시행하는 경쟁주창(competition advocacy) 정책

(다) 경제규제정책과 경쟁정책의 중첩

□ 경제규제정책과 경쟁정책이 중첩될 수 있는 영역이 존재함

- 전문 경제규제당국과 경쟁당국의 시장조사 정책
- 전문 경제규제당국과 경쟁당국의 거래규제 정책

□ 전문 경제규제당국과 경쟁당국의 시장조사 정책

- 전문 경제규제당국의 시장조사 정책
  - 경쟁상황평가 정책(전기통신사업법 제34조: 기간통신사업. 전기통신사업의 효율적인 경쟁체제의 구축과 공정한 경쟁환경의 조성을 위한 경쟁정책을 수립하기 위하여)
  - 현황 실태조사 정책(동법 제34조의2: 부가통신사업. 부가통신사업의 현황 파악을 위하여)
- 경쟁당국의 시장조사 정책
  - 독과점 시장구조 개선 정책(공정거래법 제4조. 독과점적 시장구조가 장기간 유지되고 있는 상품이나 용역의 공급시장 또는 수요시장에 대하여 경쟁을 촉진하기 위하여)
  - 거래분야 실태조사 정책(동법 제87조. 일정한 거래분야의 공정한 거래질서 확립을 위하여)

□ 전문 경제규제당국과 경쟁당국의 거래규제 정책

- 전문 경제규제당국의 사업활동 제한(line of business restrictions) 규제 정책과 경쟁당국의 거래공정화 정책
- 둘 다 사전/사후규제의 하이브리드 형태이나, 전자는 사전규제에, 후자는 사후규제에 초점

(라) 디지털 플랫폼과 인공지능에 대한 정책적 대응

디지털 플랫폼에 대한 정책적 대응

● 경제규제정책

- 2022년 출범한 ‘범부처플랫폼 정책협의체’와 ‘플랫폼 민간 자율기구’

● 경쟁정책

- 공정위의 2023년 제정 ‘온라인 플랫폼 사업자의 시장지배적 지위 남용행위에 대한 심사지침’과 2023년 행정예고 ‘기업결합 심사기준’ 개정안

● 디지털 플랫폼은 서비스 또는 서비스 기반으로 출발했지만 새로운 가치와 수익원을 창출하는 사업모델로 정립되었는데, 현재의 정책적 대응은 디지털 플랫폼을 사업모델이 아니라 사업모델에 활용되는 서비스 또는 서비스 기반으로만 보는 경향

인공지능에 대한 정책적 대응

● 경제규제정책

- 위험 기반 규제를 도입한 ‘인공지능법안’(2023. 2. 국회 법안소위원회 통과)

● 경쟁정책은 특별히 이루어지고 있지 않음

다. 생성형 AI의 등장으로 인한 플랫폼 경제의 변화에 대한 예측

(1) 생성형 AI의 등장으로 인한 새로운 시장 현상

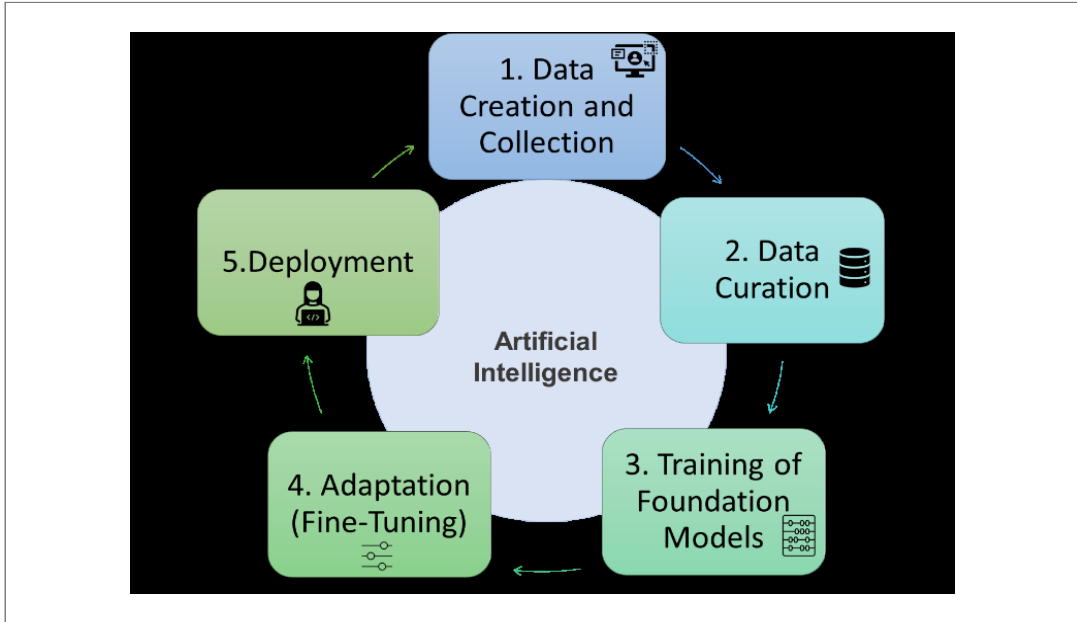
(가) 생성형 AI 시스템의 발전

AI 개발은 5가지 단계(AI 스택)와 관련됨

● 5가지 단계는 ① 데이터 생성 및 수집 단계, ② 데이터 큐레이션 단계, ③ 파운데이션 모델 훈련 단계, ④ 미세조정 단계, ⑤ 활용 단계임(Höppner & Streatfeild, 2023)

- 이와 같은 개발 단계에서 AI는 그 자체로는 상품도 독립적인 서비스도 아니지만, 일부 제공자(providers)에 의해 역동적으로 제공되고 다른 활용자(deployers)에 의해 다양한 활동에 배치되는 시스템의 역할을 함(Edward, 2022)

[그림 II-13] 개발 과정으로 본 AI 가치사슬(Value Chain)



자료: Höppner & Streatfeild(2023)

□ 파운데이션 모델의 운영 방식

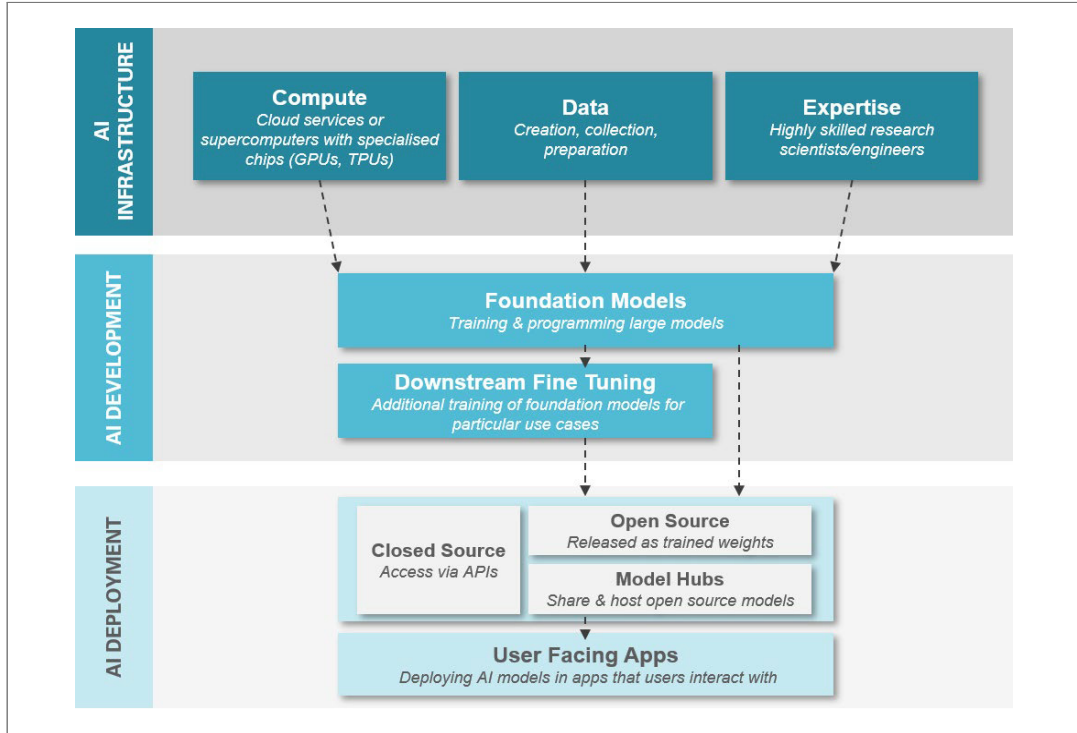
- 일부 대규모 기술 기업은 전체 AI 스택을 회사 내부에서 운영함
  - 그런 기업은 자신의 상품과 서비스를 개선하거나 AI를 동력으로 활용하는 최종 상품과 서비스를 시장에서 판매하기 위하여 폐쇄된 AI 시스템을 이용할 수 있음
  - 그러나 AI 개발은 강한 규모와 범위의 경제에 따르게 되므로, 더 큰 모델이 더 나은 결과를 가져옴
- 대부분의 AI 기업은 가치사슬의 특정한 단계에 초점을 맞추고 다른 단계는 외부에 개방함
  - 비용을 절감하고 다른 기업에 의한 조정과 개발을 증대할 목적으로 다른 사람이 이용할 수 있는 여러 오픈 소스 파운데이션 모델이 있음
  - 그러나 일부 AI 도구와 서비스는 대가를 받고 공급됨
- 파운데이션 모델의 운영 방식에는 다음과 같은 유형이 있음

- 다른 파운데이션 모델 개발자가 자유롭게 공유하여 특화된 파운데이션 모델을 만들거나 이용방식에 대한 제한이 없거나 상대적으로 적은 방식으로 공개하는 개방형, 오픈 소스 방식
- API를 통해 외부에서 제어되는 파운데이션 모델에 접근하는 등 파운데이션 모델 제공사업자가 제3자 개발자의 파운데이션 모델 접근 정도를 제어하는 폐쇄형 방식
- 파운데이션 모델을 외부에 전혀 공개하지 않고, 직접 파운데이션 모델 기반 이용자 애플리케이션을 제공하는 전유형 방식
- 개방형, 오픈 소스 방식과 폐쇄형 방식의 존재로 인해 파운데이션 모델 제공 서비스를 중심으로 하는 생성형 AI 생태계가 형성되고 있음

□ 파운데이션 모델 개발 및 제공 서비스를 중심으로 한 생태계 구성

- 산업이 아직 초기 단계에 있으나, 파운데이션 모델 개발 및 제공 서비스를 중심으로 할 때 생태계는 상류시장에 있는 컴퓨팅 자원, 데이터 및 전문가 제공 서비스, 하류시장에 있는 활용 서비스로 구성됨
- 생성형 AI 생태계는 여러 단계로 구성되는데, 우선 AI를 구동하기 위한 기초적 토양으로서 하드웨어(Hardware)와 클라우드 등 인프라(Infra)가 있고, 다음 단계로 생성형 AI의 본체인 파운데이션 모델(Foundation model)이 있으며, 마지막 단계로는 이 파운데이션 모델을 활용해 만들어진 구체적인 서비스로 애플리케이션(Application) 단계가 있음
- 파운데이션 모델 제공 사업자가 폐쇄형인 경우에는 파운데이션 모델 단계가 파운데이션 모델을 제공하는 단계로만 구성되지만, 오픈 소스형 또는 모델 허브형(Model Hub)인 경우 파운데이션 모델 단계는 파운데이션 모델을 제공하는 단계와 파운데이션 모델을 특정한 이용 목적을 위해 추가학습하여 특화된 AI 모델을 개발하는 단계로 구분될 수 있음
- 폐쇄형 모델로는 오픈AI의 GPT-3와 그 후속 모델, 구글의 BERT와 LaMDA가 있고, 오픈소스 모델로는 스테이빌리티(Stability)의 스테이블 디퓨전(Stable Diffusion), 메타의 OPT가 있으며, 모델 허브 모델로는 허깅 페이스(Hugging Face)와 레플리케이트(Replicate)가 있음

[그림 II-14] 생태계 구성으로 본 AI 가치사슬(AI 스택)

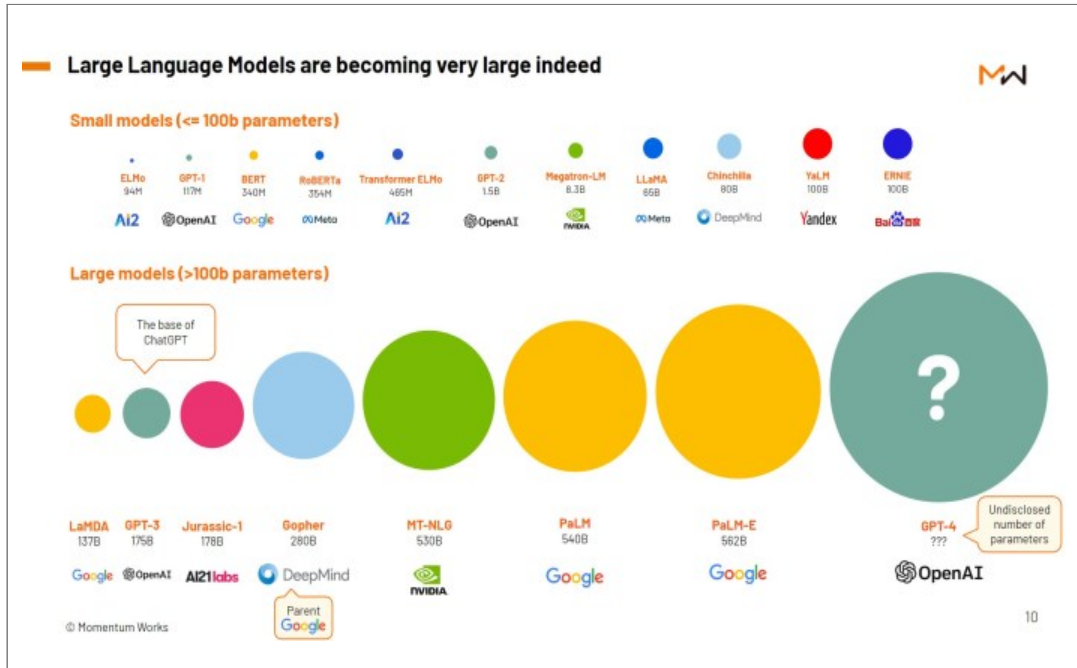


자료: CMA(2023a)

- 상류시장에 있는 컴퓨팅 자원, 데이터 및 전문가 제공 서비스 시장
  - 모델 훈련과 추론 작업량(workload)에 최적화된 가속 칩(Chips)과 클라우드 컴퓨팅과 같은 컴퓨팅 자원 시장과 데이터 생성, 수집, 거래 시장으로 구분될 수 있음
  - AI와 관련된 거의 모든 것은 컴퓨터 능력 수행에 의존하고 일정한 지점에서 클라우드 서비스(클라우드 서비스란 기업이 서버, 스토리지, 네트워크 등 IT 자원을 직접 소유·관리하는 대신, 외부 전문업체로부터 필요한 만큼을 서비스 형태로 공급받는 것을 의미하고, 서비스 유형에 따라 IaaS(IT 인프라 제공), PaaS(개발 플랫폼 제공), SaaS(소프트웨어 제공)로 구분됨)를 통해 전달되므로, 컴퓨팅 자원은 AI의 혈관과 같음
  - 이 시장의 대표적인 플레이어는 가속 칩인 GPU를 생산하는 엔비디아(NVidia), TPU를 생산하는 구글 등과 클라우드 서비스인 AWS를 제공하는 아마존, 애저(Azure)를 제공하는 마이크로소프트, GVP를 제공하는 구글이 있음

- 가속 칩 분야에서 엔비디아는 AI에 사용되는 GPU 시장의 95%를, 반도체 시장의 80%를 점유하고 있음
  - 인프라 단계에서는 아마존, 마이크로소프트, 구글 등 미국의 3대 클라우드 회사가 60% 이상을 점유하고 있음
  - 우리나라의 경우 공정위가 2022년 12월 발표한 클라우드 서비스 분야 실태조사 결과에 따르면, 최근 3년간(2019~2021) 국내 클라우드 시장점유율 현황은 1위 사업자인 아마존(AWS)이 70% 내외를 차지하여 시장이 상당 부분 집중되어 있고, 국내시장 2위에 해당되는 마이크로소프트, 3~4위에 해당되는 구글, 네이버 점유율이 점차 확대되고 있는 것으로 나타났음(공정위, 2022)
  - 데이터 분야에서 미래의 모델 성공은 관련 데이터에 대한 실시간 접근에 의존한다는 점에서 사업 데이터와 개인 데이터는 가장 관련성 있는 자원이 되고 있음(Bommasani et. al., 2021)
  - 이 시장에서는 대부분의 데이터를 수집하고 축적하는 인터넷에 연결된 기기와 서비스를 통제하는 사업자가 AI 애플리케이션 분야에서도 지도적 위치에 있게 됨. 현재까지 GPT를 훈련하는 대부분의 데이터셋은 웹에서 크롤링하는 콘텐츠인데, 구글이나 마이크로소프트와 같이 전체 앱을 계속하여 크롤링하고 색인하는 일반 검색엔진을 운영하는 사업자는 어떤 생성형 AI에 대하여도 가장 풍부한 데이터 원천이 됨
- 파운데이션 모델 개발 및 제공 서비스 시장
- 파운데이션 모델은 AI 시스템의 한 구성요소이지만, 파운데이션 모델은 창발적 능력(emergent capabilities)을 나타내면서 빠르게 실제 세계의 AI 시스템의 활용에 통합되어 사람들에게 훨씬 더 폭넓은 결과를 가져온다는 사실을 증시하여야 함(Bommasani et. al., 2021)
  - 파운데이션 모델은 이를 중심으로 한 AI 생태계를 형성하고 있음
  - 2020년 GPT-3의 출현 이후 2022년을 기준으로 개발된 11개의 민간 대규모 파운데이션 모델 중 8개가 빅테크가 운용하는 것이었음(구글과 그 자회사인 딥마인드가 개발한 것이 5개, 메타, 마이크로소프트, 바이두가 개발한 것이 각 1개씩).

[그림 II-15] 글로벌 대규모 언어 모델의 현황



자료: <https://thelowdown.momentum.asia/the-emergence-of-large-language-models-llms/>

- 우리나라에서는 글로벌 빅테크 기업에 대항하여 네이버, LG AI 연구원 등에서 연구개발이 진행되고 있으며, AI 분야 스타트업도 경량 모델을 증점적으로 개발하고 있음(유재홍 외 3, 2023)
- 대규모 자본과 인프라가 요구되는 초거대 언어 모델은 네이버, LG 등 대기업이 중심이 되어 개발되고 있음
- LG AI 연구원은 2021년 12월 엑사원을 공개한 후 2023년 7월 엑사원2를 공개한 바 있음
- 네이버클라우드는 자체 초거대 AI 모델인 ‘하이퍼클로버X’를 공개하였으며, 초거대 AI 하이퍼클로버X, 대화형 AI 서비스 ‘클로버X’, 생성형 AI 검색 ‘큐:’ 등을 소개하고 향후 자사 검색, 쇼핑, 광고 등 주요 서비스의 생성형 AI를 접목할 계획임
- 통신사인 KT도 2023년 10월 초거대 AI 플랫폼 믿음(Mi:dm)을 출시하였으며, 70억 개의 파라미터를 갖는 모델은 연구개발용으로 무료로 활용할 수 있음

- AI 활용 단계에서는 데스크탑 앱, 모바일 앱, 플러그인, 브라우저 확장자(browser extension) 등 다양한 형태로 AI 상품이 출시되고 있음
  - 이러한 사례로 마케팅/비즈니스에 특화된 문자 생성형 AI 서비스를 제공하는 재스퍼(Jasper), 콘텐츠 및 미디어 크리에이터에 특화된 이미지 및 동영상 생성 툴인 런웨이(runway) 등이 있는데, 이는 전유되지 않은 파운데이션 모델을 활용하여 개발된 최종 이용자용 생성형 AI 상품에 해당함
  - 여기에는 파운데이션 모델 개발자가 아닌 제3자가 개발한 것도 있지만, 파운데이션 모델 개발자는 모델을 API를 통해 라이선스하는 한편 자신이 직접 생성형 AI 상품을 제공하기도 함
  - GPT-3 모델과 그 후속 모델을 개발한 오픈AI가 제공하는 생성형 AI 상품으로는 텍스트를 이미지로 자동 전환하는 DALL-E 2(2022. 4.), 소프트웨어 자동 코드 생성 AI 서비스인 Copilot(2021. 6.), 작문요약 등 언어 생성 AI 서비스인 챗GPT(2022. 11.) 등이 있음(안성원 외 4, 2023)
  - 이런 서비스는 그 자체가 하나의 생성형 AI 상품이지만, 다른 생성형 AI 상품 개발을 위한 도구로도 활용될 수 있음
  - 전유적(proprietary) 파운데이션 모델 개발자는 외부에 모델을 배포하지 않으므로 이를 활용하는 생성형 AI 상품을 내부적으로 개발하여 제공하는데, 이런 사례로는 텍스트를 이미지로 자동 전환하는 미드저니(Midjourney)가 있음
  - 우리나라에서는 뤼튼테크놀로지스, 업스테이지 등의 스타트업이 네이버 하이퍼클로바, 오픈AI의 챗GPT, 오픈소스 LLM 등을 자체적으로 활용해 시험 서비스를 개발하거나 상용화를 추진하고 있으며, AI 플러그인 생태계 중심의 애플리케이션 시장을 형성하고 있음
  - 업스테이지의 AskUp은 OCR 기술을 결합해 개발한 카카오톡 AI 챗봇으로 사용자가 문서를 찍거나 전송하면 내용을 읽고 이해하고 답변을 제공하고 있다. 뤼튼테크놀로지스의 뤼튼은 SNS 광고, Push 알림, 보도자료 등 상황에 맞는 문장 생성 및 입력한 글에 대해 적절한 피드백과 참고자료를 제공함

## (2) 생성형 AI가 시장 경쟁에 미치는 영향

### (가) 긍정론

□ 생성형 AI가 기존 거대 플랫폼 외부에서 이를 기반으로 한 새로운 서비스의 가치 창출 체계를 형성하면서 이러한 새로운 서비스와 가치 창출 체계가 기존의 핵심 플랫폼 서비스를 대체하거나 독립적인 서비스와 별도 생태계로 발전할 가능성에 주목하는 견해임

● 오픈AI의 챗GPT 출시 후 빅테크인 구글의 대응은 새로운 시장을 차지하기 위한 경쟁의 양상을 보여줌(Schrepel, 2023)

- 구글은 AI 챗봇의 성장에 대응하기 위해 자체적인 초거대 AI의 개발에 박차를 가하게 될 수밖에 없는데 그로 인해 이용자가 검색엔진에서 AI 챗봇을 전환할 경우 기존 수익원의 일부가 잠재적으로 잠식될 수 있음

- 챗GPT 서비스를 출시한 사업자가 기존 빅테크가 아닌 오픈AI라는 것은 새로운 경쟁자의 출현이라는 점에서 의미가 있으나, 오픈AI가 기존 빅테크 중 하나인 마이크로소프트와 배타적 거래관계에 있다는 점은 경쟁 이슈가 될 수 있음

- 마이크로소프트와 오픈AI의 제휴 서비스가 구글이 주도하는 일반검색 서비스 시장의 외연을 확장하면서 경쟁을 촉진하는 것은 긍정적인 측면이지만, 챗GPT의 기술적 기반인 GPT 언어 모델과 그 훈련에 사용된 데이터에 대한 접근성이 줄어드는 것은 부정적인 측면임

● 이런 견해에서는 특히 오픈 소스 AI 시스템의 역할을 높게 평가함

- 주요 인공지능 도구는 모두 오픈 소스로 이루어지고 있음. 개발자는 모델 개발의 목적에 맞게 오픈 소스의 데이터를 사용하여 모델을 사용, 훈련, 수정, 재배포할 수 있고, 최신 소프트웨어 개발을 뒷받침하는 오픈 소스 라이브러리 및 패키지 생태계와 유사한 AI 개발 프로세스를 구축할 수 있음

- 이처럼 제한 없이 자유롭게 접근·사용·변경이 가능한 소프트웨어인 오픈 소스 소프트웨어(OSS)는 인공지능의 개발과 활용에 있어서 핵심적인 역할을 담당하고 있음

- Python, R, C++, Java, Scala, Javascript, Julia 등과 같은 오픈 소스 프로그래밍 언어에는 수천 개의 기계학습 알고리즘이 구현되어 있음. OSS는 AI를 사용하는 데 필요한 수학적, 기술적 지식수준의 격차를 줄여 AI 채택을 활성화하고 증가시킬 수 있음

- 오픈 소스 모델은 보다 비용효율적인 점, 고객 맞춤화가 가능한 점, 더 적은 이용 제한이 있는 점 등의 다양한 장점으로 인해 폐쇄형 모델에 대한 경쟁상 제약이 될 수 있다는 주장이 있음(Semianalysis, 2023)
- 많은 개발자가 최신 AI 도구를 사용할 수 있도록 하는 경우 더 많은 개발자가 유용한 코드 작성에 대한 접근성이 높아지고 이는 모델의 훈련이 가속화됨을 의미할 수 있음
  - 개방적 AI 기술환경은 오픈 소스의 품질을 향상하는 결과를 가져오고 이는 궁극적으로 기술 비용을 절감시키고 개발자들의 참여 기회를 넓혀줄 수 있음
  - 오픈 소스 AI 시스템은 AI 모델 및 서비스 개발에 대한 숙련된 인재들 간의 개방적인 협업 환경을 조성할 수 있음. 예를 들어, 공개된 OSS 코드의 버그에 대한 개선 가능성이 있는지 교차적으로 조사하고 피드백을 받을 수 있으며. 인기 있는 OSS에 대한 공동 개발과 참여 커뮤니티를 활성화하여 공동의 연구를 가능하도록 하여 고품질의 코드를 개발할 수 있음
  - 구글과 메타의 경우 심층학습 하위분야에 특화된 도구(각각 Tensorflow 및 PyTorch)를 오픈 소스화하고 있는데, 이는 커뮤니티 참여 및 영향력 분산과 관련이 있는 경우가 많음
- 생성형 AI 개발은 대중에게 지식을 널리 보급하고 컴퓨팅 자원이 더 저렴해지는 것을 가능하게 하여 AI 모델에 대한 시장지배력을 다소간 약화할 수 있음
  - 이를 위해서는 상당한 재정적인 자원과 방대한 규모의 데이터 소스, 독립적인 엔지니어링 인재 풀이 요구됨
  - AI에 대한 성급한 규제는 기존 시장에서 존재하는 불평등을 악화시키고 강화할 수 있음
    - EU와 같이 AI를 규제하려는 새로운 법안, 기존 데이터 보호에 대한 새로운 규제 적용, 개인정보보호, 지식재산권과 같은 기타 법률과의 관계 등과 같은 새로운 법적, 규제적 조치에 대한 사업자의 부담이 증가할 수 있음
    - 관할권 간 AI 규제의 불일치와 AI에 대한 법적 프레임워크가 관할권 전반에 걸쳐 일관성이 없는 경우 특정 관할권에서 상품 또는 서비스를 조정하기 위해 자원을 지출해야 할 수 있어 규제 부담이 커지게 됨

## (나) 부정론

- 생성형 AI가 강한 규모와 범위의 경제에 따르고 파운데이션 모델을 중심으로 자기 강화적인 성장 사이클을 가진 근본적인 집중화 성격을 보여 모델과 애플리케이션 개발 장벽을 형성할 우려가 있다고 보는 견해
  - AI 시스템의 필수 구성 요소에는 데이터, 컴퓨팅 자원 및 엔지니어링 인재가 포함됨
  - 데이터로 인해 새로운 AI 시스템의 개발을 저해할 가능성이 존재하는 진입장벽이 세워질 수 있다고 보는 견해
    - AI 시스템은 데이터가 입력되어 학습할 수 있으므로, AI 시스템에 공급되는 데이터의 양과 품질은 학습 프로세스를 개선하는 데 직접적인 역할을 하게 됨
    - 수년에 걸쳐 크고 다양한 데이터셋을 수집한 기업은 AI 훈련에 쉽게 사용할 수 있는 방대한 데이터셋을 갖고 있어 AI 상품을 개발하고 배포하는 데 큰 이점을 가질 수 있음
    - 이는 혁신의 노력에 도움이 되고, 고객을 위한 우수한 AI 상품을 창출할 수 있지만, 새로운 AI 시스템의 개발을 저해할 가능성이 존재하는 진입장벽이 세워질 수 있음(Kollnig & Li, 2023)
  - 클라우드에 접근할 수 없으면 AI 개발과 새로운 AI 사업자의 출현이 저해될 것이라는 우려
    - 컴퓨팅 자원은 AI에서도 중요한 입력으로 AI 시스템의 개발 및 실행은 처리 능력에 크게 의존하고 있음. 서버 칩에 대한 수요가 증가하는 반면 공급량은 점점 낮아지고 있어 일부 AI 회사는 다양한 유형의 칩이나 기존 기술을 사용하는 등의 해결 방법을 모색하고 있음
    - 컴퓨팅 자원과 밀접하게 연결된 또 다른 요소는 클라우드 컴퓨팅이다. 클라우드에 접근할 수 없으면 AI 개발과 새로운 AI 사업자의 출현이 저해될 것이라는 우려가 존재함(Netherlands Authority for Consumers and Markets, 2022; FTC, 2023; CMA, 2023)
    - 이러한 문제를 해결하기 위해 EU 수준에서 취하고 있는 한 가지 접근방식은 유럽위원회가 제안한 사전규제를 채택하는 것임. 예를 들어, 유럽위원회의 데이터 법 초안이 채택되면 클라우드 제공업체가 EU에서 서비스를 운영하고

가격을 책정하는 방식이 바뀔 수 있으며, 궁극적으로 제공업체는 클라우드 외부로 데이터를 이동하는 것과 관련된 수수료를 제거해야 함

- 엔지니어링 인재 부족으로 인한 경쟁법 위반의 위험
  - 회원국 경쟁당국은 엔지니어링 인재 부족이 비경쟁 조항으로 인해 근로자들의 자유로운 노동을 방해하는 효과를 가져올 수 있다는 점에 우려를 표한 바 있음 (Romanian Competition Authority, 2022)
  - 대규모 인재 풀을 보유한 기업은 직원의 직업선택의 자유를 제한하거나 경쟁법을 위반할 위험이 있음
  - 미국의 경쟁당국인 연방거래위원회(FTC)는 경쟁법이 AI의 구성요소뿐만 아니라 AI 자체 구축 프로세스에도 영향을 미칠 것으로 보면서, 오픈소스 모델이 AI 접근성 확대에 앞장설 수 있으나, 처음에 오픈소스를 먼저 공개하고 나중에 폐쇄하는 전략을 통해 오용 가능성이 있다고 지적한 바 있음(FTC, 2023)

## 라. 새로운 정책 개발 또는 정책 변화 모색을 위한 쟁점 발굴

### (1) 생성형 AI의 특성이 정책에 주는 시사점

#### (가) 전통적인 AI와 생성형 AI의 비교

##### □ 인공지능 시스템의 정의 방식

- 기존의 정의 방식
  - EU AI Act 2021년 제안
  - ‘artificial intelligence system’ (AI system) means software that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with.
- 생성형 AI의 특징을 반영한 새로운 정의 방식
  - OECD 이사회 2023년 제안(OECD, 2023)
  - An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs

such as predictions, content, recommendations, or decisions that [can] influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.

- AI 시스템이 스스로 목적을 설정하는 시나리오를 포함하기 위하여 인간이 설정한 목적이라는 요소를 삭제

#### □ 전통적인 AI와 생성형 AI의 특징 비교(Marr, 2023; Vadapalli, 2023)

##### ● 전통적인 AI의 특징

- 규칙 기반 시스템
- 직무 특화
- 지도학습
- 예측 분석
- 한정된 창조성

##### ● 생성형 AI의 특징

- 생성 모델
- 다양한 응용
- 비지도 또는 자기지도학습
- 혁신적인 창조성
- 새로운 데이터에 대한 적응성
- 예측 가능성이 적음

##### ● 전통적인 AI와 생성형 AI의 공통점

- 데이터에서 학습
- 모델 복잡성
- 실제 문제への 응용
- 전통적인 AI와 생성형 AI는 상호 배타적이지 않음. 예를 들어 전통적인 AI가 이용자 행태 데이터를 분석하면 생성형 AI은 이 분석을 이용하여 개인화된 콘텐츠를 생성할 수 있음

(나) 인공지능 기술의 특성이 정책에 주는 시사점

□ 전통적 AI 기술의 특성(EU, 2020a)

- 내부 메커니즘의 불투명성(opacity)
  - 블랙박스 효과(black-box effect): 작동 중 스스로 학습해 나가면서 새로운 상황에서 사전에 결정되지 않은 판단을 내릴 수 있는 특성
  - 인공지능 최초 설계 당시 예방하거나 예상할 수 없었던 문제 상황을 발생시킬 리스크(EU, 2020b)
- 기술의 상시적 연결성(connectivity)
- 인공지능이 내리는 결정의 자율성(autonomy)
- 데이터 의존성(data dependency)
- 제품·서비스·공급망의 복잡성(complexity)
- 생성형 AI의 발전으로 이런 특성이 강화될 수 있지만, 생성형 AI 기술에 특유한 리스크로 볼 수 있는지가 쟁점임

(2) 해외의 정책동향

(가) OECD

□ 2023년 9월 '범용 인공지능(Generative Artificial Intelligence)에 대한 최초의 정책 고려사항' 보고서(OECD, 2020)

- 보고서의 의의
  - 이 보고서는 생성형 AI의 발전이 우리 환경과 어떠한 방식으로 상호작용하여 형성되어야 하는지에 대한 디딤돌 역할을 제공하여 정부 및 입법자가 급속한 기술변화를 따라 생성형 인공지능 시스템의 단점을 완화하기 위한 다양한 전략 수립을 돕기 위하여 작성되었음
  - 훈련 데이터를 기반으로 사용자 프롬프트에 응답하여 콘텐츠를 생성하는 생성형 AI 시스템에 대한 소개와 정책입안자가 해결해야 할 긍정적 이점과 부정적 인 측면을 동시에 소개

- 생성형 AI가 야기할 수 있는 중요한 사회적, 정책적 리스크의 식별
  - 지적 재산과 관련된 우려: 비인가 콘텐츠 생성
  - 악의적인 목적을 위한 오용: 페이크 콘텐츠 생성
  - 시장 뒤엎기(disruption): 직업 대체
  - 윤리적 및 편향 우려: 편향적 콘텐츠 생성
  - 규제준수: 규제 부재로 인한 불확실성
  - 보안 위험: 데이터 보안
  - 품질 및 신뢰성: 생산물 품질 보증
  - 경제적 격차: 접근성 및 가용성(affordability)
- 생성형 AI 시장 발전에 대한 전망
  - 투자은행, 컨설팅 회사와 연구자들은 생성형 AI가 다가오는 장래에 막대한 경제적 영향을 줄 것으로 전망
  - 현재로서는 생성형 AI가 초기 발전 단계에 있어 다음의 성숙 단계로 가기 위해 막대한 R&D 투자와 숙련되지만 희소한 전문인력을 필요로 함
  - 생성형 AI가 적용되는 영역에는 칩과 부속 설계, 재료과학과 엔터테인먼트 분야가 포함됨
  - 이러한 분야에서 경쟁우위(competitive edge)가 있는 기업에는 잘 알려진 대규모 기술 기업, 기업용 소프트웨어 제공자, 주변 분야에서의 AI 기업, 반도체와 컴퓨팅 자원 인프라를 제공하는 기업이 있음
- 장래의 리스크 완화를 위한 보다 크고 체계적인 규모의 해결책에 대한 요청
  - 규제
  - 윤리적 프레임워크
  - 기술적인 AI 표준화
  - 감사
  - 모델 출시
  - 접근성 전략

(나) 영국

□ 과학혁신기술부의 백서(White Paper)(Department for Science, Innovation and Technology, 2023)

● 백서의 의의

- 개인정보보호와 안정성의 침해 없이 혁신에 초점을 맞춘 상식적인 방식에 초점을 맞추어 마련되었음
- 2022년에 당시에 인공지능정책을 담당하던 디지털문화미디어스포츠부(Department for Digital, Culture, Media and Sport)에서 발간한 정책보고서인 '친혁신(Pro-innovation) 인공지능 규제접근 수립'에서 제안된 인공지능 규제 프레임워크 수립을 위한 이행원칙과 이행방안을 구체화하는 내용을 담고 있음(한국지능정보사회진흥원, 2022)

● 규제의 목적과 방식에 대한 새로운 제안

- 인공지능 규제 프레임워크의 목적과 주요 특징은 인공지능 주도국가로서의 입지를 강화하기 위해 혁신을 도모하고 규제 불확실성을 감소시켜 인공지능에 대한 신뢰성 향상에 있음
- 이러한 규제 목적 달성을 위하여 인공지능의 특징을 '적응성'과 '자율성'이라고 정의하여 빠르게 변화하는 최첨단 기술 기반한 행위의 규제를 위한 공통적인 이해를 개념화하였음
- 맥락에 기반한 유연한(context-based and flexible) 규제: 인공지능에 기반한 규제는 모든 분야 또는 기술에 동일한 위험기준 또는 규칙을 적용하는 것이 아니라 인공지능이 구체적으로 이용되어 발생하는 결과에 기반하여 규율하도록 함
- '안전, 보안, 견고성, 투명성, 설명가능성, 공정성, 책임, 거버넌스, 이의제기, 규제' 등의 5개의 '이행원칙'에 따라 모든 규제기관들은 소관 분야에 적합한 구체적인 이행원칙을 개발, 적용이 가능하도록 제시함
- 인공지능 규제 프레임워크 이행의 전문성과 조직적인 능력을 확보하고 일관성 있는 '규제 이행'을 보장하기 위하여 동 백서는 프레임워크 '이행지원'에 필요한 정부의 핵심 기능을 구체적으로 설명하고 있음
- 범용 목적의 인공지능으로서 사회·경제 전반에 상당한 영향력을 보여주고 있는 파운데이션 모델에 주안점을 둘 필요성을 강조함(채은선, 2023)

## □ 경쟁시장청의 경쟁과 소비자 보호에 관한 최초 보고서(CMA, 2023)

## ● 보고서의 의의

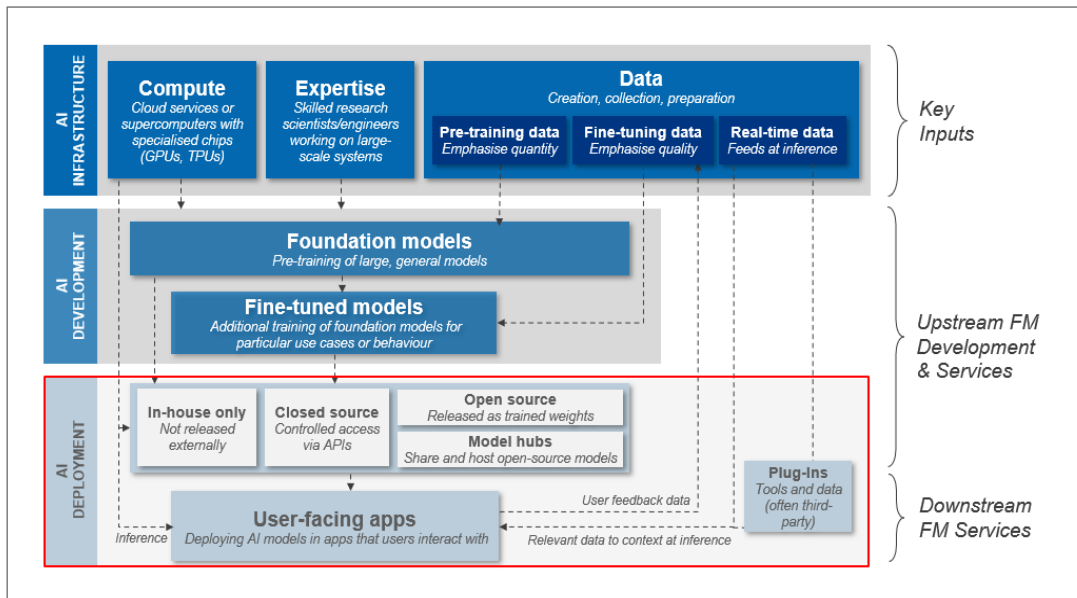
- CMA는 친혁신적인 AI 규제 프레임워크와 소비자 이용의 안정성 확보를 위하여 파운데이션 모델에 대한 개념 정의와 설계·개발 단계에 대한 설명을 다루면서 이러한 파운데이션 모델의 개발이 경쟁과 소비자 보호에 미치는 영향을 상세하게 분석함
- CMA는 파운데이션 모델이 다른 시장에 미치는 경쟁 및 소비자 보호 문제에 대하여 고려하기 위하여 규제 접근방식을 제시하면서 개방적, 경쟁적 시장 환경을 조성하기 위한 규제원칙을 명확하게 제안함

## ● 파운데이션 모델(FM)에 대한 주요 입력 및 AI 스타트업과의 수직적 관계가 출현하는 AI 시장 경쟁에 미치는 효과와 관련된 우려 식별

- 파운데이션 모델 개발 방법, 파운데이션 모델에 필요한 주요 입력 및 현재 배포 방법에 대한 CMA의 초기 견해 및 이러한 견해의 기초가 되는 수집한 뒷받침 증거를 제시하면서 파운데이션 모델 개발 경쟁의 잠재적 결과, 다른 시장에서의 경쟁과 경쟁의 잠재적 결과에 대한 파운데이션 모델의 영향, 소비자를 위한 잠재적 결과, 긍정적인 발전과 결과를 가능하게 하는 규제의 잠재적인 역할을 다루고 있음
- 특히, 파운데이션 모델 개발 경쟁의 잠재적 결과와 파운데이션 모델이 다른 시장 경쟁에 미치는 영향과 관련하여 AI 파운데이션 모델 환경은 초기 단계에 있으며, 이미 다양한 액세스 수준을 통해 광범위한 파운데이션 모델을 사용할 수 있다고 보았음
- ‘상류 FM 개발 및 공급시장’에서 FM 개발을 위한 주요 입력 사항과 배포에 사용되는 방법이 AI 공급망의 수직적인 통합관계 간에서 이루어질 경우 이것이 전환시장 간에 통합되어 경쟁이 아주 약하게 작용하여 승자독식의 결과로 이어질 수 있음을 설명하고 있음
- ‘하류 FM 서비스 시장’은 FM 배포와 관련된 시장으로 볼 수 있는데, FM 개발과 배포는 밀접하게 연결되어 있으며, 가치사슬에서 한 부분에서의 개발은 다른 부분의 경쟁에 영향을 미칠 수 있다고 봄. 이에 따라 하류시장 수준에서 경쟁 평가는 상류 FM 개발 및 공급과도 밀접하게 연관됨. 결국 상류 FM 시장의 수직적 통합은 FM 하류 서비스 시장 경쟁에 잠재적인 위험을 야기할 수 있다 봄

- 기업이 하나 이상의 시장에서 시장지배력을 갖고 그 시장지배력을 사용하여 하위 수준에서 경쟁을 약화하는 경우 경쟁 위험이 발생할 수 있다는 점을 지적하고 있음
- ‘입력’에 대한 접근을 제한하는 경우 수직적 통합의 효율성은 향상할 수 있겠지만 상류 FM에서 시장지배력을 가진 공급업체는 하류 시장 수준에서 경쟁을 제한할 의도를 갖고 액세스를 제한하거나 저하할 수 있는 능력과 동기를 가질 수 있음. 이는 하류 FM 서비스의 경쟁자 또는 잠재적인 경쟁자를 약화시킬 수 있음
- 인접한 상류 또는 하류 시장에서 상당한 시장지배력을 가진 기업이 FM 서비스 또는 다른 시장에서의 서비스와 관련된 위치를 강화하기 위해 같은 회사의 기존 서비스에만 FM 서비스를 끼워 제공하는 경우를 예시할 수 있음

[그림 II-16] 하류 FM 서비스 시장에서의 FM 활용



자료: CMA(2023b)

- 경쟁 영향에 대한 불확실성과 규제 접근 제안
  - FM 서비스가 아직 개발 초기 단계에 있다는 점을 감안할 때 FM 배포가 하류 시장의 경쟁에 어떠한 영향을 미칠지 예상하기 어렵다는 점을 언급하고 있음

- 그럼에도 불구하고 FM 서비스의 선택 아키텍처에 따른 제공 방법과 서비스 액세스가 소비자 선택과 채택에 있어서 주요한 영향력을 지닌다는 점에서 상류 FM 배포 시장의 반경쟁적인 행위에 대하여 경계하고 있음
  - 개방형과 폐쇄형을 모두 포함한 다양한 비즈니스 모델을 갖춘 파운데이션 모델에 투자와 개발이 이루어지는 단계에서 규제는 AI 시장 경쟁과 혁신을 불필요하게 어렵게 만들 수 있어 현 단계에서 규제 도입을 권장하지 않고 있음. 이는 현재의 규제당국이 집행하는 현재의 법률에 근거하는 맥락 특유의(context specific) 접근에 부합함
  - 규제당국은 법률이 아닌 기반인 원칙을 제시하고 이런 원칙에 ‘적절한 고려’(due regard)를 하도록 하는 법적인 의무를 부담함. 이러한 원칙 기반의(principle-based) 규제 접근방식은 규제의 유연성을 허용하고 빠르게 발전하는 AI 시스템에서 나타나는 위험에 대하여 초기 단계에서 규제당국의 신속한 대응을 가능하게 한다는 점에서 규제 실익을 지님
  - CMA는 AI 파운데이션 모델 발전을 계속해서 면밀하게 모니터링할 것으로 보이며, 기존 경쟁법 및 소비자 보호법의 잠재적 위반에 대하여 적극적으로 조사할 것이라는 점을 분명히 밝히고 있음
- FM 시장에서의 개발 및 활용에 대한 6개 지도 원칙
- ① 접근성: 이 원칙은 과도한 제한 없이 주요 입력(예: 데이터, 컴퓨팅, 전문 지식 및 자본)에 대한 지속적인 접근을 촉진함. 성공적인 파운데이션 모델 개발자가 ‘선발자 혜택’ 또는 확고한 시장지위를 통해 불균형한 이익을 받는 것을 방지하고 신규 진입자가 초기 진입자에게 도전할 수 있도록 보장하여 FM 개발에서 지속적, 효과적인 경쟁을 촉진하는 것을 추구하도록 함 이 원칙은 또한 다른 개발자와의 경쟁을 제한할 수 있는 ‘강력한 파트너십’을 제한하도록 함
  - ② 다양성: 오픈 소스 모델이 해당 부문의 진입 및 확장 장벽을 줄이는 데 도움이 된다는 점을 인식하여 오픈 소스 모델과 폐쇄형 소스 모델을 모두 포함하여 비즈니스 모델의 지속적인 다양성을 장려하도록 함
  - ③ 선택: 플러그인, 애플리케이션 프로그래밍 인터페이스(API), 사내 개발 및 파트너십과 같은 비즈니스에 대한 다양한 배치 옵션을 보장하고 파운데이션 모델의 사용 방법을 결정할 수 있도록 충분한 선택권을 갖도록 하는 데 중점을 둠

- ④ 유연성: 필요에 따라 기업이 다수의 파운데이션 모델을 맞춰 사용하거나 (mix and match) 활용하도록 지원하는 상호운용성(interoperability)과 이용자가 다양한 서비스를 전환 또는 이용하고 하나의 사업자 또는 생태계에 고착되지 않도록 하는 것에 중점을 둠
- ⑤ 공정거래: FM 개발자 및 배포자의 반경쟁적 자사우대(self-preferencing), 끼워팔기 또는 결합판매와 같은 반경쟁적인 행위를 방지하여 최상의 파운데이션 모델과 서비스만이 승리할 수 있도록 시장 환경을 조성함
- ⑥ 투명성: 소비자와 기업이 정보에 기반한 선택을 할 수 있도록 파운데이션 모델의 콘텐츠 생성 위험과 한계에 대한 충분한 정보를 제공하도록 함
- 6개 지도 원칙의 상위 원칙은 모델 개발자와 활용자가 소비자에게 제공되는 결과물에 대한 책무를 진다는 책무성(accountability)이고, 6개 지도 원칙 중 ①, ②는 모델 개발 단계에서, ③, ④, ⑤는 다른 시장에서의 모델 이용 단계에서, ⑥은 소비자의 모델 이용 단계에서 적용되는 원칙임

#### (다) 미국

##### 인공지능 개발·이용에 대한 행정명령(Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence)

##### ● 행정명령의 의의

- 조 바이든 미국 대통령이 2023년 10월 30일 인공지능(AI)의 안전과 보안에 대한 새로운 표준을 수립하는 행정명령에 서명함
- 미국은 주로 개발자, 그들이 사용하는 데이터 및 도구에 초점을 맞추어 안전성과 보안에 관한 조치를 통해 소비자와 같은 다양한 집단을 보호하려는 것에 중점을 두고 있으며, 이를 전제로 정부 조직의 기능을 적극적으로 활용하여 산업적 이용도와 개발의 속도를 적극적으로 개선하려는 것으로 보임
- 인공지능의 큰 편익뿐만 아니라 높은 위험성에 대하여 선도적으로 대처할 것을 목적으로 밝히고 있으며, 이를 위한 8가지 기본 정책(원칙)을 통하여 인공지능의 개발과 사용을 발전시키고 관리하려는 의도로 행정부 각 부서와 기관은 관련 법률에 따라 적절하고 일관되게 이러한 원칙을 지키면서 다른 기관, 학계, 노동조합 등의 기타 관련 단체의 견해를 고려할 의무를 부담시키고 있음

### ● 8가지 기본 정책

- 인공지능은 안전하고 보안이 보장되어야 함
- 책임감 있는 혁신, 경쟁, 협업을 촉진하여 기술의 잠재력을 발휘하도록 권장함
- 형평성과 시민권 증진을 위한 책임감 있는 AI 사용을 규정하여 이를 진보시키는 방향으로 유도함
- 소비자, 환자 그리고 학생들을 위한 지원을 지시함
- AI의 책임 있는 개발과 활용을 위한 약속이 필요함
- 혁신과 경쟁을 촉진함
- 미국의 리더십을 해외로 더욱 확장함
- 정부의 책임감 있고 효과적인 AI 사용을 보장하고 강화함
- 경쟁 촉진과 관련하여, 연방거래위원회는 AI 시장에서 공정한 경쟁을 보장하고 AI 사용으로 인해 발생할 수 있는 피해로부터 소비자와 근로자를 보호하기 위해 연방거래위원회법에 따른 규칙 제정 권한을 포함한 위원회의 기존 권한을 적절하다고 판단되는 경우 행사할지 여부를 고려할 것을 권장함

### □ 미국의 경쟁법적 규제의 현황

#### ● 구체적 입법의 부재

- 규제당국이 제기하는 경쟁 우려 중 일부는 지배적 기업이 반경쟁적 전략을 실행하기 위해 AI를 활용할 수 있다는 사실과 관련되어 있어 AI를 사용하는 지배적 기업은 규제당국으로부터 AI 사용 방식에 대한 면밀한 감독을 받아야 할 필요성이 제기됨
- 미국의 경우 생성형 AI에 특유한 법률을 제정하지 않은 상황이므로, 경쟁법 집행에 관해서 FTC의 규제 방향이 중요한 기준이 됨
- 연방 반독점법 집행을 담당하는 두 기관인 연방거래위원회(FTC)와 법무부(DOJ)는 올해 AI의 반경쟁적 잠재력에 대해 수많은 공개 성명을 발표했음

#### ● FTC를 통한 규제의 현황

- FTC는 기술 블로그(Technology Blog)를 통해서 생성형 AI에 관한 여러 정책적 관점에 관하여 블로그를 통해서 관심 방향을 공개하고 있음

- FTC는 블로그를 통해서 2023년 6월 29일 경쟁법적 관점에서의 문제점에 관해서 발표하였는데, 그 구체적인 내용으로 우선 생성형 AI의 핵심적인 경쟁요소로서 데이터, 전문 연구 인력, 컴퓨터 리소스에 대한 접근 가능성 그리고 오픈 소스 모델이 잠재적인 반경쟁적 전략을 사용할 가능성을 예로 들었음 (FTC, 2023)
- 데이터의 경우 사전 학습 단계에서 데이터의 양과 질이 중요한 역할을 담당하며 이러한 측면은 신규 시장 진입자에게 진입 장벽으로서의 역할을 담당할 수 있다고 봄
- 생성형 AI의 또 다른 필수요소로서의 노동 전문성은 높은 기술력을 가진 엔지니어와 연구진을 필요로 하고 기존의 거대 기업들이 절대적으로 유리한 상황이라고 봄
- 상당한 수준의 컴퓨팅 리소스를 필요로 하는 생성형 AI의 특성상 특수 사양의 칩들을 제조할 수 있는 능력을 가진 소수의 거대 기업의 존재는 반경쟁적 관행을 만들어낼 유인이 높다고 봄
- 오픈 소스 생태계로 발전해 온 생성형 AI의 산업이 장기적인 경쟁을 저해하는 방식으로 초기에는 공개, 나중에는 폐쇄(open-first, closed later)하는 전략을 사용하여 반경쟁적 행태를 띠 가능성이 높다고 봄
- 이에 대한 대책으로 FTC는 강력한 법 집행을 통해 새로운 비즈니스가 경쟁하고, 연구자들이 자유롭게 기술 상태를 가장 잘 발전시킬 수 있는 분야로 이동하며, 기업가들이 혁신을 지속할 수 있는 활기찬 시장을 지원하기 위해 노력하고 있으며, 생성형 AI를 둘러싼 경쟁 문제가 계속 발전함에 따라 경쟁국은 기술국과 긴밀히 협력하여 불공정 경쟁 방법을 식별하고 해결하기 위해 모든 도구를 사용할 것을 밝힘

## 마. 새로운 정책 개발 또는 정책 변화의 방향

### (1) 새롭거나 변화된 정책의 원칙과 수단으로서 다양한 규제체계 모델

#### □ 인공지능 기술 자체에 대한 규제

##### ● 유럽연합(EU)의 리스크 기반 규제(risk-based regulation)

- 리스크 기반 규제는 일반적으로 규제적 행동의 우선순위를 당사자가 규제기관의 목적 달성을 위해 제기하는 리스크의 평가에 따라 정하는 것으로 이해됨 (Baldwin, et. al., 2012)
- 리스크 기반 규제는 첫째, 목적과 규제기관이 그 목적을 달성하기 위하여 제기할 수 있는 리스크의 식별, 그 리스크를 평가하고 우선순위를 매기는 시스템의 개발, 셋째, 평가와 우선순위 시스템과 연결되는 실행 수단의 개발을 구성 요소로 함
- 이론적으로 위험(Gefahr; danger)은 가까운 장래에 손해를 야기할 충분한 상태를 의미하고, 리스크(Risiko; risk)는 위험에 이르지 않은 손해 발생의 개연성(probability)을 의미하는 것으로 구별될 수 있음(이원우, 2016; Baldwin et. al., 2012)
- 리스크를 위해(harm)의 개연성과 중대성의 불확실성 속에서 이익(benefits)을 위한 선택에 의하여 발생하는 불확실한 법익침해의 가능성으로 정의하기도 하는데(윤혜선, 2018), 내생적으로 측정이 불가능한 불확실성과 정량화로 처리할 수 있는 리스크를 구별할 수 있음(Baldwin et. alt., 2012)
- 대체로 경찰법적 위험과 경제법적 위해를 동등한 수준으로 보고 규제의 수준이라는 관점에서 볼 때, 위해는 개입의 근거가 되고(harm-based intervention) 리스크는 관리 성격을 갖는 규제의 대상이 됨(risk-based regulation)
- 유럽위원회는 인공지능 백서에서 인공지능 기술을 대상으로 리스크 기반 접근을 원칙으로 한 새로운 규제체계 도입이 필요하다고 주장함. 즉 원칙적으로 기존의 규제체계를 조정하여 인공지능 기술 리스크에 대한 적절한 대응이 이루어지도록 하되 특별히 리스크가 큰 경우에 한해서는 새로운 규제체계를 도입하여 선별적·비례적 대응이 이루어지도록 하겠다는 것임(EU, 2020b; 이상윤, 2020)

- 이런 규제체계는 고정되고 집중적인 리스크와 경감 방안 목록을 갖춘 단일한 규제체계이지만, 길보기에는 유사한 인공지능의 적용이 다른 맥락에서 이루어질 때 제기되는 다른 수준의 리스크에 유연하게 대처하지 못할 수 있음 (Edwards, 2022)
- 영국의 맥락 기반 규제(context-based regulation)
  - 영국은 인공지능에 대한 리스크 기반 규제 원칙을 표방하면서도 기술 자체를 규제하지는 않고 기술이 적용되는 맥락에 따라 규제하겠다는 맥락 기반 규제를 구체적인 원칙으로 제시함(Department for Digital, Culture, Media and Sport(2022))
  - 맥락 기반 규제는 새로운 규제를 도입하지 않고 기존 규제 구조에 의존하면서 출현하는 리스크에 따라 리스크를 식별하고 적응하는 유연성과 비례성을 보장함. 다만 이런 접근이 상황에 따라 가변적인 속성으로 인해 집중적인 접근법(centralised approach)에 비해 통일성이 부족하다는 점을 보완하기 위해 일정한 지도 원칙을 제시함
  - 영국 정부가 맥락 기반 규제를 보완하기 위한 지도 원칙으로 제시한 것은 적극적인 친혁신 접근(proactively pro-innovation approach)이고, 규제기관에게 분야 공통으로(cross-sectoral) 적용되는 이행원칙으로 OECD의 AI 원칙에 기초하여 ‘안전, 보안, 견고성, 투명성, 설명가능성, 공정성, 책임, 거버넌스, 이익제기, 구제’ 등을 제시함(채은선, 2023)
- 권리 기반 규제(rights-based approach)
  - 인공지능에 관하여 개인이 기본권에 해당하는 디지털 권리를 갖는다는 점을 전제로 하여 리스크 관리나 평가가 아닌 권리 침해 방지하기 위한 규제를 해야 한다는 견해임(Fanny, 2023)
  - 이 견해에서는 기본권의 관점에서 AI로 인한 인권에 대한 위협과 AI와 자동적 의사결정(automated decision-making, ADM)이 예측하지 못한 방식으로 행동할 수 있다는 점 때문에 독립적인 전문가와 규제기관에 의한 적절한 인권영향평가를 통해 위험(danger)을 초래하는 AI 기술 사용에 대한 금지 조치가 내려져야 한다고 주장함

- 이 견해에서는 EU의 일반데이터보호법(General Data Protection Regulation, GDPR)이 일련의 정보주체의 권리에 기반하여 그 권리 보호를 위해 정보주체가 개인정보를 통제하도록 하고 개인정보를 사용하는 기업 등의 개인정보처리자에게 의무와 요건을 부여하고 있다는 점에서 권리 기반 규제로 볼 수 있다고 주장함

## □ 인공지능 기술이 적용되는 생태계에 대한 규제

### ● 생태계 규제에 대한 기본적인 접근

- 기존에는 인공지능 기술이 기존의 디지털 플랫폼 사업자에게 채택되어 이용자에 대한 정보의 우선순위를 정하고 콘텐츠 조정(moderation)을 하는 데 이용될 경우 그로 인해 기존의 시장 리스크가 강화될 것인지의 문제에 초점이 있었음
- 파운데이션 모델을 중심으로 기존의 플랫폼 생태계 외부에서 새로운 생태계가 형성되고 이런 움직임에서 오픈 소스 모델이 의미 있는 역할을 하는 상황에 주목할 필요
- 기존의 규제가 기초를 두고 있는 사업모델이 아닌 대안 사업모델로 중대한 편익과 높은 이용자 보호 수준을 달성할 가능성을 고려하여 새로운 형태의 혁신 거래(innovation deals)에 개방적인 태도가 필요함(Renda, 2019)

### ● 구체적인 규제 원칙

- 우선허용, 사후규제 원칙
- 리스크 관리적 접근(risk management approach)
- 적응하고 학습하는 접근(adaptive and learning approach)
- 비례적 규제
- 원칙 중심(principle-based) 규제
- 결과 중심(outcome-based)규제
- 규제 거버넌스 구축을 통한 연합(joint-up)규제
- 전문적, 참여적 규제
- 자율규제, 공동규제
- 탄력적이고 민첩한(flexible and agile) 규제

## (2) 새롭거나 변화된 정책 방향 제안

### □ 파운데이션 모델 제공자에 대한 혁신 지원 정책

- 친경쟁(pro-competition) 접근 원칙 vs 친혁신(pro-innovation) 접근 원칙
  - 친경쟁 접근은 영국 정부가 퍼먼 보고서의 제안에 따라 디지털 시장에서의 새로운 경쟁정책의 원칙으로 제시한 것임. 친경쟁 접근은 디지털 시장에서의 네트워크 기반, 데이터 주도(network-based and data-driven) 플랫폼 사업 모델의 쏠림 현상에 대응하여 디지털 시장이 작동하고 경쟁, 혁신과 소비자 선택의 기회가 창출되는 방식을 변화시키는 규칙과 기준을 설정한다는 원칙임 (Report of the Digital Competition Expert Panel, 2019)
  - 영국 정부는 AI 규제에서는 친경쟁 접근이 아닌 친혁신 접근을 규제접근의 원칙으로 제시하는데, 이는 해외 업체에게 이미 지배되고 있는 디지털 플랫폼 생태계와 달리 AI 분야에서는 영국이 선도적인 역할을 갖도록 지원하는 것을 목표로 하여 혁신을 촉진하고 자국 기업이 경쟁력을 갖는 AI 생태계를 지원하려는 취지를 담고 있는 것임(Department for Digital, Culture, Media and Sport, 2022)
  - 혁신 우선 원칙은 혁신을 절대로 제한해서는 안 된다는 것이 아니라 정책입안자와 집행자가 기존 및 문서화된 위험을 해결하기 위해서만 혁신을 제한해야 한다는 것으로 혁신의 보호를 기본권보다 우선할 수 있는 근본적인 목표로 삼아야 한다는 것임(Schrepel & Pentland, 2023)
  - 규제 및 정책 의제 설정에 관해서는, 첫째, 이 분야의 새로운 규칙과 표준은 독점력을 초래하는지 여부에 대한 영향평가가 발표된 후에야 제정될 필요가 있고, 둘째, 특정 AI 전문성을 갖춘 규제기관이 영향평가를 수행하고 규제를 발표해야 하고 여러 규제기관의 구성원이 참여하는 (비공식) 협의회를 만들어 AI 정책을 조율하는 것은 좋은 방식임
  - 오픈 소스 및 오픈 액세스 기업의 연구개발을 촉진하기 위한 목적으로 반독점 규정의 예외 조항을 신설할 것을 권고하는 견해가 있음. 이를 통해 오픈 소스 및 오픈액세스 업체들이 카르텔 제재의 위험 없이 전략적 제휴를 맺을 수 있어야 한다고 봄(Schrepel & Pentland, 2023)

## □ 디지털 플랫폼 정책보다 진취적인 정책 목표 설정과 규제접근이 필요함

- 디지털 플랫폼 정책보다 진취적인 정책 목표 설정과 규제접근이 필요한 이유
  - 초거대 AI는 디지털 세상에서 새로운 시장을 얻기 위한 경쟁(competition for the market)의 새 지평을 열 것으로 기대되지만, 새로 형성하게 될 시장은 과연 경쟁적으로 형성될 것인지, 기존 디지털 플랫폼 사업분야에서 문제가 되고 있는 자사우대, 결합판매 등을 통한 시장지배력 전이 전략, 데이터 집중의 이슈가 자생적으로 해소될 것인지는 불투명함
  - 경우에 따라서는 GPT 언어 모델과 그 훈련 데이터에 대한 접근성을 높이기 위한 상호운용성(interoperability) 규제 도입이 논의될 수도 있음
  - 초거대 AI를 기반으로 한 생성형 AI 서비스 사업자가 기존 디지털 플랫폼 사업 모델 및 수익 확보 전략과 다른 어떤 사업모델 및 수익 확보 전략을 취할 것인지도 관전 포인트임
  - 생성형 AI 서비스도 플랫폼화를 통하여 양면적 또는 다면적 사업모델로 진화할 수 있는데, 이때 요금 부과 대상, 요금 책정의 방법 등 수익 확보 전략의 다변화가 예상됨 또한 오픈AI와 그 제휴관계에 있는 마이크로소프트는 이미 선발자의 우위를 갖고 있으므로, 후발주자인 기존 빅테크가 경쟁 기술을 개발하는 다른 스타트업과 협력 또는 제휴관계를 맺거나 이를 인수할 경우 이에 대해서 경쟁법이 어떤 역할을 해야 할 것인지도 논의가 필요함
- 혁신 촉진이 더 중요한 생성형 AI 분야에서 경제규제정책과 경쟁정책의 관계 재정립 필요성
  - 2022년 9월 발표된 대한민국 디지털 전략에 나타난 현재의 디지털 플랫폼 정책은 '혁신-공정이 조화로운 디지털 플랫폼 생태계 구축'이라는 정책 목표를 갖고 관련 부처 간 협의에 의한 정책 추진을 표방하였지만, 최근 공정거래위원회가 가칭 플랫폼 경쟁촉진법을 추진하는 과정에서 나타나듯이 규제 도입 단계에서 규제접근 원칙에 관한 부처 간 조율과 이해관계자 의견수렴의 과정이 원활하게 이루어지고 있다고 보기 어려움
  - 우리나라는 EU와 달리 일부 핵심 플랫폼 서비스 분야에서 국내 플랫폼 기업이 경쟁력을 갖고 핵심 플랫폼 서비스 분야에서의 경쟁력을 기반으로 플랫폼 서비스 단위를 넘어 시장을 교차하여 이루어지는 플랫폼 생태계 간 경쟁에서 해외 빅테크와 국내 플랫폼 기업이 상호 경쟁압력으로 작용하는 상황이므로, EU와

- 같이 핵심 플랫폼 서비스 분야에서 시장 진입이 어려운 약한 경합성(weak contestability)의 시장 상황에 있다는 실증적 증거가 부족함. 이는 EU의 디지털시장법(DMA)에 유사한 입법의 필요성에 의문을 제기할 수 있는 근거임
- DMA 유사 법률은 리스크 기반 접근에 따른 유연한 규제와는 거리가 먼 위해 기반 접근에 따른 사전 개입에 해당함. 따라서 디지털 플랫폼 서비스 분야에 경쟁당국의 주도로 이런 방식의 규제 도입을 논의하는 것은 전문분야 경제규제 정책이 필요한지 여부와 관계없이 시장 일반에 적용되는 사후규제를 담당하는 경쟁당국의 본원적 역할과 맞지 않고 경쟁 촉진은 물론 혁신 촉진에도 반하는 것임. 디지털 플랫폼 서비스 분야의 외곽에서 새로 출현하여 이제 막 이를 중심으로 한 새로운 플랫폼 생태계 구축 경쟁이 시작된 생성형 AI 서비스 분야 역시 디지털 플랫폼 서비스 분야와 밀접하게 연결된다는 점에서 디지털 플랫폼 서비스 분야에서의 새로운 규제 도입 여부는 매우 신중하게 판단해야 할 사안임
  - 생성형 AI 서비스 분야에서 혁신 촉진의 정책 원칙을 제시하고 자국의 혁신적인 파운데이션 모델 구축 정립 노력에 인센티브를 주는 규제체계를 수립할 필요가 있음. 이를 위해 생성형 AI 서비스 분야는 시장 일반에 적용되는 경쟁정책보다 전문분야에 적용되는 경제규제정책에 우선순위를 두고 리스크 기반 규제 중에서도 맥락 기반 규제 원칙에 따른 규제체계를 수립하고, 경쟁정책은 기존의 규제수단을 이용하여 차별(예: 플랫폼과 애그리게이터에 대한 공정한 접근), 배타적 계약(예: 경쟁 파운데이션 모델에 제공하지 않도록 하는 것), 끼워팔기(예: 다른 파운데이션 모델로 전환하지 않도록 하는 인센티브 제공) 등 이용자 접근을 제한하는 관행을 탐지하는 데 집중할 것이 권장됨
  - 생성형 AI 서비스 분야에서 선제적인 경제규제정책의 목표 설정과 규제체계 마련은 현실적으로 생성형 AI 서비스 분야의 중심 영역인 파운데이션 모델 구축 능력을 가진 국내 디지털 플랫폼 기업에 대한 경제규제정책이나 경쟁정책이 리스크 기반 규제의 원칙에서 벗어나 지나치게 개입적이 되지 않도록 피드백을 주는 정책 효과를 기대할 수 있음

## 4. AI가 광고산업에 미치는 영향

### 가. 연구목적

- AI의 발달이 광고산업에 미치는 영향
  - AI의 발달, 특히 최근의 생성형 AI의 급속한 발달과 확산은 광고와 인접 산업에 막대한 영향을 미침
  - AI는 광고의 효과와 효율성을 극대화하는 데 중요한 역할을 함
    - 광고산업에서의 AI의 영향은 광고제작, 매체 운영, 예산 최적화, 효과분석 등의 영역에서 두드러짐
- AI의 발달과 플랫폼
  - 광고가 주요 비즈니스 모델인 주요 글로벌, 로컬 플랫폼들은 AI를 적극적으로 활용하여 경쟁력을 확보하려고 함
    - AI의 독자적 개발을 위해 대규모의 투자가 필요한 상황에서 거대 플랫폼들이 광고주와 소비자에게 고품질의 광고 서비스를 경쟁 플랫폼들보다 빠르고 안정적으로 제공할 수 있음
  - 최근 출시 혹은 출시 예정인 주요 플랫폼들의 AI 기반 광고 서비스 현황을 살펴봄
- AI의 발달과 광고제작
  - AI를 활용해 다양한 브랜드의 광고가 제작되고 캠페인이 진행되어 옴
    - 대표적인 AI 기반 제작 광고 사례를 분석함
- AI를 활용하는 광고산업에 대한 전망
  - AI가 광고 제작과 매체 거래 등에 적극적으로 활용됨에 따라 관련 우려가 증대되고 있음
    - 광고의 창의성과 독창성, 소비자의 혼동, 전문인력의 대체, 개인정보 보호 등이 AI와 관련한 광고 영역의 주요 이슈로 부상함
  - 주요 플랫폼들은 소비자의 우려를 불식시키기 위한 자율규제를 도입하거나 고려하고 있음

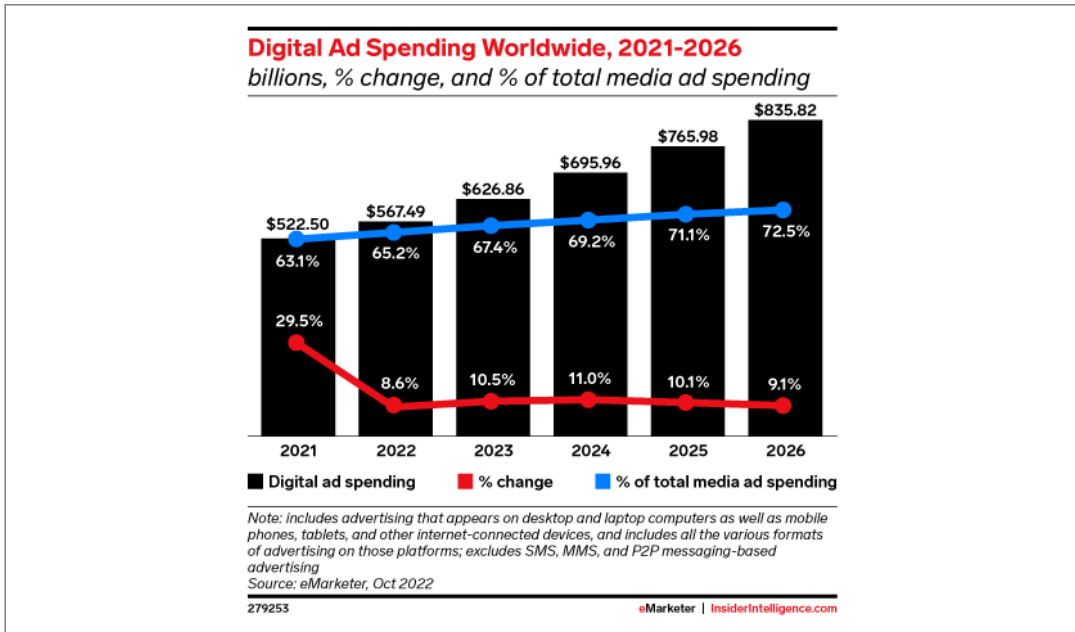
## 나. 연구배경

### (1) 광고 환경의 변화

#### □ 디지털 중심의 광고매체 시장과 파편화

- 디지털 미디어 이용의 증가와 함께 방송 등 전통적인 매체를 앞질러 디지털 중심으로 광고매체 시장이 재편됨
  - 코로나 팬데믹을 겪으며 디지털 편중은 가속화 되고 디지털 광고시장의 성장은 지속됨

[그림 II-17] 글로벌 디지털 광고비 성장



- 디지털 광고매체 시장은 모바일 앱을 비롯해 수많은 매체로 구성되어 파편화 되어 매체 기획과 집행이 복잡해지고 광고효과 입증에 어려워짐
  - 사람이 주도하는 전통적인 방식의 광고매체 집행이 어려워짐
- 디지털 환경에서 수많은 소비자 접점을 통합적으로 분석, 이해해야 하기 때문에 광고 영역의 데이터와 기술 활용의 중요성 증대

## (2) 광고 패러다임의 변화

### □ 광고의 개인화

- 개인화 광고는 디지털 중심으로의 매체환경 변화와 기술의 고도화를 반영해 효율성과 설득력을 제고한 광고로서 현재의 광고 생태계와 트렌드를 대표함
  - 개인화 광고(personalized advertising)는 맞춤형 광고(customized advertising)라고도 불리며, 소비자를 전체 혹은 다수의 집단으로 여기지 않고 소비자 개개인에 대한 정보를 이용하여 각 소비자에게 최적화하여 제작하고 전달하는 광고를 의미함
  - 소비자가 온라인 활동을 통해 남긴 행동 데이터는 온라인 맞춤형 광고(online behavioral advertising)을 가능케 함
  - 즉 소비자의 인구통계학적 속성뿐 아니라 라이프스타일, 관심사, 과거 구매 기록, 검색 기록, 위치 공간 등 행동 기반의 디지털 데이터를 수집, 분석하여 얻어낸 인사이트를 바탕으로 개별 소비자에게 맞춤형 메시지를 전달할 수 있음
  - 실시간으로 축적된 행동 기반 데이터를 기반으로 소비자의 성향과 선호를 반영한 메시지를 전달할 뿐 아니라 소비자의 필요를 예측하여 제품, 서비스를 추천하고 실제 구매로의 전환이 이루어지는지 파악이 가능한 형태로 진화하고 있음
- 다양한 플랫폼의 다양한 맥락에서 광고에 노출되는 수많은 소비자에게 개인화 광고를 전달하기 위해서는 데이터와 기술의 활용이 절대적으로 필요함
  - 타겟 소비자에게 적합한 장소와 시간에 맞춤형 메시지를 전달하는 “right person, right place, right time, right message”이 중요

### □ 광고와 커머스 결합

- 광고효과는 전통적으로 인지, 관심, 이해, 선호, 구매의도 등으로 평가되었으나, 온라인 중심의 광고매체 환경으로 전환되면서 퍼포먼스 광고의 중요성 증대
  - 퍼포먼스 광고(performance advertising)는 광고를 통해 상담, 앱 다운로드, 구매 등 구체적인 목표행동까지 도달하는지 추적하고 이러한 전환(conversion)에 기여한 광고에 대한 비용을 지불하고 데이터를 분석해 향후 광고효과를 개선하는 데 활용하는 과정을 의미함

- 나아가 코로나 팬데믹으로 인해 온라인 커머스가 급성장하면서 광고와 커머스를 직접적으로 결합하는 '미디어 커머스' 성장
  - 미디어 커머스(media commerce)는 미디어(media)와 커머스(commerce)의 조합어임
  - 미디어 관점에서, 콘텐츠를 제공하며 대가로 구독료를 받거나 콘텐츠를 무료로 제공하는 대신 광고를 수익원으로 활용해왔지만, 광고의 혼잡도(clutter)가 높아지고 소비자들의 피로감이 증가하면서 광고효과는 감소하고 직접적이고 즉각적인 매출 증대를 이끄는 광고효과에 대한 요구는 증대됨에 따라 커머스 기능을 확대함
  - 커머스 측면에서, 전자상거래가 성장하고 경쟁이 치열해지면서 사업자들은 차별화를 꾀하고 트래픽을 높이고 즉각적인 구매를 유도하는 효과적이며 새로운 방법으로서 미디어 커머스를 도입함
  - 소통이 본연의 목적인 소셜미디어도 커머스 기능을 탑재하고 전자상거래 사업자들이 브랜드드 콘텐츠, 라이브 방송, 인플루언서 등 매체의 콘텐츠와 소통방식을 도입하여 커머스에 활용하는 등 미디어와 커머스의 경계가 허물어지며 다양한 결합 형태가 나타남
  - 특히 최근 라이브 스트리밍(live streaming)과 커머스(commerce)를 결합한 라이브 커머스(live commerce)의 성장이 두드러짐. 페이스북, 인스타그램, 유튜브, 틱톡과 같은 소셜미디어 플랫폼과 함께 네이버, 카카오 등 빅테크 플랫폼, 배달의 민족, 쿠팡, 쓱닷컴 등 거의 모든 전자상거래 사업자와 개별 브랜드가 라이브 커머스를 도입함

□ 광고와 콘텐츠 융합

- 디지털 환경에서 콘텐츠 제작이 용이해지면서 다양한 개인과 생사주체가 자유롭게 콘텐츠를 제작하고 공유함
  - IT 저널리스트 톰 포렘스키(Tom Foremski)는 “모든 기업은 미디어 기업이다(“Every company is a media company)”라며 전통적인 매체 사업자가 아니더라도 기업이 다양한 콘텐츠를 직접 제작하고 유통할 수 있는 환경에 주목함<sup>8)</sup>

8) Presspage (2013. 5. 9). “Every Company Is A Media Company”. <http://www.presspage.com/news/every-company-is-a-media-company>

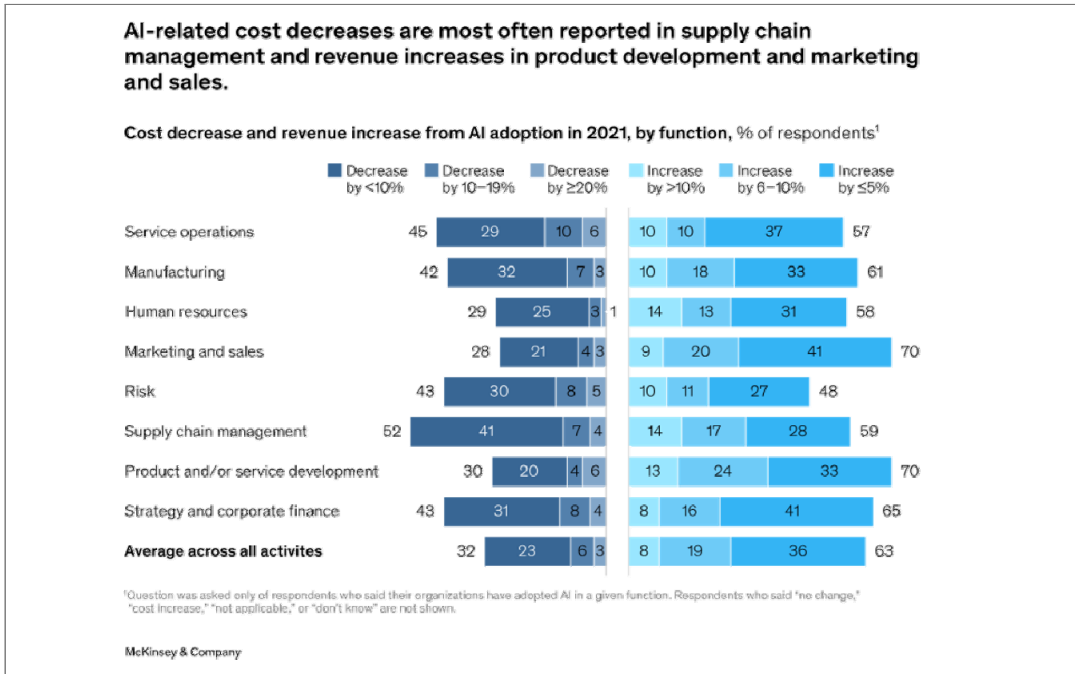
- 광고의 목적을 가지지만 콘텐츠의 형태와 가치를 지향해 브랜드 정보나 이미지를 담은 새로운 유형의 하이브리드 콘텐츠 혹은 광고가 등장함
  - 소비자는 광고는 불필요하고 성가시다는 인식 때문에 회피하는 경향을 보이지만 콘텐츠는 찾아서 소비하기 때문에 콘텐츠의 가치를 가진 광고에 대한 주목도가 높게 나타남
  - ‘수익성 있는 소비자 행동의 유발을 목적으로 가치 있고 매력적인 콘텐츠의 생산 및 배포를 통해 명확하게 정의되고 이해된 목표청중을 유도·획득하며 그들과 인게이지(engage)하는 마케팅 및 비즈니스 프로세스’인 ‘콘텐츠 마케팅’ 관점에서 이런 현상을 이해할 수 있음(Pulizzi, 2013)

### (3) 광고의 AI 활용에 대한 인식

#### □ 마케팅 분야의 AI 활용

- 마케팅/세일즈 분야는 AI 활용으로 수익 증가 효과가 발생하는 대표적인 영역으로 꼽힘
  - 생성형 AI는 인간의 질문에 대한 답을 글, 이미지(영상), 음성(음악) 등의 형태로 제시
  - 생성형 AI는 광고 기획을 포함하는 다양한 마케팅 활동에서 생산적인 도구로서 활용도가 더욱 확대될 전망(McKinsey&Company, 2022. 12. 6.; 2023.4.24.)  
광고 영역에서의 AI 활용 필요

[그림 II-18] 업무별 AI 관련 비용 증감세 추세



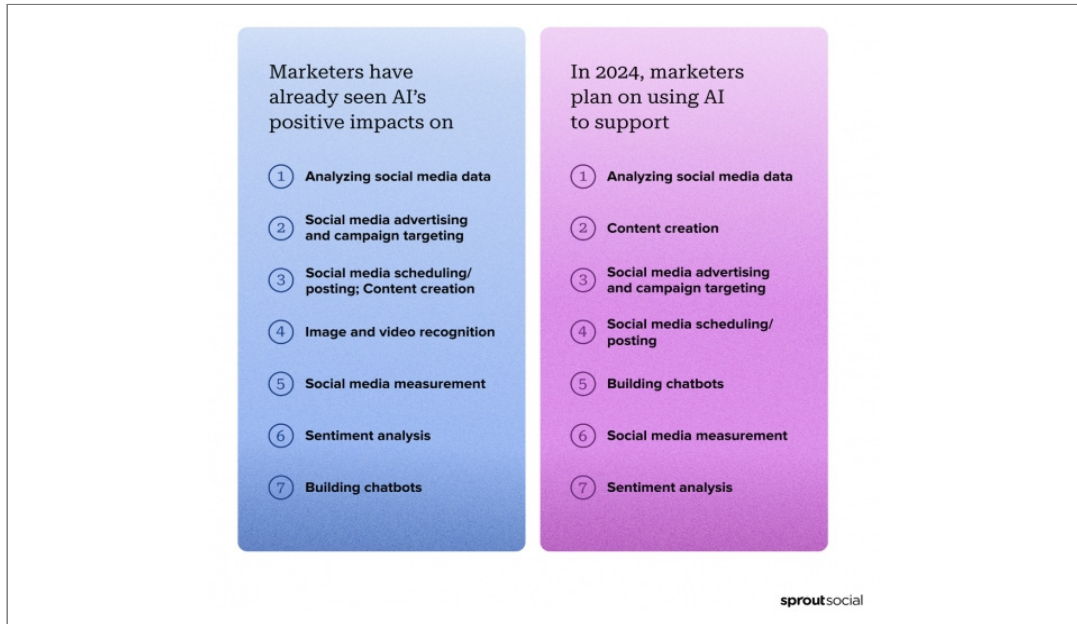
자료: McKinsey & Company (2023)

### □ 마케터의 AI 인식

- 마케터들은 소셜미디어에서 개인화된 연결에 대한 욕구를 충족하고 더 나은 고객 경험을 대규모로 제공하기 위해 AI를 활용하고 있는 것으로 나타남<sup>9)</sup>
  - 설문에 참여한 미국, 영국 900명의 마케터들 중 81%는 AI가 이미 업무에 긍정적인 영향을 미쳤다고 답했으며, 가장 큰 혜택은 창의성을 발휘할 수 있는 시간 증가(78%)와 효율성 증대(73%)라고 응답함
  - 또한 마케터들은 데이터 분석과 같이 시간이 많이 소요되는 업무를 AI가 대신함으로써 소비자가 원하는 개인화된 대화를 나누는 등 서비스를 제고할 수 있는 시간을 더 많이 확보할 수 있다는 긍정적인 효과도 경험하는 것으로 나타남
  - 2024년에는 소셜 데이터 분석(68%), 콘텐츠 제작(66%), 소셜 미디어 광고 타기팅 및 캠페인 타기팅(65%), 챗봇 구축(63%)에서 AI가 역할을 할 것으로 예상함

9) <https://www.madtimes.org/news/articleView.html?idxno=18747>

[그림 II-19] AI의 현재 영향과 미래의 영향 예측



자료: Sprout Social, 2023

□ 광고 분야의 AI 활용

- 2023년 10월 서울에서 개최되었던 애드 아시아(Ad Asia)의 키워드는 생성형 AI였으며, AI가 광고와 마케팅의 필수적인 도구라고 강조함
  - 애드 아시아 연사였던 미디어몽크스(Media Monks) 공동창업자 웨슬리 터하는 “AI의 등장으로 광고 산업의 패러다임이 바뀌고 있다. 지난 10년간 ‘최적화’가 화두였다면 앞으로는 ‘극대화’에 초점을 맞춰야 한다”고 피력함
  - 애드 아시아의 많은 연사들이 AI를 받아들이는 광고회사만 살아남을 것이라며 광고 영역에서의 적극적인 AI의 활용 중요성을 강조함
- 기술과 데이터를 적극적으로 활용해 온 광고 산업에 AI가 도입됨으로써 새로운 기회와 가능성 확대
  - 광고 기획, 제작, 매체 집행, 효과 분석 등 광고 산업의 전반적인 영역에서 AI의 활용 추진
  - 소비자 인사이트 발굴, 고객 세분화, 맞춤형 서비스를 제공하기 위한 예측 모델링과 프로세스 자동화에 이르기까지 빠르게 확장되며, 광고의 지능화를 이끌 수 있는 수단으로 AI에 주목(Salesforce, 2021; Taylor & Carlson, 2021)

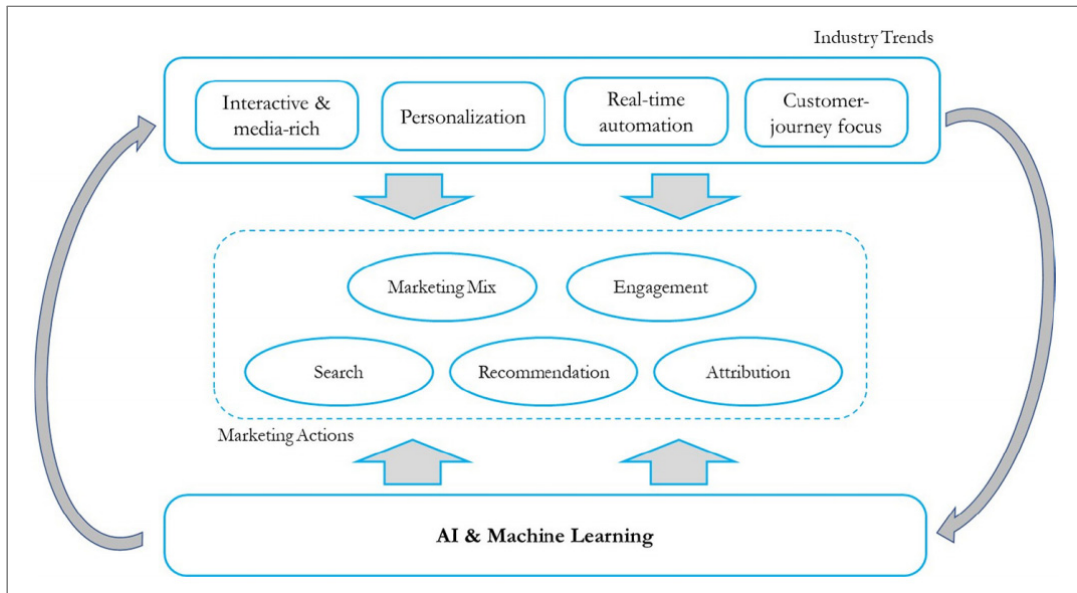
## 다. 광고와 마케팅의 AI 활용

### (1) 광고와 마케팅의 AI 활용 현황

#### □ AI가 광고와 마케팅의 고도화에 미치는 영향

- AI는 마케팅 및 광고 관행을 빠르게 변화시키며 기업이 대규모 비정형 데이터를 효율적으로 처리할 수 있도록 지원
  - 머신러닝을 기반으로 하는 AI 도구는 텍스트, 이미지, 비디오 콘텐츠가 풍부한 환경에서 인사이트와 솔루션을 생성하는 데 도움을 줌
  - 광고에서 대규모의 컨텍스트 기반 개인화를 추진하여 세분화를 더욱 정밀하게 만드는 데 도움을 줌
  - 마케팅 환경의 복잡성으로 인해 자동화된 실시간 대응에 용이함을 머신 러닝 알고리즘으로 제공함
- AI 기반 도구를 활용한 검색 최적화, 딥 뉴럴 네트워크를 포함한 머신 러닝 알고리즘을 활용한 추천 시스템의 고도화, 소비자 여정에서 다양한 접점이 실제로 기여하는 바를 파악하는 데 실질적 도움

[그림 II-20] AI 활용 마케팅 랜드스케이프

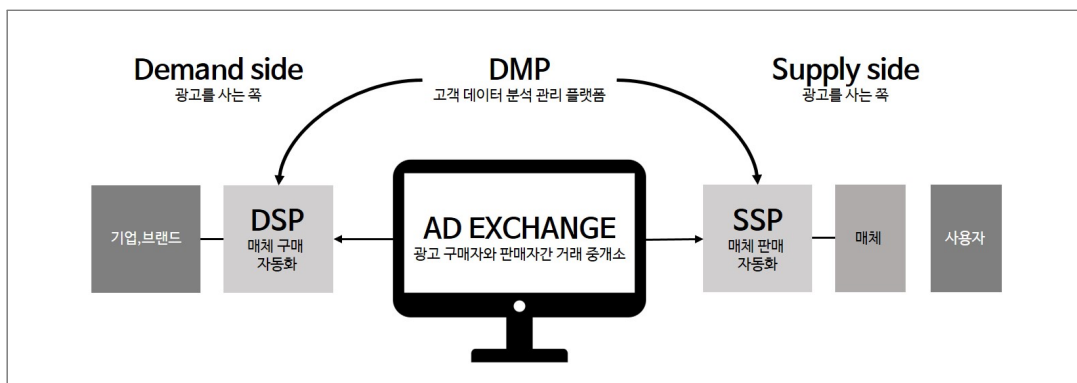


자료: Ma & Sun, 2020

□ AI가 프로그래매틱 광고에 미치는 영향

- AI는 과거의 매체 중심에서 현재의 소비자 중심으로 광고 거래를 전환하고 광고비의 효율성을 극대화하는 데 중추적인 역할
  - 전통적인 광고 매체 거래는 관련 실무자 간의 커뮤니케이션을 통해서 이루어지는 반면, 프로그래매틱 바이닝(programmatic buying)은 소비자와 매체 데이터를 바탕으로 설정된 알고리즘을 기반으로 광고 캠페인에 필요한 조건들을 반영한 자동화된 광고 거래를 통해 효율성을 제고한 시스템임
  - 온라인 광고의 중요한 기술적 발전 중 하나가 실시간 입찰(RTB: Real-time Bidding)이며, 프로그래매틱 바이닝은 디지털 환경에서 광고 게재 지면이나 시간을 판매하기 원하는 매체와 이를 구매하려는 기업 혹은 대행사 간 거래를 실시간 경매 방식으로 운영함(Wang, Zhang, & Yuan, 2016)
  - 개별 광고의 노출 당 경매가 이루어지며 실제 광고 노출이 발생하기까지 전 과정이 0.1초 이내에 이루어지며, 가용한 광고 게재 지면이나 시간을 원하는 여러 광고주나 실무자들이 원하는 가격으로 입찰에 응하고 가장 높은 가격을 제시한 광고가 노출되는 방식임
  - 실시간 입찰을 통해 많은 수의 광고 노출이 자동화된 유형을 통해 구매됨으로써 디지털 광고 시장이 신속하고 효율적이며 투명하게 운영될 수 있음

[그림 II-21] 프로그래매틱 바이닝 프로세스<sup>10)</sup>



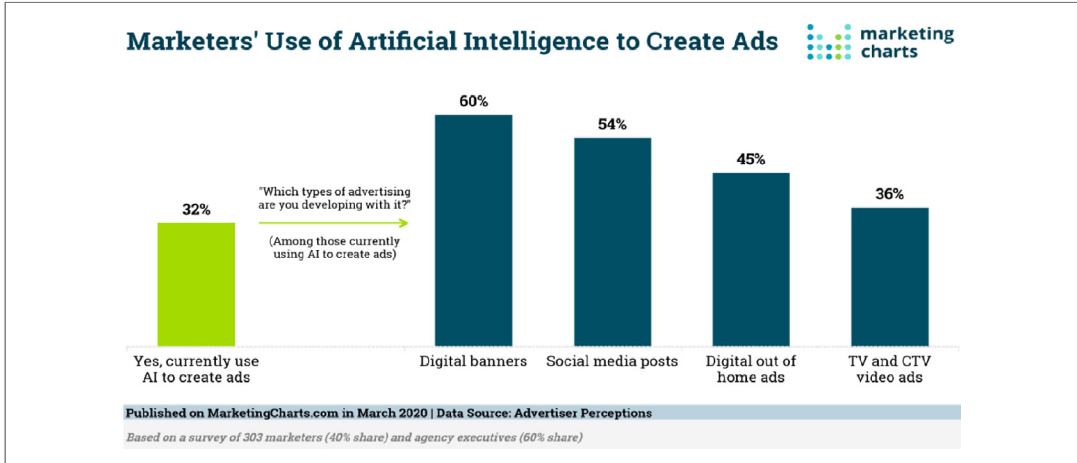
10) <http://www.dator.co.kr/bmonthly/2758473>

- AI 기술은 구매와 판매뿐 아니라 광고 메시지와 디자인에도 적용되어 소비자 데이터를 이용해 소비자가 필요로 하거나 관심을 가질 광고를 실시간으로 구성해 적시에 노출되도록 함
  - 예를 들어 DCO(Dynamic Creative Optimization)은 동영상, 이미지, 텍스트 등 광고의 소재 구성 요소를 타겟 소비자에게 적합하게 자동으로 변형해 맞춤형 광고를 노출하는 기술을 의미함
  - 데이터를 활용해 누가, 언제, 어디서 광고에 노출되는지를 파악하여 해당 타겟 소비자의 필요, 취향, 관심사, 선호와 맥락을 반영해 실시간으로 개인화 광고를 구성하여 전달함
  - 전통적인 광고가 하나 혹은 소수의 광고 소재를 제작해 노출하는 것과 대비되는 방식으로서 모두에게 같은 광고가 아닌 개개인 소비자의 특성을 고려한 해당 소비자에게 최적화된 광고를 자동으로 제작해서 바로 전달함
- AI를 활용한 프로그래매틱 바이닝과 프로그래매틱 크리에이티브를 포괄해 프로그래매틱 광고(programming advertising)라고 일컫음

□ 생성형 AI 발전이 마케팅에 미치는 영향

- 최근 특히 급속한 발달로 주목받는 생성형 AI는 고객 경험(CX) 강화, 효율성 증가, 개인화의 중요성에 영향을 미칠 가능성이 있음
  - 생성형 AI는 개별 고객 행동, 인물상 및 매 이력에 기반하여 하이퍼-개인화(hyper-personalization)된 콘텐츠와 추천을 가능하게 함으로써 고객 경험을 극대화하고, 브랜드 충성도를 높여 고객 경험 강화
  - 정확한 분석과 고객 인사이트를 기반으로 성과를 향상하고 판매와 소통 자동화로 고객과의 상호작용을 제고해 퍼포먼스 증대
  - 기업 특정 데이터와 맥락의 결합을 통해 마이크로 소비자 인사이트를 제공하고 B2C 타겟 마케팅과 제품 추천의 하이퍼-개인화 실현을 통해 개인화와 퍼포먼스 향상
- 2020년 이미 32%의 광고 실무자가 디스플레이, 동영상 등 다양한 형태의 광고 제작을 위해 AI를 활용한 것으로 나타남

[그림 II-22] AI를 이용하는 마케터 비중과 이용 용도 현황



자료: MarketingCharts (2020)

□ 소비자 구매 여정에 따른 AI 활용

- AI는 광고와 마케팅 기획과 수행의 중요한 틀인 소비자 구매 여정(customer journey)의 모든 단계에서 활용됨
  - 즉 필요/욕구 인지, 초기 고려, 적극적인 평가, 구매, 구매 후 행동으로 이어지는 소비자 구매 여정의 각 단계에서 AI를 활용해 광고의 효율성을 증대할 수 있음(Kietzmann, et al., 2018)

〈표 II-2〉 소비자 여정 단계별 AI 활용 방안 및 사례

단계	필요 / 욕구 발견	초기 고려	적극적 평가	구매 결정	구매 후 행동
소비자 행동	필요 / 욕구 인지	가능한 솔루션 검토	선택 대안 비교	가치 평가	구매 결정 평가
광고 목표	소비자 필요 욕구 탐지	목표 소비자 도달	신뢰 획득 소비자 설득	제공가치 강화 및 보증	재구매 유도
전통적 광고활동	고객정의 니즈/욕구 이해	가치 명제 제시	정보 접근 촉진	혜택 입증 구매 인센티브 제공	고객관계관리 브랜드 충성 동 제 고
AI 광고 활용 방안 및 사례	심층적인 소비자 이해	AI 기반 광고 타게팅 잠재 고객 예측	구매 가능성 점수화 광고 메시지 개인화 감성인공지능	구매 지능화 동적 가격책정 광고 리타게팅	소비자지원 챗봇 성향 모델링 1:1 CRM
	Pinterest Asko	Zendesk	Asos Kellogg's	Staples Amazon Tumi	Autodesk Wordsmith

자료: Kietzmann, Paschen, & Treen (2018, p. 265); 박종구(2021, p.174) 재인용

## (2) 광고의 생성형 AI 활용

### (가) 생성형 AI 활용 광고 증가

#### □ 생성형 AI 활용 광고 유형

- 광고·마케팅 분야의 생성형 AI를 활용한 캠페인 집행이 증가하며 다양한 시도가 이루어지고 있음
  - 혁신적인 브랜드 이미지를 형성하고 소비자의 이목을 끄는 데 효과적임
  - 소비자와 실시간으로 소통하며 즉각적으로 이미지를 생성하는 캠페인이 가능해짐
- 삼성생명의 '좋은 소식의 시작'은 업계 최초로 AI로만 제작된 영상광고로서 주목받음
  - 생성형 AI 활용 초기 단계임에도 불구하고 광고의 모든 요소를 AI를 활용해 제작하는 데 성공했지만 이에 못지않은 인력이 투입됨

[그림 II-23] AI 접목 광고 캠페인 사례 유형



자료: 메조미디어 (2023)

### (나) 광고에 활용되는 주요 생성형 AI 도구

#### □ 텍스트 생성형 AI

- 이용자와의 대화 내용을 결과에 통합하여 실제 대화를 모방하는 능력을 갖추고 있는 것이 특징

- 대표적인 텍스트 생성형 AI로는 OpenAI 사의 ChatGPT, Microsoft의 BingAI, Google의 Bard가 있음
- 국내는 네이버 사의 '클로바 X'가 네이버 및 외부 서비스 연계를 통한 생태계 확장으로 대화형 AI 서비스에 대응
- 광고에서는 스토리라인을 기획하거나 스크립트나 카피를 제작하는 데 활용됨

#### □ 이미지 생성형 AI

- 이용자의 명령에 따라 이미지를 기존 데이터베이스를 참조하여 창조, 혹은 합성하는 기술
- 2014년 Ian Goodfellow가 제안한 적대적 생성 신경망(GAN)을 기반으로 생성기와 판별기의 두 네트워크를 경쟁적으로 학습함으로써 진짜와 유사한 이미지를 생성할 수 있는 기술 토대 마련
- 스테빌리티 AI의 Stable Diffusion, OpenAI의 달리(DALL-E), 미드저니(Midjourney) 등이 대표적

#### □ 동영상 생성형 AI

- 텍스트, 이미지를 넘어서 동영상을 생성해주는 AI에 대한 관심도 이용이 급증하고 있음
- 대표적 서비스로는 메타의 메이크어비디어와 예뮤 비디오, 구글의 페나키와 이매젠 비디오를 비롯해 스테빌리티 AI, 런웨이 리프레이즈, 피카랩스, 카이버, 레이아픽스 등이 출시한 서비스도 다수
- 온라인 트래픽의 80% 이상이 동영상 트래픽일 만큼 텍스트나 이미지에 비해 동영상 콘텐츠 선호도가 높기 때문에 활용도가 높음

#### □ 텍스트 이미지 통합형 및 특수 목적 생성형 AI

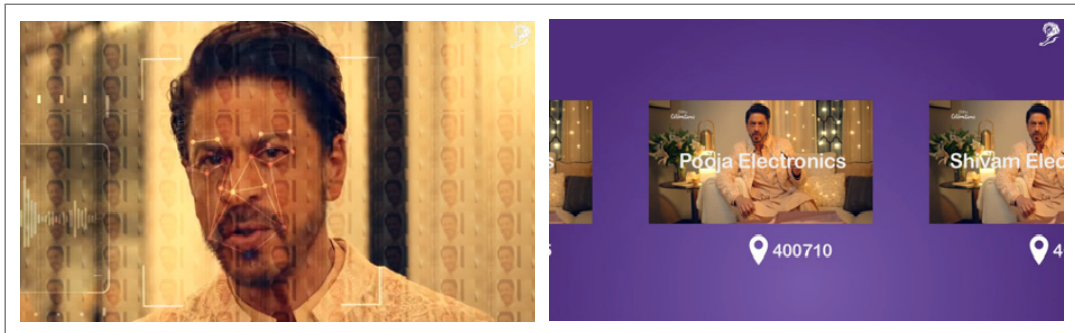
- 딥러닝 기술을 활용하여 음성 전사, 통합형 멀티 모달, 혹은 맞춤형 서비스를 개인 혹은 엔터프라이즈 대상으로 제공하는 다양한 생성형 AI가 IT 기업들을 중심으로 출시되고 있음

(다) 생성형 AI 활용 광고 캠페인 사례

□ 캐드버리(Cadbury)의 샤룩칸-마이-애드(Shah Rukh Khan-My-Ad) 캠페인

- 영국의 제과업체 ‘캐드버리’가 AI를 활용해 인도에서 집행한 캠페인
- 대기업들이 유명인을 고액으로 광고에 기용하는 것과 대조적으로, 일반 소상공인들에게는 그러한 기회가 사실상 불가능하다는 점에 착안
  - 유명한 발리우드 배우이자 가수인 샤룩 칸(Shah Rukh Kahn)의 목소리와 얼굴을 AI에게 훈련시키는 딥러닝 과정을 거침
  - 지역 상점들의 이름들을 생성형 AI로 샤룩 칸이 읽어낼 수 있게 하였음
- 다양한 매장의 브랜드와 제품을 알리는 다양한 버전의 광고를 생성, 각 지역별 핀코드를 기준으로 타겟 마케팅이 가능한 서비스 제공
- 칸느 라이온즈 2023에서 크리에이티브 효과 부문 그랑프리 수상
- 해당 캠페인을 통해 13만개의 광고 생성되었으며, 3천만회 이상의 광고 노출 달성, 35%의 기업 성장률을 기록

[그림 II-24] 캐드버리 캠페인



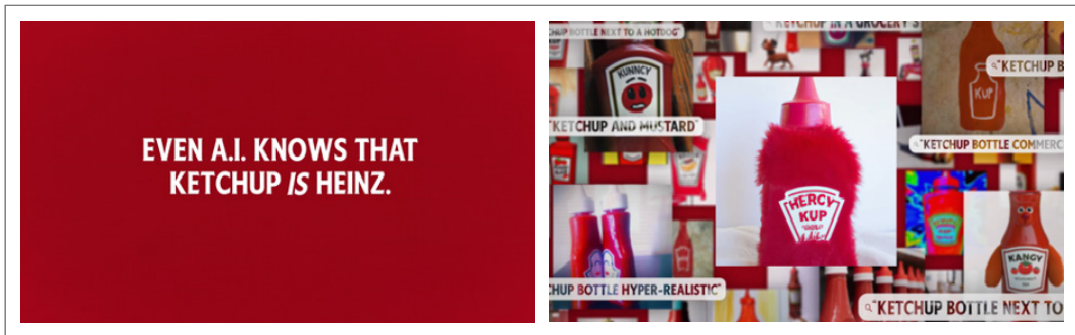
자료: Cadbury – Shah Rukh Khan-My-Ad (case study) (2023)

□ 하인즈의 ‘A.I. 케첩’ 캠페인

- 칸 라이언즈 2023 소셜 & 인플루언서 실버 부문 수상
- 케첩의 대명사 ‘하인즈’는 텍스트-이미지 변환 생성형 AI 기술 Dall-E 2 모델 이용, AI가 생각하는 케첩은 어떠한 형태인지 이미지를 생성하도록 명령

- 그 결과 케첩이라는 키워드를 받은 AI는 하인즈 케첩의 용기의 형태를 띄고 있는 케첩의 이미지를 다수 생성
- “Even A.I. knows that ketchup is Heinz.”라는 카피와 함께 케첩은 곧 하인즈라는 이미지를 AI 기반 실험을 통해 공고히 함
- 머신러닝과 딥러닝이 기반인 생성형 AI 기술을 활용한 흥미로운 사례
- 하인즈가 AI에게 ‘케첩’이라는 키워드를 제공했을 때, AI가 생성한 이미지들이 사회에서 통용되는 이념에 따라 나온 결과 도출

[그림 II-25] 하인즈 캠페인

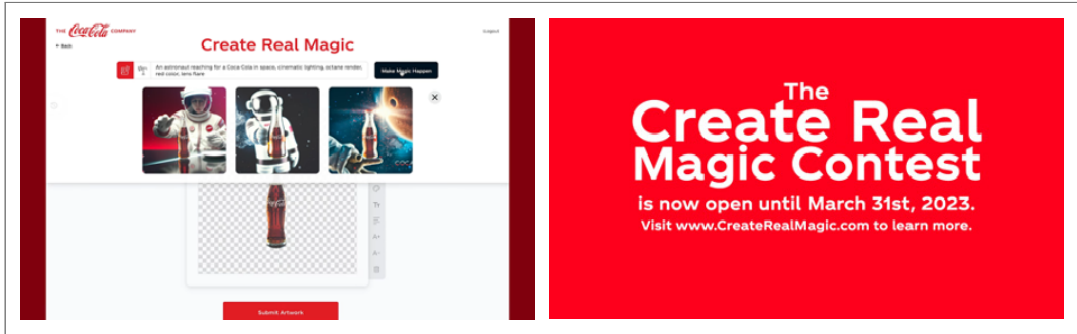


자료: HEINZ: A.I. Ketchup (Cannes 2023) (2023)

□ 코카콜라의 ‘AI 아트 경연대회’ 캠페인

- 일상의 경험을 특별한 것으로 변화시키는 예상치 못한 연결의 순간에서 마법을 발견한다는 ‘리얼 매직’이라는 글로벌 브랜드 플랫폼을 연계, 생성형 AI를 이용한 창작과 연결하여 ‘리얼 매직’ 공모전을 진행
- OpenAI, Bain & Company와 파트너십을 체결하여 오랜 역사를 자랑하는 코카콜라의 상징적 요소들을 이용하여 독창적인 예술 콘텐츠를 제작할 수 있도록 소비자에게 플랫폼을 제공
- 세계 점유율 1위 탄산음료 제조사 코카콜라가 AI 기술을 브랜드의 글로벌 마케팅 전략에 통합하려는 시도를 통해 혁신적인 이미지 획득
- AI 기술과 전통적인 브랜딩 요소의 독특한 융합을 보여준 사례

[그림 II-26] 코카콜라 캠페인

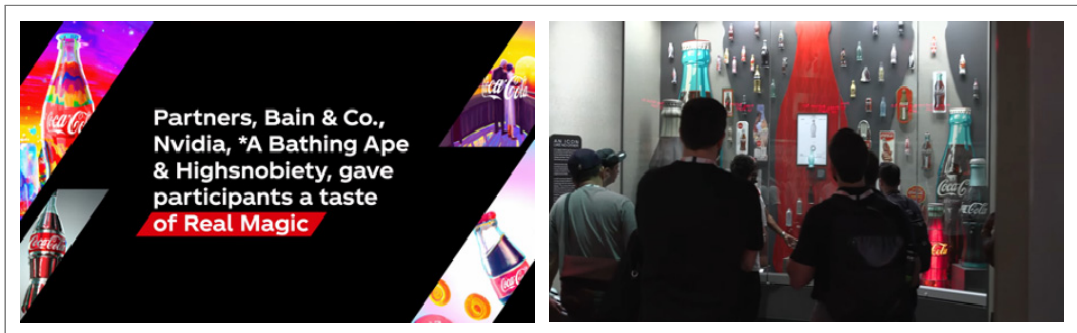


자료: Coca-Cola® | Create Real Magic. (2023)

□ 코카콜라의 '리얼 매직 크리에이티브 아카데미'

- 애틀랜타에 위치한 코카콜라 본사에서 3일 동안 '리얼 매직 크리에이티브 아카데미' 개최
  - 창의성 및 생성형 AI 혁신 포럼을 개최하여 디지털 아티스트들이 브랜드 및 다른 파트너들과 함께 배우고 협업할 수 있는 기회 제공
- 최신 생성형 AI 콘텐츠 및 기술 기반 아이디어 공동 창작 논의의 장 형성
- 캠페인을 통해 소비자 참여와 아티스트 협업을 결합함으로써, AI 기술과 기업의 상생 가능성을 보여주는 사례
- 창의적 영역에 위협을 느끼는 아티스트들에게 대화의 기회를 제공, AI 활용 가능성 확장과 동시에, 기업과 AI의 공존을 탐색하는 전략으로 평가됨

[그림 II-27] 코카콜라 캠페인 II



자료: Coca-Cola | Real Magic Creative Academy. (2023)

### (3) AI 기반 광고 플랫폼 현황

#### (가) 구글

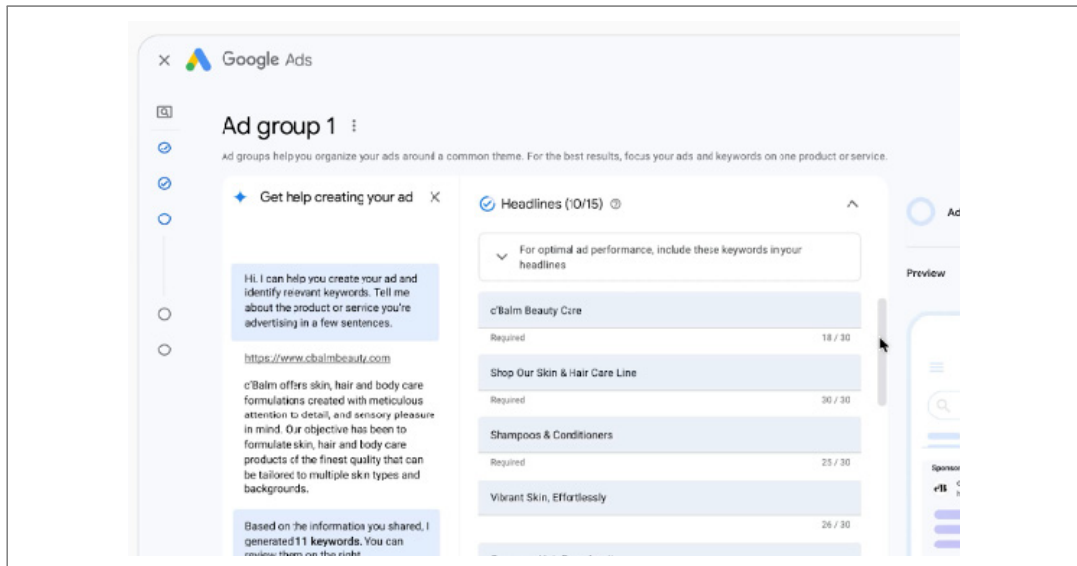
##### □ 구글 바드(Google Bard)와의 연계

- 구글은 기존에 운영하던 구글 애즈 서비스에 구글 바드(Google Bard) 기반 텍스트 생성 서비스를 접목하는 시도 진행

##### □ 자연어 대화 환경 구축

- 이용자가 채팅을 통해 복잡한 서비스도 쉽게 이용할 수 있도록 도움
- 전문 지식과 AI를 결합하여 캠페인 생성을 빠르게 시작하고 검색 광고를 간소화할 수 있도록 설계
- 웹사이트의 랜딩 페이지를 분석하여 적합한 키워드, 헤드라인, 설명, 이미지 등을 제안

[그림 II-28] 구글의 랜딩 페이지 텍스트 생성형 AI 인터페이스



자료: Google (2023)

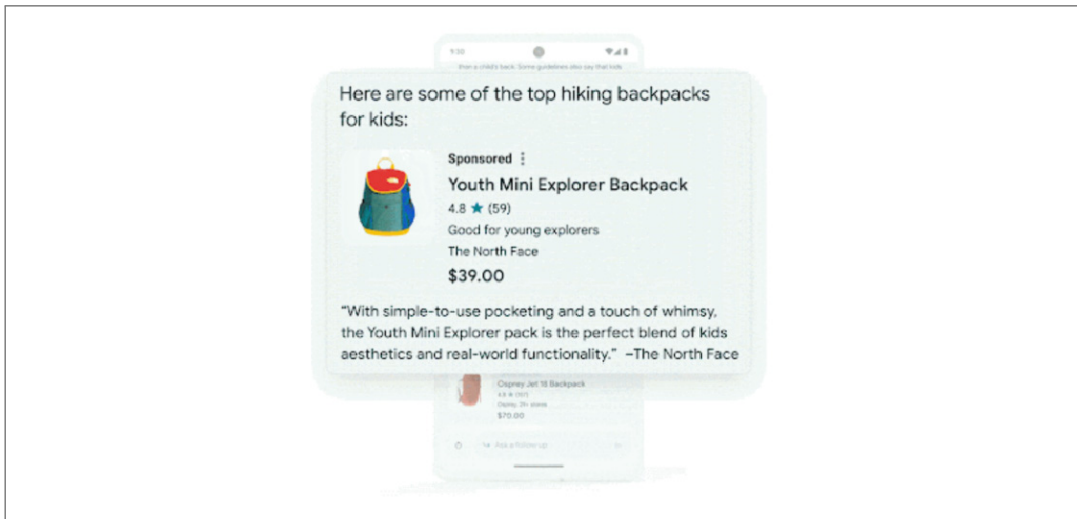
□ Automatically Created Assets(ACA)

- 자동 생성된 요소들을 활용해 광고 제작을 용이하게 함
- AI를 사용하여 랜딩 페이지와 기존 광고 콘텐츠를 기반으로 헤드라인과 설명을 자동 생성
- 검색어의 문맥에 따라 검색 광고를 더욱 효과적으로 맞춤 설정하여 브랜드에 충실하면서 광고의 관련성을 높이는 데 도움

□ Search Generative Experience(SGE, 적용 예정)

- 생성형 AI로 새로운 검색광고 경험 제공
  - AI를 사용하여 검색 및 쇼핑 광고를 맞춤화하고, 관련성 높은 고품질 광고를 생성하는 새로운 형식 실험 중
  - 검색을 간소화하고 사용자를 최적의 웹사이트로 연결하는 데 도움 주는 것이 목적
- 예를 들어, '마우이에서 할 수 있는 야외 활동'을 검색한 사용자가 검색 범위를 더 좁혀 '어린이를 위한 활동'과 '서핑'에 대해 질문할 때 어린이를 위한 서핑 강습을 홍보하는 여행 브랜드의 맞춤형 광고 표시

[그림 II-29] 구글의 자연어를 활용한 검색 경험(SGE)



자료: Google (2023)

### □ 생성형 AI를 통한 광고 제작 지원

- AI를 활용해 광고 카피, 이미지, 동영상을 제작하고 다양한 포맷과 버전을 생성해 다양한 소비자에게 맞춤형으로 제공할 수 있음

### (나) 메타

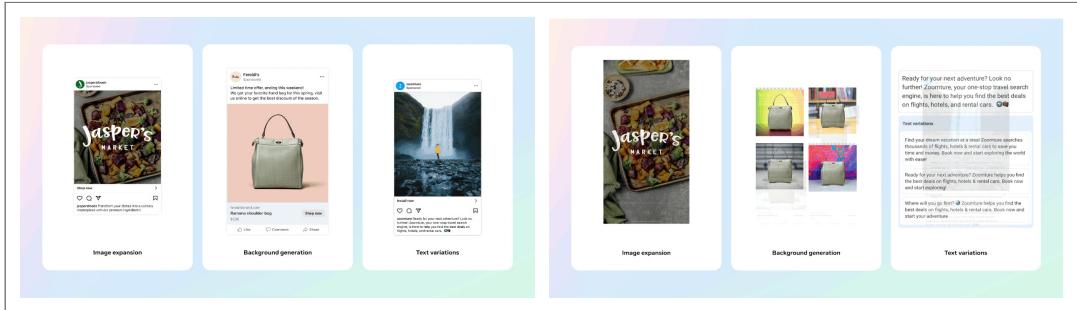
#### □ AI 기술과 마케팅 접목

- Meta는 2006년 Facebook 뉴스피드를 시작으로, AI 인프라스트럭처, AI 모델, 마케팅 경험 개선 등 크게 3가지 분야에서 AI 연구를 집중하고 있음을 밝힘
  - AI 인프라스트럭처 및 모델 고도화를 통해 세분화된 데이터 접근 없이도 정교한 성과 측정이 가능토록 함
  - 캠페인 전반을 학습하여 콘텐츠의 품질과 경험을 제고하고 성과를 향상시킴

#### □ AI 기반 새로운 마케팅 솔루션 제공

- 원클릭 자동화 캠페인 전환 지원
  - 클릭 한 번으로 기존 수동식 캠페인을 AI 기반 자동화 캠페인으로 전환
- 자동 vs. 수동 캠페인 성과 비교
  - AI 기반 캠페인과 수동식 캠페인의 성과를 비교할 수 있는 대시보드를 제공
- 자동화 캠페인으로 영상 소재 지원 (시범 운영)
  - 자동화된 캠페인 소재로 사진뿐만 아니라 영상도 활용할 수 있도록 지원 예정
- Meta Advantage+ Audience (시범 운영)
  - AI를 활용해 마케팅 타겟팅을 고도화하는 새로운 커머스 기능
- AI 샌드박스 (시범 운영)
  - 다양한 생성형 AI 기반 마케팅 툴을 실험해볼 수 있는 공간
  - 마케팅 문구 자동 생성, 마케팅 소재의 배경 이미지 자동 제작, 콘텐츠 비율 자동 재가공 등으로 활용 가능
  - 이를 통해 기존의 광고 플랫폼에 AI를 기반으로 배경 생성, 이미지 확장, 텍스트 다양화 기능을 추가할 예정

[그림 II-30] 메타 시 활용 광고 플랫폼



자료: Meta (2023)

(다) 아마존

□ 개인화 활용 상품 추천 서비스

- 결제 상품을 토대로 다양 한 배너에 구매 확률이 높은 상품 노출
  - 예측 알고리즘을 활용하여 효과적으로 개인화 경험을 제공하는데 목적

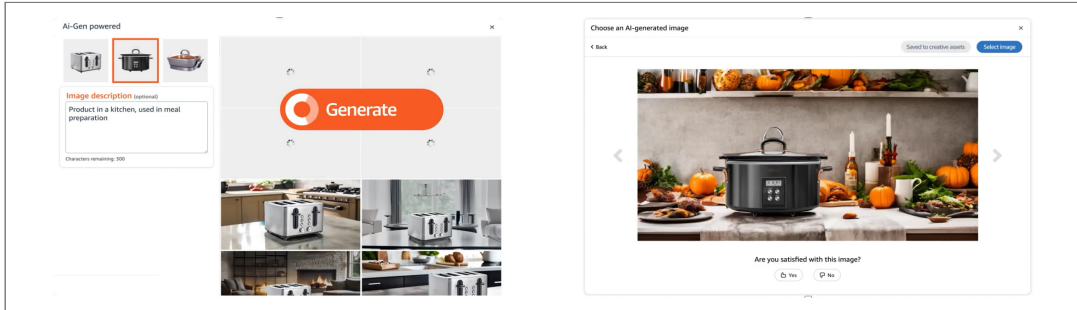
□ AI 판매자를 위한 상품 작성 AI 서비스

- 대형언어모델(LLM)을 전자상거래 플랫폼에 통합, 자동으로 상품 설명을 작성하는 생성 AI 도구 도입
  - 이를 통해 수고를 줄이는 것은 물론 정확한 설명 작성은 판매 증가와 반품 감소에도 기여할 수 있을 것으로 예측

□ 브랜드를 위한 AI 기반 이미지 생성 기능

- 이미지 생성형 AI 기술을 활용한 상품 이미지 제작 AI 솔루션 도입
- 생성형 모델과 광고주의 상품 정보를 사용하여 추가 비용 없이 브랜드와 관련된 고유한 브랜드 테마 상품 이미지 생성 지원
- 아마존 광고 콘솔에서 상품을 선택하고 '생성'을 클릭하면 자동으로 전문적인 상품 이미지 생성
  - 광고주의 지식과 전문성을 강화하는데 도움이 되는 일종의 크리에이티브 어시스턴트 제공
  - 카피, 이미지, 비디오 제작을 위해 필요한 리소스에 액세스할 수 없는 광고주에게 크리에이티브 제작의 어려움을 해소할 수 있는 AI 활용 서비스 제공

[그림 II-31] 아마존 AI 활용 광고 플랫폼



자료: Amazon (2023)

(라) 아도비

□ Sensei 서비스

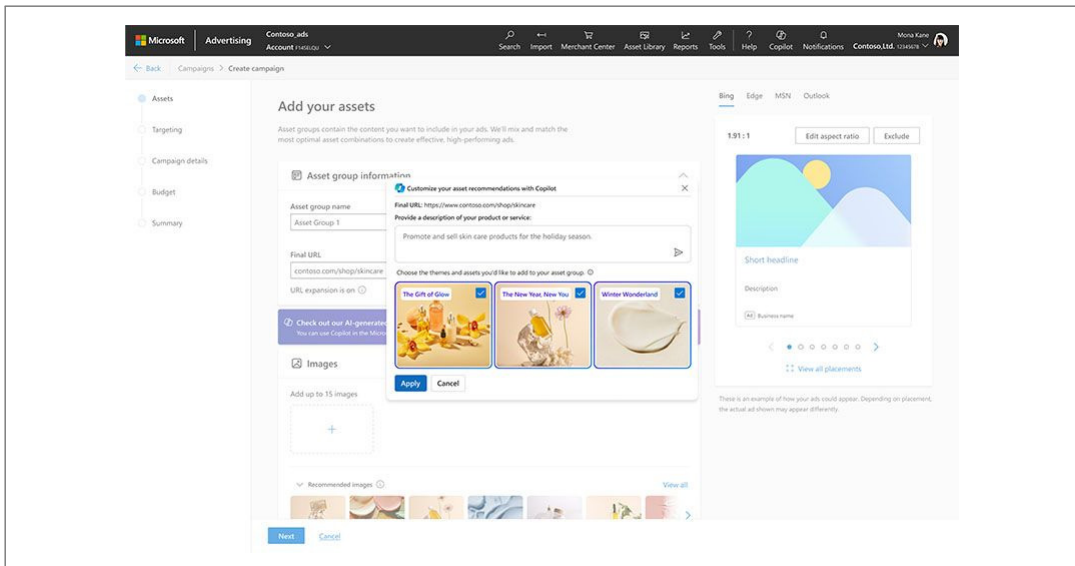
- AI 기반 서비스인 ‘Sensei’ 는 다양한 분야 마케팅 전문가들이 개인 소비자에게 최적화된 광고를 제작하고 효과적으로 전달할 수 있도록 창의적이고 매끄러운 워크플로우 제공에 초점을 둠
- 크리에이터에게 창의적인 작업에 더 집중할 수 있도록 도움을 줌
  - 원하는 콘텐츠 탐색 및 영상 소스의 불필요한 요소 삭제 등의 합성이 가능한 기능 제공
- 마케터가 타겟 마케팅과 개인화 마케팅 전략을 구현하여 ROI를 향상할 수 있도록 함
  - 다양한 예산 시나리오를 통해 광고 효과 예측 가능
  - 타겟팅 기준에 따라 자동으로 예산 조정 및 캠페인 성과에 대한 지속적인 모니터링을 통해 광고 전달 최적화
- 데이터 애널리스트에게 데이터를 이해하고 중요한 인사이트를 도출할 수 있도록 함
  - 실시간 데이터 변화 알림 및 특성, 차이점, 전환 요소를 기반으로 고객 행동 예측 서비스 제공
- 광고업체에게 광고 예산을 효율적으로 관리하고 캠페인의 성과를 최적화 할 수 있도록 함
- 비즈니스 전문가에게 비즈니스 프로세스 가속화와 디지털 문서 작업을 간소화하도록 도움

(바) 마이크로소프트

□ 코파일럿(Copilot)

- 마이크로소프트 AI 기능들은 ‘마이크로소프트 코파일럿’ 명칭으로 단일 환경으로 통합되어 윈도우 11(Windows 11), 마이크로소프트 365(Microsoft 365), 엣지(Edge) 및 Bing(Bing)에 적용되고 애플리케이션과 디바이스 전반에 걸쳐 보다 매끄러운 AI 경험 제공
  - 이용자의 모든 작업을 도울 수 있는 ‘부조종사’의 의미로서, 코파일럿은 채팅 인터페이스와 대규모 언어 모델(LLM)을 통합해 일상에서 통용되는 자연어로 원활히 소통할 수 있으며 이용자의 요청에 즉각 응답함
- 광고주와 대행사를 대상으로 특별히 제작된 코파일럿은 AI의 힘을 빌려 제품 이미지, 헤드라인 및 설명을 위한 새롭고 혁신적인 제안을 생성하는 도구를 제공함
  - 이용자들이 이미지와 같이 필요한 콘텐츠에 대한 설명만 제공하면, 코파일럿이 이러한 설명을 기반으로 자동으로 광고 캠페인 생성

[그림 II-32] 마이크로소프트 AI 활용 광고 플랫폼

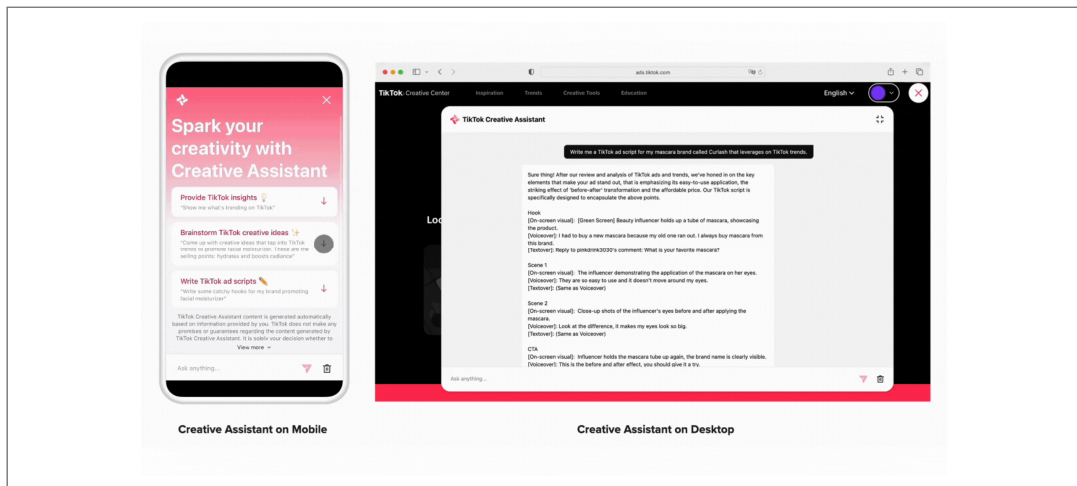


(사) 틱톡

□ 크리에이티브 어시스턴트(Creative Assistant)

- AI 기반의 새로운 도구인 크리에이티브 어시스턴트는 틱톡 크리에이티브 센터의 도구와 연계해 관련 예시, 도구에 연결하고 샘플 광고 스크립트를 제작하는 등 캠페인 제작 과정의 다양한 도움을 마케터에게 제공함
  - 틱톡에 초점을 둔 풍부한 크리에이티브 데이터에서 정보를 추출하여 틱톡용 광고 또는 동영상 제작할 때 가장 관련성 높은 답변을 제공
  - 예를 들어, 마케터가 특정 분야의 인기 광고 캠페인 사례와 모범 사례를 찾고 광고 스크립트 초안을 작성할 때 챗팅을 통해 구체적인 도움을 요청할 수 있음

[그림 II-33] 틱톡 AI 활용 광고 플랫폼

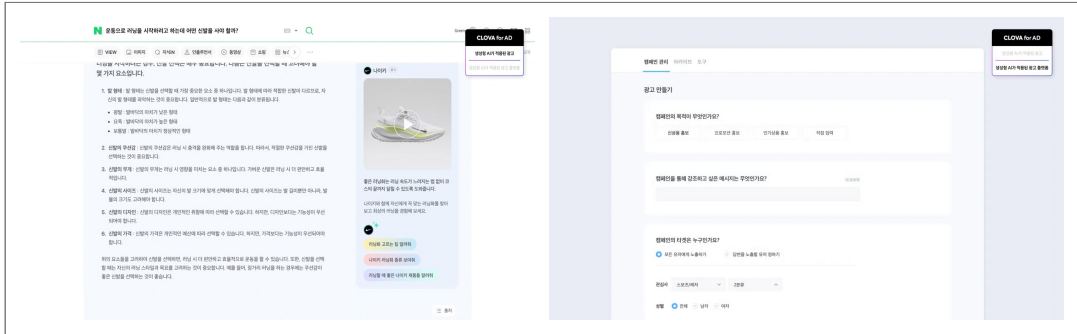


(아) 네이버

□ 소비자 맞춤형 광고 플랫폼 제공

- 자연어 기반 검색엔진과 하이퍼클로버를 연계하여 소비자의 상품 추천 요구에 따라 고려사항 및 추천 상품을 포함한 답변 생성 가능
- 판매자에게는 AI를 활용해 캠페인 목적 및 강조 메시지에 따라 답변 노출 타이밍 등을 설정할 수 있도록 함

[그림 II-34] 네이버 AI 활용 광고 플랫폼

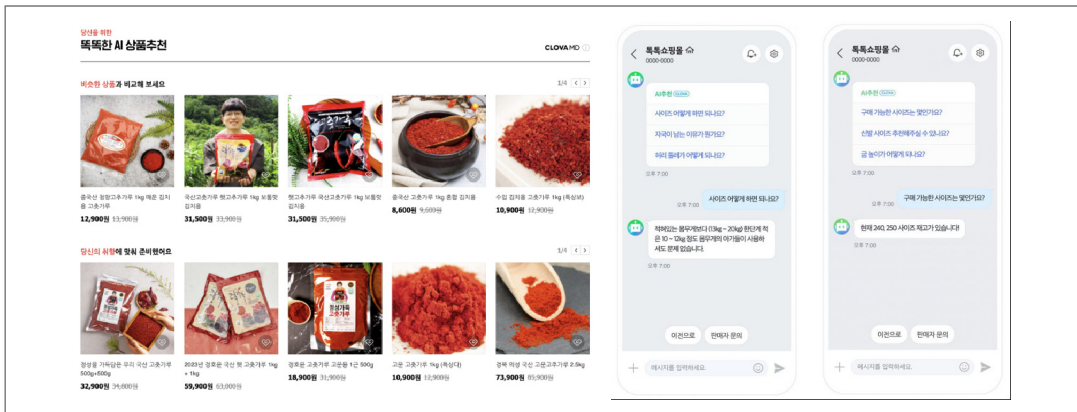


자료: 네이버 (2023)

□ 스마트스토어 판매자 온라인 사업 확장 및 직거래(D2C) 강화

- 자사 AI인 하이퍼클로바가 고객 문의를 분석해 소비자가 자주 하는 질문을 추천해주는 ‘클로바 라이브챗’ 기능으로 고객 서비스 효율을 높이도록 함
- 우수한 리뷰를 상품 페이지에 활용하는 ‘리뷰 노출 AI 매니저’는 구매 전환 유도 - 전반적인 소비자 조사 및 데이터 분석 솔루션 서비스 제공

[그림 II-35] 네이버 스마트스토어 클로바MD 적용 및 클로바 라이브챗 예시



자료: 네이버 (2023)

□ 하이퍼클로바 기반 마케팅 솔루션 성과

- AI가 타깃 고객에게 자동으로 메시지를 보내는 솔루션을 통해 마케팅 효과를 약 30% 제고

- ‘글로벌 라이브챗’과 ‘글로벌 메시지 마케팅’을 이용한 스마트스토어 입점업체는 솔루션 출시 1년 반 만에 7배 증가
- AI를 활용해 타깃한 고객이 메시지를 읽는 비율과 메시지 내 상품 정보나 스토어 등을 클릭하는 클릭률은 각각 이전 대비 44%, 31% 증가
- 광고 회사 HS애드와 커니 코리아(Kearney Korea)는 네이버의 클라우드로 하이퍼클로바 언어 모델을 이용할 수 있는 도구 ‘글로벌스튜디오’를 이용하여 생성 AI 마케팅 솔루션 개발 착수

□ SME 판매자 지원

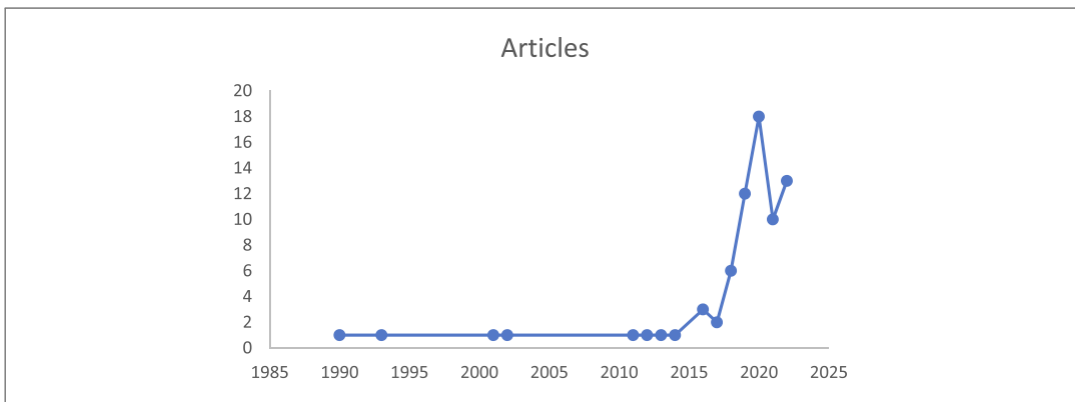
- 일손이 부족한 SME(smal and medium-sized enterprieses) 마케터와 브랜드 성장을 원하는 판매자를 중심으로 스마트스토어의 AI 솔루션 활용을 적극적으로 추진
- 이를 통해 지역 농촌 등 상품 판매자에게 하이퍼클로바 기술을 접목한 서비스로 스마트 판매 플랫폼 지원을 확대하고 있음

라. AI 활용 광고 관련 연구

□ AI 광고 연구 현황

- 1990년부터 2022년까지 AI 광고에 대한 연구가 게재된 학술논문을 분석결과, 최근 AI 광고 관련 논문의 증가가 두드러짐(Ford et al., 2023)

[그림 II-36] AI 광고 관련 논문 수



자료: Ford et al. (2023)

- 최근 몇 년 동안 AI의 발달과 함께 광고 연구는 주로 프로그래매틱 광고, 광고 캠페인 최적화, 광고 성과 측정, 광고 창의성을 다룸

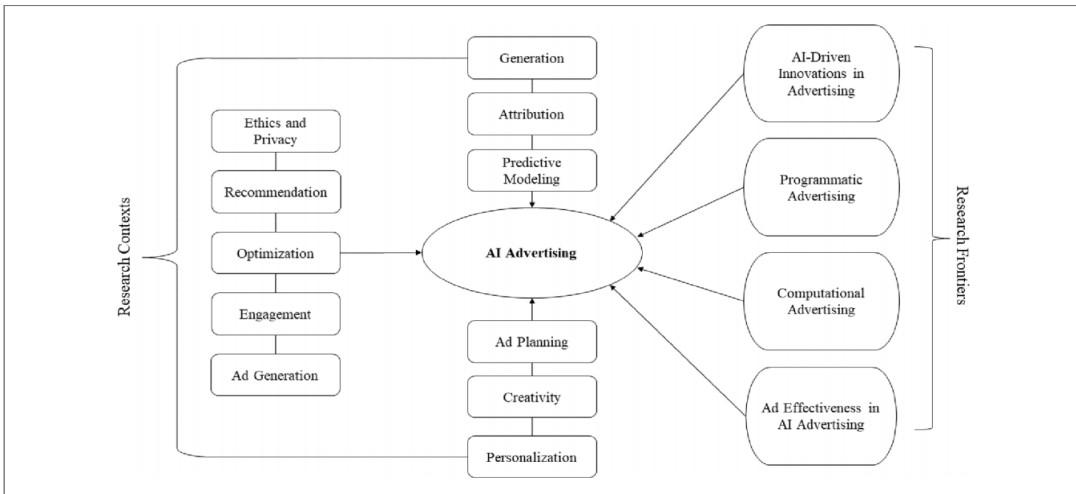
[그림 II-37] AI 광고 관련 논문 워드 클라우드



자료: Ford et al. (2023)

- 더 구체적으로 AI 광고 연구는 프라이버시와 윤리, 추천, 최적화, 인게이지먼트, 광고 생성 맥락에서 광고의 AI 기반 혁신, 프로그래매틱 광고, 컴퓨터 기반 맞춤형 광고, AI 광고의 효과 등의 주제를 다루는 그룹으로 크게 나뉘짐

[그림 II-38] AI 광고 연구 맥락과 주제



자료: Ford et al. (2023)

- AI의 발전이 광고 생태계와 소비자와 실무자의 광고 경험을 제고함에 따라 연구자와 실무자들에게 AI 광고와 관련해 혁신, 창의성, 성과, 윤리를 함께 고려하고 조화를 추구하는 시각과 접근방식이 요구됨

## 마. 광고산업에 AI가 미치는 영향

### (1) 생성형 AI가 가져올 광고 패러다임의 변화

#### □ 광고 패러다임의 현재와 미래

- 네이버 'CLOVA for AD'가 제시하는 비전을 기반으로 생성형 AI가 바꿀 광고 패러다임에 대해 전망할 수 있음

#### □ 인터랙티브 관계 구축의 중요성

- 새로운 광고 패러다임은 일방향으로 전달하던 기존 광고의 형태를 넘어서 이용자와 상호작용 관계를 구축할 수 있도록 도와줌

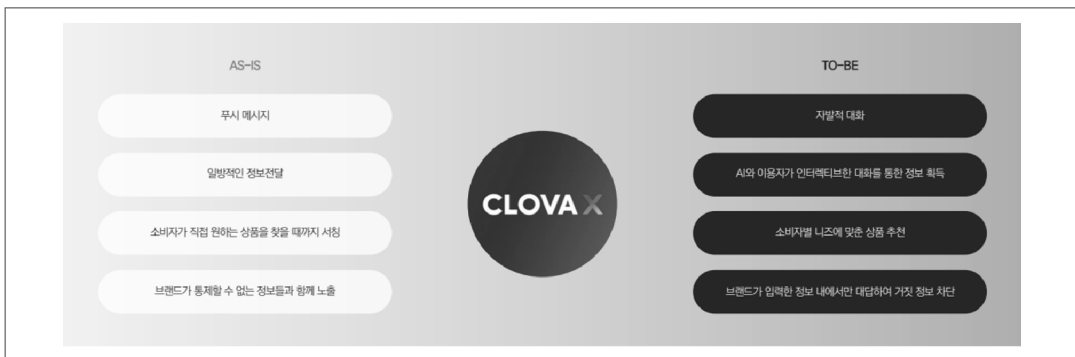
#### □ 자연어 기반 생성형 AI 활용도 증가

- 자연어 기반의 대화형 AI는 검색창에 소비자가 질문을 하거나 원하는 정보를 요청하면 필요한 정보만 취합하여 답변과 추천을 제공하고, 후속 질문과 정보를 제안하며, 개인의 선호·취향·사용 목적을 반영해 최적화된 상품을 추천

#### □ 타겟팅 없이 맞춤형 정보 제공

- 마케터는 별도의 타겟팅 없이도 소비자에게 개인 맞춤형 정보를 제공할 수 있으며 답변 내용은 모두 광고주가 입력한 정보 안에서 제공되기 때문에 brand safety가 보장되는 것도 차별점임
- 이를 통해 브랜드의 의도에 맞는 광고를 이용자에게 이론상 무한히 생성 가능

[그림 II-39] 네이버 클로바X 활용 현재와 미래



자료: 메조미디어 (2023)

## (2) 생성형 AI와 광고 창의성

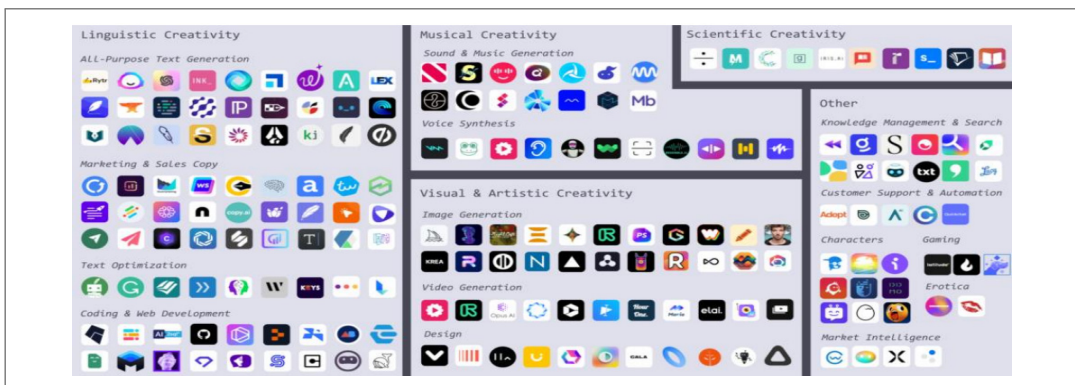
### □ 칸 라이언즈 2023

- 대표적인 국제 광고 페스티벌인 칸 라이언즈 2023의 랩업(Wrap-up) 보고서는 창의적인 프로세스에 AI가 통합되는 현상에 주목
- 2023 칸 라이언즈 출품작의 과반수가 어떤 형식으로든 AI를 활용했고 전체 수상작 중 8.3% 이상이 AI를 활용한 것으로 나타남
  - 크리에이티브 분야에서 AI의 영향력이 증가하고 있음을 보여줌
  - AI에 대한 두려움보다 개인의 호기심과 창의성을 발현하기 위해 무한한 가능성을 가진 AI의 잠재력을 인식하고 활용할 것을 강조(구글 크리에이티브 랩의 로버트 워)

### □ 생성형 AI와 인공 창의성

- ‘인공 창의성(Artificial Creativity)’은 인간이 가진 창의성과 유사한 창의성을 발휘할 수 있는 인공지능 기반 프로그램의 개발을 목표로 인간과 기계, 창의성과 생산성의 관점에서 이해할 수 있는 개념이며 연구 분야임(박종구, 2023)
  - 텍스트, 이미지, 동영상, 음성, 음악 등 다양한 형태의 결과물을 생성하는 AI의 등장은 인공 창의성의 활용에 대한 논의를 촉발함(NessLabs, 2022)
- 인공 창의성 관련 창작 도구들은 시, 소설, 광고카피 생성 등이 가능한 ‘언어적’ 창의성, 영역, 이미지, 영상을 생성하는 ‘시각·예술적’ 창의성, 작곡과 연주 및 음성 합성 등을 포함하는 ‘음악적’ 창의성 등 창의성 관련 영역에 따라 구분 가능함(NessLabs, 2022)

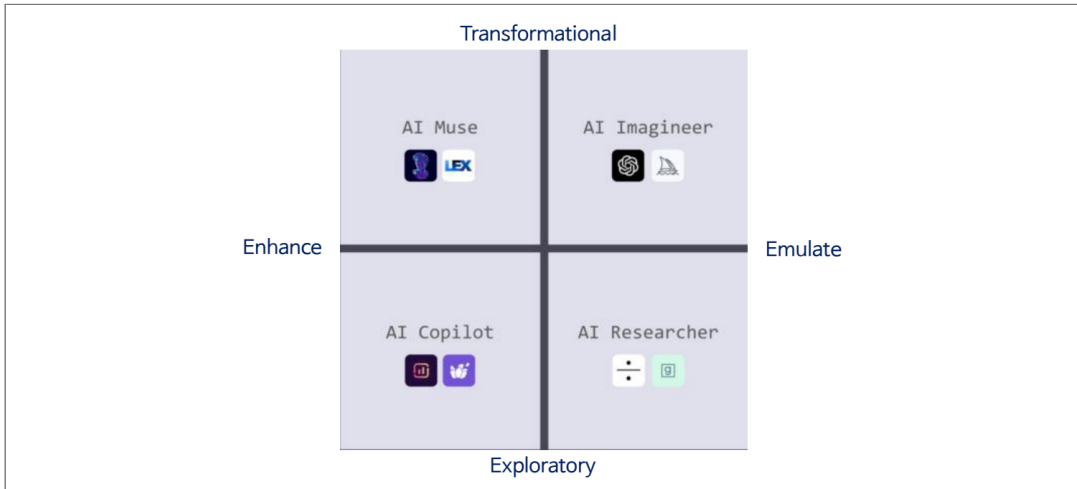
[그림 II-40] 인공 창의성 랜드스케이프



자료: NessLabs (2022)

- AI 생성도구는 인공 창의성의 목표(탐색적 또는 변형적)와 비전(인간 창의성의 향상 또는 모방)에 따라 4가지의 유형으로 구분됨
  - 먼저 인공 창의성의 목표는 탐색과 변형으로 구분됨. 탐색적 창의성은 기존의 규칙이 존재하는 창작 환경에서 정보를 탐색해 결과물을 생성하는 과정을 의미하는 반면, 변형적 창의성은 기존의 규칙에 구속받지 않고 자유롭게 변형하여 새로운 결과물을 생성하는 과정을 의미함
  - 한편, AI를 영감을 제공하고 인간의 창의성을 향상시키는 수단이거나 인간의 창의성 모방을 넘어서 자체적으로 새로운 창의성을 발현할 수 있는 존재로 인식함

[그림 II-41] 인공 창의성 도구 매트릭스



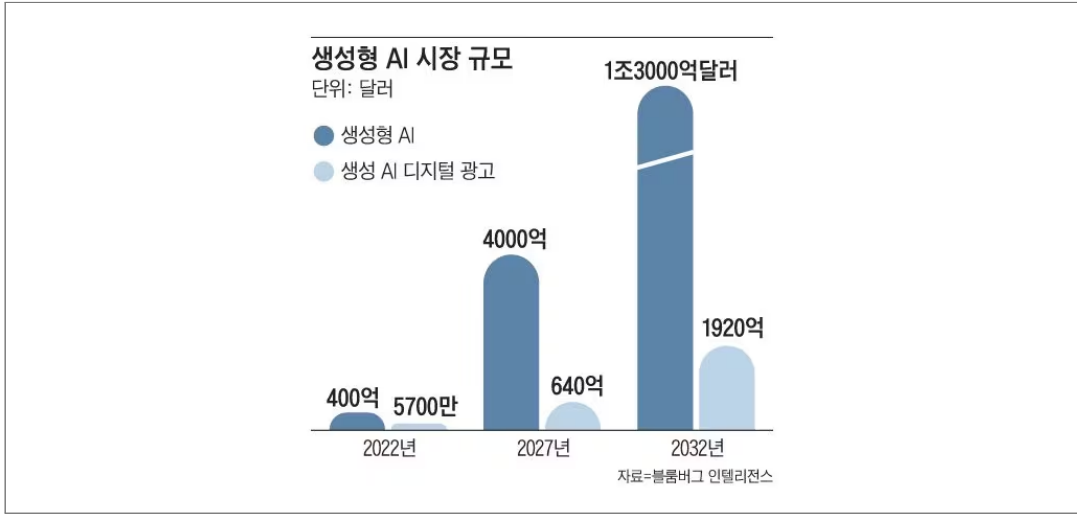
자료: NessLabs (2022)

### (3) 생성형 AI 광고시장 전망

#### □ 생성형 AI 광고시장 규모 확대

- 2022년 6천만 달러였던 생성형 AI 광고시장 규모가 매년 평균 125%의 성장을 보이며 2032년에 1천925억 달러(약 250조원)에 다다를 것으로 전망(메조미디어, 2023)
  - 블룸버그인텔리전스에 따르면, 생성형 AI 시장은 지난해 400억달러(약 53조원) 규모에서 연평균 42% 성장해 10년 뒤인 2032년 1조3000억달러(약 1740조원) 규모로 성장할 것으로 전망

[그림 II-42] 생성형 AI 시장 규모



자료: 이해인 (2023. 8. 24.)

- 생성형 AI 이용의 보편화에 따라 다양한 콘텐츠 제작이 증가할 것으로 기대됨
  - 온라인 설문에 참여한 만 15~59세 남녀 1천 명 중 약 40% 10명 중 4명은 이미 생성형 AI를 이용해본 경험이 있다고 응답함(메조미디어, 2023)
  - 특히 10대(49%)와 20대(46%)의 생성형 AI 이용 경험 비중이, 다른 연령대에 비해 높게 나타남(30대 37%, 40대 42%, 50대 36%)

□ 생성형 AI 광고 성장 원인과 우려

- 메조미디어(2023)는 “개인화에 대한 소비자 니즈가 높아지면서 광고 마케팅에도 생성형 AI의 활용은 더욱 많아질 것”이라며 “생성형 AI를 활용하면 창작과 운영에 드는 비용과 시간을 절감해 캠페인의 효율을 극대화할 수 있을 것”이라고 전망
  - 기업들이 경쟁적으로 광고에 생성형 AI를 접목하는 것은 이러한 광고가 제작 시간과 비용이 적게 소요되기 때문임
  - 예를 들어, 세계 최대 광고 대행사 WPP는 네슬레의 요구르트 제품 광고를 만들면서 요하네스 페르메이르의 그림 ‘우유 따르는 여인’을 AI로 변형한 영상을 활용함. WPP 관계자는 “저작권이 없는 이미지를 활용했고 콘텐츠를 만드는 데 비용이 거의 들지 않았다”고 함(이해인, 2023. 8. 24.)

- 현재 활용 가능한 생성형 AI의 문제는 부정확성과 편견과 고정관념 재생산 혹은 강화 가능성임
  - 제이 파티살 애널리스트는 미국 CNBC에서 “생성형 AI로 광고를 만드는 경우 대중에게 불쾌함을 주고 부정확한 시각 자료와 문구를 생성할 수도 있다”며 “기업을 특정 분야에서 편향적으로 보이게 하는 등 기업 브랜드 안전성을 해칠 가능성도 있다”고 우려를 표함(이해인, 2023. 8. 24.)
  - 실제로 앞서 살펴본 삼성생명의 “좋은 소식의 시작” 광고영상은 스토리, 모델, 이미지, 음악까지 모든 측면에서 AI를 활용해 제작되었는데, 모델의 신체 일부나 표정이 어색하다는 부정적인 평가를 받기도 함(신준용, 2023. 7. 24)

#### □ 광고에 특화된 생성형 AI 솔루션 개발

- 현재 생성형 AI의 한계를 극복하고 광고 영역에서의 활용성을 제고하기 위해 광고 영역에 특화된 생성형 AI 혹은 개별 기업의 맞춤형 AI를 도입하거나 개발하려는 시도가 증가할 것으로 전망
  - 예를 들어, 광고대행사 HS애드는 네이버클라우드의 AI 개발도구인 ‘클로바 스튜디오(CLOVA Studio)’를 활용하여 생성형 AI 기반 마케팅 솔루션 개발에 착수함. HS애드의 광고 전문가들이 그동안 축적한 양질의 콘텐츠 자산을 AI 솔루션에 학습시켜 개발하는 이 솔루션은 광고를 포함한 마케팅 콘텐츠 분야에 특화된 전문지식 역량을 갖추게 된다는 점에서 다른 AI 솔루션과 차별점을 가짐
  - 광고에 특화된 생성형 AI 솔루션은 기업들이 관리해야 할 고객 점점 채널이 증가함에 따라, 콘텐츠의 대량 생산과 고객별 맞춤 콘텐츠를 제공에 어려움을 겪는데, 양질의 콘텐츠 생산성을 향상하는 데 기여할 것으로 전망

### (4) 생성형 AI에 대한 실무자 인식

#### □ 실무자들의 적극적 AI 활용

- 생성형 AI는 잘 이용할 경우 실무자들에게 큰 도움을 줄 수 있는 좋은 도구라는 의견 다수

- 앞서 살펴본 스프라우트 소셜(Sprout Social, 2023)의 마케터 대상 설문결과에 의하면 80% 이상의 마케터들이 시간 절약, 효율성 증대 등 AI가 업무에 미치는 긍정적인 역할을 인지하고 고객 경험의 향상을 위해 AI를 활용하고 있음
  - 데이터 분석과 같이 많은 시간이 필요한 업무를 AI가 대신함으로써 소비자의 요구에 직접 응대하고 더 나은 서비스를 제공할 수 있다는 긍정적인 경험

□ 실무자들의 AI 관련 우려

- 마케터들은 AI와 관련해 일자리 감소를 가장 우려하는 것으로 나타남(Sprout Social, 2023)
  - 3분의 1은 소셜 미디어 상의 상호작용에서 AI와 인간의 접촉 사이의 균형을 유지하는 것을 우려함
- 광고, 마케팅 분야, 특히 디지털과 제작 업무에 대한 전문성이 기술로 인해 침해당하는 느낌
- 기술이 빠른 속도로 발전함에 따라 업무 수행을 위해 지속적인 학습이 필요하고 이를 수행하지 못하면 뒤처지고 설 자리를 잃을 수도 있다는 두려움
- 실제 실무자들이 직접 관리, 운영하던 광고, 마케팅 업무의 상당 부분을 AI를 활용한 디지털 플랫폼에서 대체하거나 보조함으로써 인력 감축 발생

□ 실무자들에게 필요한 자세

- 생성형 AI를 경쟁자로 보는 것이 아닌 업무를 도와줄 도구 혹은 어시스턴트의 포지션으로 분류하여 함께 갈 동반자로 생각해야 할 필요성
- 생성형 AI와 실무자 전문성의 시너지
  - 정교성과 창의성이 여전히 인간보다 부족한 생성형 AI가 실무자의 전문성과 노하우가 접목될 시, 결과물 측면에 있어 좋은 시너지를 낼 수 있을 것으로 예상
- 생성형 AI를 비롯해 AI를 잘 이해하고 활용할 수 있는 AI 리터러시의 중요성 증대
- 생성형 AI 활용에 있어 유의할 점

- 생성형 AI가 가지고 있는 윤리적 문제와 기술적 한계를 인식하고 활용하는데 주의 요망
- 활용 대상에 따른 소비자의 반응을 고려하여 AI 활용이 필요
- 예를 들어, 소비자의 42%는 소셜 미디어 상의 상호작용에 AI를 사용하는 브랜드에 대해 우려를 표함(Sprout Social, 2023). 이는 AI 기반 서비스가 인간적인 상호작용을 원하는 소비자에게는 반발을 가져올 수 있기에 효율성을 위한 지나친 탈인격화(de-humanization) 경계 필요
- 인터넷이 대중에게 보급되며 정보에 대한 특권계층의 독점이 사라진 것과 같이, 기술에 평등하게 접근 가능한 이용자들이 그 도구를 어떻게 사용하느냐에 따라 그 결과물이 긍정적일 수도 있지만, 부정적일 가능성도 배제할 수 없음. 예상치 못한 결과 가능성을 인지하고 결과물과 과정에 대한 모니터링과 개선 필요

## (5) AI가 광고 플랫폼에 미치는 영향

### □ 플랫폼 광고효과와 효율성 증진

- 2023년 AI를 기반으로 광고주와 실무자를 위한 플랫폼의 광고, 마케팅 기능개선과 새로운 서비스 소개가 두드러짐
  - 광고, 마케팅 활동을 원하는 플랫폼 내에서 생성형 AI 활용으로 이미지, 카피, 제품 설명을 다양한 버전으로 빠르고 쉽게 생성할 수 있으며 영상 소재와 스토리에 대한 추천도 받고 영상 편집도 플랫폼의 환경에 맞게 크기, 비례, 길이 등을 자유롭게 변경할 수 있음
  - 카피나 이미지 제작과 편집 관련 전문성이 없는 실무자도 원하는 내용을 대화형으로 입력하면 손쉽게 요청에 맞는 광고 소재를 생성할 수 있기 때문에 업무에 대한 장벽을 낮추고 효율성을 높임
- 데이터와 AI를 활용하는 플랫폼은 두 가지 중요한 장점을 가짐
  - 대규모 청중이 아니라 소비자 개개인의 특성과 취향을 반영해 맞춤형 광고를 제공하기 위해서는 많은 소재를 동시다발적으로 생성해야 하는데 생성형 AI 기반 플랫폼은 대규모의 다양한 버전 광고 제작을 가능케 함

- 광고 매체의 관점에서 타깃 소비자에게 맞춤형 광고를 전달해 원하는 효과를 얻어야 하는데 디지털 환경의 소비자 행동은 복잡하고 수많은 접점으로 인해 어려움이 있음. AI 기반 플랫폼은 실시간으로 소비자 행동 데이터를 분석, 활용해 목표 소비자에게 적합한 광고를 자동으로 정확하게 전달하며 목표를 달성하고 비용 효율성을 높이도록 함
- 종합적으로, 기존의 프로그래매틱 광고 시스템에 생성형 AI가 결합되어 광고 기획, 제작, 전달, 효과분석에 이르기까지 전 과정에 AI가 활용되며 광고효과와 비용 효율성을 극대화할 뿐 아니라 해당 업무의 편이성과 효율성을 높임
- 이미 매체별 광고비 점유율 1위를 차지하고 있는 디지털 광고의 장점을 확대하고 레거시 대비 경쟁력을 공고히 함

□ 플랫폼 경쟁력에 미치는 영향

- 주요 수익원인 광고 경쟁력의 강화는 디지털 플랫폼의 경쟁력 강화로 이어짐
  - 광고 성과의 증대뿐 아니라 광고 업무의 수월성을 높이는 생성형 AI의 결합은 실무자들에게 광고 플랫폼으로서의 매력을 높임
- 생성형 AI 도입 등 광고 솔루션의 개발과 개선은 구글, 메타, 아마존, 마이크로소프트, 네이버 등 주요 글로벌, 로컬 플랫폼들이 주도
  - 특히 구글의 전방위적인 AI 활용과 광고, 마케팅 솔루션 개선이 두드러짐
  - 타 플랫폼들의 AI 기반 광고 솔루션은 부분적으로 출시되거나 소개 예정인 경우가 많음
- AI 활용 광고 솔루션의 품질은 광고 매체로서의 각 플랫폼의 경쟁력 평가에 중요하며, AI 활용 정도와 우수성이 플랫폼의 차별화를 결정하고 플랫폼 간 경쟁을 가속할 가능성 존재
  - AI의 독자적 개발을 위해 대규모의 투자가 필요한 상황에서 거대 플랫폼들이 광고주와 소비자에게 고품질의 광고 서비스를 경쟁 플랫폼들보다 빠르고 안정적으로 제공할 수 있음
- 향후 AI 기반 솔루션과 서비스 관련 경쟁은 심화되고 글로벌 vs. 로컬 플랫폼 간, 거대 vs. 중소 플랫폼 간의 격차가 벌어질 것으로 전망됨

## (6) AI의 한계점 및 문제

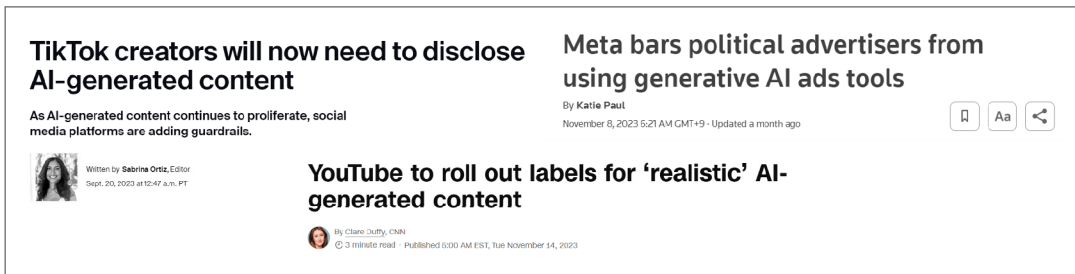
- AI의 긍정적 역할에도 불구하고 극복해야 할 한계점과 문제점은 존재하며 특히 최근 주목받는 생성형 AI에 대한 우려가 높음
  - 콘텐츠의 품질과 편향성
    - AI 생성 텍스트의 부정확성과 오류 가능성 존재
    - 이미지, 텍스트, 음성 생성에서 나타나는 저품질 및 편향된 오류 정보 전달 위험
  - 저작권 및 법적 책임
    - 원저작물 참조 및 변형으로 인한 저작권 이슈  
(저작권 침해 소송 실제 사례의 예시: 게티이미지 vs. Stable Diffusion.)
    - 저작권자 보상 미흡 및 콘텐츠 제작자 손해 발생 문제
  - AI 의사결정의 투명성
    - 미국 스탠퍼드대 기초모델연구센터(CRFM)에 의하면 OpenAI, 구글, 메타 등 빅테크 기업의 대형언어모델의 투명성 지수 낮음
    - 낮은 투명성 지수로 인한 알고리즘 타당성 신뢰 및 결과물 신뢰성 판단의 어려움 존재
  - 데이터 프라이버시
    - 개인 고객의 개인정보 사용 및 침해 위험과 기업 대상 엔터프라이즈 고객의 기밀 유출 및 손실 위험성
  - 소비자 혼동과 낮은 리터러시
    - 정보와 콘텐츠의 생성 혹은 제공 주체를 구분하기 어렵고 AI가 생성한 정보와 콘텐츠로 인해 혼동 유발
    - AI의 급속한 발전으로 인해 AI에 대한 일반 소비자들의 이해도와 리터러시 낮음

## (7) 생성형 AI 활용 광고 규제

### □ 플랫폼의 자율규제

- 플랫폼들은 생성형 AI를 활용하는 광고, 콘텐츠 제작 솔루션을 개발 및 출시했지만 관련 우려를 반영해 가이드라인을 발표함
  - 틱톡은 크리에이터들이 AI를 활용해 생성한 콘텐츠를 밝히도록 규제
  - 메타는 정치, 헬스 등 생성형 AI로 제작된 광고가 잠재적 위험을 가지고 있는 영역의 광고주들은 해당 서비스 이용을 막기로 함
  - 유튜브는 실제와 같이 보이는 AI 활용 콘텐츠는 표기를 의무화 함
  - 이상의 가이드라인은 대부분 예고는 되었지만 실시 전

[그림 II-43] 생성형 AI 활용 광고 가이드라인 예시



자료: ZDNet, CNN. (2023)

- 건강한 디지털 광고 생태계를 위해 자발적으로 준비 중인 플랫폼의 자율규제는 바람직하지만 실효성에 대한 모니터링 필요
- 특히 관련 소비자 경험이 아직 부족한 상황에서 소비자 대상 심도 있고 체계적인 조사 필요

### 1. 저작권 및 라이선스 관련 정책 방향

#### 가. 서론

- 최근 들어 ChatGPT를 비롯한 생성 AI 기술이 급속히 발전하고 법률, 회계, 창작 등을 비롯한 수많은 영역에 채용됨에 따라 생산성 향상과 새로운 서비스의 등장이 이어지고 있음.
- 생성 AI 모델은 대규모의 학습용 데이터셋을 필요로 하고, 그러한 데이터셋에는 개인정보나 저작물과 같이 타인의 권리의 대상이 되는 데이터가 포함되어 있는 경우도 많음. 또한 AI 모델을 이용하여 만들어진 생성물이 데이터셋에 포함된 개인정보나 저작물과 매우 유사한 경우도 등장하고 있음.
- 저작권자 등은 자신의 동의 없이 저작물이 모델의 학습에 이용되고, 나아가 자신의 저작물과 유사한 생성물이 만들어져 유통됨으로써 자신의 정당한 이익이 침해되었다면서 정당한 보상이 이루어져야 한다고 주장하고 있음.
- 이는 생성 AI 생태계가 확장되고 이를 통해 얻을 수 있는 잠재적인 기회와 이익을 저해하고 있음. 따라서 생성 AI의 활용 과정에서 생길 수 있는 데이터와 관련된 법적 쟁점들을 드러내고 해결방안을 모색할 필요가 있음.

#### 나. 현황과 과제

##### (1) 기술과 서비스

###### AI 모델의 유형

- 분류 모델과 생성 모델<sup>11)</sup>

11) 김병필, “인공지능 공정성 심사 기준에 관한 연구”, 서울대학교 박사학위논문 (2023) 참조.

- AI 모델은 주어진 다차원 속성변수  $X$ 를 그와 연관된 종속변수  $Y$ 에 대응시키는 분류모델과 반대로 주어진  $Y$ 에 대응하는  $X$ 를 출력하는 생성모델로 나눌 수 있음.
- 분류모델은 어떠한  $X$ 가 주어진 경우 그에 대응하는 특정  $Y$  유형에 해당할 확률, 즉  $Pr(Y|X)$ 을 산출하는 방식으로 작동함. 이에 비해 생성모델은 주어진 라벨  $Y$ 에 대응하는 특성  $X$ 가 존재할 확률, 즉  $Pr(Y|X)$ 를 산출하는 확률분포 모형을 구축한 다음, 그로부터 확률이 높은  $X^*$ 를 표본추출하는 방식으로 구현됨.

● 분류 모델

- 분류 모델은 인지적 기능을 수행하는 인지 모델과 대상 정보에 기반한 판별적 기능을 수행하는 판별 모델로 나눌 수 있음. 인지 모델의 예로는 이미지 인식, 번역 또는 음성 인식을 위한 모델을 들 수 있고, 판별 모델의 예로는 대출심사나 광고 추천을 위한 모델을 들 수 있음.
- 이들은 정답 라벨의 특성, 과제의 성격, 불확실성의 유형 등에서 차이가 있음. 인지모델의 경우 목표변수의 정답이 명확히 정의될 수 있고 고정된 성격을 지니며, 즉각적인 추단적 판단(heuristics)을 요하는 시스템 1 작업에 적합함. 이에 비해 판별 모델의 경우 목표변수의 정답을 명확히 정의하기 어렵고 확률적인 성격을 지니며, 숙고를 요하는 시스템 2 작업에 적합함. 대체로 인지 모델에 비하여 데이터 불확실성이 대체로 더 높음.

● 생성 모델

- 생성 모델은 정형 데이터 생성 모델과 비정형 데이터 생성 모델로 나눌 수 있음. 전자의 예로는 프라이버시 침해를 막기 위하여 원본 데이터와 동일 또는 유사한 통계적 속성을 갖는 합성데이터(synthetic data)를 생성하는 경우를 들 수 있으며, 후자의 예로는 언어모형, 이미지 생성 모형, 텍스트-이미지 변환 모형 등을 들 수 있음.

□ 생성 모델의 발전

● 언어모형

- 대부분의 언어모형은  $Pr(text)$ 을 계산하기 위하여 조건부 확률의 연쇄 법칙(chain rule)을 활용함. 앞선 단어 시퀀스 다음에 어떠한 단어에 등장할

것인지 예측하는 통계모델을 구축하고 그 값을 모두 곱함으로써 해당 전체 시퀀스의 확률을 산출하는 모델을 구성할 수 있음.

- 구글이 2017년 트랜스포머(Transformer) 아키텍처를 발표한 이후<sup>12)</sup> 대부분의 언어모델은 이를 활용하고 있음. ‘주어진 질문에 질문과 답이 있다’는 아이디어에 기초한 트랜스포머 아키텍처는 질문(query)을 토대로 지문(key)에서 답(value)을 찾기 위하여 ‘self-attention’을 활용하였음. 이러한 구조를 통해 모델이 입력데이터 내의 문자들 사이의 관계를 이해할 수 있게 되었을 뿐만 아니라, 종전의 RNN 모델과 달리 대규모 병렬처리가 가능하게 되어 언어모델의 규모가 커질 수 있었음.
- BERT나 GPT-3/4와 같은 대규모 언어모델이 등장하고 API로 제공되어 다양한 분야에서 사용됨에 따라 이른바 기반 모델(foundation model)로 기능하게 되었음. 기반 모델의 규모가 증가하고, 더 많은 데이터와 더 많은 연산 자원을 활용하여 이를 학습시키면서 기반모델이 수행할 수 있는 능력도 점차 확장되어 가고 있는데, 이처럼 기반모델이 새로운 능력을 갖추게 되는 현상을 ‘창발성(emergence)’이라고도 함.

#### ● 이미지 생성 모델

- 이미지 생성 모델은 어떤 주어진 이미지가 존재할 확률분포 모형, 즉  $Pr(\text{image})$ 를 활용해 새로운 이미지를 생성함. 즉, 어떤 이미지가 얼마나 그럴 듯한 것인지를 통계적으로 계산하는 모형을 만들고, 그로부터 적절한 표본을 추출해 내는 것이라고 할 수 있음.
- 이미지 생성 모델을 구현하기 위해 오토인코더(Variational Autoencoder, VAE), 적대적 생성신경망(Generative Adversarial Network, GAN), 확산(Diffusion) 모델 등이 개발되어 왔는데, 최근에는 적대적 생성 신경망 모델과 확산 모델이 널리 활용되고 있음.
- 적대적 생성신경망 구조에서는 생성 신경망과 판별 신경망을 별도로 구축하여, 생성신경망이 판별신경망을 더 잘 속이는 능력을 학습함으로써 최종적으로 원본 학습 데이터와 구별하기 힘든 수준의 유사한 데이터를 만들어 내도록 함.

12) Vaswani, Ashish, et al. "Attention is all you need." Advances in neural information processing systems 30 (2017).

- 확산모델은 학습 데이터에 포함된 이미지에 임의의 잡음을 추가한 다음, 인공신경망으로 하여금 추가된 잡음과 원본 이미지를 구별하여 잡음을 제거하는 방법을 학습하도록 함. 확산모델은 DALL-E 2나 MidJourney와 같은 대표적인 이미지 생성 서비스에서 사용됨.

● 생성모델과 암기의 가능성

- 인공지능의 내부 파라미터는 인공지능 학습 과정 중 목적 함수에 대해 최적화되는 수학적 과정을 통해 결정된 일련의 내부적 수치일 뿐이고, 학습 데이터 자체가 직접적으로 내부 파라미터에 복제되지는 않음. 따라서 인공지능이 내부 파라미터 업데이트 과정을 거쳐 학습 데이터 자체를 그대로 암기할 것이라는 우려는 크지 않았음. 오히려 생성모델을 통해 새롭게 재현된 데이터를 통해 개인정보 보호 수준을 높일 수 있을 것으로 기대되어 왔음.
- 그러나 최근 실증연구를 통해 대규모 생성모델이 상당한 학습 데이터 암기 능력을 보유하고 있다는 점이 확인됨. 대표적으로 Carlini et al. (2021)은 GPT-2의 학습 데이터 암기 사실을 확인하였음.<sup>13)</sup> 즉, GPT-2 언어모델을 활용해 생성한 문장의 최소 0.1%에서 학습 데이터가 ‘그대로(verbatim)’ 재생성되었음. 나아가 이미지 생성모델에서도 학습 데이터 암기 능력이 확인됨. Carlini et al. (2023)에 따르면 이미지 생성모델에 어떤 사람의 성명을 입력하면 위키피디아에 저장된 그 인물의 초상 사진이 “거의 동일하도록(nearly identical)” 생성되었음.<sup>14)</sup>
- 위와 같은 연구 결과를 그대로 받아들여 생성 모델이 학습 데이터를 ‘암기’한 것으로 평가할 것인지에 관하여는 견해가 대립될 것으로 보임. 그러나 암기 가능성을 인정하는 입장에서는 이미지 생성모델에 의한 저작권 침해 위험성 및 개인정보 자기결정권 침해 위험성에 대하여 우려하고 있는 상황임.

13) Carlini, Nicholas, et al. “Extracting training data from large language models.” 30th USENIX Security Symposium (USENIX Security 21). 2021.

14) Carlini, Nicolas, et al. “Extracting training data from diffusion models.” 32nd USENIX Security Symposium (USENIX Security 23). 2023.

## (2) 편익과 위험

### □ 편익

- 기존의 분류 모델은 이미지 인식, 번역, 음성 인식 또는 대출심사, 광고 추천과 같은 인지 기능 및 판별 기능을 수행함으로써 생산성을 향상시켰음.
- 생성형 AI 모델은 이에 더하여 다채롭고 풍부한 생성물을 제공해 줌으로써, 문제를 더 잘 이해하고 해결할 수 있도록 해 주고, 창작 활동을 보조하며, 사회적 관계를 개선하고 확장할 수 있도록 해 줌. 이는 궁극적으로 생산성의 향상과 새로운 서비스의 등장으로 이어져 궁극적으로 인류의 삶의 질을 향상시킬 수 있는 계기가 될 것임.
- 특히 창작을 위하여 생성 모델을 활용할 경우 노력과 비용을 대폭 감소시켜 소비자가 보다 저렴하게 다양한 창작물을 향유할 수 있음. 나아가 창작에 필요한 기술적 장벽의 감소로 창작자의 저변이 확대되고 AI 기술을 활용한 새로운 창작 영역이 생겨남에 따라 개인이 자신의 개성을 표현하고 자아를 실현할 기회가 증진되고 문화 발전이 촉진될 수 있음.

### □ 위험

- 타인의 권리 침해
  - AI 모델의 학습용 데이터세트에 타인의 저작물이나 개인정보 또는 초상과 같이 배타적 권리의 목적이 되는 데이터가 포함되어 있는 경우 학습 과정이나 생성 과정에서 제3자의 권리가 침해될 우려가 있음.
- 경쟁 저해와 소비자 권익 침해
  - AI 생성물의 판매를 통해 노력 내지 비용에 비해 과다한 수익을 얻는다고 인식될 경우 소비자 권익이 침해되고, AI 생성물이 인간 창작물로 오인될 경우 소비자의 신뢰가 훼손될 우려가 있음.
  - 대규모 언어모델의 발전으로 기반 모델이 등장하고 범용성을 지닌 일반인공지능의 개발을 위한 경쟁이 벌어지고 상황에서 그러한 모델이 경쟁 저해를 초래할 가능성에 대한 우려가 있음.

● 기존의 생태계에 대한 위협

- 값싼 AI 생성물이 시장을 잠식한다고 인식될 경우 공정한 경쟁이 훼손되고 창작자의 권익이 침해된다는 비판이 제기될 우려가 있음. 예를 들어 미국에서는 2023년 5월 미국작가조합(WGA)이 대본 작업에 AI를 활용하는 것에 반대하면서 파업을 하였고 이후 미국배우방송인노동조합(SAG-AFTRA)이 동참한 바 있고,<sup>15)</sup> 우리나라에서도 유사한 논쟁이 벌어지고 있음.<sup>16)</sup>
- 또한 창작 과정에서 AI 기술 활용이 증가할 경우 가격 경쟁에 내몰린 창작자들이 비용 절감을 위해 소수의 AI 모델을 활용하게 되는 결과 창작물의 개성 내지 다양성이 감소하고 나아가 문화적 다양성이 감소할 우려가 제기될 수 있음.

● 인간성에 대한 위협

- AI 생성물이 인간 창작물을 대체하는 것이 인간 고유의 영역을 침범하는 것으로 인식될 경우 인간성에 대한 침해라는 비판이 제기될 우려가 있음.
- 일반인공지능이 현실화될 경우 어떠한 위해를 초래할 수 있는지는 불확실하며, 장차 이에 관한 정보가 축적되면 활용 영역과 보호 법익에 따라 차별화된 규율이 이루어질 것임. 다만 일부 학자들은 인공지능이 자기개선 능력을 갖게 되면 폭발적인 성능 향상을 통해 인간의 능력을 뛰어넘고<sup>17)</sup> 인간의 통제를 벗어난 초지능(superintelligence)이 도래할 가능성이 있다면서 통제 방안을 마련할 것을 촉구하고 있는바,<sup>18)</sup> 주의 깊게 상황을 살펴볼 필요가 있음.

15) 조선일보, “파업 중인 할리우드 작가들 “AI 대본 작성 제한해야”(2023. 5. 10.자). [https://www.chosun.com/international/international\\_general/2023/05/10/CIE5ISMO6ZGKVMABKNSJJ6XR2E/?utm\\_source=naver&utm\\_medium=referral&utm\\_campaign=naver-news](https://www.chosun.com/international/international_general/2023/05/10/CIE5ISMO6ZGKVMABKNSJJ6XR2E/?utm_source=naver&utm_medium=referral&utm_campaign=naver-news)

16) 중앙일보, “작가 200명 “도둑” 이현세는 “도우미”… AI에 뒤송송한 웹툰업계”(2023. 6. 8.자). <https://www.joongang.co.kr/article/25168341>

17) GPT-4의 미공개 초기 버전을 평가한 연구자들은 이 모델이 단지 언어 뿐 아니라 수학을 비롯한 다른 영역에서도 어려운 과제들을 해결하는 창발적 능력을 보인 점에 주목하여, 일반인공지능(Artificial General Intelligence: AGI) 시스템의 불완전한 초기 버전으로 간주될 수 있다고 주장하였음. Sébastien Bubeck, et al., “Sparks of Artificial General Intelligence: Early Experiments with GPT-4” (2023).

18) Nick Bostrom, Superintelligence: Paths, Dangers, Strategies, Oxford University Press (2014).

## 다. AI 모델 학습 과정에서의 문제

### (1) 문제점

- AI 모델의 학습 과정에서는 필연적으로 학습 데이터의 복제와 전송이 이루어지므로 만약 학습 데이터에 타인의 권리의 목적이 되는 저작물이나 개인정보 등이 포함되어 있다면 저작권이나 데이터베이스권 또는 개인정보의 침해 여부가 문제됨.
- 개인정보 침해 문제는 AI 모델 활용 과정에서 주로 다루기로 하고, 여기서는 저작권이나 데이터베이스권의 침해 문제를 서술함.

### (2) 학습용 데이터세트에 타인의 저작물이 포함된 경우

#### □ 저작재산권의 침해

- AI 모델이 데이터를 학습하는 과정에서 데이터의 복제 및 전송이 이루어지므로 학습용 데이터세트에 저작권자의 동의를 얻지 않은 저작물이 포함된 경우 기술적으로는 일단 저작권의 침해가 성립됨(저작권법 제16조).
- 저작권자로부터 라이선스를 받은 경우에도 라이선스 조건에 위반하여 데이터 학습이 이루어진 경우라면 일응 저작권의 침해가 성립될 수 있음.

#### □ 저작재산권의 제한과 공정이용

##### ① 저작재산권의 제한

- 다만 저작권법은 정보에 대한 독점과 공유의 조화라는 관점에서 저작재산권 제한 제도를 두고 있는바, 위와 같은 학습 과정에서의 복제가 저작권법상 공정이용(제35조의5) 또는 공정한 인용(제28조) 등 저작재산권의 제한 사유에 해당할 경우 적법성이 인정될 수 있음.
- 이는 해당 AI 모델이 어떠한 유형의 것인가와 관련되어 있는데, 데이터 학습 결과를 바탕으로 생성물을 산출하는 생성 AI 모델의 경우 공정이용 등의 요건을 갖출 수 있는지 검토할 필요가 있음.

## ② 공정이용

### ● 공정이용 제도

- 저작권법상 저작물의 통상적인 이용 방법과 충돌하지 아니하고 저작자의 정당한 이익을 부당하게 해치지 아니하는 경우에는 공정이용에 해당하여 저작자의 허락을 받지 않고 해당 저작물을 이용하는 것이 가능함(제35조의5).
- 저작물의 공정한 이용에 해당하는지의 여부를 판단할 때는 ① 이용의 목적 및 성격, ② 저작물의 종류 및 용도, ③ 이용된 부분이 저작물 전체에서 차지하는 비중과 그 중요성, ④ 저작물의 이용이 그 저작물의 현재 시장 또는 가치나 잠재적인 시장 또는 가치에 미치는 영향을 고려해야 함.

### ● 생성 모델의 경우

- 분류 모델의 경우 검색 서비스와 유사하게 공정이용의 범위에 포함될 가능성이 높아 보이지만, 생성 AI 서비스를 위한 데이터 학습에 공정이용 원칙이 적용될 수 있는지는 예측하기 어려움.
- 이용의 목적 및 성격
  - 이용의 목적이 영리적인 경우에는 비영리적인 경우보다 공정이용에 해당하기 어려움. 다만 저작물의 이용이 변형적 이용에 해당할 경우에는 영리성 여부는 부수적인 판단요소로 취급될 수 있음<sup>19)</sup>
  - 이미지 관련 AI 모델의 경우 이미지 분류, 객체 탐지, 감정 분석, 이미지 추천 등을 위해 이용되는 분류 모델의 경우에는 변형적 이용으로 인정될 가능성이 큰 반면, 이미지 생성이나 합성 등과 같은 생성 모델의 경우에는 변형적 이용으로 볼 수 있을지는 원저작물의 목적 또는 기능과 유사한 서비스를 제공할 가능성이 높기 때문에 다툼의 여지가 많음.
- 저작물의 종류 및 용도
  - 원 저작물이 사실적 성격을 지닌 경우에는 픽션처럼 허구적인 성격을 지니는 경우에 비해 공정이용 해당 가능성이 큼.

19) 변형적 이용이란 '단지 원작의 목적을 대체하는 것이 아니라 추가적인 목적이나 다른 성격 등 무언가 새로운 것을 부가해 원작의 표현, 의미 또는 메시지를 변경하는 것'을 의미함. 예를 들면 검색을 위한 인덱스를 작성하는 과정에서 기존의 저작물을 복제하는 경우 변형적 이용에 해당하여 공정이용으로 판단될 가능성이 높음. 영리적인 목적으로 이용했음에도 불구하고 변형적 이용에 해당함을 이유로 공정이용을 인정한 사례로는 Authors Guild v. Google, Inc., 721 F.3d 132 (2d Cir. 2015).

- 생성 모델의 학습데이터의 종류와 성격은 매우 다양함. 예컨대 위키피디아 문헌이나 Common Crawl이 수집한 웹페이지 자료들 중에서 특히 명세서 관련 자료, 언론 보도 자료 등은 사실적인 요소가 강함. 컴퓨터프로그램 등 기능적 저작물은 픽션보다는 사실적 저작물에 가까움.<sup>20)</sup> 반면에 bookcorpus 등에 포함된 소설 등은 픽션이나 판타지 등의 성격을 지닌 것이 많음.
- 이용된 부분이 저작물 전체에서 차지하는 비중과 그 중요성
  - 이용된 부분의 양이 많거나 중요한 부분에 해당한다면 공정이용에 해당하지 않을 가능성이 높아짐.
  - 이용된 부분의 비중은 원칙적으로 이용된 저작물 전체를 기준으로 판단하지만 구체적인 사례에서는 만들어진 저작물을 기준으로 판단하는 것은 경우도 있음. 이용된 저작물을 기준으로 할 때 AI 학습에 이용되는 저작물들은 대부분 원저작물 전체를 복제하여 이용하는 경우가 대부분이지만, 학습용 데이터 세트 전체를 기준으로 보면 개별적인 저작물의 비중은 매우 낮음.<sup>21)</sup>
- 저작물의 이용이 그 저작물의 현재 시장 또는 가치나 잠재적인 시장 또는 가치에 미치는 영향
  - 분류 모델의 경우 변형적 이용의 형태로 활용되어 원저작물에 대한 수요를 대체하지 않기 때문에 원저작물 시장이나 가치에 미치는 영향이 적은 반면, 생성 모델의 경우 원저작물과 유사하고 대체가능한 생성물을 만들어낼 가능성이 있기 때문에 원저작물 시장이나 가치에 미치는 영향력이 상대적으로 큼.
- 라이선스와의 관계
  - 공표되지 않은 저작물도 공정이용의 대상이 될 수는 있지만, 미공표의 사실은 공정이용으로 인정하는데 불리하게 작용함.<sup>22)</sup>
  - CCL 등의 라이선스에 의하여 저작자에 의해 자유로운 이용이 허락된 저작물들을 학습데이터로 활용하는 것은 그렇지 않은 저작물에 비해 공정이용에 해당할 가능성이 높음. 그러나 자유로운 이용이 허용되더라도 비영리 조건이 부가된 경우에는 공정이용에 해당하지 않을 가능성이 높을 것으로 보임.<sup>23)</sup>

20) 다만 컴퓨터프로그램이라고 할지라도 창작성의 정도 등을 개별적으로 고려할 필요가 있음. Google LLC v. Oracle Am., Inc. 141 S. Ct. 1163 (2021).

21) Authors Guild v. Google, Inc. 사건에서는 원저작물 전체를 복제한 경우이더라도 그와 같은 복제가 변형적 이용에 불가피하고 원저작물에 대한 대체적인 효과를 가져오지 않는다면 공정이용에 해당할 수 있다고 보았다.

22) Harper & Row Publishers, Inc. v. Nation Enterprises, 471 U.S. 539 (1985)

23) 예컨대 위키피디아와 달리 나무위키는 비영리조건을 부과하고 있음.

③ 공정한 인용

- 공표된 저작물이 보도·비평·교육·연구 등을 위해 정당한 범위 안에서 공정한 관행에 합치되게 인용되는 경우 공정한 인용에 해당되어 면책될 수 있음(저작권법 제28조).
- 법원은 정당한 범위 안에서 공정한 관행에 합치되게 인용한 것인가의 여부는 인용의 목적, 저작물의 성질, 인용된 내용과 분량, 피인용저작물을 수록한 방법과 형태, 독자의 일반적 관념, 원저작물에 대한 수요를 대체하는지 여부 등을 종합적으로 고려하여 판단함.<sup>24)</sup>
- 다만 공표된 저작물이 연구개발을 위한 AI 학습을 위하여 이용되는 경우 공정한 인용에 해당할 가능성은 크지 않다고 생각됨.

④ 데이터 마이닝 면책

● 저작권법 개정안

- 데이터 경제의 발전에 따라 대규모 데이터셋을 빅데이터 분석을 위해 활용하거나 인공지능 학습을 위해 활용할 경우 데이터셋을 이루는 개별데이터 이용의 적법성을 확보하기 위한 방안이 요구됨. 이에 따라 세계 각국은 텍스트 및 데이터 마이닝(Text Data Mining; TDM, 이하 ‘데이터 마이닝’이라고만 함)을 위한 면책 규정을 저작권 법제에 도입하였으며, 우리나라에서도 데이터 마이닝 면책 규정을 담은 저작권법 개정안<sup>25)</sup>이 발의되었음. 또한, 정부(문화체육관광부)는 저작권법 개정을 통해 AI 학습을 위해 데이터를 활용하는 경우의 저작권 침해 면책 요건 및 근거를 마련하겠다고 발표하면서 AI 학습을 위한 ‘크롤링(crawling) 행위’는 저작물에 포함된 사상·감정을 향유하지 않고, 적법한 저작물 접근에 해당함을 명시하겠다고 한 바 있음.<sup>26)</sup>

24) 법원은 썸네일 이미지 사례에서 사진을 이미지 검색에 제공하기 위하여 압축된 크기의 썸네일 이미지로 게시한 것이 작품사진에 대한 수요를 대체한다고 보기 어렵다고 보아 공정한 인용으로 판단함(대법원 2006. 2. 9 선고 2005도7793 판결).

25) 2021. 1. 15. 도중환 의원 대표발의, 의안번호 제2107440호

제43조 정보분석을 위한 복제·전송

① 컴퓨터를 이용한 자동화 분석기술을 통해 다수의 저작물을 포함한 대량의 정보를 분석(규칙, 구조, 경향, 상관관계 등의 정보를 추출하는 것)하여 추가적인 정보 또는 가치를 생성하기 위한 것으로 저작물에 표현된 사상이나 감정을 향유하지 아니하는 경우에는 필요한 한도 안에서 저작물을 복제·전송할 수 있다. 다만, 해당 저작물에 적법하게 접근할 수 있는 경우에 한정한다.

② 제1항에 따라 만들어진 복제물은 정보분석을 위하여 필요한 한도에서 보관할 수 있다.

26) 2023. 7. 21. 서비스 산업의 디지털화 전략(관계부처 합동) 참조

- 다만 위 법안에 대하여는 요건으로 되어 있는 ‘적법한 접근’의 개념이 불분명하다고, 공정이용의 법리에 의한 규율로 충분하다는 등의 반론도 있음.<sup>27)</sup>

#### ● 생성 모델의 경우

- AI 모델 일반에 관한 논의 : 저작물 본래의 목적 외 다른 목적, 예컨대 검색지원이나 통계분석 등을 목적으로 하는 경우 이른바 변형적 이용(transformative use)으로서 공정이용에 해당된다고 보는 견해<sup>28)</sup>가 있음
- 생성 모델에 관한 논의 : TDM 면책 법리의 근거 중 하나는 저작물의 비표현적 이용에 있으나, 일부 생성 모델의 경우 학습데이터에 포함된 저작물의 비표현적 부분만을 사용한다고 단정할 수 없다는 의견이 있고,<sup>29)</sup> 저작권법 개정안에 의하더라도 “저작물에 표현된 사상이나 감정을 향유하지 아니하는 경우”에만 면책되도록 되어 있음.<sup>30)</sup>

### (3) 학습용 데이터세트에 타인의 데이터베이스가 포함된 경우

#### ● 데이터베이스권

- 우리 저작권법은 EU의 입법례를 따라 데이터베이스 제작자의 권리(이하 ‘데이터베이스권’이라고 함)을 인정하고 있는바, 학습용 데이터세트에 데이터베이스권의 목적이 되는 데이터베이스가 포함된 경우 데이터베이스권의 침해가 될 가능성이 있음.
- 다만, 데이터베이스권은 데이터베이스 자체를 지배하는 권리라기보다는 데이터베이스에 대한 투자를 보호하기 위한 권리의 성격이 강함. 저작권법은 ‘데이터베이스의 제작 또는 그 소재의 갱신·검증 또는 보충에 인적 또는 물적으로 상당한 투자를 한 자’만을 데이터베이스 제작자로서 보호하고 있음(제2조 제19조, 20호). 또한 데이터베이스를 이루는 개별 소재를 복제한 것만으로는 데이터베이스권의 침해가 되지 않으며 데이터베이스의 ‘전부 또는 상당한 부분’을 복제해야 침해가 됨(제93조).

27) 각국의 입법례에 관한 상세한 비교 및 우리 저작권법 개정안에 대한 평가로는 이상용, “데이터의 비계약적 이용 - 데이터 마이닝을 위한 저작권 제한을 중심으로 -”, 강원법학 제65호, 2021 참조.

28) 박성호, “텍스트 및 데이터 마이닝을 목적으로 하는 타인의 저작물의 수집·이용과 저작재산권의 제한”, 인권과 정의 Vol. 494, 대한변호사협회, 2020; 정충원, “인공지능(Artificial Intelligence : AI)시스템에 대한 저작권보호에 관한 연구”, 법학연구 제31권 제2호, 경상국립대학교 법학연구소, 2023.

29) 전용준, “인공지능 관련 저작권 침해에 관한 시론”, 경영법률 제31권 제4호, 한국경영법률학회, 2021 참조.

30) 학습데이터에 포함된 저작물의 비표현적 부분만을 이용하는 생성 모델도 있을 수 있다. 그에 해당하는지 여부는 학습단계를 놓고 판단할 수는 없고, AI 모델의 목적이나 생성물의 종류 및 용도 등을 고려하여 판단해야 할 것이다.

- 데이터베이스권의 효력은 데이터베이스를 이루는 개별 소재에 미치지 않음. 따라서 데이터베이스의 사용에 관하여 데이터베이스권자의 동의를 얻었다고 하더라도 그 소재가 되는 개별데이터의 사용 권한을 얻게 되는 것은 아님. 예를 들어 개별데이터가 개인정보나 저작물과 같은 배타적 권리의 대상이 되는 경우에는 정보주체나 저작권자의 동의를 얻거나 그 밖의 방법으로 이용 권한을 따로 확보해야 함.<sup>31)</sup>
- 참고로, 데이터베이스권 침해를 주장하는 많은 사건에서 부정경쟁방지법상 성과물 무단이용 및 데이터 부정사용(제2조 제1호 카목, 파목), 정보통신망법상 정보통신망 침입(제48조),<sup>32)</sup> 또는 형법상 업무방해 등이 함께 주장되고 있음.<sup>33)</sup>

● 데이터베이스권의 제한

- 데이터베이스권에 관하여는 공정이용을 포함하여 저작재산권의 제한에 관한 규정을 일부 준용하고 있으므로, 데이터베이스제작자의 권리 침해 주장에도 공정이용에 관한 항변이 적용 가능함.
- 데이터베이스제작자의 권리 침해 주장에 대한 항변으로 공정이용 여부를 판단할 경우, 용도 및 관련 시장이나 이용된 부분이 차지하는 비중과 그 중요성의 판단에 있어서 개별 소재가 아닌 데이터베이스 전체의 관점에서 평가해야 함. 대체로 대량의 학습용 데이터셋을 보유한 플랫폼이 원고로서 권리침해를 주장할 경우에는 개별 저작자들의 경우보다 공정이용에 해당하지 않을 가능성이 높아진다고 볼 수 있음.

(4) 라이선스 위반의 문제

□ 라이선스 위반의 효과

- 학습용 데이터셋에 저작물이나 데이터베이스가 포함된 경우 공정이용 등에 의하여 저작재산권이 제한되는 경우가 아닌 한 권리자로부터 이용허락을 받아야 함. 라이선스 조건을 위반한 경우 저작권 등의 침해가 될 수 있음.

31) 이상용, “데이터 경제 발전을 위한 데이터베이스권의 보완 및 개선”, 계간 저작권 제36권 제1호, 한국저작권위원회, 2023, 116-118면 참조

32) 누구든지 정당한 접근권한 없이 또는 허용된 접근권한을 넘어 정보통신망에 침입해서는 안되며, 이 때 접근권한을 부여하거나 허용하는 범위를 설정하는 주체는 서비스제공자이므로, 접근권한이 있는지 여부는 서비스제공자가 부여한 접근권한을 기준으로 판단해야 함(대법원 2005. 11. 25. 선고 2005도870 판결).

33) 그 밖에 콘텐츠산업법 제37조의 콘텐츠 침해행위로 주장될 수도 있을 것임.

- 데이터베이스에 관한 권리자로부터 이용허락을 얻었다 하더라도 이에 포함된 개별 소재인 저작물에 관하여 적법한 이용 권한을 확보해야 함.

#### □ 저작물에 관한 라이선스

- CCL이나 오픈소스 라이선스 등의 조건으로 이용이 허락된 경우 이용 조건을 준수해야 함.<sup>34)</sup>
- 학습용 데이터세트에 포함된 저작물에 관하여 라이선스 조건을 위반한 경우 저작권 침해가 될 수 있음.
  - 구글의 BERT 등의 학습에 사용되었던 BooksCorpus 데이터셋은 1만권 이상의 전자책을 포함하고 있었는데, 저작권 또는 라이선스 위반 문제로 더이상 활용되지 않고 있음.<sup>35)</sup>
  - GitHub는 이용약관에 AI 학습데이터로 활용하는 것에 관한 명시적인 규정을 두고 있지 않았는데, 관련 소송에서 원고는 피고들이 GitHub의 이용약관을 위반했다고 주장함.<sup>36)</sup>

#### □ 데이터베이스에 관한 라이선스

- 데이터베이스의 전부 또는 상당한 부분을 이용하는 경우 데이터베이스권자의 이용허락을 얻어야 하며 라이선스 조건을 준수해야 함.
  - 게티 이미지의 경우 이미지 자료와 관련 메타데이터를 인공지능과 머신러닝 툴의 개발 관련 분야에 사용할 수 있도록 라이선스를 제공하고 있음.
  - 대부분의 온라인 플랫폼들은 이용약관을 통해 플랫폼상의 데이터에 대한 이용 조건을 정하고 있음.
- 데이터베이스에 대한 권리자가 크롤링을 활용한 데이터 마이닝을 금지한 경우 이를 위반하여 AI 학습에 사용한 경우 이용약관을 위반한 것인지, 나아가 데이터베이스권을 침해한 것인지 여부가 문제됨.

34) 위키피디아는 CC BY-SA 조건으로 이용가능하므로, 이를 이용할 경우에는 저작자를 표시하고, 동일한 조건(CC BY-SA)으로 배포해야 함. 나무위키는 CC BY-NC-SA 조건으로 이용가능하므로, 상업적으로 활용할 수 없음. 오픈소스 라이선스로 배포되고 있는 것으로는 GitHub에 공개된 오픈소스 프로젝트들이 대표적임. 내용은 라이선스마다 다르나 GPL 등 Copyleft 라이선스들은 대부분 소스코드의 제공을 요구하는 반면, MIT 등의 Permissive 라이선스들은 소스코드를 제공할 필요가 없음.

35) 국내에서도 국립국어원의 '문어 말뭉치'에 포함된 책 콘텐츠 약 1만6000권에 대해 도서 유통사가 개별 저자나 출판사의 허락 없이 말뭉치 사업에 참여하면서 문제가 되었던 사례가 있음. 한편, 네이버는 플랫폼 이용약관에 AI 학습에 관한 명시적인 규정을 두는 형태로 블로그 이용자 등으로부터 이용허락을 받고 있음.

36) DOE 1 v. GitHub, Inc., 4:22-cv-06823, (N.D. Cal.).

- 일부 플랫폼은 크롤링 등 자동화된 수단을 통한 게시물 수집을 금지하고 있음. 예컨대 게티 이미지는 사이트 이용조건으로 데이터 마이닝, 로봇 또는 이와 비슷한 데이터 수집 또는 추출 방법을 금지하고 있고, 네이버 이용약관은 자동화된 수단을 이용하여 네이버 서비스에 게재된 게시물 등을 수집하는 등의 행위를 금지함.
- 로봇배제표준(robots.txt)<sup>37)</sup>에 위반하여 크롤링하는 것이 위법한지도 문제됨. 로봇배제표준은 자발적인 준수에 의존하고 있는데, 특정한 웹사이트에서 로봇배제표준을 이용해 크롤링을 금지하고 있다는 표시를 하고 있음에도 불구하고 크롤링을 통해 데이터를 수집하는 것이 위법한지 여부는 명확하지 않음. 검색을 위한 크롤링은 허용하지만, AI 학습데이터로 사용하는 것은 금지한다는 표시가 가능한 것인지도 문제됨.

## 라. AI 모델 활용 과정에서의 문제

### (1) 문제점

- 학습된 AI 모델을 이용하여 생성물을 만들어 배포하는 경우에 그 생성물이 학습용 데이터셋에 포함된 저작물이나 개인정보 또는 인격적 표시와 매우 유사한 경우, AI 모델 이용자에 의한 타인의 저작권 침해나 개인정보 침해 또는 인격적 법익의 침해가 발생할 가능성이 있음.

### (2) 저작권 침해의 위험성

#### □ 저작권 침해의 요건

- 저작권 침해가 인정되기 위해서는 i) 피고의 저작물이 원고의 저작물을 이용한 것일 것, 즉 '의거(依據)관계가 존재한다는 점, ii) 피고의 저작물이 원고의 저작물과 동일성이 인정되거나 종속적 관계에 있을 것, 즉 '실질적 유사성'이 있어야 함.
- 의거관계는 기존의 저작물에 대한 접근가능성 및 대상 저작물과 기존의 저작물 사이의 유사성이 인정되면 추정할 수 있음. 특히 대상 저작물과 기존의 저작물이 독립적으로 작성되어 같은 결과에 이르렀을 가능성을 배제할 수 있을 정도의

37) 웹 크롤러 및 기타 웹 로봇이 방문할 수 있는 웹 사이트 부분을 표시하기 위해 사용하는 표준임.

현저한 유사성(striking resemblance)이 인정되는 경우에는 그러한 사정만으로도 의거관계를 추정할 수 있음.

- 저작권의 보호 대상은 사람의 정신적 노력에 의하여 얻어진 사상 또는 감정을 외부에 표현한 창작적인 표현형식이고, 표현되어 있는 내용 즉 아이디어나 이론 등의 사상 및 감정 그 자체는 보호 대상이 되지 않음. 따라서 저작권의 침해 여부를 가리기 위하여 두 저작물 사이의 실질적 유사성 여부를 판단에 있어서도 창작적인 표현형식에 해당하는 것만을 가지고 대비해야 함.

#### □ AI 생성물의 저작권 침해 여부

- 학습 과정에서의 저작권 침해는 주로 복제권이나 전송권을 중심으로 문제되는 반면 생성 과정에서의 저작권 침해는 복제권 또는 2차적 저작물작성권을 중심으로 문제됨.
- AI 생성물의 경우 주로 의거관계와 관련하여 논의가 이루어짐.
  - 의거성의 인정이 어려울 것이라고 보는 견해 : “인공지능 내부의 정보처리나 창작과정을 알 수 없고 어떠한 데이터를 기초로 작품을 만들었는지 알아내기 어려우므로, 저작권 침해 주장자인 인간 창작자가 자신의 저작물에 대한 의거성을 입증하는게 거의 불가능하다”<sup>38)</sup> / “실무적으로는 의거성 입증에서 상당한 곤란이 있을 것으로 예상”<sup>39)</sup>
  - 구체적인 사안에서 의거관계의 증명이 가능할 것으로 보는 견해 : 창작 도구로 사용된 AI의 학습 데이터세트가 공개되어 있는 경우 해당 데이터세트에 원고의 저작물이 포함되어 있음을 증명할 수 있음.<sup>40)</sup> Getty Images v. Stability AI 소송에서 원고는 생성 AI가 만들어내는 이미지에 원고의 워터마크가 포함되어 있다는 점을 제시하고 있음
- 한편, 생성 AI 모델 이용자의 생성물이 저작권을 침해하는 것으로 인정될 경우 이론적으로는 AI 모델을 제공한 자 역시 방조에 의한 공동불법행위(간접침해) 책임을 질 여지가 있으나, 고의나 과실 등 주관적 요건이 인정되기는 어려울 것임.

38) 조연하, “인공지능의 콘텐츠 창작에서 저작물 이용의 법적 쟁점에 관한 고찰”, 사이버 커뮤니케이션 학보 제39권 제2호, 사이버커뮤니케이션학회, 2022.

39) 전응준, “인공지능 관련 저작권 침해에 관한 시론”, 경영법률 제31권 제4호, 한국경영법률학회, 2021.

40) LAION-Aesthetics dataset에는 1천 2백만장의 이미지-텍스트 쌍이 포함되어 있는데, 각 이미지의 작가를 분석하고 정리한 자료가 공개되어 있음.

□ 선의의 저작권 침해 문제

- 현행 저작권법상 저작권 침해로 인한 손해배상의 경우 다른 지적재산권의 경우와 달리 경과실 감액 규정이 없고, 금지청구의 경우 고의 과실은 요건으로 되어 있지 않음.
- 그러나 등록을 요건으로 하지 않아 선의의 침해가 발생하기 쉬운 저작권의 경우 다른 나라의 입법례를 참고하여(예컨대 미국은 침해소송을 위해 저작권 등록을 요건으로 하고, 독일은 금지청구에 침해자의 고의 과실을 요구하고 있음) 권리남용 등의 법리에 의하여 선의 침해자의 이익을 배려하자는 주장이 있음.<sup>41)</sup>

(3) 개인정보 침해의 위험성

□ 생성 모델의 개인정보 재현 가능성

- 기존의 이해는 AI 학습 목적의 개인정보 수집이나 AI 모델의 배포·이용은 개인정보보호에 관한 권리 침해 위험성이 없거나 낮다는 것이었음.
  - AI 학습은 인공신경망 내부 파라미터(인공 뉴런 간의 연결 강도 등)의 갱신을 통해 이루어지는데, 이는 인간이 이해할 수 없는 일련의 수치에 불과하므로 AI 모델은 개인정보를 포함하지 않음. 따라서 개인정보를 AI 학습에 사용하거나, 학습된 AI를 배포·사용하더라도 개인정보보호에 관한 권리에 대한 침해가 발생할 위험이 크지 않다고 이해되어 왔음.
- 그러나 앞에서 검토한 바와 같이 생성형 AI가 학습 데이터 중 일부를 사실상 암기할 수 있는 역량을 갖추게 되면서, 개인정보보호에 관한 권리 침해 가능성에 대한 우려가 커지고 있음.

□ ChatGPT에 대한 외국 개인정보보호 당국의 조사<sup>42)</sup>

- 캐나다 사례
  - 캐나다 연방 프라이버시 위원회는 2023. 4. 4. ChatGPT 조사에 착수하였다고 밝혔음. 이는 OpenAI에 의한 동의 없는 개인정보의 수집, 이용 및 공개에 관한 complaint에 대한 대응임.

41) 조영선, “선의의 저작권 침해에 관한 법률문제”, 저스티스 제142호, 2014.

42) 다만 아래 사례들은 모델 학습 과정에서의 개인정보보호 문제만을 다룬 것은 아니라고 여겨짐.

### ● 이탈리아 사례

- 이탈리아 개인정보 보호당국은 2023. 3. 30. 잠정조치(Provvedimento n. 112)를 통해 OpenAI에 대해 이탈리아 이용자의 개인정보의 처리를 잠정적으로 금지하고, 사실관계 조사에 착수함.
- 이탈리아 개인정보 보호당국이 지적한 주요한 GDPR 위반 사항은 다음과 같음. ① 투명성 요건 위반: 개인정보가 처리되는 이용자 및 정보주체에게 어떠한 정보도 제공되지 않음. ② 수집 요건 위반: 개인정보의 수집 및 처리에 있어 적법 근거가 결여됨. ③ 개인정보의 부정확한 처리: ChatGPT에 의해 제공된 정보가 실제 정보와 일치하지 않은 사례가 발생함. ④ 미성년자 보호: 나이 확인 시스템이 존재하지 않음.
- OpenAI의 대응 조치<sup>43)</sup>에 따라 2023. 4. 28. ChatGPT 사용이 재개되었으나 여전히 이탈리아 개인정보 보호당국에 의한 조사는 진행되고 있음.

### □ 우리 개인정보보호법

- 공개된 개인정보를 정보주체의 동의 없이 수집하여 상업적으로 이용할 수 있는지 여부에 관해서는 현행 개인정보 보호법상 명문의 규정이 존재하지 않음.
- 위 쟁점에 관해 대법원은 이른바 ‘로앤비’ 사건을 통해 정보주체의 인격적 법익과 정보처리자 등의 법적 이익 간의 형량이 필요하다는 점을 제시하였음.<sup>44)</sup> 즉, “이미 공개된 개인정보를 정보주체의 동의가 있었다고 객관적으로 인정되는 범위 내에서 수집·이용·제공 등 처리를 할 때는 정보주체의 별도의 동의는 불필요하다”고 보았음.<sup>45)</sup>
- 생성모델이 학습용 데이터세트에 포함된 개인정보를 그대로 재현한다고 볼 경우, 정보주체는 원래 개인정보를 공개한 범위를 넘어서 예상치 못한 범위에서 추가적인 공개와 확산이 이루어질 수 있고, 정보주체가 원래 공개한 목적과는 관련성이 없는 맥락에서 생성이 이루어진다고 평가할 수도 있음. 하지만, 생성된

43) ① 개인정보 처리방침을 개선하고 회원 가입 시 이를 표시하였고, ② 유럽 개인에게 개인정보 학습으로부터 옵트-아웃(opt-out) 기능을 제공하였으며, ③ 이용자 요청 시 부정확한 정보를 삭제할 수 있는 메커니즘을 도입하고, ④ “Welcome back” 페이지를 통해 연령을 확인하였고, ⑤ 회원 가입시 생일을 명시하도록 요구하였음.

44) 대법원 2016. 8. 17. 선고 2014다235080 판결

45) 나아가 위 판례는 정보주체의 동의가 있었다고 인정되는 범위 내인지는 “공개된 개인정보의 성격, 공개의 형태와 대상 범위, 그로부터 추단되는 정보주체의 공개 의도 내지 목적뿐만 아니라, 정보처리자의 정보제공 등 처리의 형태와 정보제공으로 공개의 대상 범위가 원래의 것과 달라졌는지, 정보제공이 정보주체의 원래의 공개 목적과 상당한 관련성이 있는지 등을 검토하여 객관적으로 판단”해야 한다고 명시함.

정보가 학습에 이용된 개인정보와 매우 유사하다고 하더라도 그것은 우연일 뿐 생성된 정보를 개인정보로 볼 수는 없다는 반론도 가능할 것임.

#### (4) 인격권 침해의 위험성

##### □ 퍼블리시티권 침해의 문제

- 우리 법은 개인의 인격 표지에 대해 여러 권리를 인정하고 있으며, 인격 표지에 대한 권리 보호를 강화하는 방향의 입법들이 다수 진행되고 있음.
  - 예컨대, 법무부가 제안한 민법 일부개정안은 종래에 판례에 의해 인정되어 오던 인격권을 명문화하고자 함.
  - 또한 부정경쟁방지법은 2021년 개정을 통해 유명인의 식별표시 무단사용을 부정경쟁행위로 규정하였음(제2조 제1호 타목). 나아가 현재 국회에 계류 중인 저작권법 전부개정안은 초상등재산권을 신설하여 이른바 “퍼블리시티권”을 명문화하고자 함.

##### □ 그 밖의 인격권 침해의 문제

- AI 이용자는 의도적으로 타인의 인격권을 침해하는 출력물을 생성하기 위하여 생성형 AI를 활용할 수 있음. 다른 한편, AI 이용자가 의도하지 않았으나 타인의 인격권을 침해하는 출력물이 생성되는 결과도 발생할 수 있음. 예컨대 대규모 언어모델의 할루시네이션(Hallucination)에 의해 개인의 명예를 훼손하는 허위 정보를 생성할 수 있음. 최근 ChatGPT가 미국 로스쿨 교수가 제자를 성추행하려고 시도하였다는 허위 정보를 생성한 사례가 있음.
- 따라서 생성형 AI 서비스/도구 제공자는 예견 가능한 악의적 활용의 위험성을 막기 위한 적절한 안전조치를 취할 필요가 있음(예: 모욕적인 입력 프롬프트 필터링, 모욕적인 출력물에 대한 필터링 등).

#### (5) AI 모델 라이선스 및 이용약관의 위반

##### □ AI 모델 라이선스 위반과 저작권 침해

- AI 모델의 이용과 관련하여 라이선스나 이용약관을 위반한 경우의 효과가 문제됨. 다만 AI 모델 프로그램 저작권 침해의 문제는 생성 AI와 관련된 데이터 정책과 직접 관련된 문제는 아니라고 여겨짐.

- AI 모델 라이선스가 저작권의 본래적 내용에 해당하는 경우 ‘AI 모델 프로그램’의 저작권 침해에 해당할 수 있음(오픈캡처 판결<sup>46)</sup> 및 Jacobsen 판결<sup>47)</sup>). 특히 AI 모델의 이용약관 상 상업적 이용 금지 규정이 있는 등의 경우 본래적 내용에 해당하는지 문제될 수 있음.<sup>48)</sup>

#### □ 저작권 침해 효과의 귀속

- 한편, 이용자가 AI 모델 라이선스를 위반하여 이용한 경우에도 생성물이 학습 데이터 내 저작물과 실질적 유사성이 인정되는 경우, 생성물에 관한 권리는 이용자에게 귀속되고, 학습데이터 내 저작물에 관한 저작권 침해의 효과 역시 원칙으로 이용자에게 귀속됨.<sup>49)</sup>
- 생성물에 대하여 저작권과 같은 배타적 권리가 인정된다면 그 권리변동은 물론 법적으로 규율되며 채권적 라이선스에 의하여 규율되지는 않을 것으로 보임. 다만 민법상 침부의 법리와 마찬가지로 AI 모델 제공자와 이용자는 계약에 의하여 생성물에 관한 권리의 원시적 귀속을 정할 수 있을 것임.

### 마. AI 생성물에 대한 권리와 그 귀속의 문제

#### (1) 문제점

- AI 모델을 이용하여 생성된 생성물이 저작물로 인정될 수 있는지, 만약 인정될 수 있다면 그에 대한 권리는 누구에게 귀속되는지 문제됨.
- 부수적으로, 만약 AI 모델을 이용하여 생성된 생성물에 저작물성이 인정되지 않는다면 그러한 생성물의 거래와 관련된 법률관계는 어떻게 처리될 것인지도 문제될 수 있음.

46) 대법원 2017. 11. 23. 선고 2015다1017, 2015다1031, 2015다1024, 2015다1048 판결

47) Jacobsen v. Katzer, 535 F.3d 1373 (Fed. Cir. 2008).

48) AI 모델 이용자가 결과물 판매금지 조건 등 라이선스를 위반한 경우 라이선스가 계약의 성격을 띠는 경우 해제, 손해배상, 강제이행 등 채무불이행에 기한 구제가 가능함. 그러나 위와 같은 조건은 AI 모델 프로그램 저작권의 본래적 내용에 해당하지 않으므로 라이선스를 철회하고 저작권에 기한 금지청구를 하는 방식은 곤란할 것으로 보임.

49) 전응준, “인공지능 관련 저작권 침해에 관한 시론”, 경영법률 제31권 제4호, 한국경영법률학회, 2021.

## (2) AI 생성물의 저작물성

### □ 저작물의 요건

- 현행 저작권법은 저작물을 “인간의 사상 또는 감정을 표현한 창작물”로 정의하고 있음. 이러한 정의규정을 바탕으로 저작물의 성립요건을 도출하면, 첫째는 ‘인간의 사상 또는 감정’에 관한 것으로, 둘째는 ‘표현’한 것이어야 하며, 셋째는 ‘창작물’로서의 요건을 갖추어야 함.

### □ AI 생성물의 저작물성

- 인공지능 생성물은 ‘인간’의 사상과 감정을 표현한 창작물이 아니므로 원칙적으로 저작물에 해당하지 않음.<sup>50)</sup> 같은 이유로 AI는 저작권의 귀속 주체인 창작자가 될 수 없음.
- 미국 저작권청은 저작물이 기본적으로 인간에 의해 창작되고 컴퓨터 등은 보조 도구에 불과하였는지, 혹은 문학적, 예술적, 음악적 표현이나 구성요소의 선택, 배열 등 전통적인 저자의 요건이 실질적으로 인간이 아닌 기계에 의해 구상 및 실행되었는지를 중요하게 고려함.
  - 저작물의 전통적인 창작 요소(elements of authorship)가 기계에 의해 생성된 경우 저작물에는 인간 저작물이 없으며 저작권청은 등록하지 않을 것. 예를 들면 인간으로부터 프롬프트만을 입력받은 후에 AI가 텍스트, 이미지, 음악 작품을 만드는 경우 전통적인 창작요소는 인간이 아닌 기계에 의해 결정되고 실행되는 것으로 볼 수 있음.
  - 반면 AI가 생성한 자료를 포함하고 있는 경우라도 인간의 창작이 충분히 포함된 경우에는 저작물성을 인정할 수 있음. 예를 들면 AI가 창작한 자료를 인간이 선택 또는 배열하는 과정에서 충분히 창작적인 방식으로 이루어진 경우이거나, AI가 창작한 자료를 변형하여 2차적저작물을 생성한 경우에는 저작권 보호를 받을 수 있음. 다만 저작권의 보호범위는 인간의 창작이 포함된 부분까지이며, AI가 생성한 부분은 포함하지 않음

50) *Naruto v. Slater*, No. 16-15469 (9th Cir. 2018) 참조(원숭이가 찍은 사진은 저작권 등록이 불가하다고 판시함)

● 주요 사례

- Thaler 사건 : Steven Thaler가 자신이 발명한 Creative Machine을 저자로, 자신을 청구인(기계의 소유자)으로 한 미술저작물 ‘A Recent Entrance to Paradise’의 등록을 신청한 사건에서 미국 저작권청은 사람의 개입이나 창의적 입력 등이 있었다는 증빙이 없는 한 저작권 등록을 할 수 없다며 거절함.
- Zarya of the Dawn 사건 : 미국 저작권청은 Kashtanova가 작성한 글과 서술 외에도, Midjourney가 생성한 이미지 자체는 논외로 하더라도 해당 시각적 요소와 텍스트를 선택, 조정, 배열한 편집부분에 관하여는 창의적 선택의 산물이라고 인정함. Kashtanova가 만화책의 각 페이지를 구성할 이미지 및 대사 등을 선택하고 배치하는 등의 편집을 가한 것을 창의적 선택의 산물이라고 보고 저작권을 인정함.

(3) 저작물이 아닌 생성물의 보호 및 거래

□ 저작물이 아닌 생성물의 보호

- 저작물에 해당하지 않는 생성물이라도 일정한 요건을 갖추어 경우 콘텐츠산업진흥법 상의 콘텐츠로서(제37조) 또는 부정경쟁방지법 상의 성과물 무단이용 금지조항에 근거하여 금지청구를 함으로써 배타적으로 보호받을 수도 있음(제2조 제1항 파목, 제4, 5조).
- 그 밖의 경우에는 민법상 불법행위의 요건이 갖추어진 경우 손해배상의 방법으로 구제받을 수는 있으나, 금지청구에 의하여 배타적으로 보호받지는 못함.

□ 저작물이 아닌 생성물의 거래

- 저작권이나 그 밖의 제도에 의하여 배타적으로 보호받지 못하는 생성물이라도 계약에 의한 이전이나 이용허락 등의 거래가 가능함.<sup>51)</sup>
- 이론적으로는 배타적 보호가 인정되지 않는 경우 제3자의 침해에 대한 구제수단이 제한되어 시장 가격이 보다 낮게 형성될 가능성도 있으나, 시장에서 실제로 그러한 일이 발생하는지 여부는 조사가 필요함.

51) 이상용, “데이터 거래의 법적 기초”, 법조 제67권 제2호, 법조협회, 2018 참조.

## 바. 생성형 AI 생태계 활성화를 위한 데이터 정책 방향

### (1) 독점과 공유의 조화

#### □ 데이터에 대한 배타적 권리를 인정하는 근거

- 배타적 재산권의 정당성 근거에 관하여는 전통적으로 자유 논변, 효용 논변, 그리고 필요 논변 등이 제시되어 왔음. 로크의 노동이론이나 헤겔의 인격이론과 같은 자유 논변은 데이터의 경우에도 대체로 적용될 수 있으나, 공리주의에 입각한 효용 논변의 경우에는 데이터에 그대로 적용되기 어려운 점이 있음.<sup>52)</sup>
- 즉, 데이터는 유체물과 달리 다른 사람의 소비에 의하여 제한되지 않아 경합성이 없고, 따라서 ‘공유의 비극’으로 대변되는 자원 고갈에 대한 우려가 없기 때문에 원칙적으로 유용한 데이터는 공유되는 것이 사회적 이익을 증가시키는 길임. 그러나 유용한 데이터일수록 그 생산에 많은 노력과 비용이 필요하므로 그 생산 및 개방을 위한 인센티브가 부여될 필요가 있음. 지적재산권법은 이러한 인센티브로서 배타적 권리를 부여하되, 권리 존속기간을 제한하고 권리제한 제도를 마련함으로써 독점과 공유의 조화를 꾀하고 있음.

#### □ 생성형 AI와 관련된 데이터 생태계

- 생성형 AI의 경우 학습용 데이터세트, 그 소재가 되는 개별 데이터 또는 데이터베이스, 학습을 통해 만들어진 AI 모델, AI 모델을 이용하여 만들어진 생성물 등 다양한 데이터가 등장함.
- 이러한 데이터들은 저작권이나 개인정보보호법과 같은 기존의 법률에 의하여 배타적 권리가 부여된 경우도 있고, 엄밀한 의미의 배타적 지배권은 아니지만 부정경쟁방지법 등에 의하여 금지청구권이 인정되어 사실상의 독점권이 인정된 경우도 있음. 데이터베이스에 해당하지 않는 데이터세트 또는 AI 생성물의 경우 처럼 배타적 권리의 부여 여부가 논의 중인 경우도 있음. 그리고 배타적 권리가 인정되지 않는다는 점에 이론이 없는 경우도 있음.
- 이처럼 다양한 데이터들이 잘 유통되고 이용되도록 하기 위해서는 각 데이터의 활용 맥락에 맞는 방식으로 독점과 공유의 적절한 조화가 이루어질 수 있도록 할 필요가 있음.

52) 이상용, “데이터에 대한 배타적 권리의 법철학적 정당화”, 일감법학 제50호, 2021 참조.

## (2) 법적 불확실성의 해소

### □ 예측 가능성과 효율성

- 생성형 AI 생태계가 효율적으로 작동하려면 참여자들이 자신의 행위의 결과를 예측할 수 있어야 함.

### □ 데이터의 보호 측면

- 유용한 데이터 생산을 위해서는 배타적 보호를 인정하여 인센티브를 부여할 필요가 있음. 예컨대 체계성이나 검색가능성이 없는 데이터세트의 경우에도 데이터베이스의 범주에 속하도록 하는 방안이 검토될 수 있을 것임.<sup>53)</sup>
- 다만 독점과 공유의 조화를 위하여 보호의 범위와 수준은 적정해야 하며, 적절한 권리 제한 제도가 마련되어야 함.

### □ 데이터의 유통 측면 - 계약적 이용

#### ● 사회적 잉여 극대화

- 거래비용이 낮은 경우 데이터의 거래에 관한 자발적 합의는 사회적 잉여를 극대화하는 기제가 됨.<sup>54)</sup>
- 데이터의 거래는 데이터에 대한 배타적 권리가 전제되어야 하는 것은 아니지만, 데이터에 대한 명확한 권리의 설정은 유보가치를 명확히 함으로써 거래비용을 완화할 수 있음.<sup>55)</sup>

#### ● 표준 계약서의 보급<sup>56)</sup>

- 각 데이터의 법적 유형 및 활용 맥락에 따라 타인의 권리 비침해 보장,<sup>57)</sup> 데이터의 품질 및 담보책임, 제3자 배제 특약, 보안 및 관리 등에 관하여 적절한 내용을 담은 표준계약서를 보급하는 정책이 바람직함.

#### ● 개인정보 동의 제도의 개선

53) 이상용, “데이터세트에 대한 배타적 보호”, 인권과 정의 제503호, 대한변호사협회, 2022.

54) 이상용, “데이터 거래의 법적 기초”, 법조 제67권 제2호, 법조협회, 2018 참조.

55) 이상용, “데이터 거래의 법적 기초”, 법조 제67권 제2호, 법조협회, 2018 참조.

56) 이상용, 데이터세트 보호 법제에 관한 연구, 경인문화사, 2023, 261-269면.

57) OpenAI는 2023. 11. 6. 개발자 컨퍼런스에서 구글, 마이크로소프트, 어도비 등의 전례를 따라 사용자의 AI 법적 소송을 책임지겠다고 강조했다. 자사의 제품을 중심으로 하는 AI 및 데이터 생태계를 활성화하기 위한 것으로 이해됨. 지디넷, “오픈AI, GPT-4 터보 출시 ‘한번에 300페이지 이해’” (2023. 11. 7.) <https://zdnet.co.kr/view/?no=20231107072427>

- 개인정보보호법상 동의 제도는 데이터 이용 계약의 관점에서 파악할 수도 있음. 그런데 현행법은 지나치게 후견적 입장에 치우쳐 동의의 방식이나 내용이 지나치게 경직되고 비효율적임.
- 실질적이고 정보에 입각한 동의라는 원칙을 중심으로 하여 개인정보보호의 원칙에 부합하도록 규율하되, 포괄 동의와 같은 유연한 동의 방식을 허용할 필요가 있음.
- 한편 동의는 위법성 조각사유의 측면도 지님. 정보주체의 개인정보자기결정권은 반드시 사전적 동의를 요하는 것은 아니며 사후적 승인 내지는 처리 반대권의 행사를 통해서도 확보될 수 있음. 상황에 따라서는 후자가 보다 효율적인 동시에 정보주체의 실질적 자기결정을 확보할 수 있는 방법이 될 수 있음. 유용한 데이터의 대부분이 개인정보를 포함하는 상황에서 언제나 사전적 동의를 요하도록 하는 것은 데이터의 유통과 활용에 가장 큰 장애가 되고 있음.

□ 데이터의 유통 측면 - 비계약적 이용

- 저작재산권 제한 제도의 정비 : 공정이용 제도의 불명확성을 해소하기 위하여 데이터마이닝을 위한 면책조항을 마련할 필요가 있음<sup>58)</sup>. 다만 일부 생성형 AI와 같이 시장에 미치는 영향이 큰 경우도 포함시킬 것인지 추가적인 검토 필요
- 개인정보보호 제도의 정비 : 개인정보의 비동의 이용에 관한 제도를 보다 명확하게 정비할 필요가 있음. 개인정보보호법의 모델이 되고 있는 EU의 경우 모든 경우에 사전동의를 요구하기보다는 양립가능성의 법리나 이익형량의 법리 등을 통하여 활용하는 경우가 많음. 우리나라에서는 형사처벌 조항과 연결되어 EU 식의 원칙 중심 규제가 사실상 작동하기 어려운 점을 고려할 때 기존의 비동의 이용 근거 조항들의 요건이 보다 명확하게 규정되고, 표준적 사례들이 제공될 필요가 있음.

58) 공정이용 제도에 관한 법리가 현재로서 확립되었다고 보기 어려운 점을 고려할 때 데이터 마이닝을 위한 면책조항도 별도로 요구됨

### (3) 이해관계의 조정

#### □ 적절한 보상 체계에 관한 사회적 공감대 마련

- 생성 AI 생태계에서 가장 큰 이슈는 학습용 데이터세트에 포함된 저작물의 저작자 등에 대한 보상이 적절하게 이루어지는지 여부임.
- 저작권법을 비롯한 기존의 법제도는 이러한 이해관계의 조정에 취약한 것으로 여겨짐. 어떤 맥락에서는 저작권 침해의 증거가 어려워 저작권자에게 부당한 결과를 가져온다고 여겨지기도 하지만, 다른 맥락에서는 창작의 인센티브로 부여된 저작권이 남용되어 정보의 공유와 혁신을 가로막는 것으로 여겨지기도 함. 공정이용과 같은 저작재산권 제한 제도가 있지만, 불명확성으로 인해 이러한 우려를 가라앉히기는 쉽지 않음.
- 이는 근본적으로 생성 AI 생태계가 최근에 등장하여 급속히 확산된 탓에 아직 적절한 이해관계의 균형점에 관한 사회적 합의가 무르익지 않았기 때문으로 여겨짐. 따라서 이해당사자들의 적극적이고 자발적인 참여를 통해 이에 관한 공감대가 형성될 필요가 있음. 최근에 몇몇 기업들의 주도로 자율적인 보상체계 논의가 시작되고 있는 점은 바람직해 보임.

### (4) 플랫폼의 기능과 역할

#### □ 데이터 유통과 플랫폼

- 데이터에 관하여 다수 권리자가 존재하는 경우 자원의 과소 활용으로 이어져 거래비용을 상승시킬 우려가 있음. 특히 학습용 데이터세트에 다수의 개인정보와 저작물이 포함된 경우가 많은데, 데이터세트의 이용에 관하여 데이터베이스권자로부터 허락을 받았더라도 그 효력이 소재가 되는 개별데이터에는 미치지 않기 때문에 마찬가지로 결과가 생길 수 있음.
- 이러한 문제는 기술적 수단에 의하여 완화될 수 있음. 즉, 온라인 플랫폼은 권리의 파편화에도 거래비용을 낮은 수준에서 유지할 수 있게 해 줌.<sup>59)</sup> 따라서 데이터 거래를 위한 플랫폼을 장려하는 정책이 적절함. 다만, 데이터세트에 다수의 개인정보나 저작물이 포함된 경우 이러한 방식은 한계가 있으며, 권리 비침해 보장 특약과 같은 방식으로 내부화할 수밖에 없을 것으로 보임.

59) 한편, 법적 수단에 의해서도 이 문제에 대응할 수 있는데, 예컨대 데이터세트에 대하여 배타적 권리를 부여할 경우 권리의 수를 합리적인 수준으로 낮추는 방식으로 법제도 설계가 가능함.

□ 생성 AI 모델과 플랫폼

- 생성 AI 생태계는 기반모델을 바탕으로 그 위에 여러 레벨의 서비스가 계층화되어 최종 수요자에게 최적의 AI 모델을 제공하는 형태로 구성되고 있음.
- 또한 AI 모델이나 생성물 또는 프롬프트를 공유하고 거래할 수 있는 플랫폼도 다수 등장하였음.
- 나아가 OpenAI는 2023. 11. 6. 개발자 컨퍼런스에서 맞춤형 챗GPT를 ‘GPT 스토어’에 등록할 수 있다고 밝힘. 즉, 애플의 앱스토어나 구글의 플레이스토어와 같은 플랫폼 사업자가 되겠다는 의향을 내비침.
- 이는 기존 서비스를 자연어로 사용 가능하도록 하여 GUI(Graphic User Interface)로부터 NLUI(Natural Language User Interface)로 패러다임이 변화하는 것과 맞물려 상당한 반향을 일으킬 것으로 보임.
- 이러한 플랫폼은 다수의 수요자와 공급자를 연결하여 혁신을 촉진하고 거래비용을 낮춘다는 점에서 바람직함. 이에 대하여는 경쟁 저해에 관한 우려가 제기될 가능성도 있으나, 수요자로서는 여러 플랫폼 사이에서 비교적 손쉽게 이동할 수 있는 것으로 보일 뿐만 아니라, 설령 어느 플랫폼이 독점적인 성격을 띠게 되더라도 이는 네트워크 효과로 인해 평균비용이 하락한 데 따른 자연독점의 측면이 강하여 소비자 후생을 저해하는 것으로 단정하기 어려움.

## 2. 개인정보법제 개선방향

### 가. 생성형 AI 현황

#### (1) 생성형 AI의 의의

- AI에 대해서는 다양한 정의가 시도되거나 제안되고 있지만, 생성형 AI에 대한 법적 정의는 아직 찾아보기 어려움.
  - EU 인공지능법안, 미국 2019년 회계연도 국방수권법 Section 280(g)는 AI 자체가 아닌 ‘AI 시스템’에 대해 정의하고 있으며, 일본의 개인정보활용추진기본법(2016) 제2조 제2항은 인공지능관련기술을 “인공적 방법에 의한 학습, 추론, 판단 등 지적 기능의 실현 및 인공적 방법에 의하여 실현된 당해 기능의

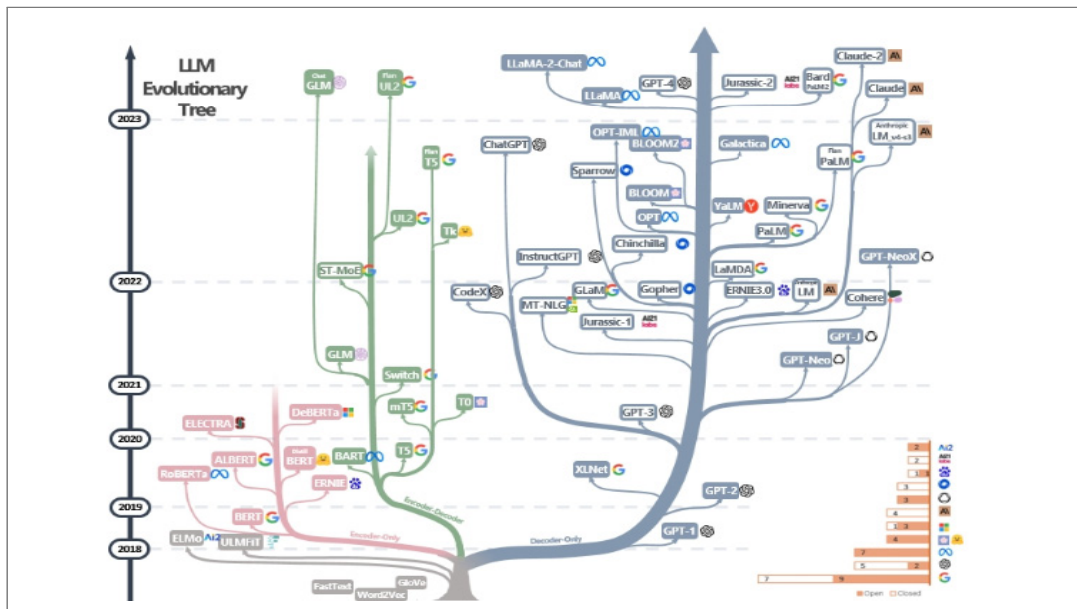
활용에 관한 기술(人工的な方法による学習、推論、判断等の知的な機能の実現及び人工的な方法により実現した当該機能の活用に関する技術)”로 정의하고 있음.<sup>60)</sup>

- Gartner에서는 생성형 AI를 ‘데이터 및 모델들로부터 인공물(artifact)들의 표시/표현(representation)을 습득하여 원본과 관련된 새로운 인공물 창출하는 인공지능 기술’ 또는 ‘어떤 모델에서 인공물들의 표시/표현을 습득하여 그와 유사한 성질을 지닌 새로운 인공물을 창출하는 인공지능 기술’로 설명.<sup>61)</sup>

(2) 생성형 AI와 플랫폼 데이터 정책의 상관관계

- 디지털 플랫폼의 가장 큰 특징 중 하나는 데이터 경제이며, 방대한 양의 데이터는 AI 발전의 핵심 기반이 됨.
- 현재 글로벌 플랫폼 기업들 중 AI를 개발하지 않는 기업은 찾아보기 어렵고(아래 [그림 III-1]<sup>62)</sup> 참조), 향후 플랫폼 생태계는 AI, 특히 생성형 AI 기반으로 재편될 가능성이 높음.

[그림 III-1] 대규모 언어모델(LLM) 발전 개요도



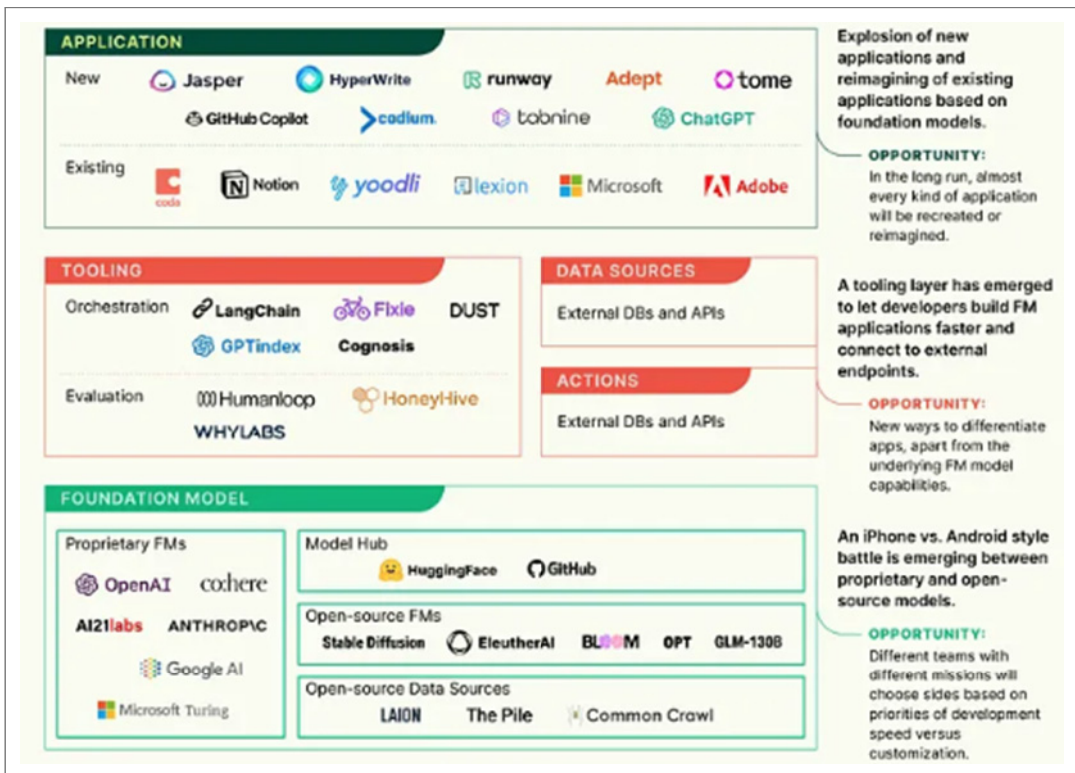
60) 장석천/이병삼, “인공지능에 의한 불법행위에 대한 손해배상책임”, 법이론실무연구, 제11권 제3호, 2023. 8., 290면 이하.

61) Gartner, “What is Artificial Intelligence?” (<https://www.gartner.com/en/topics/artificial-intelligence>).

62) Github, “The Practical Guides for Large Language Models” (<https://github.com/Mooler0410/LLMsPracticalGuide>).

- 향후 생성형 AI 생태계에서 특히 기반 모델(Foundation Model)을 둘러싸고 기존의 아이폰(i-Phone)과 안드로이드(Android) 사이의 모바일 OS 경쟁과 유사한 경쟁이 발생할 것으로 지적되고 있음.<sup>63)</sup>

[그림 III-2] 생성형 AI 생태계<sup>64)</sup>



## 나. 생성형 AI와 개인정보 관련 쟁점

### (1) 데이터셋(Data Set)을 통한 AI 학습

- 생성형 AI에서는 사전학습 기반 모델(pre-trained Foundation Model)을 활용하기 때문에, 학습을 위한 데이터셋(텍스트, 이미지, 음성 등)이 절대적으로 중요.
  - 데이터셋에는 텍스트, 이미지, 음성 등 비정형데이터가 모두 포함되며, 각종 개인정보와 위치정보가 활용될 수 있음.

63) Jon Turow/PalakGoel/Tim Porter, "Our View on the Foundation Model Stack", Madrona, 2023. 1. 27.

64) Id.

- 현재 대부분의 플랫폼은 「개인정보 보호법」<sup>65)</sup>에 따라 이용자의 동의를 거쳐 개인정보를 수집하고 있지만, 이를 AI 학습용 데이터로 활용할 수 있는지 여부, 활용하기 위한 방법 및 범위 등이 명확하지 않음.

## (2) 설명가능한 AI(Explainable AI)

- AI 시스템으로 개인정보를 처리하여 이루어지는 결정에 대한 설명의무 입법화(개정 개인정보법 제4조 제6호, 제37조의2).
  - AI 시스템에 의한 결정에 대한 합리적 추론요구권(right to reasonable inference)<sup>66)</sup>의 반영.
- 설명의무의 범위 및 실행방안이 상당히 불명확하기 때문에 산업 발전과 개인정보 보호의 조화를 도모할 수 있는 합리적 방안 마련 필요.
- 자동화된 행정결정의 경우, 설명의무 대상에서 제외되어 있어서 이 부분에 대한 보완 논의도 요구됨.

## (3) 행정조사/범죄수사, 사회적 평판 부여(Social Scoring), 취업정보활용 등 공적/공공적 활용

- 생성형 AI를 행정조사/범죄수사에 활용할 가능성이 확대되고 있는데, 이로 인해 정부기관에 의해 개인정보가 무분별하게 수집되고 편향된 결론을 도출하여 국민의 기본권을 침해할 우려가 있음.
  - 공개출처정보(OSINT) 분석: 공개정보를 바탕으로 이루어지는 OSINT(Open Source Intelligence; 공개출처정보획득)의 경우, 행정조사/범죄수사, Social Scoring, 취업정보 활용 등 다양한 곳에 활용될 가능성이 있음.
- 생성형 AI를 활용하여 사회적 평판 부여, 취업정보 등을 위한 개인정보를 수집하고, 편향된 결론을 도출하여 공공연하게 활용됨으로써 사회적 문제를 야기할 가능성도 있음.

65) 이하, '개인정보법'.

66) Sandra Wachter/Brent Mittelstadt, "A Right to Reasonable Inferences: Re-thinking Data Protection Law in the Age of Big Data and AI", Columbia Business Law Review, vol. 2019, no. 2(494), 2019.

#### (4) 자율주행 차량의 경우

- 자율주행 차량에서도 플랫폼화된 AI 활용이 증가하고 있으며, 클라우드 서비스 기반의 데이터 저장 및 이를 바탕으로 한 생성형 AI 활용 가능성이 높아지고 있음.
- 자율주행차량에 부착된 DSSAD, EDR과 같은 운행정보/사고정보기록장치를 통해 운전자와 동승자의 위치정보와 개인정보가 수집/활용될 수 있고, 자율주행차량 운행 중에 다른 차량, 보행자, 주변 환경 등을 촬영하여 인식하는 경우 불필요한 개인정보(특히, 생체정보)와 위치정보가 수집/활용될 수 있는데, 이 자체만으로는 생성형 AI와 거리가 있지만, 이를 생성형 AI의 학습데이터로 활용하여 서비스를 출시/제공하는 경우에는 문제가 발생할 수 있음.

### 다. 국내 법제도 현황

#### (1) 개인정보법

##### □ 기존 규정

- 제15조, 제16조에서는 개인정보의 수집/이용을 제한하고 있음.
  - 정보주체의 동의를 받거나(제15조 제1항 제1호), 법률에 특별한 규정이 있거나 법령상 의무를 준수하기 위하여 불가피한 경우(같은 항 제2호), 공공기관이 법령 등에서 정하는 소관 업무의 수행을 위하여 불가피한 경우(같은 항 제3호), 정보주체와 체결한 계약을 이행하거나 계약을 체결하는 과정에서 정보주체의 요청에 따른 조치를 이행하기 위하여 필요한 경우(같은 항 제4호) 등 제한적인 경우에만 이를 허용.
  - 또한, 목적에 필요한 최소한의 개인정보를 수집하고, 필요최소한의 정보 외의 개인정보 수집에 대해서는 별도의 동의를 받아야 하며, 필요최소한의 정보 수집 미동의를 이유로 한 재화/서비스 제공 거부를 금지하고 있음(제16조).
- 제18조에서는 개인정보의 목적 외 이용·제공을 제한하고 있음.
- 제28조의2에서는 개인정보처리자가 통계작성, 과학적 연구, 공익적 기록보존 등 목적이 있는 경우에만 정보주체의 동의 없이 가명정보를 처리할 수 있도록 규정하면서(제1항), 이러한 가명정보를 제3자에게 제공하는 경우에는 특정 개인

을 알아보기 위하여 사용될 수 있는 정보를 포함하지 않도록 규정하고 있고(제2항), 제28조의4에서는 가명정보가 원래의 상태로 복원되지 않도록 안전조치를 취할 의무를 부과하고 있음.

□ 2023. 3. 14. 개정에 따른 신설 규정(1) - 설명요구권

- 제4조에서는 제6호를 신설하여 “완전히 자동화된 개인정보 처리에 따른 결정을 거부하거나 그에 대한 설명 등을 요구할 권리”를 정보주체의 권리에 포함.
- 제37조의2는 자동화된 결정에 대한 정보주체의 권리에 관한 규정을 신설하였으며, 이 규정은 2024. 3. 15. 시행 예정.
  - 제1항은 “정보주체는 완전히 자동화된 시스템(인공지능 기술을 적용한 시스템을 포함한다)으로 개인정보를 처리하여 이루어지는 결정(「행정기본법」 제20조에 따른 행정청의 자동적 처분은 제외하며, 이하 이 조에서 “자동화된 결정”이라 한다)이 자신의 권리 또는 의무에 중대한 영향을 미치는 경우에는 해당 개인정보처리자에 대하여 해당 결정을 거부할 수 있는 권리를 가진다.”라고 규정하면서, 다만 자동화된 결정이 정보주체의 동의를 얻거나(제15조 제1항 제1호), 법률에 특별한 규정이 있거나 법령상 의무 준수를 위해 불가피한 경우(같은 항 제2호), 정보주체와 체결한 계약 이행, 계약 체결 과정에서 정보주체의 요청에 따른 조치를 이행하기 위해 필요한 경우(제4호)에 따라 이루어지는 경우에는 예외로 하고 있음.
  - 제2항은 정보주체는 개인정보처리자가 자동화된 결정을 한 경우에는 그 결정에 대하여 설명 등을 요구할 수 있도록 규정하고 있는데, 이 경우에는 제1항과 같은 예외 규정은 없음.
  - 제3항은 개인정보처리자가 “정보주체가 자동화된 결정을 거부하거나 이에 대한 설명 등을 요구한 경우에는 정당한 사유가 없는 한 자동화된 결정을 적용하지 아니하거나 인적 개입에 의한 재처리·설명등 필요한 조치”를 할 의무를 부여하고 있으며, 제4항은 개인정보처리자가 “자동화된 결정의 기준과 절차, 개인정보가 처리되는 방식 등을 정보주체가 쉽게 확인할 수 있도록 공개”할 의무 또한 부여하고 있음.

□ 2023. 3. 14. 개정에 따른 신설 규정(2) - 이동형 영상정보처리기기

- 개정 개인정보법은 이동형 영상정보처리기기에 대한 규정도 신설하였는데, 먼저 제2조 제7호의2에서 “이동형 영상정보처리기기”를 “사람이 신체에 착용 또는 휴대하거나 이동 가능한 물체에 부착 또는 거치(據置)하여 사람 또는 사물의 영상 등을 촬영하거나 이를 유·무선망을통하여 전송하는 장치로서 대통령령으로 정하는 장치”로 정의함.
- 제25조의2에서는 이동형 영상정보처리기기의 운영 제한에 관한 규정을 두고 있는데, 원칙적으로 공개된 장소에서 이동형 영상정보처리기기로 사람 또는 그 사람과 관련된 사물의 영상(개인정보에 해당하는 경우로 한정) 촬영을 금지하고 있음.
  - 다만, 제15조 제1항 각 호에 따라 적법하게 개인정보를 수집/이용할 수 있는 경우, 촬영 사실을 명확히 표시하여 정보주체가 촬영 사실을 알 수 있도록 하였음에도 불구하고 촬영 거부 의사를 밝히지 아니한 경우(단, 정보주체의 권리를 부당하게 침해할 우려가 없고 합리적인 범위를 초과하지 아니하는 경우로 한정) 등에는 예외 인정.
  - 또한, 이동형 영상정보처리기기로 사람 또는 그 사람과 관련된 사물의 영상을 촬영하는 경우에는 불빛, 소리, 안내판 등 대통령령으로 정하는 바에 따라 촬영 사실을 표시하고 알려야 한다고 규정.

(2) 「자율주행자동차 상용화 촉진 및 지원에 관한 법률」<sup>67)</sup>

- 자율주행자동차법 제20조는 자율주행차량 운행 과정에서 수집한 개인정보, 위치정보 등의 전부 또는 일부를 삭제하거나 대체하여 다른 정보와 결합하는 경우에도 더 이상 특정 개인을 알아볼 수 없도록 익명처리하여 정보를 활용하는 경우에는 개인정보법, 「위치정보의 보호 및 이용 등에 관한 법률」 및 「정보통신망 이용촉진 및 정보보호 등에 관한 법률」<sup>68)</sup>의 적용을 배제하고 있음.

67) 이하 '자율주행자동차법'.

68) 이하 '정보통신망법'.

## 라. 개선방향

### (1) 데이터셋을 통한 AI 학습

#### □ AI 학습데이터 수집/이용

- 현재 대부분의 플랫폼에서는 기본약관에서 개인정보 수집에 대한 당사자의 동의를 얻고 있지만, 이렇게 수집한 정보를 AI 학습용으로 활용할 수 있는지는 명확하지 않음.
- 데이터는 생성형 AI 학습에 필수불가결한 요소로 이의 확보가 전제되지 않으면 생성형 AI 모델 개발은 불가능하기 때문에 관련 규정에 대한 유연한 해석이 필요함.
- 개인정보보호위원회<sup>69)</sup>는 당초 수집 목적과 ‘합리적으로 관련된 범위’에서 ‘예측 가능’하고, 정보주체의 이익을 부당하게 침해하지 않으며, 안전성 확보에 필요한 조치를 한 경우 추가적 이용이 가능하다고 함.<sup>70)</sup>
  - 예컨대, 특정 서비스 제공 목적으로 수집한 개인정보를 해당 서비스의 개선(고도화) 목적으로 AI 개발에 이용하는 경우에는 별도의 동의 불요.
- 여기서 ‘합리적으로 관련된 범위’의 의미를 보다 유연하게 해석하여, 해당 서비스 자체 뿐만 아니라 해당 서비스와 연관된 AI 학습에 활용하는 것도 허용할 필요가 있음.
  - 예컨대, ‘해당 서비스의 개선(고도화)’을 ‘관련 서비스의 개선(고도화)’로 해석하고, AI 학습데이터 활용을 기본약관에 포함시켜 서비스 가입 단계에서 수집/이용이 가능한 범위에 포함시킬 수 있도록 해석.
- 필요하다면 opt-out 제도 도입도 고민해 볼 수도 있을 것.<sup>71)</sup>

69) 이하 ‘개보위’.

70) 개보위, 『인공지능 시대 안전한 개인정보 활용 정책방향』, 2023. 8., 6면.

71) 이유정/김민호, “인공지능 학습데이터와 개인정보보호법상 동의제도에관한 연구 - 미국의 사후동의(opt-out)제도를 중심으로”, 미국헌법연구, 제34권 제2호, 2023. 8. 참조.

□ 공개정보를 AI 학습용으로 활용할 수 있는 방법 및 범위

- 개보위는 ① 공개정보의 처리에 대한 이익 형량 결과, 처리 시 얻을 수 있는 이익이 더 큰 경우 허용하고, ② 공개정보를 크롤링 등으로 수집하여 가명처리 후 AI 학습 목적으로 이용하는 것은 가능하며(가명정보처리 목적: 통계작성, 과학적 연구), ③ 비식별화한 데이터는 AI 학습용으로 활용 가능하다는 입장.<sup>72)</sup>
- AI 학습데이터는 비식별화를 거치는 경우가 많아 개인정보법에 저촉되지 않는 경우가 다수임.
- 공개정보는 저작권법, 부정경쟁방지법, 정보통신망법 등 다른 법률에서 허용되는 경우에만 AI 학습용으로 활용할 수 있기 때문에, 개보법에서 추가적으로 규율할 필요성이 크지 않으며, 위와 같은 법들을 이미 준수했다면 처리 시 얻을 수 있는 이익이 막음으로써 얻는 이익보다 대체로 더 크다고 볼 수도 있을 것.
- 또한, AI 학습용으로 활용하는 것은 개인정보처리자의 '정당한 이익'으로 추정하여 줄 필요가 있음.

□ 가명정보 특례

- 가명정보의 경우, 과학적 연구와 같은 일정 목적 하에 정보주체 동의를 얻지 않고 활용이 가능한데, 개보위는 과학적 연구에는 기초연구/응용연구 뿐만 아니라 새로운 기술/제품/서비스 개발 및 실증을 위한 산업적 연구도 해당한다고 보고 있음.<sup>73)</sup>
- 현재, 개보위는 비정형데이터(이미지/영상, 음성, 텍스트 정보)의 가명처리 원칙, 식별 위험성 점검기준, 데이터 항목별 가명처리 방법 등 구체적 기준을 마련하고 있음.<sup>74)</sup>
- 이 외에, 개보위는 재식별가능성에 대한 걱정성 검토 수행,<sup>75)</sup> AI 출력 데이터에 대한 엄격한 필터링도 권고.<sup>76)</sup>

72) 개보위, 앞의 책, 7-9면.

73) 개보위, 앞의 책, 13면.

74) 개보위, 앞의 책, 13면.

75) 개보위, 앞의 책, 18면.

76) 개보위, 앞의 책, 14면.

- 가명처리의 경우, ‘과학적 연구’의 범위를 유연하게 해석할 필요가 있음. 즉, 산업적 연구가 궁극적으로는 영리 목적으로 활용될 가능성이 있더라도 해당 연구 자체가 일정한 수익을 대가로 행하여지는 것이 아니라면 가명처리 특례 적용.
- 현재 준비 중인 비정형데이터(이미지/영상, 음성, 텍스트 정보)의 가명처리원칙, 식별 위험성 점검기준, 데이터 항목별 가명처리방법 등에 관한 기준 또한 현실적으로 이행 가능한 것이어야 함.
- 가명/비식별처리를 위한 마스킹(masking)이나 출력 데이터에 대한 필터링 수단 사용에 관하여 정부가 가명/비식별처리 또는 필터링 수단/방법 등을 인증함으로써 기업의 향후 법 위반 가능성을 차단해 주는 방안도 고민할 필요가 있음.

## (2) AI 시스템을 통한 서비스의 투명성 제고를 위한 설명의무

- 개보위는 설명가능한 AI를 위한 AI 설명내용/방법 등 구체화 방안 마련 중.<sup>77)</sup>
- Hugging Face와 같은 AI 모델 마켓플레이스(모델 허브)에서는 해당 AI 모델이 활용한 학습데이터의 출처 및 작동원리 등을 명시하고 있음. 다만, 이러한 모델 허브의 밑바탕이 되는 기초 모델, 특히 전용 모델에서 이와 같은 학습데이터의 출처나 작동원리를 구체적으로 밝히도록 하는 것은 비현실적일 수 있음.
- 그러므로 설명의무의 발생 요건, 즉 ‘AI 시스템이 개인정보를 처리하여 이루어지는 결정이 정보주체의 권리 또는 의무에 중대한 영향을 미치는 경우’에 대한 합리적 해석 필요.
- 막연하게 모든 내용을 설명하도록 하거나인적 개입에 의한 재처리 등을 강제하기 보다는 권리/의무에 중대한 영향을 미치는 경우를 가이드라인 등에 한정적으로 열거하여 사업자의 예측가능성과 이행가능성을 보장할 필요가 있음.
  - 특히, 생성형 AI에서는 설명의무를 과도하게 요구하면 사실상 이행 불가능할 수 있음.
  - 또한, 기업들이 그럴듯한 것으로 둘러대더라도 검증할 방법도 없는 상황이기 때문에, 설명의무는 기업이 제시한 설명에 대한 자박(自縛)이라는 관점에서 신뢰성(trustworthy) 확보를 위한 도구로 활용하는 방안에 대해 고민할 필요가 있음.

77) 개보위, 앞의 책, 16면.

- 반면, 행정기본법 제20조의 자동적 처분의 경우에는 행정절차법에 따른 이유제시의무가 함부로 완화되어서는 안 될 것임.
  - 이에 대해서는 사전통지, 이유제시등에 대한 특례를 두어야 한다는 견해들이 제시되고 있지만,<sup>78)</sup> 현 단계에서는 행정의 무책임성('AI로의 회피')을 방지하고 국민의 방어권보장을 위해 신중해야 할 것.

### (3) 행정조사/범죄수사, 사회적 평판 부여, 취업정보 활용 등 공적/공공적 활용

- EU 인공지능법안에서는 행정조사/범죄수사에 AI 시스템을 활용하는 것을 금지하고 있는바, 현 단계에서는 무분별한 개인정보 수집, 공적 결정의 편향성 및 전국적 사찰 가능성 방지를 위해 이러한 행위를 아예 금지하는 것도 고려할 수 있음.
- 사회적 평판 부여, 취업정보 활용 등의 경우, 이를 공공기관이 행하는 것은 금지하고, 사기업이 행하는 것은 헌법상 대사인효를 갖는 평등권 침해 등을 근거로 제한할 필요가 있는바, 이에 대해서는 법적 근거를 마련해 두는 것이 바람직.

### (4) 자율주행차량

- DSSAD, EDR과 같은 운행정보/사고정보 기록장치는 자율주행 차량의 안전한 운행 및 향후 사고처리 및 책임소재 규명에 있어 필수적이기 때문에 익명처리 필요 없이 수집이 가능하도록 할 필요가 있고, 보관의무도 부과할 필요가 있으며, 다만 그 활용을 제한하는 방향으로 가야 할 것임.
- 자율주행차량 운전자와 동승자의 위치정보 및 개인정보의 AI 학습데이터 활용은 아직까지 그 활용 가능성이나 위험성이 검증되지 않은 측면이 있기 때문에 일반적인 AI 학습데이터의 경우와는 달리 서비스 개선 목적에 국한하여 인정할 필요가 있음.
- 자율주행 차량 운전자와 동승자의 위치정보 및 개인정보, 자율주행 차량 운행 중에 다른 차량, 보행자, 주변 환경 등을 촬영/인식하여 얻은 정보는 익명처리를 전제로 하여서만 AI 학습데이터로 허용해야 할 것임.

78) 정남철, “인공지능 시대의 도래와 디지털화에 따른 행정자동결정의 법적 쟁점 - 특히 행정기본법상 자동적 처분의 문제점을 중심으로”, 공법연구, 제50집 제2호(2021. 12.); 권은정, “자동적 처분의 행정법적 제 문제 - 처분의 형식·절차 및 구제에 관한 쟁점을 중심으로”, 법학논총, 제38집 제3호(2021. 9.) 등.

### 3. 데이터 거래 활성화 방안

#### 가. 서론

- 4차 산업혁명의 도래로 인해 정보화 혁명과 기술 융합 혁명이 주목 받고 있고, 이러한 변화로 디지털 데이터 및 IT(Information Technology) 산업의 중요성이 더욱 커지고 있으며(장영재, 2017), 디지털 데이터는 산업의 패러다임을 근본적으로 변화 시키면서 산업 간 경계를 허물고 있음.
  - 이를 통해 21세기 글로벌 경제와 산업의 주도권은 디지털 데이터에 의해 결정될 것으로 예상할 수 있음(연대성, 2017).
  - 데이터가 기존의 생산요소인 자본 및 노동을 능가하는 경쟁원천으로 부상하고 있으며, 앞으로는 데이터를 잘 수집하고 활용하는 것이 국가 및 기업의 경쟁력을 좌우할 것으로 보임(송현민, 2020).
- 미국, EU, 일본 등 주요 데이터 강국들은 미래의 데이터 기반 산업, 금융, 경제 시스템 측면에서 국제적 우위를 확보하기 위해 국가적 역량을 집중하고 있고, 이러한 과정 속에서 데이터 산업과 데이터 금융 시스템 육성에 초점을 맞추고 관련 법률 개정과 정책적, 제도적 혁신도 놓치지 않고 있음.
  - 우리 정부도 이러한 데이터 산업의 미래적 중요성을 인식하고 있으며, 데이터 인공지능경제 활성화와 데이터를 기반으로 한 산업과 경제의 육성을 미래 국정사업으로 지정하며 국가적 역량을 집중하고 있음.
- 국내에서도 데이터 중심의 생태계가 점차적으로 조성되고 있으나, 아직까지 데이터의 활용 노하우, 정보 부족 등 데이터를 활용하여 모든 산업에서 전반적으로 경쟁력을 제고하기에는 한계가 있음.
  - 이에 대한 근본적 해결책으로 업계에서는 ‘데이터 거래 활성화’를 가장 우선적 과제로 보고 있음.
- 국내 데이터 거래 및 유통 시장은 아직 초기 단계이며, 법 및 제도적 요인, 거래에 대한 인식, 물리적 기반 부족 등 사유로 사업이용자들이 플랫폼 데이터 접근에 어려움을 겪고 있음.

- 본 고에서는 먼저, 사업이용자의 플랫폼 데이터 접근 활성화의 필요성을 현재 국내 데이터 거래 및 유통 체계 상황과 함께 살펴보고, 현재 국내에서 처한 사업이용자 플랫폼 데이터 이슈와 과제에 대해 고찰해보고자 함.

## 나. 사업이용자의 플랫폼 데이터 접근의 필요성

- 2000년대 디지털 시대가 되면서 기존 아날로그 방식이었던 데이터 저장방식이 디지털로 전환되며 데이터 저장 및 처리 속도 측면에서 급격한 발전이 이루어짐.
  - 이는 빅데이터 시대의 시작을 알리는 계기가 되었고, 5G 통신 환경, 네트워크 발전, 모바일 기기의 대중화는 디지털 사회를 이끌어가는 주요 동력이 되었음.
    - 모바일 하드웨어를 통해 사람들의 일상이 데이터로 수집되어 저장되고, 데이터 그 자체는 인공지능, 블록체인 등 새롭게 등장한 IT 기술의 핵심 원자재로써 작용하고 있음.
    - 누적되고 있는 빅데이터를 분석해 제조, 유통, 금융 등 경영 분야 전반의 업무 효율성을 제고하거나 이전보다 훨씬 세분화된 고객 맞춤 전략 등을 통해 기업 생산성 향상을 이룰 수 있게 됨.
    - 해당 과정 속에서 발생한 데이터를 새롭게 활용하는 데이터 선순환 체계가 구현되는 데이터 시대가 시작되었음.
- 시대적 상황에 발맞추어, 우리 정부에서는 데이터 3법(개인정보 보호법, 신용정보법, 정보통신망법)을 개정하며, 개인 데이터의 활용도가 이전에 비해 높아짐.
  - 이전까지는 개인 데이터에 대한 특별한 규정이 존재하지 않았지만, 개인정보 보호법 개정과 데이터 기본법 제정으로 모든 산업의 개인정보 활용을 망라하는 법·제도적 기틀이 마련되었음.
- 그러나 데이터 유통 및 거래 시장은 아직까지 초기 단계의 시장으로, 법·제도적 장치 및 사업 추진 기반이 미약함.
  - 데이터 소유권 관련 법적 틀 미비, 가격 산정의 기준과 체계 부재, 데이터 품질 및 표준 문제 등 아직도 해결해야 할 문제들이 많음.

- 본 장에서는 사업이용자의 플랫폼 데이터 접근 이슈 및 과제를 논하기 이전에, 사업 이용자들이 플랫폼 데이터 접근 필요성의 배경이 되는 데이터 3법 개정에 대해 살펴보고, 현재 미국, 중국, 일본 등 데이터 강국들과 국내의 상황을 비교하는 등 국내에서의 데이터 거래와 활용의 현황이 어떠한지 살펴볼 것임.

### (1) 데이터 3법의 개정

- 2020년 1월, 국회는 개인정보 보호 관련 3개의 법률, ‘데이터 3법’으로 알려진 개인정보 보호법, 정보통신망법, 신용정보법의 개정안을 승인하였고, 이 법안들은 개인정보의 중요성에 대한 관심이 증가함에 따라 전 국민적으로 주목받았음.
  - 특히 개인정보 보호법 개정안은 안전한 환경에서 개인정보 활용을 위한 사회적 규범과 시스템을 강화하기 위해 채택됨.
    - 이로 인해, 개인정보의 개념을 재정의하고, 가명정보 및 익명정보의 적절한 활용을 가능케 하면서, 개인정보 보호 위원회의 설립으로 체계적인 관리 주체를 마련할 수 있게 되었음.
  - 정보통신망법 개정안은 정보통신망을 통해 발생하는 불법적인 정보 유출을 예방하기 위해 마련됨.
    - 법안은 불법적 정보통신망 접근을 ‘정보통신 침해 사고’로 규정하고, 이에 대비하기 위한 조치를 구체화하였으며, 침해 사고의 범위 또한 확대하여, 사전 예방에 중점을 두었다는 것이 요지임.
  - 신용정보법 개정안은 4차 산업혁명에 주요 역할을 하는 개인신용정보에 대한 접근 및 활용을 주제로 하였음.
    - 법안은 공정한 신용산업계와 시장 구조를 조성하고, 데이터 기반 산업의 발전 촉진을 골자로 하였으며, 이에 따라 개인 맞춤형 금융상품의 제공과 같은 다양한 혁신을 기대할 수 있게 되었음.
    - 이러한 노력은 정보화 시대에 필요한 법적 기반을 마련하고, 부작용을 방지하기 위한 조치를 취함으로써 국가 경쟁력을 향상시키는데 도움을 줄 것으로 예상됨.

- 데이터 3법 개정안은 데이터의 제공 및 활용에 관한 법적 규제 개선을 목표로하고 있기 때문에, 데이터 산업계에 상당한 영향을 미칠 것으로 예상됨.
  - 또한, 21세기 글로벌 정치, 경제, 사회, 문화, 일상생활 등에도 이러한 변화가 큰 영향을 미칠 것으로 예측되며, 이에 따라 데이터 3법이 국내 데이터 산업계에 미칠 영향을 예측하는 것은 매우 중대하다고 할 수 있음.
  - 특히 신용정보법 개정안 중에서 도입된 ‘본인 신용정보 관리업’ 또는 ‘마이데이터’ 관련 규정은 개인들이 자신의 신용정보를 주도적으로 관리하고, 이를 통해 보다 나은 서비스를 선택할 수 있도록 돕는 것으로 평가됨.
    - 이로 인해, 개인들은 본인의 신용정보를 선별적으로 업체들에 제공하면서 더욱 개인 맞춤형 서비스를 받을 수 있게 되고, 이는 데이터 산업의 근간이 될 것으로 예상됨.
    - 이와 더불어, 주요 데이터 관리 업체와 협력관계를 맺고 있는 기업들의 입장에서 고객 개인정보의 효과적인 활용이 비즈니스 성과와 브랜드 인지도에 큰 영향을 줄 수 있을 것으로 예상됨.
- 마이데이터 기반의 융합 데이터 산업이 전 세계적으로 성장하는 추세임을 감안할 때, 데이터 3법의 개정과 시행은 우리나라 데이터 기반 산업 활성화와 국제 경쟁력 확보를 위한 필수적 단계라고 할 수 있음.
  - 이에 따라 정부의 정책적 지원과 관련 인프라 구축, 모바일 애플리케이션 보급 등 국가적인 투자와 노력이 중요한 요소로 작용할 것으로 판단됨.

## (2) 데이터 거래 및 활용

- 본인 신용정보 관리업 또는 마이데이터의 적극적인 활성화는 데이터 거래 산업 및 데이터 거래 플랫폼의 활성화와 대중화에도 직접적인 영향을 미칠 것임.
  - 데이터 3법 개정으로 인해, 기존에는 주로 금융 분야에 한정되었던 개인정보 통합 조회와 제공, 그리고 관련 인프라가 다른 산업 분야로 확장되거나 이종 산업 간에도 합법적이고 산업적인 개인정보의 교환, 공유, 활용이 가능해질 경우, 데이터 기반 서비스는 산업 구조 전반에 큰 변화를 가져올 것으로 예상됨.

- 그러나 이를 원활하고 안정적으로 정착시키기 위해서는 개인정보의 안전한 이동을 보장할 수 있는 금융사 간, 금융사와 일반 기업 간, 그리고 이종(異種) 산업 간의 공동 시스템이 사전에 구축되어 있어야 함.
- 이를 위해 API (Application Programming Interface) 망의 설정, 금융사 내부 정보 보호 정책 수립 등 다양한 준비가 필수적임.
  - 서비스를 강제로 도입하고 문제가 발생하는 경우, 시간과 기회비용, 그리고 금전적 손실 등이 발생할 수 있으므로, 보여주기식 성과주의는 피하고 엄격하고 안정적인 기준을 가진 데이터 거래 시스템과 플랫폼을 구축해야 함.
  - 또한, 미래 선도 산업 중 하나인 데이터 거래 산업을 원활하고 합리적으로 정착시키기 위해 개인 금융 소비자의 권리 보장, 개인 신용 및 금융 정보의 통합적 관리, 대체 신용 평가 및 맞춤형 금융 상품 추천과 같은 다각적이고 포괄적인 서비스가 개발되고 제공되어야 할 것임.

### (3) 데이터 거래 현황

- 국내에서는 현재 주요한 금융 데이터 거래 플랫폼으로 한국 데이터 거래소(Korea Data eXchange, KDX)와 금융 데이터 거래소(Financial Data Exchange)가 운영되고 있음.
- 두 플랫폼 모두 통계 및 가공된 데이터를 위탁해 판매하여, 데이터 3법 시행 이후에는 개별적으로 결합 가능한 가명 처리된 데이터 판매를 새롭게 시작하였음.
  - 2019년 12월에 시작된 한국 데이터 거래소는 현재 17개의 공급업체가 15종 가량의 데이터를 제공하고 있음.
  - 금융보안원이 운영하는 금융 데이터 거래소는 2020년 5월에 시작되었고, 46개의 공급자가 12종의 데이터를 판매하고 있음.
- 데이터 3법 개정안이 시행된 이후, 두 거래소에 입점한 데이터 보유 기업들은 회사 규모에 따라 직접적 또는 간접적으로 데이터 판매 사업을 시작하거나 계획하고 있음.
- 대기업들은 자체적으로 데이터 판매를 진행하고, 중소기업들은 데이터 거래소에 위탁하는 방식을 선택할 것으로 예상됨.

- 자체 사업 부문의 운영 능력이 상대적으로 부족한 중견 및 중소기업들은 데이터 거래소의 위탁 판매 방식을 선택하는 것이 더욱 적절한 전략일 수 있음.
  - 데이터 거래소는 기업을 위한 데이터 위탁 판매, 결합, 분석 등의 서비스를 제공하여 높은 수익을 창출할 수 있을 것임.
- 앞으로 데이터 거래소나 거래 플랫폼, 그리고 데이터 제공 기업들이 안정적이고 지속 가능한 성장을 이루기 위해서는 비즈니스 전략 모색, 기술, 서비스 발전을 위한 지속적인 노력이 필수적임.
- 이러한 데이터 거래 산업 및 관련 플랫폼이 안정적으로 유지될 때, 개인정보의 명확한 가명처리와 안전한 거래 절차가 필수적이라고 할 수 있음.
- 다음 <표 III-1>은 해외 데이터 강국이라고 할 수 있는 미국, 중국, 일본의 데이터 거래소 현황과 국내 데이터 거래소 현황을 비교 및 정리한 자료임.

#### 다. 사업이용자의 플랫폼 데이터 접근 이슈 및 과제

- 데이터 경제 시대에는 데이터 생산 및 수집, 저장 및 처리, 유통 및 활용으로 이어지는 데이터 생태계 가치사슬을 통해 모든 산업 내에서 고객 맞춤 제품 및 서비스, 효율성과 생산성의 향상 등의 경제적 가치가 창출되어야 하는데, 이러한 생태계 속에서 데이터가 원자재로서 효율적으로 활용되기 위해서는 데이터 유통과 거래가 활성화되어야 함.

〈표 III-1〉 주요국 데이터 거래소 현황

구분	미국	중국	일본	한국
대표 데이터거래소	액시엄 (1969)	귀양 빅데이터 거래소 (2017)	정보은행 (2019) Eversense (2016)	금융데이터거래소 (2020) KDX 한국데이터거래소 (2019)
유형	민간 대기업 주도	정부 및 국영/민영 합작	IoT 데이터거래소, 정보은행(신탁기능)	정부주도하에 민간기업 참여
업체 수	2,500개 이상	4개	6개 시범지정	10여 개
공급자	정부, 기업	정부, 기업	정부, 기업, 개인	정부, 기업
수요자	정부, 기업, 데이터브로커	기업	기업	기업
개인 보상 지불 여부	일부 업체만 제공	×	○	×
정보제공 개인 동의 여부	×	×	○ (익명정보 제외)	×
가격	판매자 자체 결정	협약, 경매, 콜옵션	정기계약으로 월간 5만엔, 그 외 레시피 당 5천엔	협약
데이터 종류	개인 인적사항, 금융정보, 취미, 의료, 소비행태, 라이프스타일, 상품구매, IoT, 위치, 정치성향 등	금융, 정부, 의료, 기업, 전자상거래, 에너지 등 30개 분야 4,000여 종 데이터 품목 보유	부동산, 음식점, 의료, 제조업, 정보통신, 설문조사, 고용, 금융, IoT 센서 데이터 등	카드, 은행, 신용, 부동산, 보험, 통신, 기업정보, 생활 등
데이터 수집 방법	공공데이터, 사적 계약, SNS 등 웹 크롤링 및 제휴	공공데이터, 회원사 등록	공공데이터, 회원사 등록	공공데이터, 회원사 등록
장점	사후거부권 (Opt-out) 제도로 개인정보 수집 용이	정부와 민관 합작법인 설립으로 일반 사설 거래소보다 공신력과 신뢰도 확보	데이터 관리 전반을 전문가 (정보은행 사업자) 가 수행하여 데이터 효율적 활용 관리 가능	개인 동의를 얻은 정보만 제공 가능하여 개인정보의 유출 위험 낮음
단점	독과점 형성 및 개인 동의 없이 데이터 수집으로 법적 문제 발생 가능성	데이터 생태계 구축이 되는 산업 데이터 보다는 정부 데이터 중심으로 구성	개인 데이터 거래를 책임지는 정보은행 사업자의 적격성 판단이 중요	데이터 결합이 전문가를 통해서만 가능하여 데이터 활용에 제한적임

출처: 데이터 거래 활성화를 위한 제언, 제29차 4차 산업혁명위원회 보고 안건 제1호, 2022.4.18

- 그러나, 아직 초기 단계인 데이터 유통 및 거래 시장을 활발하게 이용해야 할 사업이용자들은 시장에서 아직도 쓸만한 데이터가 없다고 이야기하며, 무엇보다 모든 거래에 있어 기초라고 할 수 있는 데이터의 가치와 가격 산정의 기준과 체계가 부족하여 기업 간 데이터 활용에 제약이 있는 상황임.
- 본 장에서는 사업이용자의 플랫폼 데이터 접근의 이슈와 과제를 데이터 유통 기반, 데이터 가격, 데이터 품질 등 세 가지로 분류하여 고찰해볼 것임.
- 데이터 접근의 문제점을 데이터 유통 기반 측면, 데이터 가격 측면, 데이터 품질 측면 등 세 가지 측면으로 분류한 배경은 2021년 하반기에 실시된 데이터 산업 현황조사의 결과 보고서(과학기술정보통신부, 한국데이터산업진흥원, 2021)에서 ‘데이터 사업을 수행하였을 때 애로사항’에 대한 조사 결과에 의한 것으로, 해당 결과치에 대한 내용은 아래 <표 III-2>에 기술되어 있음.

<표 III-2> 데이터 사업 수행 시 애로사항

구분	세부 항목	비율(%)
데이터 유통 기반	데이터 유통 채널 부족	36.3
	데이터 소재 파악 및 검색의 어려움	31.5
	데이터 거래 절차 및 방법 미숙	2.1
데이터 가격	구매 데이터의 불합리한 가격 정책	37.0
데이터 품질	쓸만한 양질의 데이터 부족	52.1
	데이터 품질 문제	6.8
	데이터 가공 문제	1.4
기타	개인정보 포함 데이터 유통 활용시 법적 문제	7.5
	개인정보 처리 기술력 및 예산 부족	6.2

출처: 2021년 데이터산업현황조사 결과보고서, 2021

**(1) 데이터 유통 기반 관련 이슈 및 과제**

- 데이터를 가지고 있는 기업 또는 기관에서는 데이터를 가공하고 분석해서 도출해낸 결과에 대한 권리가 법적으로 보호되지 않음.
  - 이러한 이유로 데이터 보유 기업 및 기관이 적극적으로 데이터 관련 제품 및 서비스 개발을 추진할 유인이 부족한 상황이라고 할 수 있음.
    - 이와 더불어, 데이터 수요자는 비즈니스에 필요한 데이터가 어떤 기업 또는 기관이 보유하고 있는지, 구매가 가능한지 파악해내기 어려움.
    - 이와 반대로 데이터 공급자 입장에서도 자사의 데이터를 필요로 하는 데이터 수요자를 탐색하기 어려움.
- 이러한 문제를 해결하기 위해 정부 차원에서, 데이터 플랫폼 발전 전략 등 데이터 정책을 마련해 여러 산업 분야의 데이터 플랫폼을 구축하고 있음.
  - 예를 들어, 2019년 과학기술정보통신부는 데이터 생산, 유통, 분석, 활용 등 모든 데이터의 사이클을 지원하는 16개 분야별의 빅데이터 플랫폼을 운영하고 있음.
  - 다만, 데이터 경제 초기에는 민간의 자발적 데이터 거래를 촉진하기 위한 마중물로서 플랫폼 구축을 우선하였으나, 이후 데이터 경제가 진전됨에 따라 데이터 플랫폼 내에 축적된 데이터는 모든 분야의 기업들이 마주하고 있는 모든 문제를 해결할 수 있는 데이터를 갖추는데 한계가 발생.
- 결론적으로, 데이터 공급자 중심의 플랫폼이 아닌, 데이터 수요자의 실질적인 수요에 맞춘 데이터 소재 정보를 제공하고, 데이터 권리관계 정립 등 데이터 거래의 중개 역할이 가능한 플랫폼을 구축하고 그 역할을 담당하는 데이터 중개 사업 육성이 필수적인 과제라고 할 수 있음.
- 해당 과제에 대한 접근 방법으로는, 공급되는 데이터에 대한 정보 부족 및 이종 산업 간 데이터 융합의 한계점을 해결하기 위해 거래소 내 데이터 상품과 서비스의 다양화 및 홍보가 필요할 것이며, 데이터 검색과 거래 간 유기적인 관리가 요구될 것임.
  - 또, 데이터 중개 역할의 강화를 위해 데이터 발굴 및 수집과 관련하여 데이터를 발생시키는 다양한 미디어를 활용해 고객 상담, 데이터 기획, 데이터 산업 환경 평가 등을 수반할 수도 있을 것임.

## (2) 데이터 가격 관련 이슈 및 과제

- 데이터 거래소의 유통을 활성화시키기 위해서는 데이터 수요자가 필요한 품질의 데이터를 적절한 가격에 제공하는 것이 중요함.
  - 공급자 데이터에 대한 정보가 많을수록 데이터 수요자는 스스로에게 적합한 데이터를 선택할 가능성이 커지며 이로 인해, 데이터 유통 사례가 증가하여 데이터 거래 시장이 활성화될 수 있음.
- 그러나 국내 데이터 거래소는 대부분 데이터의 품질과 관련된 정보와 가격을 데이터 공급자가 자체적으로 제공하는 오픈 마켓의 형태로 운영되고 있음.
  - 이로 인해, 데이터 수요자는 품질을 판단할 수 있는 기회에 제약이 생기며, 데이터 거래소의 활성화를 위해서는 데이터 공급자들이 가격을 결정하는 방식에 대해 이해하고, 데이터 수요자들이 원하는 거래 조건을 제시하기 위한 유인 요인을 탐구하는 것이 선행되어야 할 것임.
- 데이터의 가격은 데이터의 가치에 다양한 시장적인 요소들이 결합하여 형성됨.
  - 하지만 데이터가 다른 자산과 달리 가지는 특징은 데이터의 활용 방식이 방대하고 데이터를 재사용하는 방식을 파악하기 어려워 데이터의 가치를 판단하고 이에 따른 적절한 가격을 산정하는데에는 어려움이 따름(Liang et al., 2018).
- 데이터 가격의 산정 방식으로는 원가접근법, 시장접근법, 소득접근법이 활용되고 있음.
  - 원가접근법은 평가 대상의 자산을 개발해내는 데 투입된 비용으로 가치를 산정하고, 산정이 어려울 경우, 동일하거나 유사한 경제적 이익을 보유한 기술 개발 또는 이와 동일하거나 유사한 가치의 기술을 구현해내는데 투입될 수 있는 원가를 추정하여 가치를 산정하는 방식임.
    - 이는 특정 데이터를 직접 구축하기 위해 현재 시점으로 소요되는 원가를 기준으로 데이터의 가치를 평가하는 방식이라고 할 수 있음.
    - 산정 시 포함되는 원가는 데이터의 기본 프로세스인 수집, 저장, 가공, 활용의 단계에 따라 데이터 중개비용, 데이터 플랫폼 운영비용, 데이터 통합비용, 데이터 보안비용, 데이터 기록비용 등이 있음.

- 원가접근법은 시장접근법의 적용이 어렵거나 매출에 대한 데이터의 기여분이 확실하지 않을 경우 사용되며, 데이터를 손실할 경우 위험 가치를 측정하는데에도 주로 사용됨.
- 시장접근법은 시장에서 거래되고 있는 가격 또는 전문가들의 조사 및 연구로 데이터의 가치를 결정하는 방식임.
  - 해당 방식은 일반적으로 시장에서 데이터 수요자 및 데이터 공급자 간 합의에 의해 가치가 평가되고, 데이터 수요자의 지불의사(Willingness-to-pay) 보유 여부가 영향을 미치기도 함.
  - 시장접근법의 데이터 거래 방식의 형태는 주로 데이터 수요자가 데이터를 구독하는 형태로 거래되지만, 시장 내에서 이전의 거래 사례가 없을 경우에는 로열티를 공제하는 방법을 사용하기도 함.
  - 시장접근법은 데이터를 통해 기업의 비용을 최소화하거나 신제품 개발, 신시장 개척, 매출을 확대하는 경우 등에 주로 사용됨.
- 소득접근법은 기업 내의 데이터 자산이 가지는 경제적 수명 속에서 발생할 것으로 예상되는 금전적 이익에 할인율을 적용해 현재가치를 추정하여 산정하는 방식임.
  - 소득접근법은 원가접근법이나 시장접근법에 비해 상대적으로 데이터의 가치가 정확하고 높게 평가될 수 있는 방법임.
- 데이터의 가치 평가 방식에 대해 기업 및 정부의 합의된 표준은 아직 없지만 2013년 OECD 보고서에 개인정보의 가치를 평가하는 가이드라인이 제시되어 있음.
  - 이는 시장가치를 기반으로 한 방법과 정보주체의 가치평가를 기반으로 한 방법으로 구분되어 있음.
    - 시장가치를 기반으로 한 방법은 개인정보 기록을 보유하고 있는 기업의 주가, 매출액, 순이익을 기준으로 하는 방법, 데이터 유출 사고 발생시 보상액을 기준으로 하는 방법, 암시장 내에서 데이터 가격을 기준으로 하는 방법으로 구분됨.
    - 이와 달리, 정보주체의 가치평가 기반 방법으로는 정보주체가 개인정보에 대해 주관적으로 부여한 가치를 조사하는 방법, 정보주체가 자신의 프라이버시를 지키기 위해 지불할 의사를 기준으로 하는 방법으로 구분됨.

- 그러나 위에서 제시한 데이터 가치 평가 및 가격 산정 방식과는 별개로, 데이터의 가격이 적절한지에 대한 이견은 데이터 공급자와 수요자 입장에 따라 큰 차이를 보일 수 있음.
  - 데이터 가격의 불합리성은 데이터 수요자는 자신의 지불의사보다 데이터의 가격이 높고, 데이터 공급자는 자신이 받아야 하는 가격보다 데이터 가격이 낮다는 인식에서 기인함.
    - 데이터 수요자의 경우 데이터를 활용하며 창출되는 가치를 쉽게 가늠하기 어려워 공급자가 제시한 가격에 직면했을 때 수요자가 활용하고자 하는 가치에 비해 데이터 가격이 상대적으로 높게 책정되었다고 인식할 수 있음.
    - 이와 반대로, 데이터 공급자의 경우, 보유하고 있는 데이터가 개인정보에 해당 하는지, 제3자의 제공 동의를 필수적으로 받아야 하는지 또는 데이터 저작권, 소유권의 귀속 여부 등 법·규제에 따른 필요 조치가 수반될 수 있기 때문에, 이와 같은 사항들을 데이터 가격을 상승시킬 수 있는 요인으로 인식할 수 있음.
- 결론적으로, 데이터를 활용하는 것에 대한 가치를 잠재적 데이터 수요자 및 공급자가 제대로 평가할 수 있는 명확한 가격산정 기준의 부재로 데이터 공급자와 수요자 모두 데이터 시장 거래가격이 불합리하다고 인식하는 것이 사업이용자의 플랫폼 데이터 접근의 이슈와 과제 중 하나라고 할 수 있음.

### (3) 데이터 품질 관련 이슈 및 과제

- 2022년 데이터 거래 및 유통을 활성화하기 위해 ‘데이터 산업 진흥 및 이용 촉진에 관한 기본법(데이터 산업법)’이 시행되면서 데이터 거래소 활성화에 필요한 법적 제도가 정비되고 양질의 데이터 공급을 위한 거래가 성사될 수 있는 시장 환경을 구축하는 기준이 마련되었음(김정연, 2022).
  - 그럼에도 불구하고, 데이터 품질에 대한 정보를 제공하거나 데이터를 표준화와 관련된 규정 등 데이터 수요자를 위한 규정은 포함되지 않아, 시장 참여자 간에 정보 비대칭 문제가 지속되고 있는 상황이며, 이에 대한 해결책을 과제로 남기고 있음.

- 현재 개인정보는 결합 전문기관 또는 데이터 전문기관을 통해 가명정보 처리 방식으로 개인정보의 결합기로 합쳐져 가명 처리 후에야 데이터로 활용할 수 있음.
  - 그러나 가명 처리는 외부 전문기관을 통해서만 가능하며 가명 처리 소요 기간이 최소 1개월에서 많게는 3개월까지 소요된다는 점에서 유형 또는 무형의 비용이 발생함(이종서, 이정현, 2023).
  
- 개인정보 이슈 및 데이터 소유권 등의 법적인 문제 때문에 비식별화된 통계형 데이터가 데이터 유통 시장 내에서 주된 거래 대상이 되고 있고, 이로 인해 데이터의 내재적인 가치가 감소하며 데이터의 활용도 역시 떨어짐.
  - 이러한 통계형 데이터 중심으로의 데이터 유통은 데이터를 분석하고 가공하기 위한 원시데이터(Raw Data)가 부족한 상황을 낳게 됨.
  
- 데이터의 실질적인 산업적 활용성을 제고하기 위해서는 통계형 데이터가 아닌 데이터의 개별적인 속성을 활용할 수 있는 원시데이터의 거래와 유통이 활성화되어야 함.
  - 이를 통해 이종 산업 간 데이터의 결합 및 연계가 가능해져 데이터 활용에 대한 가치가 더욱 증가할 수 있기 때문임.
    - 그렇다고 원시데이터 활성화가 활용성 제고를 위한 온전한 해결책이 되는 것은 아님.
    - 원시데이터 활용 가능성이 높아진다고 하더라도 많은 기관 및 기업들이 보유한 데이터의 분류, 명칭 등이 제각기 달라 데이터의 상호연계성이 떨어져 데이터를 연계하고 변환하는데 많은 시간적, 경제적 비용이 발생하기도 함.
  
- 결론적으로, 산업별로 데이터를 연계할 수 있도록 데이터의 표준화가 필요하고, 이과 더불어, 불필요한 변환 작업 최소화 및 데이터 품질 향상을 꾀해야 할 것임.
  - 데이터 수집과 유통단계에서 개인정보의 활용을 제한되어 있다는 점, 활용할 수 있는 원시데이터가 부족하다는 점, 데이터의 표준화 부재 등의 문제로 산업계에서 실질적으로 활용성이 높은 양질의 데이터가 부족하여, 사업이용자가 플랫폼 데이터에 접근하는데 커다란 장애 요인으로 작용하면서 데이터 유통 시장 활성화가 어려운 상황이라고 할 수 있음.

#### (4) 기타 이슈 및 과제

- 사업이용자의 플랫폼 데이터 접근과 관련된 기타 이슈 및 과제는, 개인정보 활용에 대한 데이터 처리 방법 및 영역 규정의 명확화, 개인정보 데이터 분석 및 처리 기술의 개발과 보완, 공정한 데이터 시장의 확립과 관련된 문제임.
  - 먼저, 개인정보 활용 개선과 관련하여, 데이터 3법 개정안 시행으로 개인정보 활용에 대한 전반적 법적 개선과 보완이 이루어졌고, 가명정보가 도입되면서 이를 적법하게 활용할 수 있게된 점은 고무적이지만, 동시에 제도의 취지나 본질을 훼손하거나 왜곡하지 않도록 철저하고 지속적인 관리 및 감독이 이루어져야 함.
    - 특히 가명정보 처리 시 암호화, 범주화, 삭제해야 할 속성 정보를 더욱 명확하게 규정하고, 이에 대한 세밀한 가이드라인이 제시되어야 함.
  - 또, 철저한 제도 보완과 실천 전략 수립을 통해 개인정보의 안전하고 적법한 처리 방식인 가명정보에 대한 국민적 이해도를 얻을 수 있도록 가명정보 처리의 장점과 효과에 대한 홍보 및 교육으로 정보화 및 기술 융복합 시대의 개인정보의 의미와 가치, 중요성 등에 대한 사회적 인식을 확산시킬 필요가 있음.
  - 이에 더해, 데이터를 제공하는 주체인 소비자의 합리적이고 안전한 데이터 거래를 보장하고, 소비자들에게 데이터 제공에 따른 대가 지급 등의 확실한 보상과 수익 환원 등을 실천해야 함.
- 가명정보 처리는 정보 보안과 안전성에 대한 확실한 기술 처리를 전제로 하는 만큼, 데이터 처리 기술의 성장과 개발 및 보완이 필수적인 요소라고 할 수 있음.
  - 최근 선진국을 중심으로 데이터 분석 및 처리 기술에 인공지능 기술과 클라우드 컴퓨팅 기술, 유비쿼터스 기술 등을 접목하면서 기술 상승의 강력한 시너지 효과가 나타나고 있음.
  - 이러한 주요 기술 중 특히 인공지능 기술은 금융 및 데이터 산업계 뿐만 아니라, 제조 및 의료 산업 내에서도 커다란 파급력을 나타내고 있음.
    - 고부가가치 산업의 발전 시대에서 한국의 인공지능 기술 경쟁력은 미국의 동기술 경쟁력의 80% 수준이며, 대략 1.8년의 기술 격차가 나타남(정보통신기획평가원, 2020).

- 이러한 기술 격차를 극복하고 데이터 처리 관련 첨단 기술을 지속적으로 발전시키기 위해 국내 산업계 및 학계의 지속적인 노력 및 투자와 그에 상응하는 제도적, 정책적 지원 등이 수반되어야 함.
- 이와 더불어, 가명정보 처리 방법 및 활용 영역에 대한 명확한 지침이 수립되어야 함.
- 데이터 3법 개정으로 인해 가명정보 개념이 도입되었지만, 데이터의 활용 범위에 관한 규정이 명확하지 않고 모호한 부분도 존재함.
    - 법률상 명시적으로 나타내고 있는 기준과 경계가 불분명하므로 세부 해석과 판단에 따라 적용 범위가 달라질 수 있는 가능성이 존재함.
    - 가명정보는 언제든지 추가 정보가 대입되어 개인 식별 정보로 전환될 수 있는 위험성이 존재하므로, 그 활용 과정에서 불법적 또는 의도적으로 남용되거나 오용되지 않도록 엄격한 조치와 합법적 가이드라인을 수립 및 적용시켜야 함.
    - 또한, 공공 및 민간 분야의 개인정보 보호활동을 수행하는 개인정보 처리자에 대한 보호 조치 점검을 위한 기술 및 교육 지원 등 실질적 인식 제고를 위한 방안을 마련할 필요가 있음.
- 데이터 관련 산업의 발전과 활성화를 유도하는데 있어 가장 큰 장애물로 예상되는 문제점은 데이터 독과점과 그로 인한 개인정보 주체의 대상화임.
- 이미 데이터를 대량으로 보유하고 있는 기업과 그렇지 못한 기업 사이에는 현실적으로 데이터 관련 산업에서의 위상 및 발전, 성장속도 등에서 명백한 차이가 있을 것임.
    - 이와 같은 현상이 심화된다면 시장이 특정 소수의 기업에 의해 좌우되고 건전한 경쟁 체제의 발전을 저해하게 되고, 개인 소비자들 역시 해당 기업의 영향력에 종속될 위험이 있음.
    - 이와 반대로, 지나치게 영세한 기업이 데이터 시장에 진입한다면 관리체계의 부실 등으로 개인정보 유출 및 불법 사용 등의 사례도 증가할 수도 있음.
    - 제반의 문제점들을 해결하기 위한 전제로 국가 소유의 막대한 공공 데이터와 현재 금융 회사 등이 보유하고 있는 신용 및 금융 정보들을 포함하여 추후 다른 산업군의 데이터에 이르기까지 정책 통합적 관점에서 기업의 자유로운

참여의 기회 보장과 규제 체계 마련 등 데이터 연관 산업 내 경쟁 원칙 확립과 분위기 조성이 이루어져야 함.

- 또, 영세한 기업에는 최소한의 데이터 관리 체계 수립을 위한 행정적 지원을 비롯해 장기적으로 데이터 시장이 건전한 자유 경쟁 체제를 유지하기 위해 데이터 접근의 공정성이 확보될 수 있도록 정책 거버넌스를 통해 효율적으로 관리되어야 함.

## 라. 결론

- 4차 산업혁명에서 핵심적인 원천이라고 할 수 있는 데이터를 활용하여 기업은 생산량 증가 및 노동시간 단축 등의 혁신을 통해 기업의 경쟁력을 향상할 수 있음.
  - 우리 정부는 데이터 3법(개인정보 보호법, 정보통신망법, 신용정보법) 개정을 통해 국내의 데이터 관련 법과 제도 체계를 4차 산업혁명 시대의 요구 및 트렌드에 맞추며 전면적, 혁신적으로 개선하기 위한 노력을 하였음.
  - 이외에도 데이터 인공지능경제 활성화를 위한 등 데이터 관련 정책 및 사업을 본격적으로 추진하면서 여러 산업 내에서의 데이터의 활용 확산을 도모하고 있음.
- 데이터 유통 및 거래 활성화를 위하여 사업이용자들이 기존보다 플랫폼 데이터에 접근을 용이하게 만들어 유인하기 위해 해결해야 할 선결 과제들을 발굴하여 이에 적합한 정책과 사업을 지속적으로 추진해야 함.
  - 데이터 유통 및 거래 시장은 초기 단계에 놓여 있으며, 많은 정부 예산이 투입되고 있으나, 아직 데이터 수요자인 사업이용자의 플랫폼 데이터에 대한 접근에는 여러 가지 제약이 따르고 있기 때문임.
- 또, 데이터의 특성을 고려한 데이터 권리관계 정립이 무엇보다 우선되어야 함.
  - 데이터는 집합재적 특성 및 연결성에 의해 가치가 증가하며 정보 주체와 수집자 사이에 다대일의 거래가 이루어질 수 있기 때문임.
  - 이에 더해, 데이터 재화의 무상 거래 관행 및 공공재적 성격 때문에 지식재산권 보호가 없을 경우 저생산의 문제에 빠질 가능성도 있음.

- 이외의 제언으로는, 첫째, 데이터의 소재 파악 및 접근을 위한 공통 플랫폼의 활용에 대한 방안이 마련되어야 함.
  - 데이터 활용을 위해 기초 데이터의 확보가 중요시되지만 필요한 데이터의 유형, 종류, 존재 여부 및 소재 등 데이터를 수집하기 위해 필수적인 정보들을 얻어내는 데 사업이용자들은 많은 어려움을 겪고 있음.
    - 분야별로 많은 데이터 플랫폼들이 만들어지고 있긴 하지만 수집된 데이터의 실질적인 산업적 활용은 어려운 편임.
    - 데이터를 거래하는 데 있어 지원기능 역시 미흡하다고 볼 수 있음.
    - 따라서 데이터 수집 단계에서 산업 내에서의 실질적인 데이터 수요를 기반으로 한 데이터 수집과 산업별로 특정된 데이터 플랫폼을 구축하고 플랫폼 간 연계를 통해 거래 지원 기능 강화를 전제로 한 데이터 유통의 탄탄한 기반 조성이 필요함.
  
- 둘째로, 데이터를 거래하고 유통하는데 명확한 기준이 될 수 있는 가격 방법론의 수립이 필요함.
  - 기존의 데이터 가치 평가 방법론은 원가접근법, 시장접근법, 소득접근법 등으로, 데이터의 가치 산정은 가능하나, 이 중 한 가지의 방법론을 채택해서 실제 데이터 거래에 적용하기에는 산업별 필요 데이터의 유형, 활용의 실질적 목적 등을 최대한 반영하기에는 제약이 있음.
    - 따라서 데이터의 적정 가격 책정을 위해 방대한 양의 데이터 중 수요가 높거나 활용 가능성이 비교적 높은 데이터를 선별하여 데이터 유형, 활용범위 등을 정의해 그에 적합한 데이터 가격 책정 방식을 실증적으로 적용해볼 필요가 있음.
    - 데이터의 가격 평가 방식과 체계를 마련하고 정착시키기 위해서는 평가 지표, 요인, 평가 결과의 적용 등 다양한 실증적 산출물을 쌓아 중장기적 관점의 데이터 가치 산정 체계를 확립시키는 것이 중요함.

- 셋째로, 데이터의 품질관리 체계 구축이 필요함.
  - 데이터 거래를 위해 데이터의 품질 요건이 중요한 요소로 작용하며, 이를 위해 더욱 많은 사업 이용자의 데이터 접근성을 늘리기 위한 데이터의 품질 관리가 중요함.
  - 정확성이 떨어지는 데이터는 신뢰도뿐만 아니라 데이터를 활용하여 산출한 결과의 왜곡을 낫게 되어 데이터를 재사용하지 못하는 경우가 발생할 수도 있음.
    - 이를 방지하기 위해, 데이터 생성 단계부터 철저한 품질관리를 할 수 있는 방안을 마련해야 할 것이며, 품질에는 데이터 표준화와 관련된 문제도 포함되기 때문에, 데이터 생성 단계부터 관리 계획을 수립해 데이터 표준 및 품질 요인을 모두 만족하는 저장, 보존, 공유를 원활하게 해낼 수 있는 방안 수립이 필요함.
    - 또, 데이터 호환을 위한 표준 정립을 마련함으로써 사업이용자들의 데이터 사용성과 활용성을 높이기 위한 데이터 포맷을 정립한 후 데이터 가치사슬 모든 과정에 해당하는 데이터 표준체계 마련이 필요함.

## 4. 데이터 독점과 부당 이용의 이슈와 과제

### 가. 데이터 독점 개념의 복합적 성격

- 홍대식(2020)에 따르면, 데이터 독점(data monopoly)이라는 용어는 주요 온라인 플랫폼 사업자들의 경쟁 위협을 강조하기 위하여 언론 매체에서 사용하기 시작한 것으로 보이나, 이런 표현은 시장의 관점에서 독점력을 논의하는 것은 아니므로 수사학적인 표현에 불과함
  - 구글, 아마존, 페이스북과 같이 특정 유형의 이용자 데이터를 많이 확보하고 이를 이용한 성공적인 사업모델을 갖고 있는 주요한 온라인 플랫폼 사업자들의 경쟁 위협을 지칭하는 하나의 표현일 뿐 엄밀한 개념은 아니라는 의견
  - 또한, 데이터 독점을 지식재산법적 의미에서의 독점에 가까운 것으로 본다면, 기술혁신에 대한 수익을 배타적으로 보호·확보할 수 있는 전유가능성(appropriability) 내지 배제가능성 등 지식재산권에 유사한 권리를 부여하게 될 수도 있음

- 한편, 이효석(2022)은 홍대식(2020)의 데이터 독점 개념의 모호성에 대한 비판적 지적을 일부 수용하면서, (1) 데이터를 배타적으로 소유·사용해 수익을 얻는다는 의미와 (2) 데이터를 바탕으로 시장지배력을 획득하거나 기존에 보유한 시장지배력을 유지·강화한다는 의미가 혼재된 것으로 정리하여 설명
  - 일반적 용례로서 독점은 타인과 공유하지 않고 혼자 전유한다는 의미를 가짐
  - 따라서, 데이터 독점은 경쟁자나 다른 사업자와 데이터를 공유하지 않고 타인의 접근이나 이용을 배제한다는 의미를 가질 수 있음
- 요컨대, 데이터 독점은 데이터를 독점의 대상으로 보는 관점과 독점적 지배력 확보의 수단으로 보는 관점의 두가지 다른 방식으로 해석될 수 있음
  - 현실에서는 아마도 전자의 상황과 후자의 상황이 연계되는 측면도 존재할 것으로 생각됨
- 본 절의 논의는 이러한 개념 정리에 근거하여 진행하고자 함
- 데이터 독점의 첫 번째 개념인 데이터의 배타적 소유·사용에 따른 문제는 곧 데이터 접근의 이슈로 연결되는데, 이에 대해서는 본 장의 나, 다에서 구체적으로 논의할 것임
- 데이터 독점의 두 번째 개념인 데이터에 기반한 시장지배력 획득·유지·강화의 문제는 경쟁제한성 관점에서의 논의로 연결되는데, 이에 대해 다음 소절(나)에서 논의할 것임
  - 이 논의는 경쟁제한성 관점에서의 데이터 접근의 이슈도 일부 포함하게 될 것임

## 나. 경쟁제한성 관점에서의 데이터 독점의 이슈들

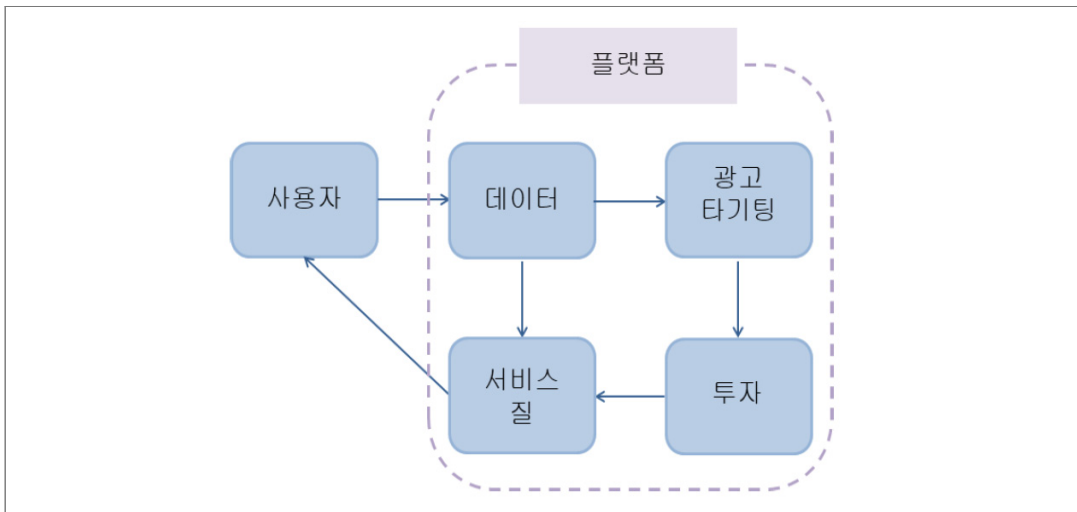
### (1) 경쟁제한성 이슈들

- 이화령·김민정(2017), 이호영(2018), 최난설현(2020) 등은 빅데이터와 관련한 경쟁제한 우려로 진입장벽의 형성, 배제적 행위를 통한 경쟁 제한, 프라이버시 보호 경쟁의 제한 등의 이슈들을 제기한 바 있음
  - 아래에서는 이러한 이슈들에 대한 기존 논의의 핵심을 정리해보고자 함

□ 진입장벽의 형성

- 최근 데이터는 필수적 사업자산으로 인식되므로 데이터 확보 경쟁은 점점 치열해지고 있으며, 이 경쟁에서 소규모 사업자나 신규 사업자는 우위를 점하기 어려움(최난설현, 2020)
  - 신규 사업자가 충분한 데이터를 수집할 수 없거나, 크기와 다양성 측면에서 기존 사업자가 사용하는 수준의 데이터에 접근하기 어렵다면, 이것이 진입장벽으로 작용할 수 있음(이호영, 2018)
  - 플랫폼 상의 다른 공급자들에게는 제공하지 않는 데이터를 이용해서 플랫폼이 시장 우위를 얻을 수 있음(이화령·김민정, 2017)
- 데이터를 매개로 하는 피드백 순환 구조에 따른 네트워크 효과(이화령·김민정, 2017)가 진입장벽으로 작용할 수 있음
  - 검색엔진이나 사회관계망 서비스와 같이 데이터의 수집 및 활용이 중요한 시장에서는 데이터의 규모가 클수록 이를 이용한 시행착오나 실행에 의한 학습에 따른 네트워크 효과가 존재하여 진입장벽으로 작용(이호영, 2018)

[그림 III-3] 플랫폼 경제에서의 데이터 피드백 순환과 네트워크 효과



출처: 이화령·김민정(2017)

- 데이터와 관련한 비용 구조 측면에서의 진입장벽도 존재할 수 있음
  - 대용량 데이터를 수집, 저장, 관리하고 (새로운 데이터 수집을 위한) 무료 서비스를 제공함에 있어 상당히 많은 비용이 소요되는 문제가 있음(최난설현, 2020)
- 그러나, 아직까지 경쟁당국이나 사법부가 직접적으로 이용자 데이터가 진입장벽을 구성한다고 본 사례는 나타나지 않았음(최난설현, 2020)

#### □ 데이터 관련 배제적 행위를 통한 경쟁 제한

- 대량의 데이터를 보유한 사업자를 인수·합병하거나 데이터에 대한 접근 거절, 차별적 접근 허용, 배타적 계약 등의 행위를 통하여 관련 시장에서의 경쟁을 제한하거나 경쟁자를 배제할 우려가 있음(이호영, 2018)
  - 기업결합이 데이터에 대한 차별적 접근을 초래하고 데이터의 집중을 증가시킬 수 있음
  - 특정한 데이터가 관련 시장에서 사업을 영위하기 위해 반드시 필요한 상황에서 이를 보유한 사업자가 경쟁자의 데이터 접근을 거절할 경우
  - 인접시장에서 경쟁관계에 있는 사업자에게 데이터에 대한 접근을 차별적으로 허용할 경우
  - 제3자와의 배타적 거래 계약을 체결하여 경쟁자의 중요 데이터 접근을 방해할 경우
  - 특정 시장에서 수집한 데이터를 다른 시장에서 사용함으로써 경쟁상 우위를 얻는 경우 등
- 데이터를 통하여 경쟁 우위를 추구하는 사업자는 경쟁사업자가 데이터에 접근하는 것을 차단하거나 유사한 데이터 획득의 기회를 봉쇄하는 반경쟁적 행위를 할 가능성이 큼(최난설현, 2020)
  - 데이터와 자사의 관련 서비스를 끼워팔기하는 경우에도 경쟁제한적 효과가 나타날 수 있음
  - 예를 들어, 기업이 자사 데이터에 데이터 분석 서비스를 결합시켜 판매하는 경우가 가능함(이화령·김민정, 2017)

□ 프라이버시 보호 경쟁의 제한

- 프라이버시 보호가 비가격 품질경쟁의 지표들 중 하나에 해당한다면, 시장 지배력의 존재에 따라 프라이버시 보호 경쟁이 제한될 수 있음
  - 시장지배력을 가진 플랫폼은 정보 비대칭성을 이용해 이윤을 증가시키는 과정에서 낮은 프라이버시 수준을 제공하는 것이 일반적 유인이 됨(이화령·김민정, 2017)
  - 독과점 시장에서는 경쟁시장과 비교해 개인 데이터가 과도하게 수집되거나 프라이버시 옵션들이 불충분하게 제공될 수 있으며, 기업결합으로 인해 데이터가 집중됨으로써 프라이버시 훼손으로 이어질 가능성도 고려할 필요(최난설현, 2020)
  - 페이스북에 대한 독일 연방카르텔청의 법집행 사례는 시장지배적 지위를 이용하여 소비자의 프라이버시를 침해한 행위로 볼 수 있음(이호영, 2018)
- 이러한 문제는 다음 소절(다)에서 살펴볼 데이터 부당이용의 관점에서 논의해볼 필요도 있음

(2) 경쟁제한성 이슈 제기에 대한 반론

- 플랫폼의 데이터 활용의 경쟁제한성 주장에 대해서는 다양한 반론들도 존재하는데, Lambrecht and Tucker(2015)의 논의를 중심으로 정리해보면 다음과 같음
- 데이터는 (경쟁자들의) 모방, 복제가 불가능하지 않음
  - 데이터는 기본적으로 비경합적(non-rivalrous)이고, 생산 및 유통의 한계비용이 대단히 낮음
  - 경쟁자들은 데이터 거래 시장을 통해 필요한 데이터를 확보할 수도 있음
- 데이터는 희소한 생산요소가 아님
  - 클라우드 기반 설비 등의 공급이 늘어나면서 데이터를 수집할 수 있는 수단들이 보편화되고 있음
  - 소비자들이 남기는 자신들의 수요와 선호에 대한 정보는 온라인 공간 어디에서나 발견할 수 있음

- 데이터는 그 자체로 큰 가치를 갖는 것은 아님
  - 호환성과 통합성을 확보하지 못하면 가치가 낮아짐
  - 그 자체로 구조화되어있지 않은 경우가 많음
  - 데이터로부터 인과관계를 도출하는 데는 어려움이 있어서, 실험적 테스트를 할 수 있을 때 데이터가 높은 가치를 갖게 되는데 이를 위해 많은 데이터가 반드시 필요한 것은 아님
- 데이터는 대체불가능하지 않음
  - 소비자를 파악하는 자체보다 결국 소비자에게 높은 가치를 제안하는 것이 더 중요함
  - 소셜미디어 시장의 동태적 경쟁을 보면, 많은 데이터를 확보한 선발 플랫폼들이 시장지배적 지위를 계속 유지하지 못한 사례들이 많음
  - 신규 스타트업 플랫폼들은 기존 기업들에 비해 데이터를 확보하지 못하고 있음에도 불구하고 투자 자금을 성공적으로 유치한 사례들이 많음
- Bourreau et al.(2017) 역시 데이터는 여러 생산요소들 중의 하나로서 중요하지만 특별하지는 않음을 지적
  - 숙련되고 혁신적인 노동력이나 자본, 유통채널 등 다른 생산요소들도 마찬가지로 중요함
- 요약하자면, 경쟁제한성에 대한 우려는 대부분 데이터가 제품·서비스 생산의 필수요소임을 전제로 하고 있으나, 데이터는 품질을 높이고 비용을 낮추는 데 사용되는 여러 요소들 중의 하나일뿐이라는 반론이 가능
  - 물론, 업종에 따라서는 데이터가 대단히 중요한 경우도 있을 것임
  - 원론적으로 생산요소들 간의 대체성을 무시해서는 안됨
  - 데이터 활용에 투자하는 기업도 하나의 전략적 선택을 하고 있을 뿐이며 그에 따른 많은 비용과 리스크가 존재하는데, 이러한 전략이 항상 비용 대비 효과적이라는 근거는 충분하지 않음
  - 데이터 자체보다는 데이터의 활용 능력과 인사이트가 더 중요할 수도 있음

## 다. 데이터 부당 이용의 문제

### (1) 데이터 독점과 데이터 부당 이용 간의 관계

- 데이터의 부당 이용 문제는 반드시 데이터 독점을 전제로 하는 것은 아님
  - 즉, 데이터 독점은 데이터 부당 이용의 필요조건이 아님
  - 데이터를 독점하거나 데이터에 기반하여 시장 지배력을 확보한 기업이 아니라 하더라도 데이터를 부당한 방식으로 이용할 수 있음
- 그러나, 데이터 독점은 데이터 이용 부당성의 한 구성요소가 되거나 부당 이용 문제를 심화하는 요인이 될 수 있음
- 첫째, 데이터를 배타적으로 소유하여 독점한 기업이 이를 자사의 이익을 위해 적극적으로 이용할 경우 부당성 문제가 제기될 수 있음
  - 경쟁자나 거래상대방, 소비자 등이 접근하기 어려운 데이터를 배타적으로 활용하였다는 점과 관련된 불공정성의 문제
  - 필수설비 원칙이 충족되지 않는다면 특정 생산요소의 배타적 사용을 불공정하다고 보기는 어렵겠으나, 아래에서 살펴볼 데이터의 공공재적 특성과 잠재적 차별성의 효과를 추가적으로 고려할 필요가 있음
- 둘째, 데이터가 시장지배력을 확보, 유지하는 수단으로 적극적으로 활용되었다면 그러한 목적·효과를 갖는 데이터 이용이 부당한 것이 아니냐는 문제가 제기될 수 있음
  - 이는 앞서 살펴본 데이터 독점의 경쟁제한성 문제의 한 측면에 해당
- 셋째, 높은 시장점유율을 가진 시장지배적 지위의 기업이 데이터를 부당하게 이용할 경우에, 그 효과가 더 심화될 수 있음
  - 그 부정적 효과의 규모가 클 뿐 아니라 거래상대방이나 소비자가 이를 회피하기 어려움
- 한편, 데이터가 일부 기업들에 집중될 때 부당 이용의 문제는 오히려 완화될 수도 있음에 유의할 필요
  - 개인정보 등 민감한 정보를 포함한 데이터가 다수의 기업들에 의해 무차별적으로 공유될 경우 부당 이용의 문제는 오히려 심화될 수 있음

- 즉, 데이터를 효율적이고 안전하게 활용할 수 있는 능력을 갖춘 기업들에 데이터가 집중되어 관리되는 것이 부당 이용의 문제를 완화할 수 있음

## (2) 데이터의 공공재적 특성과 부당 이용의 문제

- 데이터 이용의 주체가 기업일 때, 데이터는 경제학적으로 생산요소들(inputs) 중의 하나로 개념화할 수 있을 것임
- 그런데, 노동시장과 같은 예외적인 경우나 일부 경쟁제한성 효과가 있는 경우를 제외하면, 일반적으로 기업의 생산요소 사용에 있어 부당성 이슈가 제기되기는 어려움
  - 생산요소 시장을 통한 구매 또는 (수직통합적 구조에서의) 자가공급(self supply)의 과정에서 부당 이용의 이슈는 제기되기 어려움
  - 핵심적 생산요소의 판매거절이나 구매거절 등이 문제될 수 있으나, 이는 경쟁 제한성 효과에 한정된 것임
- 다른 일반적 생산요소들과 달리 데이터의 경우 특히 부당 이용의 이슈가 제기되는 한 가지 원인은 아래에서 설명하듯이 데이터의 공공재적 특성에서 찾을 수 있을 것임
  - 생산요소 공급의 주체인 인간 개인과 제공된 생산요소가 명확히 분리되지 않는다는 점에서 노동시장과 유사하며, 그로 인해 (노동시장에서와 마찬가지로) 부당 이용의 문제, 특히 잠재적 차별성의 효과가 발생한다고 볼 수 있다는 점에도 주목할 필요가 있음
  - 아래에서는 일단 공공재적 특성에 초점을 두고 논의를 전개하고자 함

### □ 데이터의 공공재적 특성

- 데이터는 비배제성(non-excludability)과 비경합성(non-rivalry)을 갖는다는 점에서 경제학적 개념상 공공재(public good)로 볼 수 있음
  - 단, 상당한 정도의 수집·선별·가공의 단계를 거친 데이터는 배제성(excludability)이 부여되어 배타적 소유권 성립이 가능해질 수 있으며, 이 경우 경제학적 용어로 클럽재(club good)의 성격을 갖게 됨

- 공공재의 경우 정부가 공공재의 공급을 담당하는 것이 원론적인 해결책인 것으로 알려져 있으나, 민간 영역의 데이터를 정부가 대신 수집, 관리, 공급하는 것은 바람직하지 않고 현실적이지도 않음
  - 조세 등 필수 공공적 목적 외의 민간 데이터를 정부가 수집, 관리하는 것은 헌법 18조(“모든 국민은 사생활의 비밀과 자유를 침해받지 아니한다”)에 반함
  - 단, 민간 데이터가 아닌 공공 영역에서의 데이터 수집, 관리, 가공, 공급의 기능은, 데이터 경제 시대의 수요에 대응하여, 정책적으로 계속하여 확대·강화될 필요가 있음
- 다른 공공재와 달리, 데이터는 과소공급의 문제가 현실적으로 심각하지 않다는 특수성이 있음
  - 원시 데이터의 경우 개인 및 집단의 경제·사회 활동의 부산물로서 끊임없이 창출되어 잠재적 수요자에게 공급되고 있음
  - 물론, 데이터의 구체적 유형에 따라 사회적 적정 수준보다 과소공급되는 경우도 있을 것임

□ 무임승차적 전유와 데이터 부당 이용의 문제

- 과소공급의 문제가 크지 않은 대신, 공공재의 더 근원적 문제라고 할 수 있는 무임승차(free riding)의 유인은 여전히 유효하므로 이에 주목할 필요가 있음
- 데이터 사용자는 무료 또는 저가로 원시 데이터를 획득하여 이를 가공한 후 배제성(소유권)을 부여하는 방식으로 전유(appropriation)할 수 있음
  - 물론 데이터의 수집과 가공의 과정에서 상당한 혁신적 방법과 비용의 투자가 요구된다면, 그 결과물로서의 데이터에 대한 배타적 권리는 어느 정도 정당화될 수 있을 것이며 무임승차로 단정하기도 어려울 것임
- 데이터 부당 이용의 문제는, 이러한 무임승차적 전유라는 법경제학적 개념의 틀에서 접근해볼 수 있음
  - 데이터 이용의 유인과 과정은 필연적으로 데이터의 공공재적 특성에 따른 무임승차적 전유의 성격을 가질 수밖에 없음
  - 무임승차적 전유 행위는 효율성과 공정성 모두를 저해하므로, 데이터 부당 이용의 문제도 이에 근거하여 규정해볼 수 있음

- 효율성의 관점에서 데이터가 과다 생산 및 소비되는 문제가 발생할 수 있음
  - 무료 또는 저가로 데이터를 획득하는 반면, 데이터 활용의 과정에서 개인정보 침해나 유출로 발생할 수 있는 피해 위험 등의 사회적 비용은 충분히 고려하지 않을 수 있음
  - 개인정보가 충분한 대가없이 활용된다는 사실 자체로 인해 소비자들이 경험하는 불만이나 무력감도 넓은 범위의 사회적 비용으로 볼 수 있는데(김성환, 2021), 이는 법경제학의 공용수용 이론 논의에서 Michelman(1967)이 제안한 사기저하 비용(demoralization cost) 개념과 유사한 것으로 이해할 수 있음
- 공정성의 관점에서는 무임승차적 전유의 과정에서 데이터 생성에 관계된 주체들의 권리가 무시되거나 침해되는 측면이 있음
- 그러나, 다른 한편으로 데이터 이용이 창출하는 여러 측면의 가치들을 고려할 때, 기업의 데이터 이용 행위 일반을 부당하다고 볼 수는 없을 것임
  - 무임승차적 전유로 인해 초래되는 효율성과 공정성의 저해가 현저히 큰 경우에만 부당한 이용이라고 보아야 할 것임

#### □ 데이터 소유권의 문제

- 데이터 이용의 근원적 문제인 무임승차적 전유를 해소하는 한가지 방법은 배제성(excludability), 즉 소유권을 기술적·제도적으로 도입하여 거래시장을 활성화하는 것임
  - 케이블 및 IP 기술을 이용해 기존 지상파 방송의 비배제성을 극복한 유료방송이 그 좋은 예임
  - DRM(Digital Rights Management) 기술에 기반한 디지털 콘텐츠 거래시장도 마찬가지임
- 그러나, 데이터 소유권은 현재 법적으로 정립되어 있지 않으며 앞으로도 이런 상황은 (특별한 기술적 환경의 변화가 없다면) 계속될 것으로 보임
- 권영준(2021)은 데이터 소유권이 정립되지 않은 국내외의 현황과 향후에도 간단히 법제화되기 어려운 여러 이유들을 자세히 설명하였음
  - 부정형성(비특정성), 귀속의 어려움(창출·가공의 과정에 복수의 주체들이 관여), 연결·결합의 부가가치 등 데이터의 속성들이 소유의 관념과 잘 부합하지 않음

- 데이터는 여러 층위의 다양한 측면에서 파악할 수 있는 다중다색의 개념으로서, 소유권의 대상이 될 수 있을 만큼 뚜렷하게 특정하거나 특정이 가능하도록 정의하는 데에 한계가 있음
- 물론, 저작권법이나 개인정보보호법 등 다른 개별법에 의해 명확히 보호받는 권리들이 존재하기는 하나, 그 적용 범위는 상당히 제한적이고, 데이터 소유권 개념으로 일반화하기 어려움
- 특히, 개인정보보호법은 개인정보 자기결정권이라는 개념에 근거한 것으로, 데이터의 재산적 측면이 아닌 인격적 측면을 규율(권영준, 2021)
- 이처럼 데이터 소유권이 정립되지 않은 제도적 환경에서 거래시장이 충분히 활성화되기 어렵고 거래시장이 일정 수준 이상 활성화되지 않으면 투명한 가격설정(pricing)이 가능할 것인가에 대한 의문이 제기될 수도 있음
- 이는 데이터 부당이용의 이슈들이 계속되는 제기되는 요인들 중의 하나라고 볼 수 있음

## 라. 데이터 부당 이용의 예시적 유형 및 사례

### (1) 데이터 기반 가격차별

- 기업의 가격차별 행위는 그 자체로 부당한 것은 아니며 현실의 기업들은 수많은 다양한 유형의 가격차별 전략을 개발하여 활용하고 있으므로, 데이터 기반 가격차별 역시 그 자체로 데이터의 부당한 이용으로 규정할 수는 없음
- 다만, 가격차별은 오늘날의 많은 기업들이 지향하는 데이터 기반 소비자 맞춤형 전략의 대표적인 예라고 할 수 있으므로, 그러한 전략이 데이터 부당 이용의 성격이나 효과를 갖게 될 가능성에 대해 검토해 볼 필요가 있음
- 데이터 기반 가격차별의 실제 사례는 해외에서는 오래전부터 다수 발생하고 있으며 우리나라에서도 마켓컬리가 동일한 상품의 가격을 소비자에 따라 다르게 제시한 행위가 언론에 보도된 바가 있음<sup>79)</sup>

79) 데이터 기반 가격차별의 사례들에 대한 보다 자세한 설명은 김성환(2021)을 참고하라.

## □ 데이터 기반 가격차별의 경제학적 특성

- 소비자들에 대한 정보가 부족한 상황, 즉 데이터 활용 이전의 상황에서 기업들의 가격차별 전략은 2차 가격차별 방식에 주로 의존함
  - 2차 가격차별은 소비자의 자기선택(self selection) 유도를 통해 기업이 소비자 잉여의 일부를 이윤화하는 전략임
  - 소비자가 누리는 잉여, 즉 정보지대(information rent)는 2차 가격차별에 의해 일부 감소하지만, 기본적인 정보 비대칭성의 구조는 유지되기 때문에 소비자 잉여는 일정 부분 확보될 수 있음
- 기업의 데이터 활용은 우선 2차 가격차별을 3차 가격차별로 전환하는 효과가 있음
  - 3차 가격차별은 소비자 정보를 이용하여 소비자들을 몇 개의 그룹으로 나누어 다음 각 그룹에 최적화된 가격설정을 통해 차별하는 방식
  - 데이터 활용을 통해 기업이 더 많은 소비자 정보를 획득하면 이를 이용해 3차 가격차별 전략의 효과성을 높일 수 있으므로, 3차 가격차별의 정교화에 따라 소비자 잉여(정보지대)는 감소하게 됨
- 기업의 데이터 활용은 (기업의 관점에서) 이상적, 장기적으로는 완전히 개인화된 가격설정, 즉 1차 가격차별로 수렴한다고 볼 수 있음
  - 완전한 가격차별에서는 소비자 잉여(정보지대)는 완전히 사라지게 됨

## □ 데이터 기반 가격차별의 부당성 이슈

- 위에서 살펴본 데이터 기반 가격차별의 경제학적 특성에 기초할 때, 근본적으로 소비자 잉여(정보지대)의 감소 효과를 초래한다는 문제가 존재
  - 자유 시장경제에서 소비자 잉여의 감소 자체만으로 부당성을 주장하기는 어렵겠으나, 데이터 활용의 증가에 따라 소비자 잉여의 체계적이고 지속적인 감소가 발생한다면 문제가 될 수 있음
- 또한, 시장을 각자가 확보한 정보에 기반한 협상 거래의 장으로 이해한다고 할 때, 기업의 소비자에 대한 정보 수집, 축적에 따라 개별 소비자의 협상력이 거의 사라지게 된다면 이를 더 이상 자유롭고 공정한 시장거래 환경으로 인정하기 어려움

- 데이터에 기반한 3차 가격차별이 갖는 일방적 할당에 따른 차별이 갖는 부당성 문제도 지적할 수 있음
  - 데이터를 활용하지 않는 2차 가격차별은 소비자 자신의 선택에 의해 결과적 차별이 발생하므로 차별의 부당성 문제가 심각하지 않음 (물론 2차 가격차별이 제시한 요금제의 구조가 실질적으로 부당한 차별을 나타낼 가능성은 존재함)
  - 3차 가격차별은 차별의 방식이 일방적일 뿐 아니라 그룹 세분화가 충분히 정교화되지 않을 경우 할당오류(misallocation)에 따른 비효율성과 불공정성이 동시에 발생할 수 있음

□ 차별 비용의 외부성에 따른 비효율성 문제

- 데이터 기반 가격차별의 과정에서 기업이 사적 비용으로 고려하지 않는 외부효과로서의 사회적 비용들이 존재함
- 예를 들어, 기업이 소비자 데이터 활용을 확대함에 따라 소비자의 개인정보가 침해됨으로써 다양한 피해가 발생할 가능성이 증가하게 됨
  - 기업의 데이터 활용이 반드시 개인정보 침해 발생을 의미하는 것은 아니므로, 개인정보 관련 피해에 따른 부당성을 사전적으로 규정할 수는 없음
  - 그러나, 그러한 피해의 위험을 소비자가 부담해야 한다는 사실이 곧 사회적 비용을 의미함

〈표 III-3〉 차별비용과 가격차별 효율성의 관계

사례	비차별 독점의 사중손실/소비자잉여와 차별 비용의 비교		차별 실행	차별 효율성
	〈-----사중손실(A)-----〉	〈-----소비자잉여(B)-----〉		
①	외부적 비용		○	○
②	외부적 비용		○	X
③	사적 비용		X	-
④	사적 비용		○	X
⑤	사적 비용	외-	X	-
⑥	사적 비용	외부적 비용	○	X
⑦	사적 비용	외부적 비용	○	X
⑧	사적 비용	외부적 비용	○	X
⑨	사-	외-	○	○

출처: 김성환(2021)

- 기업은 이러한 사회적 비용의 존재를 무시한 상태에서 가격차별 전략을 선택하게 되므로, <표 III-3>이 정리하여 보여주듯이 총후생의 관점에서 비효율성의 결과를 초래할 수 있음
  - 표는 분석의 편의상, 1차 가격차별을 가정하고 있음

## (2) 아마존의 판매자 데이터 활용

### □ Khan(2017)의 비판

- Khan(2017)은 미국 아마존 마켓플레이스의 부당한 데이터 활용의 사례를 다음과 같이 소개한 바 있음
- 아마존 판매에서 제3자 판매, 즉 중개거래의 비중은 2015년 기준 약 50%에 이룸
- 아마존은 판매자들의 다양한 상품 판매실적 데이터를 분석하여 이를 토대로 자체 PB 상품인 AmazonBasics 상품들을 지속적으로 도입, 확대해나가는 전략을 사용한 것으로 추정됨
  - 아마존이 동일한 상품을 판매하게 되면 도매협상력이나 검색결과 노출에서 확연한 우위를 갖게 됨
  - 아마존은 이러한 판매자 데이터 활용을 통해 상품의 시장도입 초기 비용이나 불확실성을 회피할 수 있음
- 오프라인 대형 유통업체들도 마찬가지로 전략을 쓸 수 있겠으나, 아마존은 그 규모와 활용하는 데이터의 구체성에서 큰 차이가 있음
- 아마존은 클라우드 컴퓨팅 서비스 사용량 데이터를 이용해 스타트업 기업들에 대한 투자 의사결정을 하는 것으로 알려져 있는데, 이 역시 비슷한 유형의 전략임
- 어떤 기업이 데이터를 집중화하여 통제할 때 발생할 수 있는 효과들에 주목할 필요가 있음

### □ 2022년 EU 집행위원회 조사 및 동의의결

- 2022년 12월 20일 EU 집행위원회는 아마존이 온라인 마켓플레이스 서비스 시장에서 시장지배적 지위를 남용한 혐의에 대한 2건의 조사와 관련하여, 아마존이 제출한 시정방안(commitments)을 최종 승인하고 동의의결 결정을 내렸음(황진우·최매화·김형준, 2023)

- 2건의 조사 중 첫 번째가 바로 Khan(2017)이 지적한 바 있는 아마존 마켓플레이스의 판매자 데이터 활용 문제에 대한 것이었음
  - 마켓플레이스 판매자들에 대한 비공개 영업 데이터(non-public business data)를 아마존이 체계적으로 활용하여 자사 상품 판매에 이용한 것을 문제삼았음
  - Vestager 부위원장은 ‘민감한 데이터(sensitive data)’라는 표현을 사용
  - 아마존이 시장 지배력을 가진 플랫폼(마켓플레이스)이면서 동시에 그 플랫폼상에서 직접 상품 판매를 하며 판매자들과 경쟁하는 이중적 지위(dual role)에 있음을 지적
- 아마존이 접근한 판매자 데이터에는, 주문 및 배송된 제품의 수, 매출액, 방문자 수, 과거 실적, 반품·교환 등 소비자 불만제기 등에 대한 정보가 포함
  - EU 지역 마켓플레이스 내의 모든 판매자(80만 이상), 모든 상품(10억개 이상), 모든 거래에 대해 이러한 ‘빅데이터’를 수집하여 축적
  - EU 집행위원회는 아마존의 유럽 마켓플레이스에서 발생한 8천만건 이상의 거래와 약 1억개의 상품 리스트를 포함하는 표본 데이터를 분석
- 위원회 조사에 따르면, 대단히 많은 양의 세부적인 거래 및 상품 데이터가 아마존 소매사업부 직원들에게 제공되고 알고리즘에 의해 자동화된 시스템으로 체계적으로 유입되었으며, 이러한 데이터들을 실시간으로 수집, 활용하여 구체적이고 목표지향적인 결론들(precise, targeted conclusions)을 도출
  - 아마존 자사 상품들의 판매 조건을 최적화하는 등, 마켓플레이스 판매자들의 이익에 반하는 아마존의 전략적 의사결정들에 데이터가 활용되었다는 것이 위원회의 판단
  - 예컨대, 아마존은 데이터를 통해 마켓플레이스상의 다양한 상품군들 중에서 가장 판매가 잘되는 상품군을 파악하여 이에 집중하고, 경쟁 판매자들의 비공개 데이터를 보면서 자사 상품 판매조건들을 변경하였음
  - 이러한 데이터와 알고리즘에 기초하여 아마존은 출시할 신상품의 선택, 개인화된 가격설정, 재고관리, 제품의 공급자 선택 등을 결정
- 이러한 데이터 활용을 통해 아마존은 시장 경쟁 과정에서 감수해야 하는 정상적인 위험(normal risk of retail competition)을 회피하였고 프랑스와 독일 마켓플레이스 서비스 시장에서의 지배적 지위를 남용한 것으로 판단됨

- 프랑스 온라인 소비자의 70% 이상, 독일 온라인 소비자의 80% 이상이 아마존으로부터 상품을 구매한 경험이 있음 (2020년 11월부터 직전 12개월의 기간 기준)
- 아마존은 마켓플레이스내의 전체 상품군들 중 10% 미만에 대해 직접 상품을 등록하고 있는데, 해당 상품군들 내에서는 50% 이상의 매출 점유율을 확보하고 있음
- 대부분의 일반적인 소매업자들은 판매할 상품을 조사, 선정하기 위해 많은 비용을 부담하고 신상품 출시를 위한 투자 및 그 가격 결정 과정에서 상당한 위험을 부담하게 되는데, 아마존은 판매자 데이터를 이용하여 이러한 비용과 위험을 회피하고 있음
- 이러한 영업환경에서, 일반 판매자들은 공정하게 경쟁하고 성장할 기회가 없음
- EU 집행위원회가 이러한 조사 결과에 따라 중간 심사보고서(Statement of Objection)를 2020년 11월 아마존에 송부하였고, 이에 대해 아마존은 시정방안(commitments)을 제출후 협의, 수정과정을 거쳐 최종적으로 동의의결에 이르게 되었음
- 아마존은 비공개 판매자 데이터를 자사 소매사업부의 이익을 위해 사용하지 않기로 하였음 (Vestager 부위원장은 이를 “data silo”의 구축이라고 표현함)
- 즉, 플랫폼에 입점한 판매자들의 영업 데이터를 이용하여 자사 상품 판매의 최적화를 하지 않기로 하였음
- 이러한 시정방안은 아마존의 사업 의사결정에 관여하는 직원과 알고리즘 툴들에 모두 적용됨
- EU 집행위원회는 이러한 조치를 통해 마켓플레이스 플랫폼 상에서의 공정한 경쟁환경(level-playing field)이 회복될 것으로 기대함

#### □ EU 집행위원회 조사에 대한 반대 의견들

- 위와 같은 EU 집행위원회의 조사 및 동의의결 조치에 대해 반대하는 견해들도 당연히 존재하는데, Malobacka-Szwast(2021)에 따르면 다음과 같은 의견들이 제기된 바 있음
- 다른 기업들의 영업 사례들을 참고하여 기회를 찾는 것은 오래전부터 이루어져 온 사업의 한 방식이고 소매유통 부문에서는 관행이나 다름없다는 주장이 있음

- 아마존의 방식과 유사하게, 오프라인 슈퍼마켓 기업들은 판매 데이터를 분석하여 PB 상품을 출시하는 전략을 사용해왔으나 유럽 경쟁법 상의 문제가 된 적은 없음
- 이러한 주장에 대해 Malobbecka-Szwast(2021)는 아마존의 경우는 수집, 활용된 데이터의 규모와 정밀성, 그 활용의 효과성에 있어 과거의 오프라인 슈퍼마켓들과는 비교가 되지 않는다는 점을 지적하였음
- 또한, 플랫폼 사업자로서 아마존이 갖고 있는 시장지배력과 그에 따른 판매자들의 아마존에 대한 의존성 측면에서도 과거의 사례들과는 차이가 있음을 지적
- 아마존의 행위는 결과적으로 경쟁을 촉진하고 가격을 낮추어 소비자들에게 이익을 주는 행위라는 주장도 있음
- 이에 대해 Malobbecka-Szwast(2021)는, 단기적 가격효과만을 보아서는 안 되고, 판매자들의 혁신과 성장 유인을 저해함으로써 중장기적으로 소비자들에게 다양한 상품들 중에서 선택할 수 있는 기회를 잃어버리게 될 수 있음을 지적

### (3) 페이스북의 이용자 정보 결합 활용

- 독일 연방카르텔청은 페이스북의 시장지배적지위 남용 혐의에 대한 심사보고서를 2017년 12월 페이스북에 송부하였는데, 페이스북이 가입자들에게 “개인정보의 광범위한 사용에 대한 동의”를 가입조건으로 부과한 것이 착취적인 조건이라고 판단하였음(공정거래위원회, 2018)
- 페이스북은 개인정보의 광범위한 사용에 대한 동의를 가입조건으로 부과하고, 이를 통해 페이스북 내에서 수집된 이용자들의 정보와 제3의 웹사이트 및 모바일앱에서 수집된 정보를 결합하여 광범위하게 활용
- 여기서 제3의 웹사이트·모바일앱은 (1) 왓츠앱, 인스타그램 등 페이스북의 자회사와 (2) 페이스북 API를 적용하여 페이스북 계정으로 로그인 가능한 웹사이트·모바일앱 등을 말함
- 이러한 웹사이트·모바일앱을 방문만 해도 해당 사이트·앱상의 모든 정보가 이용자의 페이스북 계정으로 전송됨
- 이는 이용자들의 자기결정권을 제한하는 착취적 남용행위에 해당한다고 판단

- 연방카르텔청의 결정에 대해 페이스북은 금지 명령 효력 정지 소송을 제기하였고 2019년 8월 뒤셀도르프 고등법원은 이를 받아들여 해당 결정의 효력을 정지시켰음(이상윤, 2020)
  - 데이터는 쉽게 복제될 수 있는 것으로 이용자들이 어떠한 경제적 피해를 입었다고 보기 어려움
  - 이용자들이 가입시 약관에 동의한 것은 무료서비스 이용으로 인한 편익과 그 대가로 데이터를 제공하는 것을 비교한 결과로서 자발적 동의가 없었다고 보기 어려움
  - 데이터 활용에 대한 충분한 정보 제공이 이루어지지 않은 것은 시장지배력의 결과라기보다는 이용자들의 무관심 또는 편의상의 선택에 따른 결과일 수 있음
- 그러나, 2020년 6월 독일 연방최고법원은 고등법원의 결정을 파기하고 기존의 연방카르텔청의 결정에 따라 페이스북의 이용자 데이터 수집, 이용, 결합 행위가 즉시 중단되도록 하였음(이상윤, 2020)
  - 페이스북 이용자들은 ‘제3자 등으로부터 광범위한 데이터 수집을 하지만 좀 더 개별화된 경험을 제공받을 수 있는 서비스’와 ‘덜 개별화된 경험을 제공하지만 자신이 데이터 제공을 선택할 수 있는 서비스’ 사이에서 선택권을 상실하였음
  - 후자의 서비스에 대한 수요가 있음에도 페이스북이 시장지배적 지위를 이용해 이를 거부한 것으로 볼 수 있으므로, 이를 이용자에 대한 착취로 볼 수 있음

## 마. 쟁점 정리와 과제

- 본 절에서의 논의와 사례검토를 통해 발견된 쟁점들을 정리해보면 다음과 같음
- 첫째, 데이터의 독점 또는 배타적 활용은, 거래상대방이나 소비자 등에 대한 체계적인 비대칭 우위 및 일방적 차별로 이어질 수 있다는 점에서 부당한 이용의 문제가 제기될 수 있음
  - 데이터의 무임승차적 전유에 따라, 데이터 생성에 관계되는 주체들의 권리는 무시되는 한편 데이터 전유자에게 이익이 집중되는 경우에 특히 문제가 됨
  - 그러나, 데이터의 수집 및 가공 과정에 상당한 기술과 비용이 투입되는 경우에는, 무임승차적 전유보다는 지식재산으로서의 성격이 강해지게 되므로 이용 부당성의 성립 가능성도 낮아질 것임

- 둘째, 데이터가 시장지배력을 확보, 유지하는 수단으로 적극적으로 활용된다면 그러한 목적·효과를 갖는 데이터 이용은 부당한 것이 아니냐는 문제가 제기될 수 있음
  - 이는 경쟁제한성의 문제로서 현행법적 규제의 틀 내에서 어느 정도 다뤄질 수 있을 것임
  - 다만, 경쟁에 있어서의 데이터의 필수성 및 가치에 대해 다양한 반론들도 존재하므로, 관심의 대상이 되는 개별 시장의 세부적인 조건 및 구조들에 유의할 필요가 있음
  - EU 집행위원회의 아마존 동의의결 사례에서 잘 드러났듯이, 대형 플랫폼이 데이터를 이용함에 있어서의 규모, 체계성, 정밀성 등도 경쟁제한성 효과를 판단함에 있어 주목할 필요
- 셋째, 데이터가 사회적 비용을 고려할 때 적절한 수준보다 과다하게 이용되는 비효율성의 문제가 있을 수 있음
  - 데이터의 이용이 무임승차적 전유의 성격을 가지며, 개인정보 관련 피해 등의 잠재적 비용 요소들이 외부효과로서 존재하기 때문임
  - 개인정보 보호 등의 관련 규제정책이 데이터 산업 발전을 저해한다는 시각이 존재하지만, 반대로 충분한 규제가 없을 경우 데이터가 남용되어 사회후생을 저해할 가능성도 동시에 존재함을 고려할 필요가 있음
- 본 절에서 검토한 사례들과 같이, 플랫폼 기업들의 데이터 부당 이용과 관련한 문제들은 앞으로 계속하여 발생할 것으로 예상되는데, 위에서 제시한 3가지 쟁점들을 해결해나가는 방향으로 관련 정책들이 마련될 필요가 있음
  - 단, 플랫폼 등 기업들의 데이터 활용과 관련한 기술과 사업 전략은 계속하여 빠르게 진화, 발전하는 과정에 있으므로, 이러한 변화의 가능성을 사전적, 외생적으로 제약하고 경계짓는 규제는 바람직하지 않음
  - 정부는 기본적인 문제의식과 원칙을 제시하고, 기업 스스로 데이터 이용 방식을 생태계 및 소비자 친화적으로 개선해나가도록 할 필요

## 5. 경쟁사업자의 플랫폼 데이터 접근의 이슈와 과제

### (1) 경쟁관련 데이터를 둘러싼 이해관계의 구도

- 민·상사법의 일반원칙에 비춰 보면 데이터 보유 사업자는 자기보유 데이터에 대해 소유 및 관리, 운용, 처분 권한을 갖게 됨.
  - 그런 맥락에서 보면 데이터의 공유나 접근허용 여부도 데이터 보유사업자의 의사로 중심으로 결정되어야 할 것임.
- 하지만 디지털경제하에서는 데이터 보유는 플랫폼 등 사업자의 노력과 경제활동만으로 확보 및 축적된 것이 아니라는 점에서 사적 소유의 대상 이상의 성질을 갖게 됨.
  - 이용자들은 플랫폼을 이용하는데 금전적 비용을 지불하지 않는 대신, 자신의 거래 정보 내지 경쟁관련 데이터를 남기게 되고, 플랫폼은 이를 반대급부로 수취함.
  - 데이터의 제공자와 수취자간에는 경제적 이익에 대한 심리적 균형이 어느 정도 형성됨.
  - 이에 비해, 거래상대방이나 경쟁사업자는 데이터 형성에 기여하지 않았으므로 일단 데이터 이용 및 접근에 대한 결정권을 행사할 수 없는 상황에 놓임.
- 이런 구도 속에 경쟁관련 데이터를 둘러싸고 개별 경제주체들 간에는 아래와 같은 이해관계의 기본 구도가 형성됨
  - 이용자(end user): 플랫폼 이용에 대한 금전적 대가를 지불하지 않는 대신, 자신의 거래 정보 내지 경쟁관련 데이터를 플랫폼에 유보함
  - 플랫폼(data hosting platform): 무상 서비스 제공에 대한 반대급부로 이용자의 거래정보를 취득하여 경쟁 기반으로 활용하는 한편 개인정보 보호 관련 추가 비용 발생함
  - 거래상대방(business user): 데이터에 기반하여 제고된 특정 플랫폼의 가치가 거래비용 증가 또는 플랫폼 이용 대가의 상승으로 연계됨
  - 경쟁사업자(platform competitor): 데이터 기반 경쟁 구도에서 배제되면서 자사 플랫폼의 가치 절하 및 이용자의 편익 저하를 감내해야 함

## (2) 데이터 접근에 대한 이해관계의 대립

- 데이터 확보 여부가 경쟁에 핵심적 조건으로 작용하는 환경 하에서 데이터에 대한 접근권은 추가적인 정책적 보호조치를 요하는 권리로서의 지위도 갖게 됨. 이 점에서 데이터 접근에 대한 이해관계의 대립이 발생함.
  - 이용자: 자신이 제공한 데이터의 공개에 따라 개인정보 보호의 위험이 증가하는 반면에 제공하지 않은 데이터 공개에 따른 전반적인 편익도 증가
  - 플랫폼: 데이터수집에 소요되는 비용을 회수하거나 반대급부를 보상받거나 그렇지 않을 경우 공유 및 접근권을 거절할 유인이 발생
  - 거래상대방: 데이터의 소유권을 누가 보유하는지 보다는 어느 플랫폼에 다면시장을 기반으로 데이터로 인한 네트워크 효과가 크게 발생하는 지가 중요
  - 경쟁사업자: 타 사업자 보유 데이터에 대한 접근을 시도할 유인과 함께 데이터 접근을 거절당할 경우 경쟁제한 이슈를 제기할 유인이 함께 발생
- 이런 상황에서 경쟁사업자에 대해 데이터에의 접근을 제한하거나 차별적으로 허용한다면 진입장벽으로서 경쟁이 제한 또는 왜곡되는 문제가 나타날 수 있음.
  - 특정한 사업자가 대량의 데이터를 보유·이용하는 것이 관련 시장에서 경쟁자가 필적할 수 없는 경쟁상 이점을 제공하거나 신규 진입을 저해함으로써 해당 사업자의 시장지배력을 형성·강화할 우려가 있는지가 관건임.

## (3) 데이터의 개방, 공유, 접근권 보장의 효과

- 사업자가 대용량 데이터를 수집, 저장, 관리하고 추가적인 데이터 수집을 위하여 무료서비스를 제공하려면 많은 비용이 소모됨. 그렇다면 사업자는 그들의 경쟁사업자가 데이터에 접근하지 못하도록 하거나 데이터 공유를 제한할 유인을 갖게 됨. 또한, 데이터와 자사의 관련 서비스를 끼워팔기하는 경우에도 경쟁제한적 효과가 나타날 수 있음. 아울러 소규모 사업자와 신규사업자는 많은 이용자의 데이터를 확보하기 어렵기 때문에 대형 사업자와 효과적으로 경쟁하기 곤란해짐.
- 신규사업자는 데이터를 구매함으로써 기존 플랫폼 사업자에 필적할 수 있는 데이터 축적이 가능함. 그러나 현실에서 신규사업자는 제3자 데이터만으로는 기존 기업이 보유하고 있는 데이터셋과 동일한 양, 질의 데이터 축적이 불가능할 수 있음.

- 플랫폼 특유의 피드백 순환 메커니즘으로 인한 네트워크 효과가 발생하여 신규, 중소기업자가 경쟁하기 어려운 상황이 마련됨
  - 이미 큰 이용자 기반을 보유한 온라인 플랫폼은 한편으로 더 많은 이용자 데이터를 수집·활용하여 자신이 제공하는 서비스의 품질을 개선할 수 있으므로 이를 통하여 다시 더 많은 이용자를 확보할 수 있고(이용자 피드백 고리), 다른 한편에서 더 많은 이용자 데이터를 활용하여 더 많은 수익을 올릴 수 있고 이를 서비스 개선에 투자함으로써 다시 더 많은 이용자를 확보할 수 있음.
  
- 이상을 종합할 때 데이터의 개방, 공유, 접근권 보장의 효과는 아래와 같이 정리됨.
  - 이용자: 데이터로 인한 편익과 더불어 개인정보 위험도 함께 증가할 가능성에 노출
  - 플랫폼: 단기적으로는 수집비용의 회수가능성 상실, 장기적으로는 타 플랫폼 보유 데이터에 대한 확보 비용 감소
  - 거래상대방: 플랫폼의 선택 범위가 확대되고 거래비용이 감소
  - 경쟁 플랫폼: 데이터기반 경쟁조건을 용이하게 확보하되 데이터에 기반한 상품, 서비스 경쟁은 심화
  
- 요컨대, 플랫폼이 보유한 데이터에 대한 경쟁사업자의 접근권 보장은 이용지. 거래상대방, 경쟁사업자는 물론이고 플랫폼 보유 사업자에게도 장기적으로는 이로운 선택지임.
  - 물론 그 동안 데이터의 접근권 제도화에 대해서는 상반된 주장도 제기되어 왔음. 이용자 데이터는 그 속성상 비경합적(non-rivalrous)이고 특정한 사업자가 이를 배타적으로 독점할 수도 없으므로 경쟁자 배제나 진입장벽 증가의 우려가 크지 않으며, 따라서 경쟁에서 데이터 접근권이 필수적인 것은 아니라는 인식이 그 근거임.
    - 가령, 특정한 사업자가 이용자에 관한 데이터를 수집·이용한다고 할지라도 다른 사업자 역시 동일한 이용자에 관한 데이터를 수집·이용할 수 있고, 온라인 서비스 사업자가 이용자에 대하여 다른 사업자에게 데이터를 제공하지 못하도록 할 수도 없으며, 실제로 이용자들은 자신에 관한 데이터를 다수의 온라인 플랫폼 사업자에게 제공하고 있다는 점을 근거로 제시함.

- 하지만, 플랫폼 사업자로서는 경쟁사업자에 대한 데이터의 제공, 공유, 접근을 거절하는 것이 적어도 단기적 이익에 부합함. 따라서, 플랫폼 사업자로 하여금 경쟁사업자에 대해 데이터 접근을 방해, 거절하지 않도록 정책적으로 대응하는 것은 플랫폼 시장과 경쟁, 그리고 공익에 부합할 것임.
- 경쟁사업자도 다양한 다른 방법으로 유효한 경쟁자로 생존하는 데 필요한 이용자 기반을 확보할 수 있고, 이용자 데이터를 확보하는 것만으로는 서비스의 품질을 개선할 수 없고 오히려 이를 분석하여 활용하는 것이 중요하며, 사업자들은 이용자로부터 직접 수집하는 방법 이외에도 필요한 데이터를 확보할 수 있는 다른 다양한 방법이 존재한다고 봄.

#### (4) 데이터 접근·사용 방해관련 국내외 주요 사례

- 시장 우위를 점하고 있는 플랫폼이 경쟁플랫폼이 특정 데이터(특히 경쟁 관련 데이터)에 접근하거나 수집하는 것을 방해하는 사례들이 경쟁당국에 의해 제재되고 있음
- 국내: 포털의 부동산 데이터 제공 거절 사례
  - 이용자: 부동산 정보업체(CP)에 대해 자신의 매물정보를 제공
  - 부동산 정보업체: 매물정보를 플랫폼에 올려 매매를 유도
  - 경쟁사업자: 기존에 거래관계에 있던 부동산 정보업체에 대해 데이터 제휴를 시도
  - 데이터 보유 플랫폼: 부동산 정보업체(CP)와의 계약에서 부동산 매물정보를 제3자에게 제공하지 못하도록 하는 조항을 도입
- 이에 대해 공정거래위원회는 해당 플랫폼에 대해 시정명령과 함께 10억 3,200만 원의 과징금을 부과함(2020. 9.)

#### (가) 국외: SNS 사업자의 데이터 접근 거부 사례

- 이용자: SNS를 이용하면서 플랫폼에 데이터를 남김
- 제3자 앱 개발사: 이용자 데이터를 활용하면 보다 많은 앱 이용자를 확보할 수 있으므로 데이터 접근을 희망

- SNS 플랫폼: 제3자 앱 개발사의 앱이 자신과 동일한 기능을 제공하지 않고, 경쟁 SNS사업자와 거래하지 않는 경우에만 이용자에 관한 데이터의 접근을 허용한다는 경쟁제한적 조건을 부과
- 그 결과, 향후 자사에 위협이 될 가능성이 있는 앱을 개발할 잠재적 경쟁자를 사전에 제거하였으며, 경쟁 SNS 사업자와의 거래를 제한하게 됨
- 미국 FTC는 법원에 SNS 시장에서의 데이터 접근 관련 경쟁제한적 조건 부과에 대한 금지와 동 위반행위 및 향후 유사행위에 대한 금지청구 소송을 제기함

#### (5) 여전히 제기되는 법적 이슈들

- 데이터 접근권 보장의 정책적 수요와 당위에도 불구하고 여전히 해법이 수렴되지 않은 이슈들이 잔재함. 그 이슈들은 민사, 상사, 경쟁, 그리고 정보통신 등 여러 법/정책의 영역에 분포해 있음.
- 이를 요약하면 아래와 같이 정리됨.
  - 플랫폼이 보유한 데이터의 소유권은 누구에게 귀속되어야 하는가?
  - 플랫폼이 보유한 데이터는 공공재인가, 필수설비인가, 아니면 여전히 비경합성(non-rivality)을 지니는가 ?
  - 플랫폼이 데이터를 수집하는데 따른 비용과 공개에 따른 비용은 감내할 수준인가?
  - 데이터 보유 플랫폼에게 보유 데이터를 공개하거나 접근권을 보장할 공적 의무가 존재하는가?
  - 데이터에 대한 접근권 보장 조치는 개인정보를 포함한 이용자 편의 및 후생의 총합을 증진시킬 것인가 ?
  - 데이터 접근 및 공유의 범위는 어디까지 확장될 수 있는가 ?
  - 경쟁사업자에 대한 데이터 접근권 보장을 위한 수단은 사전규제와 사후규제 중 어느 것이어야 하는가?

- 다만 데이터 접근의 필요성을 인지한 각국의 규제 및 경쟁 당국들은 경쟁사업자에 대해 데이터 접근권을 보장하는 방향으로 입법 및 정책을 추진 중임.
  - EU DMA, 독일 GWB를 위시하여 현재는 폐기된 미국의 플랫폼 패키지 법안에서는 공히 경쟁사업자에 대한 데이터 접근권 문제를 핵심적 규율사항으로 포함시키고 있어서 이를 둘러싼 논쟁의 여지는 입법적 동향으로 상당부분 해소된 것으로 볼 수 있음.
- 단, 이를 입법적으로 접근하는 방식은 데이터 거절행위를 시장지배적 지위 남용으로 서 사후규제하는 방식과 데이터접근권을 경쟁사업자에게 보장하도록 보유사업자에게 의무를 부과하는 사전규제 방식의 두 가지 유형으로 나뉨.

## 가. 데이터 접근권에 관한 주요국의 입법 동향

### (1) 독일 경쟁제한금지법

#### (가) 개정의 배경

- 2021년 독일 경쟁제한방지법 제10차 개정법<sup>80)</sup>은 디지털 경제에서 시장지배적사업자의 지위 남용행위에 대한 규율을 현실화하고, 불공정경쟁행위에 대한 독일 연방카르텔청의 규제권한을 강화하고자 하는 취지에서 추진됨. 신설된 GWB 제19a조에서는 ‘경쟁 관련 데이터에 대한 접근’ 관련 내용이 포함되었음.
- 동 개정에서는 ‘GWB-디지털화법’이라는 명칭을 사용하는 한편 그 목적을 EU 지침(RL 2019/1/EU)의 국내법 전환, 디지털 질서프레임 마련, 절차규정의 신속화로 제시함
- 이와 더불어 ‘제2장 시장지배, 기타의 경쟁제한 행위’, ‘제6장 경쟁당국의 권한과 손해배상 및 이익환수’, ‘제7장 기업결합통제’, ‘제8장 독점위원회’, ‘제9장 전기와 가스 영역에서의 도매사업자에 대한 시장투명성’ 등, 법 전반에 걸쳐 이루어짐

80) 2021년 1월 18일 독일 연방의회를 통과하여 같은 해 1월 19일에 발효됨.

(나) 데이터 관련 주요 개정사항

- 디지털 경제에서 데이터에 대한 중요성이 증대됨에 따라 사업자의 시장지배력 평가를 위한 기준에 ‘경쟁 관련 데이터에 대한 접근’을 도입하였음. 따라서 타사업자에 대한 경쟁 관련 데이터의 접근 거절은 시장지배적지위 남용행위로 금지될 수 있음.<sup>81)</sup>

[참고] 제19a조(1)에 해당하는 사업자의 금지행위

조번	내용
GWB 19a조	§ 19a 경쟁에 있어서 압도적 중요성을 가지는 시장의 경계를 넘어서는 (중요한) 의미를 가지는 사업자의 지위 남용행위
	(1) 연방카르텔청은 GWB 제18조 제3a항에 따른 시장에서 상당한 규모의 사업을 영위하는 사업자가 시장간 경쟁에 있어서 압도적 중요성을 가지는지를 처분을 통하여 결정할 수 있다. 이러한 결정에 있어서 고려요소들은 다음과 같다. ... ..
	4. <u>경쟁 관련 데이터에 대한 접근</u>
	<후략>
	(2) 연방카르텔청은 제1항의 기준에 해당하는 사업자에 대하여 다음과 같은 행위를 금지할 수 있다, ... ..
	4. 지배적 시장에서 수집된 경쟁 관련 데이터 이용은 물론 지배적 시장 외의 출처에서 기타 경쟁 관련 데이터의 결합이 다른 시장에서의 시장진입장벽을 형성 또는 강화하거나 기타의 방법으로 다른 사업자를 사업활동을 방해하는 행위, 또는 그러한 데이터의 이용을 거래조건으로 요구하는 행위
	5. 상품 또는 서비스의 상호운용성 또는 데이터 이동성을 어렵게 함으로써 경쟁을 방해하는 행위 <후략>

- 한편, 앞선 제9차 개정에서 시장지배적지위에 관한 제18조 규정과 관련하여 제3a항에서 다면시장과 네트워크 시장에서 사업자의 시장지위와 경쟁상황을 평가하기 위한 다섯 가지의 특별 기준, 즉 제1호 내지 제3호의 네트워크 효과 그리고 제4호의 데이터 접근 및 제5호의 혁신경쟁에 관한 내용이 추가되었음.
- 9차 개정 이후 경쟁 관련 데이터에 대한 접근은 사업자의 시장 내 지위를 평가하기 위한 하나의 표지로서 GWB 제18조 제3a항 제4호에서 규정되어 있었음에도 불구하고 동조 제3항에 ‘데이터 접근’을 추가적 표지로 명시한 것은 ‘모든 경제 부문에서 확대되고 있는 데이터의 중요성’을 고려하기 위함임

81) 이와 같은 내용은 독일연방경제에너지부가 시장지배적사업자 남용행위 규제체계 현대화를 위한 연구를 추진하여 2018년 8월 발간한 “시장지배적사업자에 대한 남용 감독의 현대화” 보고서상의 권고 사항 중 하나에 해당하는 내용임.

- [평가] 경쟁법상 사후규제 보완 방식. 데이터 이용 및 이동성을 경쟁법에 명시한 점에 의미가 있음. 단 지위의 남용 여부에 대한 사후적 평가가 수반되어야 하므로 데이터 접근권 전반을 창설하거나 보장하는 것은 아님

## (2) 미국의 온라인선택 및 혁신법안

- 상업적 이용자의 상품서비스를 통해 플랫폼에서 획득하거나 생성된 상업적 이용자와 최종 소비자의 비공개 데이터를 플랫폼이 자사 상품, 서비스를 제공하거나 지원하기 위해 사용하는 행위를 금지함.(§2.b.3.)
- 상업적 이용자의 상품, 서비스를 통해 플랫폼에서 획득하거나 생성된 상업적 이용자와 최종소비자의 비공개데이터에 대한 상업적 이용자의 접근을 제한하거나 방해하는 행위를 금지함.(§2.b.4.)
- [평가] 부작위 의무 부과 중심의 사후규제적 성격이 강함

## (3) EU 「온라인 플랫폼 시장의 공정성 및 투명성 강화를 위한 2019년 EU 이사회 규칙」과 데이터 이슈 - 데이터 접근권

- EU에서는 플랫폼사업자의 시장지배적지위 남용행위에 대한 경쟁법 적용만으로는 중소기업(SMEs)과 상생하는 공정한 온라인 시장환경을 조성하는데 한계가 있다는 우려가 지속적으로 제기되자 온라인 플랫폼 시장에서의 거래관행 실태조사 결과를 토대로, EU 이사회가 2019년 6월 20일 「온라인 플랫폼 시장의 공정성 및 투명성 강화를 위한 2019년 EU 이사회 규칙」<sup>82)</sup>을 제정하여 2020년 7월부터 시행되었음.
  - 이 규칙은 온라인 거래 생태계의 공정성·투명성을 규율하기 위한 사실상의 세계 최초 입법례로서 의의가 있음.
- 규칙은 19개 조항으로 이루어져 있는데, 이 가운데 데이터 접근에 관한 제9조에서는 다음과 같이 규정하였음<sup>83)</sup>

82) 「Regulation (EU) 2019/1150 of the European Parliament and of the Council of 20 June 2019 on promoting fairness and transparency for business users of online intermediation services」.

83) 외교부, “EU, 온라인 플랫폼 공정성·투명성 규정 제정” 자료 참조

1. 온라인 중개서비스 제공자는 상업적 이용자, 소비자가 온라인 중개서비스를 사용하면서 제공하거나 서비스 제공과정에서 생성된 개인정보나 다른 데이터에 대한 상업적 이용자의 기술적, 계약적 접근권에 대한 설명 또는 그러한 권한의 부재에 대한 설명을 약관에 포함해야 한다.
2. 위 1에 언급된 설명을 통해 온라인 중개서비스 제공자는 다음의 사항을 상업적 이용자에게 적절히 고지하여야 한다.
  - (a) 상업적 이용자들 또는 소비자들이 온라인 중개 서비스의 사용과 정에서 제공하거나 서비스의 제공과정에서 생성된 개인정보 또는 다른 데이터에 온라인 중개서비스 제공자가 접근할 수 있는지 여부, 접근 가능한 데이터의 범위, 접근 조건
  - (b) 상업적 이용자가 온라인 중개서비스를 사용하면서 제공하거나, 그 상업적 이용자나 그 상업적 이용자의 상품이나 서비스를 구매하는 소비자들에게 서비스를 제공하면서 생성된 개인정보 또는 다른 데이터에 해당 상업적 이용자가 접근할 수 있는지 여부, 접근 가능한 데이터의 범위, 접근 조건
  - (c) 위 (b)에 더하여, 모든 상업적 이용자와 소비자들이 제공하거나, 온라인 중개서비스를 모든 상업적 이용자와 소비자들에게 제공하는 과정에서 생성된 개인정보 또는 다른 데이터에 대해 - 집계된 형태 (aggregate form)를 포함하여 - 특정 상업적 이용자가 접근할 수 있는지 여부, 접근 가능한 데이터의 범위 및 접근 조건
  - (d) 위 (a)에 따른 데이터가 제3자에게 제공되는지 여부, 제3자에 대한 데이터 제공이 온라인 중개서비스의 정상적인 작동에 필수적이지 않은 경우에는 그러한 데이터 제공의 목적에 관한 정보와 상업적 이용자가 그러한 데이터 공유에 참여하지 않을 수 있는 가능성
3. 동 조항(제9조)은 규정 (EU) 2016/679, 지침 (EU) 2016/680, 지침(EU) 2002/58/EC의 적용을 침해하지 않는다.

[평가] 공적 규제 보다는 약관에 데이터 접근에 관한 사항을 포함하도록 유도하는 자율규제 방식이라는 점에 특징이 있음.

(4) EU 디지털시장법(DMA)상 데이터 접근권 관련 규정(2024년부터 본격 시행 예정)

- DMA에 따른 게이트키퍼는 제3조 제7항에 따라 관문 역할을 하는 것으로 확인된 각 CPS에 대해서 EU 집행위원회와 협의된 조건 하에서 11개 의무사항을 준수해야 함. 이 가운데는 데이터접근과 관련된 의무들이 아래와 같이 포함되어 있음.
  - 상업적 사용자와 경쟁관계에 있는 경우 해당 사용자의 비공개 데이터 사용이 금지됨: “사업이용자들이 제공했거나 이용사업자들의 활동을 통해 얻어진 데이터를 사업이용자들과 경쟁하는데 사용하지 말 것”(DMA 제6조 제1항(a))
  - 최종사용자가 생성한 데이터의 이동성 및 접근성을 허용해야 함: 사업이용자들 또는 최종 소비자가 데이터를 효과적으로 이동할 수 있는 수단을 제공할 것 “(DMA 제6조 제1항(h))
  - 상업적 사용자가 스스로 창출한 데이터에 대해 무료로 실시간 접근 및 사용을 허용해야 함: “사업이용자들 또는 사업이용자들이 승인한 제3자가 데이터를 무료로 효과적이고 계속적으로 실시간 접근 및 사용할 수 있도록 허용할 것” (DMA 제6조 제1항(i));
  - 제3의 온라인 검색엔진, 앱마켓 등에서 사업자가 요청하는 경우 FRAND조건으로 데이터에 대한 접근을 허용해야 함: “제3의 온라인 검색엔진 사업자가 요청하는 경우, FRAND 조건으로 랭킹, 검색, 클릭, 조회 데이터에 대한 접근을 허용할 것”(DMA 제6조 제1항(j))
- [평가] 작위의무와 부작위의무, 사전규제와 사후규제가 혼재한 형태. 수집한 데이터를 경쟁수단으로 활용하지 못하도록 하는 강력한 조치 부과하고 데이터 접근 및 사용권을 비교적 구체적으로 보장한 점이 주목됨. 단 데이터 접근권의 대상을 경쟁사업자에 국한하지 않고 거래상대방인 사업이용자들로 확장

(5) EU 디지털서비스법(DSA)

- 규제 당국은 추천 시스템을 포함한 알고리즘의 설계 방식, 작동 원리, 기능, 검증 결과 등 법의 준수 여부를 점검하고 평가하기 위해 필요한 경우, 대형 플랫폼 사업자들의 내부 데이터에 접근할 수 있음
- [평가] 규제당국의 정책 집행의 효율성을 위한 데이터 접근권 부여

## (6) 국내의 의원 발의 법안

- 이용자 수, 매출액, 제공 서비스 등이 일정 기준에 해당하는 전기통신사업자가 보유한 특정 데이터에 대해 이용자와 기업 등 접근 허용을 의무화하는 방안 논의됨
- 서비스 이용자와 온라인플랫폼 입점 사업자, 스타트업 등 제3자가 단순 집계 정보나 사물 관련 정보 등 의무사업자만 취득 가능한 정보, 기간통신서비스 커버리지와 품질 관련 정보, 이용자별 검색순위, 상품·콘텐츠 등 소비·이용 순위 등 법이 규정한 데이터에 대한 접근을 허용
- [평가] 규제당국에 의한 사전규제 방식. 입점사업자 및 경쟁사업자가 상품 판매를 증대할 수 있는 데이터 확보와 광고효과 측정이 가능해짐

## (7) 시사점 및 유의점

- 이상의 입법동향에 비추어 볼 때, 사전, 사후규제 수단의 어느 하나에 집중할 필요는 없는 것으로 판단됨.
- 개인정보 취급 민감성을 고려하고 문제 발생 최소화를 위해 개인정보보호법상 데이터 접근 규정을 준수하고 저작권법상 저작권자 권리 보호가 필요함.
- 자신의 데이터를 개방하는 과정에서 프라이버시나 영업비밀의 침해 및 보안 취약성 문제를 야기할 가능성도 제기됨.
- 플랫폼 사업자들 가운데는 데이터를 암호화하거나 서버에 접속기록을 남기지 않는 데이터도 있을 수 있는데 이런 경우에는 접근권 규제를 적용하기 쉽지 않을 수 있음

## 나. 데이터 접근 거절과 필수설비의 문제

### (1) 착안점과 유의점

- 검색엔진, 맞춤형 광고 등 새로운 경쟁환경의 등장에 따라 경쟁을 위한 필수불가결한 조건으로서 데이터의 가치 역시 증가하는 상황임.
- 데이터를 필수요소로 하는 사업의 경우 네트워크 효과와 같은 특수성으로 인하여 초기에 독점력을 확보한 시장지배적사업자의 플랫폼은 데이터 확보의 수월성, 우월성 때문에 보유한 독점력을 유지·강화할 가능성이 높기 때문에 시장 경쟁의 활성화를 위해 데이터 독점에 대한 규제기관의 개입이 필요

- 따라서 시장지배력을 보유하고 있는 사업자의 데이터 접근 및 거래거절의 문제를 해소하고 관련시장의 경쟁을 활성화하기 위해 시장지배적사업자의 거래거절의 위법성 판단기준인 ‘필수설비이론’을 데이터에도 적용할 필요가 있음.

## (2) 주요국에서의 필수설비 이론의 형성과 발전

- 미국 연방항소법원의 1977년 Hecht 판례에서 처음으로 사용된 필수설비이론(essential facilities doctrine)은 “사업의 수행에 필수적인 투입요소이자 복수의 사업자가 중복하여 구축하기 어려운 설비”로 정의하고, 필수설비이론에 대하여 “필수설비 보유자로 하여금 필수설비를 보유하고 있는 시장의 연관 시장에서 타 사업자의 활동을 위하여 적정한 가격에 무차별적인 접근을 허용하도록 의무화”하는 것을 내용으로 함.
- 미국에서 필수설비의 성립요건은 1983년 MCI판결에서 “MCI Test”를 제시하면서 정형화되기 시작함. MCI Test에서 제시한 필수설비의 성립요건은 다음과 같은 4가지임
  - 첫째, 경쟁사업자의 사업 수행에 필수적인 설비를 시장지배적사업자가 통제
  - 둘째, 경쟁사업자의 현실적 또는 합리적 이유에 의한 필수설비와 동등한 설비의 복제 불가능성
  - 셋째, 경쟁사업자의 필수설비에 대한 접근이나 거래거절
  - 넷째, 필수설비의 제공이 실질적으로 가능할 것
- EU의 경우, 시장지배적사업자의 제공 거부가 경쟁에 심각한 영향을 미칠 경우, 필수설비를 제공할 의무가 주어짐. 제공 거부가 경쟁에 미치는 영향을 파악하기 위해서는 다음 세 가지 요인을 검토함
  - 첫째, 해당 설비가 수요자의 사업활동에 얼마나 중요한지 여부
  - 둘째, 필수설비의 수요자가 필요한 재화나 서비스를 다른 경로를 통해 얻을 수 있는지 여부
  - 셋째, 필수설비의 수요자와 시장지배적사업자가 경쟁하게 될 시장에 다른 경쟁사업자가 존재하는지 여부

- 이 가운데 세 번째 요건은 필수설비 수요자의 설비접근을 통한 사업활동이 관련 시장에서 충분한 경쟁증진을 결과하지 않는다면 시장지배적사업자의 설비 제공을 강제하지 말아야 한다는 것임.
  - 만약 시장지배적사업자의 설비 제공 거부가 적은 수의 경쟁사업자들 중 한 기업의 시장 퇴출을 야기할 가능성이 크다면 시장지배적사업자는 설비를 제공하여야 함. 또한 “애로설비 등 희소한 자원을 독점적으로 보유하고 있는 사업자는 현실적, 잠재적 경쟁자에게 차별 없이 당해 자원을 제공하거나 접근을 허용”해야 함.<sup>84)</sup>
- 한국 공정거래법상 필수설비에 대한 접근 방해 행위는 현행 공정거래법 제5조 제1항 제3호의 “다른 사업자의 사업활동을 부당하게 방해하는 행위”로 취급됨. 특히 지난 2001년 개정된 동법 시행령에서는 제5조 제3항 제3호(사업활동방해의 일 유형) 및 제4항 제3호(신규진입방해의 일 유형)에서 ‘필수요소’라는 용어를 명시적으로 채택하고 그 위법성의 판단기준을 제시하여 현재에 이르고 있음.
- 시행령 제5조 제3항 제3호에서는 “정당한 이유없이 다른 사업자의 상품 또는 용역의 생산·공급·판매에 필수적인 요소의 사용 또는 접근을 거절·중단하거나 제한하는 행위”, 제5조 제4항 제3호에서는 “정당한 이유없이 새로운 경쟁자의 상품 또는 용역의 생산·공급·판매에 필수적인 요소의 사용 또는 접근을 거절·중단하거나 제한하는 행위”를 시장지배적지위의 남용행위로 규정하여 기존사업자 및 신규사업자에게 필수설비이론이 적용됨
  - 한국에서는 ‘신용카드 가맹점 공동 이용망’의 거래거절에 대한 대법원 판례에서<sup>85)</sup> 네트워크의 필수설비적 성격을 논의하였으며, ‘SK 멜론 사건’에서 서울고등법원은 ‘필수성’, ‘통제성’, ‘대체가능성’을 필수설비 여부의 판단기준으로 제시한 사례가<sup>86)</sup> 있음

### (3) 데이터에 대한 필수설비이론 적용 조건

- 시장지배적사업자가 보유하고 있는 데이터가 경쟁사업자에게 필수적인 투입요소이고 경쟁사업자가 새롭게 구축하기 어려운 요소일 경우, 거래거절이나 접근 거절의 위법성 판단기준으로 필수설비이론을 적용하기 위해서는 다음의 요건이 충족되어야 함.<sup>87)</sup>

84) 이봉의, “공정거래법상 필수설비법리의 현황과 과제-심결례 및 판례를 중심으로-”, 「상사판례연구」, 2006, 제19판 제1호, 4면 참조.

85) 대법원 2005.8.19. 선고 2003두5709 판결.

86) 서울고등법원 2007.12.27. 선고 2007누8623 판결.

87) 방정미, “데이터 경제활성화를 위한 필수요소법리 재검토”, 「경제규제와 법」, vol. 13, no. 1., 2020, 96-115면 참조

- 데이터를 독점하고 있는 시장지배적사업자가 경쟁사업자가 사용하기를 원하는 데이터의 사용 권한을 가지고 있어야 함(지배력, 통제권)
  - 해당 데이터 없이는 경쟁이 불가능하거나, 중대한 경쟁열위 상태가 지속 됨(필수성)
  - 데이터 공유를 요구하는 경쟁사업자에게 반드시 해당 데이터의 사용을 대체할 수단이 없어야 함(대체불가능성)
  - 데이터를 독점한 기업이 해당 데이터를 제공할 수 있는 방법이 있어야 함(제공 가능성)
  - 데이터를 독점한 기업의 접근 거절의 정당한 사유가 없어야 함(정당사유 부재)<sup>88)</sup>
- 위 3)에서 대체불가능성이란 당해 데이터를 재생산하거나 다른 데이터로 대체하는 것이 사실상·법률상 또는 경제적으로 불가능한 것을 의미함.
- 필수설비이론이 도입되던 초기에는 필수설비의 대상이 철도, 도로, 항만, 통신네트워크 등 물리적 설비였기 때문에 유휴설비가 없는 경우에는 경쟁사업자에게 필수설비를 제공하기 위해서는 관련 설비를 새로이 구축하여야만 가능한 경우가 상당수 존재하였음. 데이터의 경우에는 물리적 설비를 구축하여야 하는 경우는 과거에 비해 훨씬 적을 것으로 예상되며, 필요설비의 구축도 적은 비용으로 가능할 것으로 예상됨. 이 점에서 제공 가능성은 합리적 비용으로 데이터를 제공함을 의미하게 됨.

#### 다. 데이터 접근에 관한 두 가지 규제 수단

##### (1) 사전규제와 사후규제

- 데이터 접근에 관한 규제 수단으로는 i) 데이터 제공의 부당한 거절에 대한 사후규제와 ii) 데이터 접근권의 보장을 위한 사전규제의 두 가지 형태가 가능함
- 이 가운데 데이터 거절에 대한 사후규제는 경쟁법을 보유한 나라는 모두 현 시점에 집행 가능한 반면에, 데이터 접근을 위한 사전규제는 전통적 경쟁법이나 정보통신규제법상 근거규정이 미비한 상태임.
  - 단, 사후규제는 데이터 제공에 대한 거절행위가 발생한 경우에만 개입이 가능하다는 점에서 한계가 있음.

88) '정당한 사유'의 판단기준으로 「시장지배적 지위 남용행위 심사기준」의 IV.3.다.(4)를 참조.

□ 따라서 경쟁사업자에 대한 데이터 접근권을 보장하려면 사후규제를 강화하거나 사전 규제를 신설하는 두 가지의 방안을 강구해야 함

● 참고: 사전규제와 사후규제의 비교

- 사전규제: 작위의무 부과 중심. 금지행위 열거 방식. 행위요건 충족으로 위법성 도출 가능. 주로 경쟁당국이 집행해 왔으나 최근 경쟁당국의 사전규제에 대한 관할권 확장이 시도됨
- 사후규제: 부작위의무 부과 중심, 행위요건과 더불어 효과요건 충족시 비로소 위법성 도출 가능, 주로 규제당국이 집행해 왔으나 규제당국의 사후규제에 대한 관할권 확장 역시 시도되는 상황

(2) 현행법상 데이터 제공의 부당한 거절에 대한 제재 가능성과 한계

- 데이터 제공 거절은 경쟁법상 시장지배적 지위의 남용이나 불공정거래행위에 해당하므로 경쟁법상 사후규제의 대상(거래거절, refusal to deal)이 됨. 현행법상으로는 공정거래법 제5조 제1항 제3호의 사업활동방해, 제45조 제1항 제1호 부당한 거래 거절로 의율 가능함. 단 데이터 제공의 거절 모두가 위법은 아니며 부당성 판단을 거쳐 최종적인 위법여부가 결정
- 이때의 부당성의 판단기준은 국가마다 입법 및 경제 환경에 따라 다소의 편차가 존재함. 한국의 경우 2007년 대법원 전원합의체 판결에 따라 거래거절이 시장에서의 독점적 지위를 유지, 강화할 의도나 목적을 가지고 있고 실제로 경쟁을 제한할 우려가 존재하는 경우에 비로소 부당성을 인정할 수 있다는 판결을 부과한 이래 거래거절 또는 사업활동방해의 부당성 입증에 어려워진 상태임
- 데이터 거래거절에 대한 사후규제는 사안별 구체적 타당성을 확보하기에 용이. 특히 플랫폼의 개인정보 이슈를 경쟁법에 수용하려는 시도가 등장한 이래 거래거절의 경쟁제한성과 개인정보보호를 종합한 범위반여부의 판단이 가능할 것임. 기존 경쟁법을 활용할 수 있으므로 단기적 대응에 용이한 반면에 적기에 효과적 대응에는 한계가 있음.
- 또한 기존의 거래거절은 상품, 용역의 제공을 대상으로 해 왔고 플랫폼이나 데이터 경제에 특유한 논점들을 모두 포섭하는 데는 한계가 있다는 점에도 유의할 필요가 있음.

## 라. 정리 및 결론

- 현행법상 데이터 접근이나 제공 거절에 대한 방어적, 수동적, 사후적 접근법의 효용과 한계가 현재 가시적으로 드러나는 상황
  - 경쟁조건으로서 데이터의 중요성이 제고될수록 사후규제 수단의 한계 인식은 커질 것으로 전망됨
  - 이 점에서 사전규제 방식의 모색 필요성도 지속적으로 제기될 가능성이 있음
  - 주요국의 입법 동향도 국내에서의 입법 추진 동력으로 작용할 것으로 예상됨
- 단, 데이터 접근에 따른 법적, 기술적, 계약적 제약을 극복해야 하는 것이 관건이 될 것임
  - 암호화나 크롤링 방지 기술과 같은 기술적 장벽은 데이터 수집을 어렵게 하거나 불가능하게 할 수 있음에도 유의할 필요가 있음
  - 국내의 개인정보보호법 등 같은 법적 환경도 데이터 수집이나 접근권 확대에 제한요인으로 작용할 가능성이 있음.
  - 플랫폼과 거래상대방과의 계약상 데이터 활용에 대한 제약이 경쟁법상 부당성의 요건을 충족하는지 여부에 따라 계약조항도 데이터 접근권에 대한 장벽으로 작용할 수 있음.
- 요컨대, 데이터 접근에 대한 사전규제를 공적규제 방식으로만 설계할 것인지에 대해서는 신중한 접근과 모색이 필요하다고 봄.
  - 그런 점에서 2019년 EU이사회 규칙의 사례처럼, 표준약관에 관련 내용을 포함시키는 등의 연성 규범(soft law) 혹은 자율규제(self regulation)을 활용하는 방법도 대안이 될 수 있을 것임.

## 6. 국내외 디지털 플랫폼의 데이터 공유 및 개방에 대한 사례 연구

### 가. 디지털 플랫폼과 데이터 개방

#### (1) 개방적 플랫폼 생태계의 필요성

##### □ 개방적 생태계의 의미

- 개방적 생태계는 디지털 플랫폼의 다면 네트워크에 참여하는 다양한 주체들이 서비스나 기술의 개발과 혁신에 참여하고 이를 위해 자원을 공유하는 것을 의미 (Chesbrough, 2014)

##### □ 개방적 생태계가 디지털 플랫폼에 가져오는 효과

##### ● 혁신 강화

- 개방형 생태계는 아이디어와 혁신이 공유되고 공동으로 개발될 수 있는 협력적 환경 조성은 플랫폼과 플랫폼 참여자 모두에 이익을 제공
- 리드 유저 이론(Lead User Theory)에 따르면, 사용자는 종종 생산자보다 빠르게 혁신할 수 있음(Von Hippel, 1989)
- 디지털 플랫폼은 이를 통해 더 넓은 커뮤니티의 창의성과 전문성을 활용하여 빠르고 다양한 혁신을 이끌어낼 수 있음

##### ● 네트워크 효과

- 개방적 생태계는 더 많은 참가자가 플랫폼에 가입하고 기여함으로써 네트워크 효과를 증폭시키고, 이는 모든 사용자에게 더 큰 가치를 제공

##### ● 플랫폼 확장성

- 개방적 생태계는 플랫폼의 효율적인 확장을 가능하게 하며, 플랫폼 이용사업자들과 외부 개발자에게 혁신을 아웃소싱함으로써 플랫폼은 내부 자원이나 비용을 비례적으로 증가시키지 않고도 제공물을 확장할 수 있음(Tiwana, Konsynski, & Bush, 2010).

□ 디지털플랫폼 참여자와의 공진화 관점에서 개방적 생태계

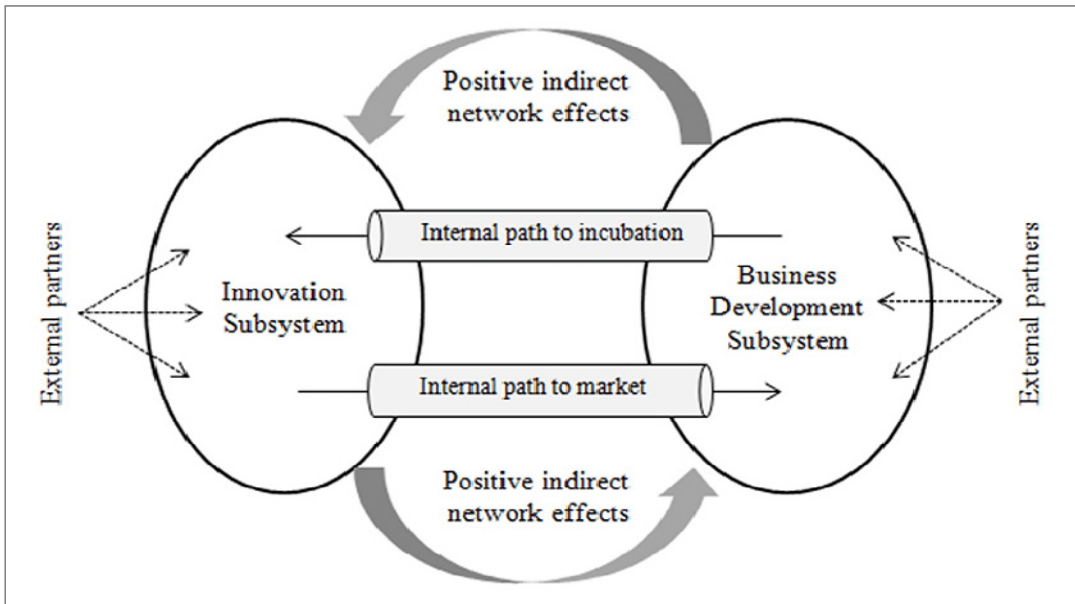
● 생태계 이론과 자원 기반 관점(Resource-Based View, RBV)

- 무어의 생태계 이론은 기업들이 새로운 혁신 주변에서 능력을 공진화시키는 다양한 참여자로 구성된 생태계의 일부로 간주되어야 함
- 협력과 공진화가 개방형 생태계의 핵심으로 바니(Barney, 1991)의 RBV는 기업이 경쟁 우위를 달성하기 위해 독특한 자원을 개발해야 한다고 주장, 개방형 생태계는 외부 기여를 통해 기업의 자원을 증가시키는 방법으로 볼 수 있음

● 참여 주체의 다양성과 자원 공유의 중요성

- 디지털 플랫폼의 개방형 생태계에는 다양한 참여 주체인 앱 개발자, 사용자, 광고주, 콘텐츠 제공자 등이 참여하며, 플랫폼의 자원의 개방될수록 참여자의 전문 지식과 역량이 증가하고 혁신과 서비스가 향상
- 디지털플랫폼의 자원 공유는 플랫폼의 기능을 향상시키고, 사용자 경험을 다양화하며, 새로운 시장 기회를 창출. 이는 플랫폼의 지속 가능한 성장과 경쟁력을 위해 필수적

[그림 III-4] 플랫폼 개방의 선순환 구조



출처: Isckia, T., De Reuver, M., & Lescop, D. (2020). Orchestrating platform ecosystems: the interplay of innovation and business development subsystems. *Journal of Innovation Economics & Management*, 197-223.

**(2) 플랫폼 개방성의 차원 분류**

- 브룩휴이젠(Broekhuizen, et.al, 2021)등의 「디지털 플랫폼 개방성: 동인, 차원 및 결과」 연구에서는 플랫폼 개방성을 개발자 또는 최종 사용자와 같은 플랫폼의 다양한 역할에 대한 참여, 개발 또는 사용에 대한 제한을 덜 두는 정도로 정의하고 있으며, 플랫폼 생태계 내의 다양한 행위자 및 비행위자에게 부여된 접근 및 권한에 대한 결정을 포함하여 설명함
- 이들은 플랫폼의 개방성은 공급자, 고객, 보완적 서비스 제공자, 제품 카테고리, 채널 등 다양한 측면에 대해 얼마나 개방적인지에 대한 일련의 결정에 의해 성립된다고 봄

**□ 공급자에 대한 개방성(Supplier Openness)**

- 플랫폼 외부의 공급자들이 플랫폼에 제품이나 서비스를 제공하거나 판매할 수 있는 정도. 다양한 독립 판매자나 공급자들이 자신의 제품이나 서비스를 플랫폼에 등록하고 판매할 수 있도록 하는 수준을 의미

**□ 고객에 대한 개방성(Customer Openness)**

- 고객이 플랫폼에서 제품이나 서비스를 검색, 비교, 구매할 수 있는 정도. 사용자들에게 다양한 상품 검색 기능이나 참여기제를 얼마나 제공하는가를 의미

**□ 보완적 서비스 제공자에 대한 개방성(Complementary Service Provider Openness)**

- 플랫폼의 핵심 제품이나 서비스를 보완하는 외부 서비스 제공자들의 플랫폼과의 상호 작용 정도. Google Play Store나 Apple App Store는 다양한 개발자들이 앱을 등록하고 판매할 수 있도록 하여 높은 보완적 서비스 제공자 개방성을 가짐

**□ 카테고리 개방성(Category Openness)**

- 플랫폼에서 제공되는 제품이나 서비스의 범위와 다양성을 의미. OTT 플랫폼의 경우 얼마나 다양한 장르와 카테고리의 영화와 TV 프로그램을 제공하는가에 따라 카테고리 개방성의 수준이 결정

채널 개방성(Channel Openness)

- 사용자가 플랫폼과 상호 작용하고 접근할 수 있는 유통 및 커뮤니케이션 채널의 범위를 의미. 웹사이트, 모바일 앱, 타사 웹사이트 및 앱을 통한 통합 등 다양한 채널을 통해 사용자와 상호 작용할 수 있는 높은 채널 개방성을 가짐

데이터 개방성(Data Openness)

- 플랫폼이 자신의 데이터에 대한 접근을 얼마나 허용하는지에 대한 정도. 이는 데이터셋 공유, API를 통한 데이터 접근, 데이터 기반 협업 등을 포함

## 나. 디지털플랫폼의 데이터 개방의 필요성과 효과

### (1) 데이터개방의 효과

혁신 및 협업 증진

- 디지털 플랫폼에서의 데이터 개방은 혁신을 크게 촉진. 외부 개발자와 연구자들에게 데이터에 대한 접근을 허용함으로써, 플랫폼은 다양한 관점과 전문 지식의 혜택을 받을 수 있으며, 이는 새로운 애플리케이션과 서비스의 개발로 이어질 수 있음(Chesbrough, 2003)

사용자 참여 및 서비스 개인화 향상

- 데이터 개방은 더 개인화되고 사용자 중심의 서비스로 이어질 수 있음. 플랫폼은 개방된 데이터를 사용하여 사용자의 선호도와 행동을 더 잘 이해하고, 사용자 경험을 향상시킬 수 있음
- 이는 순다라라잔이 디지털 플랫폼이 데이터를 활용하여 서비스를 개별적인 요구에 맞추는 방법에 대해 논의한 내용을 뒷받침(Sundararajan, 2016)

투명성 및 신뢰 구축

- 데이터 개방은 투명성에 기여하며, 이는 사용자들 사이의 신뢰를 구축하는 데 중요. 데이터를 공개적으로 공유하는 플랫폼은 종종 더 신뢰할 수 있는 것으로 인식됨(Kaplan, & Haenlein, 2010)

### □ 규제 준수 및 윤리적 데이터 관행

- 데이터 개인 정보 보호 및 윤리적 사용에 대한 우려가 증가함에 따라, 데이터 개방 관행은 플랫폼이 GDPR과 같은 규제 기준에 부합하도록 도움
- 이는 마틴이 디지털 플랫폼에서 데이터 관리의 윤리적 고려사항의 중요성을 강조한 작업에서 강조(Martin, 2016)

### □ 경제 성장 및 시장 확장

- 데이터 개방은 경제 성장과 시장 확장을 촉진할 수 있음. 가치 있는 데이터에 대한 접근을 제공함으로써, 플랫폼은 새로운 비즈니스 기회와 혁신을 제공함
- 이는 맥킨지 글로벌 연구소의 매니카 등이 빅 데이터에 대한 보고서에서 관련 내용을 논의한 바 있음(Manyika, Brown, Bughin, Dobbs, Roxburgh, & Byers, 2011)

## (2) 데이터개방의 유형

### (가) 데이터 접근 및 이동방식에 따른 분류

#### □ 접근 요청 없는 공개

- 이 방식은 플랫폼이 데이터를 누구나 자유롭게 접근하고 사용할 수 있도록 공개적으로 제공하는 것. 이는 주로 공공 데이터나 연구 데이터셋에 적용되며, 사용자가 별도의 접근 요청을 하지 않아도 데이터에 접근할 수 있음
- 데이터의 접근성을 높이고, 투명성을 증대시킴. 또한, 대중의 참여와 협력을 촉진하여, 데이터 기반의 혁신과 지식 공유를 장려

#### □ 동의에 의한 데이터 이동

- 사용자의 명시적 동의를 통해 데이터가 다른 플랫폼이나 서비스로 이동되는 방식. 이는 사용자 데이터의 소유권과 통제권을 중시하며, 사용자가 자신의 데이터를 어떻게 활용할지 결정할 수 있게 함
- 사용자의 데이터 관리에 대한 통제력을 증대시키며, 데이터의 개인화 및 맞춤형 사용 가능. 또한, 사용자의 프라이버시 보호와 데이터 보안을 강화

□ 열람만 가능한 데이터

- 사용자가 데이터를 열람만 할 수 있으며, 복제나 수정은 불가능. 이는 주로 저작권이나 기타 법적 제한이 있는 데이터에 적용됨
- 데이터의 무단 복제나 오용을 방지하며, 데이터의 원본성과 무결성 유지. 이는 특히 저작권이 보호되어야 하는 콘텐츠나 민감한 정보에 중요

□ 복제 가능한 데이터

- 사용자가 데이터를 복제하거나 다운로드할 수 있게 하는 방식. 이는 데이터의 활용도를 높이기 위해 제공되며, 사용자는 데이터를 자유롭게 활용할 수 있음
- 데이터의 활용도와 가치를 증대시키며, 연구, 교육, 혁신 등 다양한 목적으로 데이터 활용 가능. 또한, 데이터 기반의 의사결정과 분석을 촉진

(나) 데이터 이용 주체에 따른 분류

□ 이용사업자에 대한 데이터 접근성

- 디지털플랫폼의 데이터 접근은 특정 사용자나 개발자가 플랫폼의 데이터 자원에 접근할 수 있는 도구를 제공해서 데이터의 활용성을 높이는 방식을 의미
- 클라우드 기반의 서비스를 통해 디지털플랫폼이 구축하고 보유하고 있는 데이터는 원격으로 접근해서 활용할 수 있는 톨을 제공하는 방식
- 기술적인 측면에서는 API의 제공, 데이터셋 공유, 클라우드 서비스 등을 통해 이루어지며, 서비스적 측면에서는 사용자 인터페이스와 사용자 경험의 최적화를 포함(Hofman & Rajagopal, 2014)

□ 데이터와 리소스의 공개과 거래 플랫폼

- 데이터 공유는 플랫폼이 자신의 데이터와 리소스를 외부에 공공 데이터 형식으로 제공하는 것을 의미
- 이는 공동 창조와 협업을 가능하게 하며, 플랫폼의 확장성과 다양성을 강화할 도모(Soto Setzke, Böhm & Krcmar, 2019)

- 기술 및 서비스 제공 방식으로는 데이터 공유 포털, 오픈 소스 코드, 사용자 생성 콘텐츠의 통합 등을 통해 구현
- 공유된 데이터의 품질 관리와 보안이 중요한 요소(Saghiri et al., 2017)

#### (다) 데이터 개방방식

##### 클라우드 서비스의 형태로 데이터 제공

- 데이터를 클라우드 기반의 인프라에서 호스팅하고, 인터넷을 통해 접근할 수 있게 하는 방식. 이는 데이터의 저장, 관리, 처리를 클라우드 서비스 제공자가 담당
- 플랫폼은 물리적 인프라에 대한 투자와 관리 부담을 줄일 수 있으며, 사용자는 언제 어디서나 데이터에 접근할 수 있으며, 확장성과 유연성이 높아짐

##### API를 통한 데이터 제공

- API(Application Programming Interface)는 플랫폼의 데이터와 기능을 외부 애플리케이션과 연동할 수 있게 하는 인터페이스. 이를 통해 개발자들은 플랫폼의 데이터를 활용하여 새로운 애플리케이션을 개발할 수 있음
- 플랫폼의 기능 확장과 사용자 경험의 다양화에 기여

##### SDK(Software Development Kit)를 통한 데이터 제공

- 특정 플랫폼이나 애플리케이션 개발을 위해 필요한 도구, 라이브러리, 문서, 샘플 코드 등을 포함하는 소프트웨어 패키지.
- 플랫폼의 데이터와 기능을 더 쉽고 효율적으로 활용할 수 있게 하여, 애플리케이션 개발의 속도와 품질을 향상시킴

##### 기타

- 데이터 스트리밍(지도, 교통정보, 주식 시장, 소셜미디어 피드 정보 등)
- 블록체인 기반 데이터 공유

## 다. 이용사업자 대한 데이터 접근성 사례

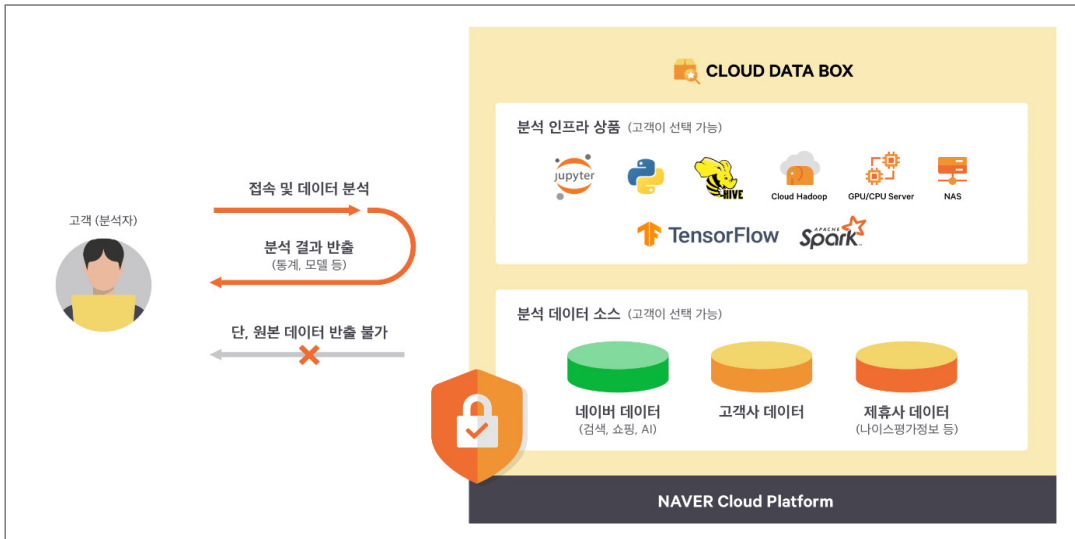
### (1) 네이버

#### (가) 클라우드 데이터 박스(Cloud Data Box)

##### 서비스 개요

- 네이버 클라우드 플랫폼의 클라우드 데이터 박스는 검색, 쇼핑과 관련된 네이버 제공 빅데이터와 AI 데이터 등을 제공하며, 대량의 온라인 행동 데이터를 분석할 수 있는 분석 플랫폼을 함께 제공

[그림 III-5] 클라우드 데이터 박스 서비스 개념도



출처: 네이버 클라우드 플랫폼(<https://www.ncloud.com/product/analytics/cloudDataBox>)

##### 데이터 유형

- 클라우드 데이터 박스에서 제공하는 분석 인프라 상품은 개별 상품 정책에 따라 판매자에게 (유료로) 제공하고 있으며, 서비스 용도와 데이터 용량(분석 인프라/패키지에 따라 가격 옵션 설정)에 따라 다음의 데이터 유형을 이용 가능
  - 네이버 데이터: 검색, 쇼핑, AI학습용 데이터 등
  - 고객사 데이터
  - 제휴사 데이터: NICE 평가정보 결합 등

□ 데이터 접근 및 이용 방식

- 데이터 공급 신청 이후에 외부망과의 통신을 차단하며 데이터 유출 위험없이 안전하게 분석 가능
- 제공받은 네이버 데이터나 클라우드 데이터 박스에서 가공된 데이터는 고객이 재판매하거나 임의로 외부에 반출 및 공개할 수 없음
- 또한 클라우드 데이터 박스에서 제공하는 데이터는 개인 재식별을 위해 이용될 수 없음
- 단, 제공된 데이터는 외부로 반출할 수 없지만, 클라우드 데이터 박스에서 만든 어넌 모델 및 집계 데이터 등의 분석 결과물은 심사를 거쳐 외부로 반출 가능함

〈표 III-4〉 이용 가능한 네이버 기본 데이터

구분		상세 데이터													
		기준 일자	성별	연령대	지역	기기	시간대	검색어 /카테고리	클릭 영역 / 키워드	검색어 /카테고리	클릭 영역 / 키워드	브랜드명	상품명	클릭 사용자 수	업데이트 주기
검색	사용자 속성별 검색/클릭 내역	일자 (일)	성별 (남/여)	연령대 (5세 단위)	지역 (거주지)	접속 기기 (PC/ 모바일)	시간대 (3시간 단위)	검색어	클릭 영역	-	-	-	-	클릭한 사용자 수	연간 업데이트
	검색어 코어커런스 매트릭스	일자 (주)						검색어 1	클릭 영역1	검색어 2	클릭 영역2	-	-	함께 클릭한 사용자 수	
	접속지별 검색/클릭 내역	일자 (일)			지역 (접속지)			검색어	-	-	-	-	-	클릭한 사용자 수	
쇼핑	사용자 속성별 상품 클릭 내역	일자 (일)	성별 (남/여)	연령대 (5세 단위)	지역 (거주지)	접속 기기 (PC/ 모바일)	시간대 (3시간 단위)	유입 키워드	-	-	-	브랜드명	상품명	클릭한 사용자 수	연간 업데이트
	사용자 속성별 상품 구매 내역	일자 (일)							상품 카테고리	-	-			-	
	상품 클릭 코어커런스 매트릭스	일자 (주)			지역 (거주지)			유입 키워드 1	상품 카테고리 2	유입 키워드 2	-	-	함께 클릭한 사용자 수		
	상품 구매 코어커런스 매트릭스	일자 (주)						상품 카테고리 1	-	-	-	-	클릭한 사용자 수		
AI 학습용	NLP 실험용 텍스트 데이터	뉴스 기사에 나온 개체명(이름) 검색 후 해당 개체명과 연관된 위키피디아 페이지(ID)에 연결한 데이터													
	어노테이션된 이미지 데이터	재활용 이미지 수집 및 레이블링													
		이미지에서 음식 영역을 바운딩 박스로 태깅한 데이터 표지판, 입간판, 메뉴판, 음식점 간판 등 이미지에 포함된 한글 OCR 어노테이션 데이터(영문, 숫자도 포함)													

※ 조회 시점 기준, 기본적으로 전년도 상반기 데이터를 이용할 수 있으며, 만약 Insight Option(12개월 이용 약정) 이용 시 전전년 1월 ~ 당해 년도 전월까지의 데이터를 이용할 수 있음

출처: 네이버 클라우드 플랫폼(<https://www.ncloud.com/product/analytics/cloudDataBox>)

□ 클라우드 데이터 박스를 통한 분석 활용

- 검색어 분석 : 브랜드 평판 관리, 연령대별 검색어를 분석하여 최적의 브랜드 콘텐츠를 제작할 수 있고 검색어 트렌드를 분석하여 고객 세그먼트별 니즈를 파악에 도움
- 고객 맞춤 서비스 제공 : 고객에 최적화된 랜딩 페이지를 구성하고, 함께 구매한 상품 추천을 통해 신규 구매율을 상승시키거나 기존 고객의 정보를 활용하여 고객별 최적화된 맞춤형 상품 추천 가능  
세부 시장 분석(자사 브랜드 핵심고객층 파악 및 세부 시장 점유율 분석), 사용자 퍼널 분석(구매 여정별 사용자 규모 및 검색어 터치 포인트 분석), 광고효과 분석(온/오프라인 광고 캠페인의 구매 효과 분석), 브랜드 이미지 분석(출시 이후 사용자의 브랜드 관련 검색어(prefix/suffix) 변화 패턴 분석) 등
- AI 분석 : 라벨링된 이미지 및 텍스트를 이용해 다양한 분류 모형을 제작하고 해당 모형을 통한 서비스 효율성 제고

(나) 데이터 애널리틱스 서비스(Data Analytics Service)

□ 서비스 개요

- 데이터 애널리틱스 서비스는 클라우드 환경에서 손쉽게 비즈니스 인텔리전스 환경을 이용할 수 있도록 도와주는 네이버 클라우드 플랫폼의 서비스 중 하나로 네이버 애널리틱스와 연동하여 사이트 방문자의 활동 데이터를 받아 볼 수 있으며, 네이버 쇼핑의 쇼핑 데이터를 받아 볼 수 있음

□ 데이터 유형

- 데이터 애널리틱스 서비스에서 제공하는 데이터는 사이트 데이터와 쇼핑 데이터가 있으며, 패키지 상품 정책에 따라 판매자에게 유료로 제공
  - 사이트 데이터 : 네이버 애널리틱스에서 수집하는 사이트 방문 로그 및 네이버 통합검색 데이터
    - 사이트 방문 로그 : 사이트 방문자의 인입 로그
    - 사이트 유입 검색어 : 사이트 유입 검색어에 대한 네이버 통합검색 수치 지표

- 방문자 검색어 : 사이트 유입 검색어를 입력하기 전/후 일정 시간 동안 사이트 방문자가 네이버에서 검색한 검색어에 대한 수치 지표
- 쇼핑 데이터 : 네이버 쇼핑에서 발생한 검색어에 대한 상품 클릭 데이터로, 상품 카테고리 및 연령대별/성별 단위로 집계되어 제공

#### □ 데이터 접근 및 이용 방식

- 데이터 파일 확인 방법은 Superset이나 Object Storage 버킷 등을 통해 확인할 수 있음
  - Superset에서 확인: Data Analytics Service에서 제공하는 Superset 대시보드에서 데이터 확인
  - Object Storage 버킷에서 확인: S3 브라우저의 Buckets > Add External Bucket 메뉴에서 버킷 이름을 입력하여 버킷을 추가한 후 데이터 파일 확인
- 제공받은 웹 로그 데이터는 네이버 클라우드 플랫폼 외부로 유출할 수 없으며, 데이터 애널리틱스 서비스에서 받은 데이터를 이용하여 특정 인물의 개인정보를 파악하거나 추정하는 행위는 금지됨
- 웹 사이트에 방문한 고객의 방문자 로그는 반드시 사이트 관리자를 인증한 후에만 제공되며, 이는 패키지에 따라 가격 옵션 설정한 이후 해당 생성 권한을 통해 데이터에 접근 가능함

〈표 III-5〉 네이버 데이터 애널리틱스 서비스 제공 데이터

구분	상세 데이터
방문자 행동 로그	내 사이트 방문자가 언제, 어디서, 어떤 검색어로 유입되고 있는지 파악 가능 방문자 행동 패턴을 분석 가능 <ul style="list-style-type: none"> <li>• 사이트 ID</li> <li>• 로그 발생시각</li> <li>• 방문 페이지 제목</li> <li>• 유입검색어</li> <li>• 국가</li> <li>• 인코딩 방식</li> <li>• 커스텀 이벤트 코드</li> <li>• 전환 분류 코드</li> <li>• 사이트 이용자 식별자</li> <li>• 방문 페이지 주소</li> <li>• 방문 페이지 레퍼러</li> <li>• 이용한 브라우저</li> <li>• 지역</li> <li>• 이용 언어</li> <li>• 커스텀 이벤트 분류코드</li> <li>• 전환 부가 값</li> </ul>
유입 키워드의 네이버 검색 통계	내 사이트로 유입되는 검색어가 네이버에서 얼마나 검색되고 있는지 절댓값으로 파악 가능 전체 트렌드를 기준으로 내 사이트의 포지션을 비교 및 분석 가능 <ul style="list-style-type: none"> <li>• 사이트 ID</li> <li>• 유입검색어</li> <li>• 성별</li> <li>• 검색수</li> <li>• 검색일자</li> <li>• PC·모바일구분</li> <li>• 연령대</li> </ul>
방문자 검색 활동 통계	내 사이트 방문자가 사이트에 방문하기 전과 후에 검색한 키워드의 통계 확인 가능 내 사이트 방문자의 관심사를 폭넓게 분석 가능 <ul style="list-style-type: none"> <li>• 사이트 ID</li> <li>• 검색 키워드</li> <li>• 검색 일자</li> <li>• 검색수</li> </ul>
쇼핑 인사이트 통계	네이버 쇼핑 카테고리별로 인입된 검색 키워드를 클릭한 횟수 확인 가능 특정 쇼핑 카테고리에 대한 관심사와 관심도 분석 가능 <ul style="list-style-type: none"> <li>• 카테고리 ID</li> <li>• 키워드</li> <li>• 연령대</li> <li>• 검색 일자</li> <li>• 성별</li> <li>• 클릭 수</li> </ul>

※ 방문자를 중복 식별하는 것을 방지하기 위해 '사이트 이용자 식별자'는 추가 익명화 처리를 거치며 이에 같은 사이트 내에서 구분된 식별자는 유일한 값이나 사이트가 달라지면 유일성 보장이 어려움

출처: 네이버 클라우드 플랫폼(<https://www.ncloud.com/product/analytics/dataAnalyticsService>)

데이터 애널리틱스 서비스를 통한 분석 활용

- 시장과 고객을 이해하기 위한 다양한 데이터 제공 : 웹 사이트의 유입 및 검색 통계에 대한 네이버 데이터를 제공, 폭넓고 정확한 타겟 시장 및 고객 분석 가능
- 쉽고 편리한 데이터 분석 환경 제공 : 매일 자동 저장되는 로그를 데이터 시각화 도구 Superset을 이용하여 쉽고 편리하게 분석
- 추천 API 제공 : 사이트 유입 통계 기반으로 분석된 인기아이템 추천 점수 제공
- 안전한 고객 정보 관리 : 엄격한 보안 인증을 받은 환경에서 데이터 보관

**(다) 네이버 애널리틱스(Naver Analytics)****□ 서비스 개요**

- 네이버 애널리틱스는 방문자의 사이트 이용 행태와 유입 정보, 인기페이지 정보 등 사이트 운영 및 마케팅 활동 개선, 온라인 비즈니스를 운영하는데 도움이 되는 분석 자료를 제공함

**□ 데이터 유형**

- 웹사이트에 대한 방문 분석, 유입 분석, 페이지 분석, 실시간 분석, 사용자 분석 등 주요 데이터 유형을 무료로 제공
  - 사이트 현황 : 방문자수, 유입정보, 인기페이지 등 가장 기본이 되는 지표를 요약하여 제공
  - 방문 분석 : 순방문자 수, 신규/재방문자 등 방문현황(UV), 페이지뷰(PV), 시간대/요일별 방문분포, 방문체류시간, 방문경로깊이, 방문지역분포 등 사이트 방문 행태에 대한 종합적인 정보를 제공
  - 유입 분석 : 검색유입현황, 유입검색어, 유입상세URL 등 방문자 유입 경로 분석
  - 페이지 분석 : 인기 페이지, 시작 페이지, 종료페이지, 반송 페이지 등 웹사이트의 개별 페이지에 대한 이용정보 제공
  - 실시간 분석 : 최근 30분 내의 방문자 수, 페이지뷰, 검색유입, 네이버 검색광고 전환 등 주요 지표를 거의 실시간으로 업데이트
  - 사용자 분석 : 방문자의 사용하는 PC/모바일 기기의 운영체제 및 웹브라우저, 방문자가 사용하는 기기의 해상도(화면크기) 등 방문자를 이해하기 위한 다양한 정보 제공
  - 이벤트 : 이벤트 수집 스크립트를 별도로 적용해 특정 웹페이지 안에서 일어나는 방문자의 행동을 추적할 수 있도록 하는 기능
- 네이버 검색광고의 프리미엄로그분석 서비스를 이용중인 경우 실시간분석 보고서에서 최근 30분 동안 네이버 검색광고로 발생한 전환을 실시간으로 보여주는 네이버 검색광고 전환 리포트 제공

- 프리미엄 로그 분석은 네이버 검색광고에서 제공하는 무료로 제공하는 자동 추적(auto tracking) 기능으로 네이버 검색광고의 광고별 체류 시간, PV, 검색 광고 전환 분석 보고서, 사이트 전체적인 유입, 방문, 페이지 분석 보고서 등의 웹 로그 분석 보고서를 제공하는 무료 서비스임
- 프리미엄 로그분석 서비스를 이용 중인 사이트에서 네이버 검색광고를 통한 전환이 발생하면 (즉, 네이버 검색광고로 유입된 방문이 전환 페이지로 연결되면) 위 리포트에 유형별 전환수가 표시됨

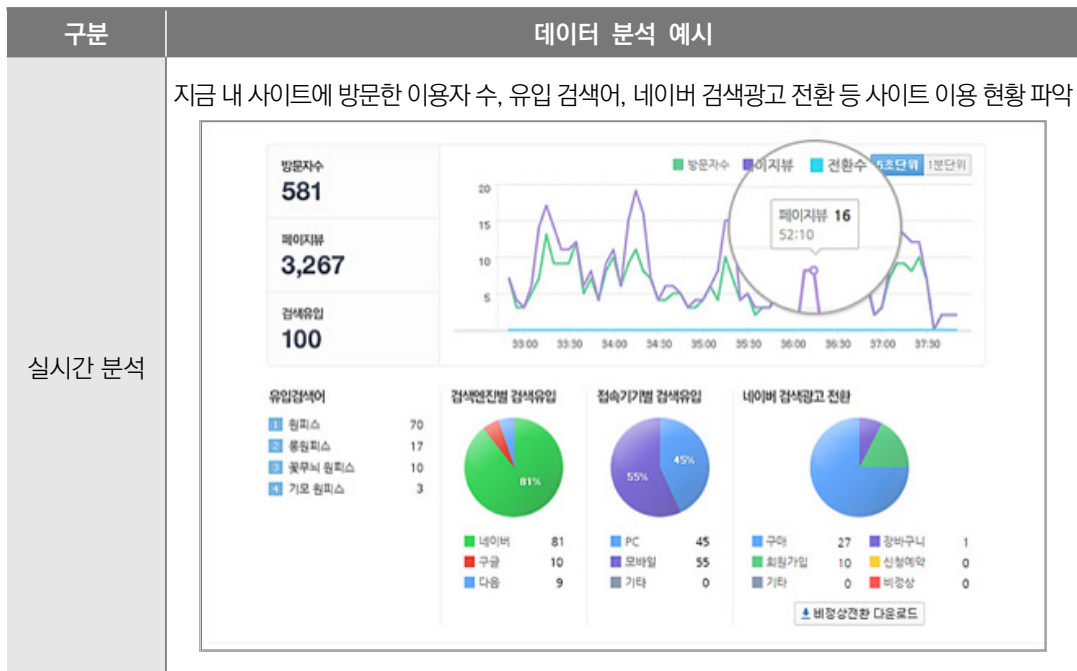
□ 데이터 접근 및 이용 방식

- 네이버 애널리틱스는 네이버에서 제공하는 무료 웹로그 분석 서비스로, 네이버 계정 또는 네이버 검색광고 계정으로 로그인하여 본인이 소유하거나 관리 권한이 있는 사이트만 등록 가능
- 애널리틱스 홈에서 이용약관에 동의한 후 사이트 등록을 완료하면 분석 스크립트가 발급되며, 발급받은 분석 스크립트(Javascript)를 원하는 사이트에 설치함
  - 일반 웹사이트 : 웹사이트의 html 소스에 설치, 사이트 이용시 항상 실행되는 HTML 코드(예: 하단 회사정보 등)를 별도의 파일로 만들고(A), 이 파일에 NA 스크립트를 심고, 이용자에게 보여줄 페이지를 구성할 때, 위의 언제나 열리는 HTML 코드(A)의 파일을 include 라는 방식을 이용하여 페이지에 포함
  - modoo! 사이트 : 네이버 애널리틱스를 이용한 웹로그 분석 서비스를 기본으로 제공함으로 별도 가입없이 분석 보고서 확인 가능(modoo! 서비스 > 홈페이지 관리 > 방문자통계를 클릭)
  - 블로그, 카페 등 : 스크립트 편집이 가능한 설치형 블로그(티스토리 등)는 서비스 이용이 가능하며, 스크립트 편집이 불가능한 네이버 블로그와 네이버 카페 등은 네이버 애널리틱스 서비스를 이용한 웹로그 분석이 불가능
- 타사의 애널리틱스를 사용하더라도 네이버 애널리틱스와 동시에 사용 가능하며, 태그 앞에 분석 스크립트 소스를 입력하면 이용 가능. 참고로 태그가 없거나 확인이 어려운 경우 페이지 소스 하단에 분석 스크립트를 삽입하여 이용 가능

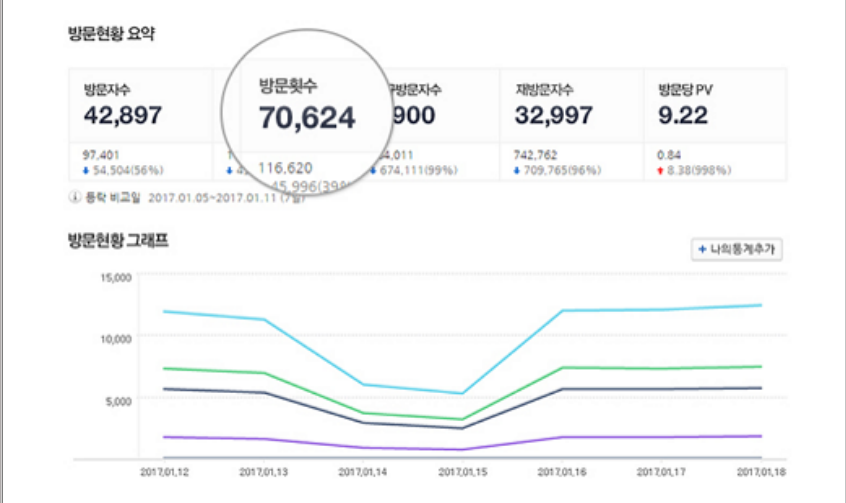

□ 네이버 애널리틱스를 통한 분석 활용

- 실시간 분석을 통한 마케팅 활동의 시작 및 효과 실시간 측정 및 개선
- 유입 분석을 통한 유입되는 검색어, 서비스에 집중하여 보다 효과적으로 사이트 방문 유도 가능
- 페이지 분석 정보를 통해 인기가 많은 콘텐츠는 강화하고, 그렇지 않은 콘텐츠는 보완하여 전체적인 사이트 품질 향상
- 중요한 고객유형(신규/재방문)이 잘 방문하는지, 방문이 집중되는 시간대는 언제인지 등을 참고하여 사이트 운영방식 개선
- 권한부여 기능을 통해 사이트 현황을 함께 분석하고, 개선점을 찾을 수 있음

〈표 III-6〉 네이버 애널리틱스 데이터 분석 자료



구분	데이터 분석 예시																																																							
<p>유입 분석</p>	<p>방문자들이 어떤 검색어를 통해 내 사이트에 방문하는지, 주로 사용하는 검색엔진은 무엇인지 파악</p> <div data-bbox="397 338 1243 870"> <h3>검색유입 요약</h3> <table border="1"> <tr> <td>검색유입수 <b>3,782</b> 3,840 ↓ 59(2%)</td> <td>전체유입수 <b>4,230</b></td> <td>검색유입율 <b>89.4%</b></td> <td>TOP 검색엔진 <b>네이버 93.07%</b> 네이버 92.66% ↓ 0.41%</td> </tr> </table> <p>① 동학 비고빌 2016.12.21</p> <h3>검색엔진 그래프</h3> <table border="1"> <caption>네이버 검색어 3,137 개</caption> <tr><td>1 원피스</td><td>1154</td><td>37.2%</td></tr> <tr><td>2 겨울 원피스</td><td>684</td><td>22.1%</td></tr> <tr><td>3 기모 원피스</td><td>489</td><td>15.8%</td></tr> <tr><td>4 롱원피스</td><td>251</td><td>8.1%</td></tr> <tr><td>5 꽃무늬 원피스</td><td>124</td><td>4.0%</td></tr> <tr><td>6 주름 원피스</td><td>101</td><td>3.3%</td></tr> <tr><td>7 니트 원피스</td><td>97</td><td>2.0%</td></tr> <tr><td>8 리본 원피스</td><td>62</td><td>1.5%</td></tr> </table> </div>	검색유입수 <b>3,782</b> 3,840 ↓ 59(2%)	전체유입수 <b>4,230</b>	검색유입율 <b>89.4%</b>	TOP 검색엔진 <b>네이버 93.07%</b> 네이버 92.66% ↓ 0.41%	1 원피스	1154	37.2%	2 겨울 원피스	684	22.1%	3 기모 원피스	489	15.8%	4 롱원피스	251	8.1%	5 꽃무늬 원피스	124	4.0%	6 주름 원피스	101	3.3%	7 니트 원피스	97	2.0%	8 리본 원피스	62	1.5%																											
검색유입수 <b>3,782</b> 3,840 ↓ 59(2%)	전체유입수 <b>4,230</b>	검색유입율 <b>89.4%</b>	TOP 검색엔진 <b>네이버 93.07%</b> 네이버 92.66% ↓ 0.41%																																																					
1 원피스	1154	37.2%																																																						
2 겨울 원피스	684	22.1%																																																						
3 기모 원피스	489	15.8%																																																						
4 롱원피스	251	8.1%																																																						
5 꽃무늬 원피스	124	4.0%																																																						
6 주름 원피스	101	3.3%																																																						
7 니트 원피스	97	2.0%																																																						
8 리본 원피스	62	1.5%																																																						
<p>페이지 분석</p>	<p>내 사이트에서 가장 인기 있는 페이지는 어디인지, 각 페이지에 방문자가 머무르는 시간은 얼마나 되는지 등 확인</p> <div data-bbox="397 991 1243 1528"> <h3>인기페이지</h3> <p>* 페이지뷰합계 508,896</p> <table border="1"> <thead> <tr> <th>순위</th> <th>페이지URL</th> <th>페이지뷰</th> <th>비율</th> <th>평균체류시간</th> </tr> </thead> <tbody> <tr><td>1</td><td>http://analytics.naver.com/1</td><td>77,053</td><td>15%</td><td>27초</td></tr> <tr><td>2</td><td>http://analytics.naver.com/2</td><td>63,757</td><td>13%</td><td>02분 42초</td></tr> <tr><td>3</td><td>http://analytics.naver.com/3</td><td></td><td>6%</td><td>11초</td></tr> <tr><td>4</td><td>http://analytics.naver.com/4</td><td>25,214</td><td>5%</td><td>34초</td></tr> <tr><td>5</td><td>http://analytics.naver.com/5</td><td>25,050</td><td>5%</td><td>11초</td></tr> <tr><td>6</td><td>http://analytics.naver.com/6</td><td>24,741</td><td>5%</td><td>05초</td></tr> <tr><td>7</td><td>http://analytics.naver.com/7</td><td>24,293</td><td>5%</td><td>09초</td></tr> <tr><td>8</td><td>http://analytics.naver.com/8</td><td>20,415</td><td>4%</td><td>28초</td></tr> <tr><td>9</td><td>http://analytics.naver.com/9</td><td>19,612</td><td>4%</td><td>03초</td></tr> <tr><td>10</td><td>http://analytics.naver.com/10</td><td>18,423</td><td>4%</td><td>06초</td></tr> </tbody> </table> </div>	순위	페이지URL	페이지뷰	비율	평균체류시간	1	http://analytics.naver.com/1	77,053	15%	27초	2	http://analytics.naver.com/2	63,757	13%	02분 42초	3	http://analytics.naver.com/3		6%	11초	4	http://analytics.naver.com/4	25,214	5%	34초	5	http://analytics.naver.com/5	25,050	5%	11초	6	http://analytics.naver.com/6	24,741	5%	05초	7	http://analytics.naver.com/7	24,293	5%	09초	8	http://analytics.naver.com/8	20,415	4%	28초	9	http://analytics.naver.com/9	19,612	4%	03초	10	http://analytics.naver.com/10	18,423	4%	06초
순위	페이지URL	페이지뷰	비율	평균체류시간																																																				
1	http://analytics.naver.com/1	77,053	15%	27초																																																				
2	http://analytics.naver.com/2	63,757	13%	02분 42초																																																				
3	http://analytics.naver.com/3		6%	11초																																																				
4	http://analytics.naver.com/4	25,214	5%	34초																																																				
5	http://analytics.naver.com/5	25,050	5%	11초																																																				
6	http://analytics.naver.com/6	24,741	5%	05초																																																				
7	http://analytics.naver.com/7	24,293	5%	09초																																																				
8	http://analytics.naver.com/8	20,415	4%	28초																																																				
9	http://analytics.naver.com/9	19,612	4%	03초																																																				
10	http://analytics.naver.com/10	18,423	4%	06초																																																				

구분	데이터 분석 예시																																				
방문 분석	<p>방문현황(UV), 신규/재방문자 수, 시간대별 방문 분포, 방문지역 등 방문자의 방문 특성을 이해하기 위한 종합적인 정보 제공</p>  <p><b>방문현황 요약</b></p> <table border="1"> <tr> <td>방문자수</td> <td>방문횟수</td> <td>방문자수</td> <td>재방문자수</td> <td>방문당 PV</td> </tr> <tr> <td>42,897</td> <td>70,624</td> <td>900</td> <td>32,997</td> <td>9.22</td> </tr> <tr> <td>97,401 ↓ 54,504(56%)</td> <td>116,620 ↓ 45,996(39%)</td> <td>4,011 ↓ 674,111(99%)</td> <td>742,762 ↓ 709,765(96%)</td> <td>0.84 ↑ 8.38(998%)</td> </tr> </table> <p>① 통학 비교일 2017.01.05-2017.01.11 17:00</p> <p><b>방문현황 그래프</b></p> <p>방문현황 그래프는 2017.01.12부터 2017.01.18까지의 방문자 수, 방문 횟수, 재방문자 수, 방문당 PV를 보여주는 선 그래프이다. 방문 횟수(파란색)가 가장 높고, 방문당 PV(보라색)가 가장 낮다.</p>	방문자수	방문횟수	방문자수	재방문자수	방문당 PV	42,897	70,624	900	32,997	9.22	97,401 ↓ 54,504(56%)	116,620 ↓ 45,996(39%)	4,011 ↓ 674,111(99%)	742,762 ↓ 709,765(96%)	0.84 ↑ 8.38(998%)																					
방문자수	방문횟수	방문자수	재방문자수	방문당 PV																																	
42,897	70,624	900	32,997	9.22																																	
97,401 ↓ 54,504(56%)	116,620 ↓ 45,996(39%)	4,011 ↓ 674,111(99%)	742,762 ↓ 709,765(96%)	0.84 ↑ 8.38(998%)																																	
권한 부여	<p>네이버 애널리틱스 보고서를 다른 사람과 공유하도록 권한부여 기능 제공</p>  <p><b>내가 권한을 준 사이트</b></p> <table border="1"> <thead> <tr> <th>아이디</th> <th>대상 사이트</th> <th>연동서비스</th> <th>관리</th> </tr> </thead> <tbody> <tr> <td>naver_01</td> <td>000사이트 (http://www.naver01.com)</td> <td>검색광고</td> <td>-</td> </tr> <tr> <td>naver_02</td> <td>000사이트 (http://www.naver02.com)</td> <td>검색광고</td> <td>권한회수</td> </tr> <tr> <td>naver_03</td> <td>000사이트 (http://naver03.modoo.at)</td> <td>Modool</td> <td>-</td> </tr> </tbody> </table> <p><b>내가 권한을 받은 사이트</b></p> <table border="1"> <thead> <tr> <th>아이디</th> <th>대상 사이트</th> <th>연동서비스</th> <th>관리</th> <th>통계보기</th> </tr> </thead> <tbody> <tr> <td>naver_04</td> <td>000사이트 (http://www.naver04.com)</td> <td>검색광고</td> <td>-</td> <td>보기</td> </tr> <tr> <td>naver_05</td> <td>000사이트 (http://www.naver05.com)</td> <td>Naver Analytics</td> <td>권한회수</td> <td>보기</td> </tr> <tr> <td>naver_06</td> <td>000사이트 (http://naver06.modoo.at)</td> <td>Modool</td> <td>-</td> <td>보기</td> </tr> </tbody> </table>	아이디	대상 사이트	연동서비스	관리	naver_01	000사이트 (http://www.naver01.com)	검색광고	-	naver_02	000사이트 (http://www.naver02.com)	검색광고	권한회수	naver_03	000사이트 (http://naver03.modoo.at)	Modool	-	아이디	대상 사이트	연동서비스	관리	통계보기	naver_04	000사이트 (http://www.naver04.com)	검색광고	-	보기	naver_05	000사이트 (http://www.naver05.com)	Naver Analytics	권한회수	보기	naver_06	000사이트 (http://naver06.modoo.at)	Modool	-	보기
아이디	대상 사이트	연동서비스	관리																																		
naver_01	000사이트 (http://www.naver01.com)	검색광고	-																																		
naver_02	000사이트 (http://www.naver02.com)	검색광고	권한회수																																		
naver_03	000사이트 (http://naver03.modoo.at)	Modool	-																																		
아이디	대상 사이트	연동서비스	관리	통계보기																																	
naver_04	000사이트 (http://www.naver04.com)	검색광고	-	보기																																	
naver_05	000사이트 (http://www.naver05.com)	Naver Analytics	권한회수	보기																																	
naver_06	000사이트 (http://naver06.modoo.at)	Modool	-	보기																																	

출처: 네이버 애널리틱스 고객센터(<https://help.naver.com/service/9864/contents/15298?osType=PC&lang=ko>)

(라) 비즈 어드바이저(Biz Advisor)

□ 서비스 개요

- 비즈 어드바이저는 온라인 비즈니스 관리에 필요한 통계 데이터를 다각도로 분석하여 보여주는 애널리틱스 도구로, 현재 네이버 스마트스토어에 한정하여 분석 서비스를 제공함

□ 데이터 유형

- 판매자 계정에 연결된 모든 사이트(스마트스토어, 개별 윈도) 데이터를 개별적으로 확인 가능하며 판매자 계정 단위로 제공됨
- 요약보고서, 판매분석 보고서, 마케팅분석 보고서, 쇼핑행동분석 보고서, 시장분석 보고서, 인사이트(판매성과예측 보고서) 등을 제공함
  - 요약보고서 : 어제의 주요 성과, 판매된(결제된)금액 관련, 결제취소/환불 관련 정보, 마케팅 채널 관련 정보, 상품별 결제취소/환불정보, 마케팅채널과의 관계, 구매한 사람들의 성별/나이 정보 등 전자상거래사업 분석에 필요한 핵심 지표 제공
  - 판매분석 보고서 : 일자별/상품별/고객별/유입경로별로 판매 성과 분석 제공
  - 마케팅분석 보고서 : 해당 마케팅 채널별 사이트에 유입/클릭 수, 방문수, 광고비, 유입 직후 당일 내 페이지 조회 수, 유입당 페이지 수, 상품ID별 결제 수, 입입당 당일 결제율, 유입당 결제금액, ROAS(마케팅채널에 사용한 비용 대비 결제 금액으로 비용 효율 %) 등
  - 쇼핑행동분석 보고서 : 상품별(상품 상세 조회수, 결제 금액, 결제 상품 수량, 상세조회 대비 결제율), 페이지별(각 페이지 URL 별, 페이지 수, 체류시간, 평균 체류시간 등) 정보 제공
  - 시장분석(벤치마크) 보고서 : 유입수, 결제수, 유입당 결제율, 결제금액, 유입당 결제금액 등 내 사이트의 지표와 다른 사이트그룹의 지표를 비교하여 내 사이트의 각 지표별 위치, 효율 등 분석 제공
  - 인사이트(판매성과예측 보고서) : 과거 90일의 결제금액 데이터를 이용하여, 오늘을 포함한 미래 일주일(7일)의 결제금액 예측을 차트로 보여주는 보고서

□ 데이터 접근 및 이용 방식

- 보고서 접속 시 우측 상단에 다운로드 버튼이 있으며, 엑셀 파일로 다운 가능함
- 엑셀 파일 다운로드가 미제공되는 보고서는 [통계 > 요약 > 전자상거래요약], [통계 > 요약 > 오늘보고서], [통계 > 시장분석 > 벤치마크] 등임

- 비즈 어드바이저 통계에서는 보고서상 제공되는 데이터 외에 추가적인 데이터의 제공, 데이터 현황 확인 등은 제공하지 않으며, 보고서는 일단위로 제공(통상적으로 로그가 발생한 다음날 8시 이후 제공)
- 기존 네이버 애널리틱스에서 사용하던 각 메뉴의 보고서의 내용은 비즈 어드바이저 각 메뉴의 보고서에서 확인 가능하며(그림 III-6 참조), 비즈 어드바이저가 네이버 애널리틱스보다 더 고도화된 서비스를 제공함

[그림 III-6] 네이버 애널리틱스와 스마트스토어 통계(Biz Advisor)의 연계 관계



출처: 네이버 고객센터(<https://help.naver.com/service/19459/contents/20848?lang=ko>)

□ 비즈 어드바이저를 통한 분석 활용

- 판매분석을 통한 일자별, 상품별, 결제 고객별, 유입 경로별 판매성과 분석
- 마케팅분석을 통한 채널별, 키워드별로 유입수, 유입고객, 결제율, 기여금액을 확인하여 마케팅채널의 효율성 파악 가능
- 쇼핑행동분석을 통한 고객의 사이트 내 행동을 분석하여 제공하며 상품별 상세 조회수와 결제를 연결하여 확인 가능

(2) 배달의 민족의 데이터 접근성 사례

(가) 가게 관리 프로그램 : 배민셀프서비스

□ 서비스 개요

- 배달의민족에 입점한 외식업 소상공인의 가게 운영 편의성 증진을 위해 제공하는 프로그램으로, B2B 고객들이 가게관리부터 사업자 정보까지 본인과 관련된 정보를 확인 및 수정 서비스를 제공

□ 데이터 유형

- 소상공인의 효과적인 음식점 운영 전략 수립을 돕기 위하여 주문 고객의 연령대·성별·지역과 같은 통계 정보는 물론 요일·시간대별 주문 추이 등 분석 정보를 무료 제공
  - 광고 통계 : 업주가 이용 중인 배달의민족 광고 상품별로 노출수, 클릭수, 주문수, 주문금액 등 통계 제공(일/주/월별)
  - 가게 운영 : 가장 주문이 많이 발생한 배달지역(동 기준)과 인기 메뉴(주문수 기준), 배달리뷰 등 확인 가능
  - 리뷰 : 최근 리뷰수(6개월) 및 별점 통계 확인, 기간별 리뷰 조회 기능(전체 리뷰, 미답변 리뷰, 차단 리뷰)
  - 배달데이터 분석 CRM(Customer Relationship Management) 솔루션 : ‘고객 관리’ 기능을 도입하여(23. 11월 신규 기능 도입) 주문 고객 수, 신규·재주문 고객 수, 연령·성별, 인기메뉴, 지역 등의 데이터 확인 가능(최근7일, 30일, 주간, 월간 기준)

□ 데이터 접근 및 이용 방식

- 배달의 민족에서 제공하는 별도의 플랫폼을 통해 관련 프로그램 및 데이터에 접근 가능함
- 개별 고객의 이름, 나이, 성별 등 개인정보 및 프라이버시 보호를 위해 특정 고객을 식별할 수 있는 정보는 제공하고 있지 않음
  - 단, 특정 기간 동안 가게에 주문한 고객들의 성별 및 연령대(10대/20대/30대/40대/50대~) 정보를 제공하고 있음
- 고객의 주소 정보의 경우 울트라콜, 오픈리스트 광고상품을 이용하는 점주에게 제공하며, 고객 연락처의 경우 실제 전화번호가 아닌 가상번호 형태로 제공(주문한 순간부터 3시간까지만 유효)
  - 배달의민족이 배달까지 제공하는 배민1 서비스의 경우, 사장님에게 고객 주소 정보가 제공되지 않음(배달원에게만 제공)

〈표 III-7〉 배민셀프서비스 제공 데이터

구분	상세 데이터
광고통계	노출수 / 클릭수 / 주문수
가게운영	인기배달지역 / 인기 메뉴 / 배달리뷰
리뷰관리	최근 리뷰수(6개월) / 별점 통계 / 기간별 리뷰
CMS 고객정보	성별 : 가게를 방문한 고객의 남성·여성 고객 수 연령 : 가게를 방문한 남성·여성의 연령대별 고객 수 신규 고객 : 최근 1년간 주문 이력이 없는 고객 재주문 고객 : 최근 1년간 1번 이상 주문 이력이 있는 고객 주문 : 고객이 주문한 총 주문 건수 인기메뉴 : 가장 많은 고객이 주문한 메뉴 인기시간대 : 요일과 시간대 별 총 주문 고객 수 지역 : 주문이 많이 발생한 지역을 5순위(법정동 기준), 신규 주문이 가장 많이 발생한 지역과 주문한 메뉴, 재주문이 가장 많이 들어오는 지역과 메뉴, 지역별로 많이 팔리는 인기메뉴 순위, 지역에서 가장 많이 주문한 고객 집단

출처: 우아한형제들

□ 배민셀프서비스를 통한 분석 활용

- 주문 고객 수, 신규·재주문 고객 수, 고객 연령대·성별, 인기메뉴, 지역 등 데이터, 주문이 높은 요일과 시간 등 고객관리데이터 제공으로 맞춤형 마케팅 지원
  - 인기 메뉴 정보를 통한 특정 고객 및 시간대 할인 전략 수립
  - 지역데이터 분석을 통해 집중해야 하는 상권 분석, 새로운 메뉴 개발 활용 가능

(나) 매출 관리 프로그램 : 배민장부

□ 서비스 개요

- 매출 정보, 식자재 및 소모품 관리, 무료 구인공고 등록 서비스 등 매출 관리 서비스 프로그램으로 배달의 민족 입점 업체 이외에 소상공인도 별도 비용 없이 이용 가능

□ 데이터 유형

- 배민장부를 이용하는 업주를 대상으로 매출 현황 및 거래 정보, 주문 분석, 광고 통계, 리뷰 분석, 가게 비교 분석 서비스, 지역 상권 정보 등 무료 제공
  - 매출 통계 : 기간에 따른 주문수, 매출액 및 상세 정산내역 제공
  - 주문 분석 : 신규 주문/재주문 건수 및 금액, 객단가 제공
  - 광고 통계 : 업주가 이용 중인 광고 상품별로 노출수, 클릭수, 주문수, 주문금액 등 통계 제공
  - 리뷰 : 등록된 리뷰의 긍정/부정 비율을 확인할 수 있도록 감정분석 정보 제공
  - 가게 비교 분석 : 업주의 가게와 배민장부를 이용하는 유사 업종 가게들 간 매출, 객단가 등 비교 분석 서비스 제공
  - 상권 정보 : 업주의 가게가 위치한 지역 상권정보 제공(내부 데이터 및 통계청, 소상공인 상권정보시스템 등 활용)

□ 데이터 접근 및 이용 방식

- 배민장부를 이용하는 업주를 대상으로 매출 정보, 다른 배달앱을 통한 매출 정보, 오프라인 카드 매출 정보 연계를 위해 타 배달앱과 여신금융협회 아이디와 패스워드 등의 수집, 이용에 대한 동의 여부 확인

### □ 배민장부를 통한 분석 활용

- 카드, 배달앱, 현금영수증 등 다양한 형태의 가게 매출 정보를 통합하여 관리, 한 화면에서 확인 가능
- 매장 인기메뉴·배달 인기지역·매출 인기시간대 정보 분석이 가능
- 배민장부에서 타 배달앱 업주 전용 사이트 아이디와 비밀번호를 입력하면 이용 중인 다른 배달앱 매출 정보를 확인할 수 있도록 구성
- 같은 업종 및 같은 메뉴의 가게 그룹과 비교 분석을 통한 가게 운영에 도움

## (3) 구글

### (가) 구글 애널리틱스(Google Analytics)

#### □ 서비스 개요

- 구글 애널리틱스는 웹사이트 및 앱에서 데이터를 수집하여 비즈니스에 유용한 정보를 제공하는 웹분석 솔루션

#### □ 데이터 유형

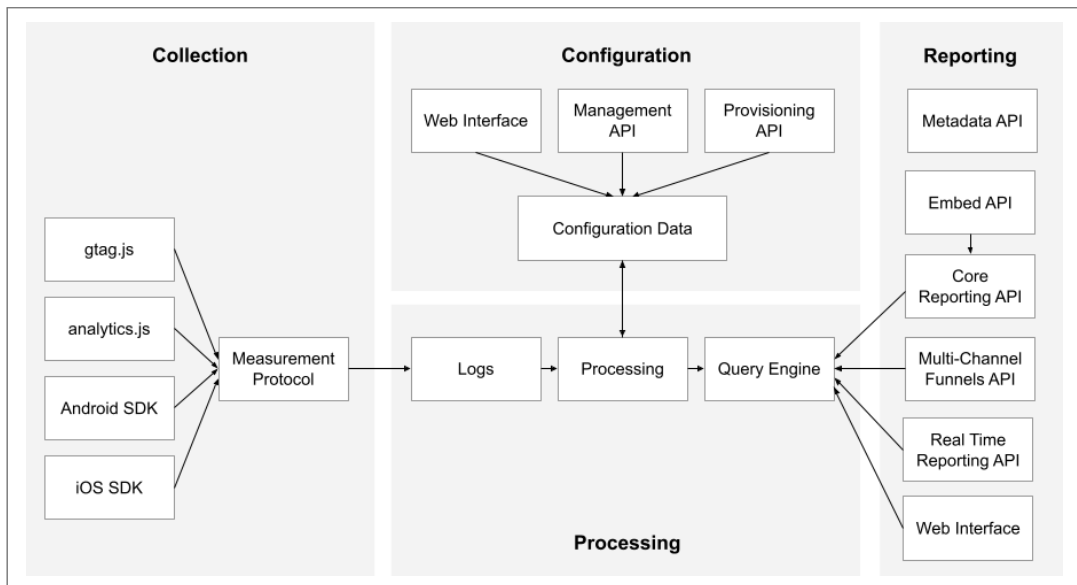
- 구글 애널리틱스는 맞춤설정, 실시간 분석, 잠재고객 분석, 유입 분석, 행동 분석, 전환 분석 등의 데이터를 제공함
  - 홈(Home) : 웹사이트의 주요 데이터를 요약적으로 보여줌. 13개의 주요 항목들이 간략하게 표시되며, 연관된 보고서로 이동할 수 있는 링크 제공
  - 맞춤 설정(Customization) : 대시보드, 맞춤 보고서, 저장된 보고서, 맞춤 알림 등 맞춤 설정 가능
  - 실시간 보고서(Real-Time): 현재 웹사이트를 이용하고 있는 사용자들의 활동에 대한 정보
  - 잠재고객 보고서(Audience) : 연령, 성별, 지역, 방문 빈도 등 웹사이트에 접속한 모든 사람들의 특성 정보
  - 획득 보고서(Acquisition) : 웹사이트 사용자가 어떤 채널 혹은 유입 경로를 통해서 웹사이트에 접속했는지 정보 제공

- 행동 보고서 (Behavior): 조회 페이지, 검색 키워드 등 웹사이트 사용자들이 웹사이트에 방문해서 어떤 행동을 하는지 분석
- 전환 보고서(Conversions) : 사이트 내에서 구매 혹은 특정 이벤트(회원 가입, 다운로드 등) 달성과 같이 목표로 한 행위에 대한 전환 행위 데이터 제공

□ 데이터 접근 및 이용 방식

- 개발자는 풍부한 사용자 인터페이스, 클라이언트 라이브러리, API를 통해 수집, 구성, 처리, 보고의 4가지 기본 구성요소로 상호작용
  - 컬렉션 : 사용자 상호작용 데이터 수집
  - 구성 : 데이터 처리 방식 관리
  - 처리 : 구성 데이터를 통해 사용자 상호작용 데이터를 처리
  - 보고 : 처리된 모든 데이터에 대한 액세스 권한 제공

[그림 III-7] 구글 애널리틱스 플랫폼 구성요소



출처: 구글 애널리틱스(<https://developers.google.com/analytics/devguides/platform?hl=ko>)

### □ 구글 애널리틱스의 활용 방안

- 웹사이트에 대한 지속적인 관리 및 퍼포먼스와 효율을 트래킹 할 수 있는 목표 설정을 통해 통찰력 제공
- 기본 스크립트 설치만으로 잠재 고객에 대한 방대한 데이터 수집이 가능하며 관련 마케팅 전략 수립에 도움

### (나) 애널리틱스 허브(Analytics Hub, 데이터 교환 플랫폼)

#### □ 서비스 개요

- 구글의 애널리틱스 허브는 데이터 공유 통합 관리형 서비스로('22.10.12 출시) 조직 내, 조직 간 데이터와 분석 자산을 안전하게 교환할 수 있도록 함

#### □ 데이터 교환 플랫폼의 기능

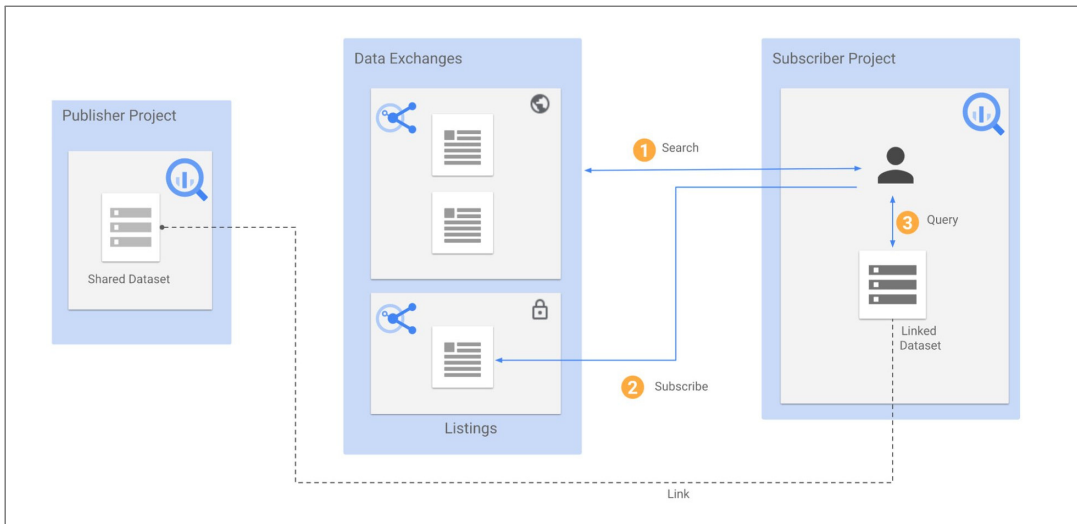
- 애널리틱스 허브는 게시 및 구독 모델을 통해 대규모 데이터를 배포하며 데이터 제공자와 가입자 모두 공유 데이터 활용 지원을 위해 다음의 기능을 추가함
  - 지역 지원 : 분석 허브 서비스는 빅쿼리가 지원되는 모든 지역에서 사용 가능
  - 구독 관리 : 데이터 공급자는 모든 공유 데이터 집합에 대한 구독을 단일 뷰에서 쉽게 보고 및 관리
  - 거버넌스 및 액세스 : 관리자는 이제 감사 로깅 및 정보 스키마를 통해 분석 허브의 사용 현황을 모니터링하는 동시에 VPC 서비스 제어를 적용하여 데이터를 안전하게 공유 가능
  - 검색 및 발견 : 가입자가 관련 목록을 빠르게 찾을 수 있도록 필터 패킷을 통해 검색 환경 개선
  - 데이터 에코시스템 : 조직이 타사 소스의 데이터를 활용할 수 있도록 금융, 지리공간, 기후, 소매업 등 다양한 산업에 걸쳐 수백 개의 새로운 공공 및 상업용 리스팅 추가

#### □ 데이터 접근 및 이용 방식

- 빅쿼리의 고유한 아키텍처를 기반으로 고객은 데이터를 이동하지 않고도 대규모로 실시간 데이터를 공유할 수 있어 데이터 관리 비용을 절감할 수 있음

- 데이터 제공자로서 안전한 데이터 교환을 생성하고 공유하려는 데이터세트가 포함된 목록을 게시할 수 있으며, 데이터 교환을 통해 목록을 보거나 구독할 수 있는 사용자 또는 그룹을 제어할 수 있음
- 기본적으로 애널리틱스 허브에서 교환은 비공개지만 모든 구글 클라우드 고객이 사용할 수 있도록 하려는 공개 또는 상업용 데이터 세트가 있는 경우 교환을 공개로 설정하도록 요청할 수도 있음
- 애널리틱스 허브는 모든 교환에서 목록을 탐색하고 검색할 수 있는 원활한 환경을 제공함
  - 데이터 구독자는 (1)관심 있는 데이터 세트를 검색하고 (2)액세스 권한을 요청하거나 액세스 권한이 있는 리스팅을 구독할 수 있음. (3)리스팅을 구독하면, 애널리틱스 허브가 프로젝트 내에서 쿼리할 수 있는 읽기 전용 링크된 데이터 세트를 생성함
  - 링크된 데이터 세트는 데이터의 사본이 아니라 공유 데이터 집합에 대한 단순한 연결 링크일 뿐이며, 소스에 대한 모든 변경 사항과 동기화 상태를 유지함

[그림 III-8] 구글 애널리틱스 허브의 데이터 교환 모델



출처: 구글 클라우드 블로그 애널리틱스 허브

(<https://cloud.google.com/blog/products/data-analytics/analytics-hub-data-exchange-is-now-generally-available?hl=en>)

## □ 애널리틱스 허브의 활용 방안

- 고객은 다양한 비즈니스 기능이나 지역에 대한 교환을 생성하여 내부 데이터를 공유할 수 있음. 예를 들어, 채널 성과, 고객 프로필, 제품 성과 등을 게시하기 위해 마케팅 교환을 설정 가능
- 조직 간 데이터를 공유할 때 고객은 각 파트너 또는 비즈니스(B2B)와 비공개 교환을 생성할 수 있음. 예를 들어 소매업체가 각 공급업체와 판매 데이터를 공유하는 경우
- 데이터 공급자는 데이터 세트로 수익을 창출하고 상업적 교환을 통해 데이터를 배포할 수도 있음. 상용 서비스 제공업체는 오프라인 자격 부여 및 청구 프로세스를 사용하고 애널리틱스 허브를 사용하여 데이터에 대한 접근 제공 가능
- 고객은 외부 또는 타사 데이터를 활용하여 새로운 통찰력을 발견하거나 경쟁 우위를 확보할 수 있음. 애널리틱스 허브와 데이터 에코시스템을 통해 분석 가능한 공개 및 상업용 데이터 세트에 쉽게 액세스할 수 있음

## (4) 아마존: AWS 마켓플레이스 상거래 분석 서비스(CAS, Commerce Analytics Service)

## □ 서비스 개요

- CAS는 아마존 AWS 마켓플레이스(AWS Marketplace)를 통해 프로그래밍 방식으로 제품 및 고객 데이터에 액세스할 수 있는 서비스로, AWS SDK를 통해 사용량, 구독 및 청구 보고서를 볼 수 있음

## □ 데이터 유형

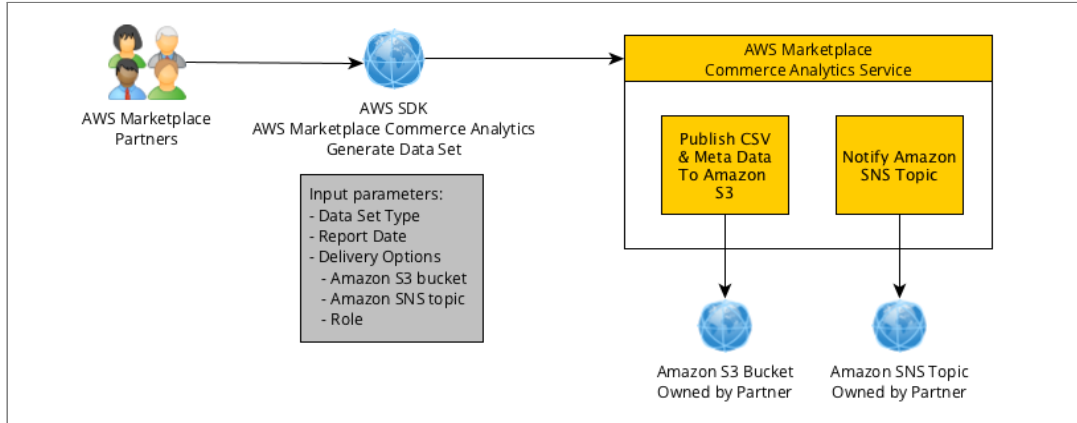
- CAS를 사용하면 API 인터페이스를 통해 보고서 데이터에 자동으로 접근 가능하며, 제품 사용, 구매자, 결제 및 지불 정보 등이 포함된 보고서(매일/매주/매월)를 제공함
- 특정 날짜에 대한 데이터 세트를 요청할 수 있으며, 데이터는 제공된 아마존 S3 버킷으로 전달, 데이터 전송에 대한 알림은 아마존 간편 알림 서비스(Amazon SNS, Simple Notification Service)를 통해 전송됨
  - 일일 비즈니스 보고서 : 고객 레퍼런스 ID, 사용자 국가/시/도, 제품명, 제품 ID, 인스턴스 유형, 예상 수익 등

- 일일 고객 구독자 보고서 : 고객 AWS 계정 번호, 제품명, 제품 ID, 제품 코드, 주문 ID, 지급인 참조 ID, 리셀러 계정 ID, 리셀러 계정 이름 등
- 지출 보고서 : 제품명, 제품코드, 제품금액, 수수료, 판매금액, 거래 참조 ID 미국 매출세 및 이용세 총액, 환급 세액, 고객 AWS 계정 번호, 고객 국가/시/도/군/구, 지급인 참조 ID, 지급인 주소 ID 등
- 월별 청구 수익 보고서 : 고객 레퍼런스 ID, 요금 청구 계정 국가/시/도, 우편번호, 제품명, 제품 코드, 대리점의 상품 비용, 파트너 수익 금액, 거래 참조 ID, 미국 매출세 고객 청구 금액, 미국 매출세 환급 금액, 제공 ID, 지급인 참조 ID, 지급인 주소 ID 등
- 영업 보상 보고서 : 고객 레퍼런스 ID, 요금 청구 계정 국가/시/도, 우편번호, 이메일 도메인, 제품명, 제품 코드, 총 수익, AWS 수수료, 총 환급금, 대리점의 상품 비용, 파트너 수익 금액, 거래 참조 ID, 지급인 참조 ID, 지급인 주소 ID 등
- 미국 매출세 및 이용세 보고서 : 항목 ID, 고객 청구서 ID, 제품명, 제품코드, 제품과세코드, 판매자 ID, 판매자명, 거래날짜, 거래가격, 총 세금, 청구지 국가/시/군/구, 거래유형코드, 표시가격 과세 유형, 세금 납부 지역, 고객 공제 인증서 ID, 고객 공제 인증서 ID 도메인 등

□ 데이터 접근 및 이용 방식

- 기존 API와 달리 데이터를 직접 반환하는 대신 프로그래밍을 통해 데이터를 비동기 방식으로 아마존 S3(Amazon S3, Amazon Simple Storage Service) 파일에 데이터를 지정할 수 있음
- SDK 도구를 사용하여 요청한 데이터는 데이터 세트 형태로 AWS 계정에 전달되며 부분의 데이터 세트는 AWS 마켓플레이스 관리 포털에서 제공되는 텍스트 기반 보고서와 동일한 데이터에 해당함
- 판매되는 소프트웨어, 서비스 및 데이터 제품의 AWS 마켓플레이스 사용에는 구매자와 판매자 간의 계약이 적용되며 제품을 무료 평가판을 통해 이용할 수 있고 사용량 기반 과금, 월간 요금 모델, 시간당 또는 시간당/연간 요금 모델, BYOL(Bring Your Own License) 모델을 통해 이용 가능

[그림 III-9] 아마존 AWS 마켓플레이스 CAS 개요



출처: 아마존 AWS(<https://docs.aws.amazon.com/marketplace/latest/userguide/commerce-analytics-service.html>)

CAS의 활용 방안

- 제품의 판매량, 수익, 재고 관리 등을 분석하여 비즈니스 의사 결정을 개선
- 고객의 행동을 분석하여 마케팅 전략 수립 및 경쟁사의 판매량, 수익 등을 분석

(5) X(구 트위터) API

서비스 개요

- X는 최대한 널리 정보를 공유할 수 있도록 기업과 개발자 및 사용자가 X API(애플리케이션 프로그래밍 인터페이스)를 통해 프로그래밍 방식으로 X 데이터에 액세스할 수 있도록 지원

데이터 유형

- X API를 사용하려면 애플리케이션을 등록해야 하며, 기본적으로 애플리케이션은 X 상의 공개 정보에만 액세스할 수 있음. 특히 쪽지를 주고 받는 등 특정 엔드 포인트는 추가로 권한을 부여받아야 정보에 접근 가능
  - 계정 및 사용자 : X는 개발자가 프로그래밍을 통해 계정 프로필 및 설정 관리, 사용자 뮤트 또는 차단, 사용자와 팔로워 관리, 승인된 계정의 활동 정보 요청을 비롯한 다양한 작업 수행
 이러한 엔드 포인트의 경우 거주자들에게 긴급 상황 대응 및 긴급 상황 알림 관련 정보를 제공하는 공공 서비스로 활용 가능

- 트윗 및 답글 : 개발자가 공개 트윗 및 답글을 이용하고 X API를 통해 트윗을 게시할 수 있도록 함. 개발자는 특정 키워드를 검색하거나 특정 계정의 트윗 샘플을 요청하는 방식으로 트윗 이용 가능  
이러한 엔드 포인트는 UN 등의 비정부기구에서 공중 보건과 관련된 잘못된 정보를 식별 및 파악하고 이에 대한 대처 수단으로 활용되며, 의료 연구자들이 발병 징후를 조기에 감지해 확산을 모니터링하는 데도 도움
- 쪽지 : X 쪽지 엔드 포인트를 통해 특정 애플리케이션에 명시적으로 권한을 부여한 사용자의 쪽지 대화 사용  
비즈니스용으로 소유 또는 관리하는 계정의 경우 이러한 사람 또는 챗봇 기반 대화 경험을 통해 고객과 직접 소통하여 고객 서비스, 마케팅, 브랜드 참여 경험을 제공
- 광고 : X는 Sprinklr처럼 개발자가 X 상의 비즈니스용 광고 캠페인을 자동으로 생성하고 관리하는 데 도움을 줄 수 있는 다양한 API 제공  
개발자들은 공개 트윗을 사용해 주제와 관심사를 파악하고, X에서 다양한 오디언스에게 도달하기 위한 비즈니스용 광고 캠페인 운영 도구 개발
- 퍼블리셔 도구 및 SDK : X는 소프트웨어 개발자 및 퍼블리셔에게 웹페이지에 타임라인, 공유 버튼, 기타 트위터 콘텐츠를 임베드하는 도구를 제공  
X에서 실시간 공개 대화를 브랜드 웹 환경으로 옮겨 자사 사이트에서 고객이 쉽게 정보와 자료 공유 가능

#### □ 데이터 접근 및 이용 방식

- X 데이터는 사용자가 공개적으로 공유하기로 선택한 정보를 반영하므로 대부분의 다른 소셜 플랫폼에서 공유되는 데이터와 달리 고유값을 가짐
- X의 API 플랫폼은 사용자가 전 세계에 공유하기로 선택한 공개 데이터에 대한 광범위한 액세스 권한을 제공하며, 사용자가 자신의 비공개 정보(예: 쪽지)를 관리하고 이러한 정보를 승인된 개발자에게 제공할 수 있는 API를 지원함
- X API와 데이터 제품 전반에 걸쳐 X는 사용자의 데이터를 제대로 보호할 책임이 있음. 이에 따라 개발자가 X 데이터를 사용하는 방법을 평가하는 엄격한 정책과 절차를 마련하여 시행 중이며, 이러한 데이터의 부적절한 사용을 제한하고 있음

## (6) 우버, 데이터 접근 정책

### □ 서비스 개요

- 우버 앱 및 웹사이트, 기타 서비스를 이용하는 이용자들에게 우버와 그 제휴사의 정보 수집 및 이용 방식을 제공하며 전 세계 모든 우버 사용자에게 적용되나 Uber Freight, Careem, Uber Carshare 또는 UT (대한민국)과 같이 별도의 개인정보 취급방침이 적용되는 서비스를 이용하는 경우는 제외됨

### □ 데이터 유형

- 사용자가 본인 정보를 관리할 수 있도록 당사는 개인정보에 대한 명확한 선택권과 간편한 관리 기능을 제공

### □ 데이터 접근 및 이용 방식

- 우버는 믿을 수 있고 편리한 이동 수단, 배달, 기타 상품 및 서비스를 제공하기 위해 정보를 사용. 사용자와 서비스의 안전과 보안 증진, 고객 지원, 연구 및 개발, 사용자 간의 커뮤니케이션 지원 마케팅 및 광고, 사용자 대상의 비마케팅 커뮤니케이션, 법적 절차 관련 데이터 사용
- 우버 일부 서비스 및 기능 이용 시 사용자의 요청이나 동의 하에 다른 사용자와 데이터를 공유함. 또한 법적 사유가 있는 경우 또는 피해 보상 또는 분쟁과 관련하여 제휴사, 자회사, 파트너와 관련 정보 공유 가능

〈표 III-8〉 우버의 정보 수집 및 이용

구분	상세 데이터	
사용자가 제공한 정보	계정 정보	이름과 성, 이메일, 전화번호, 로그인 이름과 비밀번호, 주소, 프로필 사진, 결제 정보 또는 은행 정보(관련 결제 인증 정보 포함), 사용자 설정, Uber 파트너의 로열티 프로그램 정보 포함
	신원 조회 정보 (기사 및 배달 기사)	기사의 이력이나 범죄 기록(법에 따라 허용되는 경우), 면허 상태, 알려진 별칭, 이전 주소, 근무 자격 등 기사/배달 기사 신청 과정에서 제출된 정보 포함
	신원 확인 서류 및 사진	운전면허증이나 여권과 같은 정부 발급 신분증(신분증 사진과 번호, 생년월일, 성별이 포함될 수 있음)과 셀카, 프로필 사진과 같이 사용자가 제출한 사진이 포함
	인구 통계학적 정보	특정 기능을 활성화하거나 연령 제한이 있는 제품이나 서비스에 대한 액세스 권한을 제공하기 위해 필요한 경우 생년월일/연령, 성별, 직업과 같은 인구통계학적 데이터를 수집
	사용자 콘텐츠	사용자가 Uber 고객지원팀에 문의하거나, 사용자, 레스토랑 또는 매장에 대한 별점이나 피드백을 제공하거나, 사용자가 콘텐츠를 업로드하거나 녹음(앱 내 음성 녹음 또는 대시보드 카메라 포함)을 제출할 수 있는 기능을 사용할 때 데이터(채팅 기록 및 통화 녹음 포함) 수집
	여행 정보	예약된 항공편, 숙박 시설 또는 차량 렌트 예약의 시간과 날짜를 포함한 여행 일정 정보를 Uber Travel 기능 사용자에게서 수집
서비스 이용 중에 생성되는 정보	위치 정보 (기사 및 배달 기사)	Uber 앱이 포어그라운드(앱이 열려 있으며 화면에 표시됨) 또는 백그라운드(앱이 열려 있지만 화면에 표시되지 않음)에서 실행 중일 때 기사와 배달 기사의 모바일 기기에서 정확한 위치 정보나 대략적인 위치 정보를 수집
	위치 정보 (승객 및 주문 고객)	승객과 주문 고객이 모바일 기기 설정을 통해 권한을 허용하면 이러한 기기에서 정확한 위치 정보나 대략적인 위치 정보를 수집 승객과 주문 고객은 모바일 기기의 위치 정보 수집 기능을 활성화하지 않아도 Uber 앱을 이용 가능(정확한 위치 정보를 활성화하지 않은 승객은 본인의 픽업 주소를 직접 입력) 승객이 기기에서 정확한 위치 정보를 수집할 수 있도록 설정하지 않은 경우에도 차량 서비스 중에 기사의 기기에서 수집된 정확한 위치 정보가 승객의 계정에 연결(이 정보는 영수증 생성, 고객 지원, 사기 행위 감지, 보험, 소송 등을 목적으로 사용)
	거래 정보	당사 서비스 이용과 관련된 거래 정보를 수집하며, 여기에는 요청받거나 제공된 서비스의 유형, 차량 서비스 또는 주문 세부 정보(일시, 요청된 픽업 및 도착 주소, 운행 거리, 주문 아이템 등), 결제 거래 정보(레스토랑 또는 매장의 이름 및 위치, 청구 금액 및 결제 수단) 등이 포함
	사용 정보	사용자와 당사 서비스 간의 사용 일시, 조회된 앱 기능이나 페이지, 브라우저 유형, 앱 충돌 및 기타 시스템 활동 등 상호 작용 방법에 관한 정보 수집
	기기 정보	당사 서비스 이용에 사용되는 기기에 관한 하드웨어 모델, 기기 IP 주소나 기타 고유한 기기 식별자, 운영 체제와 버전, 소프트웨어, 기본 언어, 광고 식별자, 기기 동작 정보, 모바일 네트워크 정보 등 정보를 수집
	커뮤니케이션 정보	Uber 앱을 통해 활성화된 사용자 간의 전화, 문자 메시지 또는 앱 내 커뮤니케이션에 관한 커뮤니케이션 유형(전화, 문자 또는 앱 내 메시지), 날짜/시간, 내용(사용자에게 사전에 녹음 내용을 통보받은 경우에만 통화 녹음 내용 포함) 등 정보를 수집

구분	상세 데이터
다른 출처를 통해 얻은 정보	당사의 추천 프로그램에 참여하는 사용자. 예를 들어 사용자가 다른 누군가를 추천할 경우 당사는 사용자로부터 추천받은 사람의 정보를 제공
	다른 사용자(예: 친구나 가족)를 대신하여 서비스를 요청하거나 다른 사용자가 비즈니스 계정을 통해 서비스를 요청하거나 받을 수 있도록 하는 Uber 계정 소유자(예: 기업 고객).
	피해 보상이나 분쟁에 관한 정보를 제공하는 사용자나 기타 개인
	사용자가 본인의 Uber 계정을 생성·이용하기 위해 사용하는 Uber 비즈니스 파트너(예: 결제 서비스 공급업체, 소셜 미디어 서비스, Uber의 API를 이용하거나 Uber에서 이용하는 API를 소유한 앱이나 웹사이트)
	Uber와 제휴를 맺은 금융 기관에서 발급한 체크카드나 신용카드와 관련된 Uber 비즈니스 파트너(카드 약관에 명시된 범위 내)
	사용자의 신원, 신원 정보, 근무 자격을 확인하는 데 도움을 주거나 제재, 돈세탁 방지, 고객 확인 절차(KYC)의 요건과 관련하여 사용자를 심사하는 업체
	기사 및/또는 배달 기사를 대상으로 하는 보험, 차량, 금융 서비스 제공업체
	파트너 운송업체(해당 업체와 연결된 계정으로 당사 서비스를 이용 중인 기사나 배달 기사의 경우)
	공개적으로 이용 가능한 출처
	캐시백 프로그램과 관련된 은행, 데이터 리셀러를 포함한 마케팅 파트너 및 서비스 제공업체
사법 기관 공무원, 보건 당국 공무원, 기타 정부 기관	

출처: 우버(<https://www.uber.com/>)

## 라. 데이터의 공유와 거래: 공공데이터화 사례

### (1) 네이버 데이터랩

#### 서비스 개요

- 네이버 데이터랩은 검색어 트렌드, 쇼핑 인사이트, 지역 통계, 댓글 통계 데이터를 기간별, 기기별, 성별에 따라 분류하여 제공

#### 데이터 유형

- 검색어트렌드, 쇼핑인사이트, 지역통계, 댓글통계 등 분석 데이터 제공
  - 검색어 트렌드 : 최대 5개의 키워드에 대해서 기간별, 성별, 기기별, 연령별 상대적 데이터 제공
  - 쇼핑 인사이트 : 분야 통계(쇼핑 카테고리 안에 분야를 선택하여 연령별, 기기별, 성별, 기간별 상대적 데이터 확인) 및 검색어 통계(패션 카테고리에 대해서 세부 분야와 검색어까지 입력하여 통계 확인) 제공

- 지역 통계 : 지역별 관심도(대한민국 시/군/구 지역에 대한 검색량 데이터를 10개 업종별로 분류하여 상대값 제공) 및 카드사용통계(시/군/구 지역별 결제 금액과 결제횟수에 대한 상위 데이터 제공) 제공
- 댓글 통계 : 분야별 뉴스 기사에 대한 댓글 수, 작성자 수 등 이용자 통계 제공
  - 댓글 수 : 댓글 수 및 본인 삭제 수
  - 작성자 수 : 댓글 작성자 수
  - 섹션별 분포 : 정치, 경제, 사회, 생활, 세계, IT 섹션에 대한 분포 비율
  - 시간대별 분포 : 시간별 댓글 작성 수 분포 (1시간 단위)
  - 성별, 연령별 분포 : 연령에 따른 성별 분포 (단위 10살)
  - 기기별 분포 : 모바일, PC
  - 국가별 분포 : 국내, 해외(유입이 많은 상위 국가 5개 수치와 기타로 표현)

데이터 접근 및 이용 방식

- 데이터랩 사이트 외 네이버 제공 API를 활용하여 검색어 트렌드 확인 가능

(2) 직방 RED(Real Estate Data, 부동산 정보 공공데이터 개방)

서비스 개요

- 직방이 보유하고 수집한 사용자의 행동 데이터와 △가격 △거래 △공급 △분양 등 부동산/공공 데이터를 가공·분석하여, 유용한 인사이트를 창출하고, 사용자에게 관련 주거 정보를 제공
- 부동산 전문가를 위한 인공지능(AI) 빅데이터 솔루션 직방 RED 서비스를 출시 ('23.3.29)

데이터 유형

- 부동산 경기 상황에 더해 입지와 수요, 주거 편의성 등 여러 관점에서 정보를 제공해 데이터 실용성을 강화하여 유료 서비스로 제공
- 아파트, 오피스텔, 주택 등 주거 형태에 따라 지역별 가격부터 공급 시장 분석 등 공공 데이터를 활용한 전국의 매물 정보에 대한 빅데이터 분석 서비스 제공

- 청약경쟁률, 분양가 등 분양 정보, 지역 별 분양 리스크 및 기회요인 등 아파트 분양 시장 분석과 직방 빅데이터 분석을 통해 보유 인기 많은 아파트부터 함께 찾아 본 아파트까지 최신 트렌드 분석 제공
- 정주 및 고용 여건 분석을 통해 개별 부동산 매물이 △교육 △의료 △대중교통 △쇼핑 △문화 다섯 가지 인프라를 얼마나 잘 갖췄는지 분석 데이터 제공
- 앱 거주민 리뷰를 자연어 처리기법 기반으로 분석한 데이터와 앱 이용자 로그 분석 데이터 등 직방 고유 정보 제공

□ 데이터 접근 및 이용 방식

- 정부의부동산 거래 및 청약, 건축물 및 토지 데이터, 직방 플랫폼의 이용 고객 데이터를 통해 건설, 금융사와 같은 B2B 기업 사용자가 다양한 관점에서 부동산 시장을 보다 잘 이해할 수 있도록 데이터 시각화를 통한 부동산 통계 정보 지원
- 직방 RED 웹 포털에서 다양한 필터링 옵션을 맞춤형하여 적용, 필요 데이터만 다운로드 가능

〈표 III-9〉 직방 공공데이터 활용 및 수요 현황

구분	유형	
공공데이터 활용	토지	<ul style="list-style-type: none"> <li>• 대장정보(토지, 임야), 지적, 등기 등 DB(국토교통부)</li> <li>• 개발사업현황, 내·외국인 토지소유 현황 통계 자료(국토교통부)</li> </ul>
	주택/건물	<ul style="list-style-type: none"> <li>• 대장정보(건축물, 멸실), 등기 등 DB(국토교통부)</li> <li>• 주택소유(통계청), 주택 재고, 임대주택(국토교통부), 주택보증(주택도시보증공사) 통계</li> <li>• 공동주택 관리비 통계 자료(한국부동산원)</li> <li>• 인허가, 착공, 준공, 분양 통계 자료(국토교통부)</li> </ul>
	가격, 동향정보	<ul style="list-style-type: none"> <li>• 공시가격(토지, 단독주택, 공동주택) DB(국토교통부)</li> <li>• 실거래 신고정보 DB(국토교통부)</li> <li>• 주택가격 동향, 실거래 가격지수(한국부동산원), 소비자 심리지수(국토연구원) 통계</li> </ul>
공공데이터 수요	가격, 동향정보	<ul style="list-style-type: none"> <li>• 공시가격(토지, 단독주택, 공동주택) DB(국토교통부)</li> <li>• 실거래 신고정보 DB(국토교통부)</li> <li>• 주택가격 동향, 실거래 가격지수(한국부동산원), 소비자 심리지수(국토연구원) 통계</li> </ul>

※ 2021년 부동산서비스산업 실태조사 설문 답변 기준  
출처: 직방

### (3) 직방: RED(Real Estate Data, 부동산 통계데이터)

- 부동산 시장의 통계 데이터를 한눈에 알아볼 수 있도록 시각화 서비스 제공
  - 직방이 보유하고 수집한 사용자의 행동 데이터와 △가격 △거래 △공급 △분양 등 부동산/공공 데이터를 가공·분석해 유용한 인사이트를 창출
  - 사용자에게 관련 주거 정보를 제공. 타블로(tableau)의 인터랙티브 분석 대시보드 화면을 직방 RED 웹서비스용 포털에 통합하여 임베디드 애널리틱스 서비스를 구축해서 제공

### (4) 구글 클라우드 데이터세트(Google Cloud Datasets)

- 서비스 개요
  - 빅쿼리(BigQuery, 비즈니스 데이터 웨어하우스), 클라우드 스토리지(Cloud Storage, 객체 스트로지), 어스 엔진(Earth Engine, 위성 영상 분석) 등 구글 클라우드 서비스를 통해 제공되는 사전 빌드된 데이터 솔루션 및 데이터세트를 사용하여 관심있는 데이터를 손쉽게 분석, 여러 목적으로 활용 가능
- 데이터 유형
  - 분석, 빅데이터, 데이터베이스, 머신러닝, 지도 등 유형별로 211개의 공개 무료 데이터세트에 접근 가능
- 데이터 접근 및 이용 방식
  - 신규 고객에게 구글 클라우드 전체 상품을 탐색하고 평가할 수 있도록 \$300의 무료 크레딧이 제공되며 지정된 무료 사용량 한도까지는 사용 시 요금이 부과되지 않음
  - 한 달에 최대 1TB를 무료로 탐색할 수 있으며 결제 계정이 없더라도 빅쿼리 샌드박스 사용을 시작할 수 있음.
  - 마켓플레이스(Marketplace)에서 무료로 제공되는 유형별 공개 데이터 중 관심 있는 데이터를 클릭하면 [데이터세트 보기]로 연결, 버튼을 클릭하면 클라우드 빅쿼리 플랫폼으로 이동하여 쿼리문을 작성 및 조회할 수 있음

- Legacy SQL 쿼리 또는 표준 SQL 쿼리를 사용해 빅쿼리와 클라우드 스토리지에서 구글 클라우드의 공개 데이터 세트에 액세스할 수 있음
- 이러한 데이터세트의 콘텐츠는 교육 및 연구 목적으로만 제한적으로 일반에 제공되는 경우이며, 각 데이터별로 개인정보 보호 및 데이터 보안 정책에 따라 구글은 데이터 처리 규정 및 정책을 마련하고 있음

〈표 III-10〉 구글 클라우드 추천 데이터세트

카테고리	추천데이터세트	
구글데이터세트	Google 트렌드	지난 30일 동안의 구글트렌드에서 상위 25개 및 25개 급상승 쿼리 제공 각 검색어에는 미국 내 210개의 지정 시장 지역(DMA)에 대한 5년간의 과거 데이터와 현재 전 세계 50여 개 국가의 데이터가 포함
	Google 애널리틱스 (샘플)	Google Merchandise Store에서 난독화된 12개월(2016년 8월~2017년 8월)치의 Google 애널리틱스 360 데이터 제공 전자상거래 웹사이트의 트래픽 소스, 콘텐츠, 트랜잭션 데이터 등 데이터 포함
	Google 특허 검색	Google 번역의 제목과 초록 기계 번역, 임베딩 벡터, 상위 용어 추출, 유사 문서, 순방향 참조 등 Google 특허 검색(patents.google.com)에서 사용되는 데이터 분석 작업 포함
상용 데이터 세트	Crux Informatics	Crux Deliver는 데이터 엔지니어링 및 운영을 위한 관리형 서비스로, Crux는 고객을 대신하여 기존의 모든 데이터 제공업체와 대체 데이터 제공업체를 연결하고 온보딩, 데이터 엔지니어링 및 운영의 모든 측면을 관리함 모든 데이터 세트는 유효성 검증, 정제된 실행 가능한 데이터만 제공
	Exchange Data International (EDI)	글로벌 금융 및 투자에 필요한 데이터 제공 전 세계 자기 자본 및 고정 소득 기업 조치, 배당금, 정적 참조 데이터, 최종가, 보유 주식 등 데이터 포함
	Factset	우수한 콘텐츠, 분석, 유연한 기술을 제공하는 통합 재무 정보, 분석 애플리케이션, 업계 최고의 서비스를 제공
	HouseCanary	1억 채 이상의 부동산, 대출, 가치 평가 정보에 즉시 접근 가능 ML 알고리즘은 수백 개의 데이터 소스를 처리하여 미국 주택 시장의 95% 이상을 차지하는 381개 대도시, 18,300개 우편번호, 400만 개 블록에 대한 주택 가격 지수를 제공
	LinkUp	고객이 글로벌 노동 시장에 대한 가치 있는 인사이트를 도출하고 투자자가 거시적, 부문별, 지역별, 개별 기업 수준에서 알파를 창출할 수 있도록 지원하는 독점 데이터 솔루션을 제공
	London Energy Brokers Association	LEBA의 솔루션을 통해 고객은 주요 에너지 중개업체가 제공하는 에너지 시장에 대한 고유하고 통합된 시각 제공 에너지, 석유 및 가스 생산자, 도매 사용자, 유틸리티 및 금융 트레이더는 가격 평가가 아닌 거래 활동을 기반으로 한 독립적인 시장 정보 제공

카테고리	추천데이터세트	
상용 데이터 세트	Neustar	트랜스유니온의 자회사로 가장 중요한 순간에 기업과 사람 간의 신뢰할 수 있는 연결을 가능하게 하는 데이터와 기술을 제공 마케팅, 리스크, 커뮤니케이션 분야에서 업계를 선도하는 솔루션을 제공
	RS Metrics	자산 수준의 객관적이고 검증 가능한 실시간 ESG 데이터 제공 RS Metrics는 고객이 전기차 제조업체의 공장 재고 수준에 대한 정확한 인사이트를 확보할 수 있도록 지원
	Ursa Space Systems	글로벌 위성 인텔리전스 인프라 제공업체인 Ursa Space Systems는 매주 업데이트되는 위성 이미지에서 파생된 데이터로 글로벌 경제 동향을 모니터링할 수 있는 기능을 고객에게 제공
공개 데이터 세트	Severe Storm Event Details	1950년부터 올해까지 미국 전역에서 발생한 기상이변의 위치, 방위각, 거리, 영향, 피해정도, 재산 및 농작물 피해 비용 등에 대한 통합 데이터베이스 제공
	Census Bureau US Boundaries	TIGER/라인 모양 파일에서 파생된 전체 해상도 경계 파일로, 50개 주, 컬럼비아 특별구, 푸에르토리코 및 외곽 섬 지역에 대한 정보가 포함
	American Community Survey	미국 지역사회 설문조사(ACS)는 미국 전역의 350만 가구에 연락하여 매년 미국 및 국민에 대한 중요한 정보를 제공하는 설문조사 결과 데이터는 다양한 지리적 수준에서 집계된 미국 전역의 상세한 인구통계학적 정보를 제공
	All public datasets	Google 클라우드 마켓플레이스에 등록된 200개 이상의 데이터 세트를 검색 및 이용 가능
어스 엔진 데이터 세트	Earth Engine	어스 엔진의 공개 데이터 아카이브에는 40년 이상의 역사적 이미지와 과학 데이터 세트가 포함, 매일 업데이트되어 온라인 분석에 사용 가능
캐글 데이터 세트	Kaggle datasets	데이터 과학 작업에 필요한 모든 코드와 데이터 제공 80,000개가 넘는 공개 데이터 집합과 400,000개가 넘는 공개 노트북을 통해 다양한 분석 가능
통합 데이터 집합	Cymbal Investments	심벌 인베스트먼트 그룹이 운영하는 자동화된 트레이딩 봇의 거래를 나타내며, 각 봇은 단일 알고리즘을 사용하여 거래 의사 결정을 제공 빅쿼리에 로드된 FIX 프로토콜(버전 4.4) 거래 체결 보고서에서 추출됨
연구 데이터 세트	Dataset Search	구글의 데이터 세트 검색 프로그램은 웹에서 약 2,500만 개의 데이터 세트를 색인화하여 한 곳에서 데이터 세트를 검색하고 데이터의 위치로 연결되는 링크를 찾을 수 있게 해줌 최근 날짜, 형식, 주제 등으로 필터링 가능

출처: 구글 클라우드 데이터 세트(<https://cloud.google.com/datasets?hl=en>)

### □ 구글 클라우드 데이터세트의 활용 방안

- 데이터세트를 사용하면 데이터 포인트를 탐색하고 레이어링하는 데 도움이 되며 기업 고객은 훨씬 더 간편하게 데이터 분석을 수행할 수 있음
- 유틸리티 데이터세트를 사용하면 처음부터 만들 필요 없이 유효하고 정제된 데이터 세트로 빅데이터와 데이터 분석을 탐색할 수 있음

## (5) AWS 데이터 익스체인지(Data Exchange, 데이터 전환 플랫폼)

### □ 서비스 개요

- AWS 데이터 익스체인지는 파일, API 또는 아마존 레드시프트(Amazon Redshift) 쿼리를 통해 이를 사용하는 데이터 레이크, 애플리케이션, 분석 및 기계 학습 모델에 직접 제공되는 300개 이상의 공급자가 제공하는 3,500개 이상의 제품을 갖춘 유일한 데이터 마켓플레이스
- 단일 데이터 카탈로그에서 전 세계 서드 파티 데이터를 손쉽게 찾을 수 있도록 지속적으로 혁신하고, 일관된 요금 옵션으로 간편하게 가입하고, AWS 데이터 및 분석 및 기계 학습 서비스에 원활하게 사용할 수 있도록 지원

### □ 데이터 교환 플랫폼의 이점

- 자동차, 금융 서비스, 게임, 의료 및 생명 과학, 제조, 마케팅, 미디어 및 엔터테인먼트, 소매를 비롯한 다양한 산업을 아우르는 1,000개 이상의 무료 데이터 세트 중에서 선택할 수 있음
- 광범위한 데이터 세트 선택
  - 300개 이상의 데이터 제공업체가 제공하는 3,500개 이상의 데이터 세트
  - 1,000개 이상의 무료 데이터 제품 및 맞춤형 데이터 제품
  - 새로운 데이터에 대한 자동 액세스
- 간소화된 데이터 조달 및 거버넌스
  - 공개적으로 또는 비공개적으로 데이터를 교환할 수 있는 단일 장소
  - 추가 비용 없이 기존 구독 마이그레이션
  - 단순화된 계약 및 통합 보안 결제

- 더 나은 데이터 기술
  - 데이터를 AWS에 기본 통합
  - 저장 및 전송 중 암호화된 데이터
  - AWS Identity & Access Management(IAM)과 통합
- 데이터 분석 및 기계 학습의 사용 편의성
  - 데이터 파일, 테이블, API에 대한 클라우드 기본 지원
  - 셀프서비스 옵션
  - 안전하고 규정을 준수하는 솔루션

□ 데이터 접근 및 이용 방식

- AWS 데이터 익스체인지를 사용하면 사용자나 공급자에게 추가 비용 없이 마이그레이션할 수 있는 기존 구독부터 향후 데이터 구독까지 한 곳에서 모든 타사 데이터 소비를 간소화할 수 있음
- AWS 서비스인 AWS Data Exchange는 안전하고 규정을 준수하며 AWS 및 타사 도구 및 서비스와 통합되며 통합 청구 및 구독 관리 제공

## 마. 협력적 데이터 공유

### (1) 직방 등 부동산 빅데이터 플랫폼

- 한국 부동산원과 빅데이터 플랫폼 구축에 참여해서, 자사에 데이터를 공공
- 직방 등 총 15개 공공기관, 민간기업과 함께 협업체계를 구성해서 추진

### (2) 사회 파트너십을 위한 산업 데이터(IDSP, Industry Data for Society Partnership)

□ 서비스 개요

- IDSP와 오픈 데이터 연구소(ODI, Open Data Institute)는 지역 환경을 위한 데이터 챌린지(최대 10만 달러의 상금)를 시작하여 공개적으로 이용 가능한 데이터가 지방 정부의 탄소중립 목표 달성 등 지역 환경 개선 방법을 보다 효율적으로 이해하는 데 데이터의 역할을 조명함

- 참여기업: 구글, 마이크로소프트, IBM 등의 플랫폼 기업들과 기타 통신, 금융 기업 등이 자사가 구축한 데이터를 공개

#### □ 데이터 유형

- 사회 파트너십을 위한 산업 데이터는 회원들이 공개하고 선별한 민간 부문 데이터를 포함하여 이 과제와 관련된 공개적으로 사용 가능한 데이터 제공
  - 예를 들어 에너지 전환 공간에서는 전력생산, 형평성, 운송 및 이동성, 난방 및 건물, 토지 이용과 환경, 경제적 혜택, 친환경 일자리와 기술 등 같은 주제 또는 영역에서 데이터 제공

#### □ 데이터 접근 및 이용 방식

- 참여자는 IDSP 및 오픈 데이터 연구소의 회원사가 제공하는 다음과 같은 리소스에 접근 가능함 : MS 제공
  - 참가자의 애저(Azure)를 지원하는 MS학습, 데이터과학자를 위한 MS학습 교육
- 영국 전력 네트워크 제공
  - 영국 전력 네트워크의 오픈 데이터 포털용 API 설명서
  - 영국 전력 네트워크의 그리드 및 주 변전소 대시보드 - 부하, 여유 공간, 예측 등 변전소에 대한 주요 데이터를 통합해 제공
- ODI
  - 데이터 사용자를 위한 데이터 라이선스 가이드
  - ODI 온라인 학습 모듈 : 오픈 데이터 필수 사항, 데이터에서 스토리 찾기
  - 자신의 데이터를 제공하려는 참가자를 위해 데이터 세트를 식명화하는 방법
  - 데이터 공유 툴킷에는 사례 연구, 체크리스트, 치트 시트 및 가이드를 지원하는 7개의 온라인 학습 모듈 포함
  - 데이터 환경 플레이북에는 데이터 프로젝트에 유용한 도구와 가이드라인 포함

### (3) 사회적 영향 프레임워크를 위한 공개 데이터(ODSIF, Open Data for Social Impact Framework)

#### □ 서비스 개요

- 전 세계 비영리, 다자간 조직, 정부 및 연구 기관이 데이터를 활용하여 사회 문제를 해결하는 데 기여하기 위한 리소스
- 2020년 4월, MS가 오픈 데이터 캠페인을 시작하면서 공개 및 공유 데이터를 기반으로 구축된 23개 협업을 지원, 새로운 비즈니스 솔루션 개발을 도모하고 중요한 사회적 과제를 해결하기 위해 데이터를 활용하기 위한 도구임
- 참여기업: 구글, 마이크로소프트, IBM, 오라클 등

#### □ ODSIF 이니셔티브

- 생성적 AI 기술 챌린지 : 사회적 영향을 촉진하기 위해 생성 AI에 대해 개인을 교육하고 기술을 향상시키는 조직을 위한 글로벌 기금 지원
- 용량 가속기 네트워크(CAN, Capacity Accelerator Network) : 전 세계적으로 목적 중심의 데이터 실무자로 구성된 인력을 구축하고 사회적 영향을 미치는 조직이 다양한 데이터 기반 인재, 리더십, 인프라, 전략 및 문화를 개발하는 과정에서 데이터의 힘을 활용하여 임무를 달성할 수 있도록 지원
- 에피버스(Epiverse) : 공중 보건 위기를 극복하는 데 전념하는 신뢰할 수 있는 데이터 분석 생태계를 개발하기 위해 노력하는 글로벌 협업으로, 공중 보건을 위한 전 세계적 싸움에서 표준화되고 쉽게 접근할 수 있는 역학 소프트웨어 도구 시스템을 통해 문제를 해결
- 포용적 성장 및 회복 챌린지 : 포용적 성장을 위한 마스터카드 센터와 록펠러 재단의 지원을 받아 data.org 챌린지는 특히 코로나19가 건강과 경제적 영향을 미치는 상황에서 사람들과 지역 사회가 번영할 수 있도록 데이터의 힘을 활용하는 획기적인 아이디어에 대해 기금 및 기술 지원

#### (4) 메타 데이터 포 굿(Data for Good)

##### □ 서비스 개요

- 중소기업과 지역 사회를 지원하기 위한 새로운 도구로서 페이스북은 데이터 포 굿 프로그램을 통해 경제 회복을 지원하는 새로운 데이터 세트와 인사이트 제공

##### □ 데이터 유형

- 비즈니스 활동 동향(Business Activity Trends), 통근 구역(Commuting Zones), 경기 동향 조사를 통한 인사이트(Economic Insights from the Symptom Survey), 미래 비즈니스 설문조사(The Future of Business Survey) 결과 등 총 4가지 데이터 세트를 누구나 사용할 수 있도록 공개

##### □ 데이터 접근 및 이용 방식

- 데이터 포 굿 프로그램은 코로나19와 맞서고 있는 전 세계 공중 보건 기관들에게 제공하는 기존 데이터 제품을 기반으로 구축됨
- 메릴랜드 대학의 API 혹은 Humanitarian Data Exchange의 페이스북 페이지를 방문하면 누구나 경기 동향 조사를 통해 집계된 데이터 확인 가능
- 비즈니스 활동 동향 및 통근 구역 데이터는 diseaseprevmaps@fb.com 및 futureofbusiness@fb.com으로 이메일을 보내면 관련 마이크로데이터를 받아볼 수 있음

## 바. 플랫폼 데이터 공유를 위한 제언

- 디지털 산업 생태계에서 데이터 공유는 정보의 흐름을 촉진하고, 혁신 및 협업을 가능하게 하는 핵심 요소로, 이는 기업, 연구기관, 정부, 그리고 최종 사용자들 간에 다양한 데이터를 통해 상호작용하고 가치를 창출하는 과정을 의미. 데이터 공유를 통해, 생태계 내의 기업이나 조직들은 자신의 역량을 넘어서는 통찰력을 얻고, 이를 통해 경쟁력을 강화하고, 새로운 시장 기회를 창출할 수 있음
- 디지털 산업 생태계 내에서 데이터 공유는 혁신을 가속화시키기에 지속적인 개방 및 공유 노력이 필요함
  - 예를 들어, 빅데이터 분석을 통해 얻은 인사이트는 제품 개발, 고객 서비스 향상, 그리고 비즈니스 프로세스 최적화에 활용될 수 있고, 이러한 분석은 시장 동향을 파악하고, 사용자 행동을 이해하는 데 필수적이며, 보다 맞춤형 제품과 서비스를 제공하는 데 기여
- 데이터 공유는 플랫폼 경제에서 중요한 역할을 수행. 플랫폼은 다양한 사용자와 제공자를 연결하는 중개자 역할을 하며, 데이터 공유를 통해 이러한 연결을 강화
  - 예를 들어, 소셜 미디어 플랫폼은 사용자의 데이터를 수집하고 분석하여, 개인화된 콘텐츠를 제공하고, 광고 타겟팅을 최적화. 이러한 과정은 사용자 경험을 개선하고, 플랫폼의 가치를 증가시키는 데 중요한 역할을 수행
- 데이터 공유는 개인정보 보호, 데이터 보안 및 거버넌스와 같은 도전 과제를 수반하기에 기업과 조직은 데이터 보호법, 개인정보 보호 정책 및 적절한 보안 프로토콜을 준수해야 함. 데이터 거버넌스는 민감한 데이터의 적절한 관리와 보안을 보장하고, 데이터 공유의 잠재적인 위험을 최소화하는 데 중요
- 디지털 산업 생태계에서 데이터 공유는 기업, 정부, 연구 기관 및 개인 사용자 간의 경계를 넘나들며 새로운 협업의 기회를 창출. 이는 지식의 확산, 기술 혁신의 가속화, 그리고 사회적, 경제적 가치 창출에 기여. 디지털 산업 생태계의 지속적인 성장과 발전을 위해서는 데이터 공유의 중요성을 인식하고, 이를 적극적으로 촉진하는 환경을 조성하는 것이 필수적

# 제 IV 장

## 플랫폼 경제 측정

### 1. 서론

#### 가. 연구의 필요성 및 목적

##### □ 연구의 필요성

- 플랫폼의 경제적 영향력 및 규모가 확대되고 있어 이들에 대한 정책 입안의 필요성이 증대되고 있으나, 이들의 경제적 역할이나 행위에 대한 데이터는 부족한 상황
  - 이는 비단 국내뿐 아니라 전 세계적으로 공통된 상황으로, 전통적인 경제통계나 국민 계정 등은 플랫폼의 경제활동(특히, 자체적으로 제품을 생산하거나 비축하지 않고 중개만 하는 경우)을 모니터링하는데 적합하지 않음
- 이러한 문제의식하에 유럽연합의 플랫폼 경제관측소(EU Observatory on the Online Platform Economy)는 2021년 플랫폼 경제측정에 관한 보고서(EU OOPE, 2021)를 발간
- EU 플랫폼 경제관측소 보고서의 내용은 아래와 같이 간략히 정리해볼 수 있음
  - (배경) 플랫폼의 게이트키퍼로서의 역할이 부각되고 있음에도 플랫폼 생태계의 특성을 반영한 객관적 자료의 부재로 정책 입안에 한계
  - (목표) 플랫폼 경제를 모니터링하고 증거기반의 정책 추진을 뒷받침할 수 있도록 플랫폼 경제 관련 지표의 측정방법 및 데이터 수집 방안을 연구
  - (측정지표) 3개 대주제, 9개 소주제 별 정량적 지표 및 가용데이터 현황 제시

〈표 IV-1〉 EU 플랫폼 경제관측소 경제측정 보고서의 구성

대주제	소주제
플랫폼의 경제적 중요성	플랫폼이 중개한 무역 규모
	플랫폼의 규모와 중요도
	데이터에 대한 데이터
플랫폼의 이용자 영향력	이용사업자의 플랫폼 의존도
	플랫폼의 소비자 관심 점유율
	경쟁 전략으로서의 인수
플랫폼 영향력의 결과	플랫폼의 변동성
	플랫폼의 투명성
	기타 정책 관련 실증

출처: EU OOPE(2021)

□ 연구 목적

- 본 보고서는 EU 플랫폼 경제관측소의 플랫폼 경제측정 보고서를 벤치 마크하여, 국내 실정에 맞는 체계적인 플랫폼 경제측정 방안을 제시하고자 함
- EU 플랫폼 경제관측소의 플랫폼 경제측정 보고서는 유럽연합이 처한 상황의 특수성 등으로 인해 우리나라 상황에 그대로 적용하는데 한계가 존재
  - 유럽연합의 경우, 플랫폼 시장 내 해외 플랫폼 사업자의 점유율이 높고 자국 플랫폼의 영향력이 미약한 상황인 만큼, 게이트키퍼 역할을 하고있는 해외사업자의 막강한 영향력을 어떻게 측정할 것인가가 보고서의 주안점
  - 이에 더해, 소주제별로 측정지표를 나열하는 것에 그칠 뿐, 이들을 활용한 구체적인 분석 예 등은 제시하지 않아 개별 지표의 중요성이나 활용도에 대해 체감하기 어려움
- 본 보고서는 EU 플랫폼 경제관측소의 보고서를 참고하되, 다음과 같은 점에서 차별성을 가짐
  - 첫째, 게이트키퍼의 영향력을 측정하는 것에만 목표를 두지 않고, 국내 플랫폼 산업 생태계 전반을 객관으로 조망할 수 있는 체계의 구축을 목표로 함
  - 둘째, 플랫폼 관련 선행연구 결과를 기반으로 측정지표를 제시함으로써 지표의 필요성이나 활용도에 대한 보다 구체적인 아이디어를 제공하고자 함

- 셋째, 경쟁력 있는 자국 플랫폼을 보유하고 있다는 환경적 특수성을 고려하여, 자국 플랫폼의 가치 측정방안에 대해 논의하고자 함

## 나. 보고서의 구성

- 2장에서는 플랫폼의 경쟁 측정방안에 대해 논의함
  - 전통적인 경쟁측정 방식을 플랫폼 산업에 직접 대입하기 어려운 이유를 분석하고, 기존 판례 등을 검토함으로써 플랫폼의 경쟁 측정방식 및 측정지표에 관해 논의함
  - 이에 더해, 모니터링 목적에서의 경쟁지표 수집 방안을 논의하고 이들의 한계점과 보완점에 대해 논의함
- 3장에서는 플랫폼 영향확산 측정의 필요성과 측정방안을 논의함
  - 온라인 쇼핑 플랫폼, 숙박 플랫폼 및 배달 플랫폼 관련 선행연구 결과로부터 플랫폼의 영향이 이용자, 이용사업자, 지역경제 및 노동시장에 어떻게 분배되었는지를 살펴봄
  - 이를 기반으로 플랫폼 영향확산을 지속적으로 파악하고 분석하기 위해 필요한 측정지표를 제안함
- 4장에서는 자국 플랫폼의 가치 측정방안을 논의함
  - 자국 플랫폼 보유의 중요성을 피력하고, 자국 플랫폼의 조작적 정의 및 가치 측정 방안을 논의함
  - 선행연구 현황을 정리하고, 이들의 한계점과 대안적 측정방안을 제시함
- 5장에서는 국내외 플랫폼 관련 조사를 정리하고, 2-4장에서 논의된 측정지표를 체계적으로 모니터링할 수 있는 방안을 모색함

## 2. 플랫폼 경쟁

### 가. 플랫폼 경쟁 측정의 필요성

#### (1) 플랫폼 경쟁 현황과 이슈

##### □ 국내외 플랫폼 경쟁 현황

- 플랫폼은 소셜미디어, 온라인 커머스, 클라우드 컴퓨팅, 모바일앱, 예능, 동영상 및 비디오스트리밍 플랫폼으로 분류될 수 있으며, 혁신을 바탕으로 정보 공유와 소통, 소비와 생산에서 급진적인 변화를 가져왔음
- 플랫폼의 각 분야에서 기술 발전과 새로운 시장 참가자 등의 영향으로 치열한 경쟁이 일어나고 있음에도 불구하고, 소수 플랫폼의 시장지배력이 유지되고 있음
- GAF(A(Google), A(Apple), F(Facebook), A(Amazon))의 전세계적인 시장지배력
  - 2022년 9월 기준, GAF(A)의 기업 가치는 약 5조 달러로 S&P 500 전체의 1/3 이상을 차지하였으며, 지난 10여 년간 약 500여 개의 잠재적 경쟁사업자를 인수하였음
  - GAF(A)와 Microsoft는 2020년 12월 기준 세계기업 시가총액 순위 1~5위를 차지하였고, 2023년 12월 기준 세계기업 시가총액 순위 1~6위(NVIDIA가 5위를 차지)를 차지함
- 네카쿠배(네이버, 카카오, 쿠팡, 배달의 민족)와 다양한 플랫폼 기반 스타트업 등장
  - 인터넷 포털 및 검색 엔진 시장에서는 네이버와 다음, 온라인 커머스 시장에서는 네이버와 쿠팡, 온라인 음식배달시장에서는 배달의 민족과 쿠팡이츠, 모빌리티 시장에서는 카카오모빌리티와 UT가 각각의 분야에서 경쟁하며 성장 중임

##### □ 플랫폼 관련 이슈

- 플랫폼 시장에서의 집중도 증가 및 소수 플랫폼의 시장지배력 증가
- 거래관계에서 이해상충 및 갈등, 경쟁법 위반 행위에 대한 우려 증가
  - 플랫폼이 심판과 선수를 겸하는 이중적 지위의 악용 가능

- 검색 알고리즘 조정, 자사우대, 서비스 일방적 제한, 중단, 최저가 보장 등
- 경쟁 저해 및 왜곡으로 인한 후생 및 혁신 저해에 대한 우려

## (2) 플랫폼 관련 법과 규제

### □ 유럽연합

- 2010년 시작된 유럽경쟁당국의 구글쇼핑 사건 조사를 시작으로 하여 빅테크 기업에 대한 적극적인 경쟁법 집행해 옴
- 「디지털 시장 법안(Digital Markets Act)」과 「디지털 서비스 법안(Digital Services Act)」의 사전규제 도입
- 영국은 디지털 시장을 위한 친경쟁제도(Pro-competition framework)를 수립하기로 함

### □ 미국

- 2021년 6월 하원에서 민주당과 공화당이 ‘플랫폼에 대한 반독점 패키지’ 규제 법안을 공동 발의하였으나, 그 중 ‘기업결합 심사수수료 현대화법’을 제외한 나머지 법안들은 폐기됨. 자국 온라인 플랫폼 기업에 대한 과도한 규제로 인한 플랫폼 산업의 성장 및 혁신을 저해할 것을 우려가 컸음
- 그럼에도 불구하고, 구글과 아마존 등을 상대로 한 미국 법무부, 연방거래위원회 및 주 정부의 반독점 소송을 계속되고 있음

### □ 한국

- 2022년 5월 대통령 인수위원회가 플랫폼에 대한 규제 방향을 자율규제와 필요 최소한의 제도적 장치 마련으로 선회한 이래, 플랫폼 자율규제를 추진
- ‘온라인플랫폼 중개거래의 공정화에 관한 법률’은 정무위원회 계류 중
- 공정거래위원회는 「온라인 플랫폼 사업자의 시장지배적지위 남용행위 및 불공정거래행위에 대한 심사지침」 제정 및 시행 (2023.1)하였고, 「온라인 플랫폼 기업결합 심사지침」 마련 중이며, 23년 12월 기준 이른바 ‘플랫폼 공정경쟁촉진법’을 추진 중임

### (3) 플랫폼 경쟁 지표의 필요성

- 플랫폼의 경쟁이 다각적으로 사회에 미치는 영향이 확대되고 있어, 플랫폼과 플랫폼의 경쟁에 대한 이해가 필요함
  
- 플랫폼 참여자와 관련 정책결정자의 의사결정 및 정책제안에 중요함
  - 경쟁으로 인한 소비자, 제품 및 서비스 제공자, 플랫폼 자체의 혜택을 측정하여 각자의 전략을 수립하는 데 도움이 됨
  - 정책결정자가 플랫폼의 시장지배력을 감시하고, 시장의 효율성을 평가하여 적절한 규제 및 정책을 수립하는 데 필요함
  
- 한국은 자국 플랫폼 및 플랫폼 경쟁에 관한 연구가 많지 않음
  - 유럽연합 등에서는 플랫폼에 대한 조사 및 연구가 장기간 수행되어 왔고, 이를 바탕으로 경쟁법 사건들이 처리됨
    - 경쟁지표를 포함한 플랫폼 경제에 대한 측정에 대한 논의가 축적되어 왔음
    - 유럽위원회(European Commission)의 「Final Report on Measurement & Economic Indicators」는 그동안의 논의들에 대해 총체적으로 정리함. 플랫폼 경제에 대한 지표측정의 어려움을 설명하면서도 측정의 시도 및 제안을 하고 있음
  - 한국의 경우, 플랫폼 시장의 역사가 상대적으로 짧은 반면 급격한 발전을 이루었고, 국내 로컬플랫폼이 성공적으로 자리잡았음
    - 플랫폼 규제 사건 또한 최근에 집중되어 있음
    - 유럽연합은 미국의 빅테크 플랫폼 기업이 많은 분야에서 독점적 시장지배력을 행사하고 있으나, 한국의 경우 로컬플랫폼의 시장점유율이 높은 상태이고 글로벌 플랫폼과 로컬플랫폼의 경쟁이 심화되고 있음

## 나. 경쟁과 지표

### (1) 경쟁과 전통적 경쟁 지표

#### □ 경쟁의 의미

- (사전적 의미) 같은 목적에 대하여 이기거나 앞서려고 서로 겨룸
- 시장에서 가능한 한 자기에게 가장 유리한 계약을 체결하려는 모든 시장참가자들의 노력 과정(권오승·서정, 2022)
- 사업자가 상품 또는 용역의 공급 또는 수요의 관계에서 거래상대방을 획득하기 위해서 다른 사업자와 서로 경쟁하는 과정(이봉의, 2022)

#### □ 경쟁지표의 사용

- 경쟁당국이 경쟁법 및 경쟁정책을 집행하기 위해 사용함과 동시에 일반적인 시장 또는 산업에 대한 이해를 위해서도 사용함
- 경쟁법 관련 사건에서는 사건의 성격마다 증명하고자 하는 것이 달라지며, 따라서 증점적으로 사용되는 지표 또한 달라질 수 있음
- 하지만, 일반적으로 경쟁법 관련 사건에서 측정하는 시장지배력, 경쟁제한성, 효율성은 시장참여자나 경쟁을 촉진하고자 하는 정책결정자에게도 중요한 기본 개념이므로 유사한 지표를 사용하게 됨
- 경쟁은 과정으로 정의되지만, 일반적으로 경쟁을 측정하는 지표는 경쟁 구조나 결과 변수로 제시됨

#### □ 전통적인 산업조직론의 시장구조-행태-성과 패러다임(Structure-Conduct-Performance Paradigm, 이하 'SCP 패러다임')에 근간을 둔 지표들이 사용됨

- SCP 패러다임은 시장구조(시장 내 기업의 규모 및 분포와 진입장벽 등)가 시장의 경쟁 정도(완전경쟁, 과점, 독점 등)를 결정하고 이에 따라 행태와 성과가 결정되는 일방적인 선형관계를 가정함
- SCP 패러다임 초기에는 집중도와 경쟁 정도와의 부(-)의 관계, 집중도와 시장성 과와의 정(+)의 관계를 주장함

- 하지만, 점차 단순한 선형의 일방적 관계가 아닌 서로가 영향을 주고받는 상호관련성 또는 반응 효과(feedback effect)를 고려하는 모형으로 발전됨 (정갑영·김동훈·최윤정, 2021)
- 지표 : 시장구조 지표, 행태 지표, 성과지표로 나눌 수 있음

□ 시장구조 지표

- 기업 수, 규모 및 분포, 제품 동질성(또는 차별화), 진입장벽 등에 대한 측정
- 시장집중율(concentration ratio), 허쉬-허핀달 지수(Hersh-Herfindahl Index) 등을 사용함
- 집중률 등 시장구조는 기업행태와 성과에 의해 영향을 받는 등 내생성을 가지고 있어 인과관계가 불명확하나 경쟁지표로 많이 사용됨
- 시장점유율은 시장지배력이나 경쟁제한의 직접적인 지표라고 할 수 없지만, 기업결합심사에서 합병으로 인한 경쟁제한성 추정시의 출발점으로 사용되듯이 경쟁을 판단하는 출발점으로서의 의미를 가짐
- 경쟁을 측정하는 지표들 중 하나로써, 동태적인 관점에서 집중율 지표의 변화 추이를 기업들의 진입여부 및 대체성 등을 함께 살핌으로서 경쟁정도에 대한 측정이 개선됨

□ 시장 행태 지표

- 연구개발비, 제품차별화 정도 등을 사용함

□ 시장성과 지표

- 시장 또는 기업의 이윤성, 자원배분의 효율성 등에 대해 측정함
- 마진 및 이윤 관련 지표 (러너 지수, 이윤율 등), 탄력성 등

□ 전통적 지표의 한계점

- 경쟁은 과정을 의미하며, 동태적이고 혁신적인 시장일수록 '과정'을 측정하는 것이 중요해짐. 즉 동태성 및 변동성을 잡아내야 하며, 따라서 여러 지표의 변화들이 어떻게 연결되는지를 살핌으로써 경쟁 정도에 어떠한 변화가 있는지를 유추하는 것이 더 정확함

**(2) 플랫폼의 경쟁 지표**

- 플랫폼의 가치를 증진하는 핵심 요인들을 살펴봄으로써 경쟁지표를 설정함
- 특히 플랫폼 간의 대체성과 진입장벽은 플랫폼의 시장지배력과 가치 결정에 큰 영향을 주며, 따라서 경쟁 정도를 결정하는 중요 요인임

**□ 플랫폼의 특성과 경쟁지표에 대한 시사점****● 다면시장**

- 소비자, 공급자, 광고주 등 다양한 이해관계자들을 포함함
- 각기 다른 그룹을 연결하는 플랫폼 본연의 목적에서 보면, 특히 한쪽의 거래 없이는 다른 쪽 거래도 일어나지 않는 경우에는 양면을 함께 묶어 하나의 시장으로 볼 수 있음. 대표적 예로써, 미국 연방대법원이 AMEX카드 판결에 있어서 가맹점과 소비자를 연결하는 하나의 양면 신용카드 시장으로 획정한 것을 들 수 있음. 이 경우, 신용카드 시장에서의 플랫폼과 플랫폼간의 경쟁을 측정하는 지표가 필요함
- 반면, 각 그룹과의 거래(또는 비거래)가 각기 다른 시장으로 획정될 가능성도 존재함. 각 그룹마다 거래의 종류 및 이해관계가 달라지는데, 예를 들어, 신문 광고는 신문을 매개체로 구독자에 전달되지만, 신문 구독자가 광고를 위해 신문을 구독하는 것이 아님. 따라서 이렇게 한 방향으로 작동하는 경우에는 광고자와 구독자를 포함하는 시장이 아닌, 신문광고시장과 신문구독시장이 별개의 시장으로 획정될 수 있으며, 개별시장에서의 경쟁지표를 설정할 수 있음

**● 간접네트워크효과**

- 그룹간의 연결성에 의해, 한 그룹의 사용자가 늘어날수록 다른 그룹이 얻는 가치가 늘어나고 이것이 다시 원 그룹이 얻는 가치를 높여주는, 이러한 상호작용을 통해 플랫폼의 가치가 기하급수적으로 증가할 수 있음
- 온라인마켓플레이스에서 소비자가 증가하면 제품공급자가 증가하고 다시 소비자가 증가하는 순환적인 영향이 나타남
- 각 그룹의 이용자 수 및 그 변화 추이로 간접네트워크효과를 유추할 수 있음

● 진입장벽

- 데이터에 의한 진입장벽

- 데이터가 축적될수록, 이용자의 선호 및 수요에 대한 예측이 정교화되어 기존 상품이나 서비스의 이윤극대화뿐 아니라, 새로운 상품과 서비스의 출시, 합병과 인수 등 다방면에서 이점을 가지게 됨
- 데이터 축적 및 질에 대한 지표의 필요성에는 공감하지만, 개념 정의와 측정의 어려움으로 지금까지 지표가 설정되지 못했음
- 하지만 시장구조지표에 사용되는 이용자 관련 지표 및 거래액 등이 데이터양과 정(+)의 상관관계를 갖는 경우, 그리고 인공지능, 머신러닝 등을 이용하여 데이터양과 더불어 데이터의 질이 높아지는 경우에는 해당 지표들을 데이터 진입장벽의 가변수로 사용할 수 있음. 또는 해당지표들을 바탕으로 데이터에 의한 진입장벽을 유추할 수 있음 (하나의 지표가 다양한 함의를 포함함)

- 싱글 및 멀티호밍 정도, 전환비용, 의존성

- 멀티호밍이 가능한 경우, 경쟁 플랫폼의 진입이 용이해짐
- 반면 이용자들의 플랫폼간의 전환비용이 클수록, 의존성이 클수록, 그리고 다양한 이유로 탄력성이 낮은 경우, 플랫폼 간 이동이 어려워지기 때문에 경쟁 플랫폼의 진입이 어려워짐
- 싱글 및 멀티호밍을 측정하기 위해, 이용자들을 직접 서베이하거나 이용자의 플랫폼 앱 사용을 추적하는 방법 등이 사용됨. 몇몇 플랫폼 시장의 멀티호밍 여부는 정부기관, 협회, 민간 빅데이터 기업에 의해 이미 수집되고 있음

- 플랫폼이 제공하는 서비스의 범위(다양화), 번들링과 배타적 거래의 정도

- 플랫폼이 다양한 서비스를 제공하는 경우 또는 서비스들이 번들링되어 있는 경우 이용자가 하나의 플랫폼에 고착화 될 가능성이 큼

- 다양한 진입장벽에 영향을 받는 기업의 진입과 퇴출 통계는 진입장벽의 정도를 유추할 수 있는 하나의 지표임

- 이 외에도 브랜드 가치, 이용자가 얻는 가치, 만족도, 신뢰 등도 진입장벽의 역할을 함

- 시장점유율
  - 이용자의 수, 트래픽, 서비스 체류시간(consumer attention 지표) 등과 같이 이용자 관련 지표가 고전적인 가격이나 매출, 고용 변수와 함께 사용됨
  - 반면, 이용자마다 그 가치가 다르기 때문에 유럽위원회(2021)는 플랫폼의 이용자 영향력의 지표로 인터넷 트래픽, 검색량 데이터보다 더 정교한 소비자 점유율 지표를 개발할 것을 권고하고 있음
- 이윤, 마크업과 제로 가격 정책
  - 플랫폼의 경우, 양면의 수요 특성과 교차네트워크효과에 따라 비대칭적인 요금 구조를 적용할 수 있음. 이런 경우 시장지배력을 나타내는 가격과 비용과의 비교, 러너지수 등에서 가격이 0이 되므로 지수의 의미가 없어짐
- 합병과 인수
  - 플랫폼 서비스 범위를 넓혀 이용자들의 독자적인 생태계 안에 머물게 하기 때문에, 즉 싱글호밍을 유도하기 때문에, 합병과 인수를 경쟁행태를 나타내는 지표로 설정 및 추적함으로써 경쟁과의 관계를 유추해 볼 수 있음

[그림 IV-1] 전통적 경쟁 지표와 플랫폼 경쟁 지표



다. 판례 분석<sup>89)</sup>

- 본 장은 대표적인 플랫폼 관련 판례에서 사용된 경쟁 지표를 살펴보는 데 목적이 있음
- 경쟁을 측정함에 있어서 제품 및 서비스 간의 대체성과 시장의 진입장벽 판단을 하는데, 이는 각 사례의 시장 획정 단계에서부터 시장지배력 평가 및 경쟁제한성 평가 단계에 이르기까지 고르게 고려되고 있음. 따라서 각 사례마다 사례 개요에서 시작하여 시장획정, 시장지배력과 경쟁제한성 추정시 사용되는 지표들을 살펴봄
- 본 장은 세 개의 사례를 정리 및 분석하였음: 첫 번째 사례인 배달의 민족과 요기요의 결합 건은 플랫폼의 결합신청 건으로 경제 분석이 집중적으로 사용되어 다양한 경쟁 지표들을 추정하였다는 점에서, 두 번째 사례인 네이버 쇼핑 건은 한국에서 플랫폼의 시장지배력을 판단한 첫 사례였다는 점에서 세 번째 사례인 유럽연합의 구글 쇼핑 건은 네이버 쇼핑 건과의 비교검토를 위해 선정하였음

[그림 IV-2] 플랫폼 관련 사례에서 사용된 지표

	배민·요기요	네이버 쇼핑	구글 비교쇼핑
시장	배달앱 서비스 시장	온라인 비교쇼핑 시장 온라인 오픈마켓 시장	온라인 비교쇼핑 시장 온라인 일반검색 시장
시장구조 변수	거래액, 배달앱 매출액 (광고료, 수수료 등), 앱스토어 다운로드 횟수	순이용자수, 페이지 뷰 등의 트래픽 변수, 영업수입, 거래액 등	검색연진 점유율 (질문 수, 사용자수, 페이지 뷰, 세션 수 등 불특기반 시장점유율)
경쟁 및 진입장벽 변수	멀티호밍, 매출의존도, 매출액 추정: 가격인상 가능성, 음식점의 이탈확률, 진입기업에 대한 민감도 등	상품 노출 점유율 오픈마켓 사라가기 점유율 거래액/매출액 점유율 등	일반검색에서 비교쇼핑으로의 트래픽, 멀티호밍, 진입/퇴출데이터, 투자금액 등
동태성	지표의 변화 추이		
데이터 출처	플랫폼 제공 데이터 (주문, 소비자/입점업체 관련 등), 설문조사, 재무제표, 빅데이터 기업 데이터, 경제분석 보고서 등		
출처: 공정거래위원회와 유럽위원회의 의결서 및 보도자료 등 공개 자료를 바탕으로 정리함			

89) 공정거래위원회 및 유럽위원회의 의결서와 보도자료 등을 참조함

**(1) 국내사례: 딜리버리히어로 에스이 등 4개 배달앱 사업자의 기업결합 제한규정 위반행위에 대한 건 (2020)****□ 사례에 대한 개요**

- 요기요 배달앱(이하 ‘요기요’) 등을 운영하는 딜리버리히어로 에스이(Delivery Hero SE)가 배달의민족 배달앱(이하 ‘배민’) 등을 운영하는 (주)우아한형제들의 주식 약 88%를 취득하는 기업결합을 신고함<sup>90)</sup>
- 배달앱 간의 결합은 배달앱 시장에서 거래액 기준으로 결합회사의 99.2% 시장 점유율과 구글 플레이스토어 다운로드 횟수 기준 93.1% 시장점유율로 이어짐
- P2P 측면에서의 경쟁자 인수, 플랫폼-음식점 측면에서의 수수료 문제, 플랫폼-소비자 측면에서의 노출순위 조정 등 온라인 중개 서비스의 3가지 주요 쟁점이 모두 포함된 사건임(공정거래위원회 보도자료)
- 공정거래위원회(이하 ‘공정위’)는 두 회사의 결합시 시장점유율이 99.2% DH의 국내계열회사인 딜리버리히어로코리아(이하 ‘DHK’) 지분을 전부 매각하는 조건 하에 결합을 승인함

**□ 시장획정과 시장지배력 평가**

- 공정위는 관련 상품 및 지역시장을 배달앱 서비스 및 전국 시장으로 획정함
  - 관련 상품시장은 배달앱 서비스와 전화주문, 프랜차이즈 자체 웹/앱 주문 서비스 등을 기능 및 효용의 유사성, 대체 관계 등에서 비교분석하여 획정함
    - 관련 상품시장을 획정하는데 다양한 기준을 복합적으로 사용함: 상품의 가격, 기능 및 효용의 유사성, 수요자 및 공급자(관련사업자) 측면에서의 대체성과 대체가능성에 대한 인식, 공급자의 시장인식 및 경영의사결정, 한국표준산업분류 등을 고려함(공정위 심사기준). 이 중 가장 중요한 요소는 수요대체성과 그로 인한 잠재적 경쟁압력이라고 할 수 있음
  - 주문서비스 간의 대체성을 판단하기 위해, 소비자 및 음식점 측면에서 SSNIP(Small but Significant and Non-transitory Increase in Price)테스트에 기초한 임계매출손실분석(Critical Loss Analysis)과 총전환율분석(Aggregate Diversion Analysis)을 시행함<sup>91)</sup>

90) 요기요와 배민의 수평결합 뿐 아니라, 음식배달대행 및 공유주방 서비스 등의 결합도 이루어지나, 본 판례 정리에서는 배달앱 서비스 시장에 초점을 맞춤

91) 자세한 배경 및 방법론은 원세범·이상현(2023) 참조

- 임계매출손실분석에서, 합병대상인 배민 및 요기요가 (가상의) 작지만 의미 있으며 비일시적인 5% 또는 10%의 가격을 인상(소비자 쿠폰 할인 감소)했을 때 실제가격탄력성이 임계가격탄력성보다 작았으며, 음식점 측면에서 플랫폼 수수료이 인상되는 경우 실제마진율이 임계마진율보다 낮아 배민 및 요기요를 포함하는 배달앱으로 확정되었음. 즉, 분석결과 쿠폰할인이 줄어들었을 때 타 경쟁플랫폼으로의 소비자의 이탈률이 그리 높지 않고, 플랫폼 수수료의 증가시 음식점의 매출 및 이윤감소가 그리 크지 않아 타 플랫폼으로의 이탈이 많지 않았음. 따라서 배민 및 요기요의 플랫폼과 대체(또는 경쟁)관계인 플랫폼이 없음을 의미함. 총전환율분석에서도 동일한 결과를 얻음
- 필요자료: 실제매출감소율 추정을 위한 설문조사(배달앱의 수수료 인상시 음식점 이탈 혹은 광고 감소로 인해 배달앱 거래금액이 감소하는 비율과 음식점이 수수료 인상분을 소비자에게 전가하여 발생하는 소비자 주문 감소율 필요), 마진을 계산을 위한 재무제표 자료 및 비용추정회귀분석에 사용되는 비용과 주문 데이터, 소비자의 가격탄력성을 추정하기 위한 서베이 및 실험, 결합회사의 주문자료 등

● 시장획정 논쟁점<sup>92)</sup>

- 결합 당사회사는 전화주문, 프랜차이즈 자체 앱/웹 등의 직접 주문 수단을 경쟁 대상으로 포함하여 관련 상품시장을 보다 넓게 획정할 것을 주장함
- 배민 및 요기요는 소비자측면에서 배달앱과 직접 주문 등이 함께 사용되어 대체 관계이며 음식점 측면에서 배달앱의 사용 음식점 비중이 상대적으로 적다는 현황 및 각 주문수단의 기능 수렴화로 인한 음식점 측면에서의 대체성이 강화되었음 등을 이유로 제시함
- 그 외에도 일반적으로 경쟁당국과 사건 관련 회사는 모두 시장획정을 위한 경제 분석을 시도하는데, 사용데이터, 변수의 형태 및 선정 등에서 차이가 나는 경우 결론이 달라짐
- 상대적으로 타 사건에 비해 결합 사건일 경우에 공정위 측의 시장획정 경제 분석의 내용이 의결서에 상세히 담기는 경향이 있는 반면, 결합당사회사의 주장과 이에 대한 경쟁당국의 반박은 담겨있으나 결합당사회사의 경제 분석에 대한 설명은 상대적으로 상세하지 않은 경향이 있음

92) 시장획정에 따라, 각 회사의 시장점유율에 대한 평가가 달라지고 결합승인여부에 영향을 미칠 수 있음. 시장이 넓게 획정될수록 결합 당사회사의 합계 점유율이 낮아지므로, 결합 당사회사는 넓은 시장획정을 선호함

- 반면, 공정위는 시장획정시 배달앱 이용의 급격한 증가와 배달앱과 직접주문 수단 간의 기능 및 효용의 차이, 대체재가 아닌 보조적 관계, 소비자의 쿠폰할인에 대한 낮은 탄력성, 음식점의 배달앱 의존도 등을 고려했음을 설명함

● 시장지배력과 시장구조의 측정지표

- 결합사건에서는 회사들의 시장점유율이 특정 기준을 충족하면 실질적으로 경쟁을 제한하는 것으로 추정함<sup>93)</sup>
- 배달앱 서비스 시장의 시장점유율은 거래금액으로 기준으로 산출함
  - 19년도 거래금액기준으로 결합당사회사의 점유율 합계가 99.2%임
  - 배달앱의 매출액(음식점에 부과하는 광고료 및 중개수수료 등) 기반 보다는 플랫폼 양면시장의 특성을 반영하여 소비자와 음식점 간의 음식주문 거래금액(배달비 등을 제외한 수수료 음식 관련 거래금액(부가가치세 포함))을 기준으로 산출함
  - 추가적으로 수수료 매출 기준, 소비자측면에서 배달앱 순접속자 수, 음식점측면에서의 등록음식점 수의 점유율 등을 고려함
- 시장집중도와 시장 내의 기업 간 시장점유율 격차의 변화 추이
  - 결합당사회사 점유율의 유지 및 상위 3개사의 순위 변화가 없이 유지됨을 확인함

□ 경쟁제한성 평가 I: 시장집중도와 싱글/멀티호밍 및 고착화와 경쟁압력

● 시장집중도 및 변화 추이

- 배민과 요기요의 결합과 같은 수평형 기업결합에서의 경쟁제한성 여부 판단은, 기업결합 전후의 시장집중 상황, 결합당사회사 단독의 경쟁제한 가능성, 경쟁사업자간의 공동행위 가능성, 신규진입의 가능성, 유사품 및 인접시장의 존재 여부 등을 종합적으로 고려하여 심사함(공정위 기업결합 심사기준)
- 시장집중도 및 변화 추이는 기업결합이 경쟁에 미치는 영향 분석의 출발점임

● 경쟁 플랫폼 간의 대체성: 싱글/멀티호밍 현황과 간접네트워크 효과의 크기로 추정함

93) 독점규제 및 공정거래에 관한 법률 제9조 제3항 참조

- 서베이를 통해 멀티호밍 비중 및 구매전환 가능성(특정 배달앱을 사용하지 못할 경우 차선으로 사용하는 배달앱 등)을 직접 조사하여, 소비자들의 구매전환이 대부분 결합당사회사인 요기요와 배민사이에서 발생하여 결합 시 타 플랫폼으로부터의 경쟁압력이 낮음을 제시함
- 플랫폼의 이익을 극대화함에 있어, 간접네트워크효과와 그 크기가 중요함. 거래집단(소비자와 음식점)의 규모 및 이들의 멀티호밍 정도(또는 고착화 정도)는 간접네트워크효과의 크기에 중요한 역할을 함. 즉, 플랫폼이 독점적으로 보유하고 있는 거래집단의 규모가 클수록 간접네트워크효과가 커지고 시장지배력 또한 증가하며, 경쟁 압력이 낮아짐
- 자료: 직접적으로 싱글 및 멀티호밍에 대한 질문, 각 회사간의 대체성에 관한 질문을 담은 서베이 자료 등

□ 경쟁제한성 평가 II: 가격인상가능성

- 경쟁 압력의 정량화 시도: 가격인상압력과 이탈확률 추정으로부터의 가격인상 가능성을 유추함
  - 결합당사회사가 가격을 얼마나 인상할 수 있는지를 측정함. 즉, 결합으로 인해 경쟁압력이 줄어 어느 정도 수준 이상의 가격을 인상할 수 있는지를 살펴봄
  - 소비자 측면에 대해 쿠폰할인 감소(즉 가격 인상)가 가능함을 제시함
    - 방법론: 소비자 측면에서 결합당사회사간의 총전환율(Gross Upward Pricing Pressure Index)를 도출하여 10% 가격 인상시의 기준점보다 높은 값을 가지는 것을 보였으며 이는 적어도 해당 가격만큼 인상할 요인이 있음을 의미함
    - 총전환율(GUPPI) 도출을 위해서는, 회귀분석을 통해 자체가격탄력성과 교차탄력성을 추정하고 이를 바탕으로 회사 간 전환율을 계산함. 가격탄력성의 추정에는 주문량, 가격 뿐 아니라, 음식점의 특성 및 음식점이 운영중인 지역의 특성(지역의 경쟁정도 및 거식적 특성 등) 등을 고려함
  - 음식점 측면에서 배민 및 요기요의 수수료 동시 인상 가능성을 제시함
    - 음식점 측면에서는 임계매출감소분석과 수수료 인상시의 음식점 이탈확률을 추정함. 이 경우, 배민과 요기요를 통한 음식점 매출액, 수수료 금액 등을 이용하였고, 수수료 인상시의 이탈률이 매우 미미한 것으로 나타났음을 통해 수수료 인상 유인이 있는 것으로 판단함

- 배달앱 매출의존도, 싱글/멀티호밍 매출액 비중도 고려함
- 자료: 결합 당사회사의 실제 주문 데이터와 소비자 대상 설문조사 등
- 회귀분석을 포함하여 경제 분석을 통해 다양한 경쟁압력을 추정함
- 진입기업(쿠팡이츠)의 진입에 따른 민감도를 통해 진입기업의 경쟁압력을 추정할 수 있음
- 쿠폰 할인 등과 지역 점유율 간의 상관관계를 통해 점유율이 낮은 지역 일수록 주문 건당 할인금액이 높은 것을 발견함

#### □ 경쟁제한성 평가 III: 비가격경쟁 측면의 정보자산<sup>94)</sup>

- 본 사건의 정보자산인 소비자정보, 음식점정보, 거래정보, 배달정보를 바탕으로, 결합회사는 음식점들에게 통계 제공 및 활용 교육서비스 등을 제공하여왔음
- 정보자산의 격차로 인해 마케팅 등에서 경쟁사의 경쟁압력이 낮고, 심지어 경쟁자를 퇴출시킬 수도 있음
- 정보자산을 바탕으로 시장지배력을 강화하는 경우, 경쟁이 제한될 수 있음
- 통계 및 음식점에 대한 교육은 상당한 수준의 비용이 소요될 것으로 추정되는바, 경쟁압력이 적어지면 이에 대한 지출이 줄고, 정보자산의 품질이 저하될 가능성이 큼. 즉, 정보제공의 비가격경쟁이 저해될 우려가 있음

#### □ 경쟁제한성 평가 IV: 진입과 효율성

- 공정위는 배달앱 시장에 법적, 제도적 진입장벽은 존재하지 않으나, 투자자금의 필요성 및 간접 네트워크 효과로 인해 진입이 용이하지는 않을 것으로 판단한 반면 결합대상회사는 자유로운 진입퇴출, 비즈니스 모델의 변동가능성 등의 동태적 특성을 고려시 경쟁제한성이 낮을 것으로 판단함
- 효율성

94) 기업결합 후 결합당사회사가 정보자산을 활용하여 시장지배력을 형성·강화·유지하는 경우 관련시장에서의 경쟁이 실질적으로 제한될 가능성이 있으므로, ① 결합을 통하여 얻게 되는 정보자산이 다른 방법으로는 이를 대체하기 곤란한지 여부, ② 해당 결합으로 인하여 결합당사회사가 경쟁사업자의 정보자산 접근을 제한할 유인 및 능력이 증가하는지 여부, ③ 결합 이후 정보자산 접근 제한 등으로 인하여 경쟁에 부정적 효과가 발생할 것이 예상되는지 여부, ④ 결합당사회사가 정보자산의 수집·관리·분석·활용 등과 관련한 서비스의 품질을 저하시키는 등 비가격 경쟁을 저해할 가능성이 높아지는지 여부 등을 추가로 고려하도록 함 (기업결합 심사기준)

- 결합당사회사는 배달원 당 주문밀도의 증가로 인해, 배달거리 및 시간의 단축과 소비자 만족도 증가 그리고 소비자 주문 증가로 인해 음식점과 배달원의 수익증가로 이어지는 선순환을 주장함
- 공정위는 이러한 효율성 증대효과가 반드시 결합만이 아닌 각 기업의 사업확장 등에 의해 달성될 수 있음을 주장함

□ 자료출처

- 경쟁법 사건의 경우 일반적으로 수집하기 어려운 기업의 영업비밀을 포함하여 사건 관련자들로부터 직접적으로 상세하게 자료의 수집이 가능하며, 상세한 데이터를 이용한 전문가의 경제 분석 결과가 추가됨. 따라서 경쟁상황 및 변화에 대한 심도 있는 분석이 가능해짐
- 각사 제공 주문 관련 자료, 소비자와 음식점에 대한 서베이, 유로모니터 리포트(배달외식, 포장외식의 시장규모), 통계청의 외식업 경영실태조사, 농수산물공사의 외식흐름조사, 농림축산식품부와 한국농수산물유통공사가 공동으로 발간한 「2019 국내 외식트렌드 조사보고서」각사의 서베이(경쟁 모니터링 하기 위한, 배달음식 주문 방법 등 서베이), BIT(Brand Index Tracking) 설문조사, 앱 애니 데이터(앱 다운로드 수), 코리안클릭 데이터(각 배달앱의 월별 순 접속자) 등

(2) 국내 사례: 네이버(주)[쇼핑 부문]의 시장지배적지위 남용행위 등에 대한 건(2020)

□ 사례에 대한 개요

- 공정위가 네이버가 네이버쇼핑 검색결과에서 자사 스마트스토어를 이용하는 사업자의 상품이 상위에 노출되기 유리하도록 검색알고리즘<sup>95)</sup>을 설계 및 적용한 행위를 했다고 판단하고 규제를 함
- 플랫폼 기업의 자사우대행위를 규제한 최초의 사례로, 공정위는 온라인비교쇼핑 서비스 시장에서 시장지배적 지위를 가진 네이버가 검색알고리즘을 조정함으로써 시장지배력 전이를 통해 오픈마켓서비스 시장에서의 경쟁을 제한했다고 판단하고, 위반행위의 시정조치와 265억의 과징금 부과함

95) 검색(랭킹)알고리즘의 노출순위 결정: 관련성(적합도, 인기도, 신뢰도의 3가지)과 다양성 등 검색품질을 고려한 요소를 반영함

## □ 시장획정과 시장지배력 평가

- 공정위는 관련 상품시장을 온라인 비교쇼핑서비스 시장과 오픈마켓서비스 시장으로 획정함
- 시장획정 논쟁점
  - 네이버 측은 “플랫폼 커머스 시장”, 즉 오픈마켓, 소셜커머스, 종합몰 등과 온라인 비교쇼핑 서비스를 하나의 시장으로 봐야 한다는 주장으로, 이는 해당 서비스들의 유사도가 증가하고, 소비자 측면에서 대체성이 높으며, 공급자 측면에서 서로를 경쟁사업자로 인식하고 있다는 이유를 제시함
  - 반면, 공정위는 두 서비스 자체가 상이하여 수요대체성 및 공급대체성이 높지 않으며, 서로를 대체관계(즉 하나의 시장)가 아닌 보완재 (또는 제휴관계)로 인식하고 있다고 판단함
  - 결합사건의 시장획정에 대한 경제분석은 공정위 의결서에 상세히 다루어지나, 타 사건에서는 SSNIP 테스트의 내용이나 결과가 제시되지 않는 경우가 많음. 기준가격, 마진, 모형에 따라 시장획정도 달라질 가능성이 존재하지만, 본 사건의 의결서에는 제시되지 않았음
  - 예를 들어 양면시장인 플랫폼의 특성으로 인해 양면가격의 합을 기준가격으로 할 것인지 아니면 배달앱 서비스 시장의 획정에서 소비자 측면의 쿠폰할인과 음식점 측면의 거래수수료를 기준으로 한 것처럼 각각의 그룹과 플랫폼 간의 거래가격을 기준가격으로 할지에 따라 대체성의 정도가 달라짐. 대체성이 낮으면 경쟁상대가 없는 것으로 간주하여 시장이 좁게 획정되며 대체성이 높으면 경쟁상대를 포함하기 위해 시장이 넓게 획정되기 때문에 시장획정 또한 달라질 수 있음
- 시장지배력과 시장구조의 측정지표와 출처
  - 시장지배력(또는 시장지배적 지위)은 시장성과 기준인 가격이 경쟁가격보다 얼마나 높은지로 판단하지만 가상의 경쟁가격을 추정해야 하는 어려움이 있음. 따라서 이론적으로 정의 상관관계를 갖는 시장점유율 수치를 사용해 간접적으로 시장지배력을 유추함. 그 외에도 경쟁사업자들의 상대적 규모, 진입장벽의 존재 및 정도, 경쟁사업자간의 공동행위 가능성, 유사품 및 인접시장의 존재, 시장봉쇄력, 자금력 등을 종합적으로 고려함(시장지배적 지위 남용행위 심사 기준)

- 비교쇼핑서비스시장의 시장점유율은 전통적인 영업수익<sup>96)</sup>과 거래액을 사용하였으며, 페이지뷰(PV), 순이용자수(UV:Unique Visitor)<sup>97)</sup> 등 플랫폼 트래픽 변수를 사용하여 계산함. 시장점유율이 미미한 서비스는 제외하고, 주요 4개 서비스(네이버쇼핑, 다음쇼핑하우, 다나와, 에누리)를 대상으로 함
- 진입장벽 여부는 플랫폼이 가지는 간접네트워크효과와 규모의 경제, 검색서비스 시장의 영향력 등으로 판단할 수 있는데, 공정위는 네이버쇼핑의 시장점유율 기준으로 진입장벽이 상당히 높다고 판단함
- 오픈마켓 시장점유율은 거래액을 사용하였으며, 주요 6개서비스(11번가, G마켓, 옥션, 인터파크, 네이버, 쿠팡)를 대상으로 집계함

□ 경쟁제한성 평가: 경쟁제한적 우려 내지 효과

- 시장의 경쟁과 변화정도를 측정함에 있어 시장지배적지위 남용 사건에서는 위반행위가 경쟁을 제한하거나 왜곡하는 효과를 나타내었는지를 판단하기 위해 반면, 기업과 정부는 각기 전략이나 정책을 수립하거나 개선하기 위해 사용한다는 점에서 그 목적과 관점이 상이함. 본 사건에서는 비교쇼핑서비스 검색결과에서 자사의 오픈마켓서비스의 우선노출이 가져오는 경쟁제한성을 측정하고자 함
- 경쟁 지표: 경쟁법 관련 사건 또는 일반적인 경쟁 상황을 평가하기 위한 경쟁지표는 동일하며, 가격, 산출량, 혁신, 유력한 경쟁사업자 수, 다양성 등의 지표를 사용함 (시장지배적 지위 남용행위 심사지침). 경쟁지표의 수치 변화에서 경쟁의 증가 또는 감소를 포착할 수 있음
- 공정위가 사용한 경쟁지표<sup>98)</sup>
  - 의존도: 주요 오픈마켓의 방문경로별(비교쇼핑사이트: 네이버쇼핑, 다나와, 다음, 에누리, 직접유입) 거래액 비중을 통해, 오픈마켓에 있어 네이버쇼핑의 의존도가 적지 않아 포기할 수 없는 수단임을 제시함
  - 상품노출점유율의 변화: 네이버쇼핑 검색결과에서 스마트스토어 및 오픈마켓 사업자별 상품 노출점유율 변화를 측정함

96) 비교쇼핑서비스를 제공하고, CPC(Cost Per Click), CPS(Click Per Sales)방식으로 수수하는 고정비 및 수수료, 광고비등을 포함하는 금액

97) 인터넷 특정 웹사이트의 이용자 수를 집계하는 지표로, 예를 들어 어떤 이용자가 하루에 네이버 쇼핑과 네이버 부동산을 각각 5번 접속한 경우에 순 이용자 수는1명으로 집계됨

98) 공정거래위원회의 의결서에는 경쟁지표로 사용된 지표는 제시되지만, 수치는 노출되지 않음. 따라서 각 지표의 수치 변화에 대한 부분은 공정위 의결서의 내용을 참조함

- 검색결과 상단에 스마트스토어 노출이 증가함은 타 오픈마켓 노출이 감소함을 의미함. 검색결과 첫 페이지 이후 ‘페이지뷰’와 ‘상품 클릭수’는 현저하게 떨어지며 이는 상품 거래액에 영향을 미침
- 검색알고리즘 조정에 의한 비교쇼핑검색결과가 오픈마켓시장의 경쟁을 왜곡시키는지를 판단함
- 오픈마켓 사라지기 점유율 변화: 2015년 알고리즘 변경 전후로 자사오픈마켓 사라지기 점유율의 급격한 증가와 오픈마켓 4개 사업자의 합계 점유율의 감소 관찰됨
- 거래액과 매출기준 점유율 변화: 네이버쇼핑을 통한 주요 4사 오픈마켓과 스마트스토어 거래액, “스마트스토어 대비”(상대적) 거래액 변화 및 오픈마켓의 시장점유율을 해당 기간동안 제시하여 변화 추이가 관찰됨
- 이외에도 진입장벽으로 기능할 수 있는 일반검색의 영향력을 검색서비스 사업자의 월평균 방문자 추이, 검색회수 점유율(PC, 모바일), 비교쇼핑서비스사의 매출액 및 당기순이익의 비교를 통해 파악
- 공정위의 경쟁지표에 대한 비판과 대안
  - 김성환(2022)은 경쟁제한성을 평가함에 있어서, 경쟁 오픈마켓간의 거래액 점유율 변화, 즉 스마트스토어의 점유율 증가와 타오픈마켓의 점유율 감소가 아닌 전체적인 거래액의 증가와 허쉬-허핀달 지수 등을 이용한 분포를 보아야 한다고 비판하였음
  - 공정위 의결서는 수치가 노출되지 않지만 거래액의 변화 자료 테이블을 제시하고 있어서, 위에서 지적된 수치들을 구할 수 있을 것으로 고려됨
  - 김성만송태원(2021)은 네이버페이 등 의한 소비자 규모의 증가 및 판매자 규모의 증가 등 간접네트워크효과로 인한 자연적인 네이버 스마트스토어의 거래액 증가 부분이 경쟁제한적 행위로 인한 거래액 증가로 취급되는 오류를 비판하였음

#### □ 자료출처

- 본 사건에서 사용된 자료는 플랫폼의 양면시장과 인접시장과의 긴밀도 등으로 인해 다양한 그룹의 데이터가 수집되었음. 소비자의 가격비교서비스 페이지뷰, 순방문횟수, 총체류시간, 플랫폼의 수익구조, 일별/월별 거래액, 등록상품 수, 입점사업자 수, 오픈마켓별 거래액, 등록상품 수, 입점사업자 수, 오픈마켓 카테고리

고리별 노출점유율, 클릭수, 노출수, 클릭률(CRT, Click through rate) 티넥스  
(주) 온라인 및 모바일 쇼핑 태도 및 행동에 관한 소비자 인식 조사, 코리아클릭의  
온라인 이용자 행태조사, 경제분석보고서 등을 포함함

### (3) 해외 사례: 유럽위원회의 구글 쇼핑 사건

#### □ 사례에 대한 개요

- 구글(Google)이 온라인검색시장의 독점적 지위를 남용하여 유럽 13개 국가 일반검색(Google Search) 결과에서 자사 전문비교검색사이트인 구글쇼핑(Google Shopping)을 상대적으로 유리하게 전시하고, 상단에 노출하는 자사우대를 하여 유럽위원회(European Commission)의 규제를 받는 사건
- 비교쇼핑서비스 사이트로의 트래픽에서의 구글 일반검색의 중요성, 사용자들이 첫 소수의 검색결과에 집중하는 행태, 구글쇼핑사이트로 유도된 트래픽이 고착되어 다시 일반검색으로 돌아오지 않으며, 다른 소스로부터의 대체도 불가능함을 고려하여 경쟁을 저하시켰다는 판단을 하고, 유럽위원회는 구글에 대해 24억 2천만 유로 과징금 부과함. 이는 일반법원(General Court)에 의해서도 유지됨

#### □ 시장획정과 시장지배력 평가

- 유럽위원회는 관련 상품시장으로 온라인 일반검색서비스 시장과 온라인 비교쇼핑서비스 시장으로 획정함
  - 일반검색서비스와 비교쇼핑서비스 간의 대체성이 제한되어 있음 : 비교쇼핑서비스는 특정 분야만을 다루며, 전문검색서비스를 제공하는 기업들과 일반검색서비스를 제공하는 기업들은 공급대체성이 낮아 대개 구분되어 있으며, 구글 스스로가 두 서비스를 구분하고 있음
  - 비교쇼핑서비스와 타전문검색서비스, 온라인검색광고플랫폼, 온라인마켓플레이스 플랫폼 간의 대체성이 낮음
- 시장획정 논쟁점
  - 구글은 일반검색서비스 시장 획정에 대해서는 이의를 제기하지 않았으나, 비교쇼핑서비스 시장 획정에 대해서는 온라인마켓플레이스 플랫폼도 비교쇼핑서비

스시장에 포함되어야 함을 주장함: 모두 상품검색기능을 인터넷이용자들에게 무료로 제공하며, 구글의 서베이에 따르면 소비자들이 아마존을 비교쇼핑서비스의 대체재로 인식하며, 다양한 연구에서 인터넷이용자들이 온라인마켓플레이스 플랫폼에서 상품검색을 한 후 가격비교를 통해 구매를 하며, 구글이 아마존과 이베이를 비교쇼핑서비스 시장에서의 리더로 고려함을 제시하였고, 유럽위원회가 SSNIP테스트를 했었어야 한다고 주장함

- 위원회는 인터넷 이용자와 온라인 공급자의 관점에서의 수요 특성을 분석하여 대체성이 낮음을 주장하였으며, 일반법원에서도 받아들여짐. 또한 SSNIP테스트를 할 의무가 없으며, 다른 방법들로도 시장획정이 가능하다고 설명함

#### ● 시장지배적 지위와 시장구조의 측정지표와 출처

- 측정지표: 검색엔진의 글로벌 및 유럽국가들에서의 평균 월간 시장 점유율 (데스크탑과 모바일) : 질문 수, 사용자 수, 페이지 뷰 수 또는 세션 수 등의 볼륨 기반 시장점유율 이음함
  - 볼륨 기반 시장점유율 사용이유: 인터넷 사용자에게 무료로 제공되는 서비스로 가치에 기반한 시장점유율을 구하기 힘들며, 일반검색서비스의 'Revenue Per Search(RPS)'에 대한 데이터를 얻을 수 없었으며 광고주들이 점유율을 고려함
- 진입장벽(투자금액, 네트워크효과, 진입/퇴출 데이터), 멀티호밍과 브랜드 효과(서베이) 등 다양한 요인이 고려되었음
- 출처: NetMarketShare, Nielsen Data, StatCounter, ComeScoe, 구글이 제출한 서베이, Forrester Research 등과 같은 기관들의 온라인 서베이 및 연구 등

#### □ 경쟁제한성 평가: 경쟁제한적 우려 내지 효과

- 유럽위원회는 일반검색에서 검색순위(비교쇼핑서비스 링크의 위치)가 클릭률, 그리고 이에 의한 트래픽 증감에 큰 영향을 미침을 설명함
  - 구글의 일반검색 첫 페이지 결과(상위 10순위까지)가 전체 클릭 수의 95%를 받고, 상위 5개가 전체 클릭 수의 76.65%를 받는 통계를 제시함
  - 상품의 순위가 올라갈수록 소비자의 클릭 확률, 지불 의향이나 수요가 증가하는 인과관계를 보인 학술적인 연구 결과들 또한 존재함(김민정, 2019 )

- 구글 일반검색으로부터의 트래픽이 비교쇼핑서비스들의 트래픽의 많은 부분을 차지하기 때문에 구글로의 트래픽 증감이 비교쇼핑서비스 시장의 경쟁에 지대한 영향을 미침을 설명하며, 트래픽 증감을 경쟁지표로 사용함

● 경쟁지표

- 알고리즘 변경 후의 비교 쇼핑사이트로의 트래픽 증감이 주요 지표로 사용됨 : 트래픽 규모에 대한 다양한 실증적 자료를 제공하였으며, 구글 쇼핑으로의 통화량 증가와 경쟁 비교 쇼핑사이트로의 트래픽 감소를 제시함
- 경쟁 비교 쇼핑사이트로의 전체 트래픽 중 구글의 일반검색으로부터의 트래픽 비율: 구글 일반검색으로부터의 트래픽이 다른 소스로부터 트래픽에 의해 효과적으로 대체될 수 없는 의존도가 있음을 제시함

● 경쟁지표에 대한 비판

- 구글은 자사 서비스의 품질개선을 위한 노력과 경쟁이었으며, 유럽위원회가 이러한 노력과 경쟁으로부터의 행위가 아닌 어떠한 반경쟁적인 행위 및 특징을 식별하지 못했다고 주장함. 구글이 Product Universals와 shopping Units가 적용되지 않았을 때의 구글검색 트래픽을 보여주는 Counterfactual analysis가 필요하다고 주장함
- 유럽위원회는 EU 경쟁법(TFEU) 제 102조(남용행위의 위법성 인정)에 따라 경쟁제한효과에 대한 직접적 증명이나 경쟁제한의 의도 목적의 증명을 요하지 않으며, 단지 경쟁제한의 우려에 대한 개연성을 증명, 잠재적 경쟁저하를 증명하는 것만으로 충분하다고 판단하였음. 일반법원(General Court) 또한 그런 분석이 불가능할 수 있으며, 행위와 경쟁변화 간의 상관관계 및 추가적인 정보가 인과관계를 설명할 수 있다고 판단했음

□ 자료출처

- 구글 제공 데이터와 서류, 여러 기관의 클릭데이터와 연구자료

## 라. 모니터링 관점에서의 경쟁 지표와 가용데이터의 한계점과 보완점

### (1) 경쟁 측정지표 현황 및 출처

#### □ 정부 및 공공기관의 통계

##### ● 국내 통계

- 통계청 「온라인쇼핑 동향 조사」는 1,100여개의 사업체를 대상으로 상품군별, 매체별 온라인쇼핑 거래액을 제공하며, 「서비스 조사」는 약 250,000개 사업체를 대상으로 디지털플랫폼거래여부, 온라인거래여부 등의 통계를 제공함
- 과학기술정보통신부의 「인터넷이용실태조사」는 개인들의 다양한 종류의 인터넷 이용(검색, 포털사이트, 동영상 서비스, 온라인 미디어 서비스, 이메일, 메신저 등)에 대한 빈도와 인터넷 쇼핑 빈도 및 비용 등을 제공함
- 방송통신위원회의 「경쟁상황평가」는 OTT이용자들의 온라인 리뷰에 대한 텍스트마이닝분석과 온라인 설문조사를 통해 OTT이용자의 멀티호밍 여부 등에 대한 정보를 제공함
- 고용노동부와 한국고용정보원의 「플랫폼종사자 규모와 근무실태 조사」는 플랫폼 종사자의 규모 및 유형, 근무실태 등을 제공함
- 농림축산식품부와 한국농촌경제연구원의 「외식업체경영실태조사」는 배달앱과 배달대행 관련 사용여부, 건수 및 비용 등을 제공함
- 공정거래위원회의 「클라우드 서비스 분야 실태조사」(2022)는 클라우드사 및 거래관계가 있는 관련 기업들에 대한 설문조사를 통해 시장점유율, 클라우드사에 대한 거래의존도, 경쟁사 전환 및 멀티호밍, 이동시의 제약사항 등을 조사함

##### ● 해외 통계

- 유럽연합 집행위원회의 사물인터넷('20.7~'22.1) 실태조사
- 영국 경쟁당국의 모바일 생태계(2021) 실태조사
- Australian Competition & Consumer Commission (ACCC) 서베이는 라우저, 검색, 소셜미디어 등의 플랫폼에서의 시장점유율과 소비자 행태 등을 실태 조사
- 유럽연합의 Eurostat 제공하는 4개 숙박예약 플랫폼이 제공하는 통계

※ 4개의 플랫폼(Airbnb, Booking, Expedia and Tripadvisor)으로부터 숙박일수와 숙박자수에 대한 데이터를 직접 정기적으로 받아, Eurostat의 확인 후 지역, 도시별 통계를 공개함

※ 'The 2020 non-disclosure agreement' 하에서 정보 제공을 하고 있으며, 플랫폼으로부터 직접적으로 제공 받은 신뢰성 있는 통계라는 장점을 가짐

- 유럽연합의 노동서베이에서 디지털 플랫폼 노동자들에 대한 파일럿 서베이 진행(2022)
- U.S. Census Bureau는 이커머스 매출과 추이 데이터를 제공함

#### □ 경쟁법 관련 사건의 통계

##### ● 출처

- 해당 기업과 관련 경쟁기업으로부터 직접적으로 데이터를 받음
- 해당 사건 관련하여 필요한 데이터를 얻기 위해, 특정 목적을 가지고 서베이 또는 조사를 실시함
- 정부기관의 통계, 민간빅데이터기업의 통계 등도 사용함
- 다양한 데이터를 바탕으로 회귀분석을 포함하여 경제분석을 실시함

##### ● 사용지표 : 제3장 참조

#### □ 민간 빅데이터 기업이 제공하는 지표

- 많은 민간 빅데이터 기업이 다양한 통계를 제공하고 있으며, 최근에는 플랫폼 관련 데이터와 통계를 제공하고 있음
- 민간 빅데이터 기업들은 자신이 직접 개발 및 운영하는 앱, 데이터 제공을 허용한 파트너, 고객등으로부터 데이터에 대한 동의를 받고 자신만의 네트워크와 서베이대상 등을 보유하여 이들로부터 통계를 측정함
  - 예를 들어, 빅데이터플랫폼 아이지에이웍스 모바일인덱스의 경우, 네이버와 카카오 앱 사용을 분석하여, 앱과 웹툰의 사용자수, 사용일, 시간수, 교차 사용자 수 등의 지표를 측정함
  - 온라인 거래액 등과 같은 기업정보부터 소비자의 쇼핑 행태 등도 제공함
- 해외의 경우 Statistica, StatCounter, Nielson Digital Panel, eMarketer 등이 있으며, 한국의 경우 닐슨코리아클릭, 코리아클릭, 아이지에이웍스 모바일인덱스 등이 있음

## □ 기업 및 협회 데이터

## ● 사용 지표와 출처

- 해당 기업의 데이터를 기반으로 한 기업 IR자료, 실적보고서, 재무제표 등
  - 재무제표의 매출, 영업이익 등
  - 플랫폼 기업들이 제시하는 월간 활성 이용자 수, 창출하는 가치 등
- 기업 자체의 조사
  - 아마존 구매자와 판매자에 대한 서베이(2020)는 유럽 5개국 대상 온라인마켓 플레이스에서 활동하는 판매자와 구매자의 멀티호밍 등을 조사함. 아마존에서 의뢰하고, 시장 조사 기업인 Alligator Digital이 수행했으며 Compass Lexecon의 지원을 받음
- 협회 데이터
  - 협회 구성원들로부터 직접적으로 데이터를 수집하고 공유하기 때문에 전체 시장에 대한 정보를 보유함
  - MOIBA (한국모바일산업연합회) 자료 - 플레이스토어와 원스토어의 소비자 지출액(공정위 구글-안드로이드 사건에서 사용됨)

## □ 사후적 연구 기반의 경쟁지표

- 플랫폼 관련 실증연구로는 규제의 변화가 미치는 영향을 추정하는 연구들 뿐 아니라, 플랫폼의 정책이 소비자와 공급자의 선택에 미치는 영향을 추정하는 등 다양한 연구들이 존재함
- 아직까지는 가용데이터의 한계로 연구들이 그리 많지 않지만 웹크롤링 등 데이터수집 능력이 발전함에 따라 연구가 증가하고 있으며, 공공, 기업 및 산업과의 협업이 증가함에 따라 더욱 증가할 것으로 기대됨
- 관련 기업으로부터 직접 데이터를 받거나, 웹크롤링 등을 통해 수집함
- 숙박예약 플랫폼 관련 연구
  - Mantovani, Piga, and Reggiani (2021)
    - 규제의 변화가 호텔 숙박요금에 미치는 영향을 살펴봄

- 온라인 여행사(Online Travel Agency, OTA)가 자사 플랫폼을 이용하는 호텔에게 넓은 최혜국대우(Wide Most favored nation, Wide MFN)을 적 용하여 자사플랫폼에서 동일하거나 가장 낮은 숙박요금을 책정하도록 하였으나, 프랑스에서는 규제에 의해 2015년부터 MFN을 전면적으로 금지함
  - Booking.com 예약사이트로부터 직접 수집한 데이터에 이중차분법(difference-in-difference)을 적용하여 규제 전후의 가격의 변화를 살펴봄
  - 결과: 단기적으로 Booking.com에서의 숙박요금이 감소하였으며, 특히 높은 등급의 호텔일수록 체인호텔일수록 숙박요금이 더 감소하였음.
  - Ennis, Ivaldi, and Lagos(2022)도 wide MFN이 금지로 인해 narrow MFN만 가능 했을 때 숙박요금이 더 감소했다는 유사한 결과를 얻었음. 그 외에도 다수의 연구논문이 나옴
- 김민정(2019)
- 호텔 예약 플랫폼의 데이터를 이용하여 플랫폼 내의 검색 순위, 호텔 등급, 홍보 태그 간의 관계를 분석하였음. 플랫폼이 부여하는 호텔 등급이 공식 등급보다 상향 분포되며, 그 차이가 클수록 검색 순위가 높아지는 가능성을 제시함
- 이 외에도 아마존과 구글을 포함한 온라인 쇼핑 관련 연구 등이 존재함

[그림 IV-3] 경쟁 지표의 출처 및 한계점

01	<p><b>정부 및 공공기관 통계 및 실태조사</b></p> <ul style="list-style-type: none"> <li>• 다양한 플랫폼을 다루고 있으나, <b>관련 질문 및 응답의 형태와 깊이가 다름</b></li> <li>• 정기적일 수 있으나 <b>플랫폼 역동성을 담는데 한계 존재</b></li> </ul>
02	<p><b>경쟁법 사건 관련 지표</b></p> <ul style="list-style-type: none"> <li>• 해당 위반행위, 기간 등에 특화되어 있고, <b>기업의 영업비밀 등으로 인한 공개 불가능</b></li> </ul>
03	<p><b>빅데이터 기업, 서베이 기업으로부터의 지표</b></p> <ul style="list-style-type: none"> <li>• 플랫폼 특화된 다양한 지표들을 수집 및 제시하며, <b>high frequency data 및 통계 보유</b></li> <li>• <b>서베이 대상 샘플링 및 방법론에 따라 편향적인 지표가 산정</b> 될 수 있음</li> </ul>
04	<p><b>경제분석 및 사후 연구</b></p> <ul style="list-style-type: none"> <li>• 다양한 요인을 통제하면서 경쟁정도를 추정할 수 있으나, <b>가용데이터 및 시간에서의 한계 존재</b></li> </ul>

**(2) 경쟁 측정지표와 가용데이터의 한계점과 보완점****□ 플랫폼 시장과 대상의 선정****● 플랫폼 시장의 선정**

- 경쟁을 측정하는 대상이 되는 플랫폼 시장을 선정해야 함
- 경쟁법 사건에서는 치열한 법리공방과 경제분석을 통해 시장획정 하는데, 관련 판례가 있는 경우 해당 시장획정을 참고할 수 있음. 그러나 과거에 발생한 행위를 판단하는 것이기 때문에, 역동적이고 진화하는 플랫폼의 경우 과거와 현재의 시장획정이 다를 수 있음
- 하지만 경쟁법 집행에서의 시장획정은 위반행위를 중심으로 관련 결과들을 평가하고 제재하기 위한 전 단계이고, 산업 정책적 관점에서의 시장획정은 산업의 발전과 공정한 생태계의 조성을 위해 경쟁 등 다방면으로 모니터링을 위해 선정함. 즉 시장획정의 목적과 기준이 다름
- 경쟁 모니터링의 목적에 부합하는 그리고 동태적이고 혁신적인 플랫폼 발전의 특성을 고려하여 시장을 선정하여야 함

**● 대상의 선정**

- 경쟁정도는 시장 전체 수준에서 측정될 수 있고, 대표적인 주요기업들을 중심으로 측정될 수도 있음. 예를 들어, 시장집중율이나 이익률 등은 시장 전체 수준에서 총량적으로 측정되기 용이하지만, 멀티호밍이나 데이터 등은 경쟁이 치열한 주요기업들을 중심으로 측정될 수 있음
- 따라서 개별 수준에서의 측정지표와 이를 종합한 측정지표를 모두 고려할 수 있도록 대상을 선정하는 것이 바람직함

**□ 지표의 사용**

- 현재 널리 사용되고 있는 경쟁 측정 지표들은 전통적으로 사용해 왔던 지표들에 플랫폼의 특성을 반영한 시장구조와 성과 관련 변수가 대부분임
- 플랫폼의 특성을 고려하여 전통적인 지표들을 구성하는 기본 데이터가 달라짐. 예를 들어, 매출액 기반 시장점유율 대신 이용자 수 기반 시장점유율이 더 적절할 수 있음

- 간접네트워크효과와 멀티호밍 등과 같이 타 산업에서는 드물게 쓰였던 지표들이 플랫폼 경쟁에 대한 지표로 사용되고 있음
  - 멀티호밍 여부와 조금 더 세세하게는 멀티호밍하는 플랫폼의 조합, 플랫폼간의 이동 등을 조사하기도 함. 하지만 개별 멀티호밍이 갖는 가치 및 이에 따른 플랫폼 시장지배력의 차이 및 경쟁에 대해서는 조사되지 않고 있음. 통합된 경쟁지표(aggregate index)와 더불어 좀 더 세분화할 수 있던 데이터와 그에 따른 지표는 경쟁 측정에 도움이 될 것임

□ 경쟁 측정지표의 출처, 한계점 및 보완점

- 정부, 공공기관, 협회 등의 실태조사는 최근 다양한 출처에서 플랫폼 관련 지표를 담고 있으며, 전수에 가까운 조사가 가능할 수 있다는 장점을 지님
  - (한계점) 실태조사마다 특정 산업을 다루고 있어 플랫폼 관련 질문들이 산재해 있으나, 기준이 제시되지 않아 각기 다른 형태 및 깊이의 지표를 제시하고 있음
  - (한계점) 특정 주제에 대한 조사이며, 정기적일 수 있으나 실태조사 간의 간격이 플랫폼의 역동성을 맞춰 현황을 담아내는 데 한계가 있음
  - (보완점) 실태조사 정비하여, 플랫폼 시장마다 필요한 주요지표 및 질문의 기준을 설정하고 세부 시장의 고유한 특성을 담을 수 있도록 개선할 필요가 있음
- 경쟁법 사건 집행에서 수집되는 지표들은 해당 기업과 경쟁기업으로 직접적으로 받는 세세한 수준의 데이터에서 생성되기 때문에 매우 세세하고 정밀한 지표 측정이 가능해짐
  - (한계점) 해당 위반행위, 사건 및 시장(또는 산업) 그리고 해당 기간에 특화된 데이터와 통계를 사용하기 때문에 특정 산업에 대한 경쟁 모니터링을 하는 경우보다 범위가 좁아질 수 있음
  - (한계점) 더욱이 기업의 영업비밀 등을 포함하고 있어 공개할 수 없는 경우가 많고, 해당 사건 외에는 사용할 수 없다는 약점을 가짐
- 빅데이터기관에서 수집하는 지표들은 플랫폼 간의 멀티호밍, 구매자의 주의(consumer attention), 노출정도 등 플랫폼 시장에서 중요한 지표이지만 전통적으로 수집되지 않았던 지표를 포함하고 있어 현황 파악, 전략, 정책 수립에 도움이 되며 경쟁정도의 유추에도 기여함

- (한계점) 전통적으로 사용되었던 지표들이 아니기 때문에, 전수에 가까운 데이터에 대한 통계로 수집되어 오지 않았음. 크롤링 및 서베이 대상의 일부 데이터를 통해 지표를 산정하기 때문에 샘플링 및 방법론에 따라 차이가 있을 수 있어 편향적인 지표가 산정될 수 있음
- 경제분석 및 연구는 다양한 요인들을 통제하면서 특정 요인이 시장구조나 성과, 경쟁에 미치는 영향을 추정할 수 있다는 장점을 가지고 이들을 경쟁지표로 사용할 수 있음
  - (한계점) 경제분석 또한 가용데이터의 제약이 존재하며, 분석하는 데 있어 시간이 오래 걸리며 지표가 추정되었을 때는 또 다른 경쟁이 진행되어 적기에 제공되지 못할 수 있음. 또한 정기적으로 진행하는 데 어려움이 있음
  - (보완점) 사전 모니터링, 사후 점검, 정책 개선 등에 중요한 역할을 할 수 있으므로, 공공과 산업과의 연계를 통해 활성화할 필요가 있음

□ 경쟁의 장기추적

- 장기적으로 지표의 추적 관찰 및 동태적 경쟁과정에 대한 해석이 필요함
- 시장점유율, 멀티호밍, 의존도 등 절대적인 수치가 갖는 경쟁정도에 대한 의미도 있지만, 이들의 변화가 경쟁정도의 변화를 시사함
- 동태적이고 혁신적이며 변화가 빠른 시장일수록 주기적인 지표 관찰이 더욱 중요해짐

[그림 IV-4] 경쟁 지표 출처의 보완점

01	<p><b>실태조사의 정비</b></p> <ul style="list-style-type: none"> <li>• 플랫폼의 <b>공통 주요 지표 및 질문/응답의 기준 설정, 각 플랫폼 시장의 고유 특성</b> 포함</li> </ul>
02	<p><b>정기적/전수에 가까운 정부기관의 실태조사 뿐 아니라, 빅데이터 기업의 빈도가 높은 지표의 추이를 통해 플랫폼의 동태성과 역동성을 추적할 필요 있음</b></p>
03	<p><b>경제분석 및 사후연구는 사전/사후 모니터링, 정책 개선 등에 있어 다른 경쟁지표들을 보완한 한계점에도 불구하고 활성화 필요 존재</b></p>

### 3. 플랫폼 영향 확산

#### 가. 플랫폼 영향 확산 측정의 필요성 및 범위

##### (1) 필요성

###### □ 플랫폼 경제 확산과 영향력

###### ● 플랫폼 경제 영향력 측정의 필요성

- 플랫폼 경제 확산의 중요성은 플랫폼 기업과 산업 자체에 대한 측정 및 평가뿐만 아니라 플랫폼을 이용하는 소비자와 생산자 및 산업에 대한 영향력에 대한 측정과 분석이 필요함
- 플랫폼의 영향력을 파악하기 위해서는 개별 플랫폼별로 소비자와 생산자에게 미치는 다양한 효과를 측정하고 분석해야 함
  - 플랫폼의 소비자에 대한 영향력을 파악하기 위해서는 온라인 쇼핑, 공유숙박, 배달앱 등 개별 플랫폼이 소비자 후생에 영향을 주는 메커니즘에 대한 분석이 선행되어야 함
  - 플랫폼은 개별 생산자의 성장과 다른 공급자와의 경쟁 등 해당 산업의 구조와 성장에도 영향을 줌
- 플랫폼 확산이 생산자와 소비자뿐만 아니라 관련 산업, 노동시장, 지역경제 등 다양한 부문에 미친 영향에 대한 측정 필요함

###### ● 정책적 함의

###### ● 플랫폼 경제의 확산은 관련 산업, 공급자, 소비자 간 혜택의 불균등한 배분으로 인해 이해 갈등 촉발

- 증거에 기반한 이해 갈등 조정을 위해서는 생산자와 소비자 등 개별 주체별 혜택에 대한 측정이 필요함

###### ● 한국 서비스업에서 플랫폼 경제의 확산은 서비스 부문의 자영업의 디지털 전환을 통한 저성장 극복이라는 점에서 국가 경제적으로 중요함.

- 한국은 OECD 국가 중 제조업 대비 서비스업 생산성이 가장 낮은 나라 중에 하나로 제조업 대비 서비스업 생산성이 OECD 평균 수준으로 상승하면 1인당 GDP가 현재보다 약 60% 상승할 것으로 예측(OECD 한국경제 보고서, 2022)

## (2) 범위

### □ 소비자 및 생산자

#### ● 플랫폼 이용 소비자의 후생 변화를 개별 플랫폼별로 측정

##### - 소비자 후생 효과

- 온라인 쇼핑을 통한 낮은 가격과 제품의 다양성 확대 등을 통해 소비자 후생 증가
- 공유숙박을 통해 낮은 숙박비용과 유연한 공급을 통해 소비자 후생 증가
- 배달앱을 통한 음식점까지의 이동 비용을 절감하고 다양한 음식점을 선택함으로써 소비자 후생 증가
- 낮은 가격의 경우 온오프라인 가격 정보 또는 온라인 구매 비중 자료 등으로 소비자 후생 변화 측정 가능
- 하지만 제품의 다양성 확대의 소비자 후생 효과 측정에는 신용카드 결제정보 등을 바탕으로 개별 소비자의 구매 정보와 소비자 후생 변화를 금액으로 환산할 수 있는 모형이 필요함.

##### - 소비자 혜택의 불균등

- 플랫폼 이용 소비자 혜택의 불균등은 디지털 플랫폼 이용을 위한 인프라 부족과 활용 지식 부족 등으로 발생
- 플랫폼을 이용하는 소비자 간에도 거주 지역과 소득 등에 혜택의 차이 발생

#### ● 플랫폼 이용 생산자 및 산업

- 생산자의 플랫폼 의존도는 플랫폼 이용 생산자의 비중 및 전체 매출 중 플랫폼을 이용한 매출 비중을 통해 파악
- 기업의 플랫폼 의존도는 거래액 기준뿐만 아니라 멀티호밍 여부를 통해 파악하는 것도 필요함
- 플랫폼을 이용하는 생산자는 매출을 증가시킬 수 있으며, 이러한 매출 증가의 일정 부분은 경쟁 사업체의 매출을 탈취해 발생해 기존 업체의 매출 감소와 퇴출을 유발함

- 플랫폼을 이용하는 개별 생산자의 증가는 해당 산업의 수요 확대와 구조변화를 가지고 올 수 있음

□ 노동시장과 지역경제

● 노동시장

- 플랫폼은 소비자와 생산자뿐만 아니라 소매업 및 관련 업종의 노동시장에 큰 영향을 줌
  - 기존 연구에 따르면 온라인 쇼핑 확대에 따른 일자리 감소는 지역별로 상이하  
다는 점을 고려한 정책 필요

● 지역경제

- 공유숙박 디지털 플랫폼은 음식점, 소매업 등 관광 관련 서비스업 활성화를 통해 지역경제에 기여
  - 공유숙박은 지역 기반 서비스업의 다양한 지역의 균형적인 경제 활성화에 기여

## 나. 기존 연구

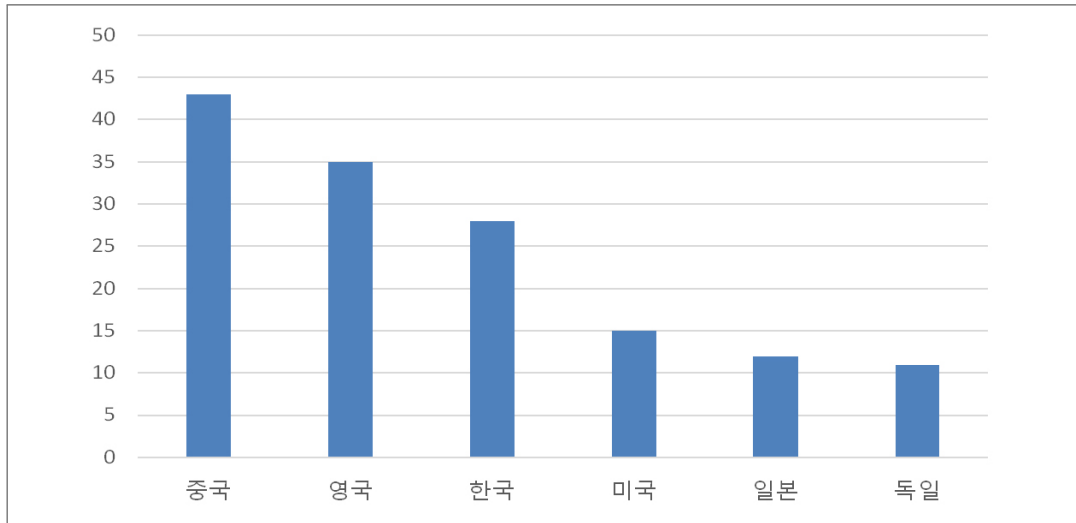
### (1) 소매업

□ 온라인 쇼핑의 확산

- 소매업에서 온라인 쇼핑 비중은 2021년 기준으로 약 28%로 소매업태 중에서 가장 큰 비중을 차지함
  - 한국의 온라인 쇼핑 비중은 미국과 일본보다는 높지만, 중국 및 영국보다는 낮은 수준임
  - 온라인 쇼핑 확대에 따른 쇼핑 플랫폼 거래 규모 및 플랫폼 간의 경쟁 구조 파악뿐만 아니라 소비자, 공급자(소매업체), 종사자에게 준 혜택과 영향에 대한 파악이 필요함

[그림 IV-5] 소매업 매출의 온라인 쇼핑 비중, 2021

(단위: %)



출처: Statista.com/statistics/1345392/e-commerce-share-retail-sales-worldwide-country

#### □ 소비자

- 소비자는 온라인 쇼핑을 이용함에 따라서 낮은 가격, 다양한 제품, 이동비용 절약 등 다양한 혜택을 받을 수 있으며, 하지만 온라인 쇼핑의 혜택은 이용자 특성에 따라서 차별적일 수 있음
  - 온라인 쇼핑은 탐색비용을 절감시켜 가격을 낮추고 업체 간 가격 차이를 줄이는 효과를 가짐(Goldfarb and Tucker 2019)
  - 온라인 쇼핑은 가격 이외에도 제품의 다양성과 이동비용 절약이라는 혜택을 소비자에게 제공함
    - 미국 VISA 신용카드 빅데이터를 분석한 Dolfen et al.(2023)의 연구에 따르면 온라인 쇼핑이 소비자에게 주는 혜택은 이동 비용 절약이라는 편리성보다는 소비하는 제품의 다양성 확대를 통해 나타난다는 것을 보여줌. 특히 온라인 쇼핑 혜택의 크기는 고소득층과 인구 밀집 지역 거주자에게 더 크게 나타남
    - 온라인 쇼핑 확대가 소비자에게 주는 혜택의 크기와 더불어 소비자 특성별로 혜택의 크기에 대한 측정이 필요함. 고소득자에게 혜택이 큰 경우 소득 불평등보다 소비 불평등이 커지지만, 반대의 경우 온라인 쇼핑이 소비를 통해 불평등을 줄이는 효과가 있음

- 온라인 쇼핑 확대에 따른 디지털 약자에 대한 파악 필요
  - 코로나 팬데믹 기간 오프라인 쇼핑이 어려워진 상황에서 일반 소비자는 온라인 쇼핑을 통해 필요한 물품을 구매함, 하지만 온라인 구매가 어려운 디지털 약자인 노령층은 팬데믹 기간 필수적인 상품 구매에 어려움에 직면함(Chun et al. 2022)
  - 장기적으로 온라인 쇼핑 확대로 기존 오프라인의 점포가 줄어들다면 온라인 쇼핑을 이용하지 못하는 고령층 등 디지털 약자에 대한 파악과 대책이 중요해짐
- 한국에서도 신용카드 빅데이터를 활용해 온라인 쇼핑이 소비자에게 주는 혜택의 종류와 크기를 측정하고, 인구통계학적 특성에 따른 소비자 혜택의 차이를 측정하는 연구가 필요함

□ 생산자

- 온라인 쇼핑의 확산은 기존 오프라인 소매업을 대체함과 동시에 기존 및 신생 소매업체가 플랫폼을 통해 성장할 수 있는 기회를 제공함
  - 오프라인 소매업체는 시장의 지리적 한계로 인해 다른 지역에 신규 점포 증설 없이는 매출 성장을 이루기 어려움, 하지만 온라인을 통한 판매는 시장의 지리적 한계가 없으므로 소매업체의 지속적인 매출 성장에 도움을 줄 수 있음
    - 온라인 소매업체는 점포개설에 따른 고정비용이 적어서 오프라인에 비해 진입이 쉽지만 생존 확률도 낮음, 하지만 생존 시 오프라인 소매업체에 비해 온라인 판매 소매업체는 매출과 생산성 성장률이 높음(Chun and Shin 2023)
    - 오프라인과 온라인 소매업체의 진입, 성장, 퇴출에 관한 특성이 다르다는 점을 고려한 소매업 규제 및 지원 정책에 대한 재설계가 요구됨
  - 소매업은 약 100만개의 사업체와 약 200만명의 종사자를 가진 서비스 산업으로 플랫폼 경제의 확산을 통한 사업체의 디지털 전환이라는 과점에서 온라인 쇼핑 플랫폼의 생산자에 대한 영향력을 측정할 필요가 있음
    - 소매업체의 90% 이상이 10인 미만의 자영업체라는 점에서 플랫폼이 자영업의 매출, 경쟁력, 그리고 생존에 미친 효과에 대한 분석이 필요함. 특히 어떤 소매업체가 나아가 온라인 쇼핑 플랫폼을 활용해 성장과 경쟁력을 확보하고 있는가에 대한 분석이 필요함

- 소매업의 디지털 전환은 거시경제 측면에서 한국 서비스업 낮은 생산성 제고와 저성장 극복에 디지털 플랫폼이 어떤 역할을 할 수 있는가와 관련
  - 소매업체의 플랫폼 이용 비중 및 의존도 등 미시 수준의 영향력 척도와 경쟁정책 차원의 접근과 더불어 산업과 경제 전반에 대한 영향력에 대한 거시경제 수준의 분석 필요

#### □ 노동시장

- 온라인 쇼핑 플랫폼은 소비자와 생산자(소매업체)뿐만 아니라 소매업 및 관련 업종의 노동시장에 큰 영향을 줌
  - 미국의 경우 온라인 쇼핑 확산에 따라 대규모 소매체인의 연쇄 파산으로 이어져 소위 소매업의 종말(retail apocalypse)에 대한 우려 발생함.
    - 한국의 경우 대형 체인점의 파산이 발생하지는 않았지만 오프라인 상점의 위축이 일어나 노동시장에 대한 영향이 발생하고 있음.
    - Chun et al.(2023)은 신용카드 빅데이터와 통계청 전수 소매사업체 자료를 이용한 실증 분석을 통해 2010-2015년 기간 온라인 쇼핑 확대로 오프라인 소매업 고용이 약 4.9% 감소함을 보임, 하지만 도시지역에서는 소매업에서 줄어든 일자리가 음식점 등 다른 서비스 부문에서 일자리가 만들어져 전체 고용에 미치는 영향은 미미했지만, 이러한 효과가 비도시 지역에서는 거의 발생하지 않음,
    - 학력, 성별 등 다양한 그룹의 고용에 미치는 영향에 대한 분석이 필요함.
- 온라인 쇼핑이 노동시장에 미친 효과에 대한 다면적인 측정과 분석 필요
  - 기존 연구에서 지적한 온라인 쇼핑 확대에 따른 일자리 감소는 지역별로 상이하다는 점을 고려한 정책 필요
    - 지역별 고용효과와 더불어 연령, 학력별로 차별적인 고용효과에 대한 분석도 필요함
  - 고용효과와 더불어 임금과 일자리 구성의 변화에 대한 효과
    - 온라인 쇼핑의 확대로 기존 일자리가 파괴되는 것뿐만 아니라 소매업의 영업 방식 자체가 변화한다는 점에 주목할 필요

- 소비자가 오프라인 점포에 와서 구매하지 않고 온라인으로 주문할 경우, 기존의 판매원은 줄어들고 대신에 풀필먼트 센터와 라스트 마일 딜리버리 등 창고업과 운수업의 종사자가 늘어남. 고용 인원 변화와 더불어 고용 구성이라는 측면에서 산업과 직업 자체의 변화가 발생함

## (2) 숙박업

### □ 공유경제의 확산

- 공유경제(sharing economy)는 유휴 자산 또는 노동력을 보유한 공급자와 이를 필요로 하는 수요자 간의 거래를 디지털 플랫폼을 통해 증대하는 것으로 정의됨
  - 디지털 플랫폼은 개인이 보유한 유휴자산과 노동력 거래를 편리하게 하여 새로운 공급자의 시장 참여를 활성화함
  - 공유경제는 공급이 비탄력적이고 제한된 숙박업, 택시 산업에서 급속하게 확산해 따라서 신규 공급자와 기존 사업자와의 경쟁이 강화되고 장기적으로는 산업 구조 자체가 변화함
  - 특히 공유경제는 공급이 비탄력적이고 제한된 숙박업, 택시 등의 산업에서 급속하게 확산해 가격을 하락시키는 등 소비자 후생에 영향을 줌
  - 공유경제는 해당 산업뿐만 아니라 연관산업과 지역경제에 외부효과(externality)를 가짐
    - 기존 숙박업과 달리 에어비앤비 등 공유숙박은 더 넓은 지역 분포함으로써 주변의 음식점과 소매업 등 관련 산업을 활성화하는 긍정적인 외부효과를 가짐
    - 기존 주택이 숙박 요도로 이용됨에 따라 주택가격 상승, 범죄 및 소음 등 부정적인 외부효과가 발생함

### □ 소비자

- 공유숙박을 통해 소비자는 단기적으로는 낮은 숙박비용 그리고 장기적으로는 다양한 숙박 서비스를 제공받을 수 있음
  - 공유숙박의 확대는 저비용 숙박 서비스 공급자의 시장진입 역할 수행을 통해 기존 호텔의 가격 하락을 유도함

- 에어비앤비 등 공유숙박의 호텔 가격 하락에 미친 효과는 뉴욕과 같이 호텔 공급이 제한된 곳이나 페스티벌, 올림픽 등과 같은 대형 행사로 단기적으로 수요가 급증한 지역에서 크게 나타남
- 호텔은 차별화된 서비스를 제공하는 산업으로 경쟁의 격화는 기업들이 비가격 전략으로써 차별화를 강화함. 기존 호텔은 식음료 등 비객실 서비스를 강화하고 진입 업체는 높은 수준의 객실서비스를 제공하는 하이엔드 호텔 전략을 취함
- 공유숙박의 확대는 숙박 서비스 가격을 하락시키고 장기적으로는 서비스가 다양화됨으로써 소비자에게 혜택을 제공함

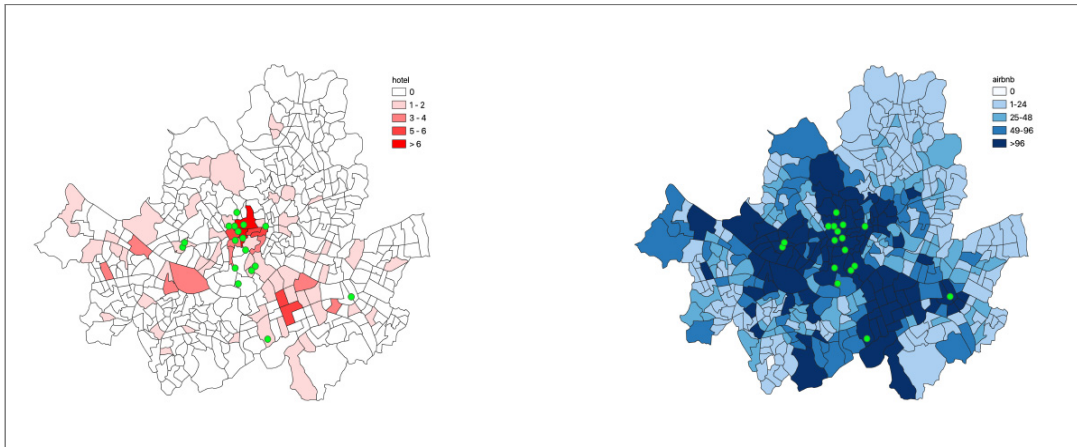
#### □ 생산자

- 공유숙박 디지털 플랫폼은 유희 주택을 보유한 일반 개인을 숙박업의 신규 공급자가 되게 함
  - 유희 주택의 활용을 높여 생산 측면에서 효율성을 높일 수 있지만 기존 숙박업체에 대한 규제 역차별 문제가 발생함
    - 공유숙박 공급업자는 기존 숙박업체에 적용되는 소방법 등 다양한 법적 규제를 안 받아 규제 역차별 이슈가 발생함
    - 디지털 플랫폼은 숙박 거래만을 중개하기 때문에 소비자와 공급자 사이에 발생하는 피해 해결에 대한 장치가 미흡함
    - 새로운 비즈니스 모델인 공유숙박에 대해 기존 규제를 적용하기보다는 공유경제의 장점을 살리고 피해를 최소화하는 새로운 규제체계를 만들 필요성이 큼
  - 공유경제 디지털 플랫폼은 거래만을 중개하므로 산업의 변화는 플랫폼 자체보다는 플랫폼을 통한 거래하는 공급자에게 달려 있음
    - 이커머스 플랫폼이 물류시설을 갖추고 기존 오프라인 업체를 대체하는 것과는 달리 공유숙박 디지털 플랫폼은 거래 중개의 역할만을 수행하므로 사업에 대한 영향력은 공급자의 구성과 특성에 크게 의존함
    - 플랫폼과 공급자 간의 수수료 분쟁보다는 플랫폼 수수료 및 리뷰정책 등이 숙박 공급자 간의 경쟁과 진입 등에 미치는 영향에 주안점을 두고 분석 필요

□ 지역경제

- 공유숙박 디지털 플랫폼은 음식점, 소매업 등 관광 관련 서비스업 활성화를 통해 지역경제에 기여
  - 특정 지역에 몰려있는 호텔과는 달리 기존 주택을 이용한 공유숙박은 도시 내에서 지리적으로 고르게 분포함으로써 숙박업소 주변의 음식점 등의 활성화를 통해 지역 경제에 기여
    - 뉴욕시에서 에어비앤비 확산은 주변 레스토랑의 매출 확대를 가지고 옴 (Alyakoob and. Rahman 2022).
    - 기존 호텔도 주변의 음식점 매출을 올리지만 효과는 특정 지역에 집중되지만 공유숙박은 도시 내 다양한 지역의 음식점 매출을 높여 관광업 혜택을 다양한 지역으로 확대하는 역할
    - 공유숙박은 지역 기반 서비스업의 다양한 지역의 균형적인 경제 활성화에 기여

[그림 IV-6] 서울 내 호텔과 에어비앤비의 지리적 분포



주: 좌측 붉은 그림은 2012년 서울 읍면동 지역의 호텔 수, 우측 파란 그림은 2019년 서울 읍면동 지역의 에어비앤비 수, 초록색 점은 주요 관광지를 표시함.

출처: Chun, Joo and Kim (2023)

**(3) 배달앱****□ 온라인 음식 배달 서비스의 확산**

- 온라인 음식 배달 서비스는 코로나 팬데믹 기간을 거치면서 급속하게 확산해 음식점업의 디지털 전환을 촉진함
  - 온라인 음식 배달 서비스의 확산은 배달앱 활용 음식점의 매출을 증가시키고 배달앱 이용 소비자에게는 편리함과 다양성을 제공함
  - 온라인 음식 배달 서비스의 확산은 배달 플랫폼과 음식점 간의 수수료 분쟁 등 첨예한 이해 갈등의 문제가 발생함
    - 음식점은 소비자의 배달 주문 수요 확보를 위해 여러 배달앱 플랫폼을 이용하는 멀티호밍이 일반적이지만 소비자는 하나의 배달앱을 이용하는 싱글호밍과 더불어 일부는 여러 배달앱을 이용함
  - 한국 자영업 부문의 대표적인 업종인 음식점업의 배달앱 확산은 디지털 전환을 통한 자영업의 구조변화라는 측면에서 이해가 필요함

**□ 소비자**

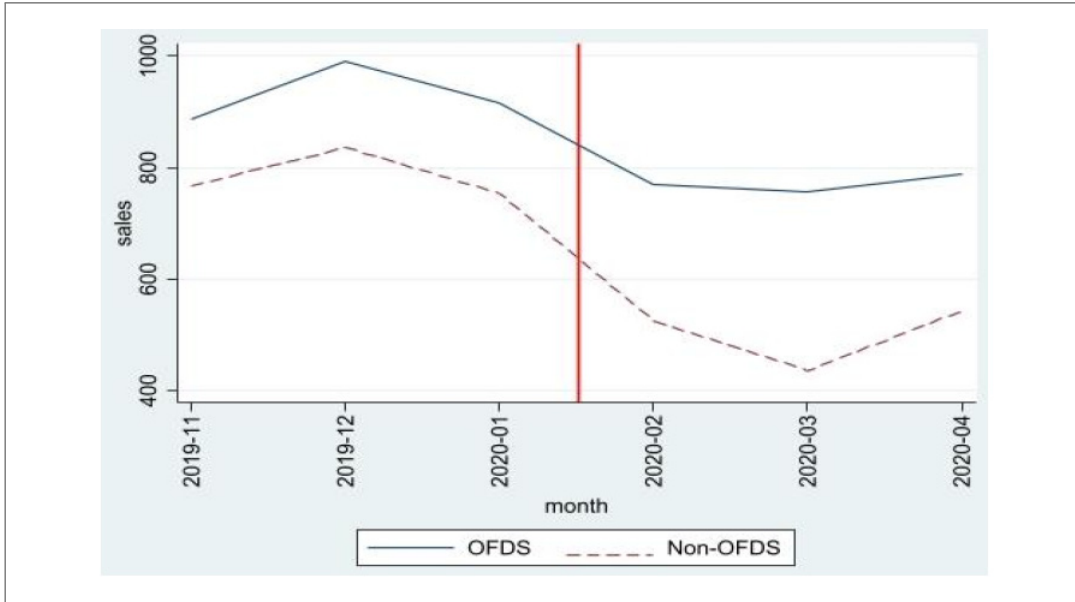
- 온라인 음식 배달 서비스는 음식점까지의 이동 비용을 절감시켜주고 많은 음식점에 대한 정보를 제공해 선택권을 확대해 소비자 효용을 증가시킴.
  - 이동 비용 절약은 소비자에게 편리함을 제공함과 더불어 멀리 있는 다양한 음식점에 쉽게 접근할 수 있게 함
    - 온라인 음식 배달 서비스의 이동 비용 절약을 통한 소비자 효용 증대 효과는 근처에 식당이 많지 않은 소비자에게 더 크게 나타날 수 있음
- 음식점에 대한 정보제공을 통해 소비자가 원하는 음식점을 쉽게 찾을 수 있도록 함
  - 음식점과 음식에 대한 정보는 소비자 리뷰에 크게 의존하기 때문에 리뷰의 정확성이 중요함

□ 생산자

- 음식점은 배달앱을 활용해 매출을 증가시킴과 동시에 배달앱 확산은 음식점업의 시장 수요를 확대함
  - 배달앱을 이용하는 업체는 그렇지 않은 업체에 비해 약 30% 정도 매출액이 높음(이공 2021)
    - 배달앱의 매출 증가 효과는 음식점별로 이질적으로 나타남
    - 이질적 매출 증가 효과는 음식점 간의 격차를 확대한다는 점과 경쟁력 있는 음식점의 성장을 촉진한다는 두 가지 측면을 가짐
  - 배달앱 확산은 음식점업 수요를 확대해 음식점업의 성장을 유도함.
    - 배달앱을 이용을 통한 매출 증가는 다른 음식점의 오프라인 매출을 줄이는 시장 탈취 효과도 가지지만 신규 수요 확대를 통해 이루어짐으로써 전체적으로 음식점 시장의 크기를 확대시킴 (Chun, Lee and Yang, 2023)
    - 배달앱 이용하는 개별 음식점이 매출 증가 효과의 일부는 시장 탈취를 통해서 발생한다는 점에서 음식점업의 창조적 파괴를 유발할 수 있음.
  - 배달앱 확산은 자영업 위주의 음식점업이 디지털 전환을 통해 창조적 파괴라는 구조적 변화가 발생할 수 있음을 시사함
    - 배달앱 확산에 따른 플랫폼과 음식점 간의 수수료 분쟁이 첨예한 문제이지만 보다 근본적인 부분은 배달앱을 이용하는 업체와 그렇지 않은 업체 간의 경쟁 구조의 변화임
    - 배달앱 이용이라는 디지털 전환을 통해 음식점의 매출 성장과 음식점 간의 경쟁 구조 변화에 대한 정확한 이해가 필요함

[그림 IV-7] 코로나19와 배달 이용 여부에 따른 음식점 매출 변화

(단위: 만원)



주: OFDS는 배달앱을 이용하는 대구지역 음식점의 코로나19 발생 전후 월간매출액이며, Non-OFDS는 배달앱을 이용하지 않는 음식점의 월간 매출액임.

출처: 이상원, 전현배(2022)

- 배달앱은 코로나 팬데믹 기간 음식점의 매출 보전에 도움을 줌
  - 코로나가 최초로 발생한 대구지역의 음식점이 배달앱을 이용하는 경우 매출감소의 약 50% 정도를 상쇄시킬 수 있는 것으로 나타남(이상원, 전현배 2022)
  - 배달앱은 또 다른 팬데믹에 대응할 수 있게 해 디지털 전환이 자영업의 생존 능력을 높이는 효과를 보여준 사례가 됨

## 다. 플랫폼 영향 확산 측정지표, 가용데이터 및 한계점

### (1) 소비자 혜택의 크기 및 불균형

#### □ 소비자 후생의 크기 측정

##### ● 플랫폼이 개별 소비자의 후생에 미친 영향 측정

- 디지털 플랫폼은 탐색비용 절감, 이동 비용 절감 등 다양한 기능을 통해 소비자에게 낮은 가격과 제품 다양성 등의 혜택을 제공(Goldfarb and Tucker 2019)
- 소비자 후생에 미친 효과를 측정하기 위해서는 소비자 후생에 영향을 주는 개별 플랫폼의 주된 채널을 파악하는 것이 선행되어야 함
  - 온라인 쇼핑은 낮은 가격과 제품 다양성, 공유숙박은 낮은 가격과 유연한 공급 등 개별 플랫폼에 따라서 소비자 혜택의 주된 채널이 다름

##### ● 소비자 혜택: 낮은 가격

- 온라인과 오프라인 가격 비교 자료를 이용해 소비자 혜택의 크기를 직접 측정 가능
  - 온라인과 오프라인 가격 비교 자료가 필요하므로 많은 품목과 장기간에 걸친 자료 구축은 어려움
- 온라인 매출 비중을 이용한 소비자 혜택 추정
  - 플랫폼을 통한 소비자 구매 비중이 증가할수록 소비자 물가 지수는 하락할 것이므로 온라인 매출액 비중과 가격 자료만으로도 간접적으로 소비자 혜택 추정 가능
  - 온라인 매출액 비중은 플랫폼 거래액 또는 신용카드 등 결제자료를 통해 온라인 거래액 파악 가능
  - 신용카드 등 결제자료의 경우 품목별 매출 또는 지출 자료 구축이 어려움
- 플랫폼 확산과 가격 하락 및 가격 격차 감소
  - 플랫폼을 통해 가격 비교 비용이 줄어 소매 품목, 보험료, 항공권 등 다양한 제품의 가격을 낮춤.
  - 온라인 판매의 특성상 지역 간 가격 차이를 줄이는 효과를 가짐(Jo et al. 2022)

● 소비자 혜택: 제품의 다양성

- 소비자 혜택이 제품의 다양성이라면 개별 소비자 수준에서 제품 구매 정보가 필요함
  - 온라인 쇼핑 플랫폼이 제공하는 제품 수 정보만으로는 소비자의 다양성 혜택 측정이 불가능함.
  - 개별 거래 수준의 신용카드 결제정보를 이용해 플랫폼을 이용하는 소비자의 구매처 수의 변화를 파악해야 함.
  - 개별 소비자의 구매처 수의 변화를 측정하는 자료가 있을지라도 제품 다양성 확대의 혜택을 금액으로 환산하기 위해서는 경제학 모형을 이용한 추정과정이 필요함.
  - 온라인 쇼핑의 소비자 혜택에서 제품 다양성이 중요하다는 최근의 연구 결과 (Dolfen et al. 2022)

□ 소비자 혜택의 불균등 분포

● 플랫폼 이용 소비자 혜택의 불균등

- 디지털 플랫폼 이용을 위한 인프라 부족과 활용 지식 부족
  - 디지털 플랫폼을 이용 자체의 차이로 인한 혜택의 차이로 소위 디지털 정보 격차(digital divide)임
  - 디지털 정보격차 자체에 대한 현황은 『디지털정보격차 실태조사』와 같은 설문조사와 디지털 플랫폼을 통한 소비자별 결제자료를 통해 파악 가능
  - <표 IV-2>는 코로나 기간 음식 배달 서비스와 온라인 쇼핑은 중요한 비대면 소비 채널로 그 유용성이 모두에게 인식되었어도 고령층의 실제 이용은 일반 국민의 절반 수준임을 보여줌.
  - 한국의 경우 연령별 플랫폼 이용 여부에 따른 혜택의 차이는 인프라보다는 활용 지식의 차이 때문

〈표 IV-2〉 코로나 기간 고령층의 음식 배달 서비스 이용

(단위: %)

구분	인지도	유용성	사용경험
일반 국민	89.1	97.7	70.9
고령층 (50대 이상)	78.1	96.0	36.7

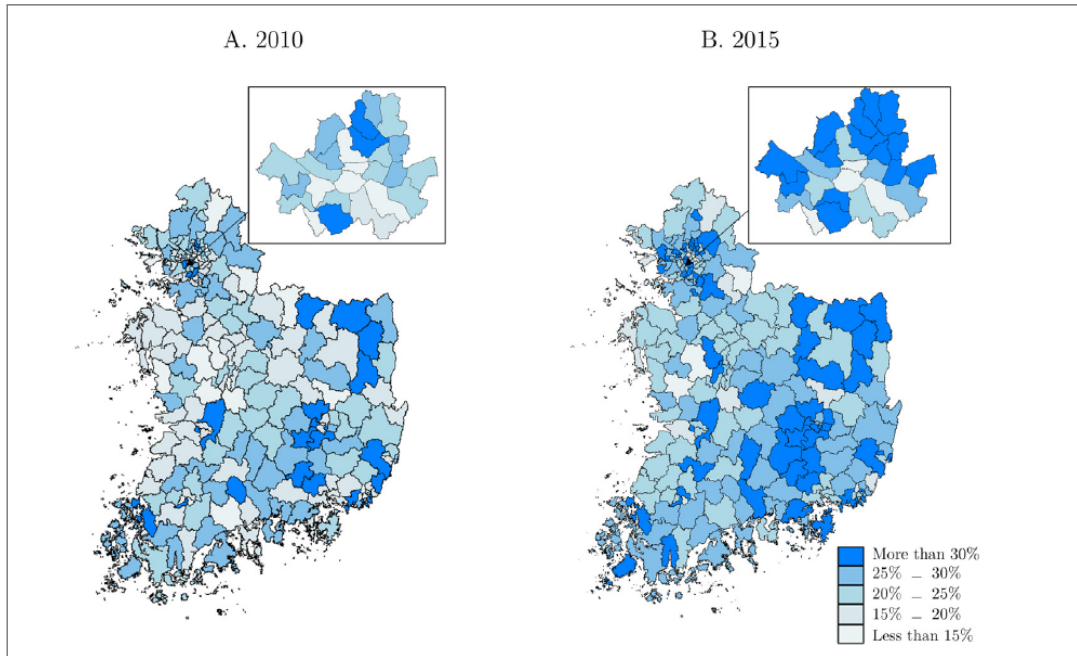
출처: 과학기술정보통신부, 한국진흥정보사회진흥원, 2022 디지털정보격차 실태조사, 2022

- 플랫폼을 이용하는 소비자 간의 혜택 차이

- 소비자의 거주 지역과 소득 등에 따라서 디지털 플랫폼 이용에 따른 혜택의 차이 발생
- 거주 지역 주변에 상점이 발달하지 않은 소비자의 경우 온라인 쇼핑의 혜택이 크다는 점에서 본다면 도시보다는 비도시 지역에서 소비자 혜택이 클 가능성 존재
- 온라인 쇼핑의 혜택의 크기가 다양성과 전체 소비액 등에 의존한다면 고소득 소비자의 혜택이 다 클 수 있으며, 실제로 이들이 거주하는 대도시 지역의 소비자 혜택이 더 클 가능성이 존재
- [그림 IV-4]는 신용카드 자료를 이용해 구축한 시군구별 소매업 지출에서 온라인 쇼핑이 차지하는 비중을 보여줌. 온라인 쇼핑 비중은 대도시 지역에서 높지만, 경북 내륙 등 상점이 잘 발달하지 않은 지역에서도 높음.
- 지역별 플랫폼 이용에 따른 소비자의 혜택은 지역 특성에 의존한다는 점에서 영향을 주는 지역 요인에 대한 정보와 지역 별 플랫폼 이용 자료 구축이 필요함
- 플랫폼 활용이 지역별 소득 격차를 축소 또는 확대하는 역할을 하는지에 대한 분석이 필요함

[그림 IV-8] 시군구 지역별 온라인 쇼핑 비중 변화: 2020년과 2015년

(단위: %)



출처: Chun et al. (2023)

## (2) 생산자 혜택의 크기 및 불균형

### □ 생산자의 디지털 플랫폼 의존도 측정

#### ● 생산자의 플랫폼 이용 여부 및 의존도

- 설문조사 또는 통신판매업자 등록 등의 행정 자료를 통해 기업의 플랫폼 이용 여부를 측정 가능
- 거래액 또는 매출액 중에서 플랫폼 특히 어떤 플랫폼을 통해 매출이 발생했는지에 대한 측정은 쉽지 않음
  - EU 권고대로 이를 정확하게 파악하기 위해서는 설문조사나 정부 공식 통계 이외에도 다양한 빅데이터의 이용이 필요
  - 플랫폼 의존 정도를 측정하기 위해서는 개별 사업자의 전체 매출액과 플랫폼을 통한 매출액 두 가지 자료가 필요하므로 플랫폼 사업자로부터의 거래자료만으로는 측정 불가능

- 온오프라인 매출 구분이 가능한 개별기업 매출에 대한 국세청 자료 또는 신용카드 결제자료 등이 필요
- 음식점의 배달앱을 이용한 매출 비중 측정은 개별 음식점의 신용카드 거래자료와 배달앱 플랫폼 사업자 자료의 결합을 통해 구축 가능(이공 2021)

〈표 IV-3〉 사업체의 디지털 플랫폼 거래 현황, 2020

사업체 업종 구분	사업체 수 (천개)	디지털 플랫폼 거래 사업체 수 (비중, %)	매출액 (십억원)	디지털 플랫폼 거래액 (비중, %)
소매업	974	106 (10.9)	443,146	33,523 (7.6)
숙박업	61	18 (29.1)	11,168	2,637 (23.6)
음식주점업	804	142 (17.6)	139,890	12,194 (8.7)

출처: 통계청, 경제총조사, 2021

- 한국 서비스업 사업체의 플랫폼 의존도
  - 〈표 IV-2〉는 통계청이 5년마다 실시하는 사업체 전수인 2020년 경제총조사에 나타난 소매업, 숙박업, 음식주점업의 플랫폼 의존도를 나타냄.
  - 3개 산업 중 숙박업의 디지털 플랫폼 이용 비중이 29.1%로 가장 높고, 음식주점업 17.6%, 소매업이 10.9% 순서임.
  - 숙박업의 경우 디지털 플랫폼을 이용한 예약이 핵심이고 소비자는 숙박업체에 방문에 소비하지만, 음식주점업과 소매업은 플랫폼을 통해 주문받고 배달 또는 택배 업체가 제품을 소비자에게 전달하는 구조임.
  - 디지털 거래액 비중의 경우 숙박업은 23.6%이지만 소매업과 음식주점업은 10% 미만임.
- 멀티호밍 또는 싱글호밍
  - 기업의 플랫폼 의존도는 거래액 기준뿐만 아니라 멀티호밍 여부를 통해 파악하는 것도 필요함.
  - 플랫폼 입점 또는 이용업체가 다른 플랫폼으로 전환해 매출을 할 수 있는 경우라면 특정 플랫폼에 대한 의존도는 낮아짐. 이러한 전환비용은 이용업체가 플랫폼으로 얻은 이익을 어느 정도 가져갈 수 있는지 즉 수수료 체계 등과 관련됨
  - 결국 사업체의 플랫폼 의존도 파악에 플랫폼 전환비용까지 고려 필요

● 한국 서비스업 사업체의 이용 중인 플랫폼 수

- <표 IV-4>는 서비스업 사업체가 이용하는 온라인 플랫폼 수를 보여줌. 소매업 사업체는 평균 4.3개의 오픈마켓, 숙박업 사업체는 평균 3.9개의 숙박앱, 음식점은 평균 2.9개의 배달앱을 이용하고 있음.
- 서비스업 사업체는 기본적으로 멀티호밍을 하고 있으며 이용 플랫폼 수는 해당 업종의 플랫폼 시장구조와 관련됨.

<표 IV-4> 서비스업 사업체가 이용 중인 온라인 플랫폼 수, 2021

온라인 플랫폼 유형	1개(%)	2개	3개	4개	5개 이상	평균(개)
오픈마켓	7.3	8.8	11.0	9.5	63.5	4.3
배달앱	10.3	20.0	46.3	18.7	4.7	2.9
숙박업	3.5	24.0	17.0	6.5	49.0	3.9

출처: 중소벤처기업부, 온라인 플랫폼 이용사업자 실태조사, 2021

□ 플랫폼 확산과 산업 구조 변화

- 플랫폼 확산은 플랫폼을 이용하는 개별 공급 사업체뿐만 아니라 해당 산업 전체의 구조를 변화시킴
  - 온라인 쇼핑의 확산은 기존 오프라인 소매업체의 매출 감소와 폐업을 촉진시킴. 온라인 판매를 하는 사업체는 오프라인 사업체에 비해 생존 확률은 낮지만 생존 시 더 높은 매출 성장률을 보임(Chun and Shin 2023). 한정된 지역 수요를 기반으로 한 오프라인 사업체와는 달리 지리적 시장의 한계가 없는 온라인 소매업체의 성장률이 높음을 의미함
  - 저가 공유 숙박 업체인 에어비앤비의 확산으로 기존 호텔은 식음료서비스를 강화하는 전략을 취하고, 신규 호텔은 더 높은 등급으로 시장에 진입하는 제품 차별화 전략을 취함에 따라서 숙박산업의 구조를 변화시킴
  - 배달앱을 이용 비중 증가는 외식산업 수요를 확대시켜 산업 성장에 기여함과 동시에 배달앱 이용 음식점이 비이용 음식점의 매출을 탈취하는 효과가 발생함

- 한국 서비스업에서 공급 사업체의 플랫폼 이용 확산은 자영업의 디지털 전환을 의미함
  - 영세 자영업 위주의 한국 서비스업의 제조업 대비 생산성은 약 50% 수준으로 OECD 평균 80%에 비해 크게 낮은 수준임
  - 서비스 부문의 자영업자가 플랫폼 이용하여 생산성을 제고할 수 있다면 서비스업 전체의 생산성 제고를 통해 저성장 극복에 기여 가능함
  - 플랫폼이 가지는 공급업자에 대한 영향력은 플랫폼과 공급업자간의 문제를 넘어서 공급업자가 속한 산업 구조의 변화와 성장에 대한 분석과 측정이 필요함

### (3) 노동시장 및 지역경제

#### □ 노동시장

- 플랫폼은 소비자와 생산자뿐만 아니라 노동시장에도 영향
  - 온라인 쇼핑의 확산에 따른 오프라인 소매업 일자리의 변화
    - 온라인 쇼핑의 확대에 기존 오프라인 소매업 일자리는 감소하지만, 소비자는 쇼핑 시간이 절약됨에 따라 레스토랑 등 다양한 서비스 이용 확대해 관련 고용 창출(Chun et al. 2023)
    - 온라인 쇼핑 확대에 따른 일자리 총량 변화보다는 일자리 파괴와 창출이 지역 및 사회인구학적 특성별로 상이할 수 있다는 점을 고려한 정책 필요
  - 한국의 경우 라이더를 중심으로 한 플랫폼 종사자 증가
    - 미국은 우버 등 승차 공유, 한국은 배달의 민족, 쿠팡이츠 등 배달 플랫폼 중심으로 플랫폼 종사자 증가
    - 기존 배달 노동자에 대한 처우 문제뿐만 아니라 보다 유연한 파트타임 노동자 소위 깃노동자(gig workers)의 등장이라는 노동시장의 변화에 주목할 필요
  - 디지털 플랫폼 확산에 따른 노동시장의 변화를 모니터링할 지표 필요
    - 개별 디지털 플랫폼이 노동시장에 미치는 영향을 모니터링하고 이들에 대한 사회적 보호 장치(안전망)에 대한 설계 필요(이화령 외 2022)

## □ 지역경제

- 플랫폼은 해당 산업뿐만 아니라 지역 경제에도 영향을 줌
  - 온라인 쇼핑은 지역적으로 분산 오프라인 상점의 일자리를 특정 지역의 물류센터 등으로 일자리의 지역적 분포의 변화
    - 미국의 경우 월마트는 50개 주에 약 5,000개의 점포를 가지고 있지만, 아마존은 50개 주 중에서 절반 이하의 주에만 약 200여개 풀필먼트 센터를 가짐
    - 단순히 매장이 물류센터로 바뀐 것뿐만 아니라, 물류 배송이 유리한 지역으로 일자리 지리적 위치도 변화하고 집중화가 일어남.
  - 공유숙박 디지털 플랫폼은 음식점, 소매업 등 관광 관련 서비스업 활성화를 통해 지역경제에 기여
    - 호텔보다 공유숙박은 다양한 지역에 분산되어 해당 지역의 음식점 및 소매업 등 지역 경제 활성화에 기여
    - 특히 공유숙박은 기존의 호텔이 들어가기 어려운 지역의 숙박이 가능해짐에 따라서 다양한 지역의 균형적인 경제 활성화에 기여
  - 디지털 플랫폼 확산에 따른 지역 경제의 변화를 모니터링할 수 있는 지표 필요
    - 디지털 플랫폼은 지역경제의 균형적인 발전에 도움이 되기도 하지만 지역간 격차를 확대할 수 있음.
    - 개별 플랫폼이 수도권과 비수도권의 불균형을 심화시키는가에 대한 엄밀한 분석을 통해 향후 지역 균형 정책에서 보완할 필요

## 4. 자국 플랫폼의 가치

### 가. 연구 목적 및 필요성

- 최근 인터넷 플랫폼은 한 국가의 경제, 사회, 문화 전반에 걸쳐 중추적인 역할을 담당하고, 혁신을 이끄는 핵심 인프라로 자리매김하고 있음
- 인터넷 플랫폼의 중요성이 주목받으면서 인터넷 플랫폼 사업자에게 거는 기대 역시 그 어느 때보다 커지고 있는 상황
  - 인터넷 플랫폼이 이제 범세계적인 산업 및 경제의 성장뿐 아니라 인간 삶의 형태까지도 극적으로 변화시키고 있음(European Commission, 2016)
- 이러한 중요성에도 불구하고, 디지털 플랫폼의 국가적 기여에 대해서는 아직 많은 연구가 진행되지 못했고, 소수의 논문만을 확인할 수 있음
  - 예를 들어, Ryu et al, (2018)은 디지털 플랫폼의 경제적 기여도와 광고 산업, 콘텐츠 산업 등 일부 산업영역에서 창출하는 경제적 부가 가치를 추정하는 시도를 한 바 있음
  - 이러한 연구들은 결국 플랫폼이 활성화되면 사회적 연결(Social connections)이 촉진되고 이를 통해 다양한 경제적 부가가치가 생성될 수 있음을 지적함 (Cui et al., 2020)
  - 한편, 플랫폼의 사회적 가치에 대한 연구들도 일부 존재하는데, 이들은 플랫폼이 안전과 보안이라는 측면에서도 사회적 가치를 창출하고 (Acheampong, R.A. (2021), 디지털 행동주의를 통해 사회적 가치를 창출(Social value creation)된다는 주장함
  - 플랫폼의 국가적 차원의 기여를 고려한 연구에서 플랫폼화(platformisation)는 국제사회에서 남북, 동서 등 전통적인 이분법을 넘어섬. 즉 온라인 플랫폼(online platforms)은 글로벌 콘텐츠 산업에서 소위 ‘문화예외 논리(cultural exception’ norm)’를 무력화시키면서 문화 제국주의를 강화한다고 주장함 (Vlassis, A, .2020)
- 이렇듯 인터넷 플랫폼은 한 국가의 경제 사회 문화 전반에 걸쳐서 큰 영향을 주는 핵심적인 역할을 담당

- 이러한 중요성에도 불구하고, 플랫폼의 다면적 특성으로 인해, 인터넷 플랫폼이 지닌 가치를 정확히 인식하기가 쉽지 않음
- 때로는 일부 영역에서는 플랫폼이 기존 산업영역과 충돌하면서 플랫폼의 반경쟁적 행위와 파괴적인 모습만 강조되기도 함
- 이러한 상황에서 특정 국가는 플랫폼의 가치를 다차원적 틀에 의해 객관적으로 분석하고, 이에 기반한 전략적 의사결정을 내릴 필요가 있음

## 나. 로컬 플랫폼의 정의

- 플랫폼이란 용어는 다양한 의미로 사용되어 왔는데, 플랫폼은 실생활에서 물리적 구조물, 예를 들어 철도 플랫폼, 우주선 발사대 등 작업을 위한 공용화된 토대의 의미로 사용되어 옴
- 이후 IT 기술이 발전하면서, 플랫폼은 응용 소프트웨어가 작동되도록 하는 컴퓨터 구조, 운영체제(OS), 프로그램 언어 관련 라이브러리, 미들웨어, 핵심 응용프로그램 등을 의미로 확장
- 플랫폼은 경제적 의미로도 해석될 수 있는데, 경제적 플랫폼은 여러 참여자가 공통된 사양이나 규칙에 따라 경제적 가치를 창출하는 토대에서 중개 기관(Intermediaries)의 역할을 함
- 디지털 플랫폼은 기술적, 경제적 의미에서의 플랫폼 역할을 수행한다고 볼 수 있음
- 디지털 플랫폼이 특정 국가에게 미치는 영향이 커지면서, 플랫폼 산업과 기업에 대한 가치 역시 재조명될 필요성이 대두되고 있음
  - 기존에도 인터넷 플랫폼과 관련해서 플랫폼 특징(Vakeel, Khadija Ali et al, 2020; Hein, A. et al, 2019), 혹은 경제적 파급효과(오재섭&송일호, 2017; 정은영, 2011), 규제적 접근의 필요성(Walker, M., 2020; 박미영, 2018; 곽관훈, 2018) 등 산업 및 경제적 관점에서 다양한 논의가 진행되어 옴
- 최근 G7 회의를 비롯한 글로벌 정상들의 회의에서 플랫폼 산업의 혁신과 경쟁을 유지하고자 하는 것은 단골 주제로 떠 오름
  - 자체 검색엔진을 보유하고 있지 못한 프랑스, 독일과 유럽연합은 뒤늦게 자국의 검색엔진을 구축하기 위해 천문학적인 돈을 투자한 국가 차원을 프로젝트

- (‘콰에로(Quaero)’<sup>99)</sup>를 진행했지만, 참여국 간에 견해차를 좁히지 못하고 2년이 채 안 된 시점에서 실패로 끝이 난 경험이 있으나(O’Brien, 2006), 이러한 노력에도 불구하고 유럽은 여전히 로컬 플랫폼이 부재한 상황이 이어지고 있음
- Hermes et al (2020)은 실질적으로 미국이 플랫폼을 장악(American platform domination)하고 있고, a self-sufficient platform economy(자급자족 플랫폼 경제)를 통해 새로운 플랫폼을 육성해 왔다고 지적함
  - 반면, EU의 경우 the risk of exploitation by American dominant platforms에 직면, 사실상 EU 자체 플랫폼 공백(the EU’s platform gap) 현상이 지속되고 있다고 지적한 바 있음
  - 세계 최대 경제주체임에도 불구하고 인터넷 대기업이 없어 미국에 비해 디지털 경쟁력이 뒤쳐진다고 평가받고 있음
    - EU의 경우 인터넷 검색시장의 90% 이상을 구글, 온라인 구매 시장의 20% 이상을 아마존, SNS 시장의 60% 이상을 페이스북 등 미국 인터넷 기업이 점유하는 등,
  - 최근 들어 생성형 AI가 화두가 되는 상황에서, 유럽은 초거대 GPT투자에 적극적으로 나서고 있으나 초거대 AI 역시 이용자 데이터를 기반으로 학습이 이루어지기 때문에, 자체적인 플랫폼을 보유하지 않고서는 그 경쟁력을 확보하기가 쉽지 않음
  - 지금 세계 플랫폼 시장은 미국과 중국이 주도
    - 이 두 나라 간 플랫폼 주도권 싸움은 2020년 도널드 트럼프 미국 대통령의 틱톡 매각과 위챗 금지 행정명령 조치를 내리는 데까지 영향을 미쳤고, 미국은 틱톡 등 중국의 플랫폼을 미리 차단하여 디지털 플랫폼 산업에서 미국의 주도권을 놓치지 않겠다는 강력한 메시지로 해석할 수 있어, 이렇듯 해외에서는 플랫폼 산업의 주도권 경쟁이 심화(Hermes et al., 2020).

99) <http://www.theguardian.com/technology/2006/apr/26/news.france>

**(1) 로컬 플랫폼 성공의 이유**

- Tian & Lan(2009)은 인터넷 마케팅에서 이용자들이 원하는 콘텐츠를 준비하는데, 이 과정에서는 지역 사람들의 정치/사회/문화적 경계를 이해하는 것이 특히 중요하다고 강조
  - 인터넷 마케팅을 완전히 소화하기 위해서는 마케터들이 다양한 문화적 가치에 대해 고려하여 그들의 목적을 분명히 해야 한다고 지적함
  - 또한 인터넷 마케팅과 전통적 마케팅의 차이점을 분명히 알고 인터넷 마케팅을 적절히 사용하기 위해서도 문화적 차이에 대한 이해가 필요하다고 볼 수 있어, 전통적 마케팅에 비해 인터넷 마케팅은 시간과 거리의 제약이 없다는 점에서 문화적 차이를 고려하는 것이 더 중요
- 문화적 차이는 인지 능력 차이를 야기함(Choong & Salvendy, 1998)
  - 예를 들어, 이모티콘의 선호도에 있어 미국인은 문자가 혼합된 모드를, 중국인들은 그림을 이용한 것을 더 잘 인식한다. 이러한 인식 능력의 차이에 따라 각 국가의 웹 서비스 디자인 역시 영향을 받을 수밖에 없음(Po Chuan, 1999; Evers, Kukulka-Hulme, & Jones, 1999).
- Fang & Rau (2003)은 언어만 다른 버전의 야후 서비스에 대한 미국과 중국 이용자들이 이용패턴과 만족도 차이를 분석
  - 야후 차이나는 미국인들을 위해 만들어진 카테고리화 서비스를 번역만 한 것이기 때문에 중국인들의 사용 만족도가 낮게 나타났고, 정보 검색에서도 중국인들이 더 많은 절차를 밟고 있다는 것이 확인
  - 사용 패턴(검색을 많이 쓰는지/카테고리를 클릭하며 읽는 것을 많이 하는지 등)에서도 차이를 보임
- 이렇듯, 인터넷 서비스는 정교한 인지적 과정이 포함되므로 인지적 특성에서 차이가 나타나는 문화적 차이를 면밀히 고려하는 것이 필수적이라 할 수 있음
- 인터넷 서비스에서 고객 신뢰를 증가시키기 위해서는 문화적인 차이를 고려해야 할 필요함(Sia et al., 2009)
  - 결과적으로 글로벌 플랫폼 사업자들이 로컬화를 서비스 확산의 핵심 성공 요인으로 보고 다양한 국가에서 현지화를 위한 노력을 해 왔음을 방증

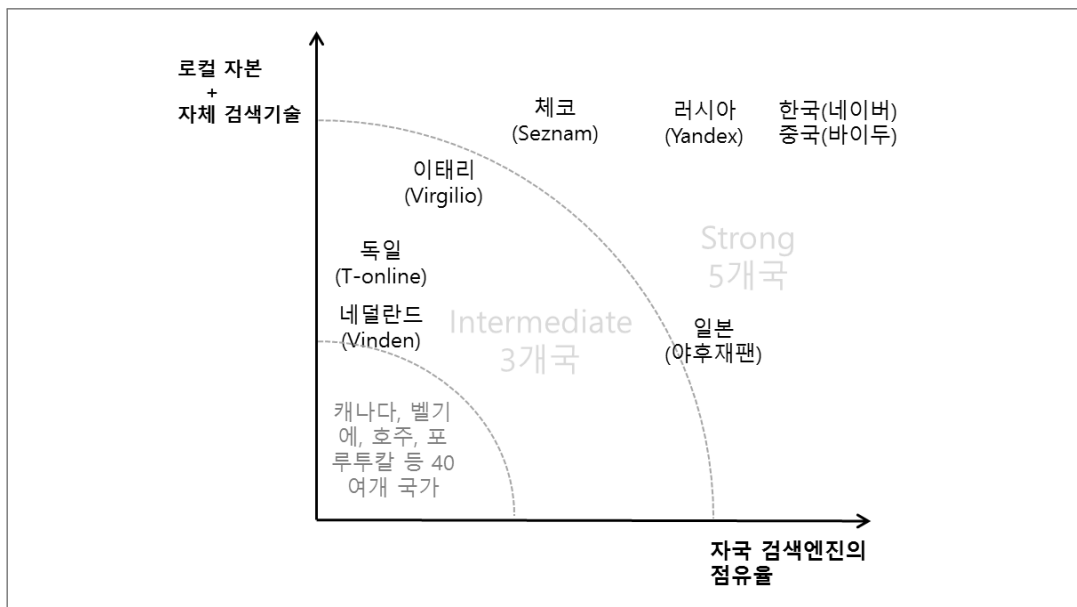
- 전 세계 대부분의 나라에서 구글 페이스북 등 미국 중심의 글로벌 플랫폼이 보편화된 것에 반해서, 국내에는 네이버 카카오 등 경쟁력 있는 로컬 플랫폼이 존재
- 국내에 이러한 로컬 플랫폼이 존재하는 이유에 대한 연구는 전무한 상황
  - 일부 중국 등의 플랫폼을 대상으로 자국의 독특한 언어가 있다는 것이 로컬 플랫폼 성공의 요인으로 분석된 바가 있으나,
  - 전 세계에서 독자적인 로컬 언어를 구사하는 나라들 모두에서 로컬 플랫폼이 성공하지 못했다는 점에서 자체 언어 보유 여부만으로는 로컬 플랫폼 성공의 이유를 충족하지 못함

## (2) 로컬 플랫폼의 구분

- 일반적으로 소프트웨어 산업에서, 해당 국가의 언어 및 지역적인 특성에 맞추는 작업을 현지화(localization)라고 일컬음
  - 이러한 기본정의에 의하면 구글 등 글로벌 검색엔진(global search engine) 역시 각국의 언어로 된 서비스를 제공하는 등 현지화된 서비스를 제공하고 있다고 볼 수 있음
- 본 보고서에서는 현지화된 글로벌 플랫폼 서비스와 자국에 기반을 둔 플랫폼을 구분
- 자국의 플랫폼은 일반적인 '로컬'의 정의보다 지역적인 개념으로, 자체 검색기술과 현지 자본 투입 여부를 기준으로 세분화해서 정의할 수 있음(Ruy et al, 2019)
- 자본과 기술 이외에도 현지 국가 중심으로 판단했을 때 본사의 위치, 조세 기준 국가 등을 기준으로 로컬 플랫폼 여부를 판단할 수 있음
- 반면, 글로벌 플랫폼은 기본적으로 한국어 이외의 언어(영어 등)를 기반으로 2개국 이상에서 인터넷 기반의 플랫폼 서비스를 제공하는 사업자로 정의할 수 있음
  - 해당 정의에 의하면 구글, 페이스북, 넷플릭스, 아마존 등이 글로벌 플랫폼으로 분류될 수 있음
  - 글로벌 플랫폼이 해당 국가에서 자회사 형태로 로컬화된 서비스를 제공하는 경우는 기술적, 경영적 독립성이 없기 때문에 글로벌 플랫폼으로 분류

- 예를 들어, 검색시장의 경우, Ji et al(2019) 연구에서 자본 및 검색 기술 차제의 Origin을 기준으로 로컬 검색을 구분하여 제시한 바 있음
- 이 기준을 적용할 때, 전 세계적으로 점유율 3% 이상의 자국 검색엔진을 보유한 국가는 한국, 일본, 중국, 러시아, 체코, 네덜란드, 독일, 이탈리아 등 총 8개국이고, 자국 검색엔진 점유율 합계가 40%를 넘는 국가는 한국, 일본, 중국, 러시아, 체코로 총 5개국으로 보고된 바 있음

[그림 IV-9] 로컬 플랫폼 구분을 위한 기준 예시



- 최근에는 스마트폰이 보편화되고 앱을 통한 서비스가 일상화되면서 특정 영역 (예를 들어, 간편송금, 공유 자동차)에서 성공을 거둔 후, 다양한 기등들을 덧붙이는 식으로 확장해 나가는 슈퍼앱 전략이 성공하고 있음
  - 국내의 경우 카카오톡이 메시징 서비스앱에서 성공을 거두고, 해당 성공을 기반으로 미디어, 페이, بانک, 쇼핑 등 다양한 서비스를 추가하는 방식으로 진화하고 있다. 토스의 경우도 간편송금으로 시작해 다양한 금융 관련 서비스를 추가하면서 금융을 중심으로 한 슈퍼앱으로 성장하고 있음
  - 해외에도 슈퍼앱이 등장하고 있는데, 대표적인 예가 싱가포르 중심의 Grab 서비스가 있는데, 우버와 같은 공유 모빌리티 서비스로 시작했지만, 현재는 택시, 음식배달, 금융 등 다양한 서비스를 제공하는 슈퍼앱으로 성장

- Grab과 유사한 슈퍼앱으로 인도네시아의 GoJek 서비스가 있음
- 슈퍼앱까지 고려하더라도 여전히 로컬 기반의 플랫폼이 성공한 경우는 아주 드물다고 할 수 있음

## 다. 로컬 플랫폼의 의미와 가치

### (1) Internet sovereignty (인터넷 주권)

- 디지털 플랫폼은 국가 경쟁우위의 핵심으로 여겨지고 있음
  - Hermes, S. et al. (2020)는 현재 전세계 플랫폼 경제는 미국과 중국 양 국가가 주도하고 있다고 지적한 바 있으며, 중국의 경우, 자급자족 플랫폼 경제를 구축했지만, 현재는 미국과 경쟁하는 유일한 국가로 성장
    - 미국: platform dominance, American platform domination(플랫폼 지배)
    - 중국: a self-sufficient platform economy(자급자족 플랫폼 경제), the emergence and growth of digital platform, platform competitiveness
    - EU: the risk of exploitation by American dominant platforms, the EU's platform gap(플랫폼 공백)
- EU의 경우 플랫폼 공백 사태를 맞고 있는데, EU는 이러한 “플랫폼 공백”을 해결하기 위해 EU는 미국 플랫폼들의 반경쟁 행위에 대해 엄격한 잣대를 통해 발 빠르게 규제해 옴
  - 특히 최근에 와서 경쟁법<sup>100)</sup>, DMA<sup>101)</sup>, DSA<sup>102)</sup> 등을 발의하고 플랫폼 사업자들이 준수해야 하는 사항들을 규칙으로 지정하며 강력한 규제 의지를 내비친 것에는 공정한 경쟁시장과 이용자 보호라는 1차적인 목적이 있지만, 대부분의 법안 규제 대상에 미국의 거대 플랫폼이 포함되는 것으로 보아 이는 기존의 미국 거대 플랫폼들을 규제하여 자국의 플랫폼 산업을 보고하고자 하는 의도 역시 그 이면에 숨어있다고 볼 수 있음

100) A New Competition Framework for the Digital Economy Report by the Commission 'Competition Law 4.0' 참조

101) 유럽연합 디지털시장법안(Digital Markets Act) 게이트키퍼 지정요건 및 의무·금지사항 참조

102) EU 디지털서비스법안(Digital Services Act)의 주요 의무사항 참조

**(2) 데이터 경제 패권(헤게모니)**

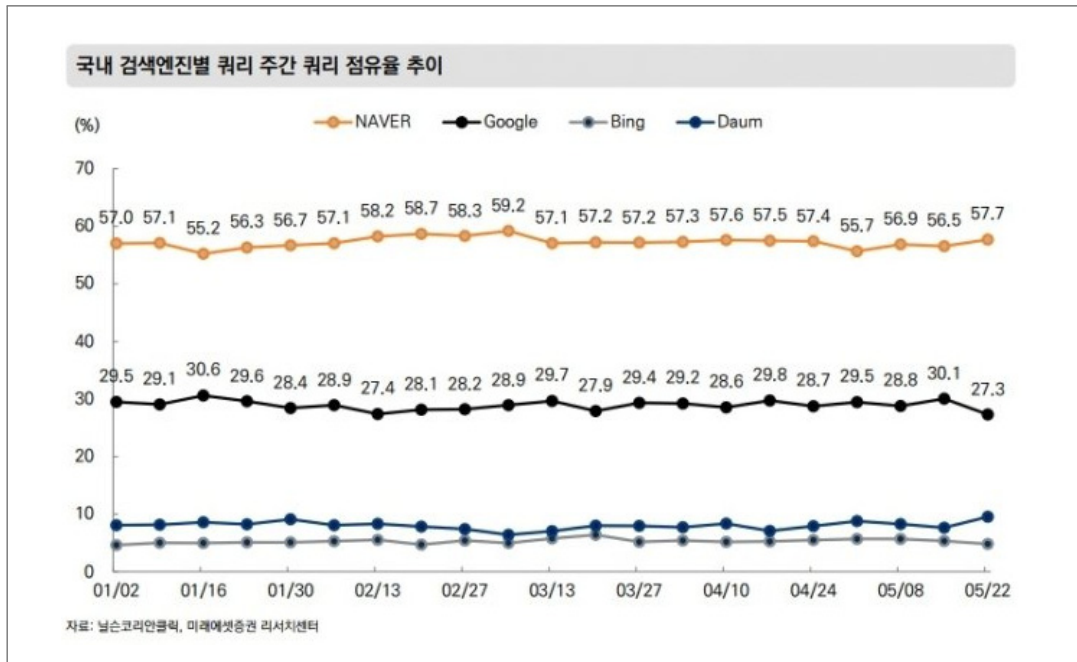
- 데이터는 인공지능과 결합하면서 다양한 산업의 경쟁력을 결정하는 핵심 원천
- 따라서, 인공지능이 활성화되면 될수록, 데이터를 가진 기업과 그렇지 않은 기업 간 격차가 점점 벌어지는 ‘데이터 외부성’ 문제가 발생할 수 있음
- 이를 해소하기 위해서는 제조업을 비롯한 국내 다양한 기업군들의 데이터 확보 및 활용에 대해 국가 주도적인 전략이 필요한데, 특히 해외 플랫폼 사업자들을 통한 국내 이용자들의 데이터 유출 문제 역시 고려해야 함
  - 유럽연합의 경우 일반개인정보보호법(GDPR)을 통해 EU 내에서 생성되는 개인정보에 대해 역외 이전이 가능한 경우를 규정하여 자국 내의 데이터를 보호하고 있음
  - 또한 미국과 중국 역시 오래전부터 데이터 경쟁력을 확보하기 위해 국가적 차원에서 노력해 옴
  - 한편, 글로벌 플랫폼 사업자를 통해 개인정보가 EU 경계를 넘어서 이동되는 경향이 가속화되자, 유럽은 개인정보 유통에 대한 통제를 강화하기 위한 개정안을 발표
    - 2012년 개인정보 보호 개정안은 EU 내 거주하는 정보주체들의 개인 정보가 영외로 이전될 경우, 이에 대한 동의를 의무화하는 한편, 정보주체가 자신의 이익을 보호하기 위해 필요하다고 판단될 경우, 필요한 조치를 요구할 수 있도록 하는 권리를 제공
  - 일부에서는 이를 두고 EU 회원국과 비회원국 사이에 이러한 정보의 유통도 엄격하게 통제하는 것을 정보봉쇄(data blockages)라고 표현하면서 비판하기도 함(박정훈, 2013)
  - 유럽이 개인정보보호의 법적 규제 수준을 강화하는 이면에는 SNS, 클라우드 컴퓨팅 서비스 등 개인정보를 활용한 인터넷 플랫폼 서비스들이 확산되는 시점에서 미국의 영향력에서 벗어나고자 하는 의도가 있다고 볼 수 있음
- 중국은 디지털 경제 패권을 위해 국가가 전략적으로 자국의 플랫폼을 육성한 사례로 볼 수 있음
  - 예를 들어, 위챗 플랫폼 모델의 인프라화는 기술국가주의(techno-nationalist) 미디어 규제와 명시적인 사이버 주권(cyber-sovereignty) 어젠다에 의해 형성된 것으로 볼 수 있음

- 디지털 플랫폼 경쟁은 검색엔진, 인공지능, 데이터 국지화, 전자상거래와 핀테크 등 다양한 분야로 확대되어 주요국의 기술 패권 경쟁양상으로 비화되고 있음 (Plantin, J. & Seta. G., 2019)
- 미국과 중국이라는 두 개의 플랫폼 사이에 있는 한국은 이러한 디지털 플랫폼 경쟁의 부상에 대응하는 미래 국가전략을 고민해야 함(김성배, 2022).

### (3) 글로벌 빅테크 견제

- 국내에 로컬 플랫폼이 존재하기 때문에, 구글 등 글로벌 플랫폼 사업자들은 국내 시장에 대한 경쟁력을 유지하고 시장을 확대하기 위해서 끊임없이 노력함
  - 실제로 구글의 경우 국내 검색 점유율이 꾸준히 증가하고 있는데, 구글이 국내 시장에 쏟는 노력은 최근의 초거대 AI 및 생성형 AI과 관련해 한글 기반의 모델을 영어를 제외한 언어 중 최우선으로 적용하는 것만 보더라도 쉽게 확인할 수 있음

[그림 IV-10] 국내 검색 점유율 변화 추이



자료: <https://news.mt.co.kr/mtview.php?no=2023061213434572239>

- 결국 국내에 로컬 플랫폼이 존재함으로써 글로벌 플랫폼 사업자들이 지속적으로 국내 시장에 대한 투자를 지속하게 하는 효과가 있다고 볼 수 있음

#### (4) 문화제국주의에 대한 대응 (Cultural Imperialism)

- 디지털 플랫폼은 표준화된 절차를 통해 확장해 왔지만, 이와 동시에 각 국가와 지역의 특성에 맞도록 분화해 옴
- Zahir, Dobing, & Gordon Hunter(2002)은 인터넷 플랫폼의 로컬화되는 과정을 문화적 맞춤(customization)의 개념으로 설명함
  - 신기술은 수렴(문화들이 서로 비슷해지는 것)과 확산(문화가 서로 다른 방식으로 기술에 적응하여 차이점이 부각되는 것)을 모두 다 야기하는데, 실제로 서로 다른 나라들의 웹 포털을 보니 두 현상이 모두 나타난다고 설명
  - 로컬 플랫폼의 경우, 기본 구조는 야후(글로벌 플랫폼)와 비슷하지만, 구체적인 기능과 특징들은 문화적 다양성에 따라 차이를 보인다고 설명함
- 특히 인터넷 서비스는 생산 비용이 상대적으로 저렴하기 때문에 문화적 맞춤(customization)의 기회가 더 클 수 있음
  - Marcus and Gould(2000)는 다양한 문화권의 웹페이지를 분석한 결과 문화는 웹페이지의 디자인에 영향을 준다고 결론을 내린 바 있음
- 반면, 최근 들어, 플랫폼이 각 국가의 문화적 특성에 맞도록 분화하고 있다는 지적에 반대되는 주장이 제기
  - 플랫폼화(platformisation)는 국제사회에서 남북, 동서 등 전통적인 이분법을 무너트리고 새로운 문화적 경제를 만들어 가고 있으며(Vlassis, A., 2020), 이는 온라인 플랫폼(online platforms)은 글로벌 콘텐츠 산업에서 소위 ‘문화 예외논리(cultural exception’ norm)’를 무력화시키면서 있다고 비판
- 이러한 비판을 가장 적극적으로 수용하고 있는 곳 역시 자체적인 로컬 플랫폼을 보유하고 있지 못한 EU임
  - 앞서 살펴본 바와 같이, EU 국가들은 미국의 글로벌 플랫폼에게 강한 의무를 부과하는 적극적인 규제를 도입하고 있는데, 이는 경제적 이익과 더불어 온라인 문화주권(cultural sovereignty online)을 지키는 조치로도 해석할 수 있음

## 라. 로컬 플랫폼 가치 측정 방법론: 선행 연구 정리

### (1) GDP 기반(시장 가치 기반) 방법론 및 한계

- GDP는 정치 경제적 이해관계자들의 절대적인 지표가 되고 있으며, 이는 생산된 가치의 합으로 소득, 소비, 생산성 측정에 기반한 방법론임
- GDP는 2차대전과 대공황 시기에 각 국가의 생산 가치를 측정하여 그에 따라 부양책을 쓰기 위해 고안된 지표로, 제조업에 초점이 맞추어져 있어, 혁신과 질적 성장을 측정하지 못함
  - 소프트웨어 무료 업그레이드, 에어비앤비의 집, 페이스북에서의 엔터테인먼트 등의 혁신도 측정 불가
- 한편, GDP는 시장 내 거래만 측정하기 때문에 장외 생산, 특히 가구에서 생산하는 가치(ex자가주택) 측정 누락은 오랫동안 논란이 되어왔음(economist, 2016).
- GDP를 디지털 플랫폼 지표로 활용하는 것에 대한 문제점을 아래와 같이 정리할 수 있음
  - 재화/서비스의 구분: 데이터는 무형이어서 일반적으로 서비스로 구분되고 있지만, 저장·변형·수집 가능하고 최초 생산 시점과 다른 시점에 소비되므로 서비스와는 다름(Mandel 2012). 인터넷 접근량 성장률은 2011년 이후부터 감소하고 있으나, 2012년 기준 앱 마켓의 앱 수(앱스토어 70%, 안드로이드 100% 성장), 모바일 데이터량(120%), 주요 온라인 서비스 이용자 수(평균 약 30%) 등은 현격히 증가한 것을 확인할 수 있음(Mandel, 2012).
  - 가격이 없거나 매우 낮은 상품. Free media: 양면시장의 형태로, 상품에 대한 직접적인 소비 지출이 없고, 대신 사람들의 주목 (Brynjolfsson&Oh, 2012)이나 광고시청동의(Soloveichik&Nakamura, 2015), 빅데이터 획득(Ahmad&Schreyer, 2016)을 수익원으로 하여 생산됨. 공짜 상품은 예전부터 있던 것이지만, 인터넷/모바일 혁명으로 이 부분이 굉장히 커졌다는 점이 중요. 가격을 지불하는 상품(주로 구독 기반)의 경우에도 매출이 산출량을 측정하지 못함. 산출량과 소비량이 비례하지 않기 때문. 같은 값을 지불하고도 훨씬 많은 소비를 할 수 있다.

- 공유경제: 물건의 생산이 없어도 가치가 창출됨. 소비자가 생산하는 가치들 (prosumer). 이전에는 GDP에 잡혔을 생산 측면에 가구(household)가 참여하게 되면서 생산으로 간주되지 않는 상황. 이것을 측정해야 하는가 논란은 있으나, 만약 측정하게 된다면 이전에 전문적 기업에서 생산하던 상품과 동일한 가격의 가치를 창출하지는 않는다는 점(가격 인하, 품질 인상 등의 변화)을 고려해야 할 것(Ahmad&Schreyer, 2016)
- 공유경제의 투자 측면: 소비자 내구재가 중복으로 사용되는 효과. GDP 측정에 변화는 없으나 자본의 활용, 생산성에는 영향을 줌(자본투자는 1만 해도 2를 투자한 효과의 생산/소비가 이루어질 수 있음). (Ahmad&Schreyer, 2016).

## (2) 인터넷의 비시장(non-market) 가치 평가 방법

### □ 조건부(가상) 가치평가 Contingent valuation/ Conjoint analysis

- 가격이 없는 상품의 경우 소비자잉여를 계산함으로써 그 가치를 측정할 수 있음
- 조건부 가치평가 모델은 환경재의 가치평가에서 많이 사용되는데, 해당 방법론은 표본 조사를 통해, 개인들로부터 가상의 가치들을 직접적으로 도출
  - 전통적으로 조건부가치측정 연구들은 기름유출, 화학사고와 같은 미래의 불운한 환경 사건 가능성을 감소시킬 수 있는 가상적인 정부사업에 관한 정보를 제공하고, 응답자들이 그 환경사업을 지지하기 위해 어느 정도의 경제적 희생을 감수할 수 있는지를 알아보기 위해 사용되었음(Arrow et al., 1993)
  - 조건부가치측정 도출의 가장 이상적인 것으로 권고하는 것은 양분선택형 질문법임(Arrow et al., 1993)
  - Rosston, Savage, Waldman(2010)은 인터넷 사용자 5799명을 대상(비사용자 472명) 대상으로 총 8개 시나리오에 대한 선택을 조사하고, 이 선택을 자신의 현재 인터넷 서비스와 비교하여 선택하고 이 결과로부터 소비자의 효용곡선과 WTP 도출(Random utility model로 측정)한 바 있음
  - 또한, 브로드밴드의 가치(속도/안정성 등)를 평가하면서 그 안에 추가적인 사항으로 미디어 다운로드, 인터넷 전화 등 인터넷 서비스의 가치도 평가하였는데, 그 결과, 인터넷을 활용한 건강서비스 이용에는 4.39\$, 비디오전화는 5.06\$, 영화 다운로드에는 3.29\$의 평균 한계 WTP가 도출됨

□ Hedonic price model

- Hedonic price model은 다양한 속성을 지닌 상품에 내재되어 있는 각 속성의 가치(암묵적 가격)를 실제 소비 패턴을 통해서 측정하는 방법
  - 이는 소비자들이 상품의 각 속성을 잘 이해하고 있고 자신에게 최적화된 소비를 하고 있다고 가정함(Lee,Park&Kim, 2002)
  - Lee,Park&Kim(2002)은 일주일간의 인터넷 사용 조건을 기록. 임의의 가격함수(Box-Cox) P에 각 속성(이메일, 쇼핑, 검색, 은행, 뉴스, 건강 등)의 일주일간 사용패턴을 집어넣어서 각 요소의 계수를 도출
  - 이후, 탄력성(가격 올릴 때 사용 정도의 변화) 계산해 가격함수에 대입함으로써 WTP 도출했다. 그 결과, 검색 서비스 이용이 1% 늘어날 때 이용자들의 평균 한계 WTP는 16원, 쇼핑의 경우 8원, 건강의 경우 -5원으로 도출된 바 있음

□ Welfare analysis

- 인터넷 서비스에 대해 가격과 시간 측면에서의 효용을 추정하기 위한 몇 가지 시도가 있었음
- 가격지출 측면에서, Greenstein & McDevitt(2009), Brynjolfsson & Oh(2012)은 인터넷 관련 서비스에 지출한 가격 변화를 놓고 탄력성 계산해서 이용자 잉여(CS, Consumer Surplus)를 계산한 바 있음
- 시간측면에서, Goolsbee & Klenos(2006)은 가격이 거의 들지 않는 상품의 소비자 잉여(consumer surplus) 계산하는 방법을 제안함(소비자 후생을 알기 위해서는 수요탄력성을 알아야 하지만, 가격이 0인 경우엔 다른 방법이 필요)
  - 수요의 가격탄력성을 소비에 들이는 시간, 그 시간에 대한 노동 기회비용을 활용하여 도출. Ex) 인터넷 이용에 1시간을 투자한다면, 1시간 동안 노동을 했다면 별 수 있었던 수입을 포기하는 것으로 가정하고, 인터넷 이용 시간, 인터넷 외의 다른 상품에 지출하는 비용, 노동 임금 간의 효용함수를 계산해서 인터넷 이용의 가격탄력성 산출. 탄력성을 알면 소비자잉여를 구함
  - 그 결과, 인터넷 소비자 잉여는 평균 일인당 총 수입의 2~3%(탄력성 계산에 따라 달라짐), 이는 일인당 \$2,500~\$3,800 정도로 추정됨

- Brynjolfsson & Oh (2012)는 온라인 이용 시간의 2/3이 무료 사이트 이용에 해당한다고 가정하고, 인터넷에 사용한 시간적 가치를 노동-여가 대체재로 놓고 탄력성 계산해서 CS 도출
  - 그 결과 인터넷 사용의 평균 소비자잉여는 8,380억 달러(GDP의 5.83%), 그 중 무료 사이트 이용의 잉여는 5,590억 달러(GDP의 3.89%)로, GDP 성장에 기여하는 바가 TV보다 높게 나타남
  - 특히, 페이스북, 유튜브, 트위터, 위키피디아 등 주요 인터넷 서비스의 연간 소비자잉여를 계산한 점을 주목할 만함

[그림 IV-11] 주요 인터넷 서비스의 연간 소비자잉여를 계산 예 Brynjolfsson&Oh (2012)

	Reach%	Minutes	Time share	Yearly CS (\$Billion)
Facebook	0.434	24	16.00%	16.7
YouTube	0.330	17	8.62%	9.1
Twitter	0.093	7	0.99%	1.1
Wikipedia	0.144	4	0.88%	0.9
LinkedIn	0.050	7	0.53%	0.6
Craigslist	0.015	13	0.30%	0.3

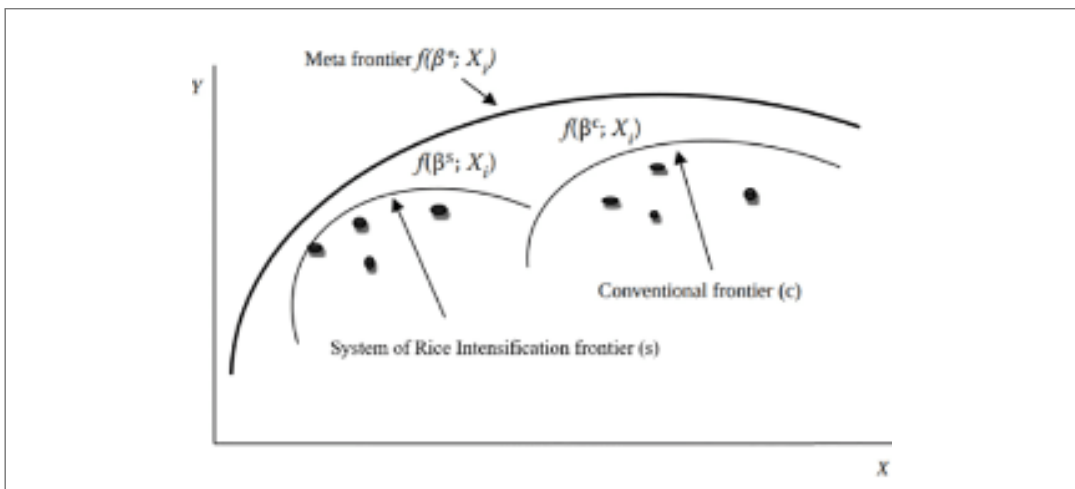
- Dean et al(2012) 이용자들이 인터넷 장치, 어플리케이션, 서비스, 인터넷 접근을 이용하면서 얻는다고 체감하는 가치를 묻는 서베이 실시한 바 있음
  - 현실 물가에 맞추기 위해 위험 혐오 효과(loss aversion technique)를 활용하여 숫자를 조정
    - 그 결과, 한국의 소비자잉여는 연간 일인당 453 달러, 그 중 검색 서비스는 96달러, 이메일은 87달러, 온라인 금융은 74달러로 나타남
- 한편, 가격을 임의로 부여하여(shadow price 설정) 가치의 측정이 가능
  - 이는 소비자 잉여를 측정하는 것은 소비자 가치평가와는 다름

- Non-market product나 free-product가 아닌 전통적인 시장에 있는 상품 가치는 소비자 잉여로 측정되는 것이 아니다. 생산성과 소득에 대한 측정은 사회 후생 측정과는 다른 목적을 띠고 있어(Ahmad & Schreyer, 2016), shadow price의 측정이 이루어져야 함
- Soloveichik & Nakamura(2015)은 엔터테인먼트 소비자와 광고주 간의 광고 시청에 대한 바터무역(barter trade, 물물교환)을 간주하여 이 중 여가에 활용되는 미디어의 광고비(투입)를 광고시청 생산가격으로 보고 GDP 누락분 계산한 바 있음
- 미디어의 무료 상품의 가치는 광고 서비스의 가치(생산자의 가치평가)와 같다고 가정하는 것으로, 생산자의 가치평가는 소비자에 의한 가치평가와 일치하지 않을 수 있다는 단점을 내포함(Ahmad & Schreyer, 2016).

□ 확률적 변경접근법 (Stochastic frontier model)

- 메타프론티어(MetaFrontier)의 개념은 확률적 변경접근(SFA: Stochastics Frontier Approach)을 사용하여 그룹 사이의 상대적 효율성을 비교하는 분석 방법(O'Donnell et al., 2008)
- 아래 [그림 IV-12]와 같이, 각 그룹의 효율성을 확률 모형으로 표현하고, 이론적인 효율성 극대치와의 거리를 비교함으로써, 그룹 간 효율성을 비교

[그림 IV-12] 메타프론티어(MetaFrontier) 그룹 간 효율성 비교 예시



- 예를 들어, Ryu et al.(2019)은 의류 판매상을 대상으로, 온라인 채널만을 이용하는 그룹과 온라인과 오프라인을 동시에 활용하는 그룹의 효율성을 비교하고, 온/오프라인을 모두 활용하는 그룹의 효율성이 높다는 것을 계량적으로 분석
- 해당 방법론은 (이용자 및 사업자 대상 연구에서) 로컬 플랫폼을 사용하는 그룹 vs 글로벌 플랫폼을 사용하는 그룹에 대한 효율성 비교하는 연구 방법으로 활용 가능

〈표 IV-5〉 온라인 vs 온/오프라인 채널 이용 상점 간 효율성 비교 예시 (Ryu et al., 2019)

Statistics by group (unit: ten thousand KRW<sup>a</sup>).

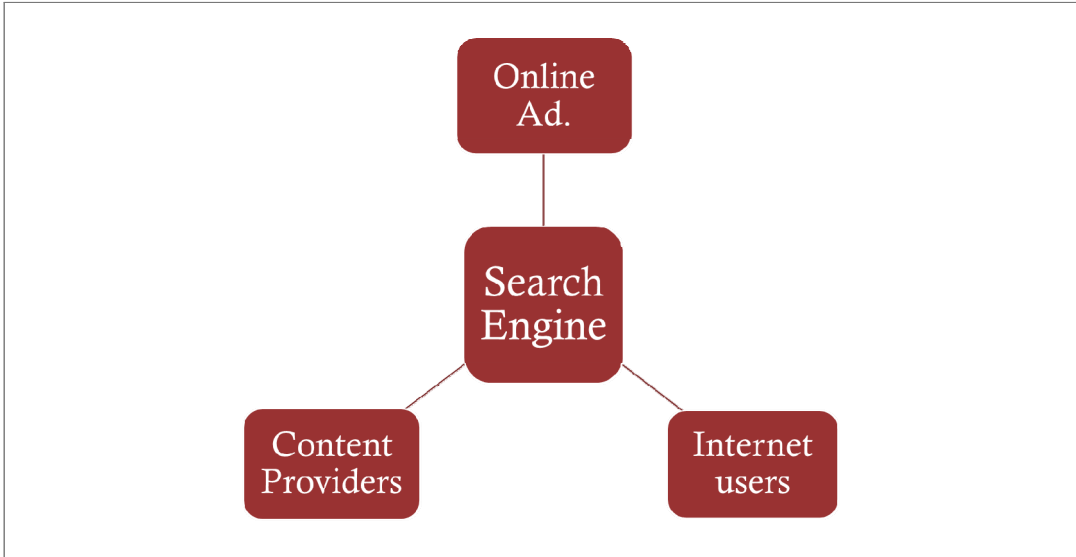
Classification	Group with offline channel	Group without offline channel
Sample size	163	119
Average sales	2510.1227	13,801.0504
Average labor costs	290.6380	755.1261
Average other spending	724.5092	834.6639
Average capital	3106.6196	4621.4286

<sup>a</sup> As of September 6, 2018, 1 USD is 1118.50 KRW.

□ 파급효과 기반 분석

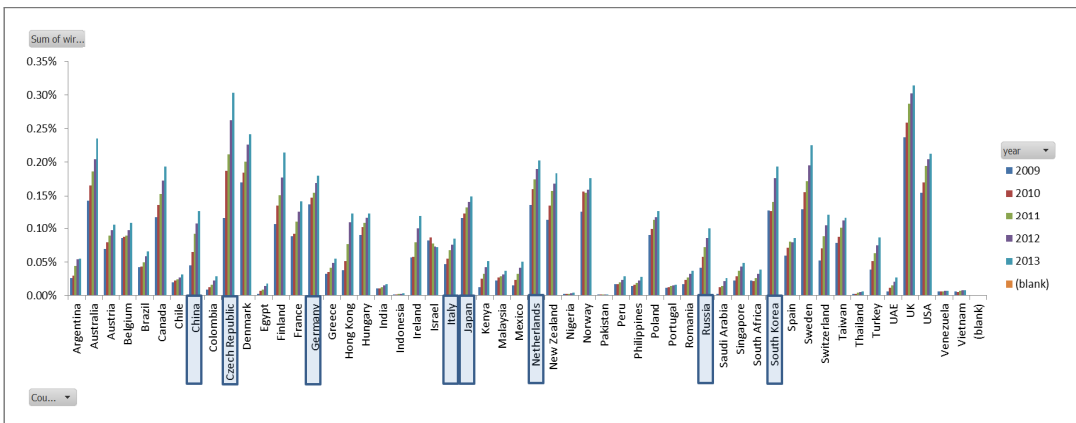
- 디지털 플랫폼 시장은 양면 시장 (two-sided) 혹은 다면 시장(multi-sided)으로 특징지을 수 있음(그림 IV-13) 참조
  - 예를 들면 네이버는 중심에 위치한 플랫폼으로 온라인 광고주 그룹, 인터넷 사용자 그룹, 그리고 콘텐츠 공급자 그룹들을 매개하는 역할을 함
  - 디지털 플랫폼 시장에서는 양면시장의 특징인 간접적 네트워크 외부성이 작동하여 인터넷 사용자의 증가는 온라인 광고시장에 영향을 주고, 동시에 온라인 콘텐츠 공급자의 증가에 영향을 주어, 각 그룹이 속한 시장의 매출을 증가시키는 효과가 있을 수 있음
  - 즉 온라인 광고자, 온라인 콘텐츠 공급자, 인터넷 유저 사이에 발생하는 네트워크 외부성을 통하여 각각의 그룹들이 이득을 얻는 선순환 효과가 발생
- 플랫폼의 이러한 특징을 이용해서, 로컬 플랫폼을 보유하는 것에 대한 관련 시장에 대한 파급력을 추정하는 접근법이 가능

[그림 IV-13] 검색시장 관련 연계 산업의 예시



- Ji et al.(2019)은 전세계 46개국의 검색점유율(자국 검색엔진 보유유무<sup>103</sup>), 온라인 광고산업의 크기, GDP, 인터넷 이용률, 컴퓨터 보급률, 고유언어 유무 등의 변수를 이용, 계량경제학 모델을 만들어 실증 분석을 진행
  - 분석 대상 국가는 아래와 그림과 같이, 파란 박스로 처리된 국가가 자국 검색엔진을 보유한 국가이며, 그래프의 Y 축은 해당 국가의 온라인 광고시장의 규모(GDP 대비 비율)를 나타냄

[그림 IV-14] 온라인 광고시장의 크기 (자료: Ji et al., 2019)



103) 해당 국가의 자국 검색엔진 보유 여부는, 로컬자본 비율과 자체 기술 보유 여부로 판단

- 분석 결과, 자국 검색엔진을 보유한 국가의 ‘온라인 광고시장의 GDP 점유율(인터넷 광고집약도)’은 그렇지 않은 국가들보다 0.029%높게 나타남
- 해당 결과를 2013년 기준 한국 상황에 적용해 보면, 전체 온라인(유선) 광고 시장의 크기인 약 2조 7천억원(\$2.5 billion)의 14%에 해당하는 3천6백억원 (\$350.89 million)의 순증가를 의미함

〈표 IV-6〉 자국 검색 보유 여부에 따른 광고시장 규모 추정 결과 (자료: Ji et al., 2019)

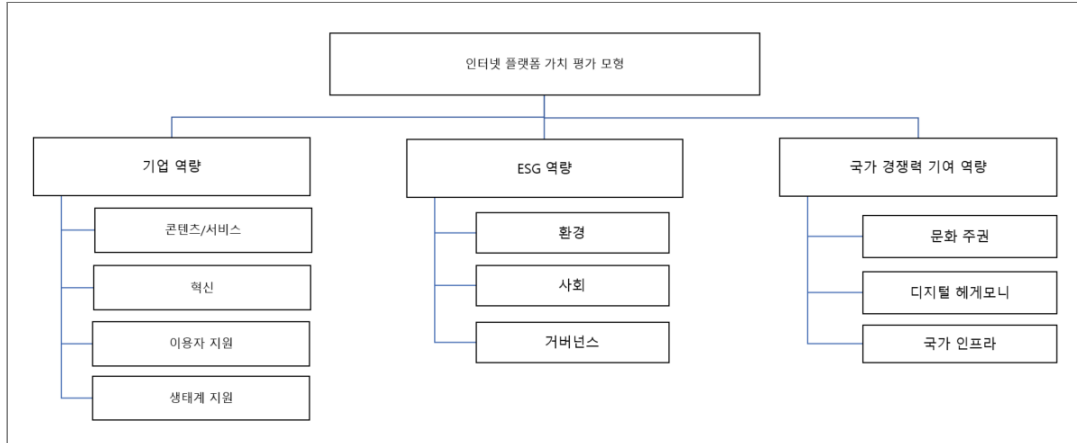
D.V = Online advertising intensity	(1) Pooled- OLS	(2) Fixed effect	(3) Random effect	(4) Arellano- Bover1	(5) Arellano- Bover2
<i>브로드밴드 이용률</i> (Broadband penetration)	-0.047 (-1.09)	0.07 (1.03)	-0.001 (-0.03)	0.015 (0.68)	0.016 (0.73)
인구수(Population)	0.000006 (0.40)	-0.00005 (-0.27)	0.00002* (1.74)	-0.00002 (-1.02)	-0.00001 (-0.47)
영어 외 타언어 이용 (Other language)	-0.06* (-1.81)		-0.057* (-1.74)	-0.004 (-1.41)	-0.029 (-1.45)
<i>컴퓨터 이용률</i> (Computer penetration)	0.221*** (5.87)	0.176*** (2.80)	0.206*** (4.15)	0.024 (0.97)	0.016 (0.61)
자국 검색엔진 보유 주요 5개국 Domestic search engine 5	0.053** (2.60)		0.047** (2.27)	0.029** (2.11)	
자국 검색엔진 보유 광의 8개국 Domestic search engine 8					0.018* (1.94)
상수 (Constant)	0.049 (1.39)	-0.037 (-1.50)	0.028 (0.81)	0.026 (0.93)	0.021 (0.94)
(1년전 온라인 광고시장 크기, GDP 비중) 1-year lagged online Advertising. Intensity				0.91*** (15.48)	0.942*** (13.39)
Observations	230	230	230	184	184
R <sup>2</sup>	0.6816	0.438			
F statistics / Wald Chi <sup>2</sup>	F(5, 45) = 29.51***	F(3, 45) = 15.09***	Chi <sup>2</sup> (5) = 92.62***	Chi <sup>2</sup> (6) = 4221.15***	Chi <sup>2</sup> (10) = 2962.88***
Arellano-Bond test AR(1)				-1.30	-1.37
Arellano-Bond test AR(2)				-1.46	-1.34
Sargan test				24.43	24.22

- 구체적인 분석 방법은 아래와 같음
  - 추정 방법은 전년도의 시장에서 보이는 경향성과 시장의 동태적 발전을 통제할 수 있는 Arrellano-Bover 추정법을 사용하였으며, 다른 추정법(Pooled OLS, Fixed Effect, Random Effect)에서 도출된 결과도 명시하여 Arrellano-Bover 모형의 robustness를 확인
  - 추정기법의 특성상 각 나라별 고유의 특성 (country-fixed effect)와 연도별 특성(year-fixed effect)도 추정모델에서 통제됨
  - 온라인 광고와 관련된 선행연구를 바탕으로, 자국 검색엔진의 존재 유무 외 온라인 광고산업에 영향을 줄 수 있는 변수들을 통제함 (Jeon, Jullien, & Klimenko, 2012; Van der Wurff, Bakker, and Picard, 2008; Zentner, 2012; Zhao & Tse, 2007)
  - 종속변인은 온라인 광고집약도 (online advertising intensity)로 온라인 광고 시장이 GDP에서 차지하는 비율로 정의함
- 해당 연구결과는 자국 검색엔진이 해당 국가의 온라인 광고산업의 성장에 긍정적인 영향을 준다는 연구가설을 입증
  - 논문은 이러한 결과가 도출된 이유를 양면시장 이론을 들어 설명
  - 즉, 양면시장 이론에 의해, (자국 검색이 존재하면 해당 국가의 검색엔진 시장 내 경쟁 수준이 높아지고) 검색시장 내 경쟁 증가로 인해 광고주가 양질의 광고를 다양한 사업자를 통해서 제공 가능해지면서 온라인 광고시장 전체 크기가 커질 수 있다고 설명함

□ AHP 분석

- 디지털 플랫폼이 국가에 기여하는 바를 평가하기 위한 프레임워크를 제시한 논문을 소개하고자 함
- Ryu et al (work in process) 논문은 인터넷 플랫폼의 가치를 평가하는 major 항목으로 기업, 지역사회 및 환경(ESG), 국가적 차원의 역할을 중심으로 평가할 수 있다고 주장
  - 즉 플랫폼의 역량을 평가할 때, 국가 경쟁력에 대한 기여분을 충분히 고려해야 한다는 것

[그림 IV-15] An AHP model for evaluating internet platforms



<b>기업 역량</b>	플랫폼 기업의 기업 본연의 경쟁력 및 경제적인 성과와 관련된 항목으로 콘텐츠 및 서비스의 경쟁력, 기술의 혁신성과 플랫폼의 양면 시장의 특징이 고려된 이용자와 공급자 생태계 관리 역량을 의미한다.
<b>ESG 역량</b>	플랫폼 기업의 장기적인 지속가능성과 사회에 미치는 영향 평가하는 항목으로 환경(E), 사회(S), 기업 지배구조(G)의 세 가지 요소에 대한 관리 역량을 의미한다.
<b>국가 경쟁력 기여 역량</b>	플랫폼 기업이 국가 경쟁력에 기여하는 정도를 측정하는 항목으로 문화 주권, 디지털 경제 헤게모니, 국가 인프라 강화 차원의 기여를 포함한다.

- 기업역량은 플랫폼 기업의 기업 본연의 경쟁력 및 경제적인 성과와 관련된 항목으로 콘텐츠 및 서비스의 경쟁력, 기술의 혁신성과 플랫폼의 양면 시장의 특징이 고려된 이용자와 공급자 생태계 관리 역량을 의미
- 지역사회 및 환경(ESG) 역량은 플랫폼 기업의 장기적인 지속가능성과 사회에 미치는 영향 평가하는 항목으로 환경(E), 사회(S), 기업 지배구조(G)의 세 가지 요소에 대한 관리 역량을 포함
- 국가 경쟁력 기여 역량은 플랫폼 기업이 국가 경쟁력에 기여하는 정도를 측정하는 항목으로 문화 주권, 디지털 경제 헤게모니, 국가 인프라 강화 차원의 기여를 포함
- 국가 경쟁력 기여 역량에 대한 하위 평가 기준은 다음과 같이 제시될 수 있음

문화 주권	플랫폼 기업이 해당 국가의 문화적 정체성을 유지하기 위해 어느 정도 기여하는지, 더 나아가 해당 국가의 문화가 세계 곳곳으로 전파될 수 있도록 기여하는 정도를 평가한다. 예를 들어, 플랫폼이 우리 문화의 정체성을 유지할 수 있도록 돕고, 서구 중심의 문화제국주의에 대응할 수 있도록 기여하는 정도가 평가될 수 있다.
디지털 헤게모니	데이터 경제의 주도권 확보 차원에서 데이터 주권을 유지하는데 필요한 역량을 평가한다. 또한, 특정 국가의 콘텐츠/서비스가 글로벌 영토로 확장하는 데 도움을 주는 정도를 평가한다. 예를 들어, 해당 플랫폼으로 인해 데이터 주권이 지켜질 수 있는지, 국내 이용자들의 이익이 최우선적으로 고려되는지, 국내 이용자들의 개인정보를 안전하게 보관하고 관리하는지 등이 평가될 수 있다.
국가 인프라	플랫폼이 재난 상황에 대한 대응 등 국가의 공적인 하부구조의 역할과 시민 참여를 통한 혁신을 유도하고 문화 예술의 지원 활동을 평가한다. 또한 국가 공동체 일원으로써 세금을 회피하지 않고 성실히 납부하는지를 통해 평가될 수 있다.

## 마. 시사점 및 결론

### (1) 로컬 플랫폼 경쟁력 강화를 위한 지원 방향

- 미국(FAANG)이나 중국(BAT)의 플랫폼 기업들은 빅데이터, 인공지능으로 대표되는 지능정보사회의 핵심 플랫폼을 장악하고 자체 글로벌 생태계 양분하고 생태계를 주도
  - 디지털 경제 체계에서는 플랫폼 기업의 경쟁력이 곧 국가 차원의 경쟁력으로 연결됨
- 현재 전 세계에서 자국의 플랫폼을 보유한 국가는 손에 꼽을 만큼 소수
- 다행히 현재 국내에는 네이버(라인)과 카카오 등 글로벌 경쟁력을 갖춘 플랫폼 사업자가 존재하고 있어, 해당 기업들이 해외 사업자와 경쟁할 수 있는 규모로 성장 할 수 있는 지원 정책이 필요
- 미국, 중국 정부는 안보나 국익차원에서 자국의 기업들과 직, 간접적으로 공조하면서 자체적으로도 정보전쟁을 치밀하게 준비해 옴
  - 미국 정부는 플랫폼 기업들에 대한 자국 내 규제를 느슨하게 유지해 왔는데, 이를 통해 미국의 플랫폼 기업들은 다양한 혁신적인 아이디어를 자유롭게 시험하면서 경쟁력을 키워왔고, 글로벌 시장으로도 적극적으로 진출하고 있음
  - 중국 정부는 자국 내에서는 강한 규제로 외국 기업들의 손발을 묶는 한편 자국의 기업들의 설립과 성장 그리고 해외 진출에 적극적으로 지원해 옴

- 반면, 자국의 플랫폼 대표 선수가 없어 ‘플랫폼 공백’ 사태를 겪고 있는 EU의 경우, 미국이나 중국의 플랫폼 기업을 견제하기 위해 규제를 강화하는 것 이외에 다른 해법이 없는 상황

## (2) 자국 플랫폼의 영향력에 대한 체계적인 연구지원 필요

- 국내 플랫폼 시장은 경쟁력 있는 자국 플랫폼을 보유하고 있고, 이들 플랫폼들이 (동남아시아를 중심으로) 글로벌 진출을 시도하는 상황
- 자국 플랫폼 경쟁력을 강화하면서, 이들이 글로벌 시장에서 성과를 확산할 수 있도록 하는 ‘한국형규제 및 진흥 모델’에 대한 고민 필요
  - 국내는 EU처럼 글로벌 플랫폼에 대한 무조건적인 견제를 진행하기 어려움
- 이를 위해, 자국 플랫폼이 국가 경제 및 관련 산업에 미치는 영향력에 대한 체계적인 분석 필요하지만, 선행 연구는 거의 전무한 수준
- 국가, 기업, 개인 차원에서의 효용을 분리해서 접근-경제적(계량적, 비계량적), 사회적, 문화적 측면 등 다양한 차원에 대한 분석 필요

## (3) AI 생태계 확장을 위해서는 자국 플랫폼의 규모의 경제 달성 지원

- 플랫폼 산업의 핵심이 되는 인공지능 기술의 개발은 기술 자체뿐 아니라 데이터 및 피드백이 중요한 인공지능의 특성을 고려해야 함
- 플랫폼 기업이 성공하기 위해서는 개방형 혁신을 통한 협력 생태계를 구축하는 것이 필수
- 최근에는 인공지능 기술을 좀 더 효율적으로 개발할 수 있도록 협업 플랫폼이 적극적으로 도입되고 있고, 생성형 AI가 확산되면서 거대한 분량의 데이터를 확보하기 위한 전방위적인 노력을 진행하고 있음
  - 국내 사업자들 역시 인공지능 소스를 공개하고, 거대 생성형 모델을 공개하면서 경쟁력을 높이고 있지만, 글로벌 사업자와 비교했을 때 규모와 질적인 측면에서 앞서 있다고 말하기 힘든 상황
- 국내 기업의 오픈 생태계 확장을 지원하기 위해서는 먼저 글로벌 사업자와 경쟁할 수 있는 규모의 경제를 실현하는 것이 필요

- 플랫폼은 연결과 확장을 통해 성장
- 현재 규제 프리존(샌드박스) 등 기존 규제를 걷어내기 위한 노력이 진행되고 있지만, 미국 중국 등 플랫폼 선진국과 비교하면 아직도 국내의 플랫폼 기업 및 스타트업들은 규제 때문에 시도하지 못하는 많은 일과 영역이 존재
- 최근 공정위를 중심으로 한 지배적 플랫폼 사업자에 대한 규제 강화 시도는 규모의 경제를 통한 경쟁력 확보에 걸림돌로 작용할 수 있음

#### (4) 플랫폼 산업의 산업규제 적용의 필요성을 원점부터 재 점검

- 중국의 플랫폼 기업들이 짧은 시간에 이렇게 급성장할 수 있었던 데에는 중국 정부의 자국 산업 보호를 위한 강력한 리더십이 큰 역할을 하였음
  - 중국 정부는 자국 기업들에 한해 자신들이 원하는 서비스를 마음껏 만들어 낼 수 있는 환경을 조성해 주고 있음
- 그동안 한국은 자국 플랫폼 사업자의 확장을 골목상권 등의 논리로 풀어내고, 플랫폼 사업자를 통신과 방송과 같은 사전 규제의 대상으로 삼고 정부의 관리와 통제를 강화하려는 시도를 이어감
- 플랫폼 산업을 통신과 방송과 같이 사전적인 산업규제의 대상으로 보고 접근하는 마인드 셋 자체가 이 산업의 미래를 가로막는 가장 큰 걸림돌일 수 있음
- 플랫폼 산업의 혁신성에 대해 기존과는 다른 시각으로 이해하고자 하는, 사회적 공감대 형성을 위한 노력도 필요

## 5. 국내 실태조사 현황 및 향후 과제

- 플랫폼 경제 측정 지표를 지속적으로 관측하기 위해서는 정기적인 실태조사를 통해 기초 자료를 마련할 필요
- 재화 및 서비스를 스스로 생산하지 않고 이용자 간 거래를 매개하는 플랫폼의 특수성으로 인해 기존 산업에서 실시하던 실태조사로는 플랫폼 경제 현황 파악 어려움
- 본 장에서는 국내 플랫폼 경제 현황 파악을 위해 현재까지 이루어진 실태조사를 정리하고, 향후 필요한 연구를 제시
- 구체적으로, 본 장의 구성은 다음과 같음
  - 벤치마크로서 해외 연구를 통해 플랫폼 경제 측정 지표를 관측하기 위해 필요한 기초자료 정리
  - 국내 플랫폼 경제 실태조사 현황 정리
  - 앞 두 절의 내용을 바탕으로 지표별 현존 실태조사를 통해 파악 가능한 정도를 평가하고 향후 필요한 조사 제시

### 가. 플랫폼 경제 측정 지표 및 방법론

- 본 절에서는 해외 연구로부터 제시된 플랫폼 경제 파악을 위해 필요한 자료 정리
- EU Platform Observatory, 「Measurement and Economic Indicators」 (2023)
  - 여기서는 앞서 언급된 EU Platform Observatory에서 제시한 9개의 플랫폼 경제 지표 측정 방법 및 가용한/필요한 데이터를 요약
  - 중개한 거래 규모
    - 플랫폼이 경제에 미치는 영향력의 중요한 지표
    - 측정단위는 거래액, 웹 트래픽 등
    - 플랫폼 기업은 자사가 중개한 거래에 대해 세부적 정보를 지니고 있으나, 이러한 데이터는 공개가 의무화되지 않아 구하기 어려움
    - 웹 트래픽 데이터는 SimilarWeb, Mavens, Alexa 등에서 제공

- 플랫폼 규모
  - 기업 규모의 전통적 지표(수입, 시장점유율 등)는 플랫폼 기업의 규모 지표로서 적합성이 떨어짐
  - 플랫폼 부문 및 규제 목적에 따라 플랫폼 규모 측정에 적합한 지표가 다름
  - 부문 및 규제 목적 각각에 적합하게 플랫폼 규모를 이론적으로 개념화하기 위한 노력 필요
- 데이터에 대한 데이터
  - 데이터의 종류 및 가치가 분야에 따라 고도로 이질적이기 때문에 획일화된 측정 방법이 부재
  - 향후 데이터 거래시장이 활성화되면 시장가치를 통해 데이터를 측정 가능할 수 있음
- 이용사업자의 플랫폼 의존도
  - 이용사업자 매출 중 각 플랫폼을 통한 매출이 차지하는 비중을 통해 플랫폼 의존도 측정 가능
  - 현재 특정 플랫폼의 비중이 높더라도 필요할 경우 다른 플랫폼으로 전환하는 비용이 낮을 수 있다는 점에서 의존도에 대한 완벽한 지표는 아님
  - 대안적으로 이용사업자의 멀티호밍 정도, 즉 사용하는 서로 다른 판매경로의 수를 통해 플랫폼 의존도 측정 가능
  - 하지만 각 경로가 서로 다른 시장에 도달할 수 있다는 점, 특정 경로가 다른 경로보다 현저하게 중요할 수 있다는 점에서 완벽한 지표는 아님
- 소비자 관심 점유율
  - 플랫폼의 소비자 관심 점유율을 통해 플랫폼의 경제적 영향력 측정
  - 소비자 관심을 측정하기 위한 방법론은 아직 잘 정립되지 않았고, 다양한 시도가 이루어지고 있음
  - 소비자 설문조사를 통해 특정 범주의 상품 구매를 위해 떠오르는 대표적인 플랫폼을 질문함으로써 측정
  - 웹 분석 도구를 이용해 이탈률(bounce rate), 페이지당 소비시간, 클릭률(click-through rate) 등을 수집함으로써 측정

- 플랫폼의 웹 페이지로 향하는 인터넷 트래픽 데이터를 측정
- 경쟁전략으로서의 인수
  - 현행 인수합병 규제가 거대 플랫폼 기업에 의한 반경쟁 행위를 제대로 감지하는 지에 대한 우려 제기
  - EU Merger Regulation은 매각기업의 매출액이 일정 기준 이상인 경우를 대상으로 하는데, 스타트업은 매출보다 사용자층 확대 및 데이터 수집에 집중하므로 규제를 빠져나가는 경우 발생
  - 이러한 규제 공백을 메우기 위해 독일과 오스트리아는 매출액 기준에 더해 일정 규모 이상의 거래규모를 지닌 인수합병을 규제 대상으로 지정
  - 뉴스, 신문기사, 마켓 인텔리전스 보고서 등으로부터 많은 자료가 존재하나, 장기간 동안 다양한 산업에 걸쳐 일관적으로 축적된 데이터는 부재
  - 인수합병 현황 파악을 위해 주요 플랫폼이 시행한 모든 인수합병 보고를 의무화 하는 법령 도입 권고
- 플랫폼 변동성
  - 변동성이란 기술, 정책, 운영에서 발생하는 급격하고 예측하기 어려운 변화를 의미
  - 플랫폼 공지, 약관 등을 스크레이핑(scraping)함으로써 데이터 수집은 가능하지만 텍스트 프로세싱 등의 방법으로는 실제 변화를 파악하기 어려움
  - 플랫폼 이용자 대상의 설문조사 등 고비용의 방식을 통해서만 유의미한 측정이 가능하다는 한계점
- 플랫폼 투명성
  - 투명성이란 이용자가 플랫폼의 세부적 규칙, 규제, 메커니즘을 이해할 수 있는 정도를 의미
  - 투명성의 모든 차원을 정량화하기는 어렵지만, 명확하게 정의된 일부 차원들(예컨대 검색/추천/필터 알고리즘, 데이터 수집 및 이용 방식 등)에 대해 채점표(scorecard) 형식의 지표를 구축한 사례 존재
- 기타
  - 상술한 주제 이외로 특별히 관리가 필요한 문제가 발생할 수 있음

- Platform Observatory는 부당 사례를 신고하는 온라인 포털을 운영함으로써 이러한 문제를 감지하고자 시도할 예정
  - P2B Regulation Article 11이 온라인 중개 서비스의 이용자 불만 접수 시스템 운영정보 공개를 의무화함에 따라 플랫폼 이용자 불만에 대한 정보가 축적될 것으로 예상
- OECD·International Labour Organization·European Union, 「Handbook on Measuring Digital Platform Employment and Work」 (2023)
- 플랫폼 노동 현황 파악을 위해 가용한 다양한 기초 데이터 및 설문조사 설계 방법론 기술
  - 노동인구조사
    - 미국을 포함한 여러 국가에서는 정기적으로 실시하는 노동인구조사에 플랫폼 노동에 관한 질문을 추가
    - 플랫폼 노동 관련 응답이 여러 인구통계학적 특성과 함께 기록되기 때문에 이러한 특성이 플랫폼 노동 참여에 미치는 영향 분석이 가능하다는 장점
    - 하지만 전체 노동자 중 플랫폼 노동자의 비중이 낮기 때문에 표본 크기가 크더라도 포함된 플랫폼 노동자 수가 적음
    - 따라서 플랫폼 노동의 세부적 특성(수입, 노동시간 등) 추정에 활용하기에는 표본오차에 취약하다는 한계점
  - 정보통신기술 이용자 조사
    - 정보통신기술 이용자조사는 일반적으로 정보통신기술 이용 여부, 목적, 빈도, 능숙도 등을 조사
    - 정보통신기술 이용자조사에 적절한 질문을 추가함으로써 플랫폼 노동자 및 플랫폼 클라이언트(음식배달, 택시 등의 경우)에 대한 정보 수집 가능
    - 그러나 전체 인구 대비 플랫폼 노동자의 비중이 낮기 때문에, 노동인구조사와 같이 표본에 포함된 플랫폼 노동자 수가 적다는 한계점
  - 기업 대상 설문조사
    - 플랫폼 기업 또는 플랫폼 클라이언트 기업을 대상으로 실태조사를 실시함으로써 플랫폼 노동 수요에 대한 정보 수집 가능

- 플랫폼 기업을 대상으로 설문조사를 설계할 경우 표본 구축을 위해 플랫폼 등록부가 필요하나, 많은 국가에서 이러한 등록부 부재
- 한편 플랫폼 클라이언트 기업을 대상으로 한 실태조사는 상대적으로 수월하게 설계 가능
- 세무 등록부
  - 개인의 플랫폼 노동 여부 또는 기업의 플랫폼 여부를 식별 가능할 경우, 개인 및 기업 세무 등록부가 플랫폼 노동에 대해 유용한 정보 제공
  - 특히 플랫폼 노동 소득에 대해 정확도가 높다는 장점
  - 정규 근로계약만이 기록되며 세금 면제자가 제외된다는 한계점
- 플랫폼 노동자 대상 설문조사
  - 플랫폼을 통해 일자리를 구한 경험이 있는 개인을 대상으로 설문조사
  - 수입, 근로조건 등 플랫폼 노동의 질적 측면에 대한 정보 수집에 효과적
  - 모집단을 확정하고 대표성을 지닌 표본을 설계하기 위해 양질의 선형적(a priori) 정보가 요구됨

## 나. 국내 플랫폼 경제 실태조사

- 본 절에서는 다음과 같이 조사대상에 따라 국내 실태조사를 분류해 정리
  - 플랫폼 대상 실태조사
  - 플랫폼 이용사업자 대상 실태조사
  - 플랫폼 노동자 대상 실태조사
  - 플랫폼 이용 소비자 대상 실태조사
  - 둘 이상의 이용자 그룹 대상 실태조사

### (1) 플랫폼 대상 실태조사

#### 필요성 및 목적

- 플랫폼 기업을 직접 대상으로 한 실태조사를 통해 매출액, 영업이익, 매출 구성 요소, 이용자 수, R&D 투자 등의 기초 자료 수집

□ 과학기술정보통신부, 「부가통신실태조사」 (2021)

- 부가통신사업 신고/등록/신고간주 사업자 16,708개사 대상 설문조사
- 주요 결과
  - 플랫폼 사업자의 정의: 중개 플랫폼 또는 플랫폼 인프라 서비스를 제공한다고 응답한 기업
  - 플랫폼 기업 총 매출액은 378조 원
  - 플랫폼 매출 구성은 중개수수료 50.7%, 직접 판매 25.3%, 이용자 수수료 15.0%, 광고 수수료 5.9%, 기타 3.1%
  - 평균적으로 가입한 소비자 수는 대기업 1,258만 명, 중견기업 288만 명, 중소기업 98만 명
  - 평균적으로 가입한 공급자 수는 대기업 14만, 중견기업 9.9만, 중소기업 4.9만
  - 3개월 동안의 월 평균 이용자 수는 대기업 501만 명, 중견기업 145만 명, 중소기업 58.9만 명
  - 플랫폼 기업 총 종사자는 15.4만 명, 기업 평균 143명
  - R&D 개발 인력을 보유한 기업의 비중은 46.4%
- 플랫폼 기업뿐 아니라 모든 종류의 부가통신사업자를 대상으로 설계한 조사이기 때문에 플랫폼 특성에 고유한 자료 수집에는 한계점 존재

(2) 플랫폼 이용사업자 대상 실태조사

□ 필요성 및 목적

- 개별 사업자는 플랫폼을 통해 넓은 범위의 소비자에게 수월하게 접근 가능
- 한편 플랫폼 거래 비중이 높아짐에 따라 이용사업자의 플랫폼 의존도가 높아지고, 플랫폼이 이용사업자에게 부당행위를 저지를 수 있음
  - 여러 플랫폼 부문에서 과도한 수수료, 무리한 가격 인하 요구, 입점업체 간 차별 등의 문제가 제기됨
- 플랫폼이 이용사업자에게 제공하는 혜택을 보전하면서 과도한 영향력을 지니는 것을 방지하는 정책을 수립하기 위해 이용사업자 실태 파악 필요

## □ 중소벤처기업부, 「온라인 플랫폼 이용사업자 실태조사」 (2020, 2021)

- 온라인 플랫폼 이용사업자 1,000표본을 대상으로 전화를 통해 설문조사
- 주요 결과 (2021년 보고서 기준)
  - 각 사업자가 이용하는 플랫폼의 수는 5개 이상이 36.7%, 4개가 11.2%, 3개가 25.6%, 2개가 18.8%, 1개가 7.7%
  - 창업 후 플랫폼 이용까지 걸린 기간의 경우, 창업과 동시에 이용한 기업이 70.1%, 1년 미만이 2.5%, 1~5년 미만이 11.4%, 5~10년 미만이 5.4%, 10년 이상이 4.8%
  - 플랫폼 이용 이유는 ‘플랫폼을 이용하지 않으면 영업을 지속하기 어려움’이 59.2%, ‘온라인 시장으로 사업 범위 확대’가 54.4%
  - 연 매출액 중 온라인 플랫폼 매출액 비중은 100%가 23.5%, 75~100% 미만이 25.9%, 50~75% 미만이 24.7%, 25~50% 미만이 13.6%, 10~25% 미만이 7.5%, 10% 미만이 4.7%
  - 창업 이후 플랫폼 이용을 시작한 사업자의 71.2%는 플랫폼 이용 후 매출 증가, 25.1%는 매출 변화 없음, 3.7%는 매출 감소
  - 매출이 증가한 사업자의 경우, 증가 비율은 50% 이상이 27.7%, 40~50% 미만이 4.7%, 30~40% 미만이 13.1%, 20~30% 미만이 16.4%, 10~20% 미만이 24.4%, 10% 미만이 13.6%
  - 창업 이후 플랫폼 이용을 시작한 사업자의 58.9%가 플랫폼 이용 후 영업이익 증가, 30.4%가 영업이익 변화 없음, 10.7%는 영업이익 감소
  - 영업이익이 증가한 사업자의 경우, 증가 비율은 50% 이상이 24.4%, 40~50% 미만이 2.3%, 30~40% 미만이 11.9%, 20~30% 미만이 16.5%, 10~20% 미만이 25.0%, 10% 미만이 19.9%
  - 플랫폼 매출액 대비 중개수수료 비중은 20% 이상이 6.5%, 15~20% 미만이 15.8%, 10~15% 미만이 46.6%, 5~10% 미만이 24.7%, 5% 미만이 6.4%
  - 플랫폼 이용료가 적정하다는 응답은 9.2%, 부담된다는 응답은 71.3%
  - 플랫폼 매출액 대비 플랫폼 광고비 비중은 20% 이상이 13.0%, 15~20% 미만이 7.4%, 10~15% 미만이 10.8%, 5~10% 미만이 24.2%, 5% 미만이 20.5%, 별도 광고비 없음이 24.1%

- 광고비를 부담한 사업자 중, 광고비가 적정하다는 응답은 10.2%, 부담된다는 응답은 69.9%
- 플랫폼에 대한 만족도의 경우 만족이 31.6%, 보통이 49.3%, 불만족이 19.1%
- 플랫폼으로 인한 혜택은 매출증대 14.9%, 광고홍보 효과 11.0%, 예약/주문관리 편리성 8.5% (복수응답)
- 플랫폼에 대한 불만은 수수료 정책 67.9%, 정산 절차 19.2%, 계약 조건 13.3% (복수응답)
- 응답 업체의 53.4%가 플랫폼으로부터 부당행위를 경험
- 부당행위 유형별 경험 비율은 수수료 및 거래절차 관련 91.2%, 타 플랫폼 이용 제한 및 차별적 취급 26.4%, 부당 요구 24.0%
- 부당행위 경험 시 대처방법은 플랫폼 업체에 시정 요구 69.2%, 행정기관에 신고 8.1%, 기타 중재기관 활용 4.3%, 법원 소송 1.0%, 대응하지 않고 참는다 17.4%

### (3) 플랫폼 노동자 대상 실태조사

#### □ 필요성 및 목적

- 온라인 플랫폼을 통한 일자리 증가는 가계가 추가 소득을 벌 수 있는 유연한 일자리를 제공
- 한편 플랫폼 노동은 (i) 안정성이 낮고, (ii) 사회적으로 보호받기 어려우며, (iii) 진로 계발에 불리하고, (iv) 집단 교섭이 어려우며, (v) 알고리즘에 의한 차별에 취약하다는 특성을 지님
- 이에 더해, 플랫폼 종사자는 사업자와 근로자의 경계에 위치한다는 사실로 인해 기존 법 제도를 적용하기 어려움
  - 플랫폼 종사자를 선불리 사업자로 규정하면 이들이 근로자에게 보장되는 권리 및 보호의 사각지대에 놓일 수 있음
- 이에 따라 플랫폼 노동시장 현황을 파악하고, 이를 기반으로 플랫폼 노동자 권익 보호를 위한 정책을 수립할 필요성

## □ 한국노동연구원·고용노동부, 「플랫폼 노동 종사자 보호방안 마련을 위한 실태조사」 (2020)

- 전국 만 15세~64세 90,000명 대상 설문조사
- 주요 결과
  - 플랫폼을 통해 일자리를 구한 경험이 있는 응답자는(구인구직앱 이용자, 전자상거래/임대업 종사자 제외) 전체 취업자의 0.92%인 542명(이하 플랫폼 노동자)
  - 플랫폼 노동자 중 현재 플랫폼을 이용한 일자리에 종사하고 있는 사람은 76.6%
  - 플랫폼 노동자 중 플랫폼을 이용한 일자리가 주업인 사람은 56.8%
  - 플랫폼 노동자 중 작업을 온라인에서 수행하는 사람은 20.0%, 오프라인에서 수행하는 사람은 80.0%
  - 온라인에서 수행하는 작업의 종류는 단순문서 작업이 23.6%, 창작활동이 26.2%, IT 개발 및 기술지원이 19.7%, 전문 서비스가 15.2%
  - 오프라인에서 수행하는 작업의 종류는 운송이 67.8%, 전문서비스가 11.8%, 지역기반 서비스가 5.0%, 기타가 15.5%
  - 수행하는 일 결정 방식은 본인이 선택하는 경우가 57.5%, 플랫폼이 배정하는 경우가 23.3%, 플랫폼 노동자가 제안하고 고객이 선택하는 경우가 18.4%, 무응답이 0.9%
  - 서비스 가격 결정 주체는 플랫폼이 41.6%, 소속회사가 16.5%, 플랫폼 노동자가 14.8%, 고객이 12.4%, 플랫폼 노동자와 고객이 협의해 결정하는 경우가 14.0%, 기타가 0.8%
  - 근로시간 결정 주체는 플랫폼 노동자가 69.1%, 플랫폼/소속회사가 30.9%
  - 근무시간은 수시로 바꿀 수 있는 경우가 86.2%, 한번 정하면 일정 기간 유지해야 하는 경우가 13.8%
  - 성과/서비스 만족도에 대한 별점 평가가 이루어지는 경우는 53.5%
  - 평가가 이루어지는 경우, 평가 결과가 좋지 않을 때 발생하는 불이익은 일감 감소가 52.0%, 자격 일시 정지가 19.1%, 영향 없음이 17.9%, 자격 박탈이 11.6%, 수당 하락이 10.6%

- 지난 한 달간 플랫폼 노동 근무일 수는 평균 15.3일, 근무일 하루 평균 근무시간은 6.3시간
- 플랫폼 노동을 통한 월 평균소득은 145만 원
- 전체 소득 대비 플랫폼 노동을 통한 소득 비중은 91% 이상이 40.3%, 76~90%가 3.8%, 51~75%가 3.4%, 26~50%가 16.7%, 11~25%가 12.1%, 10% 이하가 23.6%

□ 한국고용정보원·고용노동부, 「플랫폼 노동종사자 근로실태와 정책과제」 (2021)

- 설문조사와 심층집단인터뷰 병행
- 조사 대상
  - 모빌리티 플랫폼 노동: 음식배달라이더, 킥서비스라이더, 대리운전기사
  - 여성 중심 직종: 가사도우미, 간병인
  - 웹 기반형 플랫폼노동: 번역, IT 프리랜서, 미디어창작자 등
- 주요 결과
  - 2020년 2월 이후 플랫폼 노동을 통한 월평균 수입은 음식배달 234만원, 킥서비스 195만원, 대리운전 167만원, 웹 기반형 플랫폼 노동 118만원, 가사도우미 100만원
  - 노동에 소요되는 월평균 비용은 음식배달 53.1만원, 대리운전 47.4만원, 킥서비스 44.6만원
  - 주당 평균 근로일은 모빌리티 플랫폼 노동자 6일, 가사도우미 4.2일, 웹 기반형 플랫폼 노동 3일
  - 민간 상해/재해보험 가입율은 음식배달 80.6%, 킥서비스 76.8%, 대리운전 26.5%, 가사도우미 37.4%
  - 민간 상해/재해보험 연간 보험료는 음식배달 171.7만 원, 킥서비스 165.0만 원, 대리운전 47.0만 원, 가사도우미 2.9만 원
  - 앱이 규정/규칙을 정한다는 응답의 비율은 음식배달 45.7%, 대리운전 46.3%, 킥서비스 15.2%, 가사도우미 62.6%

- 규정/규칙 내용은 고객을 대할 때의 복장/말투/태도, 최소 근무 시간, 근무일, 출퇴근 시간, 일일 최대 일감 취소 건수 등
- 규정/규칙 위반 시 불이익은 앱 이용 일시 차단, 좋은 일감 배당 제한 등
- 고객으로부터 평가를 받는 비율은 음식배달 및 대리운전 50% 이상, 퀵서비스 42.1%, 가사도우미 75%, 웹 기반형 플랫폼 노동 66%
- 산재보험 가입률은 음식배달 36.1%, 퀵서비스 17.9%, 대리운전 16.0%, 가사도우미 40%, 웹 기반형 플랫폼 노동 39%
- 고용보험 가입률은 음식배달 17.7%, 퀵서비스 8.9%, 대리운전 22.2%, 가사도우미 34%, 웹 기반형 플랫폼 노동 37%
- 직업훈련을 받은 비율은 음식배달 12.8%, 퀵서비스 5.8%, 대리운전 8.0%, 가사도우미 39.6%

#### (4) 플랫폼 이용 소비자 대상 실태조사

##### □ 필요성 및 목적

- 플랫폼은 소비자에게 낮은 가격, 제품 다양성, 구매의 편리성, 정보 습득의 편리성 등의 혜택을 제공
- 한편 거래 중개자와 상품/서비스 공급자가 분리됨에 따라 소비자가 잘못된 정보, 사기 거래, 허위 매물 등에 취약해짐
  - 예컨대 숙박 예약 플랫폼의 경우, 숙박업체가 앱에서 제공된 정보와 다른 방을 제공하는 사례 발생
- 이에 더해, 플랫폼 간 경쟁이 약한 산업의 경우 독과점적 지위를 지닌 플랫폼이 시장지배력을 남용할 수 있음
  - 예컨대 배달 플랫폼의 경우 배달비가 지나치게 높다는 지적이 제기
- 플랫폼으로 인한 소비자 혜택은 보존하면서 플랫폼의 부작용을 해소하는 정책을 설계하기 위해 플랫폼 이용 소비자 실태조사 필요

□ 한국소비자원 실태조사

- 한국소비자원에서는 다양한 플랫폼 부문 각각에 대해 소비자 실태조사 실시
  - 「신유형 1인 미디어 콘텐츠 소비 실태조사」 (2017)
  - 「디지털간행물(웹툰 등) 플랫폼 서비스 실태조사」 (2018)
  - 「모바일 인앱결제서비스 실태조사」 (2018)
  - 「배달서비스 플랫폼(앱) 소비자문제 실태조사」 (2019)
  - 「디지털 콘텐츠 플랫폼 소비자문제 실태조사」 (2020)
  - 「배달앱 이용 실태조사」 (2022)
  - 「숙박 예약 플랫폼 소비자문제 실태조사」 (2022)
  - 「택시 플랫폼 이용 실태조사」 (2022)
  - 「명품 플랫폼 이용 실태조사」 (2022)
  - 「중고거래 플랫폼 소비자문제 실태조사」 (2022)
  - 「가격비교사이트 운영 실태조사」 (2022)
- 상기한 연구들은 전반적으로 1372소비자상담센터에 접수된 소비자상담 자료 분석에 초점을 맞춤
  - 소비자 불만 및 피해 사례를 피해 유형, 피해 품목, 피해액, 처리 결과 등에 따라 분류해 분포 제시
  - 대표적인 소비자상담 사례를 요약해 수록
  - 관련된 소비자 보호 법규를 조사하고 기업의 법규 준수 여부 검토
  - 소비자 보호를 위한 정책 개선안 제시
  - 일부 연구는 필요에 따라 소비자 대상 설문조사, 이용사업자 대상 설문조사 등을 추가적으로 실시
- 플랫폼 부문별 소비자 불만 및 피해에 대해 정량적 및 정성적인 분석을 통해 구체적인 현황 제시
- 한편 소비자 불만 및 피해에 초점을 맞춘 연구 성격상 플랫폼으로 인한 소비자 혜택은 거의 다루어지지 않음

**(5) 둘 이상의 이용자 그룹 대상 실태조사**

- 일부 연구는 하나의 플랫폼 부문에 초점을 맞추어 둘 이상의 이용자 그룹을 대상으로 실태조사 실시

## □ 국토교통부, 「배달업 실태조사」 (2022)

- 음식배달업 종사자 1,200명 및 음식배달 서비스 이용 소비자 1천명 대상 설문조사
- 주요 결과: 배달원 설문조사
  - 월 평균 근로일은 25.3일, 월 평균 수입은 381만원, 월 평균 95만원을 보험료 및 렌탈료 등으로 지출
  - 하루 평균 배달건수는 주중 37.4건, 주말 42.3건
  - 하루 평균 운행거리는 주중 103km, 주말 117km
  - 근로여건 개선을 위해 필요한 조치는 배달 수수료 체계 개선 43.8%, 노동자 지위 인정 13.7%, 갑질 완화 12.9%, 위험 보상 12.5%
- 주요 결과: 소비자 설문조사
  - 배달주문 방식은 배달앱이 88%
  - 이용 빈도는 월 3~5회 42%, 1~2회 38%, 6~10회 15%
  - 소비자 지불 배달료는 평균 2천원대 57%, 3천원대 31%
  - 소비자 인식 적절 배달료는 음식값의 약 10%
  - 신속성, 편의성, 선택 다양성에 대해서는 대체로 만족하지만 높은 배달료에 대해서는 개선이 필요하다는 의견

## 다. 국내 플랫폼 실태조사 현황 평가 및 향후 과제

- 본 절에서는 앞 두 절의 내용을 바탕으로 플랫폼 경제 영역별로 현재까지 수행된 실태조사를 통해 파악 가능한 정도를 평가, 향후 필요한 조사를 제시

### □ 플랫폼 산업 지표

#### ● 플랫폼 산업 성장

- 「부가통신사업 실태조사」는 플랫폼의 매출액 및 이용자 수 데이터를 수집
- 이로부터 플랫폼 산업의 총 매출액/이용자 수 등을 추정, 시간에 따른 추이를 추적함으로써 플랫폼 산업 성장 관측 가능
- 그러나 플랫폼만을 대상으로 한 조사가 아니기 때문에 증개 거래량(건수/액수) 등 플랫폼 규모 파악에 유용한 자료가 수집되지 않았다는 한계점

#### ● 플랫폼 기업 간 경쟁

- 「부가통신사업 실태조사」로부터 수집한 플랫폼 매출액 및 이용자 수 데이터로부터 경쟁 상황을 파악하는 것은 상대적으로 한계가 뚜렷함
- 이는 (i) 경쟁 상황 파악은 산업 성장 파악에 비해 기업 누락에 취약하고 (ii) 플랫폼 특성상 매출액 기준 시장점유율이 시장지배력의 지표로서 적합성이 부족하며 (iii) 이용자 수 또한 이용빈도 및 지출액에 대한 정보 없이는 시장지배력을 대리하기 어렵기 때문
- 경쟁 현황을 정확히 파악하기 위해서는 각 플랫폼 부문마다 부문별 특성에 맞추어 플랫폼 기업 전수 대상 실태조사 설계 필요
- 플랫폼 전수 실태조사가 어려울 경우, (i) 입점업체를 대상으로 각 플랫폼을 통한 거래 건수 및 매출액을 조사하거나 (ii) 이용 소비자를 대상으로 각 플랫폼 이용 여부, 이용빈도, 각 플랫폼을 통한 지출액 등을 조사함으로써 보완 가능

### □ 플랫폼 이용사업자 지표

#### ● 플랫폼으로 인한 이용사업자 혜택

- 「온라인 플랫폼 이용사업자 실태조사」에서는 입점사업자의 플랫폼 이용 후 매출 및 영업이익 증감 여부 및 증감율 자료를 수집

- 플랫폼 이용 후 매출 및 영업이익이 상승했다는 응답이 지배적
- 그러나 매출과 영업이익은 다양한 변수에 의해 결정되므로 내생성 문제로 인해 단순히 플랫폼 이용 전후 값을 비교해서는 플랫폼 이용의 혜택을 정확히 추정할 수 없음
- 플랫폼으로 인한 이용사업자 혜택을 측정하기 위해서는 이용사업자의 매출, 영업이익 및 다양한 특성에 대한 패널 데이터를 구축, 이를 바탕으로 보다 정교한 통계적 기법을 도입할 필요성
- 표본이 대표성을 지닐 경우, 플랫폼으로 인한 이용사업자 이윤 변화 추정치로부터 플랫폼이 공급자 후생에 미친 영향 추정 가능
- 이용사업자의 플랫폼 의존도
  - 「온라인 플랫폼 이용사업자 실태조사」에서 조사한 (i) 총매출액 중 플랫폼을 통한 매출액 비중, (ii) 이용하는 플랫폼 개수 데이터를 통해 이용사업자의 플랫폼 의존도를 어느 정도 파악 가능
  - 조사에 따르면 연 매출액 중 플랫폼을 통한 매출액 비중이 절반 이상인 사업자가 75%로, 이용사업자의 플랫폼 전반에 대한 의존도는 상당히 높음
  - 이용 플랫폼의 수가 3개 이상인 사업자는 75%로, 이용사업자들이 멀티호밍을 통해 개별 플랫폼에 대한 과도한 의존을 기피함을 시사
  - 이용사업자의 개별 플랫폼에 대한 의존도를 측정하려면 이용사업자의 판매경로별 매출액 자료가 필요하나, 이를 수집한 실태조사는 부재

#### □ 플랫폼 노동자 지표

- 「플랫폼 노동 종사자 보호방안 마련을 위한 실태조사」는 업종, 소득, 주업/부업 여부, 임금근로자/개인사업자 여부, 4대 보험 가입 여부, 가격/근로시간 결정 방식, 수수료 등 다양한 질적 특성을 조사
  - 그러나 90,000명의 응답자 중 플랫폼 노동자는 542명으로 수가 많지 않음
  - 업종에 따라 세분할 경우 플랫폼 노동자 표본 크기가 50명 미만인 경우가 많아 업종별 플랫폼 노동 특성 추정치가 표본오차에 취약할 수 있음

- 「플랫폼 노동종사자 근로실태와 정책과제」는 모빌리티 플랫폼 노동, 가사도우미/간병인, 웹 기반 플랫폼 노동의 세 부문에 초점을 맞추어 사용 플랫폼 개수, 수수료, 노동일, 노동시간, 수입, 비용, 지시/규제 등 다양한 질적 특성 조사
  - 업종별로 따로 표본을 설계하여 수백 명 단위의 종사자 조사

□ 플랫폼 이용 소비자 지표

- 플랫폼으로 인한 소비자 혜택
  - 「배달업 실태조사」 등에서는 소비자 설문조사를 통해 서비스의 신속성, 편의성, 선택 다양성 등을 리커트 척도로 측정. 이러한 자료는 전반적인 소비자 만족도 파악에는 유용하나 소비자 혜택의 정량화에는 한계점
  - 경제학적 관점에서 플랫폼으로 인한 소비자 혜택에 대한 가장 중요한 정량적 지표는 소비자후생임
  - 일반적으로 소비자후생 추정은 소비자 구매내역에 대한 미시자료가 필요하나, 이러한 자료를 실태조사를 통해 수집하는 것은 일반적으로 어려움
  - 플랫폼으로 인한 소비자후생 변화를 추정하기 위해서는 기업 협조를 통해 신용카드 이용내역, 플랫폼 거래내역 데이터 등을 얻어야 할 것으로 보임
- 플랫폼에 대한 소비자 불만
  - 한국소비자원은 보도자료 및 보고서 등을 통해 1372소비자상담센터 소비자상담 자료를 플랫폼 부문별로 분석한 결과를 발표
  - 이를 통해 플랫폼 부문별 소비자 불만을 구체적으로 파악 가능

## 참고문헌

### | 제2장 참고문헌

#### [국내 문헌]

- 밀라노비치, 브랑코 (2017) 『왜 우리는 불평등해졌는가』 서정아 역, 21세기 북스.
- 한지우·오삼일 (2023) “AI와 노동시장 변화” BOK 이슈노트, 한국은행, 2023. 11. 16.
- 남지연 / Nam, Jiyeon. (2023). 예술작품 - 제품 결합 형태의 적합성이 소비자 평가에 미치는 영향 : 인공지능(AI) 생성 이미지와 작가 작품을 중심으로. 커뮤니케이션 디자인학연구 / Journal of Communication Design, 85, 472.
- 박종구 (2023. 5.). 생성형 인공지능과 광고, Ad & Media Trend Report, Vol. 2.
- 손영화 / Young-hoa Son. (2023). 생성형 Ai에 의한 창작물과 저작권. 법과 정책연구 / Journal of Law and Politics Research, 23(3), 357.
- 유성희 Y. S. (2023). ‘생성 AI(Generative-AI)’가 던지는 윤리적 쟁점 - ‘좋은 AI 사회 (Good AI Society)’를 위한 대응 전략. 생명윤리, 24(1), 1-29.
- 이성엽 / Seong-yeob Lee. (2023). 생성형 인공지능(AI) 규제원칙에 관한 연구. 고려법학 / Korea Law Review, 110, 149. <https://doi-org-ssl.oca.korea.ac.kr/10.36532/kulri.2023.110.149>
- 홍진아 / Hong, Jina, & 구자준 / Koo, Ja Joon. (2023). 이미지 생성 AI의 효율성과 창의성이 지속사용의도에 미치는 영향 - OpenAI 달리(DALL-E)2를 중심으로. 디자인융복합연구 / Design Convergence Study, 22(5), 121. <https://doi-org-ssl.oca.korea.ac.kr/10.31678/SDC102.7>
- [삼성생명 TVCF] 좋은 소식의 시작, 삼성생명. (2023). YouTube. Retrieved from [https://youtu.be/Wxrw5sy5\\_Hk](https://youtu.be/Wxrw5sy5_Hk)
- DB금융네트워크, 2020년 광고 촬영 현장 스케치. (2020). DB 블로그. <https://www.dbblog.co.kr/1729>
- KT. (2023). 초거대 AI 믿음. KT Enterprise. Retrieved from [https://enterprise.kt.com/pd/P\\_PD\\_NE\\_00\\_316.do](https://enterprise.kt.com/pd/P_PD_NE_00_316.do)
- 구아현. “투명하지 않은 생성형 ai, 사용해도 될까.” THE AI. (2023). Retrieved from <https://www.newstheai.com/news/articleView.html?idxno=4391>
- 이상덕. 샘 올트먼 “AI 위협 과장 ... 빈곤층에 더 혜택.” (2023). 매일경제. <https://www.mk.co.kr/news/it/10836932>

칸 라이언즈. (2023). AI, 크리에이티비티의 가능성을 확장시키다. . . 칸 라이언즈 2023 랩업 리포트2. Retrieved from <https://www.canneslions.co.kr/news/view?aid=2545>

클로바. (2023). 나의 생산성을 높여주는 초대규모 AI 하이퍼클로바 X. CLOVA. Retrieved from <https://clova.ai/hyperclova>

관계부처 합동(2022), 대한민국 디지털 전략.

김태원(2023), ChatGPT는 혁신의 도구가 될 수 있을까?: ChatGPT 활용 사례 및 전망, The AI Report 2023-1, 지능정보사회진흥원.

법무법인(유) 지평(2021), 플랫폼 경쟁법, 박영사.

안성원, 유재홍, 조원영, 노재원, 손효현(2023), 초거대 언어모델의 부상과 주요 이슈 -ChatGPT의 기술적 특징과 사회적·산업적 시사점, 소프트웨어정책연구소 이슈리포트.

오장민(2023), “생성형 AI의 기술 진화 및 연구 동향: 언어 모델을 중심으로”, KISDI AI Outlook Vol. 13.

유재홍 외 6(2023), 생성 AI의 부상과 산업 변화, 이슈 리포트 IS-161, 소프트웨어정책연구소.

유재홍, 안성원, 안미소, 노재원(2023), 생성 AI 산업 생태계 현황과 과제, 소프트웨어정책연구소 이슈리포트.

윤혜선(2009), 리스크 규제에 관한 공법적 연구, 서울대학교 법학박사 학위논문.

윤혜선(2018), “빅데이터 규제의 새로운 패러다임 모색을 위한 연구-개인정보 보호법제의 한계와 리스크 규제체제로의 전환 가능성에 대한 검토를 중심으로-”, 경제규제와 법 제11권 제1호.

이상윤(2020), “유럽연합 디지털 정책의 동향과 전망: “유럽의 디지털 미래”·“유럽 데이터 전략”·“인공지능 백서”의 주요 내용과 의의”, 고려법학 제97호.

이원우(2016), “혁신과 규제”, 경제규제와 법 제9권 제2호.

이원우 외(2019), 4차 산업혁명 시대의 기술혁신과 규제정책, 홍문사.

이효석(2022), “데이터 독점과 경쟁법의 적용”, 법학논고(경북대학교 법학연구원) 제77집.

채은선(2023), “영국 인공지능 규제 백서 주요 내용 및 시사점”, 디지털법제 Brief, 한국지능정보사회진흥원.

최난설현(2018), “AI 등을 활용한 사업자간 담합과 경쟁법의 대응”, 경쟁법연구 제38권.

한국지능정보사회진흥원(2022), “영국의 친혁신 AI 규제 추진 동향 : AI 규제 원칙 및 규제 역량 강화방안 보고서를 중심으로”, THE AI REPORT 2022-9.

홍대식(2019), “인공지능과 경쟁법”, 한국인공지능법학회 편, 인공지능과 법, 박영사.

공정거래위원회 2022. 12. 28. 보도참고자료 “클라우드 서비스 분야 실태조사 결과 발표”.  
 공정위 2021. 1. 27.자 제2021-027호 의결(2018서감2521)(네이버 쇼핑 부문)  
 공정위 2023. 6. 13.자 제2023-093호 의결(2021시감0760)  
 서울고법 2022. 12. 14. 선고 2021누36129 판결

## [해외 문헌]

- H. Touvron et al., Llama2: Open Foundation and Fine-Tuned Chat Models, <https://arxiv.org/abs/2307.09288>, 2023.
- A. Q. Jian et al., Mistral 7B, <https://arxiv.org/abs/2310.06825>, 2023.
- Official repository of project Yi, <https://github.com/01-ai/Yi>, 2023.
- Y. Wang et al., Self-Instruct: Aligning Language Models with Self-Generated Instructions, <https://arxiv.org/abs/2212.10560>, 2022.
- R. Taori et al., Stanford Alpaca: An Instruction-following LLaMA model, [https://github.com/tatsu-lab/stanford\\_alpaca](https://github.com/tatsu-lab/stanford_alpaca). 2022.
- S. Mukherjee et al., Orca: Progressive Learning from Complex Explanation Traces of GPT-4, <https://arxiv.org/abs/2306.02707>, 2023.
- L. Tunstall et al., Zephyr: Direct Distillation of LM Alignment, <https://arxiv.org/abs/2310.16944>, 2023.
- N. Ding et al., Enhancing Chat Language Models by Scaling High-quality Instructional Conversations, <https://arxiv.org/abs/2305.14233>, 2023.
- G. Cue et al., UltraFeedback: Boosting Language Models with High-quality Feedback, <https://arxiv.org/abs/2310.01377>, 2023.
- T. Dettmers et al., LLM.int8(): 8-bit Matrix Multiplication for Transformers at Scale, <https://arxiv.org/abs/2208.07339>, 2022.
- E. Frantar et al., GPTQ: Accurate Post-Training Quantization for Generative Pre-trained Transformers, <https://arxiv.org/abs/2210.17323>, 2023.
- J. Lin et al., AWQ: Activation-aware Weight Quantization for LLM Compression and Acceleration, <https://arxiv.org/abs/2306.00978>, 2023.
- E. J. Hu et al., LoRA: Low-Rank Adaptation of Large Language Models, <https://arxiv.org/abs/2106.09685>, 2021.
- A news article, <https://ark-invest.com/newsletters/issue-392/>, 2023 <https://carboncredits.com/how-big-is-the-co2-footprint-of-ai-models-chatgpts-emissions/>, 2023.

- A news article, <https://news.skhynix.com/sk-hynix-to-supply-industrys-first-hbm3-dram-to-nvidia/>, 2023.
- A. Wan et al., Poisoning Language Models During Instruction Tuning, <https://arxiv.org/abs/2305.00944>, 2023.
- A. Wei et al., Jailbroken: How Does LLM Safety Training Fail?, <https://arxiv.org/abs/2307.02483>, 2023.
- X. Qi et al., Visual Adversarial Examples Jailbreak Aligned Large Language Models, <https://arxiv.org/abs/2306.13213>, 2023.
- <https://huggingface.co/chat/>, 2023.
- <https://github.com/Hannibal046/Awesome-LLM>, 2023.
- <https://github.com/bentoml/OpenLLM>, 2023
- <https://www.langchain.com/>, 2023
- Acemoglu, D. and D. Autor (2011) “Skills, Tasks and Technologies: Implications for Employment and Earnings” in David Card and Orley Ashenfelter eds. *Handbook of Labor Economics*, Elsevier
- Acemoglu, D. and T. Lensman (2023) “Regulating Transformative Technologies,” MIT Working Paper.
- Acemoglu, D., & Restrepo, P. (2018) The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review*, 108(6), 1488–1542.
- Acemoglu, D., & P. Restrepo (2019) Automation and new tasks: How technology displaces and reinstates labor. *Journal of Economic Perspectives*, 33(2), 3–30.
- Acemoglu, D., & Restrepo, P. (2020a) Robots and jobs: Evidence from U.S. labor markets. *Journal of Political Economy*, 128(6), 2218–2243.
- Acemoglu, D., & Restrepo, P. (2020b). The wrong kind of AI? Artificial Intelligence and the future of labour demand. *Cambridge Journal of Regions, Economy and Society*, Cambridge Political Economy Society, 13(1), 25–35.
- Aghion, P., Jones, B. F., & Jones, C. I. (2019) Artificial intelligence and economic growth. In G. Agrawal & Goldfarb (eds.), *The economics of artificial intelligence: An agenda* (pp. 237–282). Chicago, USA: National Bureau of Economic Research, Inc.

- Agrawal, A., Gans, J., & Goldfarb, A., (2019). *Economic policy for artificial intelligence. Innovation Policy and the Economy*, 139-158.
- Agrawal, A. Gans, J., & Goldfarb, A. (2023) “Do we want less automation?” *Science*, 381(6654), pp.155-158. July14, 2023.
- Allen, W., Berg, C., Ilyushina, N. & J. Potts (2023) “Large language models reduce agency costs,” RMIT Univ. Working Paper.
- Autor, D., D. Dorn, & G. Hanson(2013) “The China Syndrome: Local Labor Market Effects of Import Competition in the United States.” *American Economic Review*, 103 (6): 2121-68.
- Autor, D., Dorn, D., Katz, L. F., Patterson, C., & Reenen, J. V. (2020) The fall of the labor share and the rise of superstar firms. *The Quarterly Journal of Economics*, 135(2), 645-709.
- Autor, D., & Salomons, A. (2017) Robocalypse Now-Does productivity growth threaten employment? MIT working paper.
- Bessen, James E., How Computer Automation Affects Occupations: Technology, Jobs, and Skills (October 3, 2016). Boston Univ. School of Law, Law and Economics Research Paper No. 15-49
- Bresnahan, T. and M. Trajtenberg (1995) “General Purpose Technologies ‘Engines of Growth’?” *Journal of Econometrics*, 65 (1995) 83-108.
- Brynjolfsson, E., Rock, D., & Syverson, C. (2019) Artificial intelligence and the modern productivity paradox: A clash of expectations and statistics. In G. Agrawal & Goldfarb (eds.), *The economics of artificial intelligence: An agenda* (pp. 23-57). Chicago, USA: National Bureau of Economic Research, Inc.
- Brynjolfsson, E. “The Turing Trap: The Promise and Peril of Human-Like Artificial Intelligence,” *Daedalus* 151, 272 (2022).
- Brynjolfsson, E. Li, D. and L. Raymond (2023) “Generative AI at Work,” NBER Working Paper 31161.
- Carugati, Christophe (2023) : Competition in generative artificial intelligence foundation models, Bruegel Working Paper, No. 14/2023, Bruegel, Brussels
- Choi, J. and D. Schwarcz, “AI Assistance in Legal Analysis: An Empirical Study,” August 2023

- Cockburn, I. M., Henderson, R., & Stern, R. (2019) The impact of artificial intelligence on innovation: An exploratory analysis. In G. Agrawal & Goldfarb (eds.), *The economics of artificial intelligence: An agenda*, pp. 115-146. Chicago, USA: National Bureau of Economic Research, Inc.
- David, P. (1990). "The Dynamo and the Computer: An Historical Perspective on the Modern Productivity Paradox," *American Economic Review*, vol. 80(2), 355-361.
- Dell'Acqua, F. et al. (2023) "Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality," Harvard Business School Working Paper 24-013.
- DellaVigna, Stefano, and Matthew Gentzkow. 2017. "Uniform Pricing in US Retail Chains." NBER Working Paper no. 23996, Cambridge, MA.
- Dube, J.-P., & Misra, S. (2017). Scalable price targeting. NBER Working Paper no. 23775.
- Eloundou, Tyna, Sam Manning, Pamela Mishkin, and Daniel Rock, "GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models," March 2023. arXiv:2303.10130.
- Felten, Edward W., Manav Raj, and Robert Seamans, "Occupational Heterogeneity in Exposure to Generative AI," April 2023.
- Frey, C. B., & Osborne, M. A. (2017) The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change*, 114, 254-280.
- Furman, J. et al., *Unlocking Digital Competition*, Report of the Digital Competition Expert Panel (Mar 2019)
- Graetz, G., & Michaels, G. (2015) Robots at work. CEP Discussion Paper No 1335.
- Goldin, C. and L. Katz(2010), *Race between Education and Technology*, Belknap Press
- Goos, M. and A. Manning (2007) "Lousy and Lovely Jobs: The Rising Polarization of Work in Britain," *Review of Economics and Statistics*, 89(1) pp.118-133
- Griliches, Z. (1957) "Hybrid Corn: An Exploration in the Economics of Technological Change," *Econometrica*, 25(4), 501-522.

- Hatzius, Jan, Joseph Briggs, Devesh Kodnani, and Giovanni Pierdomenico. 2023. "The Potentially Large Effects of Artificial Intelligence on Economic Growth." Goldman Sachs Economic Research <https://www.gspublishing.com/content/research/en/reports/2023/03/27/d64e052b-0f6e-45d7-967b-d7be35fabd16.html>
- Hui, X., O. Reshef, and L. Zhou (2023) "The Short-Term Effects of Generative Artificial Intelligence on Employment: Evidence from an Online Labor Market," SSRN Working Paper 4527336.
- Jones, C. (2023) "The A.I. Dilemma: Growth versus Existential Risk," Working Paper, July 26, 2023.
- Kaldor, N (1957), "A Model of Economic Growth," *The Economic Journal*, 67 (268), 591-624.
- Karabarbounis, L., & Neiman, B. (2014) The global decline of the labor share. *Quarterly Journal of Economics*, 129(1),61-103
- Kavuri, A. S., & McKibbin, W. J. (2017) Technology and leisure: Macroeconomic Implications. CAMA Working Paper, No. 43/2017.
- Kollnig, K. and Q. Li (2023) "Exploring Antitrust and Platform Power in Generative AI," arXiv Working Paper 11342v3, Jul. 2023.
- Korinek, A., & Stiglitz, J. E. (2019) Artificial intelligence and its implications for income distribution and unemployment. In G. Agrawal & Goldfarb (eds.), *The Economics of artificial intelligence: An agenda*, (pp. 349-390). Chicago, USA: National Bureau of Economic Research, Inc.
- Christoph Lakner, Branko Milanovic, Global Income Distribution: From the Fall of the Berlin Wall to the Great Recession, *The World Bank Economic Review*, Volume 30, Issue 2, 2016, Pages 203-232,
- Frank Levy and Richard Murnane (2013) *Dancing with Robots: Human Skills for Computerized Work, Third Way*.
- Lou, B. H. Sun, and T. Sun. (2023) "GPTs and Labor Markets in the Developing Economy: Evidence from China," SSRN Working Paper 4426461.
- Lu Y, Zhou Y. A review on the economics of artificial intelligence. *Journal of Economic Survey*. 2021;35:1045-1072.
- Milanovic, B. (2015) "Global Inequality of Opportunity: How Much of Our Income Is Determined by Where We Live?" *Review of Economics and Statistics*, 97(2), pp. 452-460.

- Milanovic, B. (2022) “Global income inequality: time to revise the elephant.” *Social Europe*, Retrieved from <https://www.socialeurope.eu/global-income-inequality-time-to-revise-the-elephant>
- Mookherjee, D. and D. Ray, “Capital and Inequality in the Long Run: Automation Without Technical Progress,” 2017. Working paper.
- Nordhaus, William D. (2021). Are we approaching an economic singularity? Information technology and the future of economic growth. *American Economic Journal: Macroeconomics*, 13(1), 299-332.
- Norton, A. (2017) Automation and inequality: The changing world of work in the global South. Issue Paper. Retrieved from <http://pubs.iied.org/pdfs/11506IIED.pdf>
- Noy, Shakked and Whitney Zhang, “Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence,” Available at SSRN 4375283, 2023.
- Schrepel, T. and A. Pentland (2023) “Competition Between AI Foundation Models: Dynamics and Policy Recommendations,” MIT Connection Science Working Paper.
- Tirole, J. (2017). *Economics for the common good*. Princeton University Press, 2017.
- Trammell, P. and A. Korinek (2023) “Economic Growth under Transformative AI,” NBER Working Paper, 31815.
- Webb, M. (2020). The impact of artificial intelligence on the labor market. Stanford University Working Paper.
- Varian, H. (2019). Artificial intelligence, economics, and industrial organization. In A. Agrawal, J. Gans, & A. Goldfarb (Eds.), *The economics of artificial intelligence: An agenda*. University of Chicago Press.
- Vipra, J. and A. Korinek (2023) “Market concentration implications of foundation models: the invisible hand of ChatGPT,” Brookings Center on Regulations and Markets WP 9, September 2023.
- World Bank (2017) *Trouble in the making? The future of manufacturing-led development*. Online available at <https://openknowledge.worldbank.org/handle/10986/27946>
- Yasar et al. (2023) “AI and the EU Digital Markets Act: Addressing the Risks of

- Bigness in Generative AI,” Proceedings of the 40 th International Conference on Machine Learning, Honolulu, Hawaii, USA. PMLR 202, 2023
- Youdkowsky, E. (2023) “Pausing AI Development Isn’t Enough. We Need to Shut it All Down,” *Time*, 2023. 3. 29.
- Zeira, J., (1998) “Workers, Machines, and Economic Growth,” *Quarterly Journal of Economics*, 113 (4), 1091-1117.
- Zingales, L. et al. (2019) *Stigler Committee on Digital Platforms: Final Report*, Chicago Booth Stigler Center for the Study of the Economy and the State.
- Ariel Ezrachi and Maurice E. Stucke(2016), *Virtual Competition*, Harvard University Press.
- Autorité de la Concurrence and Bundeskartellamt(2016), *Competition Law and Big Data*.
- Baldwin, Robert, Martin Cave and Martin Lodge(2012), *Understanding Regulation*, 2<sup>nd</sup> Ed., Oxford University Press.
- Rishi Bommasani, Percy Liang et. al.(2021), *On the Opportunities and Risks of Foundation Models*, Center for Research on Foundation Models (CRFM) Stanford Institute for Human-Centered Artificial Intelligence (HAI) Stanford University. Available at: <https://crfm.stanford.edu/report.html>
- Bostoen, Friso(2023), “Understanding the Digital Markets Act”, *Antitrust Bulletin* 2023, Vol. 68(2), 263-306, Tilburg Law School Research Paper, Available at SSRN: <https://ssrn.com/abstract=44440819>
- Competition & Markets Authority(CMA)(2023a), *AI Foundation Models: Initial Review*.
- Competition & Markets Authority(CMA)(2023b), *AI Foundation Models: Initial Report*.
- Department for Digital, Culture, Media and Sport(2022), *Establishing a pro-innovation approach to regulating AI*, Presented to Parliament by Command of Her Majesty.
- Department for Digital, Culture, Media and Sport and Department for Business, Energy and Industrial Strategy(2022), *A new pro-competition regime for digital markets - government response to consultation*, Presented to Parliament by Command of Her Majesty.

- Department for Science, Innovation and Technology(2023), A pro-innovation approach to AI regulation, Government Policy Paper, Presented to Parliament by Command of His Majesty.
- Lilian Edwards(2022), Regulating AI in Europe: four problems and four solutions, Ada Lovelace Institute.
- European Commission(2020a), Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics, COM (2020) 64 final.
- European Commission(2020b), White Paper on Artificial Intelligence-A European approach to excellence and trust, COM (2020) 65 final.
- Fanny et. al.(2023), “The EU should regulate AI on the basis of rights, not risks”, AccessNow.
- FTC(2023), “Generative AI Raises Competition Concerns”. Available at: <https://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2023/06/generative-ai-raises-competition-concerns>
- Harvard Business Review(2021), How Organizations Can Mitigate the Risks of AI.
- Thomas Höppner and Luke Streatfeild(2023), “ChatGPT, Bard & Co.: an introduction to AI for competition and regulatory lawyers”, 9 Hausfeld Competition Bulletin, Article 1.
- K. Kollnig and Q. Li(2023), “Exploring Antitrust and Platform Power in Generative AI”, the 1 st Workshop on Generative AI and Law, colocated with the International Conference on Machine Learning, Honolulu, Hawaii, USA.
- Paul Johnson et. al.(2023), “How New Regulations are Shaping Europe’s DigitalLandscape”, Baker McKinsey, GCR.
- Philippe Lorenz, Karine Perset, Jamie Berryhill(2023), Initial Policy Considerations for Generative Artificial Intelligence, OECD Artificial Intelligence Papers.
- Bernard Marr(2023), “The Difference Between Generative AI And Traditional AI: An Easy Explanation For Anyone”, Forbes.
- OECD(2017), Algorithms and Collusion: Competition Policy in the Digital Age.
- OECD(2019, 2023), Recommendation of the Council on Artificial Intelligence,

OECD/LEGAL/0449.

Andrea Renda(2019), Artificial Intelligence Ethics, governance and policy challenges, Report of a CEPS Task Force, Centre for European Policy Studies (CEPS), Brussels.

Report of the Digital Competition Expert Panel(2019), Unlocking digital competition. Available at: [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/785547/unlocking\\_digital\\_competition\\_furman\\_review\\_web.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/785547/unlocking_digital_competition_furman_review_web.pdf)

Schrepel, Thibault(2023), “Competition Is One Prompt Away”, Network Law Review.

Schrepel, Thibault and Pentland, Alex(2023), “Competition between AI Foundation Models: Dynamics and Policy Recommendations”, MIT Connection Science Working Paper 1-2003, Available at SSRN: <https://ssrn.com/abstract=4493900>

Semianalysis(2023) Google “We Have No Moat, And Neither Does OpenAI”. Available at: <https://www.semianalysis.com>

Darren Tucker et. al.(2023), “Foundational Regulation: UK Competition & Markets Authority Outlines Vision for Competition Regulation of AI Foundation Models”, Vinson&Elkins.

Pavan Vadapalli(2023), “Generative AI vs Traditional AI: Understanding the Differences and Advantages”, upGrade.

Yasar, Ayse Gizem et. al.(2023), “AI and the EU Digital Markets Act: Addressing the Risks of Bigness in Generative AI”, Proceedings of the 40th International Conference on Machine Learning, Honolulu, Hawaii, USA. arXiv:2308.02033 [cs.CY]. Available at: <https://arxiv.org/abs/2308.02033>

BGH(2020), June 25, 2020, Ref. KVR 69/19. Press Release No 080/2020 published by the Bundesgerichtshof (Federal Court of Justice) on 23/06/2020 provided by the Bundeskartellamt. Available at <https://www.bundesgerichtshof.de/SharedDocs/Pressemitteilungen/DE/2020/2020080.html>

Bundeskartellamt(2019), Bundeskartellamt prohibits Facebook from combining user data from different sources, Press Release, 2019. 2. 7. Available at <https://www.bundeskartellamt.de/SharedDocs/Meldung/EN/Pressemitteil>

ungen/2019/07\_02\_2019\_Facebook.html?nn=3591568

European Commission, Case AT.39740 Google Search (Shopping) (27 June 2017)

European Court of Justice, Meta vs Bundeskartellamt Case C-252/21, Judgment - 04/07/2023 Case T-612/17

European General Court, Google and Alphabet v Commission (Google Shopping), Judgment 10 November 2021.

White House(2023). Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence. Available at: <https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/>

Alqahtani, Hamed & Kavakli, Manolya & Kumar Ahuja, Dr. Gulshan. (2019). AN ANALYSIS OF EVALUATION METRICS OF GANS.

Cadbury - Shah Rukh Khan-My-Ad (case study) Creative Effectiveness Grand Prix Cannes Lions 2023. (2023). YouTube. Retrieved from [https://youtu.be/US\\_1qLyOmUc](https://youtu.be/US_1qLyOmUc)

ChatGPT Writes a Mint Mobile Ad. (2023). YouTube. Retrieved from [https://youtu.be/\\_eHjifELI-k](https://youtu.be/_eHjifELI-k)

Coca-Cola | Real Magic Creative Academy. (2023). YouTube. Retrieved from <https://youtu.be/JYFhQneku-g>

Coca-Cola® | Create Real Magic. (2023). Youtube. Retrieved from <https://youtu.be/FPGDMj1QUBE>

De Cremer, D., Bianzino, N.M. and Falk, B. (2023), "How generative AI could disrupt creative work". Harvard Business Review. Retrieved from <https://hbr.org/2023/04/how-generative-ai-could-disrupt-creative-work#:~:text=5In%20this%20scenario%2C%20generative%20AI,less%20new%20art%20and%20content>

Ford, J., Jain, V., Wadhvani, K., & Gupta, D.G. (2023). "AI Advertising: An Overview and Guidelines". Journal of Business Research 166, 114124

Hammett, E. (2018). How Lexus programmed a machine to write the world's first AI-scripted ad. Marketing Week. Retrieved from <https://www.market>

- ingweek.com/lexus-artificial-intelligence-advert/  
 HEINZ: A.I. Ketchup (Cannes 2023) (Social & Influencer Silver Lion). (2023).  
 YouTube. Retrieved from <https://youtu.be/cFJD9A3oqP8>  
 Hunger Station - The Subconscious Order (case study) Creative Commerce  
 Grand Prix Cannes Lions 2023. (2023). YouTube. Retrieved from <https://youtu.be/HuAQ3UyyjaU>  
 Lexus “Driven By Intuition”: the rationale behind the script. (2018). YouTube.  
 Retrieved from <https://youtu.be/4FydsARgmtg>  
 Mark Hachman. (2023). Getty sues Stability AI, and the future of AI art could be  
 at stake. PCWorld. Retrieved from <https://www.pcworld.com/article/1475597/getty-sues-stability-ai-and-the-future-of-ai-art-could-be-at-stake.html>  
 Microsoft holiday 2023 | Make your a holiday masterpiece. (2023). YouTube.  
 Retrieved from <https://youtu.be/KYqu359oY6o>  
 PseudoLab Tutorial Team (2022). GAN 소개. PseudoLab. Retrieved from  
<https://pseudo-lab.github.io/Tutorial-Book/chapters/GAN/Ch1-Introduction.html>  
 Reset with AI - The Cannes Lions 2023 Official Wrap-up Report. (2023). Info.  
 canneslions.com. Retrieved from <https://info.canneslions.com/the-cannes-lions-2023-official-wrap-up-report/reset-with-ai>  
 Spangler, T. (2018). Lexus debuts first AI-scripted commercial, directed by  
 Kevin Macdonald. Variety. Retrieved from <https://variety.com/2018/digital/news/lexus-ai-scripted-ad-ibm-watson-kevin-macdonald-1203030693/>

## | 제3장 참고문헌

### [국내 문헌]

- 권영준 (2021) “데이터 귀속·보호·거래에 관한 법리 체계와 방향,” 「비교사법」 제28권 제1호.  
 공정거래위원회 (2018) “독일 1. 연방카르텔청, 페이스북의 시지남용행위 관련 심사보고서 송부,” 해외경쟁정책동향 제 137호.  
 김성환 (2021) “데이터 기반 가격차별: 소비자 후생과 차별 비용의 문제,” 「산업조직연구」 제29집 제2호.

- 이상윤 (2020) “디지털 플랫폼 사업자의 소비자 착취 행위에 대한 경쟁법의 적용: 독일 페이스북 사건,” 「선진상사법률연구」제91호.
- 이호영 (2018) “빅데이터의 경쟁법적 함의에 관한 연구,” 「법경제학 연구」제15권 제3호.
- 이화령·김민정 (2017) “플랫폼 경제의 시장기제와 정부정책,” 연구보고서 2017-07, KDI.
- 이효석 (2022) “데이터 독점과 경쟁법의 적용,” 「법학논고」제77집.
- 최난설현 (2020) “빅데이터 관련 경쟁이슈 및 경쟁법적 과제,” 데이터 독점과 경쟁·소비자 이슈 학술 심포지엄, 발표자료.
- 홍대식 (2020) “데이터 독점 관련 경쟁법 이슈 연구,” 2020년 법·경제분석그룹(LEG) 연구보고서, 한국공정거래조정원.
- 황진우·최매화·김형준 (2023) “아마존의 자사우대행위에 대한 EU 집행위원회 동의를 결,” 「경쟁저널」제 217호.
- 과학기술정보통신부(2021). <데이터플랫폼 발전전략>. 과학기술정보통신부
- 관계부처합동(2018). <데이터산업활성화 전략>.
- 김은찬, 김은영, 이효찬, 유병준(2021). 데이터 3법 개정안의 내용과 전망: 국내 금융 및 데이터 산업계의 변화를 중심으로. <정보화정책>, 28권 3호, 49-72.
- 김정연(2022). 데이터거래소에서의 개인정보 데이터 가격 형성에 대한 연구: 데이터 공급자 중심으로. <한국산업보안연구>, 12:3, 7-25.
- 송현민(2020). 국내 데이터 산업 동향 및 유통·거래 활성화를 위한 정부 정책 방향. <전자공학회지>, 제47권 제6호, 13-17.
- 유준수(2022). 데이터 거래 및 중개에 관한 연구: 데이터 가격을 중심으로. <경영컨설팅 연구>, 22:6, 279-288.
- 연대성(2017). <디지털 트렌드: 사물인터넷과 인공지능이 바뀌놓을 일상의 패러다임 대 전환>. 서울: 책들의정원.
- 이종서, 이종현(2023). 데이터 유통, 거래 활성화를 위한 제언. <한국통신학회지(정보와 통신)>, 제40권 제3호, 10-15.
- 장영재(2017). 4차 산업혁명 시대의 데이터과학 교육 방향성 모색: 인공지능과 데이터 주도권에 관한 이해를 중심으로. <통합인문학연구>, 9권 1호, 155-180.
- 정보통신기획평가원(2020). <인공지능 기술청사진 2030>. 대전: 정보통신기획평가원.
- 제29차 4차산업혁명위원회(2022). <데이터 거래 활성화를 위한 제언>. 보고 안건 제1호.
- 한국데이터산업진흥원(2022). <2021년 데이터산업현황조사>. 한국데이터산업진흥원.

**[해외 문헌]**

- Bourreau, Marc, Alexandre de Stree, and Inge Graef (2017) "Big Data and Competition Policy: Market Power, Personalised Pricing and Advertising," Available at SSRN: <https://ssrn.com/abstract=2920301>.
- Khan, Lina M. (2017) "Amazon's Antitrust Paradox," *Yale Law Journal*, 126(3).
- Lambrecht, Anja and Catherine E. Tucker (2015) "Can Big Data Protect a Firm from Competition?" Available at SSRN: <https://ssrn.com/abstract=2705530>.
- Malobecka-Szwast, Iga (2021) "Using Competitors' Data - a Role for Competition Law? Some thoughts on the Amazon Marketplace Case," *Yearbook of Antitrust and Regulatory Studies*, 14(24).
- Michelman, Frank (1967) "Property, Utility, and Fairness: Comments on the Ethical Foundations of Just Compensation Law," *Harvard Law Review*, 80(6).
- Liang, Yu, An, Yang, Fu, & Zhao (2018). "A Survey on Big Data Market: Pricing, Trading and Protection", *IEEE Access*, 6:15132-15154.
- McAfee & Brynjolfsson (2012). "Big Data: The Management Revolution", *Harvard Business Review*, 61-68.
- Chesbrough, H., & Bogers, M. (2014). *Explicating open innovation: Clarifying an emerging paradigm for understanding innovation*. *New Frontiers in Open Innovation*. Oxford: Oxford University Press, Forthcoming, 3-28.
- Barney, J. (1991). Firm Resources and Sustained Competitive Advantage. *Journal of Management*, 17(1), 99-120.
- Broekhuizen, T.L.J., Emrich, O., Gijzenberg, M.J., Broekhuis, M., Donkers, B., & Sloot, L.M. (2021). Digital platform openness: Drivers, dimensions and outcomes. *Journal of Business Research*, 122, 902-914.
- Chesbrough, H. (2003). *Open Innovation: The New Imperative for Creating and Profiting from Technology*. Harvard Business School Press.
- Hofman, W., & Rajagopal, M. (2014). A technical framework for data sharing. *Journal of theoretical and applied electronic commerce research*, 9(3), 45-58.
- Isckia, T., De Reuver, M., & Lescop, D. (2020). Orchestrating platform ecosystems: the interplay of innovation and business development subsystems. *Journal of Innovation Economics & Management*, 197-223.

- Kaplan, A.M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons*, 53(1), 59-68.
- Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A.H. (2011). *Big data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute.
- Martin, K. (2016). Ethical issues in the big data industry. *MIS Quarterly Executive*, 15(2), 67-85.
- Moore, J.F. (1993). *Predators and Prey: A New Ecology of Competition*. Harvard Business Review.
- Shapiro, C., & Varian, H.R. (1999). *Information Rules: A Strategic Guide to the Network Economy*. Harvard Business School Press.
- Soto Setzke, D., Böhm, M., & Krcmar, H. (2019). Platform openness: A systematic literature review and avenues for future research.
- Sundararajan, A. (2016). *The Sharing Economy: The End of Employment and the Rise of Crowd-Based Capitalism*. MIT Press.
- Tiwana, A., Konsynski, B., & Bush, A. A. (2010). Research commentary—Platform evolution: Coevolution of platform architecture, governance, and environmental dynamics. *Information systems research*, 21(4), 675-687.
- von Hippel, Eric, (1989) "New Product Ideas from Lead Users," *Research Technology Management* 32, No. 3 (May/June), pp 24-37.

## | 제4장 참고문헌

### [국내 문헌]

- 강지원·임지영 (2021), "온라인 플랫폼 사업자의 자사우대에 대한 경쟁법상 허용 범위의 한계- 네이버쇼핑 사건과 EU Google Shopping 사건, 영국 Streetmap사건을 중심으로," 《경쟁법 연구, 44》, 177-223.
- 공정거래위원회 (2021), "네이버(주)[쇼핑 부문]의 시장지배적지위 남용행위 등에 대한 건," 전원회의 의결 제2021-027호.
- 공정거래위원회 (2021), "딜리버리히어로 에스이 등 4개 배달앱 사업자의 기업결합 제한 규정 위반행위에 대한 건," 전원회의 의결 제2021-032호.
- 과학기술정보통신부(2021) "2021년 부가통신사업 실태조사"

- 권오승·서정 (2022), 《독점규제법》, 법문사.
- 김민정 (2019), “플랫폼 시장의 경쟁 이슈에 대한 연구: 검색 순위를 중심으로,” 《정책연구 시리즈 2019-05》, 한국개발연구원.
- 김민정·이화령 (2018), “소비자 행동 편향에 대한 경제학적 분석과 정책적 함의: 순차공개 가격 책정(drip pricing)을 중심으로,” 《연구보고서 2018-06》, 한국개발연구원.
- 김상배 (2021), “디지털 플랫폼 경쟁의 국제정치경제: 미중기술패권 경쟁의 진화,” 《국제지역연구, 30(1)》, 41-76.
- 김성만·송태원 (2021), “네이버쇼핑 자사서비스 우대 심결사건에 대한 법적 검토,” 《경제법 연구, 20(1)》, 149-179.
- 김성환 (2022), “네이버쇼핑 사건에서의 경쟁제한성 판단에 대한 비판적 검토,” 《산업조직연구, 30(3)》, 35-65.
- 박상곤 (2019), “공유경제가 관광산업 및 지역에 미치는 영향\_공유숙박 문헌연구를 중심으로,” 《정책연구 2019-18》, 한국문화관광연구원.
- 신동한·전현배·조장희 (2020), “한국 소매업의 생산성 성장: 사업체 진입과 퇴출의 역할을 중심으로,” 《한국경제의 분석, 26(3)》, 131-183.
- 원세범·이상현 (2023), “공정거래위원회의 경제분석: 배달앱과 유리 제조업에서의 기업 결합,” 《산업조직연구 31(2)》, 75-111.
- 이공 (2021), “온디맨드 플랫폼 시장에서의 입점업체 매출분포 변화에 관한 연구: 배달앱 시장을 중심으로,” 《정책연구시리즈, 2021-03》, 한국개발연구원.
- 이봉의 (2022), 《공정거래법》, 박영사.
- 이상원·전현배 (2022), “배달앱 이용과 음식점 매출: 대구광역시 코로나19 효과를 중심으로,” 《경제학연구, 70(4)》, 69-91.
- 이화령·한요셉 (2022), “플랫폼 노동에 관한 경제학적 분석: 플랫폼 노동수요독점력과 사회적 보호 설계를 중심으로,” 《연구보고서 2022-08 KDI》.
- 중소벤처기업부(2020), “2020년 온라인 플랫폼 사용기업 실태조사 결과보고서”
- 중소벤처기업부(2021), “2021년 온라인 플랫폼 이용사업자 실태조사 결과보고서”
- 지성욱·최영준·류민호 (2015), “네이버 연구보고서 ‘자국 검색엔진이 온라인 광고 시장에 미치는 영향.’”
- 한국노동연구원·고용노동부(2020), “플랫폼 노동 종사자 보호방안 마련을 위한 실태조사”
- 한국고용정보원·고용노동부(2021), “플랫폼 노동 종사자 근로실태와 정책과제”
- 한국소비자원(2017), “신유형 1인 미디어 콘텐츠 소비 실태조사: 인터넷 개인방송을 중심으로”
- 한국소비자원(2018), “디지털간행물(웹툰 등) 플랫폼 서비스 실태조사,” 「시장조사」 18-9

- 한국소비자원(2018), “모바일 인앱결제서비스 실태조사”
- 한국소비자원(2019), “배달서비스 플랫폼(앱) 소비자문제 실태조사”
- 한국소비자원(2020), “디지털 콘텐츠 플랫폼 소비자문제 실태조사”, 「시장조사」 20-22
- 한국소비자원(2022), “배달앱 이용 실태조사”
- 한국소비자원(2022), “숙박 예약 플랫폼 소비자문제 실태조사”
- 한국소비자원(2022), “택시 플랫폼 이용 실태조사”
- 한국소비자원(2022), “명품 플랫폼 이용 실태조사”
- 한국소비자원(2022), “중고거래 플랫폼 소비자문제 실태조사”
- 한국소비자원(2022), “가격비교사이트 운영 실태조사”

### [해외 문헌]

- Alyakoob, Mohammed, and Mohammad S. Rahman, “Shared Prosperity (or Lack Thereof) in the Sharing Economy,” *Information System Research*, 33(2), 2022, 638-658.
- Argenton, C., & Prufer, J. (2012). Search engine competition with network externalities. *Journal of Competition Law and Economics*, 8(1), 73-105.
- Armstrong, M. (2006). Competition in two-sided markets. *The RAND Journal of Economics*, 37(3), 668-691.
- Bughin, J., Corb, L., James, M., Nottebohm, O., Chui, M., Barbat, B. d. M., & Said, R. (2011). The i technologies: Search. [http://www.mckinsey.com/in\\_sights/marketing\\_sales/measuring\\_the\\_value\\_of\\_search](http://www.mckinsey.com/in_sights/marketing_sales/measuring_the_value_of_search)
- Choi, J. P. (2010). Tying in two-sided markets with multi-homing. *The Journal of Industrial Economics*, 58(3), 607-626.
- Choi, S. (2010). The Implications of the Formation of Local Networks in the Global Online Knowledge Network: Case Study of South Korea. *Global Media Journal*, 9(16).
- Chun, Hyunbae and Donghan Shin, “The Entry of Online Retailers and Productivity Dynamics in the Retail Trade Sector,” *Economics Letters*, 222, January 2023, 110957.
- Chun, Hyunbae, Gong Lee, and Dongyun Yang, “How Do Online Food Delivery Services Increase Restaurant Sales?” Presented at the European Association for Research in Industrial Economics (EARIE) Conference, August 2023.

- Chun, Hyunbae, Hailey Hayeon Joo, Ian Jaeyeon Kim, Sangkon Park, and Chamna Yoon, "Competing with Airbnb: Strategies of Incumbent and Entrant Hotels," Presented at the European Association for Research in Industrial Economics (EARIE) Conference, August 2019.
- Chun, Hyunbae, Hailey Hayeon Joo, Jisoo Kang, and Yoonsoo Lee, "E-Commerce and Local Labor Markets: Is the Retail Apocalypse Near?" *Journal of Urban Economics*, 137, September 2023, 103594.
- Chun, Hyunbae, Eunjee Kwon, Dongyun Yang, "The Rise of E-commerce and Generational Consumption Inequality: Evidence from COVID-19 in South Korea," *Regional Science and Urban Economics*, 104, 2024, 103971.
- CIA (2015) World Factbook. <https://www.cia.gov/library/publications/the-world-factbook/>
- Cohen, P., Hahn, R., Hall, J., Levitt, S., & Metcalfe, R. (2016) Using Big Data to Estimate Consumer Surplus: The Case of Uber. NBER Working Paper No. 22627.
- Collison, Jack, "The Impact of Online Food Delivery Services on Restaurant Sales," Working Paper, 2020.
- Cramer, J. & Krueger, A.B. (2016) Disruptive Change in the Taxi Business: The Case of Uber. *American Economic Review: Papers & Proceedings* 106(5): 177.182.
- Davies, J., Khodjamirian, S. Giallombardo, F. and P. Aletti (2022), Survey Evidence on User Multihoming in Online Retail Businesses, *Compass Lexecon*.
- Dolfen, Paul, Liran Einav, Peter J. Klenow, Benjamin Klopach, Jonathan D. Levin, Larry Levin, and Wayne Best. "Assessing the gains from e-commerce." *American Economic Journal: Macroeconomics*, 15(1), 2023, 342-370.
- Einav, Liran, Chiara Farronato, and Jonathan Levin, "Peer-to-Peer Markets," *Annual Review of Economics*, 8, 2016, 615-635.
- Ennis, S., M. Ivaldi, M. and V. Lagos (2022), Price Parity Clauses for Hotel Room Booking: Empirical Evidence from Regulatory Change, working paper.
- European Commission (2017), Final Report of the Expert Group for the Observatory on the Online Platform Economy on the Measurement and

- Economic Indicators.
- European Commission (2017), Google Search (Shopping), case AT 39740, Commission Decision.
- European Commission (2019), Online Platforms: New Rules to Increase Transparency and Fairness.
- European Union (EU), Expert Group for the EU Observatory on the Online Platform Economy: Final reports: Measurement and Economic Indicators of the Online Platform Economy, 2021.
- Fang, X., & Rau, P. L. P. (2003). Culture differences in the design of portal sites. *Ergonomics*, 46(1-3), 242-254.
- Farronato, Chiara, and Andrey Fradkin, “The Welfare Effects of Peer Entry: The Case of Airbnb and the Accommodation Industry,” *American Economic Review*, 112(6), 2022. 1782-1817.
- Goldfarb, Avi, and Catherine Tucker, “Digital Economics,” *Journal of Economic Literature* 57(1), 2019, 3-43.
- Hermes, S., Clemons, E., Schrieck, M., Pfab, S., Mitre, M., Böhm, M., Wiesche, M., and Krcmar, H. (2020), BREEDING GROUNDS OF DIGITAL PLATFORMS: EXPLORING THE SOURCES OF AMERICAN PLATFORM DOMINATION, CHINA’S PLATFORM SELF-SUFFICIENCY, AND EUROPE’S PLATFORM GAP. *Research Papers*. 132. [https://aisel.aisnet.org/ecis2020\\_rp/132](https://aisel.aisnet.org/ecis2020_rp/132)
- IIPA. (2015). 2015 Special 301 Report on Copyright Protection and Enforcement. [http://www.iipa.com/2015\\_SPEC301\\_TOC.htm](http://www.iipa.com/2015_SPEC301_TOC.htm)
- Jeon, D.S., Jullien, B., & Klimenko, M. M. (2012). Language, Internet, and Platform Competition: the case of Search Engine: Centre for Economic Policy Research.
- Jo, Yoon, Misaki Matsumura, and David E. Weinstein, “The Impact of Retail E-Commerce on Relative Prices and Consumer Welfare,” *Review of Economics and Statistics*, 2022, 1-45.
- Mantovani, A, Piga, C.A. and C. Reggiani (2021), Online platform price parity clauses: Evidence from the EU Booking.com case, *European Economic Review* 131, 103625.
- OECD (2022), *Economic Surveys: Korea*.
- OECD·International Labour Organization·European Union(2023), “Handbook

- on Measuring Digital Platform Employment and Work”
- O'Brien, K. J. (2006). Europeans weigh plan on Google challenge. *International Herald Tribune*. 18 January 2006. Retrieved January 4, 2007, from <http://www.iht.com/articles/2006/01/17/business/quaero.php>
- Plantin, J. & Seta. G. (2019) WeChat as infrastructure: the techno-nationalist shaping of Chinese digital platforms, *Chinese Journal of Communication*, 12:3, 257-273.
- PwC. (2014). Global entertainment and media outlook 2014-2018. from Price waterhouseCoopers <http://www.pwc.com/gx/en/global-entertainment-media-outlook/>
- Rochet, J. C., & Tirole, J. (2003). Platform competition in two-sided markets. *Journal of the European Economic Association*, 1(4), 990-1029.
- Schmalensee, R., & Evans, D. S. (2007). Industrial Organization of Markets with Two-Sided Platforms. *Competition Policy International*, 3(1).
- Van der Wurff, R., Bakker, P., & Picard, R. G. (2008). Economic growth and advertising expenditures in different media in different countries. *Journal of Media Economics*, 21(1), 28-52.
- Zahir, S., Dobing, B., & Gordon Hunter, M. (2002). Cross-cultural dimensions of internet portals. *Internet Research*, 12(3), 210-220.
- Zhao, W., & Tse, E. (2011). Competition in the search engine market. *Journal of Business Strategies*, 28(2).
- Zervas, Georgios, Davide Proserpio, and John W. Byers, “The Rise of the Sharing Economy: Estimating the Impact of Airbnb on the Hotel Industry,” *Journal of Marketing Research*, 54(5), 2017, 687-705.



## ● 저 자 소 개 ●

### 박 유 리

- 서울대학교 기술정책과정 경제학 박사
- 현 정보통신정책연구원 선임연구위원

### 최 계 영

- Univ. of California, Davis 경제학 박사
- 현 정보통신정책연구원 선임연구위원

### 김 현 수

- 고려대학교 법학 박사
- 현 정보통신정책연구원 플랫폼정책연구실장

### 이 경 선

- Purdue University 경영학 박사
- 현 정보통신정책연구원 연구위원

### 김 경 은

- University of Southern California 경제학 박사
- 현 정보통신정책연구원 연구위원

### 백 소 성

- 한국과학기술원 경영공학 박사
- 현 정보통신정책연구원 부연구위원

### 이 은 민

- 성신여자대학교 경제학 석사
- 현 정보통신정책연구원 부연구위원

### 이 경 남

- 고려대학교 경영학 석사
- 현 정보통신정책연구원 부연구위원

### 이 경 은

- 서울대학교 경영학 석사
- 현 정보통신정책연구원 부연구위원

### 이 선 희

- 성균관대학교 신문방송학 박사
- 현 정보통신정책연구원 부연구위원

### 손 가 녕

- 성신여자대학교 경제학 석사
- 현 정보통신정책연구원 전문연구원

### 권 남 훈

- 스탠포드대학교 경제학 박사
- 현 건국대학교 경제학과 교수

### 김 성 환

- Johns Hopkins University 경제학 박사
- 현 아주대 경제학과 교수

### 류 민 호

- 한국과학기술원 기술경영학 박사
- 현 동아대학교 경영정보학과 교수

### 신 영 수

- 서울대학교 법학 박사
- 현 경북대 법학전문대학원 교수

### 오 장 민

- 서울대학교 전기컴퓨터공학부 박사
- 현 성신여자대학교 AI융합학부 교수

### 유 병 준

- Carnegie Mellon University 박사
- 현 서울대 경영전문대학원 교수

### 이 상 용

- 고려대학교 법학과 박사
- 현 건국대 법학전문대학원 교수

### 이 성 업

- 서울대학교 대학원 법학 박사
- 현 고려대 기술경영대학원 교수

### 이 승 민

- 서울대학교 법학 박사
- 현 성균관대 법학전문대학원 교수

### 전 현 배

- 뉴욕대학교 경제학 박사
- 현 서강대학교 경제학과 교수

### 최 세 정

- 미시간 주립 대학교 매스미디어 박사
- 현 고려대학교 미디어학부 교수

### 최 윤 정

- 코넬대학교 경제학 박사
- 현 연세대학교 경제학부 교수

### 홍 대 식

- 서울대학교 법학 박사
- 현 서강대학교 법학전문대학원 교수

### 황 용 석

- 성균관대학교 신문방송학 박사
- 현 건국대 미디어커뮤니케이션학부 교수



정책자료 23-24

## 온라인 플랫폼 정책 수립체계 운영

발행일 2023년 12월

발행인 과학기술정보통신부 장관

발행처 과학기술정보통신부

주소 세종 가림로 194 세종파이낸스센터

홈페이지 [www.msit.go.kr](http://www.msit.go.kr)

인쇄 경성문화사

