현안연구 | 24-01

디지털 전환 시대의 디지털 산업 현황 분석

정현준/박소연/김민진/박재홍

2024. 8

KISDI 정보통신정책연구원
KOREA INFORMATION SOCIETY DEVELOPMENT INSTITUTE

현안연구 | 24-01

디지털 전환 시대의 디지털 산업 현황 분석

정현준/박소연/김민진/박재홍

2024. 8

KISDI 정보통신정책연구원
KOREA INFORMATION SOCIETY DEVELOPMENT INSTITUTE

서 언

디지털전환이 가속화하고, 산업현장뿐만 아니라 일상에서도 디지털화가 확산하면서 디지털 경제에 관한 관심이 높아지고 있습니다. 특히, 최근의 생성형 인공지능 등과 같이 디지털 경제의 생산 측면에서 디지털 산업의 발전과 성장이 여타 산업의 성장뿐만 아니라 우리의 삶의 질을 좌우하는 단계에까지 이르고 있습니다.

디지털 전환 시대에 국내 산업구조 역시 빠르게 재편되고 있으며, 전 세계적으로도 이러한 변화를 경험하고 있습니다. 디지털 전환에 따른 변화를 포착하고 대응하기 위해 전 세계적으로 디지털 경제의 측정에 관한 연구가 진행 중입니다. 이번 연구 역시 객관적인 지표에 근거해 디지털 산업의 현황을 파악하고 대응하기 위한 노력의 일환으로 디지털산업 현황을 파악하고 분석하기 위한 기초 자료를 제공하고 있습니다. 디지털산업은 무엇인지, 구분의 기준은 무엇인지, 어디까지를 포함하여야 하는지에 관한 합의된 정의는 없습니다. 하지만 국제기구를 중심으로 이러한 합의 노력이 진행 중이고, 본 보고서에서는 이러한 논의를 소개하고 그 결과 각 국가에서 디지털산업의 측정을 위해 노력한 통계 및 데이터를 소개하고 있습니다.

본 연구는 정보통신정책연구원의 정현준 연구위원이 총괄책임을 맡아 수행하였으며, 박소연 전문연구원, 김민진 부연구위원, 그리고 경희대학교의 박재홍 교수님이 집필진으로 참여하였습니다.

본 연구 결과를 통해 디지털산업에 대한 이해도가 높아지고 보다 정확한 디지털산업 현황을 측정하고 파악할 수 있기를 기대합니다.

2024년 8월

정보통신정책연구원

원 장 배 경 율

목 차

서 언 ····· **1**

요약문 ····· **9**

제 1 장 서 론 ····· **13**

제 2 장 디지털산업의 정의 및 분류체계 ····· **15**

제 1 절 디지털산업의 정의 ····· 15

1. 디지털 경제의 정의 ····· 15
2. 디지털 경제 측정을 위한 프레임워크 ····· 17
3. 디지털 산업의 정의 ····· 18

제 2 절 디지털 산업의 분류체계 ····· 21

1. 디지털 산업 분류체계 ····· 21
2. 디지털산업의 포함 범위 ····· 26
3. 디지털산업 실태조사의 모집단 현황 ····· 29

제 3 장 디지털산업 현황 및 산업구조 분석 ····· **33**

제 1 절 해외 주요국의 디지털산업 통계 조사 및 분석 ····· 33

1. 개 요 ····· 33
2. 해외 주요국의 디지털산업 통계 현황 ····· 33

제 2 절 한국의 디지털산업 통계 조사 및 분석 ····· 43

1. 개요 ····· 43

2. 시장규모 ····· 43
3. 산업 구조 ····· 44
4. 디지털화 ····· 48
5. 기술 ····· 51

제 4 장 디지털플랫폼 사업자 현황 분석 ····· 54

제 1 절 디지털플랫폼 제공 사업자의 식별 ····· 54
제 2 절 디지털플랫폼 제공 사업자 분류 자동화 ····· 55
1. 디지털 중개 플랫폼 제공사업체 분류 자동화 필요성 ····· 55
2. 디지털 중개 플랫폼 제공 사업체 분류 모델 개발 ····· 57
3. 소 결 ····· 80

제 5 장 결 론 ····· 81

1. 주요 내용 ····· 81
2. 디지털산업 관련 통계 및 데이터 개선을 위한 시사점 ····· 81

참고문헌 ····· 85

Abstract ····· 87

표 목 차

〈표 2-1〉 디지털산업 분류체계 ····· 25
〈표 2-2〉 디지털산업 관련 분류체계 비교 ····· 28
〈표 3-1〉 캐나다 디지털 경제의 GDP ····· 36
〈표 3-2〉 호주 디지털 산업의 성장률 ····· 38
〈표 3-3〉 호주 디지털 산업의 부가가치 ····· 39
〈표 3-4〉 호주 디지털 활동 총부가가치의 디지털 산업 기여도 ····· 39
〈표 3-5〉 주요 디지털 상품 ····· 40
〈표 3-6〉 디지털산업 부문별 매출액 규모 ····· 44
〈표 3-7〉 디지털산업 해외 진출 유형 ····· 45
〈표 3-8〉 고객 유형별 거래 형태 ····· 45
〈표 3-9〉 디지털플랫폼 제공 방식 ····· 46
〈표 3-10〉 디지털플랫폼 비즈니스 유형(업체 수 비중) ····· 46
〈표 3-11〉 영업활동에 활용하는 디지털플랫폼 유형 ····· 47
〈표 3-12〉 디지털산업의 디지털화 수준 ····· 49
〈표 3-13〉 인공지능 기반 기술 이용 및 계획 ····· 50
〈표 3-14〉 디지털화 추진 이유(1순위 기준) ····· 50
〈표 3-15〉 디지털산업의 디지털 주문 현황(매출액 기준) ····· 51
〈표 3-16〉 디지털 신기술 개발 및 도입 현황 ····· 52
〈표 3-17〉 디지털 신기술 개발 현황 ····· 52
〈표 3-18〉 디지털 신기술 도입 현황 ····· 53
〈표 4-1〉 키워드별 검색어 사전 ····· 61
〈표 4-2〉 주요 오픈마켓 URL ····· 65

〈표 4-3〉 초기 행 설정 ····· 65
〈표 4-4〉 키워드 검색 결과 삽입 ····· 66
〈표 4-5〉 키워드 검색 결과 삽입 ····· 66

그 림 목 차

[그림 2-1] 디지털 경제: 계층적 접근법(a 'tiered' approach) ····· 16
[그림 2-2] 디지털 경제 측정을 위한 개념 프레임워크 ····· 17
[그림 2-3] OECD 정의를 반영한 디지털 산업의 범위 ····· 19
[그림 3-1] 디지털 경제의 부가가치 ····· 35
[그림 3-2] 캐나다의 산업별 GDP 비중 ····· 37
[그림 3-3] ADB의 디지털경제 측정 방법론 ····· 41
[그림 3-4] 전체 경제에서 디지털경제가 차지하는 비중 ····· 42
[그림 3-5] 디지털플랫폼 활용 시작 시점(연도별) ····· 48
[그림 4-1] 모델 개발 프로세스의 도식화 ····· 57
[그림 4-2] 선정된 키워드 분류 ····· 58
[그림 4-3] 업체별 표현 방식 차이 ····· 59
[그림 4-4] 전체 프로세스 도식화 ····· 62
[그림 4-5] Selenium(techsearch, 2023. 4.) ····· 63
[그림 4-6] 네이버 쇼핑 UI 차이 ····· 63
[그림 4-7] OCR의 예시 ····· 67
[그림 4-8] 최종 산출물 예시 ····· 69
[그림 4-9] 통신판매사업자 표기 예시: 네이버 웹툰 ····· 70
[그림 4-10] 사업자등록번호가 표기되었지만 수집되지 않은 경우별 개수 ··· 71
[그림 4-11] 주요 컬럼 Confusion Matrix ····· 73
[그림 4-12] 통신판매 중개사업자 오분류 ····· 74
[그림 4-13] 보조키워드 수집 개수 ····· 76
[그림 4-14] 통신판매 중개 키워드와 당사자 키워드 벤다이어그램 ····· 77

[그림 4-15] 통신판매 중개 키워드와 책임 키워드 벤다이어그램 ····· 78
[그림 4-16] 통신판매 중개 키워드와 입점 키워드 벤다이어그램 ····· 79

요 약 문

1. 제 목

○ 디지털 전환 시대의 디지털 산업 현황 분석

2. 연구 목적 및 필요성

○ 디지털 전환 가속으로 인한 국내 산업구조의 변화를 점검하고, 전 세계적으로 진행되는 디지털 경제 확산에 대응하여 디지털 기술의 공급과 수요의 양측면 모두를 반영한 디지털 산업 현황을 분석하고, 실증기반 디지털 산업의 정책수립을 지원할 필요가 있음

○ 본 연구는 디지털산업 현황을 조사한 선행연구의 연구결과를 바탕으로 디지털 전환 시대의 디지털산업 현황을 분석 및 검토하고 변화 방향 도출

○ 본 연구는 디지털 전환 시대의 디지털산업을 반영하는 다양한 현황 분석과 선행연구 결과를 바탕으로 디지털 전환 시대의 디지털산업 현황을 분석 및 검토하고, 디지털산업 정책을 지원할 수 있는 기반을 제공

3. 연구의 구성 및 범위

○ 디지털산업 정의 및 분류체계
 - 디지털산업 정의 및 분류체계
 - 디지털산업 조사범위 및 특성

○ 디지털산업 현황 및 산업구조 분석

- 해외 주요국의 디지털산업 통계 조사 및 분석
- 한국의 디지털산업 통계 조사 및 분석

○ 디지털화 사업자 현황 분석
- 웹사이트 등 미시데이터 분석을 통한 디지털산업 현황 파악

○ 디지털산업 데이터 개선 방향

4. 연구 내용 및 결과

○ 본 연구에서는 디지털 전환 시대의 디지털산업 현황을 파악하고자 디지털산업에 대한 통계 및 데이터 현황, 그리고 디지털산업의 현황을 보다 효율적으로 파악하기 위한 방법론을 검토함

○ (디지털산업 정의 및 분류체계) 제2장에서는 OECD를 중심으로 논의 중인 디지털산업의 정의와 분류체계를 살펴봤고, 이 과정에서 OECD가 제시한 디지털경제 측정을 위한 개념모델을 살펴봄

○ (디지털산업 현황 및 산업구조 분석) 제3장에서는 미국, 캐나다, 호주 등 해외 주요국의 디지털산업 통계 현황에 대해 살펴봄
- 그리고 한국의 디지털산업 현황을 최근 공표된 디지털산업 실태조사의 항목을 중심으로 소개함

○ (디지털플랫폼 사업자 현황 분석) 제4장에서는 디지털산업에 속한 사업자 중 디지털플랫폼 제공업체를 식별하기 위한 분류 자동화 방안에 대한 연구 결과를 제시

○ (디지털산업 관련 통계 및 데이터 개선을 위한 시사점)
- (디지털산업 통계 조사 확대) 각국은 디지털산업 관련 통계를 개발, 시험추계, 조사 중인 것으로 나타났고, 한국도 과학기술정보통신부와 정보통신정책연구원에서 디지털산업 실태조사를 국가통계로 승인받고 조사결과를 공표
- (디지털산업 분류체계 개선 방향) 현재 디지털산업 실태조사를 추진하고 있

으나, 우리나라 디지털산업의 현황을 보다 잘 반영하기 위한 개선이 필요하며, 통계청의 경제 분야 특수분류로 제정하는 등의 노력이 요구됨

- (디지털 경제 진전과 산업구조의 변화) 디지털화로 인한 거래방식의 변화는 산업구조의 변화로까지 이어지고 있으며, 기존의 산업분류체계에서 변화된 산업 현황을 파악하는 것은 한계가 있기 때문에 인터넷, 모바일 등 온라인 인터페이스인 웹사이트 크롤링을 통한 키워드 분석 등 디지털산업 내 다양한 사업체 현황을 파악하기 위한 새로운 방법 필요

5. 기대효과

○ 디지털산업에 대한 정의와 분류체계 등을 검토해 디지털산업에 대한 이해도 제고

○ 디지털산업에 대한 통계 및 데이터를 검토하여 디지털산업의 현황 및 실태를 파악하고 디지털산업에 대한 이해를 제고

○ 해외 주요국과 우리나라의 디지털산업 통계를 소개하여 실증기반 디지털산업 정책수립 기반 조성

○ 디지털산업 실태조사 등 통계조사의 기초자료로 활용

제1장 서 론

○ 코로나19 팬데믹을 거치면서 비대면 서비스가 일상화되고 디지털 활용이 산업 전반으로 확산하면서 전 세계적으로 디지털 경제가 확산되는 추이

- 디지털화는 제품의 생산에서부터 유통, 판매에 이르기까지 우리나라를 비롯하여 전 세계의 경제 산업 구조 전반에 많은 영향을 끼침
- 디지털플랫폼을 통한 디지털 주문이 보편화되면서 국민 생활 전반에 침투함에 따라, 다양한 디지털플랫폼 제공·활용 산업군 또한 성장 추세

○ 이에 따라, 기존의 ICT 산업뿐 아니라 최근 확산 추세인 디지털산업의 현황과 산업구조의 변화까지도 파악하고 분석할 필요성이 대두

- 디지털플랫폼 제공·활용 산업과 디지털 관련 산업을 모두 포괄하는 디지털산업 전반에 대한 현황 분석 및 검토, 변화 방향에 대한 모색 필요

○ 본 연구는 문헌 연구를 통해 디지털산업의 정의 및 분류체계에 대해 파악하고, 이를 근거로 실용적 측면에서 디지털산업의 정의, 분류, 모집단, 포함 범위 등에 관한 현황을 파악하고자 함

○ 본 연구에서는 디지털산업 현황을 조사한 국내외 선행연구 및 통계조사의 연구 결과를 바탕으로 디지털산업의 현황을 분석, 검토

- 미국, 캐나다, 호주 등의 디지털산업 통계를 소개하고, 관련 사항을 검토
- 한국의 디지털산업 현황을 파악할 수 있는 정보는 단편적이었으나, 2023년 디지털산업 실태조사가 국가통계로 승인을 받고, 2024년 6월 공표됨에 따라 이를 활용해 한국의 디지털산업 현황을 소개

○ 나아가 빠르게 변화하는 디지털산업의 특성을 반영하여, 산업 현황에 대한 통계조사를 위한 기초자료로서 업체 현황 정보를 수집하기 위한 방법론을 연구함

- 구체적으로 디지털화 사업자 현황 분석을 위해 웹사이트 등 미시데이터를 활용함으로써, 디지털산업 전반에 대한 현황을 파악할 수 있는 방법론을 살펴봄

○ 디지털산업 현황에 대한 선행 연구 검토 및 실질적인 통계 분석을 통해 보다 실효성 있는 디지털산업 정책 수립에 기여할 수 있도록 정책적 시사점도 함께 도출하고자 함

- 우리나라 디지털산업의 현황을 잘 파악하고 조사할 수 있는 기초 정보를 제공하고, 디지털산업의 분류체계 개선 방향에 대한 시사점을 제시

제 2 장 디지털산업의 정의 및 분류체계

제 1 절 디지털산업의 정의

1. 디지털 경제의 정의

○ 디지털 경제에 대한 다양한 정의가 제시되었으나 아직까지 합의된 정의는 없는 상황(IMF, 2018)

- OECD는 아래 세 가지 관점으로 디지털 경제를 정의하고 있음

[OECD의 디지털 경제 정의]
○ OECD(2020a)는 디지털 경제를 세 가지 관점에서 정리
- 상향식 정의: 디지털 경제를 디지털 투입물을 생산하거나 디지털 투입물에 핵심적으로 의존하는 산업
- 하향식 정의: 디지털 경제를 디지털 전환 과정 및 디지털 기술의 도입으로 강화된 경제활동
- 유연한 정의: 핵심 영역으로 ICT 산업, 좁은 영역으로 디지털플랫폼 등 디지털 경제, 넓은 영역으로 디지털화된 경제로 정의

자료: 과기정통부·KISDI(2024), 「2023 디지털산업 실태조사」 p.4 재인용

○ OECD(2020a)는 'G20 디지털 경제 측정을 위한 프레임워크 로드맵'에서 계층적 접근 방식의 정의를 제시

○ 여기서 OECD는 측정의 관점에서 유연하고, 정확한 측정을 위해 계층적 접근법(a tiered approach)을 제시하고, 경제적 활동으로 측정되지는 않으나 디지털화로 인한 사회적 활동의 변화를 반영하기 위해 디지털 사회 부문도 추가

○ 계층적 접근법(a tiered approach)은 디지털 경제를 핵심 영역과 좁은 영역, 넓은 영역으로 구분

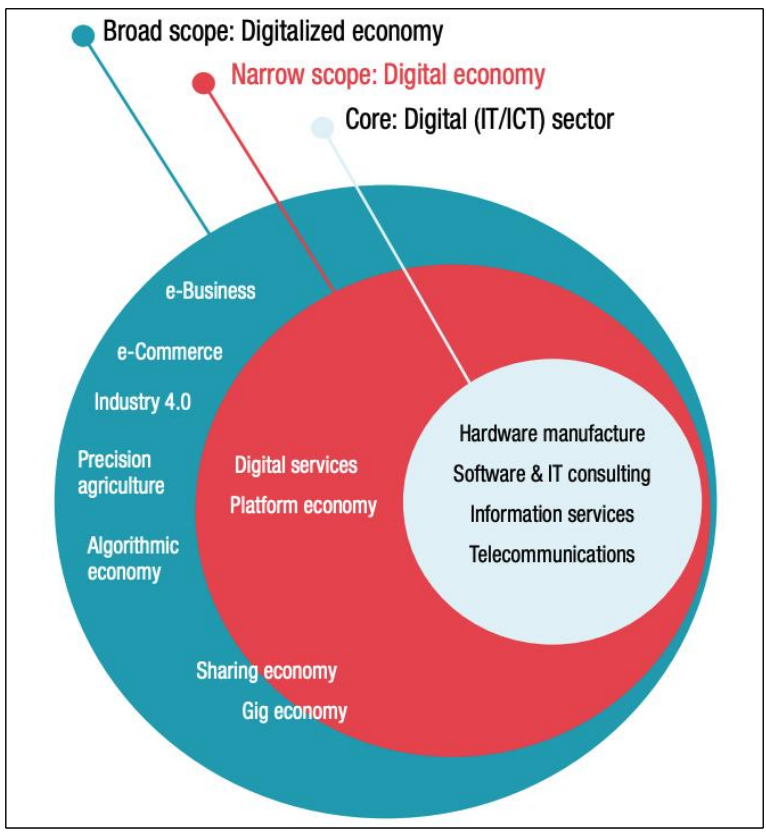
- 핵심 영역: 디지털콘텐츠와 ICT 재화 및 서비스 생산자의 경제적 활동만을

포함

- 좁은 영역: 핵심 영역에 디지털 투입 요소에 의존하는 기업에서 발생하는 경제적 활동을 포함
- 넓은 영역: 핵심 영역과 좁은 영역에 더하여, 디지털 투입 요소의 활용을 통해 현저하게 향상된 기업의 경제적 활동을 포함

※ 디지털 사회: GDP 생산 범위에 포함되지 않지만, 디지털화된 상호작용 및 활동을 포함(예: 무료 디지털 서비스)

[그림 2-1] 디지털 경제: 계층적 접근법(a 'tiered' approach)



자료: UNCTAD(2019), Bukht and Heeks(2017), OECD(2020a) p.37 재인용

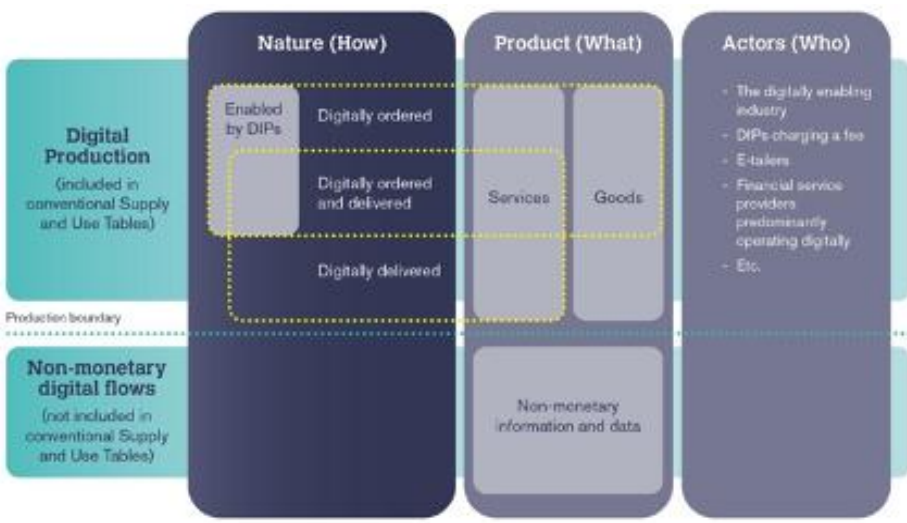
2. 디지털 경제 측정을 위한 프레임워크

○ OECD는 디지털 경제를 파악하기 위한 개념적 프레임워크를 도입하였고, 디지털 경제를 거래의 본질, 생산물의 특성, 그리고 산업 측면에서 검토

○ 디지털 경제 측정 프레임워크는 가장 중요한 요소로 거래의 본질(nature of the transaction), 즉 '어떻게 거래되는가(how)'에 주목

○ 그리고 디지털 경제의 전체 모습을 파악하기 위해 디지털 경제의 생산물, 즉 '무엇을 생산하는가(What),' 그리고 디지털 산업, 즉 '디지털 경제의 생산자로 누가 참여하는가'에 관심을 보임

[그림 2-2] 디지털 경제 측정을 위한 개념 프레임워크



자료: OECD(2023) p.20

○ (거래) 디지털 경제의 개념 요소로 거래의 본질을 살펴보고 있는데, 여기서 디지털 주문과 디지털 배송을 주목

- 디지털 주문은 크게 2가지로 구분되는데, 하나는 생산자에게 직접 주문하는 것, 그리고 다른 하나는 디지털 중개 플랫폼을 통해 주문하는 것으로 구분

- 한편, 디지털 중개 플랫폼을 통한 주문은 국내 디지털 중개 플랫폼을 통한 주문과 해외 디지털 중개 플랫폼을 통한 주문으로 구분

○ (생산물) 디지털 생산물은 크게 ICT 재화와 디지털 서비스로 총 10가지로 구분
- ICT 재화는 컴퓨터 및 주변기기, 통신장비, 소비자 가전, 그 외 ICT 부품 및 상품 등 4가지로 구분
- 디지털 서비스는 ICT 장비를 위한 제조 서비스, 비즈니스 및 생산성 SW 및 라이선스 서비스, 정보통신 기술 자문 및 서비스, 통신서비스, ICT 장비 리스 및 렌털 서비스, 그 외 ICT 서비스 등 6가지로 구분

○ (생산자) 디지털 경제에서 생산자를 구분하는데, OECD는 디지털 산업으로 아래의 7가지 산업을 제시[1)]
- 디지털기반산업
- 수수료를 부과하는 디지털 중개 플랫폼 산업
- 데이터 및 광고 기반 디지털 플랫폼
- 디지털 중개 플랫폼 의존 생산자
- 디지털 도소매업
- 디지털 금융업
- 그 외 디지털로 운영하는 업체

3. 디지털 산업의 정의

○ 본 연구에서는 디지털 산업을 디지털 경제의 생산자로서 디지털 산업을 파악하고 OECD의 계층적 접근법을 적용
- 과기정통부와 정보통신정책연구원의 디지털 산업 실태조사는 OECD의 논의를 기반으로 디지털 산업의 디지털 정의와 범위에서 설정

○ (정의) 디지털 산업은 디지털 재화 및 서비스를 생산하는 산업과 여기서 생산된 디지털 투입물의 활용을 통해 생산과 거래 활동이 디지털화된 산업으로

1) 산업 분류에 관한 자세한 내용은 2장 2절에서 자세히 소개

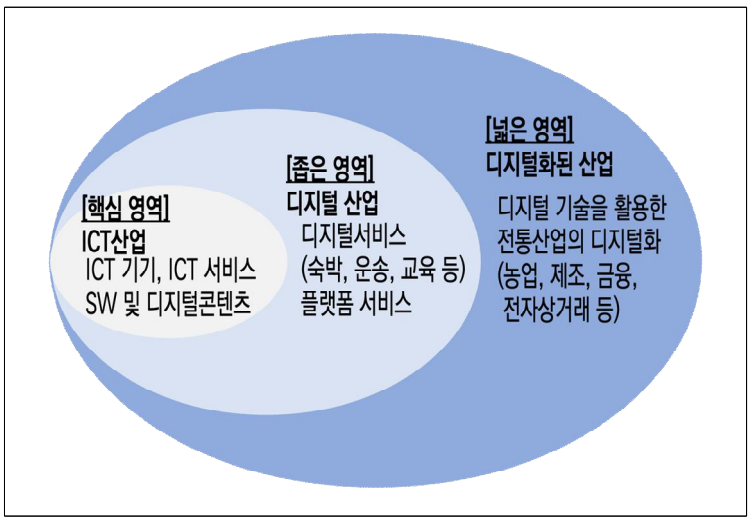
정의(「2023 디지털산업실태조사」)

- 정의에 언급되는 디지털 재화 및 서비스로는 디지털 기기 및 부품, 통신, 클라우드서비스, 디지털 플랫폼 제공 등을 포함하며,
- 디지털 투입물은 디지털 기술, 디지털 장비(하드웨어, 소프트웨어 및 통신장비) 및 생산 활동의 중간재로 사용되는 디지털 상품 및 서비스(디지털 플랫폼 서비스 등) 포함

○ (범위) 위와 같은 정의에 따르면, 디지털 산업은 OECD에서 제시한 핵심 영역, 좁은 영역, 그리고 넓은 영역을 포함

- 핵심 영역은 ICT 기기, ICT 서비스, SW 및 디지털콘텐츠와 같은 ICT 산업을 포함
- 좁은 영역은 숙박, 운송, 교육 등의 디지털 플랫폼 서비스 등을 포함
- 넓은 영역은 디지털 기술을 활용하여 디지털화된 전통산업으로 디지털화된 산업을 포함

[그림 2-3] OECD 정의를 반영한 디지털 산업의 범위



자료: OECD(2020a) 보완, 「2023 디지털산업실태조사」 p.4 수정 재인용

○ 종합하여 살펴보면, 디지털 산업의 정의는 개념적으로는 상향식 정의를 택하고 있는 것으로 보이나, 실질적으로 범위 설정 등 실질적 부분에서는 유연한 정의를 택하고 있음

※ 상향식 정의: 디지털 경제를 디지털 투입물을 생산하거나 디지털 투입물에 핵심적으로 의존하는 산업

※ 유연한 정의: 핵심 영역으로 ICT 산업, 좁은 영역으로 디지털 플랫폼 등 디지털 경제, 넓은 영역으로 디지털화된 경제로 정의

제 2 절 디지털 산업의 분류체계

1. 디지털 산업 분류체계

가. OECD의 디지털 산업 분류

□ 개요

○ OECD는 디지털 공급사용표와 관련하여 7가지 디지털 산업 분류를 제시하였는데, 아래에서는 각각의 정의를 살펴봄

- 디지털 기반 산업(Digitally enabling industry)
- 수수료를 부과하는 디지털 중개 플랫폼 산업(Digital intermediation platforms charging a fee)
- 데이터 및 광고 기반 디지털 플랫폼(Data- and advertising-driven digital platforms)
- 디지털 중개 플랫폼 의존 생산자(Producers dependent on DIPs)
- 디지털 도소매업(E-tailers)
- 디지털 금융업(Financial service providers predominantly operating digitally)
- 그 외 디지털로 운영하는 업체(Other producers only operating digitally)

□ OECD의 디지털 산업 분류[2)]

○ (디지털 기반 산업)

- (정의) 전송 및 디스플레이 등 전자적 수단으로 정보처리 및 통신 기능을 가능하게 하는 산업
- (분류) ISIC 26 컴퓨터, 전자, 광학제품 제조, 4651 컴퓨터, 컴퓨터 주변기기 및 소프트웨어 도매업, 4652 전자 및 통신장비, 부품 도매, 5820

2) OECD의 디지털 산업 분류의 정의는 OECD(2020a; 2023)의 내용이고, 한글 문구는 정현준 외(2021)를 인용하였고, 개별인용은 생략함

소프트웨어 출판업, 61 통신, 62 컴퓨터 프로그래밍, 컨설턴트 관련 활동, 631 데이터 처리, 호스팅, 관련 활동, 951 컴퓨터, 통신장비 수리

○ (수수료를 부과하는 디지털 중개 플랫폼 산업)

- (정의) 인터넷을 통해 두 명 이상의 상호 의존적인 사용자 간의 상호작용을 가능하게 하는 디지털 서비스
- 생산자 또는 소비자로부터 중개 대가로 명시적인 수수료를 받아야 함
- 명시적인 중개 수수료가 아닌, 사용자의 데이터만 지급을 받았을 경우 사용자의 정보를 광고 또는 판매하여 수익을 창출하기 때문에 데이터 및 광고 기반 디지털 플랫폼으로 분류
- 중개 플랫폼이 공급자로부터 독립적이지 않고, 하나의 생산자에 의해 플랫폼이 운영되는 경우 중개 플랫폼에 대한 정의가 충족되지 않으며, 전자상거래(e-commerce)에 해당
- (분류) ISIC 4799 점포, 노점 또는 시장에서 판매되지 않는 기타 소매업, 4791 우편주문업, 인터넷을 통한 소매업, 7990 기타 예약 서비스 및 관련 활동

○ (데이터 및 광고 기반 디지털 플랫폼 산업)

- (정의) 온라인으로 운영되며 데이터 판매하거나 광고를 통해 수익을 창출하는 플랫폼

  ※ 소셜 미디어 플랫폼, 검색엔진, 지식공유 플랫폼, 무료 애플리케이션 제공 등 포함

- 방문자를 다른 웹사이트로 안내하여 수익을 얻는 웹사이트, 플랫폼이 데이터 및 광고 기반 디지털 플랫폼으로 분류

  ※ 수수료를 받더라도 두 사용자 간의 거래를 촉진하지는 않기 때문에 수수료를 부과하는 디지털 중개 플랫폼의 정의 미충족

○ (디지털 중개 플랫폼 의존 생산자)

- (정의) 제품 및 서비스의 수요가 주로 중개 플랫폼을 통해 발생하는 기업
  ※ 호텔 체인과 같은 다국적기업에서부터 음식 배달원과 같은 작은 독립 계약자를 포함
- 플랫폼을 보조 채널로 사용하는 산업에 대해서도 디지털화의 영향을 파악할 수 있게 추정될 필요가 있음
  ※ OECD는 중개 플랫폼을 통한 수요가 50% 미만 발생하는 경우, 플랫폼을 보조 채널로 사용하는 산업으로 간주하고 있음

○ (디지털 도소매업)

- (정의) 디지털 방식으로 주문을 받아 제품 및 서비스의 구매 또는 재판매 과정에 관여하는 도소매업자
- (분류) ISIC 4791 우편주문업 또는 인터넷을 통한 소매업

○ (디지털 금융업)

- (정의) 소비자와 물리적으로 상호작용할 수 있는 방법이 제한적이거나 없는 금융 서비스 제공자를 의미하며, 금융 서비스 제공자가 주로 디지털 채널을 통해 소비자와 거래하는 주로 디지털 방식으로 운영되는 금융 서비스 제공자

○ (그 외 디지털로 운영하는 업체)

- (정의) 그 외 디지털 방식으로 운영되는 산업으로 구성되며, 판매를 위한 자체 디지털 서비스를 생산 및 디지털 방식으로 독점 운영하는 비즈니스
  ※ 디지털 콘텐츠 온라인 게임, 스트리밍 서비스 등 포함

나. 디지털 산업 실태조사의 분류체계

□ 분류체계[3)]

○ 디지털산업실태조사는 OECD의 디지털 경제 정의와 디지털 공급사용표의

3) 디지털 산업 실태조사의 분류체계는 과기정통부·KISDI(2024)의 개요 내용을 소개함

산업분류를 이용하여 디지털 산업을 정의하고, 디지털 산업 분류체계를 작성하고 있음(과기정통부·KISDI, 2024)

- 디지털 산업의 정의는 OECD(2020a)[4] 등의 디지털 경제 정의를 반영했고, 디지털 산업 분류체계는 OECD(2020b)[5]의 디지털 공급사용표 분류를 한국의 현실에 맞게 반영하고 있다는 특징이 있음

○ 디지털산업 분류체계는 OECD의 디지털 공급사용표 산업분류를 적용해 디지털화된 산업구조를 종합적으로 포괄하는 ① 디지털기반산업, ② 디지털 플랫폼 제공산업, ③ 디지털 플랫폼 활용산업, ④ 디지털 관련산업의 4개 대분류와 10개 중분류 산업으로 구성(과기정통부·KISDI, 2024)

○ 디지털산업 분류체계는 기기-서비스-SW 중심의 ICT 산업구조에서 벗어나 경제 전체에서 디지털화된 산업구조를 포괄적으로 반영한다는 특징(과기정통부·KISDI, 2024)

- 핵심 영역으로 ① 디지털기반산업, 좁은 영역으로 ② 디지털플랫폼 제공산업, ③ 디지털플랫폼 활용산업 그리고 넓은 영역으로 ④ 디지털 관련 산업을 포괄하여 OECD가 제시한 디지털 경제 산업구조를 반영(과기정통부·KISDI, 2024)
- 한편, 공급 측면에서 ① 디지털 재화 및 서비스를 생산하는 산업을 디지털기반산업(대분류1), 디지털플랫폼 제공산업(대분류2)으로 구분하고, 수요 측면에서 ② 디지털 투입물의 활용을 통해 생산과 거래 활동이 디지털화된 산업을 디지털플랫폼 활용(의존) 산업(대분류3), 디지털 관련 산업(대분류4)으로 분류할 수 있음

---

4) OECD(2020a), “A roadmap toward a common framework for measuring the digital economy”

5) OECD(2020b), “Guidelines for Supply-Use tables for the Digital Economy”, Working Party on National Accounts, SDD/CSSP/WPNA(2019)1/REV1

〈표 2-1〉 디지털산업 분류체계

분류		산업 정의 / 주요 제품 및 서비스
[디지털기반산업] 디지털 기술을 활용해 제품 및 서비스를 생산 또는 판매할 수 있게 도와주는 장비, 통신, 방송, 소프트웨어 기술 등의 디지털 기반을 생산 혹은 제공하는 산업		
	101. 디지털 기반 기기·부품	디지털화를 위한 장치, 기기에 사용되는 디지털 기기 및 부품 생산(ICT 제조) 예) 컴퓨터, 스마트폰, 반도체, 디스플레이, 트랜지스터, 회로기판
	102. 디지털 기반 서비스	디지털화를 위한 기반서비스로서 통신, 방송, 클라우드, 정보, SW 및 IT 서비스를 개발 및 제공, 그리고 관련 서비스 제공 예) 통신서비스, 초고속인터넷서비스, 지상파 방송 및 유료방송, 클라우드 서비스(서버자원, 데이터센터, 호스팅 및 도메인 서비스, 애플리케이션, 코로케이션, IaaS, SaaS, PaaS 모두 포함), 디지털 기술 개발 공급업(응용 SW, 게임 SW, 패키지 SW, IT 서비스)
[디지털플랫폼 제공산업] 디지털 기술을 통해 다수의 이용자에게 포털, 소셜미디어 등 정보서비스를 제공하거나 생산자와 소비자가 서로 원하는 제품과 서비스를 거래할 수 있도록 중개 기능을 제공하는 산업, 그리고 디지털 정보를 다수의 이용자가 거래할 수 있도록 기획, 애그리게이팅, 공급하는 디지털 서비스를 제공하는 산업		
	201. 데이터 및 광고 기반 디지털 플랫폼	포털, 메신저 및 사회관계망서비스(SNS) 등을 통해 다수의 이용자에게 무료서비스(일부 유료)를 제공하고 이를 통해 수집되는 이용자 데이터나 광고를 통해 수익 창출. 예) ① 정보서비스 기반 플랫폼: 웹서치 포털, 메신저, SNS, 커뮤니티 등 광고 기반의 플랫폼 서비스, ② 매칭정보 제공 플랫폼: 구인구직, 인력시장, 중고거래, 비교검색, 결혼 및 매칭 정보(소개팅) 등 정보검색 매칭 서비스
	202. 디지털 중개 플랫폼	디지털 기술을 통해 다수의 생산자와 소비자가 서로 원하는 제품과 서비스를 거래할 수 있도록 중개 기능을 제공하는 디지털 플랫폼을 운영하며 중개 수수료 또는 구독료를 통해 수익 창출 예) 배달, 심부름, 온라인 쇼핑몰(오픈마켓), 여행상품, 공연, 숙박, 공간예약, 부동산, 운송서비스, 모바일 앱마켓 운영 등)
	203. 디지털 영상음향 정보 플랫폼	디지털 기술을 통해 디지털 영상, 음향, 정보 등 디지털 콘텐츠를 다수의 이용자가 거래할 수 있도록 기획, 애그리게이팅, 공급하는 디지털 서비스를 제공하며, 주로 구독료 또는 라이선스를 통해 수익 창출 예) 방송/음악/웹툰/교육 등 다양한 디지털 콘텐츠를 기획/애그리게이팅/공급하는 서비스 플랫폼: 영상/음향 스트리밍 서비스 등

분류		산업 정의 / 주요 제품 및 서비스
[디지털플랫폼 활용(의존)산업] 대부분의 제품과 서비스의 수요가 디지털 플랫폼에 의존하여 발생하는 상품을 생산하는 산업		
	301. 실물 상품 기반 플랫폼 활용(의존) 산업	디지털 플랫폼에 입점하거나 활용하여 대부분의 상품 수요가 발생하는 실물 기반 제품/서비스 산업 예) 식당, 호텔, 오픈마켓 입점 업체, 우버 기사, 배달앱 라이더 등
	302. 디지털 상품 기반 플랫폼 활용 (의존) 산업	디지털 플랫폼에 입점하거나 활용하여 대부분의 상품 수요가 발생하는 디지털 제품/서비스 산업 예) 이북 제작자, 웹툰 작가, 음원 제작업체, 디지털 영상 제작업체 등
[디지털 관련 산업] 디지털 도소매, 디지털 금융 등 디지털 기술을 활용하여 전통산업의 제품 및 서비스를 디지털로 생산 또는 제공하는 산업		
	401. 디지털 도소매	자체 인터넷 쇼핑몰(e-커머스)을 운영하며 제품/서비스 유통(단독 전자상거래 전문 쇼핑몰)
	402. 디지털 금융	온라인을 통한 은행, 증권, 보험, 결제 서비스 등을 제공하는 산업 분야로, 주로 디지털 기술을 통해 비대면으로 제공하는 금융 서비스 예) 인터넷 은행, 인터넷 증권, 인터넷 보험, 전자금융, 신용검색, PG, VAN 등 결제대행 서비스
	403. 그 외 디지털 관련 산업	위의 산업에 속하지 않으나 자신이 생산 및 제공하는 제품 및 서비스를 디지털로 생산, 운영하는 산업

자료: 과기정통부·KISDI(2024), 「2023 디지털산업 실태조사」 pp. 5-6

2. 디지털산업의 포함 범위

□ 개요

○ 여기서는 OECD의 디지털 공급사용표 산업 분류체계, 디지털산업실태조사 분류체계 그리고 기존의 ICT 산업 분류체계(ICT 통합분류체계)를 살펴봄

– 디지털 산업의 범위에 대한 전체적인 시각을 제시하기 위해 여기서는 디지털산업 실태조사의 분류체계를 중심으로 나머지 2가지 분류를 비교함

□ 디지털 산업 관련 분류체계별 포함범위

○ 앞서 살펴본 바와 같이 디지털산업 실태조사의 분류체계는 OECD의 디지털 산업 분류를 준용하여 크게 4개 대분류로 재분류하여 작성

- OECD의 경우 디지털 플랫폼을 데이터 및 광고 기반 디지털 플랫폼과 수수료 기반 중개 플랫폼의 2가지로 구분
- 반면, 디지털산업 실태조사의 분류체계는 데이터 및 광고 기반 디지털 플랫폼, 디지털 중개 플랫폼, 디지털 영상음향 정보 플랫폼의 3개 중분류로 구성

○ OECD와 디지털산업 실태조사 분류체계는 산업 현실을 반영하여 서로 차이점도 존재

- 디지털 영상음향 정보 플랫폼은 방송/음악/웹툰 등 다양한 디지털 콘텐츠를 기획/애그리게이팅/공급하는 서비스 플랫폼이며 주로 영상/음향 스트리밍 서비스 등을 대상으로 하는데, 이러한 서비스는 OECD에서는 디지털 방식의 생산업자로 구분한다는 점에서 차이가 있음
- 다만, 디지털산업 실태조사 분류체계는 필요에 따라 재분류를 통해 OECD 분류와의 연계 분석이 가능하도록 구성되어 있음

○ 디지털산업 실태조사의 경우 디지털 플랫폼 활용(의존)산업을 실물상품 기반과 디지털상품 기반으로 구분하고 있음

○ ICT 통합분류체계는 ICT 산업을 대분류 수준에서 정보통신방송기기, 정보통신방송서비스, SW 및 디지털콘텐츠, 정보통신인프라서비스로 분류하고 있는데, ICT 산업의 상당 부분인 정보통신방송기기, 정보통신방송서비스, SW의 경우 디지털산업 실태조사의 디지털기반산업 분류에 포함됨

- ICT 산업 분류체계는 디지털콘텐츠 등을 포함하여 디지털산업 실태조사의 디지털기반산업과 디지털플랫폼 제공산업보다 소폭 넓은 범위를 포괄하지만, 디지털산업 실태조사와 달리 디지털 플랫폼을 명시적으로 구분하지 않음
- ICT 통합분류체계는 최근 정보통신인프라서비스가 추가되어 정보통신공사업을 포함하고 있다는 점에서도 차이가 있음

- 다만, ICT 산업 분류체계는 도소매 부문을 포함하지 않는다는 점에서 디지털산업 실태조사 분류와 차이가 있음

〈표 2-2〉 디지털산업 관련 분류체계 비교

자료: OECD(2020b), KISDI 내부자료

3. 디지털산업 실태조사의 모집단 현황[6)]

가. 디지털산업의 모집단

○ 앞에서 디지털산업 실태조사의 포괄 범위는 분류체계를 통해 살펴보았고, 여기서는 모집단을 통해 보다 구체적으로 파악하도록 함

○ 디지털산업 실태조사의 모집단은 디지털산업을 영위하는 사업체로 정의하고 있으며, 목표모집단과 조사모집단을 각각 아래와 같이 정의하고 있음(과기정통부·KISDI, 2024)

- 목표모집단: 디지털산업을 영위하는 사업체
- 조사모집단: 전국사업체조사와 행정자료에서 선별한 매년 작성일 기준 현재 전국에 분포하는 디지털산업을 영위하는 1인 이상의 사업체

나. 디지털산업 분류별 모집단

□ 디지털기반산업

○ 디지털기반산업의 모집단은 'ICT 통합모집단 중 디지털 기술을 활용해 제품 및 서비스를 생산 또는 판매할 수 있게 도와주는 장비, 통신, 방송, 소프트웨어 기술 등의 디지털 기반을 생산 혹은 제공하는 산업에 속한 사업체'로 구성(과기정통부·KISDI, 2024)

- 한국표준산업분류에서 전자부품, 컴퓨터, 영상, 음향 및 통신장비 제조업, 출판업, 영상·오디오 기록물 제작 및 배급업, 방송업, 우편 및 통신업, 컴퓨터 프로그래밍, 시스템 통합 및 관리업, 정보서비스업이 포함됨

※ 디지털 산업과 관련성이 떨어지는 581. 서적, 잡지 및 기타 인쇄물 출판업과 611. 공영 우편업을 제외

※ 63120. 포털 및 기타 인터넷 정보 매개 서비스업은 디지털플랫폼 제공 산업의 모집단에 포함되도록 조정

6) 디지털산업 실태조사 모집단은 과기정통부·KISDI(2024) 개요의 내용을 기반으로 소개함

□ 디지털플랫폼 제공산업

○ 디지털플랫폼 제공산업의 모집단은 데스크리서치 및 부가통신사업자 신고 리스트 등을 이용해 별도의 판별조사를 통해 선별하고, 그 외에 한국표준산업분류 63120을 추가하여 구성(과기정통부·KISDI, 2023)

- 한국표준산업분류에서 63120. 포털 및 기타 인터넷 정보 매개 서비스업, C. 제조업, G. 도매 및 소매업, J. 정보통신업(63120 제외), M. 전문, 과학 및 기술 서비스업(70~73), N. 사업시설 관리, 사업 지원 및 임대 서비스업(74~76), P. 교육 서비스업(85), 그리고 기타 업종이 포함됨

□ 디지털플랫폼 활용산업 및 디지털도소매

○ 디지털플랫폼 활용(의존)산업과 디지털관련산업 중 디지털도소매의 모집단은 '전국사업체조사 결과 중 도소매업(45, 46, 47) 및 숙박·음식점업(55, 56) 중 대부분의 제품과 서비스의 수요가 디지털 플랫폼에 의존하여 발생하는 상품을 생산하는 산업에 속한 사업체'로 구성(과기정통부·KISDI, 2024)

- 경제총조사와 서비스업조사의 정보를 활용하여 디지털 플랫폼 활용 집중 업종[7]을 선정
  ① 도소매업: 통신판매업(KSIC 4791)
  ② 숙박업: 일반 및 생활숙박시설 운영업(KSIC 551)
  ③ 음식점업: 기타간이음식점업 등(한식 일반 음식점업(KSIC 56111), 피자, 햄버거, 샌드위치 및 유사 음식점업(56192), 치킨 전문점(56193), 김밥 및 기타 간이 음식점업(56194))
- 통신판매업(4791) 중 자체 인터넷 쇼핑몰(e-커머스)을 운영하며 제품/서비스 유통 활용하는 것이 주요 사업인 사업체는 디지털 도소매업으로 구분하고, 이외의 사업체는 디지털 플랫폼 활용산업으로 구분

7) 디지털플랫폼 활용률이 높거나 활용 업체 수가 많은 업종을 선정

- 실사 과정에서 디지털중개플랫폼을 활용하지 않거나, 인터넷 쇼핑몰 운용을 하지 않는 업체의 경우 비율을 조사하고, 이들을 제외하는 방식으로 모집단 규모를 추정

□ 디지털 금융

○ 디지털관련산업 중 디지털금융업의 모집단은 '온라인을 통한 은행, 증권, 보험, 결제 서비스 등을 제공하는 산업(인터넷은행, 금융 지원 PG(66199) 포함)에 속한 사업체'로 구성(과기정통부·KISDI, 2024)

- 한국표준산업분류 금융업(64), 보험 및 연금업(65), 증권 및 선물 중개업(6612), 금융지원(PG)(66199) 등에 속한 사업체를 포함
- 디지털 금융의 경우 금융업의 실적 자료 관리 등의 특성을 반영하여 사업체 조사가 아닌 기업체 조사로 설계[8)]

다. 정리

○ 디지털산업 실태조사는 모집단 구축에 있어서 한국표준산업분류를 최대한 활용하고 있으며, 이 과정에서 몇 가지 조정을 통해 모집단 및 조사대상을 명확하게 하고 있음

○ 디지털산업과의 관련성이 낮은 산업을 제외

- 디지털기반산업의 경우 581. 서적, 잡지 및 기타 인쇄물 출판업과 611. 공영 우편업을 제외

○ 디지털산업 분류 간 겸업이 가능한 부문을 조정

- 디지털플랫폼 제공산업에 한국표준산업분류 63120. 포털 및 기타 인터넷 정보 매개 서비스업을 포함하고 디지털기반산업에서는 제외

○ 목표 모집단이 사실상 전 산업으로 확장될 수 있다는 점을 고려하여, 디지털플랫폼활용 집중업종을 선정하는 등 조사모집단을 명확하게 정리

- 디지털플랫폼 활용산업과 디지털도소매의 경우 도소매업, 숙박업, 음식점

8) 산업의 특성상 단독사업체 및 본사·본점만 포함

업의 범위를 디지털플랫폼 활용률과 활용업체 수를 기준으로 한국표준산업분류 5자리 기준으로 집중업종을 선정

- 집중업종으로 선정된 표본틀 내에서도 실제 디지털중개플랫폼 입점, 인터넷 쇼핑몰 운용 등의 비율을 조사하여 디지털산업 조사모집단의 크기를 추정하는 방식을 적용

제 3 장 디지털산업 현황 및 산업구조 분석

제 1 절 해외 주요국의 디지털산업 통계 조사 및 분석

1. 개 요

○ 본 절에서는 해외의 디지털산업의 통계를 살펴봄으로써 각국의 디지털산업의 현황과 산업구조 그리고 통계 작성 현황에 관한 시사점을 얻고자 함

○ 미국 BEA, 캐나다 통계청(Statistics Canada), 호주 통계청(ABS)의 사례를 살펴봄

- 미국과 호주는 미국의 BEA 방법론을 적용 중이며, 캐나다는 OECD의 디지털 공급사용표 방법론을 적용

○ 후속 절에서 한국의 디지털산업 조사 결과를 살펴보기로 함

- 한국에서는 OECD 디지털 공급사용표의 정의와 분류체계를 준용한 디지털산업 실태조사를 실시 중이며, 여기서 조사된 주요 결과를 살펴봄

2. 해외 주요국의 디지털산업 통계 현황

□ 미국 디지털 경제[9)]

○ 미국의 BEA(U.S. Bureau of Economic Analysis)는 미국 경제의 변화 양상을 지속적으로 파악하기 위해 스페셜 토픽의 형태로 새로운 통계를 연구 중

- 기술 변화가 미국의 디지털 경제와 글로벌 공급망에 미치는 영향을 파악하기 위해 다양한 노력을 기울이고 있음

9) 미국의 경우 아래 웹사이트 및 BEA(2023)을 참고하여 작성(24. 8. 19 검색)
https://www.bea.gov/data/special-topics/digital-economy

- 예를 들어 디지털 제품 및 서비스, 공유경제(sharing economy), 무료 디지털 콘텐츠, GDP 통계에 자산으로서의 데이터를 어떻게 측정, 반영할 것인지, 그리고 수출입 측정 방법론 개선 등을 연구

○ 미국은 GDP 통계에 BEA의 방법론을 통해 온라인 쇼핑, 디지털 미디어, 공유경제, 전자상거래 등의 경제적 효과를 반영 중

○ BEA는 디지털 경제를 아래 4가지 재화와 서비스를 통해 측정[10]

- 인프라(Infrastructure): ICT 재화 및 ICT 서비스 등 컴퓨터 네트워크와 디지털 경제의 기반이 되는 물리적 장비와 서비스를 포함
- 전자상거래(E-commerce): 컴퓨터 네트워크를 통한 재화와 서비스의 판매 활동으로, 소매 형태의 B2C 전자상거래와 도매 형태의 B2B 전자상거래를 포함
- 유료 디지털 서비스(Priced Digital Services): 클라우드 컴퓨팅 서비스, 통신서비스, 인터넷 및 데이터 서비스, 그리고 기타 유료 디지털 서비스와 같은 소비자에게 요금을 부과하는 컴퓨팅 및 통신 서비스
- 비국방 연방 디지털 서비스(Federal Nondefense Digital Services): 디지털 경제를 지원하는 서비스에 직접 관련된 비국방 부문의 연방 정부 기관의 예산

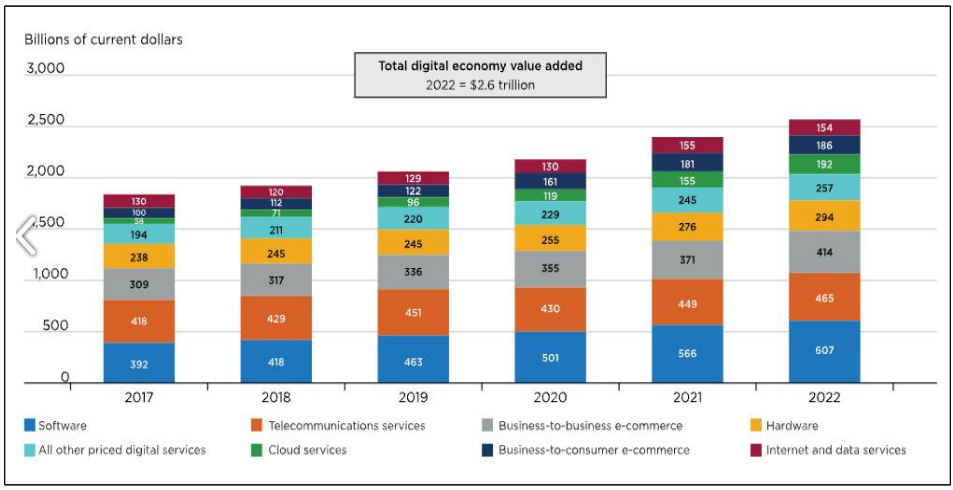
○ 2022년 미국의 디지털 경제 부가가치는 2조 6,000억 달러 수준으로, GDP의 10.0% 수준(명목 달러 기준)

- 미국의 디지털 경제는 2017~2022년 연평균 7.1% 성장
- 2022년 디지털 경제 성장률은 6.3%로 미국 GDP 성장률(1.9%)보다 3배 이상을 기록

○ 2022년 디지털 경제 부가가치에서 소프트웨어가 차지하는 비중은 24% 수준으로 가장 높으며, 통신서비스(18%), 전자상거래(16%) 등의 순으로 나타남

10) 디지털 경제 위성 계정은 예산 문제로 더 이상 생산하지 않고 있으며, 2023년 12월 6일 업데이트가 마지막 통계임(2024년 8월 현재)

[그림 3-1] 디지털 경제의 부가가치



자료: BEA(2023)

□ 캐나다[11)]

○ 캐나다 통계청은 OECD의 디지털 공급사용표 방식에 따라 디지털 경제를 측정하고 있음

○ 캐나다의 디지털 경제는 2017년 GDP의 5.2%(1,040억 달러, 이하 캐나다 달러)에서 2020년 5.9%(1,230억 달러)로 팬데믹 동안에도 안정적으로 성장

○ 캐나다의 2020년 디지털 산업에서 ICT 산업은 전체 GDP의 5.3% 수준으로 디지털 경제에서 가장 큰 비중을 차지

- ICT 산업에서는 소프트웨어와 통신업의 기여가 각각 GDP의 2.6%, 1.9%를 차지해 디지털 경제에 가장 큰 기여
- 그다음으로는 디지털 도소매업(0.27%), 디지털 금융업(0.14%), 그리고 디지털 중개 플랫폼(0.12%) 순으로 나타남

11) 캐나다의 경우 아래 문헌을 참고하여 작성
Statistics Canada(2023), "Digital supply and use tables,2017to2020", Released at 8:30 a.m. Eastern time in The Daily, Tuesday, July 25, 2023.

〈표 3-1〉 캐나다 디지털 경제의 GDP

(단위: 백만 달러, 캐나다 달러)

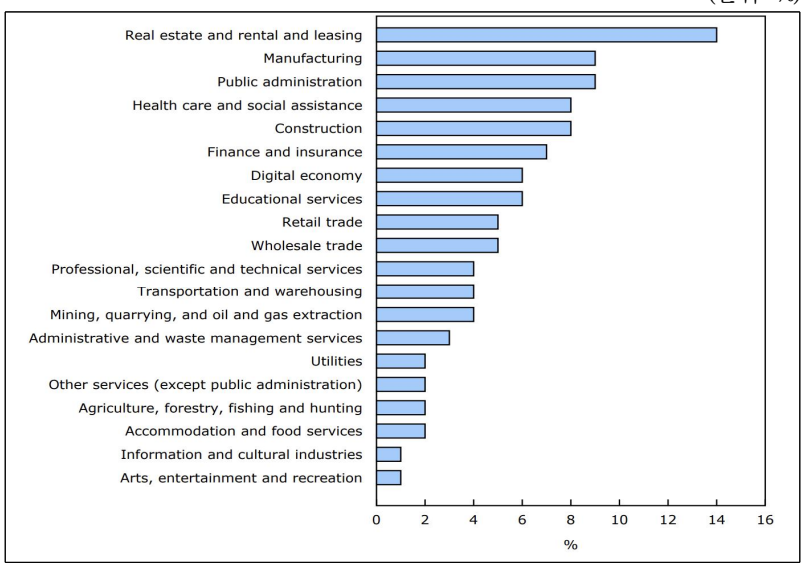
구분	2017	2018	2019	2020
전체 산업	1,991,534	2,083,379	2,161,924	2,076,634
디지털 산업	104,356	110,633	122,018	122,628
ICT 산업				
-하드웨어	6,536	6,913	7,454	6,575
-소프트웨어	41,891	46,067	52,840	54,565
-통신서비스	36,166	36,399	38,133	38,526
-기타 서비스	9,912	9,981	10,151	9,966
디지털 중개 플랫폼	1,762	2,446	3,025	2,504
데이터 및 광고기반 디지털 플랫폼	1,024	1,106	1,326	434
온라인 소매 및 도매업	3,793	4,017	4,611	5,699
디지털 금융 및 보험서비스	2,204	2,476	2,947	2,944
그 외 디지털 운영 업체	1,069	1,229	1,530	1,415

자료: 캐나다 디지털공급사용표; Statistics Canada(2023). p.2

○ 2020년 캐나다 GDP에서 디지털경제가 차지하는 비중은 5.9% 수준(1,230억 달러)으로 캐나다의 산업 간 규모를 비교할 때 금융 및 보험업(1,470억 달러)보다는 작으며, 교육서비스(1,170억달러)보다는 높은 수준임

[그림 3-2] 캐나다의 산업별 GDP 비중

(단위: %)



자료: 캐나다 디지털 공급사용표 2020; Statistics Canada(2023). p.3

○ 한편, 디지털 방식으로 주문된 제품은 전체 공급의 7.6%이며, 디지털 방식으로 주문된 제품 중 69.6%는 디지털 도소매 산업, 29.2%는 국내 도소매 산업, 그리고 디지털 중개 플랫폼은 나머지 1.2%를 담당

□ 호주[12)]

○ 호주는 디지털 경제를 측정하기 위해 2019년 미국의 BEA 방식을 적용한 추정치를 제시

- 호주 통계청(Australian Bureau of Statistics, ABS)은 호주 경제에서 디지털 활동의 기여도를 평가하고, 디지털 경제의 여러 측면을 분석하고 그 영향을 측정하여, 호주 내 디지털 경제 활동의 범위와 경제적 영향을 보다 명확하게 이해하고자 한다고 밝힘

12) 호주의 경우 아래 문헌을 참고하여 작성
Australian Bureau of Statistics(2023), “Digital activity in the Australian economy”, 2021-22 27/10/2023

○ 호주 통계청은 디지털 활동으로 다음 항목의 생산을 포함함
- 디지털 지원 인프라(Digital enabling infrastructure)
- 디지털 미디어(Digital media)
- 전자상거래(E-commerce)

○ 2021~2022년 호주의 디지털 산업은 8.6% 증가했는데, 하드웨어 부문을 제외한 모든 부문의 규모가 증가(같은 기간 경제성장률은 11.7% 수준)
- 디지털 산업은 하드웨어 부문을 제외한 모든 부문의 규모가 증가
- 하드웨어는 6.5% 감소했지만, 지원 서비스(11.4%), 소프트웨어 구매(25.0%), 전자상거래-도매(4.5%), 전자상거래-소매(12.6%)가 성장을 견인

〈표 3-2〉 호주 디지털 산업의 성장률

(단위: %, 명목가격)

구분	2019~2020	2020~2021	2021~2022
전체 경제	2.0	5.0	11.7
디지털산업	8.0	7.7	8.6
디지털 미디어	-10.3	7.8	3.5
소프트웨어	14.0	17.0	25.0
통신서비스	3.1	-0.1	4.6
이커머스-도매	13.4	8.6	4.5
이커머스-소매	25.3	41.8	12.6
지원서비스	7.5	5.8	11.4
하드웨어	6.0	-0.1	-6.5

자료: ABS(2023)

○ 호주의 디지털산업은 2021~2022년 전체 총부가가치의 6.3%를 차지

〈표 3-3〉 호주 디지털 산업의 부가가치

(단위: 백만 달러, %, 명목가격)

구분	디지털 산업의 비중	디지털 산업 부가가치
2019~2020	6.3	116,755
2020~2021	6.4	125,693
2021~2022	6.3	136,563

자료: ABS(2023)

○ 2021~2022년 디지털 활동의 부가가치는 J.정보, 미디어 및 통신업, M.전문, 과학 및 기술 서비스업, F.도매업, G.소매업의 4가지 산업에 집중된 것으로 나타남(4개 산업 비중은 92.1%)

〈표 3-4〉 호주 디지털 활동 총부가가치의 디지털 산업 기여도

(단위: %, 명목가격)

구분	2020~2021	2021~2022
J Information, Media and Telecommunications	**32.2**	**31.1**
M Professional, Scientific and Technical Services	**31.2**	**28.8**
F Wholesale Trade	**19.4**	**21.2**
G Retail Trade	**9.2**	**11.0**
C Manufacturing	3.1	3.1
E Construction	1.8	1.7
O Public Administration and Safety	1.1	0.9
S Other Services	0.9	0.9
I Transport, Postal and Warehousing	0.2	0.3
K Financial and Insurance Services	0.3	0.2
Other	0.6	0.7

자료: ABS(2023)

□ 분석적 방법에 의한 디지털 경제의 파악(ADB 방법론)

○ ADB(2021)는 2021년 아태지역의 디지털 경제를 파악하기 위한 지표를 개발

- ADB의 방법론은 디지털경제에 대한 명확한 정의도 어려운 상황에서 이를

측정하기 위한 방법론까지 제시했다는 점에서 의의가 있음

○ 상대적으로 쉽게 접근할 수 있는 국민계정 데이터와 산업연관분석 프레임워크를 이용하여 디지털 경제를 통해 발생하는 국내총생산(GDP)을 추정

- 디지털 경제의 부가가치 기여도는 디지털 산업 전체의 GDP와 비디지털 산업의 GDP 중 디지털 산업에 의해 발생한 부분으로 구성한다고 파악

○ ADB는 디지털 경제의 주요 디지털 상품으로 하드웨어(Hardware), 소프트웨어 출판(Software publishing), 웹 출판(Web publishing), 통신서비스, 전문지원서비스(Specialized and support services)의 5가지 그룹을 제시

〈표 3-5〉 주요 디지털 상품

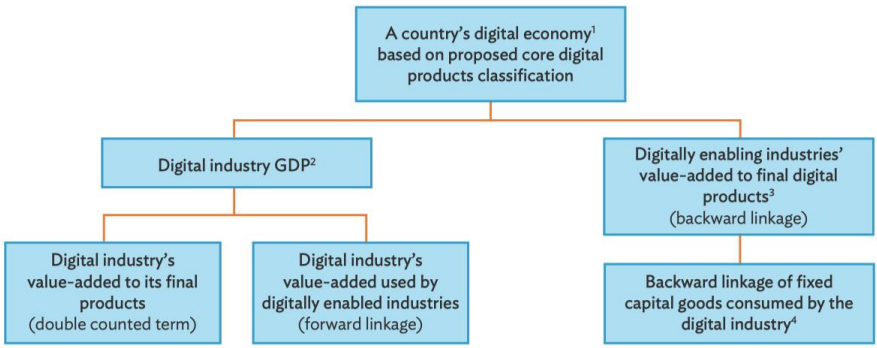
주요 활동 그룹	CPC2 코드	상품
하드웨어	452	Computing machinery and parts and accessories thereof
	475	Disks, tapes, solid-state nonvolatile storage devices, and other media, not recorded
소프트웨어 출판	38582	Software cartridges for video game consoles
	478	Packaged software
	83143	Software originals
	8434	Software downloads
	84391	Oline games
	84392	Online software
웹 출판	83633	Sale of internet advertising space(except on commission)
	843	Online content*
통신서비스	841	Telephony and other telecommunications services
	842	Internet telecommunications services
전문지원 서비스	8313	IT consulting and support services
	83141	IT design and development services for applications
	83142	IT design and development services for networks and systems
	8315	Hosting and IT infrastructure provisioning services
	8316	IT infrastructure and network management services

주: CPC2 843은 웹 출판에서 제외하고, SW 출판(8434, 84391, 84392)에 포함
자료: ADB(2021) p.6

○ ADB의 디지털경제 측정 방법론은 투입산출분석을 기반으로 아래의 단계로 구성

- (1) 주요 디지털상품에 근거한 디지털 경제의 측정
- (2) 디지털산업의 총부가가치의 측정
- (3) 디지털기반산업의 최종 디지털상품 부가가치의 측정
- (4) 디지털기반산업에서 구매한 고정 자본재의 후방연쇄효과의 측정

[그림 3-3] ADB의 디지털경제 측정 방법론



GDP = gross domestic product.
[1] Given by the $GDP_{digital}$ equation, $i^T \hat{V}B\hat{Y}\varepsilon_1 + i^T (\hat{V}B\hat{Y})^T \varepsilon_1 - [diag(\hat{V}B\hat{Y})]^T \varepsilon_1 + (i - \varepsilon_1)^T \hat{V}B\hat{Y}\hat{r}\varepsilon_2$.
[2] Given by the second term of the $GDP_{digital}$ equation.
[3] Given by the first term of the $GDP_{digital}$ equation.
[4] Given by the fourth term of the $GDP_{digital}$ equation.

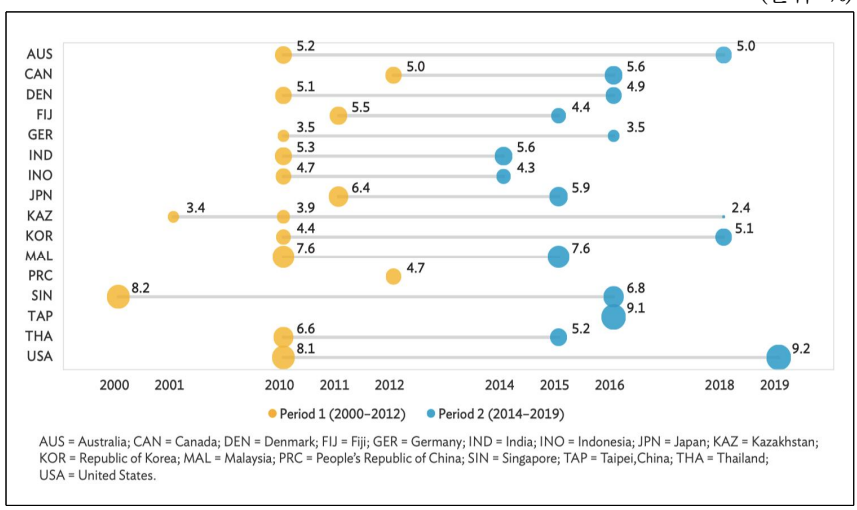
자료: ADB(2021) p.14

○ ADB는 각국의 GDP 대비 디지털경제의 비중을 추정하였는데, 미국은 2019년 9.2%, 캐나다 2016년 5.6%, 호주 2018년 5.0% 등으로 추정되었고, 한국은 2010년 4.4%에서 2018년 5.1%로 증가한 것으로 추정[13)]

13) ADB 방법론으로 측정한 디지털경제의 비중은 한국의 경우 다소 과소평가된 것으로 판단됨. 한국은행이 조사하는 ICT 부문의 명목 부가가치 비중은 2018년 11.0% 수준임

[그림 3-4] 전체 경제에서 디지털경제가 차지하는 비중

(단위: %)



자료: ADB(2021) p.31

제 2 절 한국의 디지털산업 통계 조사 및 분석[14)]

1. 개요

□ 디지털산업 실태조사 통계 발표

○ 디지털화는 다양한 산업군에서 제품 생산 및 판매에 많은 영향을 미치고 있으며, 특히 디지털 주문이 보편화되는 현재 디지털 플랫폼 서비스는 우리 국민의 생활과 밀접하게 연관되어 있는 것을 알 수 있음

○ 한국에서는 디지털산업 실태조사가 2024년 6월에 공표되었는데, 이 조사는 기존의 ICT 산업의 범위에서 나아가 최초로 디지털산업 전반까지 그 범위를 확장한 조사이며, 아래에서는 디지털산업 실태조사의 주요 결과를 요약하여 소개

2. 시장규모

□ 디지털산업 시장 규모

○ 2022년 기준 국내 디지털산업의 시장규모는 1,142조 원 규모로 2022년 전체 산업 매출액 8,772조 원의 13.0%를 차지[15)16)]

※ 디지털산업 실태조사는 사업체의 매출액을 조사하여 개별 사업체 단위에서의 매출을 전체 모집단으로 추정한 값임

○ 대분류 부문별 시장규모를 살펴보면, 디지털 기반 산업 부문은 437.3조 원

---

14) 본 절에서는 「2023 디지털산업 실태조사」의 주요 결과를 인용하여 소개함. 본 절의 모든 표와 조사결과는 「2023 디지털산업 실태조사」의 조사결과를 인용함

15) 22년 우리나라 전체 산업 매출액은 통계청의 전국사업체조사 전체산업 매출액(8,772조 원)을 적용하여 산출

16) 디지털산업에 대한 조사들이 총산출, 부가가치, 매출액 등 다양한 기준으로 조사되고 있기 때문에 주의를 요함. 예를 들어 2022년 기준 한국은행에서 발표하는 우리나라 GDP는 2,323.8조 원, 총부가가치는 2,138.6조 원(GDP에서 순생산물세 제외), 총산출은 5,565.6조 원이며, 통계청의 전체매출액은 8,772조 원 수준임

(38.3%), 디지털 플랫폼 제공 산업 부문은 162.9조 원(14.3%), 디지털플랫폼 활용 산업 부문은 151.9조 원(13.3%), 디지털 관련 산업 부문은 389.3조 원(34.1%)으로 조사

〈표 3-6〉 디지털산업 부문별 매출액 규모

(n=475,315개, 단위: 조 원)

구분	디지털기반 산업 부문	디지털플랫폼 제공 산업 부문	디지털플랫폼 활용 산업 부문	디지털관련 산업 부문
디지털산업	**437.3**	**162.9**	**151.9**	**389.4**
디지털기반산업	412.1	43.5	28.4	71.7
디지털플랫폼 제공산업	9.6	88.1	11.0	30.6
디지털플랫폼 활용산업	1.0	20.5	88.5	14.2
디지털관련산업	14.6	10.8	24.0	272.9

자료: 과기정통부·KISDI(2024),「2023 디지털산업 실태조사」p.39
* 조사범위: 디지털산업 전체(종사자규모 1인 이상 사업장)

3. 산업 구조

□ 해외 진출 현황 및 유형

○ 조사결과에 따르면 2022년 전체 디지털산업의 17.4%가 해외 진출을 한 것으로 나타남

※ 디지털 기반산업이 27.3%로 가장 높게 나타났고, 디지털플랫폼 제공산업 역시 17.0%로 나타남

-종사자 규모가 커질수록 해외 지사 및 해외 법인 설립 유형이 증가했고, 디지털기반산업의 경우 규모가 클수록 해외 진출의 정도가 높은 것으로 조사

〈표 3-7〉 디지털산업 해외 진출 유형

(n=215,887개, 단위: %, 복수응답)

구분	해외 판매 및 수출	라이선스, 개발, 유지보수	해외지사 및 해외법인 설립	기타	해당사항 없음
디지털산업	10.0	6.1	5.5	1.5	82.6
산업분류					
디지털기반산업	16.7	10.2	8.8	1.4	72.7
디지털플랫폼 제공산업	6.9	7.2	8.7	2.7	83.0
디지털플랫폼 활용산업	2.6	1.6	2.6	5.6	87.6
디지털관련산업	0.3	0.1	0.3	0.1	99.3

자료: 과기정통부·KISDI(2024), 「2023 디지털산업 실태조사」 p.42
* 디지털산업(음식·숙박업의 10인 미만 사업체 제외)을 대상으로 조사

□ 고객 유형별 거래 형태

○ 고객 유형별 거래 형태를 살펴보면, 디지털산업 전체에서 B2B(49.3%)가 절반가량을 차지하고, B2C(39.4%), B2G(11.3%) 순으로 나타남

○ 디지털기반산업은 B2B(69.9%) 비중이 상대적으로 크고, 디지털플랫폼 활용산업과 디지털관련산업의 경우 B2C(각각 80.8%, 71.7%) 비중이 큼

〈표 3-8〉 고객 유형별 거래 형태

(n=215,887개, 단위: %)

구분	B2B (기업체)	B2C (소비자)	B2G (정부/공공)
디지털산업	49.3	39.4	11.3
산업분류			
디지털기반산업	69.9	13.3	16.8
디지털플랫폼 제공산업	42.6	47.8	9.6
디지털플랫폼 활용산업	15.8	80.8	3.3
디지털관련산업	24.2	71.7	4.1

자료: 과기정통부·KISDI(2024), 「2023 디지털산업 실태조사」 p.43
* 디지털산업(음식·숙박업의 10인 미만 사업체 제외)을 대상으로 조사

□ 디지털플랫폼 제공 방식

○ 디지털플랫폼 제공산업에서 플랫폼 서비스를 제공사는 방식은 인터넷 웹사이트(85.9%) 비중이 가장 높으며, 그다음은 모바일 웹사이트(47.2%), 모바일 앱(37.0%) 순으로 나타남

- 한편, 종사자 규모가 증가할수록 모바일을 통한 제공방식이 증가

〈표 3-9〉 디지털플랫폼 제공 방식

(n=6,114개, 단위: %, 복수응답)

구분		모바일 앱	인터넷 웹사이트	모바일 웹사이트	기타
디지털플랫폼 제공산업		37.0	85.9	47.2	2.8
산업별 종사자규모					
디지털플랫폼 제공산업	1~9인	28.3	89.3	45.9	1.6
	10~49인	40.9	82.6	44.0	5.7
	50~299인	54.0	83.3	57.2	1.9
	300인 이상	75.6	69.9	56.0	4.5

자료: 과기정통부·KISDI(2024), 「2023 디지털산업 실태조사」 p.46
* 디지털플랫폼 제공산업을 대상으로 조사

□ 디지털플랫폼 제공산업의 비즈니스 유형

○ 디지털플랫폼 제공산업은 광고기반(46.1%) 비즈니스를 가장 많이 수행 중이며, 수수료기반(41.5%), 무료 서비스(21.4%), 직매입 판매(18.8%) 등의 순으로 비즈니스를 수행

〈표 3-10〉 디지털플랫폼 비즈니스 유형(업체 수 비중)

(n=6,114개, 단위: %, 복수응답)

구분	광고 기반	수수료 기반	구독료 기반	라이선스 기반	유료 아이템	직매입 판매	무료 서비스	기타
디지털플랫폼 제공산업	46.1	41.5	17.8	12.3	9.4	18.8	21.4	3.1

자료: 과기정통부·KISDI(2024), 「2023 디지털산업 실태조사」 p.48

□ 디지털플랫폼 활용산업의 영업활동에 활용

○ 디지털플랫폼 활용산업에서 영업활동에 가장 많이 활용하는 플랫폼 유형은 배달 기반 플랫폼(71.2%)이며, 웹서치 포털(54.8%), 커뮤니티(14.4%) 등이 그 뒤를 이어 활용됨[17)]

〈표 3-11〉 영업활동에 활용하는 디지털플랫폼 유형

(n=284,813개, 단위: %, 복수응답)

구분	디지털플랫폼 활용 비중
모바일 앱마켓	10.9
웹서치 포털	54.8
커뮤니티	14.4
메신저 기반 플랫폼	2.9
소셜 미디어 기반 플랫폼	2.7
데이터 기반 매칭 플랫폼	1.5
영상음향 정보 콘텐츠 제공 플랫폼	0.2
전자상거래 기반 플랫폼	5.5
배달 기반 플랫폼	71.2
숙박·여행 플랫폼	13.5
기타	0.9

자료: 과기정통부·KISDI(2024), 「2023 디지털산업 실태조사」 p.63
* 디지털플랫폼 활용산업을 대상으로 조사

□ 디지털플랫폼 도입 시점

○ 디지털플랫폼 활용산업은 2015년 이후 본격적으로 디지털플랫폼을 영업활동에 활용하여 코로나19 팬데믹을 계기로 대폭 확대[18)]

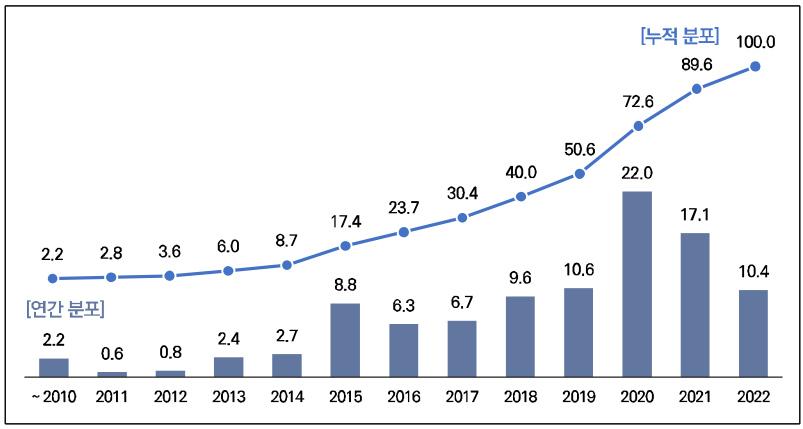
○ 2020년 이후부터 소규모 업체(1~9인 규모)의 디지털플랫폼 활용이 빠르게 확산된 것으로 나타남

17) 이는 조사모집단의 특성이 반영된 현황으로 판단

18) 2022년도 디지털플랫폼 활용 누적분포가 100%인 점은 2022년 기준 디지털플랫폼 활용업체를 대상으로 조사하였기 때문임

[그림 3-5] 디지털플랫폼 활용 시작 시점(연도별)

(n=284,813개, 단위: %)



자료: 과기정통부·KISDI(2024), 「2023 디지털산업 실태조사」 p.66

* 디지털플랫폼 활용산업을 대상으로 조사

* 현재 디지털플랫폼을 활용하는 업체의 활용 시작 시점을 조사

4. 디지털화

□ 디지털화 수준

○ 디지털산업의 디지털화 단계 조사는 '디지털화 단계'(33.2%)에 속한다는 응답이 가장 높게 나타났고, 그 뒤를 '전산화-디지털화 단계'(27.4%), '전산화 단계'(20.2%)로 나타남

- 디지털산업의 디지털화 수준 평균점수는 2.59점이며, 세부산업별로 디지털관련산업(3.23점), 디지털플랫폼 제공산업(3.00점), 디지털기반산업(2.74점), 디지털플랫폼 활용산업(2.37점) 순으로 나타남[19)]
- 한편, 종사자 규모가 클수록 디지털화 수준이 높은 것으로 나타남(디지털화 단계 이상 응답 비율 및 평균 점수)

19) 디지털관련산업의 점수가 높게 나타난 점은 디지털도소매와 디지털금융의 산업 특성으로 판단됨

〈표 3-12〉 디지털산업의 디지털화 수준

(n=475,315개, 단위: %, 점/5점 척도)

구분	전산화 단계 (1점)	전산화 - 디지털화 단계 (2점)	디지털화 단계 (3점)	디지털화 - 디지털 전환 단계 (4점)	디지털 전환 단계 (5점)	평균
디지털산업	20.2	27.4	33.2	11.6	7.5	2.59
산업분류						
디지털기반산업	25.0	18.3	28.4	14.4	13.9	2.74
디지털플랫폼 제공산업	16.2	12.2	43.2	12.3	16.1	3.00
디지털플랫폼 활용산업	21.6	35.5	31.8	6.1	4.9	2.37
디지털관련산업	5.4	10.1	47.3	30.9	6.3	3.23

자료: 과기정통부·KISDI(2024), 「2023 디지털산업 실태조사」 p.79
* 디지털산업을 대상으로 조사

□ 디지털플랫폼 제공산업의 인공지능 기반 기술 이용 현황

○ 디지털플랫폼 제공산업을 대상으로 인공지능 기반 기술 이용 현황을 조사한 결과, 16.9%가 '이용하고 있다'고 응답했고, 52.1%는 '이용하지 않지만, 향후 이용계획이 있다'고 응답

- 한편, 사업체 규모별로 살펴보면, 종사자 규모가 증가하면서 인공지능 기반 기술 이용률은 증가하는 것으로 나타남

※ 인공지능 기반 기술 이용 현황: 50~299인 29.0%, 300인 이상 35.0%

〈표 3-13〉 인공지능 기반 기술 이용 및 계획

(n=6,114개, 단위: %, 복수응답)

구분		이용하고 있다	이용하지 않지만, 향후 이용계획이 있다	이용하지 않고, 향후 이용계획도 없다
디지털플랫폼 제공산업		16.9	52.1	31.0
산업별 종사자규모				
디지털플랫폼 제공산업	1~9인	9.8	54.3	35.9
	10~49인	23.2	53.0	23.7
	50~299인	29.0	40.0	31.0
	300인 이상	35.0	50.6	14.4

자료: 과기정통부·KISDI(2024), 「2023 디지털산업 실태조사」 p.47
* 디지털플랫폼 제공산업을 대상으로 조사

□ 디지털화 추진 이유

○ 디지털화를 추진하는 이유를 조사했고, 디지털산업은 업무 프로세스 효율화(72.3%), 거래 채널 다양화(10.1%), 마케팅 및 소비자 분석 강화(9.1%) 등의 이유 순으로 디지털화를 추진한다고 응답

〈표 3-14〉 디지털화 추진 이유(1순위 기준)

(n=215,887개, 단위: %)

구분	내부 업무 프로세스 효율화	마케팅, 소비자 분석 강화	제품/서비스 거래 채널 다양화 (고객 확대 등)	신규 제품/서비스 기획, 사업 다각화를 통한 매출 증대	비용 절감 및 자동화	고객접점 확대 및 피드백/환류 강화	기타
디지털산업	72.3	9.1	10.1	4.3	3.2	0.9	0.0
산업분류							
디지털기반산업	85.6	6.3	4.3	2.4	0.9	0.5	0.1
디지털플랫폼 제공산업	56.7	19.6	8.3	7.5	5.1	2.9	0.0
디지털플랫폼 활용산업	35.7	22.7	19.6	15.5	5.4	1.0	0.1
디지털관련산업	63.3	8.0	17.6	3.3	6.5	1.4	0.0

자료: 과기정통부·KISDI(2024), 「2023 디지털산업 실태조사」 p.85
* 디지털산업(음식·숙박업의 경우 10인 미만 사업체 제외)을 대상으로 조사

□ 디지털 주문

○ 디지털산업에서 디지털 주문, 즉 인터넷, 모바일을 통해 상품을 주문받는 현황을 매출액 기준으로 살펴보면, 디지털 주문을 받은 매출액 비율은 58.4%로 조사됨

〈표 3-15〉 디지털산업의 디지털 주문 현황(매출액 기준)

(n=475,315개, 단위: %)

구분	디지털을 통한 주문	비디지털을 통한 주문
디지털산업	58.4	41.6
산업분류		
디지털기반산업	49.1	50.9
디지털플랫폼 제공산업	83.0	17.0
디지털플랫폼 활용산업	66.4	33.6
디지털관련산업	60.7	39.3

자료: 과기정통부·KISDI(2024), 「2023 디지털산업 실태조사」 p.86

* 디지털산업을 대상으로 조사

* 디지털주문 매출액을 산출한 후 구성비율 산출

5. 기술

□ 디지털 신기술 개발 및 도입

1) 디지털 신기술 개발 및 도입

○ 디지털 산업 중 디지털 신기술을 개발 혹은 개발 중이라고 응답한 업체는 25.6%, 외부 기술을 도입한 업체는 38.2%임

- 디지털기반산업과 디지털플랫폼 제공산업의 경우 응답 비중이 타 산업 대비 높게 나타남

〈표 3-16〉 디지털 신기술 개발 및 도입 현황

(n=215,887개, 단위: %)

구분	개발 혹은 개발 중인 기술		도입한 기술	
	있음	없음	있음	없음
디지털산업	25.6	74.4	38.2	61.8
산업분류				
디지털기반산업	41.1	58.9	47.7	52.3
디지털플랫폼 제공산업	42.0	58.0	40.5	59.5
디지털플랫폼 활용산업	5.3	94.7	8.7	91.3
디지털관련산업	2.6	97.4	31.7	68.3

자료: 과기정통부·KISDI(2024), 「2023 디지털산업 실태조사」 p.101

* 디지털산업(음식·숙박업의 경우 10인 미만 사업체 제외)을 대상으로 조사

2) 디지털 신기술 개발

○ 디지털산업에서 개발하고 있는 디지털 신기술은 인공지능(46.1%), 빅데이터(44.5%), 클라우드컴퓨팅(30.3%), 가상증강현실(20.3%), 모바일(18.9%), 사물인터넷(16.2%) 등의 순으로 나타남

〈표 3-17〉 디지털 신기술 개발 현황

(n=15,247개, 단위: %, 복수응답)

구분	인공지능	AI 반도체	사물인터넷	클라우드컴퓨팅	빅데이터	모바일(5G/6G)	로봇(공학)	가상·증강현실	블록체인	3D 프린팅	무인항공(드론)	자율주행	양자컴퓨터
디지털산업	46.1	5.3	16.2	30.3	44.5	18.9	4.7	20.3	9.2	2.0	2.8	11.2	1.2
산업분류													
디지털기반산업	51.2	5.7	17.1	36.0	46.6	12.1	4.6	16.4	10.1	2.3	3.2	5.4	1.5
디지털플랫폼 제공산업	47.1	6.8	12.9	16.3	59.8	9.7	7.4	10.0	6.3	1.7	1.8	6.7	0.5
디지털플랫폼 활용산업	12.1	4.4	39.4	3.8	17.9	41.5	4.8	1.6	1.9	0.5	0.8	0.4	0.0
디지털관련산업	18.0	1.8	4.8	8.9	24.1	70.6	3.1	65.3	7.3	0.2	1.4	63.4	0.1

자료: 과기정통부·KISDI(2024), 「2023 디지털산업 실태조사」 p.103

* 디지털 신기술 설문에서 신기술을 하나라도 개발 혹은 개발 중이라고 응답한 업체(10인 미만 사 업체 제외)를 대상으로 조사

3) 디지털 신기술 도입

○ 디지털산업에서 도입하고 있는 디지털 신기술 유형은 클라우드컴퓨팅(52.5%), 인공지능(48.2%), 빅데이터(40.9%) 등의 순으로 나타남

〈표 3-18〉 디지털 신기술 도입 현황

(n=18,267개, 단위: %, 복수응답)

구분	인공지능	AI 반도체	사물인터넷	클라우드컴퓨팅	빅데이터	모바일(5G/6G)	로봇(공학)	가상·증강현실	블록체인	3D 프린팅	무인항공(드론)	자율주행	양자컴퓨터
디지털산업	48.2	6.1	10.5	52.5	40.9	16.2	4.5	17.2	22.6	2.4	3.8	7.4	0.6
산업분류													
디지털기반산업	47.4	8.6	14.9	45.4	41.9	13.9	6.4	17.2	13.2	3.6	5.4	2.8	0.9
디지털플랫폼 제공산업	51.2	6.9	12.5	34.6	49.4	12.0	4.7	7.6	6.0	1.7	1.6	2.9	0.2
디지털플랫폼 활용산업	14.4	4.2	5.7	60.7	31.9	17.0	3.7	8.3	2.9	0.4	2.1	1.0	0.0
디지털관련산업	51.9	0.6	0.7	72.6	37.0	22.5	0.0	20.4	49.9	0.1	1.0	19.4	0.0

자료: 과기정통부·KISDI(2024), 「2023 디지털산업 실태조사」 p.106

* 디지털 신기술 설문에서 신기술을 하나라도 도입하였다고 응답한 업체(10인 미만 사업체 제외)를 대상으로 조사

제 4 장 디지털플랫폼 사업자 현황 분석

제 1 절 디지털플랫폼 제공 사업자의 식별

○ 디지털플랫폼 제공산업은 디지털화로 인해 발생한 거래방식의 변화를 반영한 분류이기 때문에 특정한 표준산업분류가 존재하지 않음

○ 최근 디지털플랫폼에 관한 관심이 높아지는 상황에서 디지털플랫폼에 관한 보다 정확한 통계가 필요하며, 디지털플랫폼 제공 사업자를 식별할 수 있는 방안 마련이 필요

○ 디지털산업 실태조사도 디지털플랫폼 제공산업의 모집단 구축을 위해 판별조사를 수행하고 있음

○ 중분류 수준의 201. 데이터 및 광고 기반 디지털 플랫폼의 경우 한국표준산업분류 63120. 포털 및 기타 인터넷 정보 매개 서비스업을 영위하는 사업체를 모집단으로 포섭하고 있으나, 202. 디지털 중개 플랫폼의 경우 다양한 산업에 분포하는 상황임

○ 202. 디지털 중개 플랫폼이 인터넷, 모바일 등 온라인 인터페이스를 통해 다수의 생산자와 소비자가 서로 원하는 제품과 서비스를 거래할 수 있도록 중개 기능을 제공한다는 점에서 공정거래위원회에 있는 통신판매사업자 신고 리스트를 활용하고자 함

○ 여기서 찾으려는 업체는 디지털 중개 플랫폼 사업자이며, 이들은 법률상 통신판매중개사업자이며, 통신판매중개사업자는 통신판매사업자 신고 리스트에 포함되어 있음

- 다만, 통신판매중개사업자는 통신판매사업자 신고를 하여야 하지만, 별도의 통신판매중개사업자 신고 의무가 없어 이들을 웹사이트 정보를 통해 구분하고자 함

○ 아래에서는 웹사이트 정보를 통해 디지털플랫폼 제공 사업자를 식별하는 방법론에 관해 살펴보기로 함

제 2 절 디지털플랫폼 제공 사업자 분류 자동화

1. 디지털 중개 플랫폼 제공사업체 분류 자동화 필요성

가. 통신판매업 규모의 증가

○ 통신판매 및 통신판매 중개업의 규모는, 인터넷의 발전에 따라 계속해서 커지고 있음

- 통신판매는 자신의 재화를 직접 판매하는 업태를 말하고, 통신판매 중개업은 판매자와 구매자를 연결하는 업태를 의미함

○ 통계청의 2023년 12월 온라인 쇼핑 동향에 따르면, 온라인 쇼핑의 규모는 전년 동월 대비 7.7% 상승한 20조 3,574억 원에 달함. 또한 2023년 연간 온라인 쇼핑 규모는 227조 3,470억 원으로 추정됨

○ 통계청의 2013년 4/4분기 전자상거래 및 사이버쇼핑 동향 보도자료를 살펴보면, 2013년의 사이버쇼핑 거래액은 연간 38조 4,940억 원으로 추정되었음. 또한 10년 동안 온라인 쇼핑몰의 규모는 590% 상승함

○ 이처럼 통신판매업의 규모가 증가함에 따라 이에 따른 실태조사를 포함한 관리가 필요해짐

나. 통신판매업의 특성과 조사의 어려움

○ 통신판매업종은 업체 수가 많은 업태 중 하나임. 2020년까지 사업자 수 1위였던 한식 음식업을 제치고, 2021년에는 통신판매업이 1위로 올라섰음 (신진욱, 2023)

- 특히 2018년과 비교했을 때 통신판매업의 업체 수는 2022년 말까지 148.4%나 증가하는 모습을 보임

○ 통신판매업종의 또 다른 특성은 평균 존속 연수가 굉장히 짧다는 것임
- 코로나19 이후로 2021년 한 해에만 73,100여 개의 의류 쇼핑몰이 생겨났고, 같은 해 20,311개의 쇼핑몰이 폐업했음
- 2022년 말 기준 100대 생활업종 업종별 존속 연수 상하위 현황을 보면, 통신판매업종의 존속 연수는 평균 2년 7개월로 가장 짧은 모습을 보이고 있음

○ 이는 매년 실태조사 시 업체가 변경될 수 있음을 뜻하고, 자동적으로 통신판매업자의 현황을 조사할 필요가 있음

○ 통신판매업은 영세한 경우를 제외하고는 신고가 필수이며, 통신판매업 신고 사업체의 목록은 공정거래위원회에서 확인할 수 있음

○ 그러나 일부 통신판매 중개사업자의 경우 법적인 문제를 예방하기 위해 통신판매업 등록을 병행하는 경우가 존재함

○ 이는 통신판매업 조사를 위해 공정거래위원회의 데이터베이스를 사용할 때 100% 신뢰할 수 없음을 뜻함

다. 통신판매 중개업의 실태 조사 현황

○ 통신판매 중개사업자는 별도의 신고가 필요하지 않음

○ 전자상거래 등에서의 「소비자보호에 관한 법률」 제2조 제4호에 따르면 통신판매 중개에 대한 정의가 나와 있으며, 이는 통신판매를 알선하는 행위임

○ 통신판매 신고 의무는 위 법 12조에 따라 통신판매자에게 부과된 의무이기 때문에, 통신판매 중개업자는 해당하지 않음

○ 즉, 통신판매업처럼 데이터베이스를 받아볼 수 없음을 의미함

○ 이 때문에 통신판매중개자에 대한 실태조사를 진행할 때 업체 선정 등 많은 어려움이 존재함
- 한편 실태조사는 해당 업태의 객관적인 현황을 파악하고 기초 데이터를 구축하기 위한 중요한 업무임. 통신판매 관련 사업체들에 대한 실태조사 또한 세수 확보 등 기초 데이터 구축을 위해 매우 중요함

○ 현재는 조사를 위해서 조사원을 고용해 데이터를 수집하고 있으나, 매년 이들의 업태를 분류하기 위해 인력을 투입하기보다는 분류 모델을 연구 개발해 자동화한다면 비용과 시간, 그리고 오류를 줄일 수 있을 것임

2. 디지털 중개 플랫폼 제공 사업체 분류 모델 개발

가. 모델 개발의 목표

○ 현 모델의 목적은 사업체의 웹사이트를 확인하여, 해당 사업체가 어느 분류체계(통신판매 사업자, 중개사업자, 그 외)에 해당하는 것인지 분류하는 것임

- 수집된 키워드를 기반으로 사업체의 업태(통신판매 사업자, 통신판매 중개사업자, 그 외)로 분류하고자 함

○ 추가적으로 해당 홈페이지에 자체적인 결제 시스템이 존재하는지 확인하는 것을 목적으로 함

[그림 4-1] 모델 개발 프로세스의 도식화

나. 모델 개발의 실제: 키워드 선정 및 검색어 사전 생성

1) 키워드 선정

○ 키워드는 사업체 웹사이트를 확인해 어느 업태 분류에 해당하는지를 확인

하는 목적에 적합하게 선정함

- 이를 위해서 통신판매업자를 확인하기 위한 키워드, 통신판매 중개사업자를 확인하기 위한 키워드가 필요함

[그림 4-2] 선정된 키워드 분류

○ 수집할 키워드는 사업자등록번호, 통신판매업, 통신판매 중개사업자, 당사자, 거래에 대한 책임, 입점, 결제 시스템의 총 7개로 결정하였음

○ 주요 키워드는 사업자등록번호, 통신판매업, 통신판매 중개사업자로 총 3개

- 주요 키워드는 해당 사업체에 대한 주요 정보를 담고 있음
- 주요 정보란 사업자등록번호와 해당 사업체의 업태 구분을 의미함

○ 보조 키워드는 당사자, 거래에 대한 책임, 입점, 결제 시스템으로 총 4개임

- 보조 키워드는 주요 키워드, 특히 해당 사업체의 업태를 구분할 때 도움을 주는 키워드임
- 당사자, 거래에 대한 책임, 입점이라는 키워드는 통신판매중개사업자 홈페이지에 자주 등장하는 키워드이고, 결제 시스템은 통신판매업자와 통신판매중개사업자 홈페이지에 존재하는 기능임

2) 검색어 사전 생성

가) 웹사이트 표기 방식의 특성

○ 해당 키워드를 검색할 때 각 홈페이지별로 표현되는 방식이 상이함

- 사업자등록번호를 표기에는 사업자등록번호, 사업자번호, 사업번호, BIZNO 등 굉장히 많은 표기 방식이 존재함
- 사업자등록번호 형식 또한 XXX-XX-XXXXX, XXX XX XXXXX, XX XXXXXXXX으로 여러 가지 방식으로 표기되어 있음
- 이 문제는 통신판매 중개사업자 키워드에서 많이 나타남

[그림 4-3] 업체별 표현 방식 차이

c_name	homepage1	세부사항 homepage1homepage2homepage4
랜트	https://www.lafent.com/index.html?udmp=P	통신판매중개서비스의 경우 라펜트는 통신판매의 당사자가 아니며 개별 판매자가 제공하는 서비스에 대한 이행,
올버스	https://allbus.kr/	올버스는 통신판매중개사업자로 전세버스 운행계약의 거래당사자가 아니며, 운행계약 및 이용약관은 버스회사나 버스기사가 소속되어 있는 버스회사의 이용약관을 따르고 있습니다
주식회사 멀치	http://www.smelchi.com/	주식회사 멀치는 통신판매 중개자로서 오픈마켓 멀치쇼핑의 거래당사자가 아니며 , 입점판매자가 등록한 상품정보 및 거래에 대해 주식회사 멀치는 일체 책임을 지지 않습니다.
심플	https://www.simpol.co.kr/main/main.php	심플에 등록된 판매플품과 물품의 내용은 심플이 아닌 개별 판매자가 등록한 것으로서, 심플은 중계시스템만 제공하며 그 등록 내용에 대하여 일체의 책임을 지지 않습니다.
엑스플랜트	https://www.xplant.co.kr/	엑스플랜트에 등록된 판매플품과 물품의 내용은 엑스플랜트가 아닌 개별 판매자가 등록한 것으로서, 엑스플랜트는 중계시스템만 제공하며 그 등록 내용에 대하여 일체의 책임을 지
마켓보로	https://www.foodspring.co.kr/	(주)마켓보로는 통신판매의 당사자가 아닙니다. 따라서 (주)마켓보로는 상품 거래정보 및 거래에 대하여 책임을 지지 않습니다.
필웨이	https://www.feelway.com/	필웨이에서는 장품만을 거래할 수 있으며, 불확실한 상품의 거래는 엄격히 제한됩니다. 상품은 개별 판매자가 등록한 것으로, 필웨이는 중개시스템을 제공하며, 그 등록내용에 대하여는 일체의
핀다	https://finda.co.kr/about-us	핀다는 금융감독원의 Open API를 사용하고 있으며 일부 제휴 기관들로부터 제공받은 금융 상품 정보를 전달하고 있습니다. 핀다에서는 제공된 정보에 대한 법적인 책임을 지지 않습니다. 핀다에서 제공하는 금융
인앤드림	https://inndream.co.kr/	(주) 인앤드림은 통신판매 중개자로써 통신판매 당사자가 아니며 등록된 매물의 정보 및 계약등과 관련한 의무와 책임은 해당 중개사무소에 있습니다.
콜버스	https://callbus.com/	콜버스업은 통신판매의 당사자가 아니라 중개자이며 상품, 거래정보 및 거래와 관련하여 판매자와 고의 또는 과실로 소비자에게 발생하는 손해에 대해 일체 책임을 지지 않습니다.
우리샵	https://woorishop.com/	우리샵에 등록된 판매플품과 물품의 내용은 우리샵이 아닌 개별 판매자가 등록한 것으로서, 우리샵은 통신판매중개업자로서 거래의 당사자가 아니며, 중개 시스템만 제공합니다.
주식회사 트리플	https://triple.guide/	㈜인터파크트리플은 통신판매중개로서 통신판매의 당사자가 아니며 상품 거래정보 및 거래등에 대해 책임을 지지 않습니다.
㈜위민	http://pension-king.com/	펜션킹은 펜션예약중개자로 펜션입점주가 등록한 정보및 현장에서 발생되는 일에 대해 책임을 지지 않습니다.
㈜한투한	https://www.worldshopping.co.kr/	주)한투한은 해외구매 및 경매를 중개 및 대행하는 서비스를 제공하며 상품의 이미지 및 등록 내용에 관해서는 일체 책임을 지지 않습니다.
㈜ 수입육거래소	https://www.meat-asia.com/	https://www.meat-asia.com/dsm_comm/img/footer_quick/copyright_lice.jpg
㈜ 멀치	https://www.smelchi.com/	주식회사 멀치는 통신판매 중개자로서 오픈마켓 멀치쇼핑의 거래당사자가 아니며 , 입점판매자가 등록한 상품정보 및 거래에 대해 주식회사 멀치는 일체 책임을 지지 않습니다
주식회사 트렌비	https://www.trenbe.com/	사이버몰 내 판매되는 상품 중에는 개별 판매자가 판매하는 마켓플레이스 상품이 포함되어 있습니다. (주)트렌비는 통신판매 중개자로서 통신판매의 당사자가 아니므로 개별 판매자가 등록한 상품
다나와	https://www.danawa.com/	(주)커넥트웨이브는 상품판매와 직접적인 관련이 없으며, 모든 상거래의 책임은 구매자와 판매자에게 있습니다. 이에 대해 (주)커넥트웨이브는 일체의 책임을 지지 않습니다.
도매창고	https://www.wholesaledepot.co.kr/	https://www.wholesaledepot.co.kr/shop/images/main/5_bar2_text.png
㈜오케이몰	https://www.okmall.com/#benefits	오케이몰은 아무런 책임 없이 판매만 중개하는 일반적인 쇼핑몰과는 완전히 다릅니다. 모든 상품을 100% 직매입하는 오케이몰은 상품의 입고부터 검수, 촬영, 등록, 포장, 재고
㈜필웨이	https://www.feelway.com/	필웨이에서는 장품만을 거래할 수 있으며, 불확실한 상품의 거래는 엄격히 제한됩니다. 상품은 개별 판매자가 등록한 것으로, 필웨이는 중개시스템을 제공하며, 그 등록내용에 대하여는 일체의
쿠쿠전자(주)	https://www.cuckoo.co.kr/#	여러 회사의 사업자등록번호가 존재
(주)더우리샵	https://woorishop.com/	우리샵에 등록된 판매플품과 물품의 내용은 우리샵이 아닌 개별 판매자가 등록한 것으로서, 우리샵은 통신판매중개업자로서 거래의 당사자가 아니며, 중개 시스템만 제공합니다.

- [그림 4-3]은 사업체가 통신판매 중개사업자임을 나타내는 방식들을 모아놓은 표임
- 이를 모두 탐지해서 검색하기 위해 하나의 키워드에 여러 가지 검색어를 포함해야 함

나) HTML 수집 시 전처리 방식 선정

○ 위 문제를 해결하기 위해, 사업체들의 모든 HTML을 수집해 TEXT로 변환 후 추가적인 전처리 작업을 진행하기로 함

- 앞서 설명한 형식에 맞추어 검색어 사전을 선정하기 위해서는 굉장히 많은 단어가 필요함. 따라서 추가적인 전처리 방법이 훨씬 효율적임
- 사업자등록번호를 예시로 들면, 사업자등록번호라는 문구 표기에 5가지 경우, 숫자 표기에 3가지 경우가 존재하게 되면 총 15가지의 단어를 집어넣어야 함
- 이를 단순화 하기 위해 TEXT 내부의 특수문자와 띄어쓰기를 제거하는 과정을 거침
- 사업자등록번호 XXX-XX-XXXXX, 사업자등록번호 XXX XXXXXXX로 다르게 표기된 경우에도 위 과정을 거치면 같은 단어로 변경됨. 이를 통해 검색어를 단순화할 수 있었음

다) 검색어 사전의 생성

○ 검색어 사전은 휴리스틱한 방식을 활용하여 계속해서 검색어 업데이트를 진행하였음

- 기초 사전을 연구진 조사를 통해 작성한 후, 사용할 검색 알고리즘을 실행함
- 검색 알고리즘 실행 결과를 확인하며 사업체 홈페이지에서 새로운 검색어를 찾아 사전에 추가함

○ 위와 같은 방식으로 분류 목적에 맞는 검색어 사전을 아래와 같이 구성

〈표 4-1〉 키워드별 검색어 사전

키워드	검색어
사업자등록번호	'사업자등록번호\\d{10}', '사업자등록번호안내\\d{10}', 'bizno\\d{10}', 'no\\d{10}', '사업자번호\\d{10}','\\|\\d{10}\\|', '사업자등록\\d{10}', 'LICENSE\\d{10}', 'BN\\d{10}', \'number\\d{10}', '사업자등록자번호\\d{10}'
통신판매업	'통신판매업', '통신판매'
통신판매 중개자	'통신판매중개자', '통신판매중개', '개별판매자', '중계시스템', '개별판매자', '중개시스템', '중개자', '통신판매의 당사자', '마켓플레이스', '상거래의 책임'
당사자	'거래당사자', '판매 당사자', '거래당사자', '통신판매의 당사자'
거래에 대한 책임	'책임을지지 않', '거래.*책임'
입점	'입점', '입점안내'
결제 시스템	'안전거래', '장바구니', '지급보증', 'cart'

○ 파이썬의 re 패키지를 사용하여 검색할 예정이기 때문에 정규 표현식을 사용해서 검색어를 표현함

- 정규 표현식이란 10개의 무작위 숫자를 표현할 때 d{10}으로 표현하는 방식을 의미함
- 이를 통해 여러 표현을 하나로 압축해 검색하는 것이 가능함
- re의 findall 기능은 한 문자열에서 완전표현식에 해당하는 부분이 있는지를 검색하게 해줌
- 예를 들어 '통신판매중개서비스의 경우 라펜트는...' 이라는 문자열이 있을 때 '통신판매중개'라는 부분이 있기 때문에 re. findall 기능에 검색되어 이를 찾을 수 있게 됨

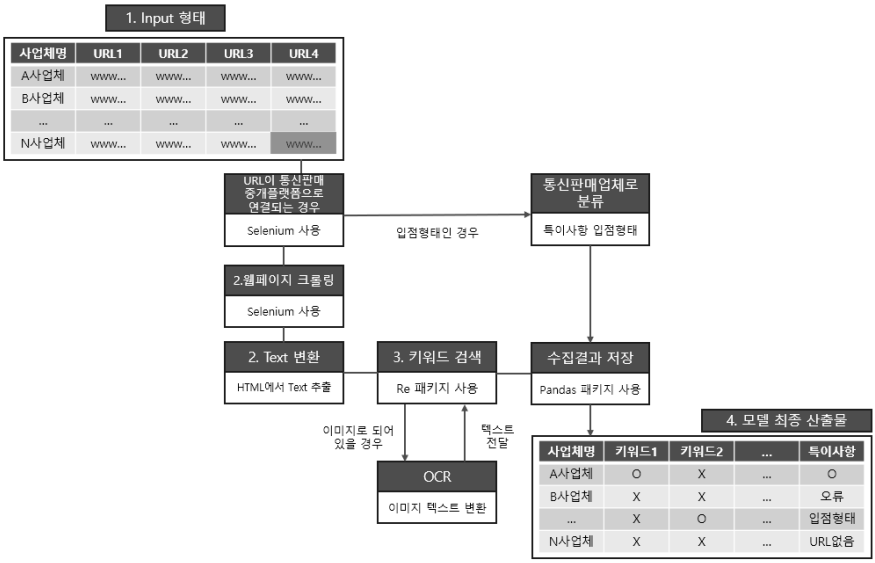
○ 해당 검색어 사전은 언제든지 수정될 수 있으므로, 이를 통해 모델 성능의 향상을 기대할 수 있음

다. 모델 개발의 실제: 주요 알고리즘 설명

○ 모델의 전체 프로세스는 다음과 같이 도식화하였음

○ 각 프로세스별 상세 설명을 통해 자세한 알고리즘을 설명함

[그림 4-4] 전체 프로세스 도식화



1) 웹페이지 크롤링(HTML 수집) 및 텍스트 변환

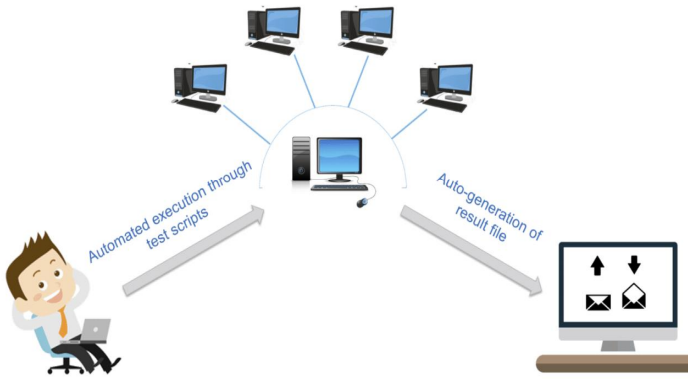
○ 하나의 URL을 단위로 해당 기능이 수행됨

○ Selenium 패키지를 사용해서 이 기능을 수행함

- selenium은 python 패키지 중 하나로, 웹사이트로 접속과 웹사이트 내부에서의 활동을 코드를 통해 자동화할 수 있게 해줌
- 최근 웹사이트는 대부분 특정 activities(스크롤, 클릭)에 반응해 웹사이트 내용이 달라지는 동적 웹사이트임
- 따라서 웹사이트의 모든 정보를 수집하기 위해서 selenium을 사용하였음

○ 본 연구에 이용한 사업체의 웹사이트들은 상당수 동적페이지로, selenium을 사용하지 않으면 페이지 내부의 모든 정보를 가져올 수 없음

[그림 4-5] Selenium(techsearch, 2023. 4.)



자료: “How Does Selenium Automation Testing Work?”, Infinigeek(2018), 2024. 7월, https://infinigeek.com/selenium-automation-testing-work/

○ 또한 window 창의 크기에 따라서 UI가 바뀌는 경우가 존재함

[그림 4-6] 네이버 쇼핑 UI 차이

자료: https://shopping.naver.com/home

○ 위의 [그림 4-6]은 네이버 쇼핑의 window 크기에 따른 하단 정보의 차이를 나타내고 있음
○ 전체 창이 아닐 경우 입점 문의나 사업자 정보가 사라진 모습을 볼 수 있음
○ 이를 해결하기 위해 window 창을 전체 창으로 놓은 뒤 스크롤 다운을 통해 홈페이지 전체가 로딩되도록 만든 후, 해당 페이지의 HTML을 가져옴
○ 위와 같은 기능은 Selenium과 같은 자동화 소프트웨어가 아니면 사용할 수 없음

가) 텍스트 변환

○ 이렇게 가져온 HTML은 문서의 구조를 모두 포함하고 있음
○ 문서의 구조가 아니라 내부에 적혀 있는 텍스트만이 필요하므로, HTML을 텍스트로 변경하는 과정을 거쳐야 함
○ 검색어 사전 생성에서 언급했듯이, 텍스트로 변경 후에 특수문자 및 띄어쓰기를 제거하는 과정을 거쳐 전처리를 완료하였음

나) URL이 통신판매중개 플랫폼으로 연결되는 경우

○ 규모가 작은 통신판매 중개플랫폼의 경우와 달리 오픈마켓은 하나의 시장과 같은 형태를 띠고 있음
○ 오픈마켓에 입점한 사업체는 해당 URL을 본인의 점포로 사용하기 때문에, 이를 본인의 페이지로 소개하는 경우가 존재함
○ 이 경우 해당 사업체는 통신판매사업자임에도 불구하고, 오픈마켓의 페이지이기 때문에 통신판매 중개사업자로 분류됨
  - 오픈마켓에 입점해 있는 경우 해당 페이지 하단에 오픈마켓이 통신판매중개사업자임을 나타내는 문구가 존재하기 때문
○ 이 문제를 해결하기 위해 주요 오픈마켓 URL을 따로 등록하고, re패키지를 사용해 이를 탐지함

〈표 4-2〉 주요 오픈마켓 URL

오픈마켓	URL
옥션	'https?://stores.auction.co.kr/\w+'
네이버 스마트스토어	'https?://smartstore.naver.com/\w+'
인터파크	'https?://store.interpark.com/\w+'
롯데온	'https?://www.lotteon.com/p/display/seller'
쿠팡	r'https?://shop.coupang.com'
지마켓	r'https?://minishop.gmarket.co.kr/'
홈플러스	r'https?://mfront.homeplus.co.kr/sellershop/'
SSG	r'https?://www.ssg.com/sellerhome/'

○ 해당 URL을 게시한 사업체의 경우 통신판매업체로 분류하고 URL에서 별도의 키워드 수집은 진행하지 않음

○ 특이사항에 입점 형태로 표기함

2) 키워드 검색

○ 변환된 텍스트에서 해당 키워드가 존재하는지 검색

○ 키워드 검색을 위해 다음과 같은 초기 행을 설정함

〈표 4-3〉 초기 행 설정

c_name	사업자 등록번호	통신 판매업	통신판매 중개자	당사자	거래에 대한 책임	입점	결제 시스템	특이 사항
A 사업체	X	X	X	X	X	X	X	X

○ 사업체 명에는 현재 검색을 진행하고 있는 사업체 상호명을 삽입. 나머지 행은 아직 검색이 진행되지 않았으므로 모두 X로 표기됨

○ A 사업체의 URL의 HTML을 먼저 텍스트 변환 후, 특정 키워드 존재 여부 확인

○ 예를 들어, 텍스트 내에 사업자등록번호를 검색하기 위해 re 패키지를 활용. 이 re 패키지를 통해 사업자등록번호가 123-45-678910이라는 사업자등록번호가 검색되었다면 다음과 같이 해당 행이 업데이트됨

〈표 4-4〉 키워드 검색 결과 삽입

c_name	사업자 등록번호	통신 판매업	통신판매 중개자	당사자	거래에 대한 책임	입점	결제 시스템	특이 사항
A 사업체	123-45-678 910	X	X	X	X	X	X	X

○ 이후 같은 과정을 통해 모든 행을 업데이트함

〈표 4-5〉 키워드 검색 결과 삽입

c_name	사업자 등록번호	통신 판매업	통신판매 중개자	당사자	거래에 대한 책임	입점	결제 시스템	특이 사항
A 사업체	123-45-678 910	O	X	X	X	X	X	X

○ A 사업체의 모든 URL에 대한 검색이 완료되었으면 해당 행을 저장 후 다음 사업체로 넘어감

○ 그러나 HTML을 텍스트로 변환해 이를 진행하지만, 일부 사업체는 사업체 정보를 이미지 파일로 게시해 놓는 경우가 존재함. 이를 해결하기 위해 아래와 같이 검색 프로세스를 두 단계로 나눔

* 텍스트로 되어 있는 경우

○ 수집된 웹페이지의 정보가 text 형태이기 때문에 별도의 처리없이 키워드의 존재를 확인

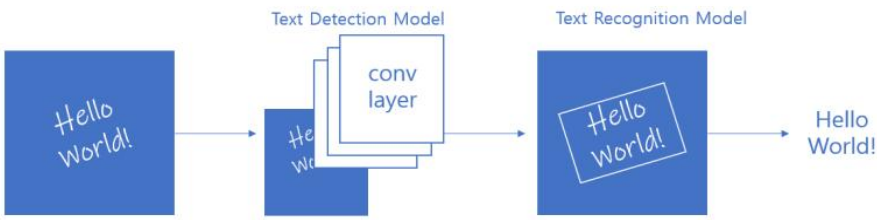
* 이미지로 되어 있는 경우

○ 이미지로 되어 있는 경우 웹페이지의 정보에 키워드 검색 시 어떠한 키워드도 감지되지 않음

○ 이 경우 OCR 기술을 적용해 이미지 내부의 정보를 text로 전환

- OCR은 광학 문자인식 기술로, 사진이나 이미지 내부의 글자를 인식해서 text 형태로 바꿔주는 딥러닝 기술

[그림 4-7] OCR의 예시



자료: 데이터메이커(2020. 10.)

가) 로컬환경

○ OCR 패키지 중 정확도가 높은 Tesseract를 사용함

- Tesseract는 다양한 운영체제에서 사용할 수 있는 광학 문자 인식 엔진임
- 아파치 라이선스 2.0에 따르는 무료 소프트웨어이며 2006년부터 구글이 개발을 후원하고 있음

나) 가상환경: EasyOCR

○ Tesseract는 로컬 환경에 설치해야 한다는 단점이 있기 때문에, 가상환경에서 실행하기 위해서는 다른 OCR로 대체해야 할 필요가 있음

○ EasyOCR을 대체 OCR로 사용할 수 있음

○ EasyOCR은 OCR 오픈소스로 Detection과 Recognition 부분으로 나누어져 있으며 Text Detection은 CRAFT를 사용하고 Recgnition은 CRNN 사용함

다) EasyOCR 문제점

○ Recognition 기능을 사용해야 하기 때문에, CPU를 사용하게 되면 속도가 현저히 느려진다는 문제가 발생하며, 또한 Tesseract보다 한글 인식에 취약한 편임

라) OCR 수행 방식

○ OCR을 사용해서 이미지를 텍스트로 변환한 후에는 앞선 키워드 검색 과정을 거쳐 해당 키워드가 이미지에 존재하는지 확인함

- 이미지는 HTML에 src 태그를 가진 모든 jpg,png를 수집함으로써 얻을 수 있음
- 그러나 이렇게 수집된 이미지의 양이 너무 방대하기 때문에, 모두 OCR을 수행하기에는 많은 시간이 소요됨
- 대부분의 사업자 관련 정보는 페이지 하단에 존재하므로, 페이지 하단 20개의 이미지만 OCR을 통해 검색해 사용함. 필요한 정보들은 주로 웹페이지 하단에 있음

3) 모델 최종 산출물

○ 사업자등록번호와 키워드가 포함되며, 해당 키워드를 통해 통신판매업자, 통신판매 중개업자 여부와 결제 시스템 존재 여부를 확인할 수 있음

○ 특이사항은 문제가 없음을 뜻하는 X, URL이 존재하지 않음을 뜻하는 URL 없음, 비정상적 URL이나 접속이 불가함을 뜻하는 오류, 해당 사업체가 통신판매 중개플랫폼에 입점해 있음을 알리는 입점형태로 구분함

○ 최종 산출물은 다음과 같은 엑셀 형태로 출력될 예정

[그림 4-8] 최종 산출물 예시

c_name	사업자등록번호	통신판매업	통신판매중개자	당사자	책에대한 책	입점	결제시스템	특이사항	homepage1
(주)위즈컴퍼니	6068632551	O	X	X	X	X	X	X	https://www.luxboy.com/
주식회사 에프에스에스	3728100976	O	X	X	X	X	X	X	https://www.fssuniverse.com/
(주)에이모	X	X	X	X	X	X	X	X	https://www.aimmo.ai/ko
농업회사법인유한회사 행복한	4118144624	O	X	X	X	X	O	X	https://happysweet.co.kr/
그린화훼영농조합법인	X	O	X	X	X	X	X	입점형태	https://smartstore.naver.com/gjgreei
쿠팡페이 주식회사	1208800767	O	O	X	O	O	O	X	https://www.coupang.com/
디퍼런트밀리언즈(주)	X	X	X	X	X	X	X	X	https://dmil.kr/
주식회사 로지올	X	X	X	X	X	X	X	X	https://logiall.com/
주식회사 하트잇	2278701833	O	X	X	X	X	X	X	https://heartit.kr
(주)아파트너	1048659942	O	X	X	X	X	X	X	https://www.aptner.com/
행복쇼핑 주식회사	1058789183	O	O	X	O	O	O	X	https://www.pping.kr/
(주) 미미박스 (memebox inc.	2118847525	O	X	X	X	X	O	X	https://www.memebox.com/
(주)바다코리아	1198688361	X	X	X	X	X	X	X	https://www.vadakorea.com/
주식회사 렌딧소셜대부	2588700070	X	X	X	X	X	X	X	https://www.lendit.co.kr/loan
네오프랜드(주)	1068613409	O	X	X	X	X	O	X	http://www.online24.co.kr/
DB손해보험	2018145593	O	X	X	O	X	X	X	https://www.idbins.com/FWMAIV000
(주) 유니정보	1208791771	O	X	X	O	X	O	X	https://www.uniinfo.co.kr/
유니온컴(주)	X	X	X	X	X	X	X	X	http://unioncom.kr/bbs/
(주)한얼기술정보이십일	X	X	X	X	X	X	X	X	http://hanol21.co.kr/

라. 모델 분석의 결과

1) 데이터 수집 결과

○ 키워드 검색 알고리즘 성능을 평가하기 위해 수집이 완료된 통신판매 사업자 1,350개 사업체를 대상으로 초기 성능평가를 수행함

○ 주요 키워드인 사업자등록번호 키워드와 통신판매업 키워드를 활용해 검색 알고리즘의 성능을 평가하였음. 이와 함께 '사업자등록번호 표기의 특성,' '중개플랫폼 입점형태,' '사업자 등록번호 미표기,' '사업자등록번호 미수집' 등 예외사항 처리에 대한 모델 결과를 확인 후, 그 해결 방안을 제시하였음

가) 사업자등록번호 표기의 특성

○ 일반적으로는 [그림 4-9]와 같이 통신판매업자는 사업자등록번호와 통신판매업 신고번호를 함께 기재

[그림 4-9] 통신판매사업자 표기 예시: 네이버 웹툰

자료: https://comic.naver.com/index

○ 이런 특성을 활용해 통신판매업자로 분류되었지만, 사업자등록번호가 없는 경우를 수집이 제대로 되지 않은 경우로 상정할 수 있음

○ 통신판매업자 중 사업자등록번호가 수집되지 않은 사업체의 개수와, 수집되지 않은 이유를 확인해 모델의 초기 성능을 판단함

○ 사업자등록번호가 수집되지 않았는데, 통신판매업자인 경우는 83개로 집계됨

○ 밑에서 서술한 중개플랫폼 입점 형태 특성과 사업자등록번호가 표기되어 있지 않은 경우를 제외하면 실제로 제대로 수집되지 않은 경우는 20개 사업체임. 또한 추후 사전 업데이트를 통해 이 오류를 대부분 수정하였음

나) 사업자등록번호 표기의 특성

○ 입점 형태의 경우에는 페이지 하단에 해당 통신판매사업자의 정보가 아닌 중개 플랫폼의 정보가 표기되어 있음

○ 입점 형태의 페이지에서 정보를 수집하게 되면 중개 플랫폼의 정보가 수집되게 됨

○ 그러므로 통신판매사업자 중 중개 플랫폼에 입점한 형태의 페이지에서는 사업자등록번호를 수집하지 않음

○ 이러한 중개 플랫폼 입점 형태의 이유로 사업자등록번호가 수집되지 않은 통신판매사업자의 개수는 51개 업체임

○ 입점 형태가 아니면서 사업자등록번호가 수집되지 않은 통신판매업자의 개수는 전체 83개 중 32개 업체로 확인됨

다) 사업자등록번호가 표기되어 있지 않은 경우

○ 통신판매업자가 사업자등록번호를 표기하는 것은 일반적인 경우지만 법적으로 의무화되어 있지 않음

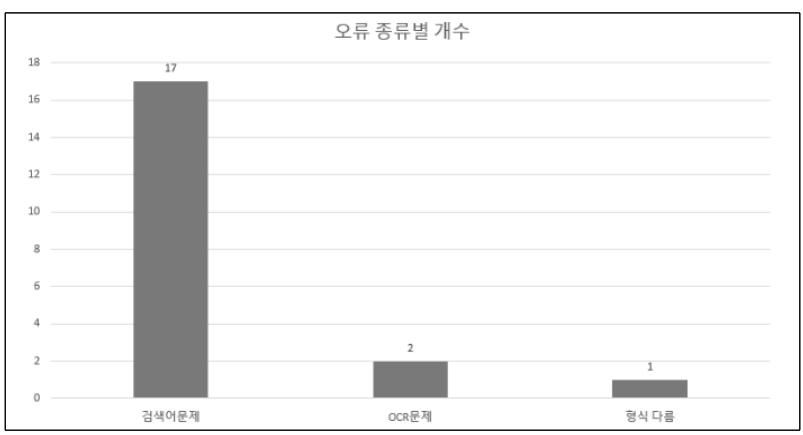
○ 입점 형태가 아니면서 사업자등록번호가 수집되지 않은 사업체 32개 중 12개 업체가 사업자등록번호를 표기하지 않는 것이 확인됨

○ 즉, 20개 업체에서 실제로 사업자등록번호가 표기되었지만, 본 연구의 검색 알고리즘을 통해 수집되지 않았음을 확인

라) 사업자등록번호가 표기되었지만 수집되지 않은 경우

○ [그림 4-10]은 사업자등록번호가 표기되었지만 수집되지 않은 경우 총 20개의 오류별 개수를 나타내고 있음

[그림 4-10] 사업자등록번호가 표기되었지만 수집되지 않은 경우별 개수



□ 검색어 문제

○ 검색어 문제의 경우 검색어 사전에 등록되어 있지 않아 사업자등록번호가 수집되지 않은 경우를 의미함

○ 검색어 문제는 검색어 사전에 사업자등록|\\d{10}과 같은 추가적인 검색어를 등록함으로써 오류를 해결했음

□ OCR 문제

○ OCR의 문제는 OCR 성능의 한계로 이미지로 표기되어 있는 사업체의 정보를 수집하지 못하는 경우를 의미함

○ OCR의 경우 성능 개선이 제한적임

- 현재 오픈소스 모델을 사용하고 있기 때문에 개선하기 위해선 이미지 데이터를 통해 파인 튜닝을 진행해야 함
- 파인 튜닝을 위해서는 정답이 라벨링된 데이터가 대량으로 필요함
- 이는 하나의 개별적인 프로젝트로 진행되어야 할 규모이기 때문에 현재 상황에서는 불가능함

○ 또한 OCR 문제는 극히 일부 사업체에서 발생하기 때문에 성능 개선이 필수적이지 않음

- 1,350개의 통신판매업자 중에 2개의 경우에 OCR 문제가 나타났음

□ 형식 다름

○ 형식 다름 오류의 경우 사업자등록번호의 일반적인 형식인 XXX-XXX-XX-XXXXX 형식이 아닌 XXXXXX-XXXXXXX 형식의 사업자 등록번호를 가지고 있는 경우를 뜻함

○ 검색어 문제와 마찬가지로 XXXXXX-XXXXXXX 형식을 검색어 사전에 등록하여 문제를 해결했음

마) 통신판매업자 분류에 대한 종합적인 판단

○ 1,350개의 통신판매업자 중 사업자등록번호가 제대로 되지 않은 경우는 총 20개임

○ 이 중 18개는 검색어 사전 업데이트를 통해 문제를 해결했음

○ OCR 문제는 현재 상황에서 해결이 불가능하지만 1,350개의 통신판매업자 중 단 2개로 극히 일부분을 차지하고 있음

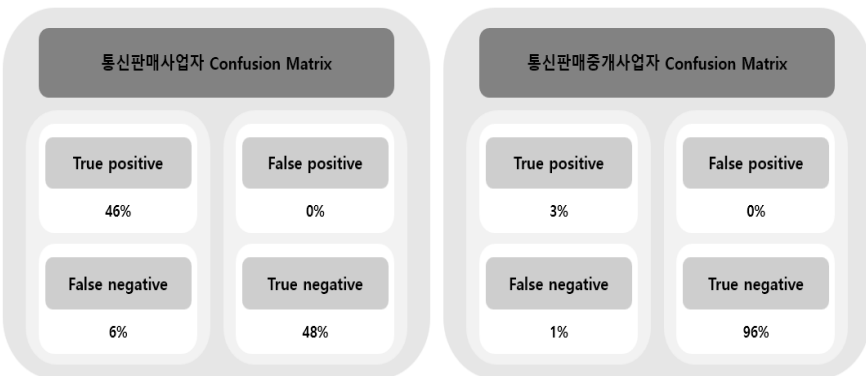
○ 종합적으로 통신판매업자 검색 알고리즘의 성능은 문제가 없다고 판단됨. 추후 검색어 사전을 계속 업데이트함으로써 추가적인 성능도 개선될

수 있다고 판단됨

2) 샘플링을 통한 전체 데이터 모델 결과 분석

○ 3,130개의 수집 결과 중 난수 생성을 통해 무작위로 뽑은 100개의 수집 결과에 대한 Confusion matrix를 생성했으며, 주요 컬럼의 Confusion matrix는 다음과 같음

[그림 4-11] 주요 컬럼 Confusion Matrix



○ 통신판매 사업자의 경우 94%의 정확도를, 통신판매 중개자의 경우 99%의 정확도를 보였음

3) 전체 데이터를 활용한 결과 오류 분석

○ 통신판매업, 통신판매중개자 두 키워드의 3,130개 전체 결과에 대한 오류를 각각 분석하였음

가) False Positive

○ False Positive 오류는 실제 통신판매업자(혹은 통신판매중개업자)가 아니나, 통신판매업자(혹은 통신판매중개업자)가 맞다고 오분류하는 경우를 말함

- 실태조사를 위한 대상 선정을 위해 자동 수집을 진행하는데 잘못된 대상이 선정될 수 있음

○ 통신판매사업자: 통신판매사업자로 분류된 1,350개 사업체 모두 확인한 결과 오분류는 없었음. 즉, 통신판매업자가 아닌데 통신판매업자로 분류된 건수는 0건으로 매우 높은 정확도를 보임

○ 통신판매 중개사업자: 반면, 통신판매 중개사업자로 분류된 206건에서는 다음과 같은 총 10건의 사업체에 대해 오분류가 발생함

○ 즉, 통신판매 중개사업자는 196개 사업체로 분류됨

[그림 4-12] 통신판매 중개사업자 오분류

c_name	사업자등록번호	통신판매업	통신판매중개자	당사자	거래에대한 책임	입점	결제시스템	특이사항
에이티솔루션즈	1078171649	X	O	X	X	X	X	X
㈜엔투비	X	X	O	X	X	X	X	X
(주)아톤	1078171649	X	O	X	X	X	X	X
Zoom	X	X	O	X	X	X	X	X
(주)해시스크래퍼	4458801059	X	O	X	O	X	X	X
(주)이지스	5038149156	X	O	X	X	X	X	X
(주)인터파크비즈마켓	2148880576	X	O	X	O	O	X	X
한국아이디지	2148722467	X	O	X	O	X	X	X
(주)플러스앤	2208178800	X	O	X	X	X	X	X
인터넷우체국	2068302050	X	O	X	X	X	X	X

○ [그림 4-12]는 통신판매 중개사업자가 아닌데 통신판매 중개사업자로 오분류된 사업체를 모아놓은 표임

- 통신판매 중개자로 오분류된 경우에는 통신판매업자로 분류되지 않았음

○ 따라서 통신판매 중개사업자의 경우 통신판매업 키워드가 수집되지 않았다면 통신판매업자 혹은 중개업자가 아닌 것으로 판단할 수 있음

나) False Negative

○ False Negative 오류는 해당 업태(통신판매사업자, 통신판매중개자)가 맞지만, 해당 업태가 아니라고 오분류된 경우를 의미함

○ 모델 평가를 위해 통신판매 사업자가 아니라고 분류된 사업체, 통신판매 중개사업자가 아니라고 분류된 사업체 100개씩을 샘플링해 확인했음
○ 통신판매사업자가 맞지만 아니라고 오분류된 사업체는 100건 중 7건임
- 반면 통신판매 중개사업자가 맞지만 아니라고 오분류된 사업체는 100건 중 12건임
○ 통신판매사업자: 통신판매 사업자는 통신판매 신고번호를 메인 페이지에 (혹은 첫 페이지) 게시할 의무가 없음. 물론 통신판매사업자는 통신판매 신고번호를 자사 홈페이지에 게시해야 함. 그러나 메인 페이지(혹은 첫 페이지)에 없는 경우가 100건 중 6건이 있었음. 현재 분류 알고리즘은 메인 페이지(혹은 첫 페이지)만 검색하여 찾아내는 방식임. 따라서 현재 분류 알고리즘으로는 이 오류를 해결하기가 쉽지 않음. 추후 인공지능 알고리즘 등을 활용해 이런 문제를 해결할 수는 있음
○ 통신판매 중개사업자: 통신판매 중개사업자의 경우 사업체가 통신판매 중개사업자임을 나타내는 문구를 게시해야 함(공정거래위원회)
○ 그러나 아래 서술할 사업체 분류군은 중개사업자를 나타내는 문구를 메인 페이지에 게시하지 않고 있음(100건 중 12건)

4) 보조 키워드를 활용한 통신판매 중개자 분류 고도화

○ 통신판매중개자가 아닌데 통신판매중개자로 오분류된 경우(False Positive)와 통신판매중개자인데 통신판매중개자가 아닌 것으로 오분류된 경우를 (False Negative) 보완하기 위해 보조 키워드를 활용하기로 결정함

가) 통신판매중개자 보조 키워드 활용

○ 보조 키워드는 '통신판매 중개자' 키워드가 감지하지 못한 통신판매중개 업체를 감지하거나 통신판매중개자로 오분류된 업체를 정정하기 위해서 사용되는 키워드임
- 통신판매 중개사업자의 경우 하단에 본인이 통신판매 중개사업자임을

표시하는 것이 의무화되어 있음

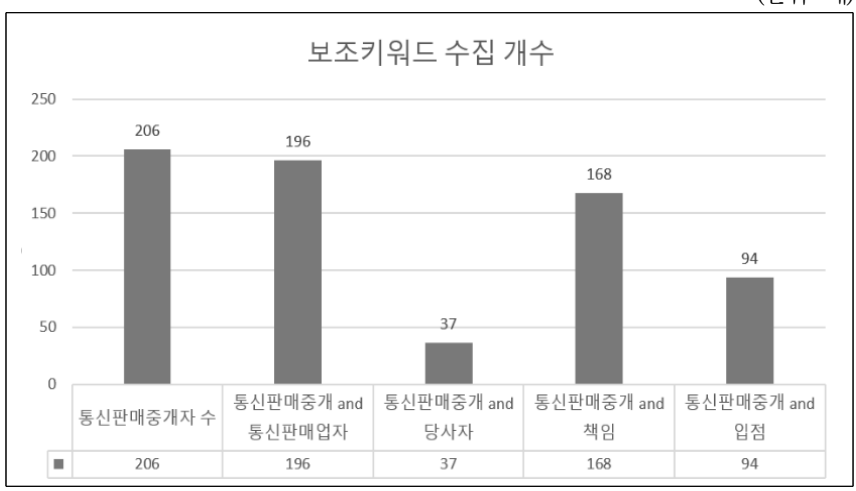
- 그러나 외국 사이트나 통신판매업을 겸업하는 통신판매 중개사업자의 경우 이를 게시하지 않는 경우가 존재함
- 이 경우를 감지하기 위해 3가지 보조 키워드와 그 조합들을 사용해 통신판매 중개사업자 분류 성능을 보완함

○ 통신판매 중개자 분류 성능을 보다 고도화하기 위해 보조 키워드를 추가로 활용

- 이때 사용한 보조 키워드는 당사자, 거래에 대한 책임, 입점 등을 사용

[그림 4-13] 보조키워드 수집 개수

(단위: 개)



○ [그림 4-13]은 통신판매중개자로 분류된 사업체의 홈페이지에서 각 보조 키워드가 검색된 경우의 개수를 나타냄

- 통신판매중개 and 책임이 168개라면 168개의 사업체에서 통신판매중개라는 키워드와 책임 키워드가 동시에 수집되었다는 의미임

○ 통신판매중개자이면서 통신판매업자인 경우는 196개로 확인됨. 이에 관련해서는 False positive에서 상세히 서술함

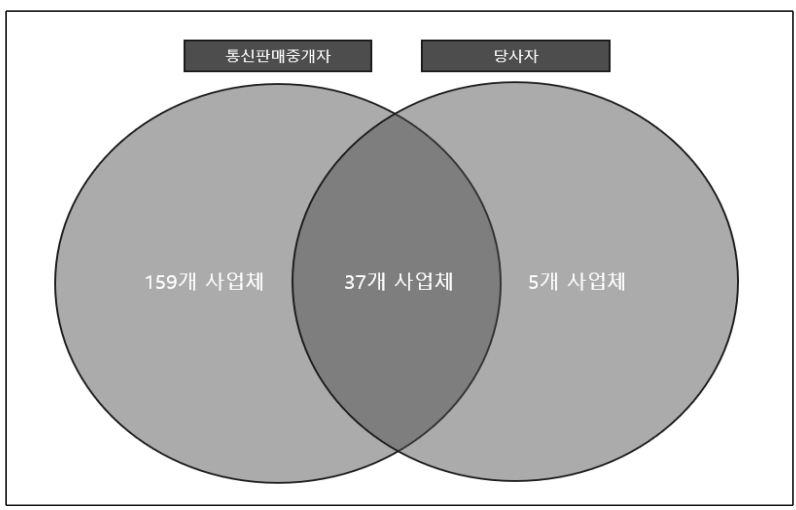
* ‘통신판매 중개’ and ‘당사자’

○ 통신판매 중개업자 196개 중 ‘통신판매 중개자 and 당사자’ 키워드가 수집된 사업체는 37개로 확인됨

○ 전체 3,130개 사업체 중 ‘당사자’가 수집된 사업체는 42개이며, ‘통신판매중개자’ 키워드가 검색되지 않았으면서 ‘당사자’가 검색된 사업체는 모두 5개였고, 직접 확인 결과 5개 모두 통신판매중개자가 아니었음

○ 따라서, 통신판매중개자가 아닌데 통신판매중개자로 오분류된 경우(False Positive)와 통신판매중개자인데 통신판매중개자가 아닌 것으로 오분류된 경우를(False Negative) 정정하기 위한 보조 키워드로서 적절하지 않아 보임

[그림 4-14] 통신판매 중개 키워드와 당사자 키워드 벤다이어그램



* ‘통신판매 중개’ and ‘책임’

○ 통신판매 중개사업자 196개 중 ‘통신판매 중개’ and ‘책임’ 키워드가 수집된 사업체는 168개로 확인됨

○ 전체 3,130개 사업체 중 ‘책임’이 수집된 사업체는 405개이며, 통신판매

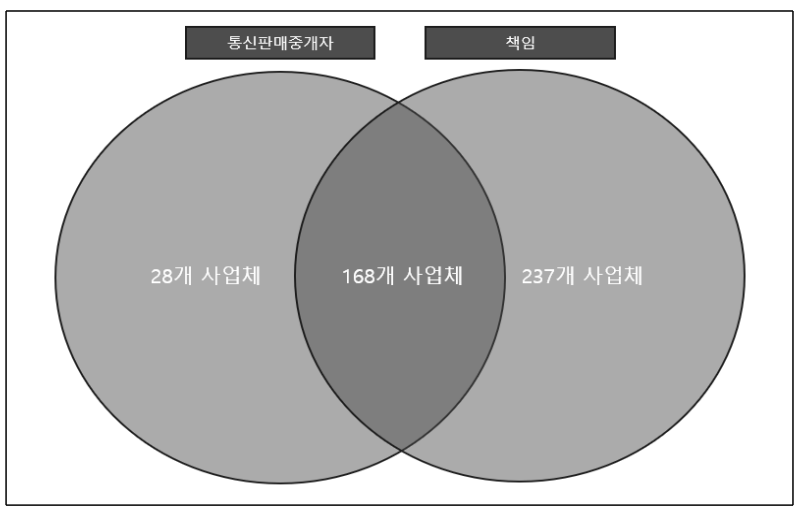
중개자 키워드가 수집되지 않았으면서 ‘책임’ 키워드만이 수집된 사업체는 237개임

○ ‘책임’ 키워드만이 수집된 사업체 237개의 경우 금융업종사업체나 사업체의 홈페이지에 개인정보, 청소년보호 등의 책임자를 기입해 놓은 경우였으며 통신판매 중개업체는 존재하지 않았음

○ 그러나 206개의 통신판매 중개업체([그림 4-15] 참고) 중에 168개의 업체에서 ‘책임’ 키워드가 함께 검색되었음

○ 그러므로 통신판매 중개자가 아닌데 통신판매 중개자로 오분류된 경우(False Positive)를 보완하는 보조 키워드로서 사용 가능하다고 판단됨

[그림 4-15] 통신판매 중개 키워드와 책임 키워드 벤다이어그램



* ‘통신판매 중개’ and ‘책임’

○ 통신판매 중개사업자 196개 중 ‘통신판매중개’ and ‘입점’ 키워드로 검색된 사업체는 94개로 확인됨

○ 전체 3,130개 사업체 중 ‘입점’으로 검색된 업체는 295개이며, ‘통신판매 중개자’ 키워드가 검색되지 않았으면서 ‘입점’이 검색된 사업체는 201개임

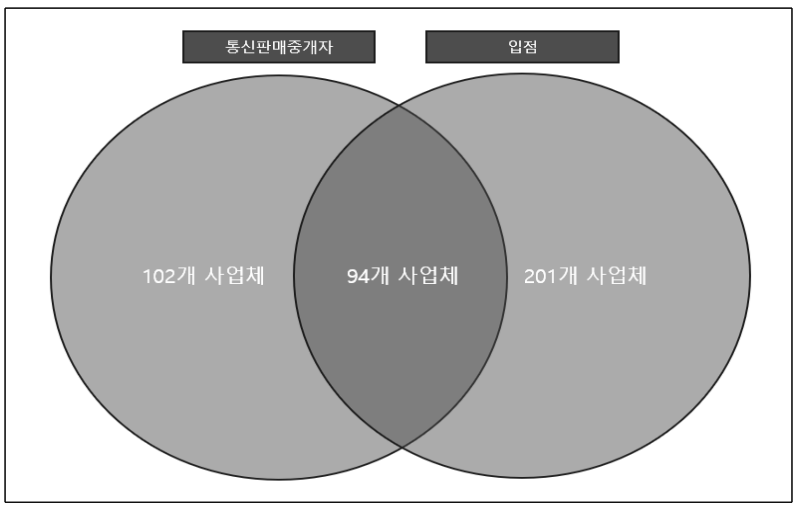
○ 201개의 업체 확인 결과, 대부분이 통신판매 중개자 형태를 띠고 있음이 확인됨

○ 통신판매 중개자에 검색되지 않은 통신판매 중개업체를 찾아(False Negative) 보완하기 위한 보조 키워드로서 유용하게 사용할 수 있다고 판단됨

○ 그러나 입점 문의를 받는 경우에, 업체가 물건을 구매해 파는지, 구매하지 않고 마켓 플레이스만을 제공하는지에 대한 여부는 알 수 없음

- 통신판매자이면서 입점 키워드가 존재하는 경우에는 이용약관을 통해 판매자에게 책임이 존재하는지 추가적인 판단을 통해 통신판매 중개자 여부를 확인해야 함

[그림 4-16] 통신판매 중개 키워드와 입점 키워드 벤다이어그램



3. 소 결

가. 웹사이트 분류 자동화 프로그램

○ 현재 완성된 모델은 웹사이트에 특정 키워드를 검색해 이를 기반으로 업체 분류에 도움을 주는 모델임

○ 현재 모델은 완전한 자동화가 아닌, 업태 분류에 보조적인 도움을 줄 수 있는 한정적인 자동화 모델임

- 즉, 특정 키워드가 웹사이트에 존재한다는 결과를 가지고 해당 페이지를 분류하기에는 변수가 굉장히 많음
- 더 자세하고 정확한 분류를 위해서는 더욱 많은 정보를 모델에 입력할 필요가 있음

○ 현재 모델은 이후 완전 자동화 프로그램을 만드는 데 도움을 줄 수 있는 보조적 모델로 적용할 수 있음

- 웹사이트에 특정 키워드가 존재한다는 사실은 해당 웹사이트를 분류하는 데 큰 힌트가 될 수 있음
- 이를 딥러닝 모델에 input으로 사용하여 정확도를 높일 수 있음
- 현재 모델은 간단한 알고리즘을 사용하기 때문에, 추후 다른 모델을 업데이트해서 사용해도 비교적 간단한 코드의 수정으로 적용할 수 있다는 장점이 있음

나. 보조 키워드

○ 각 보조 키워드의 성능을 분석해 확인할 수 있었음

- 입점과 같이 유용한 키워드도 존재했던 반면, 당사자와 같이 효용이 없는 키워드도 존재함

○ 현재 사용하는 보조 키워드 외에도 더 나은 보조 키워드를 찾아 모델의 성능을 높일 수 있음

제5장 결 론

1. 주요 내용

○ 본 연구에서는 디지털 전환 시대의 디지털산업 현황을 파악하고자 디지털산업에 대한 통계 및 데이터의 현황, 그리고 디지털산업의 현황을 보다 효율적으로 파악하기 위한 방법론을 검토함

- 구체적으로, 제2장에서는 OECD를 중심으로 논의 중인 디지털산업의 정의와 분류체계를 살펴봤고, 이 과정에서 OECD가 제시한 디지털경제 측정을 위한 개념모델을 살펴봄
- 제3장에서는 미국, 캐나다, 호주 등 해외 주요국의 디지털산업 통계 현황에 대해 살펴봄
- 그리고 한국의 디지털산업 현황을 최근 공표된 디지털산업 실태조사의 항목을 중심으로 소개함
- 제4장에서는 디지털산업에 속한 사업자 중 디지털플랫폼 제공 업체를 식별하기 위한 분류 자동화 방안에 대한 연구 결과를 제시

2. 디지털산업 관련 통계 및 데이터 개선을 위한 시사점

□ 디지털산업 통계 조사 확대

○ 각국은 디지털산업 관련 통계를 개발, 시험추계, 조사 중인 것으로 나타남

○ 한국도 2023년 과학기술정보통신부와 정보통신정책연구원에서 2022년 기준 디지털산업 실태조사를 국가통계로 승인받고, 2024년 6월 조사결과를 공표했으며, 한국은행에서도 2019년도 기준 디지털공급사용표를 시험추계하여 그 결과를 연구결과로 제시

□ 디지털산업 분류체계 개선 방향

○ 디지털산업 실태조사의 경우 OECD의 디지털 산업 분류체계를 준용하여 분류체계를 작성하였으나, 일부 부문에서 차이가 존재

- 한국은 OECD와 달리 방송서비스업을 ICT 산업으로 구분하고 있는데, 디지털산업 실태조사의 분류 역시 국가별 환경과 보유한 정보 및 데이터의 차이를 반영한 결과임

○ 통계청은 디지털산업 분류체계를 통계청의 경제 특수분류로 제정하는 것에 대한 수요 제기를 검토 요청

- 디지털산업의 확산과 중요성을 염두에 둘 때 주무부처인 과학기술정보통신부와 관련 기관인 정보통신정책연구원에서 디지털산업 특수분류를 제정할 필요성이 존재함
- 디지털산업 실태조사를 국가통계로 승인받은 만큼 경제 특수분류로 운영할 필요성과 실익이 있을 것으로 판단

○ 디지털산업의 분류체계 관점에서는 조사 가능 여부와 별개로 산업의 소분류 구분이 필요

- 디지털기반산업의 경우, 중분류 수준에서 101. 디지털기반 기기·부품, 102. 디지털기반 서비스로 구분하고 있으나, 소분류 수준에서 보다 세분화된 산업분류가 가능할 것으로 판단
- 캐나다 통계청은 디지털기반산업을 하드웨어, 소프트웨어, 통신서비스, 그 외 서비스로 구분 중이며, 한국은 반도체, 디스플레이 등 부품 분야의 경쟁력이 높은 점을 고려할 때 하드웨어를 기기와 부품으로 세분화가 가능할 것으로 보임
- 디지털플랫폼 제공산업의 경우 201. 데이터 및 광고 기반 디지털 플랫폼, 202. 디지털 중개 플랫폼, 203. 디지털 영상음향 정보 플랫폼으로 구분하고 있는데, 202. 디지털 중개 플랫폼의 경우 주요 영역별로 구분하는 것도 정보 제공 관점에서 검토할 필요가 있음
- 203. 디지털 영상음향 정보 플랫폼에서는 디지털 콘텐츠의 유형에 따라 세

분할 수 있을 것으로 보이며, J. 정보통신업에서 생산 중인 방송, 음악, 웹툰 등의 디지털 콘텐츠와 M. 전문과학기술서비스업, P. 교육서비스 등의 디지털 콘텐츠로 구분하는 등의 분류가 가능할 것으로 보임

- 디지털 플랫폼 활용(의존)산업의 경우, 301. 실물 상품 기반 플랫폼 활용(의존) 산업과 302. 디지털 상품 기반 플랫폼 활용(의존) 산업으로 구분하고 있는데, 디지털 중개플랫폼 활용과 관련한 산업영역인 도소매, 음식, 숙박, 교통, 운송, 배달 등의 영역으로 세분화하는 것이 필요
- 마지막으로 디지털관련산업의 경우, 현재 401. 디지털 도소매, 402. 디지털 금융, 403. 그 외 디지털 관련 산업으로 구분되어 있는데, 중분류 이하 소분류를 세분할 실익이 있는가에 대한 검토 필요

□ 디지털 경제 진전과 산업구조의 변화

○ 디지털 전환으로 개별 기업뿐만 아니라 경제 내에서 산업구조까지 변화하고 있음

○ 앞서 살펴본 바와 같이 디지털산업과 관련한 새로운 분류체계가 요구되며, 이들 분류체계는 기존의 통계 분류를 활용하여 연계하기에 한계가 존재

○ 디지털산업 현황 파악을 위해서는 새로운 통계조사가 필요하며, 개별 산업에 대한 판별조사 역시 필요함

○ 고비용의 판별조사 수요를 최소화하기 위해 웹사이트의 특성을 반영하여 디지털산업의 기초 정보를 파악하는 것이 중요하며, 본 연구에서는 웹사이트 크롤링을 통해 키워드 분석을 수행

○ 향후 인터넷과 모바일을 통한 정보 분석, 그리고 행정자료의 적극적인 활용 등을 통해 디지털 경제가 진전함에 따라 변화하는 산업구조를 파악할 필요가 있음

□ 디지털산업 내 다양한 사업체 현황 파악 연구 필요

○ 디지털화로 인한 거래방식의 변화는 산업구조의 변화까지 이어짐

○ 디지털 기술의 발전은 거래비용의 감소로 이어져, 기존의 도소매업자들이

디지털중개플랫폼을 이용한다거나, 독자적인 웹사이트를 구축하여 전자상거래를 하는 등의 변화가 일상화됨

- 이에 디지털산업 내 다양한 사업체들을 기존의 산업분류로 구분하기 곤란한 상황이 되었고, 디지털경제를 파악하기 위해 디지털산업 분류와 같은 별도의 분류체계가 필요하게 됨
- 그중 대표적인 영역이 인터넷, 모바일 등 온라인 인터페이스를 이용해 상품을 판매하거나 중개하는 사업자들이 활동하는 산업임

○ 통계조사 관점에서 디지털 경제에서 변화한 사업자들은 다양한 산업군에 혼재되어 분포하여 표준산업분류를 이용한 분류도 쉽지 않은 상황

- 특히, 특정 산업에 포함되어 있더라도 하나의 분류 안에서도 상이한 성격을 지닌 업체가 혼재하며, 이들을 식별하기 위한 판별조사가 필요

○ 본 연구에서는 웹사이트 크롤링을 통한 키워드 분석을 수행함. 특히 통신판매사업자를 대상으로 키워드 분석이 가능한가에 대한 연구를 수행

○ 통신판매사업자 중 입점을 받는 경우에, 물건을 구입해 판매하는 경우와 플랫폼만을 제공하는 경우를 구분할 수 없기 때문에 향후 딥러닝 모델을 이용해 거래 방식의 차이를 구분하는 방법을 검토할 필요가 있음

- 예를 들어 웹사이트 내에 결제 시스템의 존재 여부를 파악하는 것은 해당 웹사이트의 특성을 파악하는 데 중요하며, 결제 시스템 운영 여부를 파악하기 위해서는 키워드보다 HTML 자체를 이해할 수 있는 딥러닝 모델의 구축이 추후 필요할 것으로 보임

참 고 문 헌

[국내 문헌]

과학기술정보통신부·KISDI(2024), 「2023 디지털산업 실태조사」.

통계청(2014), 2013년 연간 및 4분기 전자상거래 및 사이버쇼핑동향 보도자료.

신진욱, (2023. 6. 1.), "통신판매업, 2년 연속 100대 생활업종 사업자수 1위", 《한국경영자신문》.

정현준·김성옥·이학기·정부연(2021), 『지역내총생산(GRDP)에 디지털 경제 반영을 위한 연구』, 정보통신정책연구원; 통계청.

[해외문헌]

Australian Bureau of Statistics(2023). "Digital activity in the Australian economy", 2021-22 27/10/2023.

ADB(2021). "Capturing the Digital Economy: A Proposed Measurement Framework and it's Applications – Special Supplement of Key Indicators for Asia and the Pacific 2021".

BEA(2023). "U.S. Digital Economy: New and Revised Estimates, 2017-2022", *Survey of Current Business*, December 6, 2023.

Bukht R. and Heeks R.(2017). "Defining, conceptualising and measuring the digital economy". *GDI DevelopmentInformatics Working Papers*, no. 68. University of Manchester, Manchester.

IMF(2018). "Measuring the Digital Economy".

OECD(2020a). "A roadmap toward a common framework for measuring the digital economy", Report for the G20 DigitalEconomy Task

Force, SAUDI ARABIA, 2020.

_____(2020b), "Guidelines for Supply-Use tables for the Digital Economy", Working Party on National Accounts, SDD/CSSP/WPNA (2019)1/REV1.

_____(2023). *OECD Handbook on Compiling Digital Supply and Use Table*", OECD Publishing, Paris.

Statistics Canada(2023), "Digital supply and use tables, 2017 to 2020", Released at 8:30 a.m. *Eastern time in The Daily*, Tuesday, July 25, 2023.

UNCTAD(2019), "Digital economy report 2019," United Nations Conference on Trade and Development, 2019

[온라인자료]

"How Does Selenium Automation Testing Work?", Infinigeek(2018), 2024. 7월, https://infinigeek.com/selenium-automation-testing-work/.

"Digital Economy", Bea(2023), 2024. 7월, https://www.bea.gov/data/special-topics/digital-economy.

Abstract

■ Title

Analysis of the current status of digital industries in the era of digital transformation

■ Purpose of Research

This study examines the changes in the domestic industrial structure due to the accelerated digital transformation. It analyzes the current state of the digital industry, reflecting both the supply and demand aspects of digital technologies, in response to the global spread of the digital economy. By reviewing and analyzing the current state of the digital industry in the era of digital transformation, the study seeks to explore the direction of these changes. The study derives implications for improving statistics and data related to the digital industry, based on an analysis of the current state of the digital industry, industrial structure, and digital platform operators, providing a foundation to support digital industry policies.

■ Main Outcomes of Research and Policy Implications

This study aims to understand the current state of the digital industry

in the era of digital transformation by reviewing the status of statistics and data related to the digital industry and examining methodologies for more effectively capturing the industry's status. Chapter 2 examines the OECD's ongoing discussions on defining and classifying the digital industry, along with their model for measuring the digital economy. Chapter 3 reviews digital industry statistics from countries like the US, Canada, and Australia, and introduces the current state of Korea's digital industry based on recent survey data. Chapter 4 discusses research on automated methods for classifying digital platform providers within the digital industry. Efforts are being made to expand digital industry statistics surveys. the study addresses directions for improving the digital industry classification system and conducts keyword analysis through website crawling. This study reviews and analyzes the definition, classification, statistics, and data of the digital industry to understand its current status and improve comprehension. This study also analyzes and reviews the digital industry's definition, classification, and data to enhance understanding and support evidence-based policy development.

● 저 자 소 개 ●

정 현 준

- 고려대 경제학과 졸업
- 고려대 경제학과 석사
- 고려대 경제학과 박사
- 현 정보통신정책연구원 연구위원

김 민 진

- 연세대 경영학과 졸업
- KAIST 경영학 석사
- 현 정보통신정책연구원 부연구위원

박 소 연

- 고려대 산업공학과 졸업
- 서울대 경영학과 석사
- 현 정보통신정책연구원 전문연구원

박 재 홍

- 경희대 경영학과 졸업
- Stanford University 통계학 석사
- University of Texas at Austin 경영학 박사
- 현 경희대 경영학과 교수

현안연구 24-01

디지털 전환 시대의 디지털 산업 현황 분석

2024년 8월 31일 인쇄
2024년 8월 31일 발행

발행인 배 경 율
발행처 정 보 통 신 정 책 연 구 원
충청북도 진천군 덕산읍 정통로 18
TEL: 043-531-4114 FAX: 043-535-4695～6
인 쇄 인 성 문 화
ISBN 979-11-7000-389-2 93320

〈비매품〉