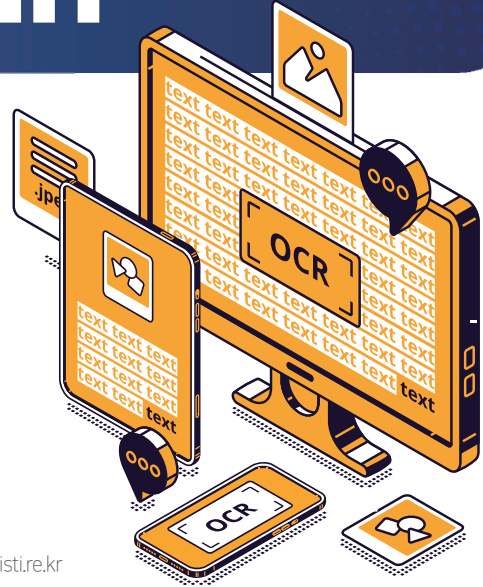


ASTI MARKET INSIGHT

광학 문자인식



데이터분석본부 대구경북지원 선임연구원 노대경 Tel: 053-601-5171 e-mail: ndk@kisti.re.kr

KEY FINDING

1. 광학문자인식은 데이터 구축 기반 기술로 인공지능 및 빅데이터 관련 산업과 함께 성장하며 응용 분야가 매우 다양한 산업적 특징을 갖고 있다.
2. 광학문자인식의 상위시장인 '자동인식 및 데이터 취득' 세계시장은 2020년 기준 401억 달러에서 연평균 14.9 % 성장해 2025년 803억 달러의 규모가 될 것으로 전망된다.
3. 광학문자인식의 세계시장은 2020년 1,330백만 달러 규모에서 연평균 15.4 % 성장해 2025년 2,720백만 달러의 규모가 될 것으로 전망되며, 특히 광학문자인식기술 응용 서비스 시장은 2020년 618백만 달러에서 2025년 1,385백만 달러의 규모로, 연평균 17.5 %의 고성장이 예상된다. 국내 시장은 2020년 110백만 달러에서 연평균 15.4 % 성장해 2025년에는 225백만 달러의 규모가 될 것으로 전망된다.
4. 2020년12월10일 전자문서 및 전자거래 기본법 개정안이 시행되어 디지털문서가 원본 종이문서와 동일한 법적 효력이 인정되고 있으며, 이로 인한 문서의 디지털화에 대한 사회적 요구가 증가해 광학문자인식 기술의 적용분야 확대가 기대되고 있다.
5. 광학문자인식 시장은 경쟁이 과열되고 있는 시장으로 경쟁력을 갖추기 위해서는 다른 분야와의 융합을 통한 특화된 기술개발 전략(회계장부 디지털화 → 재무 분석 시스템 연계 등)이 필요하다.

1) 시장의 개요

최근 인공지능(AI)과 빅데이터 관련 산업의 비약적 발전으로 디지털화가 되지 않은 문서들의 디지털화에 대한 사회적 요구가 증가하

면서 광학문자인식(Optical Character Recognition, OCR) 기술이 다시금 주목받고 있다. OCR이란 다양한 필체를 이미지 스캐너 등의 장비를 통해 컴퓨터가 편집할 수 있는 문자코드로 변환하는 기술을 말한다. 과거에는 렌즈와 같은 광학장비를 이용한 광학문자인식과

이미지 스캐너를 이용한 디지털문자인식은 다른 기술로 분류되었으나, 현재는 광학문자인식으로 언어적 개념이 통일되었다. 광학문자인식은 1928년에 패턴매칭(Pattern Matching)기법을 이용한 문자 인식 방법이 특허로 등록되면서 활성화되었다고 알려져 있다. 최근 플랫폼 대기업에서도 광학문자인식 서비스를 오픈하는 등 점점 해당 시장에 대한 관심이 증가하는 추세이다.

광학문자인식 시장은 자동인식 및 데이터취득(Automatic Identification and Data Capture, AIDC) 시장의 하위시장으로,

AIDC는 객체, 사운드, 이미지 또는 개인으로부터 데이터를 수집해 여러 종류의 자산, 문서, 재고, 납품 및 보안 등을 관리하는 목적을 가진다. 이는 AI와 함께 성장하는 산업으로서 응용분야가 매우 다양한 산업적 특징을 갖는다. 최근들어 금융기관에서 광학문자인식 솔루션 채택이 증가하고 있고, 아울러 일반기업과 공공기관도 디지털 데이터 변환의 수요가 증가하는 추세를 보이면서 OCR 시장의 확대에 탄력을 받고 있는 상황이다.

표 1 광학문자인식 솔루션의 응용범위

종류	종류별 특징
신분증인식	• 신분증 이미지에서 원하는 필드의 정보를 인식 • 주민등록증/운전면허증/여권 등 다양한 형식 인식 가능
명함인식	• 명함 이미지를 정확히 인식해 활용 • 자동 명함 영역 탐지, 양면 촬영 등 편의기능
신용카드인식	• 신용카드 이미지에서 필요한 필드의 정보를 인식 • 다양한 형태의 신용카드 이미지 처리 가능
서식인식	• 서식이 있는 문서를 촬영해 필요한 필드의 정보를 인식 • 새로운 서식과 인식할 필드를 등록 가능
정보마스킹	• 민감한 정보를 마스킹 처리 • 특정 양식 및 필드를 지정해 마스킹 처리 가능
필기체인식	• 손으로 작성된 날자 필기체를 인식 • 지정된 서식에 쓴 글씨를 카메라로 인식해 정보 추출
문서인식	• 촬영한 문서의 텍스트 정보를 인식 • 인식된 텍스트 정보를 활용할 수 있도록 다양한 형태로 제공
문서복원	• 촬영한 이미지를 오피스 파일로 복원 • 촬영한 문서 이미지의 구조를 그대로 복원, 편집 가능한 파일로 제공
실시간 단어 뜻/번역	• 카메라로 단어를 인식하여 사전/번역과 연동된 결과를 제공 • 카메라를 비추는 것만으로 단어 인식 가능

출처: <https://www.selvasai.com/>

2) 국내 법률 및 정책

과학기술정보통신부와 법무부가 적극적으로 추진해 온 '전자문서 및 전자거래 기본법 개정안'이 2020년12월10일 시행되면서 디지털 문서가 원본 종이문서와 동일한 법적 효력을 갖게 되었다. 이에 광학문자인식기술을 통해 문서의 디지털변환을 필요로 하는 곳이 많아졌다. 예를 들어 환자의 차트 등을 종이문서로 오랫동안 보관해온 병의

원, 회계자료 등을 디지털화하려는 기업, 대한민국 근현대사를 기록하는 대통령기록관, 수사기록과 같은 중요한 종이문서를 많이 보유한 검찰 등이 대표적인 광학문자인식기술의 수요처라고 할 수 있다. 해당 법에서는 디지털문서의 위변조를 방지하기 위하여 종이문서의 스캔 본을 공인전자문서센터에 보관하고, 종이문서를 폐기하는 내용을 포함한다. 과거 스캔한 문서는 그림파일로 인식되었으나, 광학문자인식을 통해 텍스트를 디지털화해 컴퓨터에 저장이 가능해지므로

추후 문서 내용의 분석과 같은 후속작업이 용이해졌다. 광학문자인식기술은 다양한 기업 및 기관의 디지털전환을 가속화시킬 기술이며, 더 나아가 거시적 관점으로는 국가 데이터 구축의 기반기술로 볼 수 있다.

광학문자인식기술은 데이터 축적과 가장 밀접하며, 정책적으로는 AI 및 데이터 관련 육성정책과 관련이 있다. 현재 관계부처들이 합동으로 공표한 계획과 전략은 2019년에 발표된 2023년까지의 계획인

'데이터·시경제 활성화 계획'과 2021년 발표한 '신뢰할 수 있는 인공지능 실현 전략'이 대표적인데, 이들 계획 및 전략은 데이터의 신뢰성 확보를 중요시한다. 광학문자인식은 현재 기술적으로 한글분야에서 99% 이상의 인식률을 달성하였고, 필기습관에 따라서 구분이 어려울 수 있는 'ㄹ'과 'ㄷ', 'ㄴ'과 'ㅇ'과 같은 경우도 AI를 이용해 문맥에 맞는 것을 정확하게 선택할 수 있는 수준이 되었다.

표 2 전자문서 제도 주요 제·개정 현황

시기	핵심 내용
1999.07 제정	• 전자문서의 효력 및 보관 등에 대한 [전자거래기본법] 제정
2005.03 개정	• 공인전자문서보관소 지정제도 도입
2007.05 개정	• 전자화문서 효력 및 인증제도 도입
2012.06 개정	• 전자화문서 효력 및 인증제도 도입 • 공인전자주소 제도 및 공인전자문서증거자 지정 제도 도입 • 공인전자문서보관소 명칭 변경(공인전자문서보관소 → 공인전자문서센터)
2018.04 개정	• 전자거래 사업자에 대한 인증제도 폐지
2020.12 개정	• 공인전자문서증거자 지정제도 인증제도로 전환 • 전자화제도 관련 종이문서 폐기 규정 신설

출처: 전자문서통합지원센터

3) 시장동향

1) 시장 규모

AIDC 시장은 광학문자인식, RFID 등 세분화된 다양한 시장을 포함하고 있는 상위시장이다. AIDC의 세계시장은 2020년 401억 달러에서 연평균 14.9% 성장해 2025년 803억 달러가 될 것으로 전망되고 있다.

광학문자인식은 <표 3>에서 알 수 있듯이 AIDC의 하위 시장으로서 시장의 플레이어들이 소비자에게 제공하는 형태에 따라 소프트웨어와 서비스로 구분이 가능하다. 광학문자인식 소프트웨어 세계 시장은 2020년 713백만 달러에서 연평균 13.4% 성장해 2025년에는 1,335백만 달러가 될 것으로 전망된다. 광학문자인식 응용 서비스 관련 세계 시장은 2020년 618백만 달러에서 연평균 17.5% 성장해 2025년 1,385백만 달러의 규모가 될 것으로 전망된다. 서비스

시장은 소프트웨어 시장에 비해 고성장이 예상되며, 2024년부터는 서비스 시장이 소프트웨어 시장의 규모를 넘어설 것으로 예상된다.



표 3 AIDC의 세계 시장

(단위: USD BILLION)

제품구분	2020	2021	2022	2023	2024	2025	CAGR(%)
스마트카드 (Smart Cards)	14.57	18.01	21.63	25.42	29.38	33.51	18.1
RFID 시스템 (RFID Systems)	8.31	9.53	10.59	11.5	12.27	12.88	9.2
바코드 솔루션 (Barcoding Solutions)	6.36	7.13	7.74	8.18	8.45	8.56	6.1
생체인식 시스템 (Biometric Systems)	5.02	6.4	7.89	9.5	11.24	13.1	21.1
헤드업 디스플레이 (Heads-up Displays)	2.03	2.7	3.45	4.28	5.21	6.21	25.1
광학문자인식 (OCR Systems)	1.33	1.6	1.88	2.15	2.44	2.72	15.4
마그네틱 카드 (Magnetic Stripe Cards)	1.19	1.18	1.1	0.93	0.69	0.36	-21.1
음성안내 WMS (Voice-directed WMS)	0.79	0.96	1.14	1.33	1.52	1.71	16.7
웨어러블 (Wearables)	0.38	0.48	0.58	0.69	0.8	0.92	19.0
VR 솔루션 (VR Solutions)	0.07	0.1	0.13	0.16	0.19	0.23	25.5
구글 글라스 (Google Glass)	0.05	0.05	0.08	0.09	0.1	0.11	14.8
계	40.1	48.2	56.2	64.2	72.3	80.3	14.9

자료: "Automatic Identification and Data Capture Market - Global Forecast to 2025", MarketsandMarkets, 2020.08

표 4 광학문자인식 소프트웨어 및 서비스의 세계 시장

(단위: USD MILLION)

형태	2020	2021	2022	2023	2024	2025	CAGR(%)
소프트웨어	713	841	970	1,090	1,216	1,335	13.4
서비스	618	759	910	1,060	1,224	1,385	17.5
계	1,330	1,600	1,880	2,150	2,440	2,720	15.4

자료: "Automatic Identification and Data Capture Market - Global Forecast to 2025", MarketsandMarkets, 2020.08
 산출근거: AIDC 시장의 세그먼트에서 하드웨어 부분을 제외하고, 소프트웨어와 서비스의 비중을 산출하여 광학문자인식 시장에 적용

좀 더 상세하게 대륙별 시장규모를 살펴보면, 아시아태평양 지역이 2020년 기준 487백만 달러로 가장 크고, 연평균 17.9 % 성장해 2025년 1,109백만 달러의 규모가 될 것으로 전망된다. 아시아태평양 지역은 데이터 산업에 대한 관심이 타 지역에 비하여 높기 때문에 고성장이 예상되며, 향후 성장세가 계속될 것으로 예상된다. 북미 지역은 2020년 기준 376백만 달러에서 연평균 13.9 % 성장해

2025년 721백만 달러가 될 것으로 전망된다. 유럽은 2020년 기준 335백만 달러에서 연평균 13.5 % 성장해 2025년에는 629백만 달러의 규모가 될 것으로 전망되며, 기타 국가들은 2020년 기준 130백만 달러에서 2025년 261백만 달러로 연평균 15 %의 성장이 전망된다.

표 5 광학문자인식의 대륙별 시장

(단위: USD MILLION)

지역구분	2020	2021	2022	2023	2024	2025	CAGR(%)
아시아태평양	487	601	721	841	976	1,109	17.9
북미	376	447	520	586	655	721	13.9
유럽	335	397	459	515	575	629	13.5
기타	130	155	182	208	234	261	15.0
계	1,330	1,600	1,880	2,150	2,440	2,720	15.4

자료: "Automatic Identification and Data Capture Market - Global Forecast to 2025", MarketsandMarkets, 2020.08
 산출근거: 광학문자인식 시장의 상위시장인 AIDC시장의 대륙별 비중을 산출해 광학문자인식 세계시장규모에 적용

국내 광학문자인식 시장 역시 AIDC의 세계 시장규모를 기반으로 2020년 국내 데이터산업이 세계 시장에서 차지하는 비중 8.26 %를

일괄 적용해 도출한 결과, 2020년 110백만 달러에서 연평균 15.4 % 성장해 2025년에는 225백만 달러가 될 것으로 전망된다.

표 6 광학문자인식의 국내 시장

(단위: USD MILLION)

지역구분	2020	2021	2022	2023	2024	2025	CAGR(%)
국내	110	132	155	178	202	225	15.4

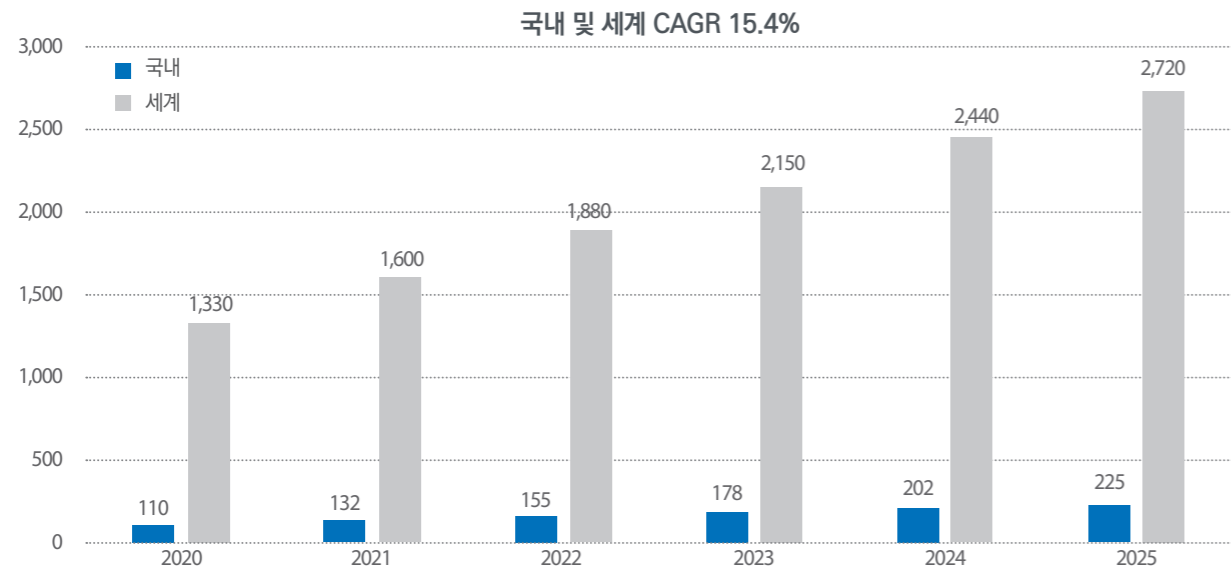
자료: "Automatic Identification and Data Capture Market - Global Forecast to 2025", MarketsandMarkets, 2020.08
 "모든 길은 데이터로 통한다... 마이데이터 사업 도전하는 금융권", 위키리스크한국, 2020.02.21.
 "2021 데이터산업현황조사", 한국데이터산업진흥원, 과학기술정보통신부, 2021.12.
 산출근거: AIDC 세계시장 규모에 2020년 국내 데이터산업이 세계에서 차지하는 비중(20조원/242조원=8.26%)을 일괄적용

경쟁현황

광학문자인식 관련 기업은 국내외 다수가 존재하며, 경쟁이 과열되고 있는 상태이다. 국내기업으로는 (주)모바일리더, (주)셀바스에이아이, (주)레티아, (주)페이스시스템, (주)코테크시스템, (주)에이브레인, (주)드라마엔컴퍼니, (주)악어디지털 등이 있으며, 해외기업에는 러시

아 아비(ABBYY), 일본의 파나소닉(Panasonic), 미국의 리드테크놀로지(LEAD Technologies Inc.), 뉴앙스커뮤니케이션(Nuance communications), 벨기에의 아이리스에스에이(I.R.I.S. SA.) 등이 있다. 일반적으로 광학문자인식 시장에서 활동하는 기업은 3가지 형태로 구분이 가능하다.

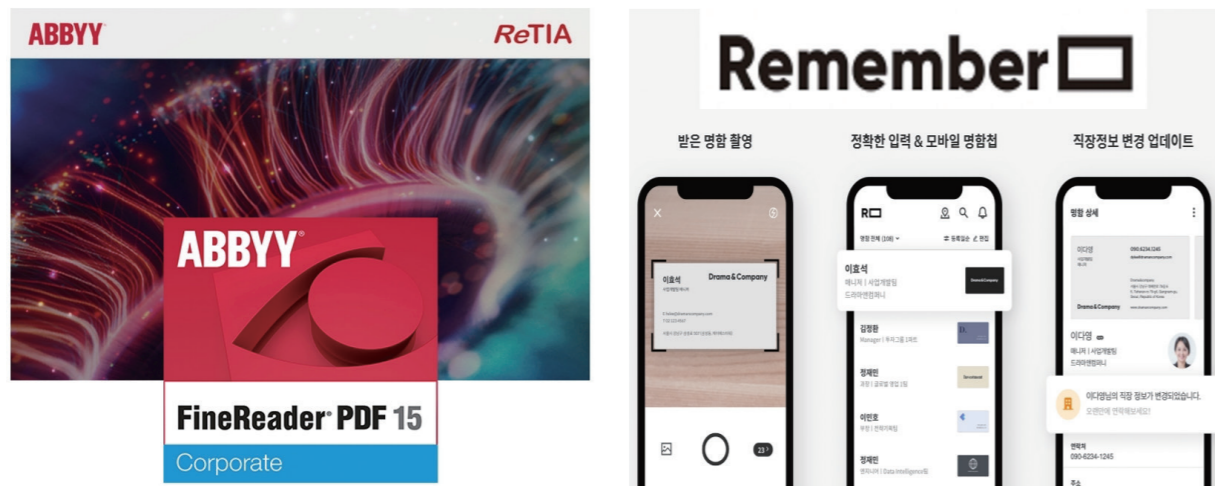
그림 1 광학문자인식의 세계 및 국내 시장 (단위: USD MILLION)



①광학문자인식 엔진을 개발해 판매하는 기업, ②광학문자인식기술로 문서의 디지털화 서비스를 영위하는 기업, ③광학문자인식기술을 기반으로 타 영역의 서비스까지 확장하는 기업으로 구분할 수 있다. 엔진을 직접 개발해 판매하는 대표적 기업은 러시아의 아비인데, 1989년 모스크바에서 설립되었고, 광학문자인식 및 언어기반 기술,

데이터 캡처 분야에서 업계를 선도하고 있는 소프트웨어 개발 및 서비스 공급업체이다. 문서의 디지털화 서비스를 영위하는 기업은 국내의 (주)악어디지털이고, 2014년 설립되어 서류 수거부터 스캔, 전자화, 원본보관·파기까지 디지털문서 전환 과정에 필요한 서비스를 통합 제공하고 있다.

그림 2 광학문자인식 애플리케이션



출처 : (좌)ABBY 사의 광학문자인식 소프트웨어 FineReader PDF 15, (우)드라마앤컴퍼니 사의 광학문자인식기술 기반 명함관리 애플리케이션 Remember.

총 300억 원의 투자를 받았으며, 현재는 일본 시장을 개척하기 위하여 일본 현지법인을 세웠다. 광학문자인식 기술을 기반으로 타 영역까지 서비스 범위를 확장한 기업으로는 (주)드라마앤컴퍼니를 들 수 있다. 동사는 광학문자인식기술 기반 명함관리 애플리케이션인 '리멤버'를 개발해 현재는 HR(Human Resources)분야로 서비스 범위를 확장하고 있으며, 링크드인과 같은 종합 비즈니스 전문 소셜 미디어로 성장하고 있다.

광학문자인식 시장의 경쟁 현황을 요약하면, 경쟁자가 다수이기 때문에 경쟁강도(Competitive Intensity)가 높고 공급자 교섭력이 낮으며, 부가 가치 창출을 위하여 특화된 기술이 필요한 시점에 놓인 시장인 것으로 판단된다. 광학문자인식 기술은 숙련된 소프트웨어 개발 경험을 가진 인력과 문자코드로 변환된 언어에 대한 이해, AI 기술을 통한 이미지 처리, 텍스트 위치 검출, 문자 인식 기술의 고도화 등이 경쟁요소로 작용한다고 할 수 있다.

그림 3 광학문자인식 시장의 사업기회 분석

정책(Policy)

디지털문서의 법적 효력 인정 및 육성정책 시행

- 전자문서 및 전자거래 기본법 개정안 시행
 - 종이문서의 스캔본을 공인전자문서센터에 보관
 - 종이문서 원본의 폐기
 - 디지털문서의 법적 효력 인정
- AI 및 데이터 관련 산업 육성정책 시행
 - 데이터 구축 관점에서 데이터의 신뢰성 문제 대두

기술(Technology)

문자 인식기술의 고도화와 특화기술 개발

- AI를 활용한 글자 인식기술 고도화
 - 훼손되거나 분간이 어려운 글자를 앞뒤 문맥을 분석하여 정확하게 추정 및 복원
- 광학문자인식 기술과 타 분야 분석도구의 융합을 통한 특화기술 개발
 - 회계장부(변환 대상) 재무 분석 시스템과 연계
 - 병원차트(변환 대상) 환자 이력관리 시스템 연계

사회(Society)

비대면 문화 확산에 따른 금융 분야 수요 증가

- 인터넷 전문은행(케이뱅크, 카카오뱅크) 및 가상화폐거래소 이용자 증가
 - 본인 인증(신분증)을 위한 광학문자인식 기술 수요 증가
- 모바일 결제 시스템에 광학문자인식기술 적용
 - 스마트폰을 이용한 신용카드(실물) 촬영 및 인식

시장(Market)

광학문자인식 서비스 시장 확대

- 광학문자인식 엔진을 통한 응용 서비스 분야 확대
 - 회계장부, 병원차트, 수사기록, 역사기록, 명함, 영수증 등
- 광학문자인식 기술을 기반으로 타 분야 서비스 확장
 - 명함인식에서 종합 HR 분야로 서비스 확장

4) 분석자 인사이트

현재 국내는 2020년12월부터 전자문서 및 전자거래 기본법 개정이 시행되면서 종이문서와 동일한 수준으로 디지털문서에 대한 법적 효력이 인정되는 상황이다. 본 법안은 네거티브 방식이 적용된 법안으로 특별한 규정이 있는 문서를 제외하고, 그 외 모든 전자문서는 법적 효력을 인정받게 된다는 내용을 포함한다. 네거티브 방식의 적용은 다양한 분야에서 광학문자인식 기술의 적용을 촉진시키는 요인으

로 작용할 것으로 기대된다. AI 및 데이터 산업 육성을 위한 신뢰성 있는 데이터 구축에 대한 요구의 증가 역시 데이터 구축 기반기술 가운데 하나인 광학문자인식 기술의 수요를 증가시킬 것으로 사료된다. 수기로 작성되는 등 디지털화가 되지 않은 종이문서를 정확하게 디지털화하기 위해서는 훼손되거나 흐려 쓴 글씨의 정확한 추정이 필수적이다. 이를 위해 다양한 문장을 시로 학습시켜 앞뒤 문맥을 분석하여 목표문자를 추정하면 정확하게 복원이 가능하며, 이러한 부분에 기술개발의 포커스를 맞출 필요가 있다고 판단된다.

광학문자인식 시장에서 활동하는 플레이어들의 활동영역은 소프트웨어 부문과 응용 서비스 부문으로 구분이 가능하며, 앞으로 서비스 부문의 시장 확대가 기대된다. 서비스 부문에서 차별화를 갖기 위해서는 특화된 기술개발이 필요하다. 예를 들어 수기로 작성된 기업의 회계장부를 디지털화할 때 재무 분석 시스템을 자동으로 연계한다면, 수기로 작성된 병원차트를 디지털화할 때 환자 이력관리 시스템으로 자동 연계하는 등의 특화기술 개발은 경쟁이 과열된 시장에서 타 업체와의 차별성을 갖는 좋은 전략이 될 수 있다. 광학문자인식을 통한 명함관리 애플리케이션에서 시작하여, 이력관리, 맞춤형 기업 정보 제공 등으로 연계한 서비스를 진행하고 있는 (주)드라마엔컴퍼니가 이러한 전략을 적용한 좋은 사례라고 할 수 있다.

사회적으로 코로나19로 인한 비대면 문화의 확산은 인터넷 전문은

행의 성장을 가속화시켰으며, 신분증 촬영 및 인식을 통한 본인인증 절차로 인하여 광학문자인식 기술의 수요 역시 함께 증가하였다. 가상화폐거래소 이용자의 폭발적인 증가 역시 동일한 이유로 기술의 수요를 증가시키는 요인이 되었다. 또한 비대면 문화의 확산은 사람이 재택에서 상주하는 시간을 길게 하는 요인이 되었고, 식자재, 생필품 등을 온라인에서 구매하는 경우가 늘어나면서 모바일 결제 시스템을 활용하는 인구가 증가하였다. 현재 일부 모바일 결제 시스템은 스마트폰으로 신용카드(실물)를 촬영하면 광학문자인식을 통해 결제 시스템으로 연동되어 동작한다. 이러한 기술의 보급은 광학문자인식 시장에 긍정적으로 작용할 것으로 기대되며, 향후 비대면 문화의 확산으로 인해 광학문자인식 기술이 필요한 부분에 대한 선제적인 분석과 비즈니스 모델 발굴을 고민해야 할 시기라고 사료된다. [ASTI](#)



ASTI MARKET INSIGHT



본원 (우)34141 대전광역시 유성구 대학로 245 한국과학기술정보연구원
T. 042) 869-1004, 1237 F. 042) 869-1091

분원 (우)02456 서울특별시 동대문구 회기로 66 한국과학기술정보연구원
T. 02)3299-6114 F. 02)3299-6244

