

REPORT

**『KAKAO AI REPORT & AI 생태계, 스타트업이 말한다』
에 관하여**

대학원 지속성장데이터사이언스학부 경제학전공(박사과정)

AD20211902 김윤희

1. 서론 : 보고서의 성격

이 두 보고서는 중 하나는 kakao 그룹의 AI 사업의 진행상황과 이에 따른 문제점과 해결방안 등의 고민을 담고 있는 보고서임.

다른 하나는 국내 AI 스타트업의 생태계의 확장을 위해서, 무엇을 추구하고 무엇을 해야하는 지, 그리고 그 선결조건들이 무엇인지에 대하여 심층 면접 등을 통해 그 결과를 도출한 보고서임.

특이한 것은 KAKAO 보고서는 시간적으로 2018년에 쓰여진 보고서이며, AI 생태계 보고서는 시간적으로 2021년에 쓰여진 보고서라 간격이 있을수 있음

2. KAKAO AI REPORT

카카오 report에서는 AI에 대한 윤리현장으로부터 시작해 AI 기술이 일상생활에 적용되고 있는 분야로 보고서가 전개된다.

우리 생활과 밀접하게 연관되어있는 분야인 스마트 홈, 음식, 자동차, 오피스, 엔터테인먼트 분야에서의 AI의 현재와 미래 모습을살펴보고 AI 확산을 방해하는 허들을 넘기 위해 필요한 것들을 살펴보고 있다.

카카오에서 이용자 보호와 고객관리를 위해 어떤 철학과 신념으로 AI 기술을 활용하고있는지 알아보고, OCR 시스템 구성과 모델 그리고 OCR 기술을 활용한 아날로그 데이터의디지털화 방법들을 살펴보며 EMNLP 2018의 주요 내용과 카카오에서 머신러닝으로 문제해결을 위해 오픈한 공개 플랫폼을 소개하고 있다

1부 스페셜 토픽

1부 스페셜 토픽에서는 유네스코에 초청된 카카오의 AI윤리 현장의 발표내용이 있으며, 유네스코의 인터넷 사용에 대한 일반적원칙인 ROAM을 소개하고 있다

1-1) 유네스코에 초청된 카카오의 AI 윤리

- 2018년 11월 유네스코에서 카카오는 동사의 알고리즘 혹은 인공지능(artificial intelligence, AI)에 대한 내부 윤리 규정을 만들고, 이를 발표한 것은 국내에서는 카카오가 처음이라고 함
- 카카오 알고리즘 윤리 현장은 국내 기업에서는 최초로 밝힌 것으로 간단히 요약하면 다음과 같다

<1> 카카오 알고리즘의 기본 원칙

*카카오는 알고리즘과 관련된 모든 노력을 우리 사회 윤리 안에서 다하며, 이를 통해 인류의 편익과 행복을 추구한다.

<2> 차별에 대한 경계

*알고리즘 결과에서 의도적인 사회적 차별이 일어나지 않도록 경계한다.

<3> 학습 데이터 운영

*알고리즘에 입력되는 학습 데이터를 사회 윤리에 근거하여 수집·분석·활용한다.

<4> 알고리즘의 독립성

*알고리즘이 누군가에 의해 자의적으로 훼손되거나 영향받는 일이 없도록 엄정하게 관리한다.

<5> 알고리즘에 대한 설명

*이용자와의 신뢰 관계를 위해 기업 경쟁력을 훼손하지 않는 범위 내에서 알고리즘에 대해 성실하게 설명한다.

- 이는 카카오의 AI에 대한 방향성을 가리키는 것으로 볼 수 있으며, 대부분의 개인이나 기업, 정부기관 등등 AI에 대한 이용방안에 대해 윤리적 준칙이 필요할 것으로 보임

1-2) 유네스코 AI 윤리의 토대가 될 ROAM

- 유네스코는 2018년 11월 인터넷 보편성 지표의 최종 초안을 발표함, ROAM이라는 단어로 인터넷 보편성의 원칙을 정의했고, 인터넷 보편성은 4가지 원칙간의 상호작용을 포함에 전체적으로 인터넷을 발전을 이해하는 것이 중요하다고 강조함
- 인터넷 보편성의 원칙인 ROAM의 의미는 다음과 같음
 - R: 인터넷은 인권을 기반으로 한다(that the Internet is based on human Rights).
 - O: 인터넷은 열려 있어야 한다(that it is Open).
 - A: 모든 사람이 접근할 수 있어야 한다(that it should be Accessible to all).
 - M: 다중 이해관계자들에 의해 육성돼야 한다(that it is nurtured by Multistakeholder participation).
- 좀 더 자세히 들어가자면 지표의 구현(Implementation of the Indicators)을 위한 6가지 카테고리, 25개 테마 및 124가지 질문으로 개발된 303가지 지표(110가지 핵심 지표 포함)가 포함되어 있으며, ROAM 카테고리 상단에는 79개의 크로스커팅 지표가 개발되어 있음
- 유네스코는 이 준칙들을 활용해로 각 국가별 인터넷 문제에 적용해 큰 가치를 창출할 수 있기를 희망하고 있음.

2부. In-Depth

2부에서는 AI의 다양한 활용상태를 주거, 음식, 교통, 오피스, 여가, 범죄 등의 일상생활에서 흔히 적용되는 분야로 나누어 소개하고 있음

2-1) (주거) 스마트홈이 성공하기 위한 필요충분조건

- 스마트 홈을 구현하는 방법에 ICT뿐 아니라 사회·경제·문화적인 다양한 측면을 고려해야 한다고 보고서에서는 말하고 있음
- 과거 인텔리전트 홈의 활성화 실패에도 불구하고 현재 스마트홈이라는 이름으로 확산되고 있음
- 과거의 실패를 되풀이하지 않을 스마트홈 성공의 조건이 무엇인지 밝히고 있음
- 스마트 홈의 충분조건은 무엇인가?
 - 스마트 홈의 핵심적인 문제는(주로 아파트 같은 건축분야에서) 분양 원가가 상승한다는 문제도 있지만, 사용자의 피드백을 받을 수 있는 체계의 부재로 인해 일방적인 스마트 홈 서비스 공급이 안게되는 경직성과 비소통의 문제라 할 수 있음
 - 필자는 애플(Apple)의 앱스토어에서 스마트폰과 같은 개념의 플랫폼 비즈니스를 기반으로 하는 스마트 홈 공급 체계가 스마트 홈의 충분조건으로서 고려될 필요가 있다고 전망하고 있음
- 스마트 홈을 플랫폼 비즈니스 관점에서 접근하는 것은 관련된 다양한 주체들 간의 공감대와 합의가 전제되어야 하며, 공유경제 관점에서 협력(collaboration)해야 하는 사안임(쉬운일은 아니다라고 밝히고 있음)

2-2) (음식) 맛있는 인공지능 이야기, 스마트하게 먹기 위한 방법

- 음식분야 AI의 혁신적인 도입사례로 2018 AI 서밋 SF에서 만난 T.G.I. 프라이데이스의 우수 직원인 AI 바텐더 플래너건을 우선 소개하고 있음
- 특히 식품분야 AI에 포함된 스마트센서는 식품의 전 주기에 대한 유통 환경 정보 계측, 모니터링, 품질 정보 예측,

상황 제어 등의 여러 역할을 한 번에 수행한다고 알려주고 있음.

- AI를 기반으로 첨단기술을 식품산업에 융합한다는 점에서 식품 안전 관리를 위한 스마트센서는 IT 융합의 결정체라고 할 만하다고 하며 기술 융합의 중심은 단연 AI이다.

- 식품 유통에서 AI가 담당하는 또 하나의 중요한 일은 수요 예측으로 수요 예측이 잘못된다면 식품의 품귀 또는 손실로 연결되며, 식품의 품귀 또는 손실이 반복될 경우 품질의 저하는 불을 보듯 뻔하다며 AI수요예측의 중요성을 밝히고 있음

2-3) (교통) 자율주행, 스마트 카를 위한 딥러닝의 진화와 확산의 허들

- 이 글에서는 자율주행과 인공지능, 스마트 카의 딥러닝 적용 현황에 대한 정리와 함께 향후 딥러닝 기술의 확산을 위한 자율주행차의 과제를 정리하고 있음
- 현재 자율주행에서 딥러닝 기술은 주로 인식 분야에 적용되고 있다. 자율주행 레벨 2 정도로 상용화된 대부분의 부분 자율주행 차량들에서는 딥러닝 기술이 적용되지 않음
 - 현재 상용화된 자율주행 차량은 대부분 고속도로 자율주행 수준의 레벨 2 차량임. 향후 레벨 3 자율주행차가 본격화되는 시점은 2021년 정도로 예상된다고 말하고 있음
- 딥러닝 기술이 필수적인 레벨 4와 레벨 5로 레벨 4부터는 도심 자율주행 기능이 필요하게 되며 고속도로 주행에서 필요로 했던 주요 기능인 한 차선 주행, 차선 변경을 넘어서기 때문에, 물체 인식과 주행 측면에서 몇 단계 더 높은 수준의 기술을 필요로 하게 된다고 말하고 있음.
 - 교차로의 신호등 인식, 교차로 주행, 합류 차선 주행, 분기 차선 주행, 끼어들기, 보행자 및 자전거 인식 등 다양한 기능을 필요로 함. 더불어 차량의 진행 방향 예측, 보행자 및

자전거의 진행 방향 예측, 기타 정형화되지 않은 문제 등에 대한 해법도 필요하며,

- 도심 주행을 위한 문제를 해결하기 위해서 센서 시스템, 인식 시스템 등에서 업그레이드가 필요한 상황이며, 인식 및 판단의 정밀도 향상을 위해서 딥러닝 기술이 필수적으로 적용될 것으로 예상된다고 말하고 있음

○ 스마트 카에서의 딥러닝 적용 동향은?

- 스마트카를 도입하고 있는 이들 업체의 공통점은 GPU 기반의 인공지능 하드웨어를 사용한다는 점.
- 딥러닝 기술은 인식 성능이 뛰어나지만, 차량에 내장하기 위해서는 처리 속도와 성능 향상을 위해 전용 하드웨어를 필요로 함. 현재는 주로 GPU 기반 보드가 사용되고 있으며, 앞으로는 인공지능 전용 칩의 사용도 예상되는 상황
- 현재 개발 중이거나 상용화된 기술로 볼 때 합성곱 신경망(convolutional neural network, CNN), 순환 신경망(recurrent neural network, RNN), 심층 Q-네트워크(deep q-networks, DQN) 등 인공지능 기술 측면에서는 서로 다른 응용 영역을 가지고 있는 분야가 주로 사용될 것이라고 말함.
- CNN 기술은 주로 영상 인식 및 상태 인식에, RNN 기술은 음성인식 및 교통 상황 예측에, DQN 기술은 주로 주행 학습과 인공지능 시뮬레이터에 각각 사용되고 있으며, CNN 기술은 차선 인식, 차량, 보행자 및 객체 인식, 자율주행을 위한 end-to-end 학습, 운전자 상태 인식, 탑승자 인식 등의 분야에 적용되거나 연구되고 있음, RNN 기술은 대표적으로 사용되는 음성 인식뿐 아니라, 교통 상황 예측, 운전자 상태 인식 등에 사용. DQN 기술은 다양한 상황에서의 주행 대응 방법 학습, 인공지능 시뮬레이터에서의 주행에서 학습 알고리즘 등에

사용되고 있음

2-4) (오피스) 인공지능과 함께 일하게 될 미래

- 딥러닝(deep learning) 기술로 대변되는 머신러닝(machine learning) 기술이 우리 생활 곳곳에서 사용되고 있음
 - 미래 우리의 사무 환경은 더 많은 AI가 일하는 곳이 될 것이며, 모든 사무실에 컴퓨터가 있듯이 AI가 일하는 모습 역시 현실로 오고 있음.
 - AI가 인간보다 더 잘하는 영역을 정리하면 다음과 같음
 - 같은 일을 지속적으로 반복하는 일
 - 빅데이터에서 패턴을 찾아내고 분류하여 이를 처리하는 일
 - 방대하고 정형·비정형화된 데이터 분석과 관련된 문제를 해결하는 일
- 분야에서 AI가 인간의 업무를 대신하게 될 것이다라고 말하며, 우리 인간은 더 창의적이고 감성적인 업무 등에 집중할 수 있을 것이다라고 말하고 있음

2-5) (여가) 게임산업을 위한 인공지능 기술의 활용 가능성

- 현재 인공지능 기반의 게임산업을 위해 연구되고 있는 기술 개발 트렌드는 크게 게임 플레이가 가능한 게임과 인공지능 엔진의 구축과 지능형 분석 도구의 2가지로 나누어 생각해볼 수 있음
- 게임 분야를 대상으로 핵심 인공지능 기술을 분류하면 다음과 같이 4가지로 나누어볼 수 있음
 - <1> 심층 강화 학습 기술 <2> 절차적 콘텐츠 생성 기술 <3> 게이머 행동 기반 게임 애널리틱스 기술 <4> 인게임 행동 데이터 마이닝 기술 로 볼수 있음

- 향후 연구개발 트렌드의 변화는 딥러닝과 같은 기계학습 기술의 고도화에 따르는 게임 인공지능 연구의 본격화와 기계학습을 위한 학습 데이터의 활용 측면으로 나누어 생각해볼 수 있음
- 마지막으로, 국내외 산업의 활성화를 위해 고려되어야 할 기술적 이슈와 현황을 정리함
 - <1> 심층 강화 학습 기반의 게임 인공지능
 - <2> 인공지능 학습 데이터 구축 표준화
 - <3> 스스로 변화에 적응하며 진화하는 콘텐츠
 - <4> 다양해지는 사용자 니즈 변화에 빠르게 대응
- 게임 속에서 빚어지는 상황의 문제점 해결 방법은 실세계를 대상으로 연계선상에서 적용이 가능하며, 인게임 행동 예측은 군중심리 · 집단행동 · 사회현상 예측 등 인간 행동이 관여된 현상 분석에 활용하는 것이 가능함.
 - 특히 게임 인공지능 기반의 기술 활용 측면에는 이러한 예측에 의해 예상되는 위기 대처를 위한 최적화 대응책의 자동 생성 기술이 포함되어 그 활용 가능성은 무궁무진할 것으로 보임
- 게임 인공지능의 활용은 게임으로만 제한되는 것이 아니라, 단편적으로는 소비자의 성향 변화 예측에 따르는 쇼핑 제안부터 행동 예측 기반의 사회 안전망 구축, 전술 · 전략 예측 및 최적화된 대응 전략 생성, 인간 성향의 대응 로봇 행동 생성 등 거의 모든 실세계 문제를 게임 속 세상을 통해 풀어낼 수 있을 것으로 예상됨.
- 따라서 이미 딥마인드를 비롯한 게임 인공지능 연구기관들의 목적은 ‘게임을 통해 인공지능의 가능성을 확인하고자’ 하는 것이지 게임 자체만을 목적으로 하지 않음. 이처럼 게임 인공지능 기술은 게임 분야에만 한정하지 않고 전반적인 기술을 도약시키기 위한 전방위적 산업 활성화의 테스트베드로서 그

역할과 파급력이 큰 기술적 도메인으로 판단하는 것이 합당할 것으로 생각됨

2-6) (치안) 인공지능, 치안의 일선에 서다

- 치안 분야에서의 인공지능 활약상에 대해 소개하는데 4가지 범죄군 중 현재 인공지능을 활용해 발생 가능성이 높은 범죄를 예측하여 예방하거나 수사하는 데에 가장 효과적인 성과를 보이고 있는 사례를 범죄군별(범죄수사, 재산범죄, 성범죄, 테러대응)로 나누어 살펴보고자 함
- 범죄 수사
 - 부산지방경찰청은 올해 1월부터 인공지능 기술로 범죄자의 여죄를 추적하고 있다. 인공지능은 미제로 남은 사건이 기록된 입장일지와 범죄자의 입장일지를 비교하면서 그가 저질렀을 가능성이 높은 사건들(여죄)을 찾아냄
 - CCTV 영상을 분석하는 기술이 고도화되는 점도 주목할 만함. 일본 오사카대학교는 2017년 사람의 걸음걸이를 분석해 높은 정밀도로 개인을 식별할 수 있는 인공지능 시스템을 개발
 - CCTV 영상에 촬영된 사람의 걸음걸이에서 나오는 보폭이나 손을 흔드는 방식, 자세 등에서 개인의 특징을 추출해 식별하는 것이다
- 재산범죄: 보이스피싱, 불법 대부, 불공정 거래를 예방하는 인공지능으로, 현재 금융감독원과 국립과학수사연구원은 인공지능을 활용해 보이스피싱 사기범의 음성 특징을 비교·분석해 사기범의 목소리를 데이터베이스화함
- 성범죄 : 피해자 진술을 대신 받는 챗봇과 성범죄자를 감시하는 전자발찌 그리고 음란물을 차단하는 인공지능으로 개발되고 있음

- 성범죄 사건의 피해자는 진술 및 재판에서의 신문 과정을 거치며 2차로 정신적 피해를 입을 수 있음. 경찰청과 과학기술정보통신부는 성범죄 피해자의 심리적 부담감을 덜어줌과 동시에 제대로 된 진술을 확보하기 위한 기술을 개발했는데, 바로 성범죄 피해자의 진술을 돕는 비대면(非對面) 인공지능 상담 챗봇
- 정밀한 기계 학습으로 훈련된 인공지능 상담 챗봇이 사람을 대신해 수사에 필요한 피해 특성별 시나리오에 기반한 피해자의 진술을 확보하고 추가 피해를 막는 것임
- 테러 대응 : 페이스북은 테러리스트들의 콘텐츠를 찾아내고 삭제하는 작업에 인공지능 기술을 적용.
 - 현재 페이스북이 삭제하는 이슬람국가(IS) 및 알카에다(Al-Qaeda) 관련 콘텐츠의 99%가 인공지능에 의해 탐지 및 삭제되고 있다. 나머지 채워지지 않는 1%는 사람의 힘을 빌림
 - 인공지능이 콘텐츠의 맥락(context)까지 파악할 수 없다는 기술적 한계에 따라 사람이 이 틈을 메우는 역할을 하는 것

3부. KAKAO INSIDE

3부에서는 KAKAO 회사내에서의 AI의 실제적인 사용예를 1) 인공지능 기반의 고객센터를 향한 첫걸음의 시작 2) 카카오, 유해 콘텐츠 차단을 위해 AI 방패를 펼치다 3) 문제 해결을 위한 머신러닝 오픈 플랫폼, 카카오 아레나로 이야기를 풀어가고 있다

3-1) 인공지능 기반의 고객센터를 향한 첫걸음의 시작

- 이 글에서는 카카오에서 인공지능 기반 고객센터의 첫 시작이라 할 수 있는 카카오톡의 상담 챗봇과 상담톡을 고객센터에

적용하기까지 과정을 설명하고 있음

- 고객들로부터 접수되는 문의 유형을 육하원칙(5W1H)에 따라 분류해보면, 왜(Why) 유형에 해당되는 문의가 꽤 많은 비중을 차지한다는 것을 알 수 있음.
 - 문의 유형 중 해결하는 데 가장 많은 시간이 필요한 질문이 ‘왜’에 대한 유형이다. 문의 사항을 해결하고 답변을 하기까지 많은 시간이 필요하게 됨. 그리고 문의에 대한 해결과 답변을 받기까지 가장 많은 시간이 필요한 문의가 ‘전화’라는 수단을 통해서 접수되고 있다는 문제점이 있음
- 상담 챗봇에서 답을 구하다
 - 카카오톡의 상담톡과 상담 챗봇을 카카오톡 고객센터에 도입
 - <1> 고객이 채팅창에 입력하는 질문과 정확히 일치하는 내용이 있을 경우, 고객이 자주 질문하는 서비스 및 기능에 대한 소개 내용을 ‘기능 메뉴 바로가기’, ‘동영상을 이용한 설명’, ‘안내 이미지’ 등 다양한 콘텐츠를 활용해 답변을 제작하고, 채팅창에 입력되는 내용을 분석해 이에 대응하는 콘텐츠를 챗봇이 답변하도록 함
 - <2> 고객이 채팅창에 입력하는 질문과 정확히 일치하는 내용이 없을 경우, 국내 최초 챗봇 내 검색 기능인 ‘사만다R’를 도입하여 유사한 질문 목록을 보여주고 고객이 가진 질문에 대한 답을 최대한 찾도록 안내하고 있음

3-2) 카카오, 유해 콘텐츠 차단을 위해 AI 방패를 펼치다

- 지금까지는 육안을 통해서 사람이 콘텐츠 이상 여부를 판단하였지만, 최근에는 머신러닝 기술을 활용하여 콘텐츠의 유해성을 판단하고 걸러내는 작업이 진행되고 있음.
- 현재 카카오에서는 이미지 외에 글, 동영상, 음악, 파일 등

다양한 콘텐츠의 유해성을 관리

○ 카카오의 유해 콘텐츠 필터링 기술의 원리와 현황

- 카카오의 유해 콘텐츠 필터링 기술은 텐서플로를 기반으로 하고 있음. 예를 들어 하나의 이미지를 숫자로 바꾸고, 이미지를 하나의 행렬로 표현.
- 텐서플로 모델의 규칙에 따라서 행렬 데이터를 잘게 나누고 별집처럼 나누어서, 적용 알고리즘에 따라 특징을 분석
- 특히 이미지의 경우 카카오는 합성곱 신경망(convolutional neural network, CNN) 모델이라는 학습 알고리즘을 사용하고, 알고리즘 기술은 각 회사들마다 다른 특징 기술을 가지고 있음

3-3) 문제 해결을 위한 머신러닝 오픈 플랫폼, 카카오 아레나

○ 지난 2018년 11월 카카오는 머신러닝으로 문제를 해결하고 최적의 솔루션을 찾는 대회를 시작

- 이 대회는 ‘카카오 아레나’라는 공간을 통해 오픈됐고, 이곳은 머신러닝 대회를 상시 개설하고 운영할 수 있는 플랫폼

○ 첫 번째 대회는 쇼핑몰에 등록된 상품의 텍스트, 이미지 정보 등을 활용해 카테고리 분류의 정확도를 높이는 ‘쇼핑몰 상품 카테고리 분류’를 주제로 진행되었으며 머신러닝 기술을 활용해 다음(Daum) 쇼핑에 존재하는 수억 개의 상품을 더 정확하게 카테고리화할 수 있는 기술을 만드는 것을 목표로 함

- 이 대회를 위해 카카오는 다음 쇼핑에 누적된 상품 이름, 브랜드, 이미지 등 약 1000만 건에 달하는 데이터와 4000건 이상의 카테고리 데이터를 오픈

4부. TECH & CONFERENCE

4부에서는 1) 카카오 OCR 시스템 구성과 모델 2) 아날로그 기상 데이터를 OCR로 디지털화할 수 있을까? 3) EMNLP 2018 참관기 4) 2019년 주요 AI 콘퍼런스로 보고서를 마무리 하고 있음

4-1) 카카오 OCR 시스템 구성과 모델

- 카카오의 다양한 서비스에서 가장 필요로 하는 핵심적인 기술 중 하나는 이미지 내의 글자를 자동으로 인식하는 것.
- 예를 들어 카카오 드라이버에서 자동차 번호판을 자동으로 인식하거나 카카오뱅크에서 자동으로 신분증의 글자를 뽑아내는 등, 이미지에서 자동으로 글자를 추출해 검색 및 광고에 활용하게 된다면 모바일 서비스의 사용성을 더욱 증대시킬 수 있음
- 이러한 카카오 사내 수요의 증가에 발맞춰 최근 카카오 AI 기술팀에서는 사내에 광학 문자 인식(optical character recognition, OCR) 서비스를 베타 론칭함.
- OCR란 사람이 쓰거나 컴퓨터로 입력되어 인쇄된 문자의 영상을 기계가 읽을 수 있는 문자로 변환하는 것을 말하는데, 이 분야 역시 딥러닝으로 인해 많은 발전을 이루고 있음

4-2) 아날로그 기상 데이터를 OCR로 디지털화할 수 있을까?

- 전통적인 OCR 파이프라인(pipeline)은 텍스트 라인(text line)을 찾는 모듈과 문자(letter)를 분할하는 모듈 등 세분된 모듈로 동작한다. 서로 다른 패턴을 인식하는 특징(feature)은 인간 개발자가 직접 설계해야 함. 아울러 평판 스캐너로 촬영한 고품질 이미지에서만 제한적으로 동작
- 다행히 컴퓨터 비전(computer vision)에서 이뤄진 기술적 발전으로 OCR 기능 또한 크게 향상

- CNN(convolutional neural networks)이 바로 그 주역으로 대량의 데이터 학습을 통해 이미지에서 텍스트를 인식하는 규칙(feature extraction)을 스스로 만들어냄.
- 오늘날의 딥러닝 기반 OCR는 이미지에서 텍스트 영역(text box)을 검출하는 부분과 텍스트 영역의 텍스트를 개별적으로 인식하는 부분 2가지로 나누어볼 수 있음

3. AI 생태계, 스타트업이 말하다 T

이 레포트는 report에서는 AI 생태계의 확장을 위한 조건들이 무엇이 있는지, 무엇을 해야하는지를 전문가 집단의 심층면접 분석등을 통해서 밝히고 있음

CHAPTER 1. Overview

1-1) Background & Purpose 배경 및 목적

- 국내 스타트업으로서는 유일하게 위이드가 2021년 ‘AI 100’에 이름을 올림. 그러나 일부 전문가들은 국내 인공지능 스타트업들이 내수시장을 중심으로 사업을 영위하고 있어 글로벌 스케일 확보는 다소 부족한 것이 아니냐는 지적
- 기술과 사업의 최전선에서 분투하고 있는 국내 인공지능 스타트업들은 어떤 고민을 하고 있는지 국내 인공지능 스타트업들이 바라는 선결과제를 정리해봄

1-2) Methodology & Process 연구방법론

- 인공지능 스타트업 현황 분석 : 기존 데이터 분석 - 314개 스타트업에 대한 언론 보도기사, 기업 홈페이지 등을 통해

산업군, 비즈니스 현황, 투자유치 현황 등에 대한 자료를 수집하고 분석

- 인공지능 스타트업의 생태계 인식 분석 : 사회표상 및 구조 분석
 - 인공지능 스타트업 관계자들이 느끼는 국내 인공지능 스타트업의 현안에 대해 알아보기 위해, 우선 심층인터뷰를 진행해 사회표상과 그 구조를 도출-인공지능 스타트업의 대표 또는 관리자 16명을 대상으로 한 심층인터뷰를 바탕으로 사회표상 코드를 도출하고 이를 바탕으로 전체 사회표상의 구조를 그려냄
- 인공지능 스타트업 생태계 활성화를 위한 선결과제 제안 : 분석적 계층화 기법(AHP)-인공지능 생태계 활성화를 위한 여러 과제의 우선순위를 도출하기 위한 방법으로, 여러 계층으로 구성된 항목의 상대적 중요도를 비교하는 기법인 ‘계층적 분석기법(Analytical Hierarchical Process, AHP)’을 사용-인공지능 스타트업의 대표자 및 관리자 27명을 대상으로 “인공지능 생태계 활성화를 위해 무엇이 더 중요한지” 비교해 응답하는 설문조사를 시행하고, 이를 분석한 결과를 담음

CHAPTER 2. Status of AI Startups

2-1) 산업 현황

- 2021년 2월을 기준으로, ‘한국 스타트업 투자 데이터베이스’에 주요 기술이 인공지능으로 등록된 스타트업은 총 314개사
 - 국내 인공지능 스타트업의 산업군 분포는 의료(20.6%), 엔터테인먼트(16.8%), 교육(7.2%), 광고마케팅(6.5%), 금융(6.2%) 등이 각각 1~5위를 차지

2-2) 투자유치 현황

- 전체 인공지능 스타트업의 상당수(260개사, 전체의 89.4%)는 4회

이하의 투자를 유치한 것으로 나타났는데요. 분포가 높았던 5개 산업군별 총 투자유치 횟수를 비교한 결과 교육 분야(평균 3.4회), 의료 및 금융 분야(각 평균 2.8회), 엔터테인먼트 분야(평균 2.4회), 광고마케팅 분야(평균 2.1회) 순으로 나타남

2-3) 비즈니스 유형별 현황

- 전체 인공지능 스타트업의 비즈니스 유형을 B2B(Business to Business), B2C(Business to Customer), B2G(Business to Government)로 나누어 분석한 결과, 4곳 중 3곳이 B2B 비즈니스를 하고 있는 것으로 나타남

CHAPTER 3. Ecosystem seen by AI Startups

3-1) 조사 및 분석방법

- 본 연구에서는 인공지능 스타트업의 대표 또는 관리자 16명을 대상으로 인공지능 생태계에 대한 인식을 묻는 심층인터뷰를 통해 이러한 사회표상 코드를 도출
 - 또한 도출된 코드를 바탕으로 전체 사회표상의 구조를 분석, 구체적으로는 각 사회표상 코드를 중심부 코드와 주변부 코드로 나누는 방법을 사용함

3-2) 사회표상 코드 분석

- 16명의 인공지능 스타트업 대표 또는 관리자를 대상으로 인공지능 기술 및 비즈니스, 인공지능에 대한 사회인식, 인공지능 스타트업 생태계와 관련한 질문을 중심으로 심층인터뷰를 진행하고, 그 결과를 분석해 총 18개의 코드로 분류
 - 또한 네트워크분석기법을 통해 각 코드별 중심성(Coreness)을 계산해 중심부 코드와 주변부 코드를 도출
- 그 결과, ‘인적자원 확보 및 관리의 필요성(C16, Coreness=0.547)’, ‘데이터 확보와 공유 필요(C12,

Coreness=0.536)’ , ‘인공지능 기술 응용분야의 확대(C11, Coreness=0.302)’ 등 3개의 중심부 코드가 도출

- 이는 인공지능 스타트업 생태계를 나타내는 여러 사회표상 가운데 위 세 가지 코드가 다른 코드와 연결되는 중심 역할을 하고 있음을 의미

3-3) 사회표상 구조 분석

○ 도출된 중심부 코드와 주변부 코드를 바탕으로 전체 사회표상의 구조를 나타냄.

- ‘사회’ , ‘비즈니스’ , ‘산업환경’ , ‘정부’ , ‘기술’ 등 5가지 영역으로 나누어짐

--사회) : 만연한 윤리 이슈에 비해 AI에 대한 사회적 이해도는 낮다고 생각

--비즈니스) : 인적자원 확보 통해 기술력 높이고 응용분야 넓히는 것이 중요

--산업환경) : 데이터 확보를 위해 스타트업 규제 최소화되어야

--정부) : 데이터 확보를 위한 정부의 효과적 지원정책 필요

--기술) : 기술적 승자독식에 대한 우려, 가이드라인이 필요

CHAPTER 4. Priorities for AI Ecosystem

4-1) 조사 및 분석방법

○ 여러 계층으로 구성된 항목에 대해 상대적 중요도를 비교하기 위해 사용되는 대표적인 연구방법으로 ‘분석적 계층화 과정 (Analytical Hierarchical Process, AHP)’ 이 있음

- 본 리포트에서는 인공지능 스타트업의 대표자 및 관리자 27명을 대상으로 “인공지능 생태계 활성화를 위해 무엇이 더 중요

한지” 비교해 응답하는 설문조사를 시행하고, 이를 분석한 결과를 다음

4-2) 상/하위기준 도출 및 계층구조화

- 목적 : 인공지능 생태계 활성화를 위한 선결 과제
- 상위과제 : 인공지능 인프라, 시장 잠재력, 정책환경
- 하위과제
 - 인공지능 인프라 : 개발인력, 데이터, 기반기술
 - 시장잠재력 : 시장수요, 사회적 수용, 비즈니스모델
 - 정책환경 : 정부지원정책, 자율적규제, 법적규제

4-3) 상대적 중요도 분석 결과

- [상위기준 상대적 중요도 : 인공지능 인프라 > 시장 잠재력 > 정책 환경 순]
- [‘인공지능 인프라’ 의 하위기준 상대적 중요도 : 개발 인력 > 데이터 > 기반기술 순]
- [‘시장 잠재력’ 의 하위기준 상대적 중요도 : 시장 수요 > 사회적 수용 > 비즈니스 모델 순]
- [‘정부 정책’ 의 하위기준 상대적 중요도 : 정부 지원정책 > 자율적 규제 > 법적 규제 순]
- [전체 하위기준별 우선순위 : 시장 수요 > 개발 인력 > 데이터 > ... > 법적 규제 순]

CHAPTER 5. Proposals for AI Ecosystem Activation

5-1) 사회적 수용을 바탕으로 한 충분한 시장 수요 창출

- 인공지능 스타트업들은 뛰어난 기술을 기반으로 성장하고 있지만, 지속적인 성장을 위해서는 충분한 시장의 수요가 창출되어야 함
- 한편, 인공지능 기술과 서비스에 대한 사회적 수용도 뒷받침되어야 함

5-2) 인공지능 개발&연구 전문인력 확보를 위한 기반 마련

- 기술을 습득하고 활용하는 인력을 확보하는 것이 국내 인공지능 생태계 활성화에도 매우 중요
- 앞선 중요도 분석 결과에서도 전체 하위기준 중에서도 ‘개발 인력’은 전체 우선순위에서 2위(상대적 중요도는 18.6%)를 차지할 정도로 중요한 요소
- 교육과정 중 일부는 전문인력을 양성하기보다는 ‘청년 취업난’을 해소 하는 데에 초점이 맞춰져 있거나, 코딩과 같은 기초적인 데이터 활용능력 중심 교육에 그치고 있음
- 개발자 인력난을 해소하고, 궁극적으로 인공지능 생태계 전반을 활성화하기 위해서는 인력의 양적인 증대도 필요하지만, 더불어 숙련된 개발자와 연구인력을 양성하는 것도 중요한 과제

5-3) 양질의 데이터를 안전하게 확보하고 거래하기 위한 환경 조성

- 인공지능 기술의 핵심은 ‘양질의 데이터’에 있음
- 양질의 데이터를 거래하고 유통할 수 있는 안전한 기반이 마련되어야 함.
- 대기업 및 스타트업이 보유하고 있는 데이터 공유를 촉진하기 위해 적절한 보상 체계를 마련하고 이를 지원하는 것이 한 방법

5-4) 법적 규제 최소화, 참여형 규제 환경 조성

- 인공지능 기술과 관련한 규제는 법·정책 전문가뿐만 아니라, 인공지능 기술 분야 전문가, 인공지능 기술의 사회적 파급 효과를 연구하는 사회과학자 등 인공지능을 둘러싼 이해관계자의 참여와 사회적 합의를 통해 만들어져야 함
- 또한 지속적으로 최신 기술과 정보를 업데이트하여 시장 수요를 반영한 규제방안을 마련할 수 있도록 정책환경을 조성하는 것이 필요