

인공지능의 새로운 패러다임 온디바이스 AI 기술 동향

제 1 장 온디바이스(On-Device) AI 기술 개요

1. 새로운 인공지능 트렌드 생성형 AI

1-1. 인공지능 기술 개요

1-1-1. 클라우드 기반 디지털 세상

1-1-2. 클라우드 컴퓨팅 기반 AI

1-2. 로봇과 인간의 공존

1-2-1. 인간과 AI의 공존 개요

1-2-2. 로봇기술의 발전과 인간의 삶

(1) 로봇 기술의 발전

(2) 로봇 기술이 인간의 삶에 미치는 영향

1-2-3. 인간과 로봇의 공존

(1) 로봇공학 혁명

(2) 자동화를 넘어선 AI 기반 학습 로봇

(3) 인간-로봇 협업(Human-Robot Cooperation)으로 인해 변화하는 업무 환경

가. 인간-로봇 간의 협업

나. 인간-로봇 협업의 예

① 제조업

② 의료 로봇 시스템

③ 농업

다. 일자리 형태와 근무 환경 변화

(4) 인간-로봇 공존의 미래

(5) 인간-로봇 공존의 이점

가. 경제적 관점

나. 사회적 관점

1-2-4. 인간-로봇 공존을 위한 과제

(1) 일자리 손실 문제

(2) 인간-로봇 간 상호작용의 복잡성

(3) 윤리 문제

가. 미래 산업 판도를 바꾸게 될 게임체인저 인공지능

나. 다양한 윤리적 문제 야기

1-3. 인간-로봇의 공생을 위한 미래

1-3-1. 로봇의 자율성과 휴머노이드 로봇

(1) 로봇의 자율성

(2) 휴머노이드 로봇

1-3-2. 인간-AI의 공존을 위한 규칙

1-3-3. 인간-로봇의 공생을 위한 과제와 고려 사항

1-3-4. 인간-로봇 공존을 위한 발전 방향

2. 인공지능(AI)의 진화

2-1. 미래 사회의 핵심 동력 인공지능(AI)

2-1-1. AI 기술 트렌드 변화

2-1-2. AI 인류의 등장 생성형 AI(Generative AI)

2-2. 인공지능 전환(AX) 시대 생성형 AI의 환각 현상(Hallucination)

2-2-1. 환각 현상(Hallucination) 개념

2-2-2. 환각 현상 발생 원인

2-2-3. 내재적(Intrinsic) 환각과 외재적(Extrinsic) 환각

2-2-4. 생성형 AI 의 환각 사례

2-2-5. 환각 현상 해결 방안

(1) 환각 현상 해결

(2) 검색 증강 생성(RAG) 기술

2-3. 인간 중심 AI

2-3-1. 인공지능의 새로운 접근 방식 인간 중심 AI

2-3-2. 인간 중심 AI 의 핵심 요소

(1) 윤리적 가이드라인 수립

가. 로봇 윤리 기준 마련

나. 로봇 윤리 매뉴얼

다. 국가별 AI 윤리 가이드라인

(2) 책임있는 AI 를 위한 신뢰성 확보

(3) 설명가능한 AI(eXplainable Artificial Intelligence)

3. 온디바이스(On-Device) AI 기술

3-1. 온디바이스(On-Device) AI 기술 개요

3-1-1. 온디바이스 AI 등장 배경

(1) 클라우드 기반 AI 기술의 고민

(2) 생성형 AI 의 확산

(3) 온디바이스 AI 필요성

3-1-2. 온디바이스(On-Device) AI 기술 개념

(1) 인공지능의 게임체인저 온디바이스 AI

(2) 온디바이스 AI 개념

(3) 온디바이스 AI 의의

3-1-3. 온디바이스 AI 장·단점

(1) 장점

(2) 단점

3-1-4. 온디바이스 AI 와 클라우드 AI 차이점

3-2. 온디바이스(On-Device) AI 의 하드웨어 구성 요소 및 작동원리

3-2-1. 온디바이스(On-Device) AI 핵심 기술 구성 요소

(1) 온디바이스 AI 칩

(2) AI 모델 경량화 기술

(3) 머신러닝 프레임워크와 엣지컴퓨팅

3-2-2. 온디바이스 AI 의 작동 방식

3-2-3. 온디바이스(On-Device) AI 기술 특징

(1) 낮은 네트워크 의존도로 인한 효율성과 접근성 향상

(2) 데이터 보안 및 개인정보 보호 강화

(3) 지연시간 감소와 빠른 응답

(4) 비용 절감 및 에너지 효율성

(5) 개인화되고 맞춤형 AI 모델

제 2 장 온디바이스(On-Device) AI 기술 현황 및 시장 전망

1. 온디바이스(On-Device) 핵심 기술

1-1. 새로운 기술 트렌드와 반도체 산업

1-2. AI 반도체 개요 및 특징

1-2-1. AI 반도체 개요

1-2-2. AI 반도체 특징

1-3. AI 반도체 발전 방향

1-3-1. NPU(신경망 칩)

(1) NPU(신경망 칩) 개요

(2) NPU(신경망 칩) 특징

1-3-2. 지능형 반도체(PIM)

(1) 지능형 반도체(PIM) 개요

(2) 지능형 반도체(PIM) 특징

1-4. 인간 뇌를 모방한 뉴로모픽(Neuromorphic) 기술

1-4-1. 뉴로모픽(Neuromorphic) 반도체 개념

1-4-2. 뉴로모픽(Neuromorphic) 반도체 등장 배경

(1) 폰노이만 구조의 한계

(2) 뉴로모픽 반도체 등장

(3) 폰노이만과 뉴로모픽 아키텍처의 비교

1-4-3. 뉴로모픽 반도체의 구조

1-4-4. 뉴로모픽 반도체의 작동 방식

1-4-5. 뉴로모픽 반도체의 장단점

1-4-6. 뉴로모픽 컴퓨팅(Neuromorphic Computing)

1-4-7. 뉴로모픽 컴퓨팅(Neuromorphic computing) 특징

1-4-8. 뉴로모픽(Neuromorphic) 컴퓨팅 파급효과

1-4-9. 뉴로모픽 반도체 개발 동향

(1) 해외 뉴로모픽 반도체 개발 현황

가. 인텔(Intel)

나. 퀄컴(Qualcomm)

다. IBM

(2) 국내 뉴로모픽 반도체 개발 현황

1-4-10. 뉴로모픽 컴퓨팅 향후 전망 및 해결 과제

(1) 뉴로모픽 컴퓨팅 파급효과 및 향후 전망

(2) 뉴로모픽 칩 구현을 위한 해결 과제

2. 온디바이스 AI 응용 분야 및 국내외 정책 현황

2-1. 온디바이스 AI 생태계

2-1-1. 온디바이스 AI 시대

2-1-2. 온디바이스 AI 의 미래

2-2. 온디바이스 AI 응용 분야

2-2-1. 스마트폰

2-2-2. 자율주행 자동차

2-2-3. 의료 산업

2-3. 국내외 산업 현황

2-3-1. Google

2-3-2. Apple

2-3-3. 인텔

2-3-4. 퀄컴(Qualcomm)

2-3-5. 삼성전자

2-4. 온디바이스 AI 시장 전망

2-4-1. 시장 성장 요인

2-4-2. 시장 전망

2-4-3. 시장 구현을 위한 온디바이스 AI 의 과제

참고문헌

그림목차

[그림 1] 클라우드 기반 시스템

[그림 2] 클라우드 AI 의 이점

[그림 3] AI 와 클라우드 컴퓨팅의 결합

[그림 4] 로봇의 구성 요소

[그림 5] AI가 사회에 미치는 영향

[그림 6] 로봇 프로세스 자동화 도구

[그림 7] 노동시장에서 인간-AI 공존

[그림 8] 로봇공학 혁명

[그림 9] AI 기술 환경

[그림 10] 인간-로봇 간의 협업 응용 분야

[그림 11] 제조업에서의 인간-로봇 협업 프레임워크

[그림 12] AI와 로봇공학을 기반으로 한 헬스케어 시스템 분류

[그림 13] 농업용 로봇의 구성 요소

[그림 14] 인간-로봇 간의 상호 인지적이고 공감적인 협업 방식

[그림 15] 로봇과의 협업 유형

[그림 16] 인간-로봇 상호작용

[그림 17] HRC 시스템의 구조 및 구성 요소

[그림 18] 인간-로봇 관계

[그림 19] 인간-로봇 협업 패러다임

[그림 20] 인간 컴퓨터 상호작용

[그림 21] 주요 산업의 인공지능 사용 사례

[그림 22] 로봇과 AI에 대한 다양한 차원의 안전

[그림 23] 로봇의 자율성을 위한 딥러닝 접근 방식

[그림 24] 휴머노이드 로봇의 자유 이동 모드와 참여 모드

[그림 25] 상호작용 동작의 개략도

[그림 26] 지능형 자동화(IA)와 인공지능(AI)

[그림 27] 인간-로봇 협업 시스템의 구조적 구성 요소

[그림 28] 최근 주목해야 할 AI 트렌드

[그림 29] 생성형 AI(Generative AI)

[그림 30] 생성형 AI(Generative AI)

[그림 31] 인공지능(AI)의 잠재력

[그림 32] AI의 환각이 작동하는 방식

[그림 33] 2023년 AI 파운데이션(기초 모델) 개발 기업·국가 상위 1~4위

[그림 34] 생성형 AI의 환각 사례

[그림 35] LLM 환각 유형

[그림 36] RAG 아키텍처

[그림 37] 인간 중심의 AI 개발 접근 방식

[그림 38] 로봇공학을 둘러싼 윤리적 고려 사항

[그림 39] 윤리적인 AI 프레임워크

[그림 40] 책임감 있고 신뢰할 수 있는 AI

[그림 41] 신뢰할 수 있는 AI 기준

[그림 42] 설명가능한 AI의 특징

[그림 43] 대화형 AI

[그림 44] 생성형 AI 애플리케이션

[그림 45] 인간 중심 AI의 균형

[그림 46] 미래 AI 주요 활용 사례

[그림 47] 클라우드 AI vs. 온디바이스 AI

[그림 48] AI 파이프라인

[그림 49] 온디바이스 AI의 이점

[그림 50] 온디바이스 AI processing

[그림 51] 하이브리드 AI

[그림 52] 스마트폰에 도입된 온디바이스 AI 추론 과정

[그림 53] 데이터와 AI 기술 계층

[그림 54] 머신러닝(ML) 프로세스

[그림 55] 머신러닝 아키텍처

[그림 56] 온디바이스 AI 의 작동 원리

[그림 57] 온디바이스 AI 주요 이점

[그림 58] 온디바이스 AI 작동 방식

[그림 59] 인공지능(AI)의 보안 위협

[그림 60] 네트워크 지연의 원인

[그림 61] 클라우드에서 엣지로 이동하는 AI

[그림 62] 인공지능의 작동 방식

[그림 63] 엣지에서의 데이터 융합과 AI 를 위한 프레임워크

[그림 64] 반도체 산업 주요 사용 사례

[그림 65] AI 반도체의 구분과 특징

[그림 66] 세대별 반도체 분류

[그림 67] NPU 아키텍처

[그림 68] 지능형 반도체(PIM) 개념

[그림 69] 기존 폰노이만 구조 vs. PIM 구조

[그림 70] AI 모델의 진화

[그림 71] 시냅스에서 정보가 전달되는 과정

[그림 72] 폰노이만 아키텍처

[그림 73] 반도체 프로세싱 유닛의 발전방향

[그림 74] 폰노이만과 뉴로모픽 아키텍처의 비교

[그림 75] 인간의 뇌 신경시스템의 계층 구조와 인공신경망(ANN)

[그림 76] 인간 뇌를 모방한 뉴로모픽 반도체

[그림 77] 인간 두뇌 신경망

[그림 78] 뉴로모픽 컴퓨팅

[그림 79] 뉴로모픽 기술

[그림 80] 뉴로모픽 컴퓨팅 파급 효과

[그림 81] 뉴로모픽 컴퓨터 접근 알고리즘

- [그림 82] AI 반도체 구분
- [그림 83] 퀄컴(Qualcomm) AI 연구 방향
- [그림 84] 노스폴(NorthPole) 아키텍처의 일부
- [그림 85] 차세대 뉴로모픽 반도체
- [그림 86] 뉴로모픽 컴퓨팅 접근 알고리즘
- [그림 87] 인공지능 생태계
- [그림 88] 온디바이스 AI 주요 장점
- [그림 89] 글로벌 AI PC 지원
- [그림 90] 온디바이스 AI 활용 비즈니스 생태계
- [그림 91] 전통적인 AI 처리 방식
- [그림 92] 에지 AI 처리 방식
- [그림 93] 온디바이스 AI 활용 사례
- [그림 94] 하이브리드 AI
- [그림 95] Google AI Edge SDK, AICore 및 제미나이 나노(Gemini Nano)
- [그림 96] 애플의 지능화 전략
- [그림 97] Apple 기초 모델에 대한 모델링 개요
- [그림 98] 인텔 AI 에브리웨어 스택
- [그림 99] 퀄컴(Qualcomm)의 풀스택 AI 최적화
- [그림 100] 삼성전자 AI 솔루션
- [그림 101] 모바일 AI 시장 규모
- [그림 102] 대화형 AI 시장 성장 요인
- [그림 103] 글로벌 엣지 AI 시장 규모
- [그림 104] 온디바이스 학습을 위한 연합학습

표 목차

[표 1] 데이터 거래시 애로사항(단위: %, 복수응답, 데이터 구매 경험 기업)

[표 2] AI 주도권 확보 위한 주요 빅테크의 전략

[표 3] 온디바이스 AI와 클라우드서버 AI 장단점 비교

[표 4] 온디바이스 AI의 소프트웨어 기술

[표 5] 주요 테크업체들의 LLM 및 온디바이스 AI 준비 현황

[표 6] 뉴로모픽 반도체와 기존 반도체 비교

[표 7] 주요 기업 뉴로모픽 특허 현황

[표 8] 인공지능 1~3세대별 기능과 특성

[표 9] AI 기술 트렌드 변화

[표 10] 각 기업의 온디바이스 AI 활용 사례