AI(Artificial Intelligence)

Definition

AI(인공지능, Artificial Intelligence)는 인간의 학습, 추론, 지각, 문제 해결 능력 등을 컴퓨터 과학 기술로 구현하여 기계가 지능적인 행동을 하도록 만드는 기술로 데이터 기반의 머신러닝(기계학습)과 딥러닝을 통해 인간처럼 배우고 판단하며, 창의적인 콘텐츠를 생성하거나 데이터를 예측하는 등 다양한 분야에서 활용되는 핵심 기술

종류

AI는 학습 및 기능에 따라 데이터를 분류하는 분류형 AI와 새로운 콘텐츠를 만들어내는 생성형 AI(Generative AI)로 나뉨. 크게 머신러닝/딥러닝 기술, 데이터 분석 기반의 모델, 텍스트·이미지·영상 등을 생성하는 서비스형 모델로 구분할 수 있음.

1. 생성 방식 및 기능에 따른 분류
    - 생성형 AI (Generative AI): 대규모 데이터를 학습하여 텍스트, 이미지, 오디오, 코드 등 새로운 콘텐츠를 생성합니다. (예: ChatGPT, Midjourney, Claude)
    - 분류형/구분형 AI (Discriminative AI): 데이터를 바탕으로 패턴을 인식하고 분류합니다. (예: 스팸 메일 분류, 이미지 객체 탐지)
2. 기술적 구현 수준에 따른 분류
    - 약인공지능 (Narrow AI): 특정 분야의 작업(예: 바둑, 음성인식)만 수행 가능하며 현재 수준의 AI.
    - 강인공지능 (General AI/AGI): 인간과 동등한 지능을 가지고 스스로 학습 및 해결할 수 있는 수준 (미래 목표).
3. 데이터 학습 방식에 따른 분류
    - 머신러닝 (Machine Learning): 컴퓨터가 데이터를 통해 스스로 학습하고 패턴을 찾아내는 방식.
    - 딥러닝 (Deep Learning): 인간의 뇌 구조(인공신경망)를 모방하여 복잡한 데이터를 처리하는 머신러닝의 하위 집합.
4. 생성형 AI의 유형
    - 언어 모델 (LLM): 텍스트 생성, 번역, 요약 (예: GPT-4, Llama).
    - 이미지 모델: 텍스트를 이미지로 변환 (예: DALL·E 3, Stable Diffusion).
    - 영상/음성 모델: 비디오 생성 및 음성 합성.

로컬LLM

> Qwen, DeepSeek, Codestral (Mistral AI), Phi-4 (Microsoft), StarCoder2, Ollama, Gemma, Llama, Exaone, ...

MCP(Model Context Protocol)

> MCP(Model Context Protocol)는 앤트로픽(Anthropic)이 개발한 개방형 프로토콜로, LLM(거대 언어 모델)과 외부 데이터/도구(API, 데이터베이스 등)를 표준화된 방식으로 연결해 주는 기술. “AI 시대의 USB-

> C 포트"로 불리며, 서로 다른 AI 모델과 애플리케이션을 손쉽게 연동하여 실시간 데이터 접근 및 작업을 자동화하는 데 사용.

핵심 특징 및 구성

- 표준화된 연결: 파편화된 외부 도구 연결 방식을 하나로 통합하여 편리성 향상.
- 주요 구성 요소:
  - MCP 호스트: 사용자가 명령을 입력하는 도구(예: 클로드, 커서).
  - MCP 클라이언트: 호스트와 서버를 연결하는 브릿지.
  - MCP 서버: 특정 데이터 소스나 도구(예: 노션, 피그마)를 외부와 연동.
- 장점: 플러그 앤 플레이(Plug-and-Play) 방식으로 새로운 도구 연결 시 코딩 작업이 최소화됨.

MCP의 필요성

기본적인 LLM은 학습된 데이터에만 의존하여 정적인 정보를 제공하지만, MCP를 통해 실시간 데이터와 외부 시스템(회의 예약, 데이터베이스 업데이트)을 안전하게 연결하여 더 유능한 AI 에이전트를 구축할 수 있음.

AI Agent

> AI 에이전트(AI Agent)는 대규모 언어 모델(LLM)을 기반으로, 사람의 개입 없이 자율적으로 목표를 설정하고 계획을 수립하여 복잡한 작업을 수행하는 지능형 소프트웨어 시스템. 단순한 챗봇을 넘어 API, 검색 엔진 등 외부 도구를 활용해 실시간 환경과 상호작용하며 문제를 해결하는 능동적인 인공지능.

AI 에이전트의 핵심 특징 및 구성 요소

- 자율성 (Autonomy): 목표 달성을 위해 스스로 판단하고 행동함.
- 추론 및 계획 (Reasoning & Planning): 복잡한 문제를 분해하고 단계적인 계획을 세움.
- 기억 (Memory): 이전 상호작용과 데이터를 기억하여 맥락을 유지함.
- 도구 사용 (Tool Use): 웹 검색, 코드 실행, 외부 API 연동 등 실제 행동을 수행함.

기존 AI(챗봇)와의 차이점

특징	기존 AI/챗봇	AI 에이전트
작동 방식	수동적 (질문에 답변)	능동적/자율적 (목표 달성)
작업 범위	단발성, 정보 제공 위주	복잡한 다단계 워크플로 수행
도구 연동	제한적	외부 API/도구 활용 가능

주요 활용 사례

- 개인 비서: 일정 관리, 이메일 작성, 여행 예약.
- 업무 자동화: 고객 상담, 데이터 분석, 마케팅 캠페인 실행.
- 소프트웨어 개발: 코드 작성, 버그 수정, 테스트 자동화.

AI 에이전트는 향후 멀티모달 기술과 융합되어 텍스트, 음성, 이미지를 통합 처리하며 산업 전반의 운영 효

율성을 크게 향상시킬 것으로 기대됨.

RAG(Retrieval-Augmented Generation, 검색 증강 생성), 기타 등등.

> RAG(Retrieval-Augmented Generation, 검색 증강 생성)는 LLM(거대 언어 모델)이 답변을 생성하기 전, 외부 지식 베이스(데이터베이스)에서 관련 정보를 검색하여 참고하게 함으로써 정확성과 최신성을 높이는 AI 기술입니다. 모델의 내부 학습 데이터에 의존하지 않고 신뢰할 수 있는 외부 데이터를 활용해 할루시네이션(거짓 답변)을 줄이고 최신 정보나 특정 기업 데이터를 AI가 답변에 반영할 수 있게 합니다.

RAG 핵심 특징 및 장점

- 할루시네이션 방지: 외부 문서에 기반한 정확한 정보 제공으로 AI가 모르는 것을 지어내는 현상을 줄임.
- 최신 정보 반영: 모델을 재학습시키지 않고 데이터베이스만 업데이트하여 실시간 정보 반영이 가능함.
- 보안 및 맞춤화: 기업 내부 문서, 개인화된 데이터 등을 기반으로 맞춤형 챗봇 및 검색 엔진 구현이 가능함.
- 비용 효율성: 대규모 언어 모델을 전체 재학습(Fine-tuning)하는 것보다 훨씬 저렴하게 정확도를 높일 수 있음.

RAG의 작동 방식 (4단계)

1. 데이터 임베딩 (Indexing): 사내 문서, 데이터베이스 등 외부 지식을 숫자 형태(벡터)로 변환하여 벡터 데이터베이스에 저장함 [13, www.ncloud-forums.com].
2. 질문 검색 (Retrieval): 사용자의 질문과 가장 관련성 높은 문서를 데이터베이스에서 검색함 [2, YouTube].
3. 프롬프트 증강 (Augmentation): 사용자 질문에 검색된 문서를 추가하여 AI 모델에게 전달할 프롬프트를 보강함 [2, YouTube].
4. 답변 생성 (Generation): 모델이 제공된 외부 정보를 바탕으로 정확한 답변을 생성함 [1, Airbridge.io].

RAG 활용 사례

- 기업용 AI 챗봇: 사내 규정, 매뉴얼을 학습한 상담 봇.
- 전문 분야 검색: 법률 문서 분석, 의료 진단 보조.
- 지식 관리 시스템: 방대한 문서에서 핵심 정보 요약 및 검색.

RAG는 LLM의 창의성에 신뢰할 수 있는 데이터를 더해 실제 비즈니스 환경에서 AI를 유용하게 활용할 수 있도록 만드는 핵심 기술 [1, 4, Airbridge.io].

References

Tuesday, January 27, 2026

- 100일 동안 총 22개의 AI를 써봤습니다 | AI로 기획 분석 마케팅 디자인 자동화 개발까지 - YouTube @www.youtube.com



  - 분야별 분류
    - 기획 : perplexity, chatgpt, gemini, grok
    - 분석: Gemini Deep Research
    - 글쓰기; Claude, Gemini
    - 이미지 생성; Nanobanana, Imagen, Midjourney
    - 영상생성; Veo, Hailuo, Kling, Grok Imagine, Sora
    - 음원생성; Suno AI
    - 자동화; n8n, make, Opal, (+Chatgpt AI Agent Kit)
    - 디자인; Figma
    - 개발; Copilot, Cursor, a0, v0, replit
  - 사이트
    - Perpelexity Perplexity @www.perplexity.ai
    - ChatGPT ChatGPT @chatgpt.com
    - Gemini Google Gemini @gemini.google.com
    - Grok Grok @grok.com
    - Claude Claude @claude.ai
    - Nanobanana Google AI Studio @aistudio.google.com
    - Imagen Google AI Studio @aistudio.google.com
    - Midjourney Midjourney @www.midjourney.com
    - Veo Google AI Studio @aistudio.google.com
    - Hailuo Hailuo AI: AI Video Generator from Text & Image @hailuoai.video
    - Kling Kling AI: Next-Gen AI Video & AI Image Generator @app.klingai.com
    - Grok Imagine Grok Imagine @grok.com
    - Sora Sora | OpenAI @openai.com
    - Suno Suno | AI Music @suno.com
    - n8n AI Workflow Automation Platform & Tools - n8n @n8n.io
    - make Make | AI Workflow Automation Software & Tools | Make @www.make.com
    - opal Opal [Experiment] @opal.google
    - figma Figma | 협업을 위한 인터페이스 디자인 도구 @www.figma.com
    - copilot GitHub Copilot · Your AI pair programmer · GitHub @github.com
    - cursor Cursor @cursor.com
    - a0 a0.dev - Ship Mobile Apps to App Store & Google Play with AI @a0.dev
    - v0 v0 by Vercel - Build Agents, Apps, and Websites with AI @v0.app
    - replit Replit – Build apps and sites with AI - Replit @replit.com
    - (+ Canva 모두가 사용할 수 있는 Canva 비주얼 스위트 @www.canva.com)

    - (+ Banana Pro AI Banana Pro AI: The Best Free AI Image Generator Online @bananaproai.com)
    - (+ NotebookLM Google NotebookLM | AI 조사 도구 및 사고 파트너 @notebooklm.google)
- 2025년 개발자가 직접 써본 코딩용 로컬 LLM 5가지 추천 - 피카부랩스 블로그 @peekaboolabs.ai
- 다모아 AI - AI 사이트 모음 @www.damoa.ai

From:
https://theta5912.net/ - **reth**

Permanent link:
https://theta5912.net/doku.php?id=public:computer:ai_artificial_intelligence

Last update: **2026/01/27 22:27**