

KDT

LLM과 Model Context Protocol (MCP)

- LLM과 MCP

본 문서는 SK AX의 콘텐츠 자산으로, 무단 사용 및 불법 배포 시 법적 조치를 받을 수 있습니다.

MCP (Model Context Protocol)

개념

- 개방형 표준 통신 규약
- LLM이 외부 다양한 데이터 소스와 도구들과 안전하고 효율적 소통을 할 수 있도록 만들어진 규약
- AI 모델을 위한 통합 USB Port 역할

MCP 역할

- 환각(Hallucination) 감소
→ 학습 데이터에 없는 최신 정보, 정확한 정보를 외부데이터베이스에서 가져와 더 신뢰성 있는 답변을 생성
- 실제 행동 수행
→ 단순 답변 외 실제 작업까지 진행 (eg, 일정/회의실 예약까지 병행처리)
- 개발 효율성 증가
→ **MCP** 하나로 어떤 **AI** 모델이든 재사용 가능 (개발시간 감소)

MCP (Model Context Protocol)

【 컨텍스트(Context)의 확장 】

- 기존 LLM은 학습 데이터에만 의존하여 답변을 생성
→ 최신 정보나 개인화된 정보에는 접근 불가
- 이런 한계 극복을 위해 외부의 실시간 데이터(예: 주식 시장 데이터, 최신 뉴스)나 개인 데이터(예: 사용자의 캘린더, 이메일)를 '컨텍스트'로 가져와 모델에게 제공
- 이를 통해 모델은 훨씬 더 풍부하고 정확한 정보를 바탕으로 답변 및 작업까지 가능

【 도구(Tool) 사용의 표준화 】

- AI 모델이 단순한 답변을 넘어, 실제 행동을 하려면 외부 도구(API, 프로그램 등)를 호출
→ "오늘 회의 일정 잡아줘"
. 캘린더 API 사용
- 각 도구의 API 규격에 맞춰 복잡한 코드를 따로 작성해야 했으나
- MCP는 '도구 사용'에 대한 통신 규격을 표준화
- MCP를 통해 어떤 도구든 표준화된 방식으로 호출하고, 결과 확인 가능

MCP (Model Context Protocol)

【 MCP의 작동 방식: 클라이언트-서버 구조 】

MCP 호스트 (Host)

- AI 모델을 활용하는 애플리케이션
(예: AI 비서 앱, IDE의 AI 기능 등)
- 사용자의 명령을 받고, AI 모델의 추론을 바탕으로 MCP 클라이언트에게 요청

MCP 클라이언트 (Client)

- 호스트 내부에 존재하며, MCP 서버와 1대1로 연결을 관리 역할
- 호스트의 요청을 표준화된 MCP 프로토콜에 맞춰 서버로 보내고, 서버의 응답을 받아 다시 호스트에게 전달

MCP 서버 (Server)

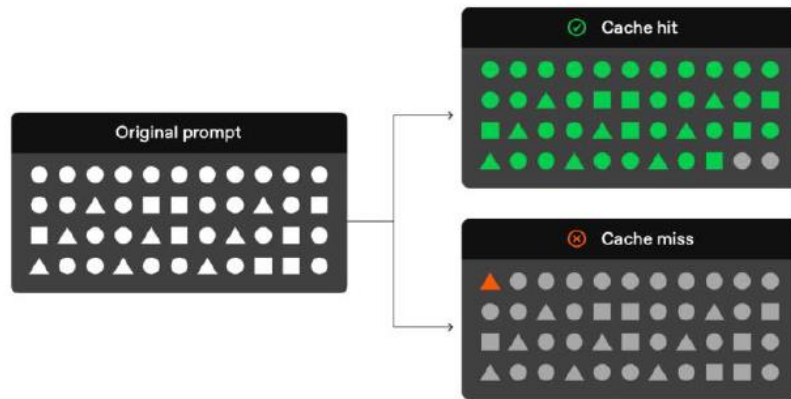
- 특정 기능이나 데이터를 제공하는 외부 시스템
 - 캘린더 서버, 파일시스템 서버, DB서버
- 클라이언트로부터 요청을 받으면 해당 기능을 수행하고, 그 결과를 클라이언트에게 다시 전달

“범용 리모컨”을 가진 것처럼
MCP 서버라는
“다양한 기기”들을 쉽게 제어”

- 개발자 입장에서는 AI 모델에 새로운 기능을 추가할 때마다 복잡한 연동 코드를 만들 필요 없이, MCP 서버만 구축하면 되므로 개발 효율성이 크게 향상

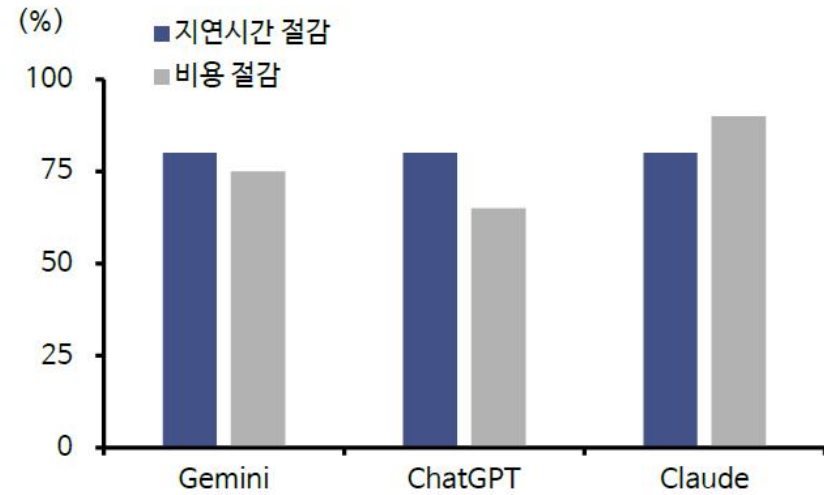
MCP (Model Context Protocol)

프롬프트 캐싱 작동 원리



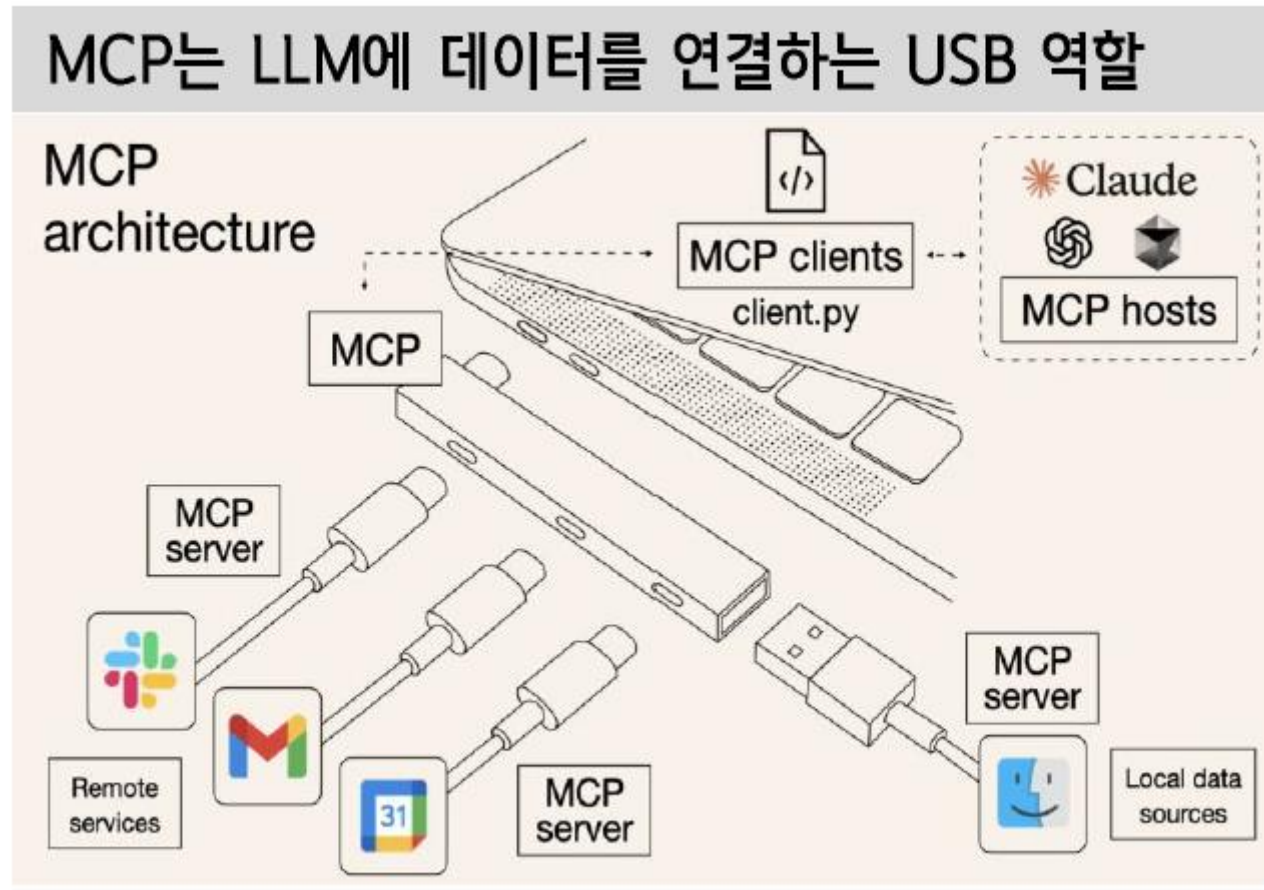
자료: 회사 자료, 신한투자증권

프롬프트 캐싱 적용 효과



자료: 회사 자료, 신한투자증권

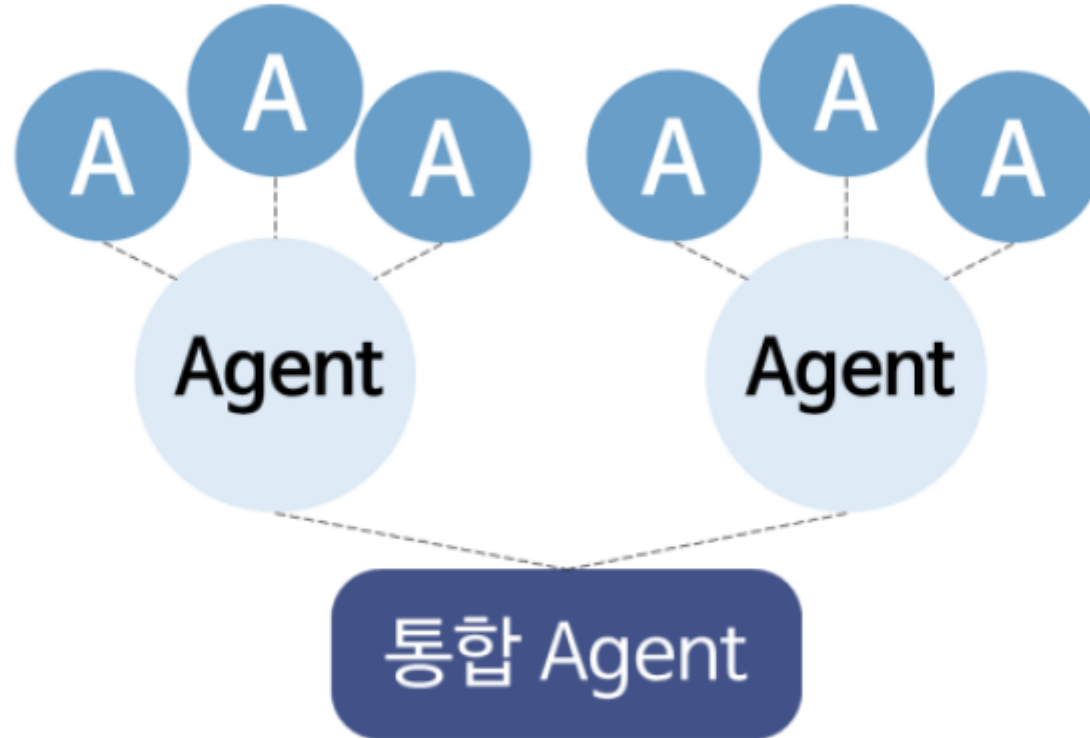
MCP (Model Context Protocol)



자료: 엔트로픽, 신한투자증권

MCP (Model Context Protocol)

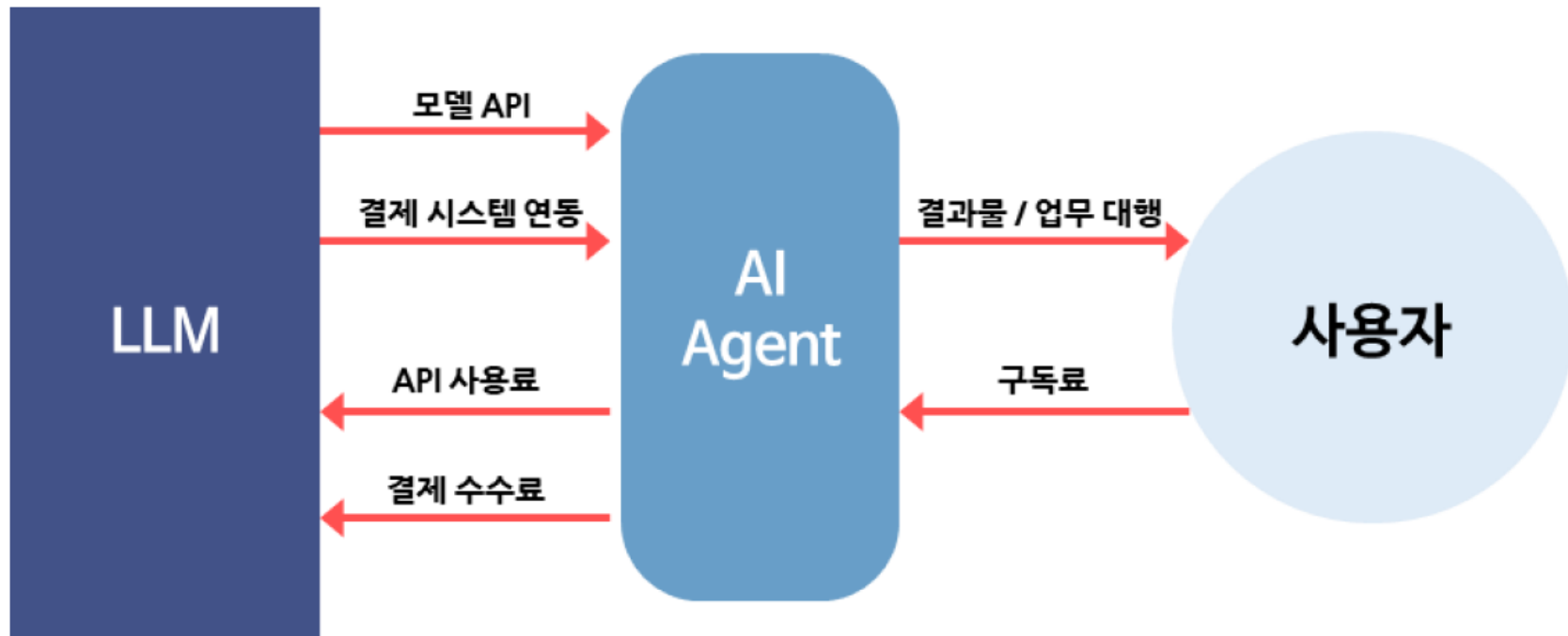
구글의 A2A(Agent to Agent) 규격



자료: 신한투자증권

MCP (Model Context Protocol)

LLM 기반의 AI Agent 수익 구조



자료: 신한투자증권

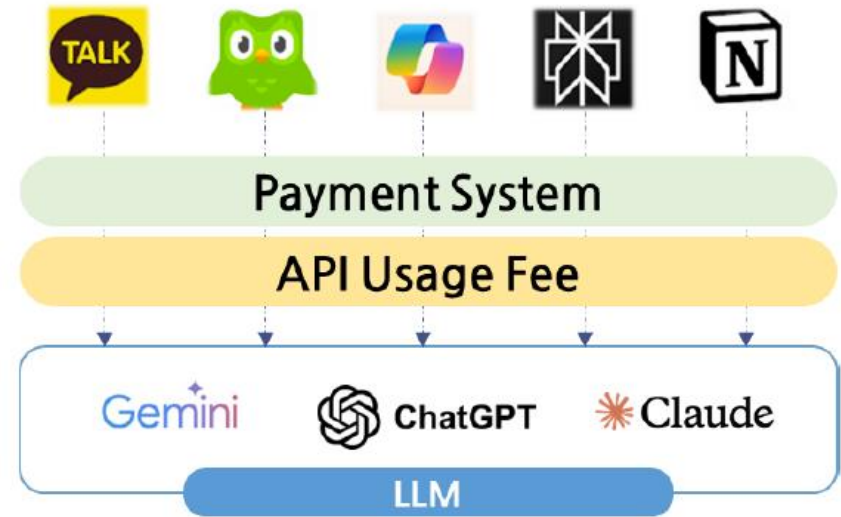
MCP (Model Context Protocol)

모바일 시대 구글/애플의 수수료 수익 구조



자료: 신한투자증권

AI 시대 LLM들의 수수료 수익 구조



자료: 신한투자증권