MCP로 데이터 엔지니어 업무 자동화하기

이왕원

발표자 소개

· 패션 AI 스타트업 코디미 CTO

· 경영, 프로젝트 관리 및 디렉팅, AI 연구, 데이터 수집 및 정제, 서비스 및 기능 기획, 웹개발, 마케팅 등 다양한 업무를 함

· 몸이 10개라도 모자란 상황이라 자동화/효율화에 관심이 많음

· 근본은 AI 엔지니어. 최신 AI 기술이나 흐름을 항상 주의깊게 보는 중

발표 흐름 선공유

- 최근 핫한 MCP를 현업에서 직접 적용해 본 과정을 공유
- 단계별 개선 과정과 현 시점의 현실적인 한계점을 공유
- 단순 개념 설명이 아닌 트러블 슈팅 위주의 경험 공유

Presentation Overview

개요


왕원님, 이번 달 신규가입자 몇 명인지 DB 조회해서 알려주세요.
이거 자주 필요할 것 같은데, 이드민 페이지 기능으로 추가해주세요!


누가 나 대신 해줬으면 좋겠다

개요

요즘 핫한 MCP 써서 한 번 만들어보자

목표

1. 간단한 DB 조회 및 집계를 대신 해주는 AI 봇 만들기
2. 비개발 직군이 쓸 수 있을 정도로 쉬워야 함
3. 시간 투자 많이 안하고 만들기

과정

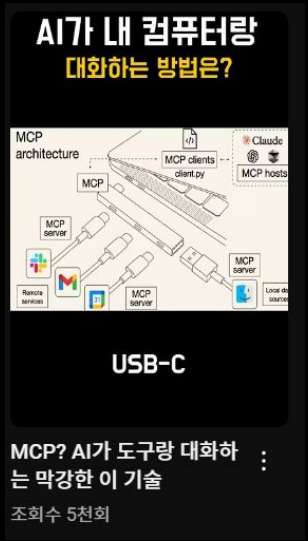
Gradio 에이전트 제작

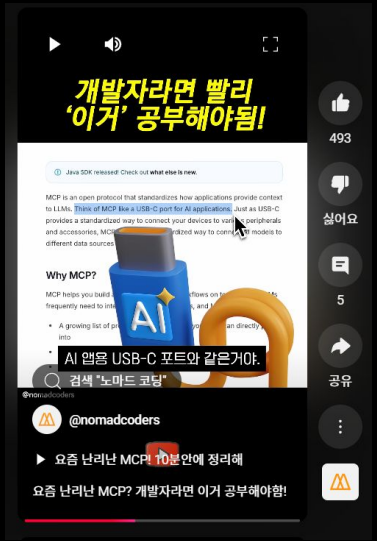
```python
with gr.Tab("SQL DB MCP Agent (개발 중)"):
    ### TODO: Implement this
    query_text = gr.Textbox(label="질의문", placeholder="예
    ask_button = gr.Button("물어보기")
    output_status_message = gr.Text(label="상태 메시지")
    output_message = gr.Text(label="결과")

    ask_button.click(
        ask_to_sql_mcp,
        inputs=query_text,
        outputs=[output_status_message, output_message]
    )
```

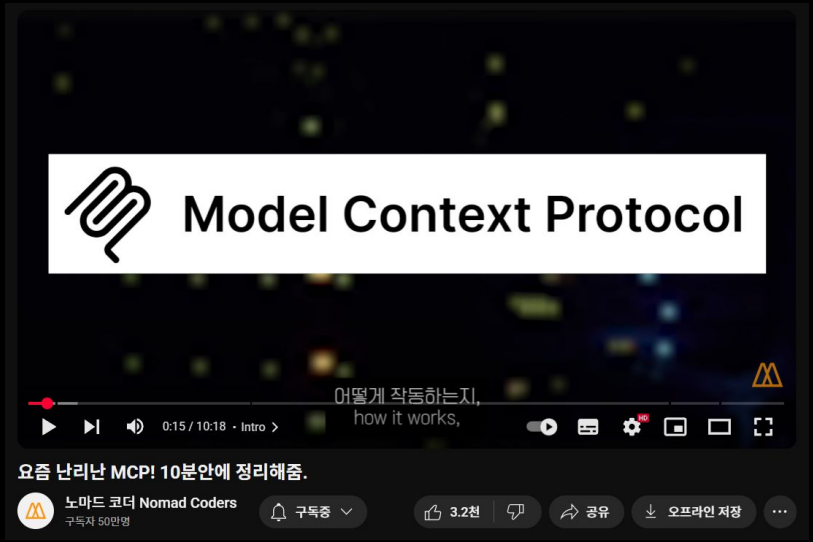
과정

MCP 공부

Source : MCP 검색 시 나오는 유튜브 영상들 캡쳐





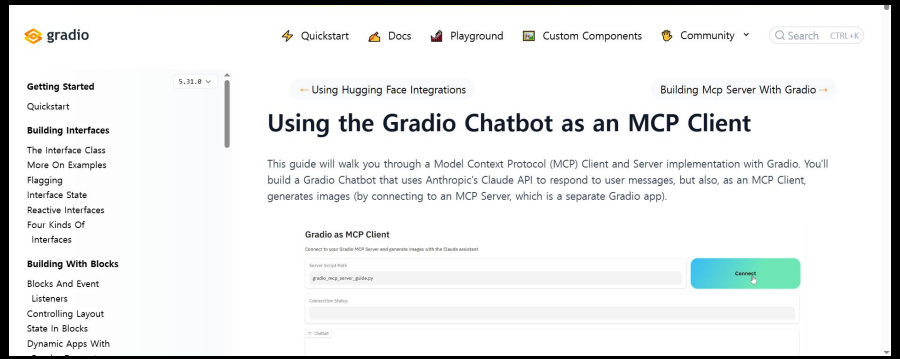


유튜브에 다양한 설명이 있으나, 쉽게 이해가 되진 않았음 ㅠ

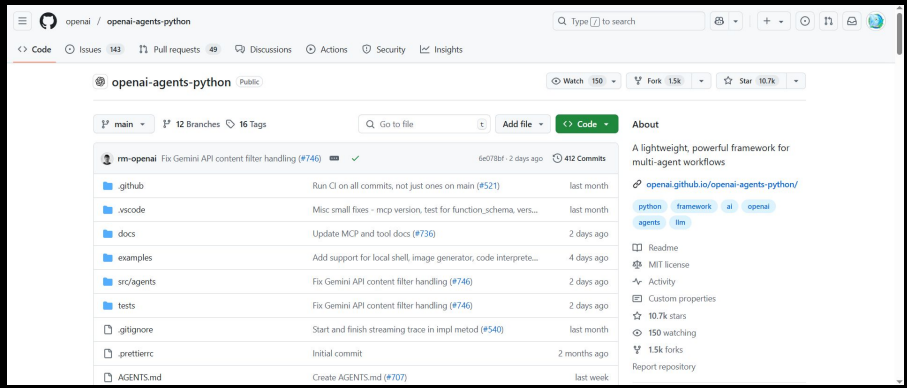
과정

MCP 공부

Source : Gradio 공식 MCP example



Source : OpenAI-agents-python 깃헙 레포



직접 example 돌려보고 로그 찍어보면서 이해하는 게 가장 도움이 되었음

과정

프로토타입 제작

```python
import asyncio
import os
from agents import Agent, Runner
from agents.mcp import MCPServerStdio
from dotenv import load_dotenv

load_dotenv()


def init_mcp_server():
    mcp_server_base_00 = MCPServerStdio(
        cache_tools_list=True,
        params={
            "command": "uvx",
            "args": ["--from", "mysql-mcp-server", "mysql_mcp_server"],
            "env": {
                "MYSQL_HOST": os.getenv('MYSQL_HOST'),
                "MYSQL_PORT": os.getenv('MYSQL_PORT'),
                "MYSQL_USER": os.getenv('MYSQL_USER'),
                "MYSQL_PASSWORD": os.getenv('MYSQL_PASSWORD'),
                "MYSQL_DATABASE": os.getenv('MYSQL_DATABASE'),
            }
        }
    )

    return mcp_server_base_00
```

```python
def run_query(message):
    async def async_run(message):
        async with init_mcp_server() as mcp_server_00:
            mcp_agent = Agent(
                name="MCP Assistant",
                instructions="You are a helpful assistant with access to MCP tools.",
                mcp_servers=[mcp_server_00],
            )

            result = await Runner.run(
                starting_agent=mcp_agent,
                input=message,
            )
        return result.final_output

    return asyncio.run(async_run(message))
```

바퀴를 2번 발명할 필요는 없으니, uvx로 괜찮은 오픈소스 MCP 서버를 가져옴

이런 식으로 아주 빠르게 프로토타입을 완성

과정

문제점 발생 1

```
User: read_user
Database: Potoo_Service
2025-05-25 16:16:51,155 - mysql_mcp_server - INFO - Starting MySQL MCP server...
2025-05-25 16:16:51,155 - mysql_mcp_server - INFO - Database config: potoo-datab
ase.cp3ompmjgggf.ap-northeast-2.rds.amazonaws.com/Potoo_Service as read_user
2025-05-25 16:16:51,165 - mcp.server.lowlevel.server - INFO - Processing request
 of type ListToolsRequest
2025-05-25 16:16:51,165 - mysql_mcp_server - INFO - Listing tools...
2025-05-25 16:16:53,747 - mcp.server.lowlevel.server - INFO - Processing request
 of type CallToolRequest
2025-05-25 16:16:53,748 - mysql_mcp_server - INFO - Calling tool: execute_sql wi
th arguments: {'query': "SELECT user_id FROM users WHERE name = '이왕원';"}
2025-05-25 16:16:54,119 - mysql_mcp_server - ERROR - Error executing SQL 'SELECT
 user_id FROM users WHERE name = '이왕원';': 1146 (42S02): Table 'Potoo_Service.
users' doesn't exist
```

- 실제 유저 정보 테이블의 스키마 예시 -

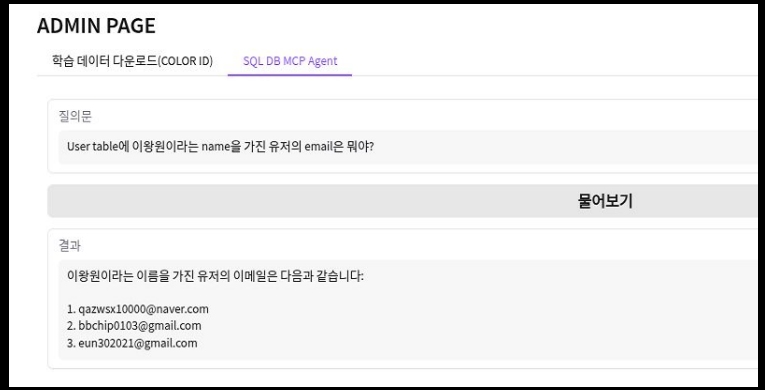
```
1  CREATE TABLE `User` (
2      `userId` varchar(64) NOT NULL DEFAULT (uuid()),
3      `email` varchar(50) NOT NULL DEFAULT '',
4      [redacted]
5      `name` varchar(30) NOT NULL,
```

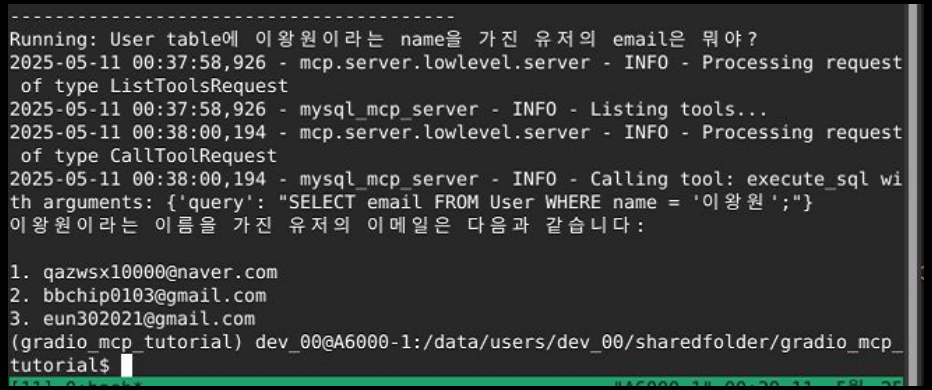
“이름이 이왕원인 유저의 user_id를 알려줘”라는 질의했더니,

존재하지 않는 테이블(users)과 속성(user_id)을 대상으로 쿼리를 날림

과정

1차 해결 방안(?)





"User table에 이왕원이라는 name을 가진 유저의 userId를 알려줘"처럼,

테이블과 속성을 명시해서 물어보면 잘 동작

과정

한계점

1. 이대로면 직접 쿼리 짜서 날리는 것과 다를 것이 없음
2. DB에 대한 이해 없이 AI가 단순히 연결만 되어 있는 상황
3. 비개발 직군이 쓸 수 있을 정도로 쉽지 않음

과정

2차 해결 방안

```
message = '내 질문에 대답하기 전에 우리 SQL DB 구조를 먼저 파악하고, 이를 기반으로 대답해 줘.\n' + m

result = await Runner.run(
    starting_agent=mcp_agent,
```

컨텍스트에 관련된 문제이기 때문에, AI가 그 맥락을 파악할 수 있으면 됨

전체적인 DB 스키마 구조를 파악할 수 있도록 사전 프롬프트 추가

과정

2차 해결 방안

드디어 원했던 형태로 동작

과정

2차 해결 방안

```
2025-05-24 19:07:34,816 - mysql_mcp_server - INFO - Listing tools...
2025-05-24 19:07:36,675 - mcp.server.lowlevel.server - INFO - Processing request
 of type CallToolRequest
2025-05-24 19:07:36,675 - mysql_mcp_server - INFO - Calling tool: execute_sql wi
th arguments: {'query': 'SHOW TABLES;'}
2025-05-24 19:07:38,609 - mcp.server.lowlevel.server - INFO - Processing request
 of type CallToolRequest
2025-05-24 19:07:38,609 - mysql_mcp_server - INFO - Calling tool: execute_sql wi
th arguments: {'query': 'DESCRIBE User;'}
2025-05-24 19:07:40,248 - mcp.server.lowlevel.server - INFO - Processing request
 of type CallToolRequest
2025-05-24 19:07:40,248 - mysql_mcp_server - INFO - Calling tool: execute_sql wi
th arguments: {'query': "SELECT email FROM User WHERE name = '이왕원';"}
```

로그를 보니, AI가 테이블 리스트와 스키마 구조를 파악하고 올바른 쿼리를 작성

과정

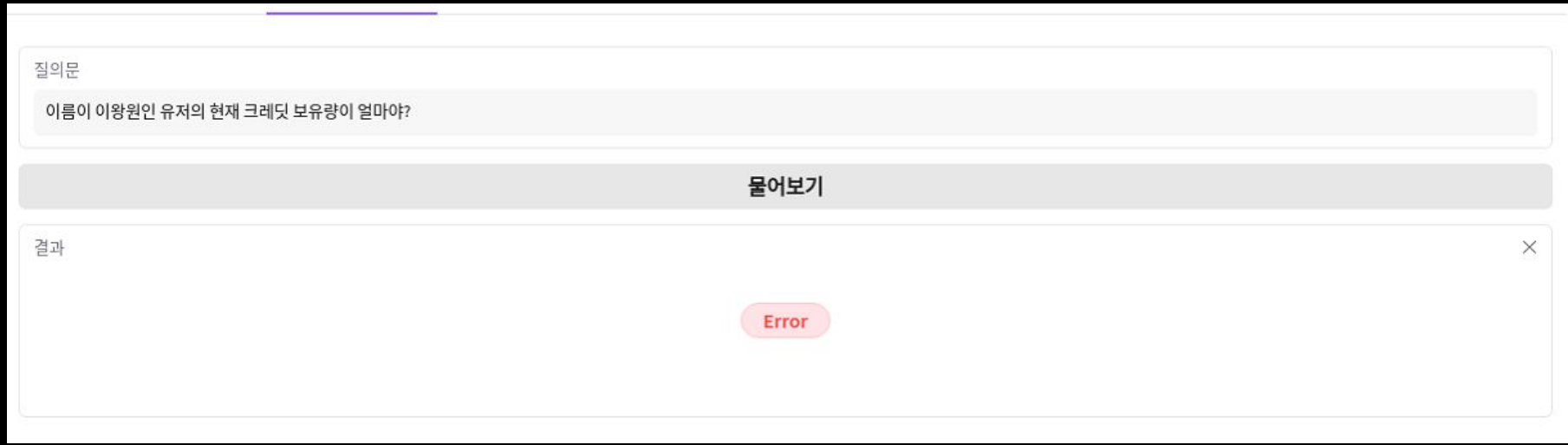
추가 테스트

동시에 여러 가지를 물어봐도 잘 정리해서 답변해 줌

여러번 물어보면서 테스트 해봐도 오류 없이 잘 동작

과정

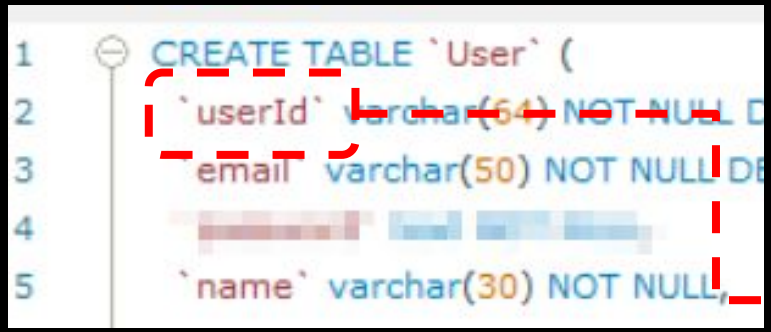
추가 테스트

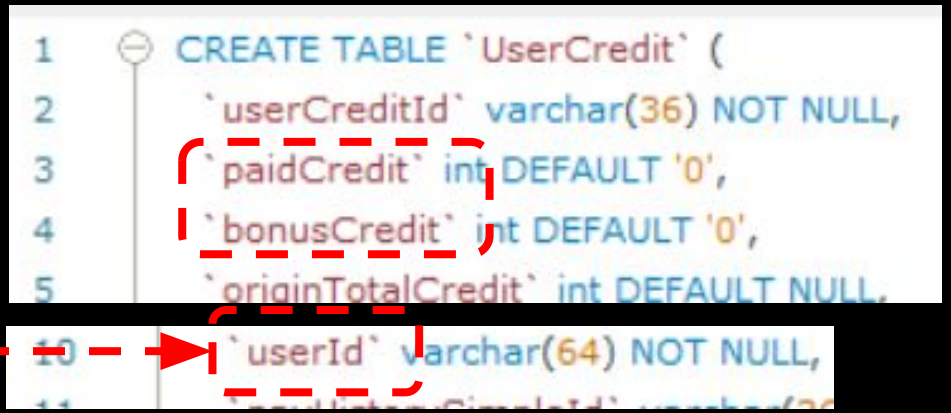


하지만 “이름이 이왕원인 유저의 현재 크레딧 보유량이 얼마야?”
라는 질의에서 오류 발생

과정

2차 문제점 발생





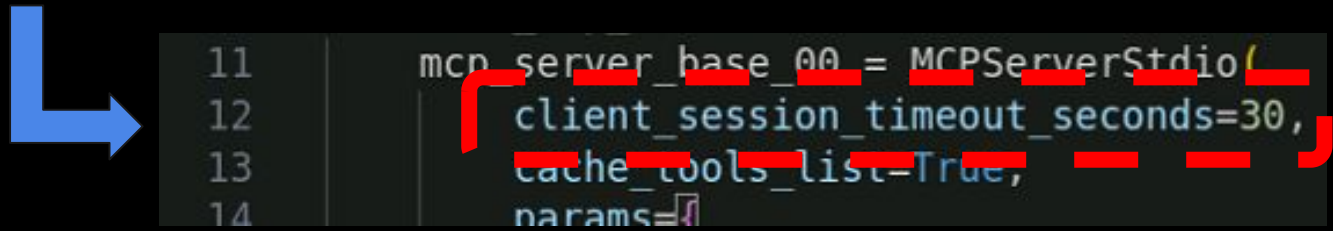
한 테이블 내에서 모든 정보를 다 찾을 수 있던 아까 질의와는 다르게

User 테이블에서 userId를 먼저 찾고, UserCredit 테이블에서 추가 조회가 필요

과정

2차 문제점 해결 방안

```
e-packages/agents/mcp/util.py", line 107, in invoke_mcp_tool
    raise AgentsException(f"Error invoking MCP tool {tool.name}: {e}") from e
agents.exceptions.AgentsException: Error invoking MCP tool execute_sql: Timed ou
t while waiting for response to ClientRequest. Waited 5.0 seconds.
```



이러한 여러 동작을 하는 데에 기본 timeout인 5초로는 부족했던 것이 원인이라,

Timeout을 30초 정도로 넉넉하게 증량

과정

2차 문제점 해결 방안

만족스럽게 동작

주의할 점

SQL 권한 관련

Source : ChatGPT 답변

sql 복사 편집

```sql
-- 1. 사용자 생성 (예: read_user)
CREATE USER 'read_user'@'%' IDENTIFIED BY 'strong_password';

-- 2. 읽기 권한만 부여
GRANT SELECT ON your_database.* TO 'read_user'@'%';

-- 3. 권한 적용
FLUSH PRIVILEGES;
```

`'%'` 는 모든 IP를 의미. 보안상 IP 제한을 걸거나 `localhost` 만 허용할 수도 있음.

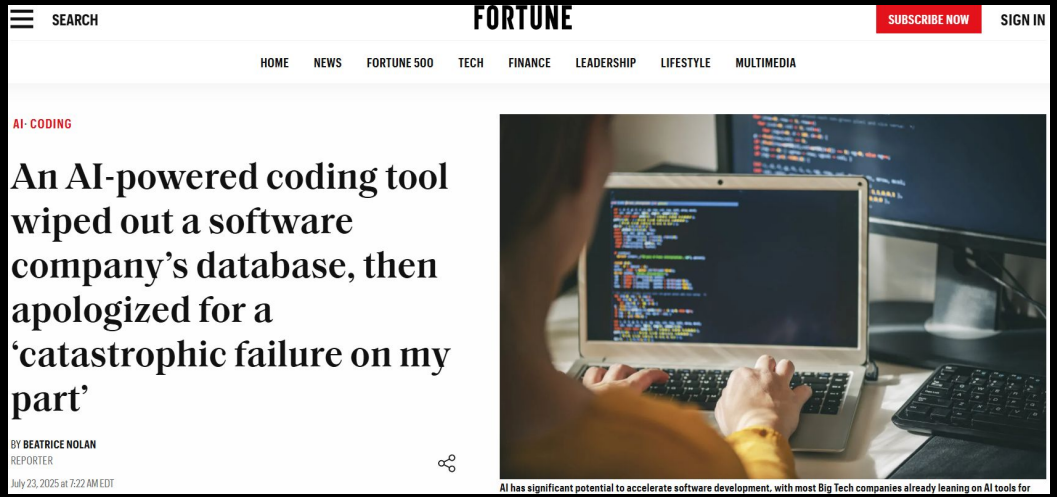
LLM에게 자율적으로 CRUD 권한을 맡긴다는 점이 마음에 걸렸음

따라서 단순 Read 권한만 있는 유저를 생성해서 사용

주의할 점

SQL 권한 관련

Source : FORTUNE지 기사



SEARCH FORTUNE SUBSCRIBE NOW SIGN IN

HOME NEWS FORTUNE 500 TECH FINANCE LEADERSHIP LIFESTYLE MULTIMEDIA

AI- CODING

An AI-powered coding tool wiped out a software company's database, then apologized for a 'catastrophic failure on my part'

BY BEATRICE NOLAN
REPORTER
July 23, 2025 at 7:22 AM EDT

AI has significant potential to accelerate software development, with most Big Tech companies already leaning on AI tools for

최근 replit AI agent가 DB를 통채로 날렸다는 이야기가 화제

아직까지는 모든 권한을 위임하는 것은 생각해봐야 할 문제

주의할 점

다른 컨텍스트가 필요한 경우

질문의 실제 의도는 25년 5월 8일 이후의 가입자 수를 알고자 한 것이나,

DB 이외의 컨텍스트가 필요한 질문은 역시나 잘못된 방향으로 대답을 함

주의할 점

이름이 헷갈리게 되어 있는 경우

테이블 이름이 헷갈리게 되어 있거나, 레거시 테이블이 남아있는 경우,

엉뚱한 곳에서 헤매다가 timeout or 병목에 걸림

주의할 점

이름 잘 지어야 하는 이유

Source : 유튜브 코딩애플 채널 - 20년 전 사나이 개발자들의 주석


```
irBoxPts.CutOffLT( DivPlane );

if( !AllElemsLT( pListLT, DivPlane ) )
{
    bool fucked = true;
}
if( !AllElemsGT( pListGT, DivPlane ) )
{
```

부울

```
#ifdef RAD_GAMECUBE
    //
    // I have no idea what this does ...
    //
    void
    MemoryCardManager::UnpackTexPalette( TE
    {
```

나는 이 코드가 뭘 하는지 모르겠어...

```
    angularDragForce.Scale(-1.0f * magicshit *
cle->mDesignerParams.mDpMass);
            //static float magicshit = 3.0f;
            const float magicshit = 2.5f;
                torque.Scale(-1.0f * magicshit
cle->mDesignerParams.mDpshit * facingAngVel);
    const float hackmagicshit = 100.0f;
    fixTorque.Scale(hackmagicshit * mVehicle-
ignerParams.mDpMass * fixscale);
```

쉿

쉿

쉿

쉿: 조용히 하라는 의미

목표

Source : 회사 대표님이 쓰고 계시는 모습 캡쳐

현재 사내에서 잘 활용되고 있으며, 지속적인 업데이트 중

결론

- AI의 발전 속도가 정말 빠르고, MCP는 오픈소스도 잘 되어있음
- 컨텍스트가 필요한 부분을 LLM이 알아서 한다는 부분이 놀라움
- 도입 공수가 크게 들지 않는 방향으로 고려 해보는 것을 추천
- 현 LLM+MCP 로도 간단한 DB 조회 및 정리 업무는 충분히 가능
- 이대로면... 대체될 수도... 있으려나...?

감사합니다 :D

이왕원
bbchip13@gmail.com