

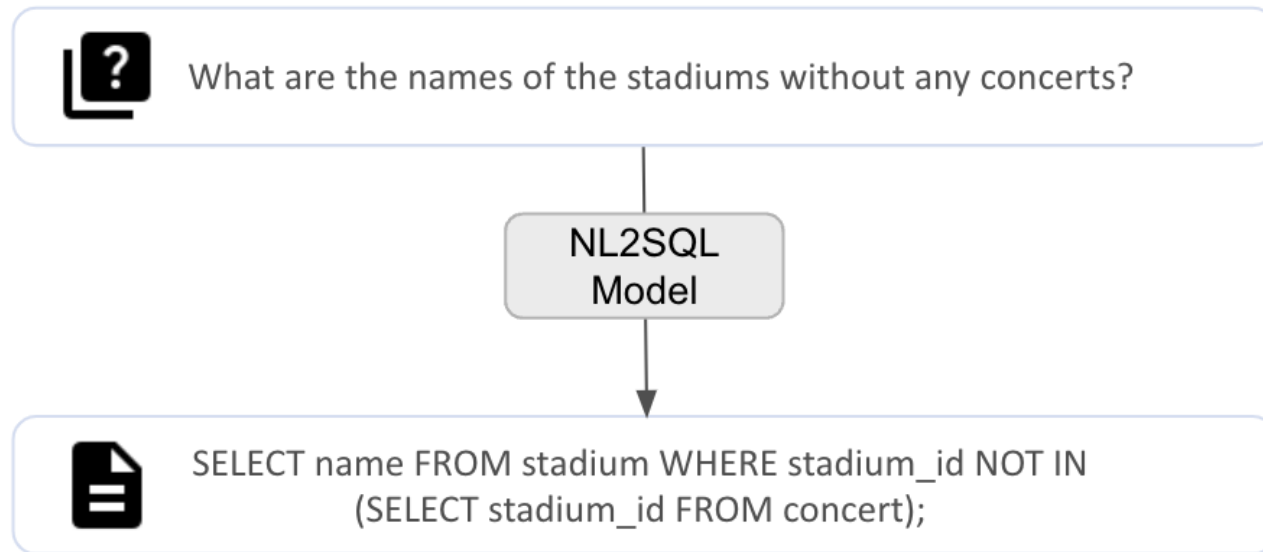
NL2SQL

VLDB Lab.

Professor Sang-Won Lee

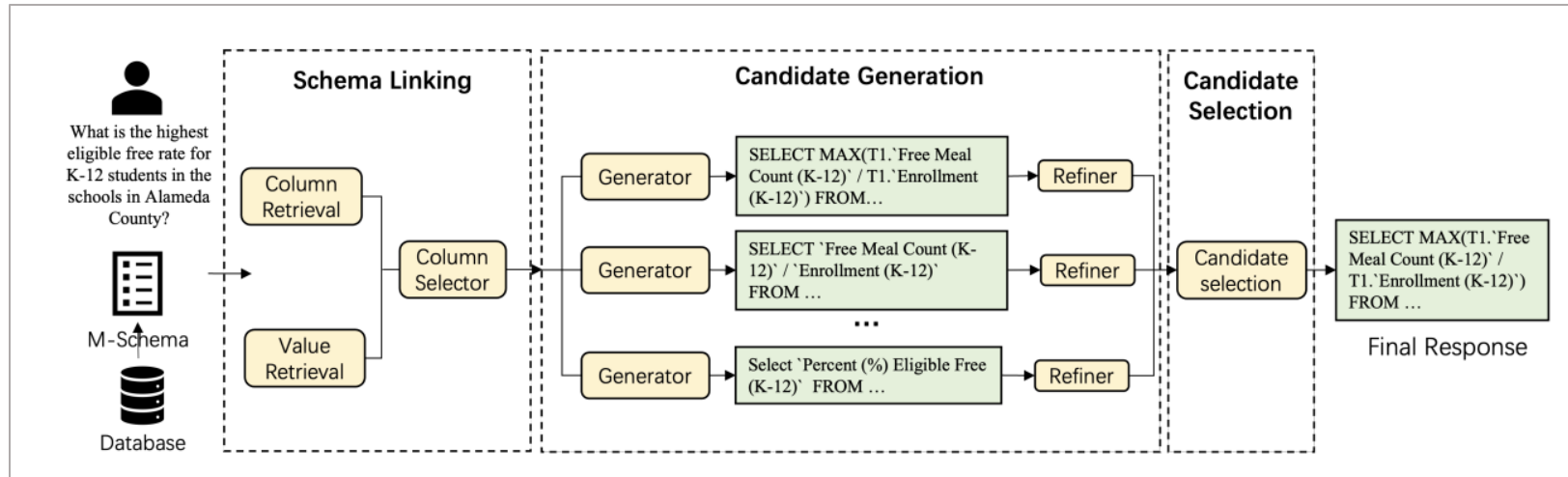
What is NL2SQL?

- 자연어 질문을 SQL 쿼리로 변환하는 작업 (Text-to-SQL이라고도 함)
- 비전문가들이 SQL 쿼리를 쉽게 생성하고 데이터를 분석하기 쉽게 하는 것이 목표

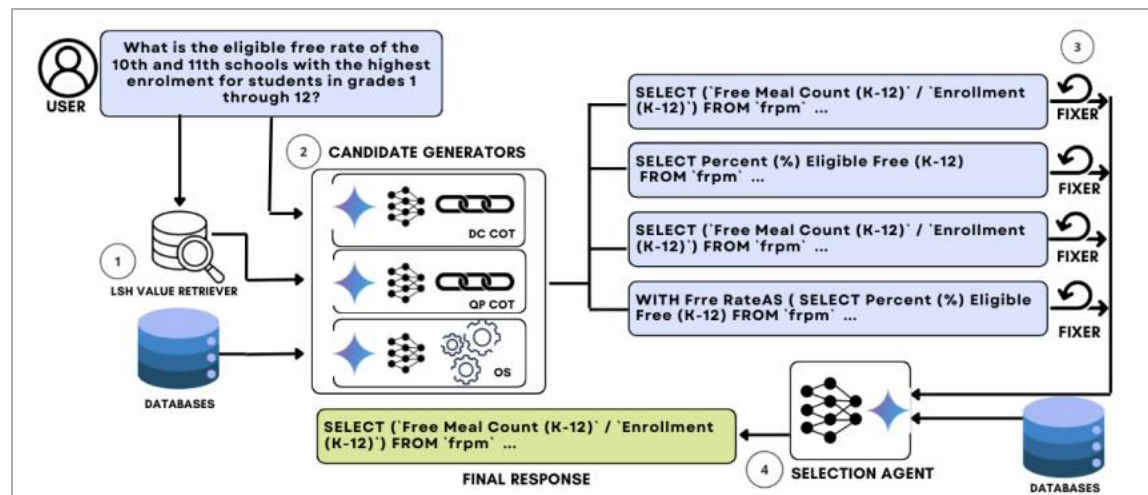


What is NL2SQL?

- 학계에서는 다양한 방법을 도입한 파이프라인을 구축



XiYan-SQL: A Multi-Generator Ensemble Framework for Text-to-SQL



CHASE-SQL: Multi-Path Reasoning and Preference Optimized Candidate Selection in Text-to-SQL

NL2SQL Platform

- 최근에는 NL2SQL을 지원하는 많은 플랫폼들이 생겨남.
 - <https://www.eversql.com/text-to-sql/>
 - <https://www.text2sql.ai>
 - <https://www.sqlai.ai/text-to-sql-ai>
- Oracle같은 DBMS에서도 NL2SQL을 개발중임.
 - <https://blogs.oracle.com/machinelearning/post/introducing-natural-language-to-sql-generation-on-autonomous-database>
- 우리는 ChatGPT를 사용해서 간단하게 NL2SQL을 체험하는 것이 목적임.
 - <https://chat.openai.com/>

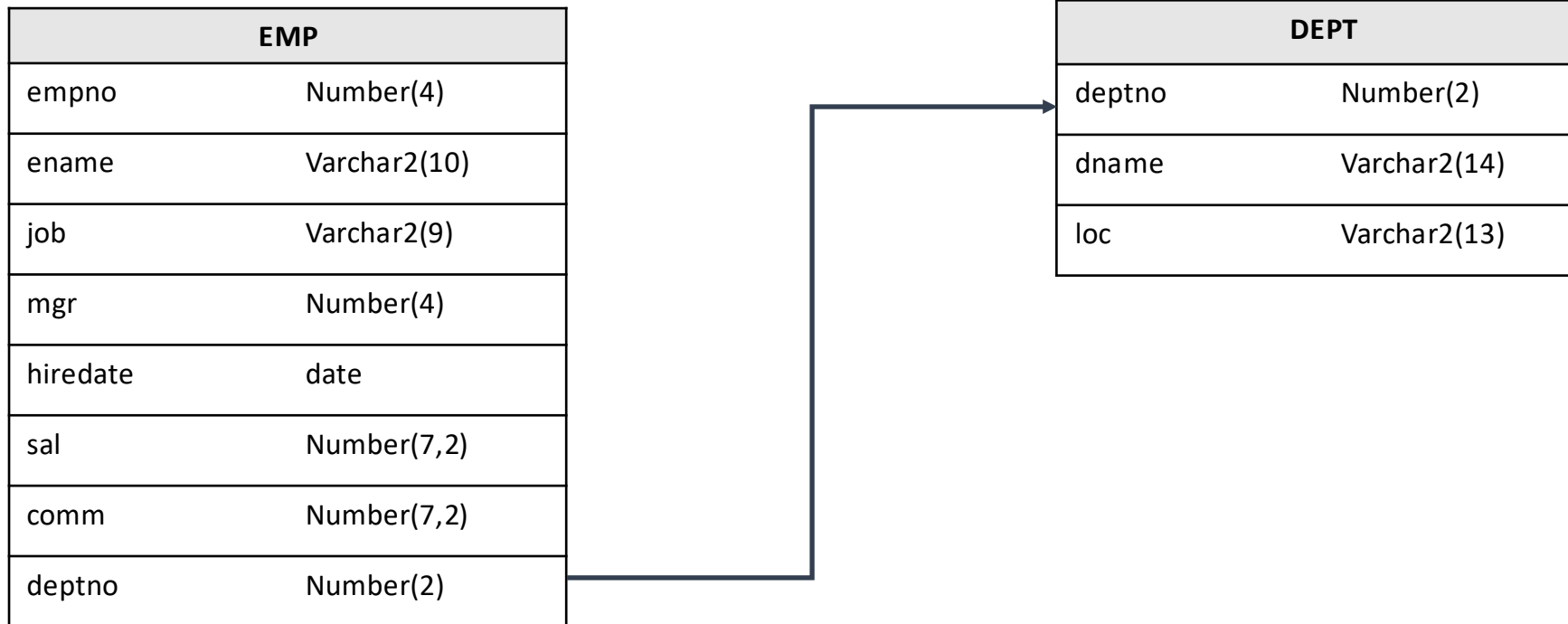
NL2SQL Prompt

- ChatGPT 사용시, ChatGPT가 대답을 잘 생성할 수 있도록 적당한 프롬프트를 넣어주어야 함.
- **프롬프트란?** 누군가(생성형 AI)의 특정한 작업 수행을 도우려 전달하는 메시지.



- 프롬프트를 어떻게 구성하느냐에 따라 다른 결과가 출력될 수 있음.
- NL2SQL에서는 **데이터베이스 스키마와 사용자 질문**을 사용해서 프롬프트를 구성해야함.

NL2SQL Practice (Scott Schema)



NL2SQL Practice (Schema Information)

- 다음과 같은 스키마가 존재한다고 가정할 때, 스키마 정보에 대한 프롬프팅 방법이 여러가지 있을 수 있음.

1.

```
CREATE TABLE DEPT (DEPTNO NUMBER(2) CONSTRAINT PK_DEPT PRIMARY KEY, DNAME VARCHAR2(14) , LOC VARCHAR2(13) ) ;
```

```
CREATE TABLE EMP (EMPNO NUMBER(4) CONSTRAINT PK_EMP PRIMARY KEY, ENAME VARCHAR2(10), JOB VARCHAR2(9), MGR NUMBER(4), HIREDATE DATE, SAL NUMBER(7,2), COMM NUMBER(7,2), DEPTNO NUMBER(2) CONSTRAINT FK_DEPTNO REFERENCES DEPT);
```

2.

```
DEPT (DEPTNO, DNAME, LOC)  
EMP (EMPNO, ENAME, JOB, MGR, HIREDATE, SAL, COMM, DEPTNO)
```

NL2SQL Practice (NLQ)

- 자연어 질문 예시

1. Display all employee whose location is DALLAS?
2. Display all the departments where department has 3 employees?
3. Delete all employees those who are reporting to BLAKE?
4. Display average salary for job SALESMAN
5. Display all ename, empno, dname, loc from emp, dept table without joining two tables?

NL2SQL Practice (Final Prompt)

- 최종 프롬프트는 다양하게 생성 가능.

1.

```
### Create the Postgres SQL Query using the following schema:  
# CREATE TABLE DEPT (DEPTNO NUMBER(2) CONSTRAINT PK_DEPT PRIMARY  
KEY, DNAME VARCHAR2(14) , LOC VARCHAR2(13) ) ;  
# CREATE TABLE EMP (EMPNO NUMBER(4) CONSTRAINT PK_EMP PRIMARY  
KEY, ENAME VARCHAR2(10), JOB VARCHAR2(9), MGR NUMBER(4), HIREDATE  
DATE, SAL NUMBER(7,2), COMM NUMBER(7,2), DEPTNO NUMBER(2)  
CONSTRAINT FK_DEPTNO REFERENCES DEPT);  
### Display all employee whose location is DALLAS?
```

2.

```
### Create the Postgres SQL Query using the following schema:  
### [Schema]  
# DEPT (DEPTNO, DNAME, LOC)  
# EMP (EMPNO, ENAME, JOB, MGR, HIREDATE, SAL, COMM, DEPTNO)  
# EMP.DEPTNO references DEPT.DEPTNO  
### [NLQ]  
### Display all employee whose location is DALLAS.
```