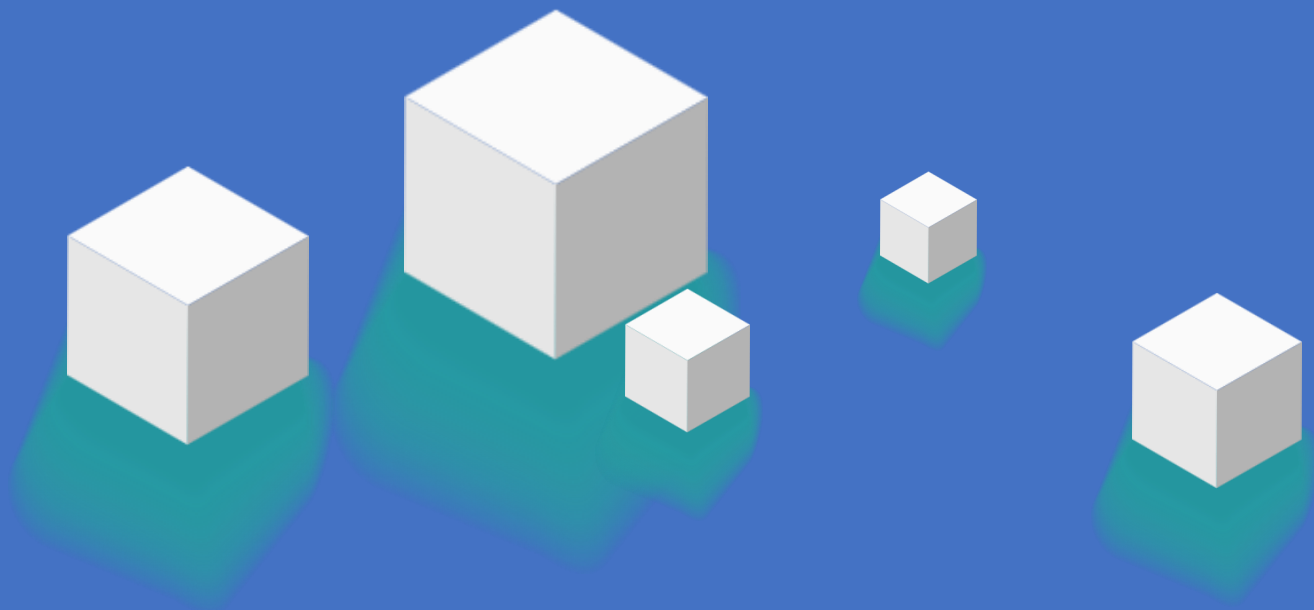
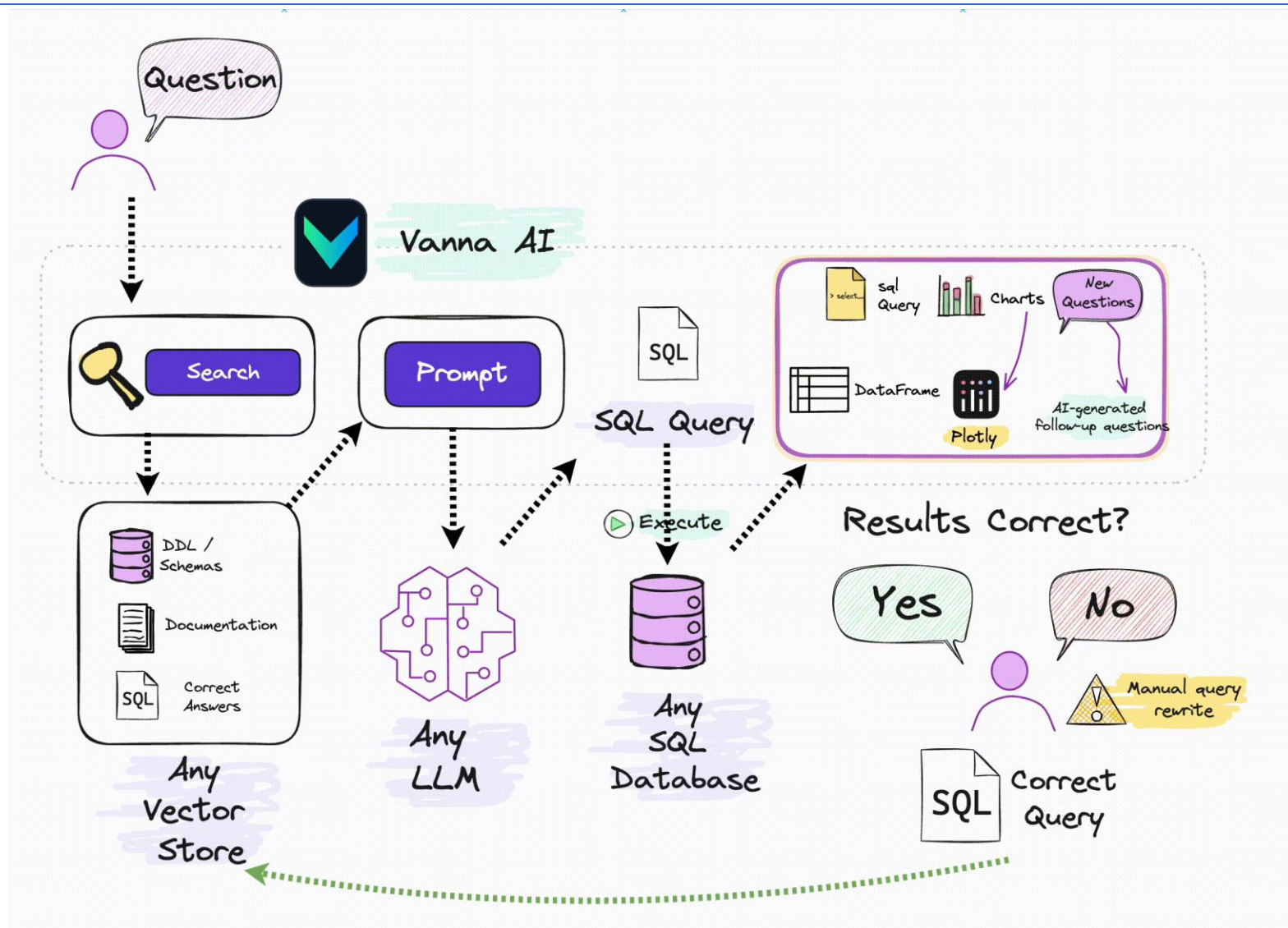


# Vanna使用



# Vanna使用



Vanna 是开源的 RAG 框架，专注于将自然语言转换为 SQL 查询（Text-to-SQL），并支持与数据库的交互

# Vanna特点

---

Vanna特点:

- **开源与可定制化:** Vanna 提供完整的 Python 库, 支持本地化部署, 允许用户自定义大语言模型 (LLM)、向量数据库 (VectorDB) 和关系型数据库 (如 MySQL、PostgreSQL 等)。
- **RAG 增强的准确性:** 通过检索增强生成技术, 结合数据库的元数据 (如 DDL 语句、表注释、示例 SQL 等) 训练模型, 显著提升复杂查询的准确率。
- **多场景支持:** 适用于企业数据分析、智能客服、电商搜索、金融报告生成等场景, 非技术人员可通过自然语言直接查询数据库。
- **灵活的基础设施:** 支持多种 LLM (如 OpenAI、本地部署的 Ollama)、向量数据库 (如 ChromaDB), 并可扩展至非默认支持的数据库。

# Vanna工作原理

---

## Vanna工作原理

- 训练 RAG 模型

输入数据库的元数据（如 INFORMATION\_SCHEMA）、DDL 语句、文档或示例 SQL。

模型将这些信息转换为向量并存储到向量库中，用于后续检索。

- 生成 SQL

用户提问时，系统从向量库中检索相关上下文，组装成 Prompt 发送给 LLM。

LLM 生成 SQL 后，自动执行并返回结果（表格或图表）。

# Vanna使用步骤

---

Vanna使用步骤:

- **vanna安装**

`pip install vanna`, 可选扩展如 `vanna[chromadb,ollama,mysql]` 支持本地化部署。

- **连接数据库**

自定义 `run_sql` 方法 (如 MySQL 需通过 `mysql.connector` 返回 Pandas DataFrame) 。

- **训练模型**

通过 DDL、文档或 SQL 示例训练, 例如:

```
vn.train(ddl="CREATE TABLE users (id INT PRIMARY KEY, name VARCHAR(100))")
```

- **提问与查询**

调用 `vn.ask("查询销售额最高的产品")`, 生成并执行 SQL。

# Vanna使用 (ask函数)

---

## ask函数

作用：用户通过自然语言提问时调用此函数，它是查询的核心入口，会依次调用`generate_sql`、`run_sql`、`generate_plotly_code`、`get_plotly_figure`四个函数来完成整个查询及可视化的过程。

工作流程：

- 首先将用户的问题转换成向量表示，然后在向量数据库中检索与问题语义最相似的DDL语句、文档和SQL查询。
- 将检索到的信息和用户的问题一起提供给LLM，生成对应的SQL查询。
- 执行生成的SQL查询，并将查询结果以表格和Plotly图表的形式返回给用户。

比如：`vn.ask("查询heros表中 英雄攻击力前5名的英雄")`

# Vanna使用 (generate\_sql函数)

---

## generate\_sql函数

作用：根据用户输入的自然语言问题，生成对应的SQL语句。

工作流程：

- 调用get\_similar\_question\_sql函数，在向量数据库中检索与问题相似的sql/question对。
- 调用get\_related\_ddl函数，在向量数据库中检索与问题相似的建表语句ddl。
- 调用get\_related\_documentation函数，在向量数据库中检索与问题相似的文档。
- 调用get\_sql\_prompt函数，结合上述检索到的信息生成prompt，然后将prompt提供给LLM，生成SQL语句。

比如：sql=vn.generate\_sql("查询heros表中 英雄攻击力前5名的英雄")

# Vanna使用 (generate\_sql函数)

---

run\_sql函数

作用：执行generate\_sql函数生成的SQL语句，并返回查询结果。

工作流程：将生成的SQL语句发送到连接的数据库中执行，获取并返回查询结果。

比如：

```
sql=vn.generate_sql("查询heros表中 英雄攻击力前5名的英雄")
```

```
vn.run_sql(sql)
```






# 打卡：Text2SQL



在你的工作中，都有哪些SQL查询的场景？（对应的数据表、SQL查询语句，LLM能否完成，是否有临时SQL的需求）

使用 LangChain 或者 vanna 或者自己调用大模型来完成  
可以使用 heros数据表，或者用自己本地的 MySQL数据表



Thank You  
Using data to solve problems