

Sjálfsgreiðsla gagna með hjálp mállíkana

Pálmar Gíslason – CCP Games



Hvað er sjálfsafgreiðsla gagna?

- Tól fyrir notendur til að vinna með gögn án aðstoðar.



Hvaða þekkingu þarf í raun

- Mælaborðalausnir (Tableau, PowerBI, ...)
- Gagnavélar (Analysis Services, Fabric, ...)
- Aðrar lausnir (Jupyter, Databricks, ...)

Hvar liggur þekking starfsfólksins

- Hversu margir geta skrifað SQL fyrirspurnir?
- Hversu margir eru með BI leyfi og kunna á forritin?
- Hversu margir geta reddað sér í excel?

Tableau – Ask Data

The screenshot displays the Tableau 'Ask Data' interface. At the top, there are two buttons: '+ Add Field' and '+ Add Filter'. Below these is a search bar with the placeholder text 'Search fields or values to create a visualization' and a blue arrow button on the right. Underneath the search bar, a 'Suggestions' section shows a list of field-based suggestions: 'sum of Sales', 'count of Orders', 'by Category', 'sum of Profit', and 'filter State to South Carolina'. Each suggestion is accompanied by a small icon representing its data type (e.g., a green square for sum, a blue square for count, a red square for by, a green square for sum, and a blue square for filter). Below the suggestions is a section titled 'Recommended Visualizations'. This section includes a brief instruction: 'To add or replace links in this list, go to a visualization and click the pin icon in the toolbar.' Below the instruction is a list of four recommended visualization types, each with a right-pointing arrow icon: 'Basic Data Analysis', 'Date and Time', 'Filters', and 'Viz Type'.

+ Add Field

+ Add Filter

Search fields or values to create a visualization

Suggestions | sum of Sales | count of Orders | by Category | sum of Profit | filter State to South Carolina

Recommended Visualizations

To add or replace links in this list, go to a visualization and click the pin icon in the toolbar.

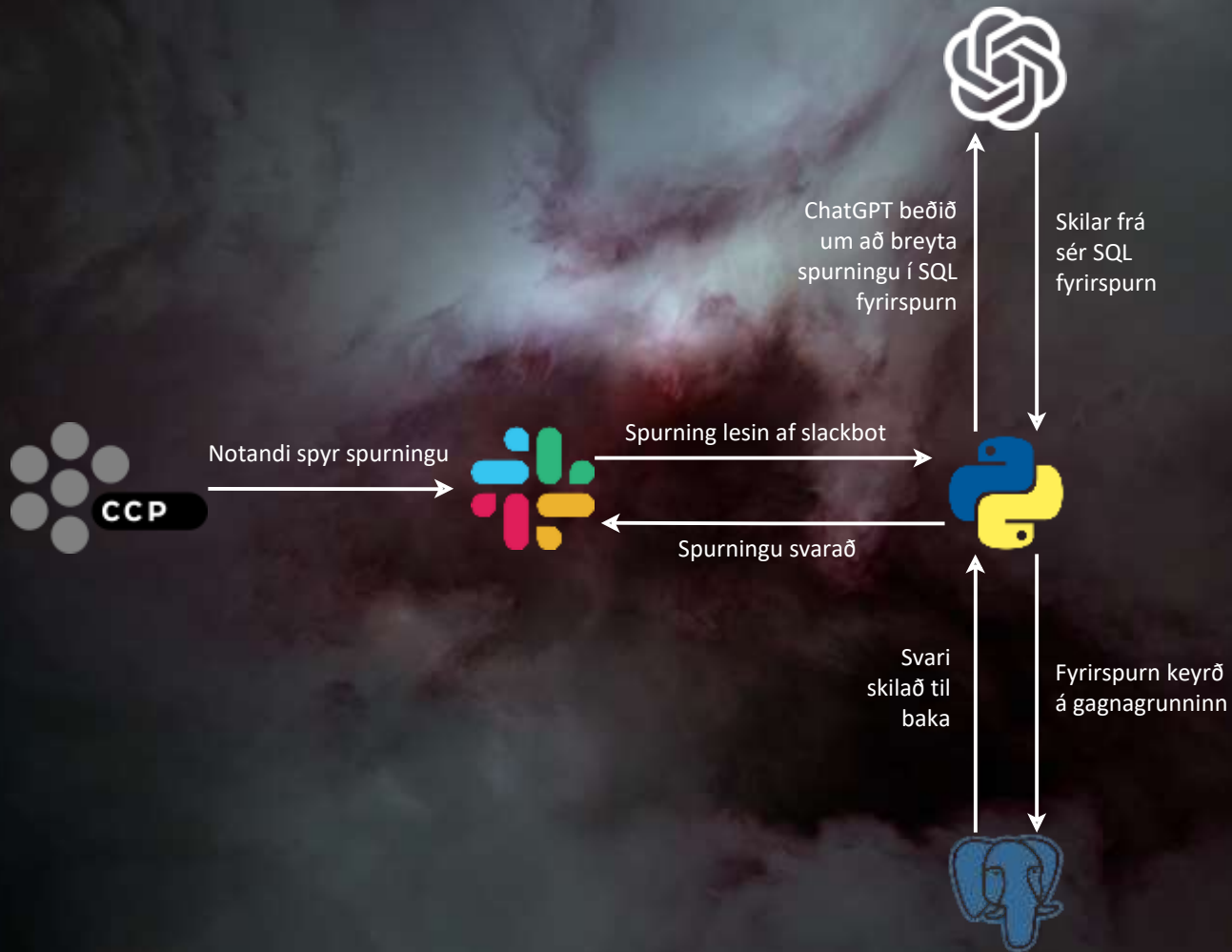
- > Basic Data Analysis
- > Date and Time
- > Filters
- > Viz Type

Mállíkön

- Góð í að skrifa og skilja texta
- Sæmilega góð í að skrifa kóða
- Erfitt að þjálfa líkönin á sérhæfðum gögnin (er að verða betra)

Gagna spjallmennið

- Tekur við fyrirspurnum á Slack
- Breytir fyrirspurn í SQL kóða
- Skilar niðurstöðu SQL fyrirspurnarinnar



Demo

- Allir viðburðir S
- Öll erindi og fyr



```
//Given the following SQL tables in a sqlite database,  
//your job is to write queries given a user's request  
//and only return the SQL query statement.
```

```
CREATE TABLE events (  
  ...  
);  
CREATE TABLE presentations (  
  ...  
);  
CREATE TABLE presenters (  
  ...  
);
```

muži mēdži val foto

- ggr-3-3-turbo
- ggr-4-1286-preview

Sparring

Itersu mēdži val foto itals verit le

Syrja sparringi

Sijna graf


	idp	idp_2021	idp_2022
0	idp1, idp2, idp3, idp4, idp5, idp6, idp7, idp8	2021-01-18	14:00:00:000000 18:00:00:000000
1	idp1, idp2, idp3, idp4, idp5, idp6, idp7, idp8	2021-01-12	09:30:00:000000 12:30:00:000000
2	idp1, idp2, idp3, idp4, idp5, idp6, idp7, idp8	2021-01-19	09:00:00:000000 12:00:00:000000
3	idp1, idp2, idp3, idp4, idp5, idp6, idp7, idp8	2022-01-19	09:00:00:000000 12:00:00:000000
4	idp1, idp2, idp3, idp4, idp5, idp6, idp7, idp8	2021-01-05	09:00:00:000000 12:00:00:000000

Aðferðin er ekki fullkomin

- GPT-3.5 höndlar margar töflur illa
- GPT-4 er heldur hægt, og token limit eru vandamál
- Hvenær er maður tilbúinn að treysta niðurstöðunum?

Framtíðar betrubætur

- Sérhæfð líkön með vitneskju um stöðug gögn
- Hægt að þjálfa líkönin sérstaklega í ákveðnum hlutum



Takk fyrir áheyrnina!