Paper Title

Jaouhara Chanchaf
UM6P

Karima Echihabi
UM6P

ABSTRACT

300 word description of the project

1 INTRODUCTION

- **Use case 1: Keyword Query**
  A data scientist wants to analyze the impact of food cost inflation on food consumption. Initially The user decides to start the search with a keyword-query $Q_0 = (\{"food", "consumption"\}, k = 10)$. The search engine returns Table 2 which contains data from year 1990 to 2009 about *"Per Capita Consumption of Principal Foods (in pounds)"*. The user decides to keep Table 2 for the study and continue to search for other relevant tables.
- **Use case 2: Join Query** Table 2 is a good first result as it contains a complete list of the main food types, however this result lacks information on food prices. For that the use perform a join query on the food column to explore other tables that may have information about food prices for the years 1990 to 2009.
  Attempt 1: To speed up search the user submits $Q_1 = $ (Table 2, Join column : "Food", $k = 10$) with a small k value. The search engine returned 775 tables. However, after skimming through the list of returned tables nothing seemed relevant to the user.
  Attempt 2: The user decides to increase k to get more results from the search engine. He/she submits a second query $Q_2 = $ (Table 2, Join column : "Food", $k = 20$). This time the search engine returned 161 tables, because the number of results is big the user could notice Table **??** ranked at position 55.
  Attempt 3: As a last attempt the user gave up on getting any fast meaningful result so he/she decide to increase k significantly in hope that a relevant table will appear in the list of results. He/she submits $Q_3 = $ (Table 2, Join column : "Food", $k = 200$). Finally and after several attempts, the search engine returned Table **??** at position 11 which contains information on food prices from the 2007 WIC program.

2 DEFINITIONS AND TERMINOLOGY

Dataset Discovery.

Keyword and Join Queries.

Incremental Query Answering.

3 SYSTEM ARCHITECTURE

4 DEMONSTRATION

REFERENCES

[1] K. Echihabi, K. Zoumpatianos, T. Palpanas, and H. Benbrahim. The Lernaean Hydra of Data Series Similarity Search: An Experimental Evaluation of the State of the Art. *PVLDB*, 12(2), 2018.

[2] PhDComics. Graduate Student Work Output. https://phdcomics.com/comics/archive.php?comicid=124, 2022.

Company	Plant	Location	Feedstock	Capacity (MW)
Wheelabrator Technologies Inc.	Wheelabrator Shasta Energy Co. Inc.	Anderson - CA	Logging and Mill Residue/Ag Residue	50
Greenleaf Power LLC	Desert View	Mecca - CA	Ag Residue/Urban Wood Waste	47
Covanta	Covanta Delano	Delano - CA	Orchard and Vineyard Prunings/Nut Shells/Stone Fruit Pits	58
Greenleaf Power LLC	Honey Lake	Wendel - CA	Mill and Logging Residue/Forest Thinning/Urban Woodwaste	30

Table 1: U.S. Biomass Power Plants

Category	Plant ID	Plant Name	Unit	Status	Start Date	Retire Date	Prime mover ID	Prime Mover Description	Capacity	net MWh

Table 2: U.S. Biomass Power Plants