dePsypher

Developed By:

Hamnah Khalid 237888

Muhammad Haris 217735

Shifa Shariq 211686

Taha Abdullah 213818

Work Distribution

➢ Parser:
- Hamnah Khalid
- Taha Abdullah

➢ GUI and Presentation:
- Muhammad Haris
- Shifa Shariq

The choices we made..

➢ **The Dataset**

- Wikipedia articles
    - To actually implement a "web" search engine
    - More suitable to extend in future
    - To explore how web searching actually works

The Language

- Python
  - High-level, programmer-friendly language
  - Provides various modern libraries
  - Syntax- less coding

Database Language

- SQLite
  - Easy to integrate with Python
  - SQLite3 module used instead of "pymysql" and "mysqlconnector"
  - No need of local host, works on a single server

Python Modules

- BeautifulSoup
  - Easily pulls data from HTML and XML files
  - Compatible to work with parsers for navigating and modifying parse trees

nltk

- Natural Language Toolkit
  - Easy-to-use interface
  - Variety of text-processing libraries

Lemmatizing over Stemming

- Stemming partitions the word, doesn't return the root word

- Lemmatizer takes into account the context information i.e. part of speech

- Returns valid root words

Commit History

Branch: master ▾	New pull request	Create new file Upload files Find file Clone or download ▾

TahaAbdullah20 Add files via upload   Latest commit 8a9a54a 14 hours ago

📄 GUI_v2.py	Add files via upload	14 hours ago
📄 Parser_v4.py	Rename Parser1.py to Parser_v4.py	11 days ago
📄 README.md	Initial commit	14 days ago
📄 forward.db	Add files via upload	3 days ago
📄 parser.py	Add files via upload	14 days ago
📄 parser2.py	Add files via upload	13 days ago
📄 parser_v2.py	Rename parserr.py to parser_v2.py	11 days ago
📄 parser_v3.py	Add files via upload	11 days ago
📄 parser_v4.2.py	Add files via upload	11 days ago
📄 parser_v4.3.py	Add files via upload	5 days ago
📄 parser_v4.4.py	Add files via upload	3 days ago
📄 store.py	Add files via upload	5 days ago

THANK YOU!