

Information Retrieval with PostgreSQL

Alexander Hebel

Heidelberg University
Institute of Computer Science
Database Systems Research Group
vx228@uni-heidelberg.de

Mai 6, 2020

Outline

- 1 Introduction
- 2 Approach and realizations
- 3 Custom C-functions in PostgreSQL
- 4 Rating sections vs. rating pages
- 5 Conclusion

Outline

- 1 Introduction
- 2 Approach and realizations
- 3 Custom C-functions in PostgreSQL
- 4 Rating sections vs. rating pages
- 5 Conclusion

Task definition

- How looks and performs an IRS made of a relational database
- similar to Apache Solr
- finding different database models
- Python api for the database creation and communication
- Crawl Wikipages to gather some text data
- special type in PostgreSQL named tsvector (full text search)

First goal

support some boolean search queries like AND

Outline

- 1 Introduction
- 2 Approach and realizations**
- 3 Custom C-functions in PostgreSQL
- 4 Rating sections vs. rating pages
- 5 Conclusion

Database models

Outline

- 1 Introduction
- 2 Approach and realizations
- 3 Custom C-functions in PostgreSQL**
- 4 Rating sections vs. rating pages
- 5 Conclusion

Adding your custom C-functions to PostgreSQL

Prerequisites

- Developer version of PostgreSQL
- Installation of make
- Root privilege on database

Folder structure

Extension

- ─ function.c
- ─ Makefile
- ─ function.control
- ─ function--1.0.sql
- ─ README.function

Steps

- (1) make install
- (2) CREATE EXTENSION "extension"

Outline

- 1 Introduction
- 2 Approach and realizations
- 3 Custom C-functions in PostgreSQL
- 4 Rating sections vs. rating pages**
- 5 Conclusion

Idea

- The Beamer class – CTAN
`http://ctan.math.utah.edu/ctan/tex-archive/macros/latex/contrib/beamer/doc/beameruserguide.pdf`
- “A Beamer Tutorial in Beamer” von Charles T. Batts
`https://www.uncg.edu/cmp/reu/presentations/Charles%20Batts%20-%20Beamer%20Tutorial.pdf`
- The Beamer class for \LaTeX
`http://www.mathematik.uni-leipzig.de/~hellmund/LaTeX/beamer2.pdf`
- Beamer Theme Matrix
`http://www.hartwork.org/beamer-theme-matrix/`

Outline

- 1 Introduction
- 2 Approach and realizations
- 3 Custom C-functions in PostgreSQL
- 4 Rating sections vs. rating pages
- 5 Conclusion**

Conclusion and future work

Questions

Questions