

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Факультет прикладной математики и информатики

Кафедра дискретной математики и алгоритмики

Аннотация к дипломной работе

«Разработка модуля полнотекстового поиска для PostgreSQL»

Новиков Ярослав Кириллович

Научный руководитель – старший преподаватель кафедры дискретной математики и алгоритмики ФПМИ Молочко И. П.

Минск, 2025

АННОТАЦИЯ

Дипломная работа, 43 стр., 15 лист., 14 источников

Ключевые слова: ПОЛНОТЕКСТОВЫЙ ПОИСК, POSTGRESQL, PYTHON, FTS, TSVECTOR, TSQUERY, GIN ИНДЕКС, UNACCENT, ЛОКАЛЬНЫЙ ПОИСК, ИНДЕКСАЦИЯ ДОКУМЕНТОВ, РАНЖИРОВАНИЕ РЕЗУЛЬТАТОВ, ПАРСИНГ ФАЙЛОВ, PDF, TXT, MD.

Объектом исследования является полнотекстовый поиск в PostgreSQL.

Предметом исследования является задача разработки и реализации системы локального полнотекстового поиска для персональных знаний.

Целью работы является разработка системы локального полнотекстового поиска на основе PostgreSQL FTS и Python, анализ ее компонентов и демонстрация практического применения для поиска по локальным документам.

Методами исследования являются методы полнотекстового поиска.

Полученные результаты и их новизна: была разработана и реализована система локального полнотекстового поиска, интегрирующая возможности СУБД PostgreSQL для хранения и индексации персональных данных и использующая Python для обработки файлов.

Достоверность материалов и результатов дипломной работы: использованные материалы и результаты дипломной работы являются достоверными. Работа выполнена самостоятельно.

Областью возможного практического применения является система для управления персональными знаниями, для организации локальных архивов документов.

АНАТАЦЫЯ

Дыпломная праца, 43 стар., 15 ліст., 14 крыніц.

Ключавыя словы: ПОЎНАТЭКСТАВЫ ПОШУК, POSTGRESQL, PYTHON, FTS, TSVECTOR, TSQUERY, GIN ІНДЭКС, UNACCENT, ЛАКАЛЬНЫ ПОШУК, ІНДЭКСАЦЫЯ ДАКУМЕНТАЎ, РАНЖЫРАВАННЕ ВЫНІКАЎ, ПАРСІНГ ФАЙЛАЎ, PDF, TXT, MD

Аб'ектам даследавання з'яўляецца задача распрацоўкі і рэалізацыі сістэмы лакальнага поўнатэкставага пошуку для персанальных ведаў.

Прадметам даследавання з'яўляюцца распрацоўка сістэмы лакальнага поўнатэкставага пошуку на аснове PostgreSQL FTS і Python, аналіз яе кампанентаў і дэманстрацыя практычнага прымянення для пошуку па лакальных дакументах.

Мэтай даследавання з'яўляецца сістэмны падыход, аналіз тэхнічнай дакументацыі і літаратуры, праектаванне праграмага забеспячэння, распрацоўка і тэсціраванне праграмных модуляў, апрацоўка тэкставых даных.

Метадамі даследавання з'яўляюцца метады поўнатэкставага пошуку.

Атрыманыя вынікі і іх навізна: была распрацавана і рэалізавана сістэма лакальнага поўнатэкставага пошуку, якая інтэгруе магчымасці СУБД PostgreSQL для захоўвання і індэксацыі персанальных даных і выкарыстоўваючая Python для апрацоўкі файлаў.

Даставернасць матэрыялаў і вынікаў дыпломнай працы: выкарыстаныя матэрыялы і вынікі дыпломнай Працы з'яўляюцца даставернымі. Праца выканана самастойна.

Вобласцю магчымага практычнага прымянення з'яўляецца сістэма арганізацыя лакальных архіваў дакументаў, падтрымка індывідуальнай даследчай дзейнасці.

ANNOTATION

Diploma thesis, 43 p., 15 list., 14 sources.

Keywords: FULL-TEXT SEARCH, POSTGRESQL, PYTHON, FTS, TSVECTOR, TSQUERY, GIN INDEX, UNACCENT, LOCAL SEARCH, DOCUMENT INDEXING, RESULT RANKING, FILE PARSING, PDF, TXT, MD.

The object of the research is the full text search in PostgreSQL.

The subject of the study is the task of developing and implementing a local full-text search system for personal knowledge.

The purpose of the research is to develop a local full-text search system based on PostgreSQL FTS and Python, analyze its components, and demonstrate practical application for searching local documents.

Methods of research are full text search methods.

The results of the work and their novelty: a local full-text search system was developed and implemented, integrating the capabilities of the PostgreSQL DBMS for personal data storage and indexing with the flexibility of Python for file processing.

Authenticity of the materials and results of the diploma work: the materials used and the results of the diploma work are authentic. The work has been put through independently.

Recommendations on the usage: can be used in personal knowledge management, organization of local document archives.