

Система для анализа библиометрических показателей CEUR Workshop Proceedings

Шитов Егор Александрович

Научный руководитель: Чернышев Георгий Алексеевич

СПбГУ, Математико-механический факультет
Системное программирование

Цель — разработать систему для анализа библиометрических показателей открытого репозитория CEUR-WS. Задачи, поставленные в рамках данной работы:

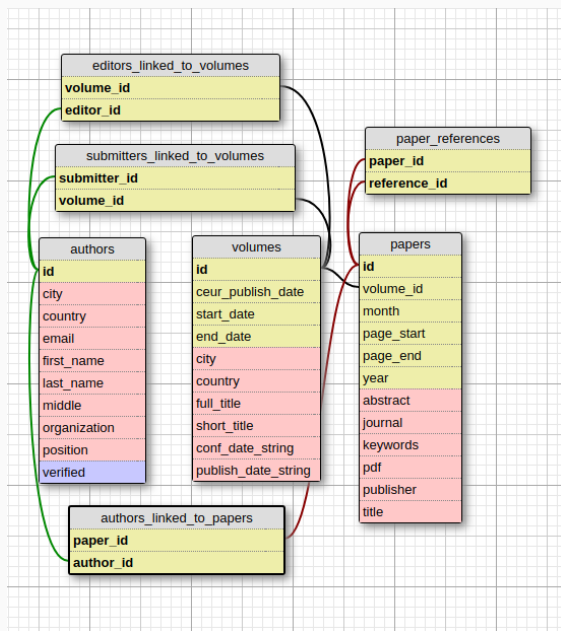
- реализовать забор данных с сайта
- разработать схему хранения данных в БД
- извлекать информацию из pdf файла статьи
- реализовать модуль дедупликации статей
- реализовать пользовательский интерфейс

- ACM Digital Library
- Springer Link
- IEEE Xplore Digital Library
- DBLP

- Kotlin
- Golang
- Spring Boot
- MySQL
- Freemarker

- трехуровневая
- модульная
 - отказоустойчивость
 - легкость замены модулей
 - распределение нагрузки

Структура базы данных



- AnyStyle
- FreeCite
- GROBID

- сравнить года публикации
- сравнить авторов по инициалам
- сравнить названия статей

- реализован забор данных с сайта CEUR-WS.org
- организовано хранение данных в БД
- найден и использован эффективный инструмент для извлечения информации из pdf файла статьи
- реализован модуль дедупликации статей
- реализован базовый пользовательский интерфейс