



Санкт-Петербургский государственный университет
Кафедра системного программирования

Система поиска учебно-методической документации с поддержкой метаданных

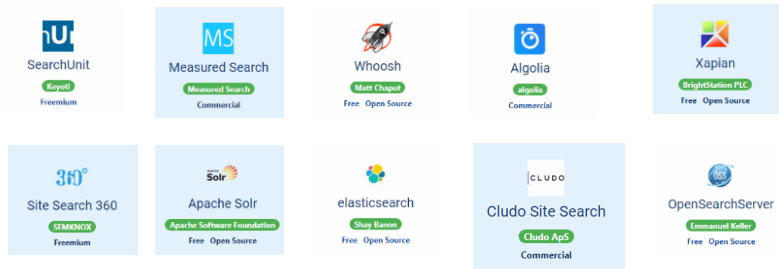
Ахмедов Гаджи Омар оглы, 21.М07-мм группа

Научный руководитель: к.ф.-м.н., доцент кафедры СП Д.В.Луцев

Санкт-Петербург
2022

- Современные компании, фирмы нуждаются в удобном и быстром поисковике для поиска информации в своей документации.
- Есть большой выбор поисковиков но у каждого есть свои плюсы так и минусы.

Существующие решения (инструменты, подходы, алгоритмы)



- Скорость индексации
- Стабильность для всех типов запросов
- Доступность

В данной работе будет использоваться elasticsearch для поиска метаданных на рабочем столе.

Выводы

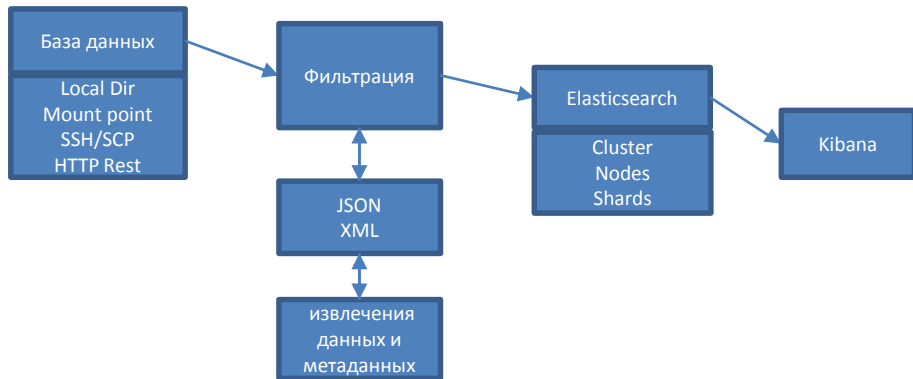
-) Подвести итог
-) Указать недостатки существующих подходов, на борьбу с которыми направлена данная работа
-) Чётко сформулировать существующую проблему, которая будет решаться в данной работе

Постановка задачи

Целью работы является создание десктопного поисковика, ориентированный на конкретный вид документации (например: Учебно-методическую документацию). Для достижения обозначенной цели были поставлены следующие задачи:

1. Выполнить анализ предметной области — существующих решений и подходов.
2. Сформулировать требования к поисковой системе.
3. Спроектировать архитектуру поисковой системы.
4. Реализовать поддержку специфических метаданных при поиске.

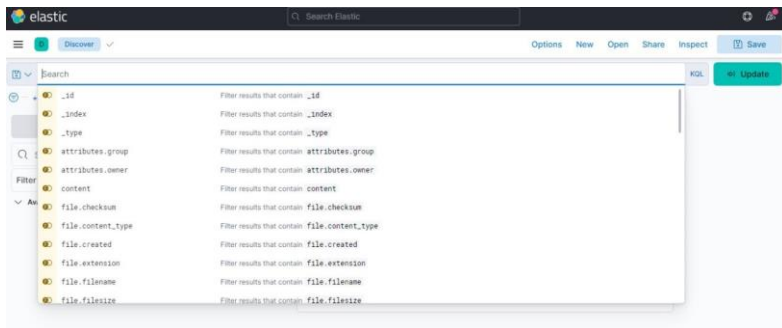
Архитектура решения



Экспериментальное исследование

Постановка эксперимента

- Был использован файл с учебно-методической документацией на моем локальном диске
- В сравнении с elastic не требуется писать все вручную то есть добавление, удаление, поиск данных стал удобен и автономен.



Результаты

- На данный момент изучена предметная область и стек технологий, сделан обзор, сформулирована цель и поставлены задачи.
- Возможность поиска на локальном диске.
- Реализовал задачу поиска по разным форматам файлов.
- Раскрытие сведений о признаках и свойствах, характеризующих какие-либо сущности, позволяющие автоматически искать и управлять ими в больших информационных потоках.
- Чтение, извлечение и обработка данных.

Дальнейшие планы

Реализовать интеграцию пользовательских метаданных