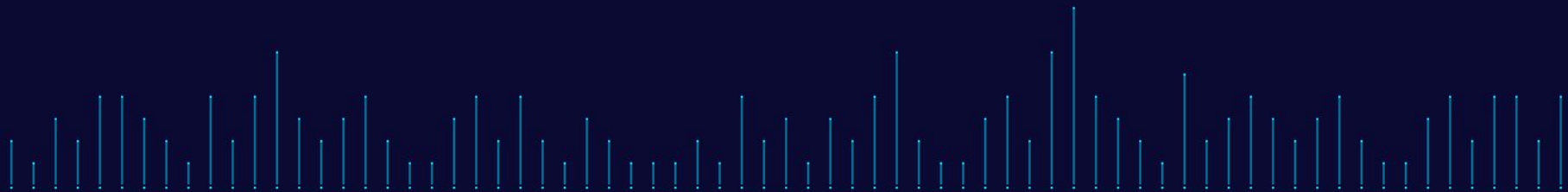


ВАС ПОНЯЛ

Киберфизическая поисковая ЭКОсистема

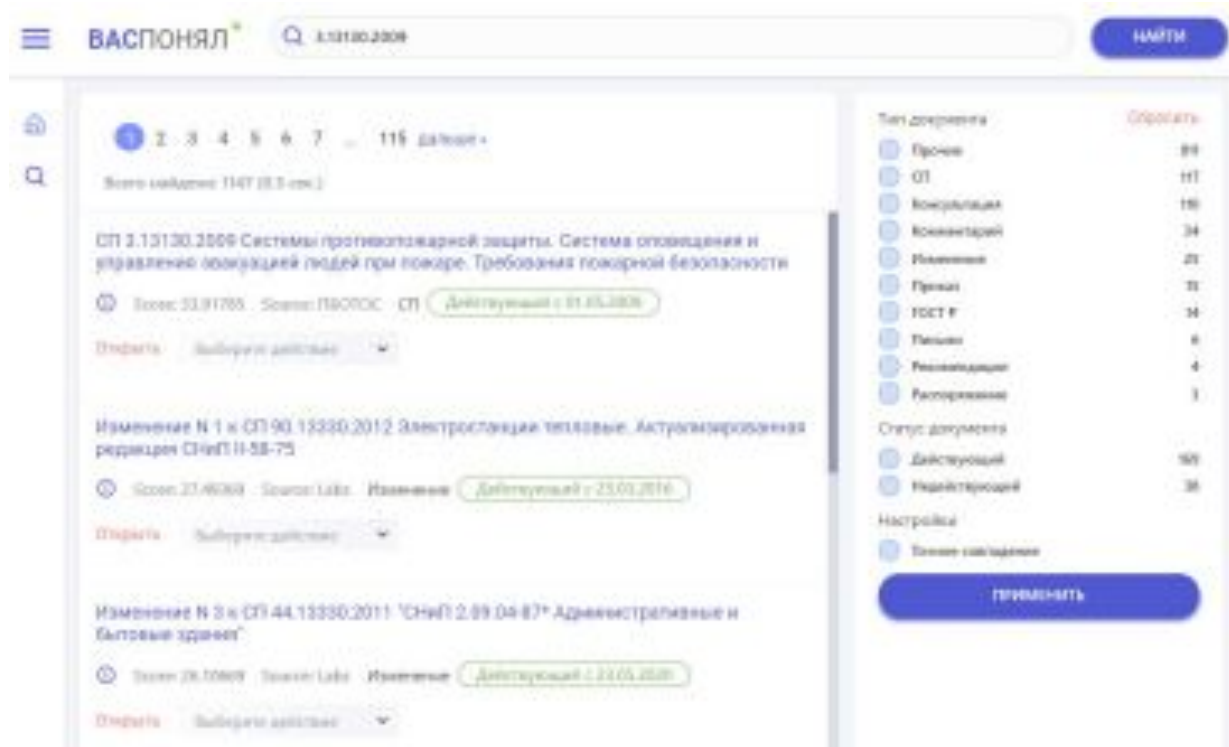


Суть проекта

Цифровая платформа для хранения и поиска по нормативно-техническим, справочным и методическим документам «ВасПонял»

Состоит из:

- Управляющая оболочка с системой точного поиска
- Обновляемая база нормативно-технических, отраслевых документов, комментариев, разъяснений, справочной, консультационной информации
- Чат-бот помощник для быстрого подключения пользователя к единой горячей линии



Требования к серверам (в разработке)

1
Аутентификация
4 core
8 GB RAM
200 GB SSD
1000 Mbit/s Internet

2
Загрузка данных
8 core
16 GB RAM
3 TB HDD SATA
1000 Mbit/s Internet

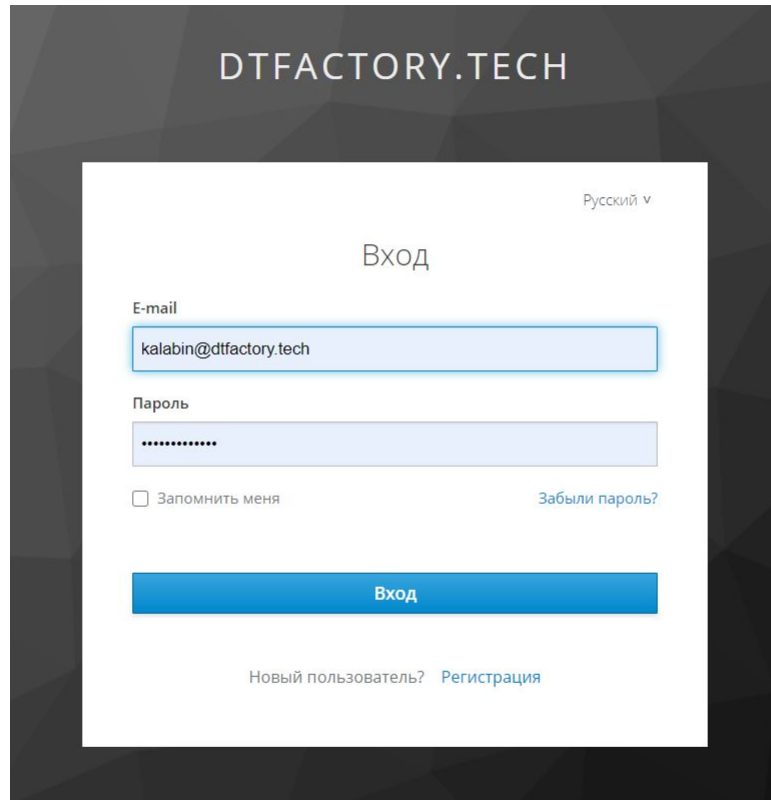
3
Клиент/сервер/gitlab/
chat
16 core
16 GB RAM
200 GB SSD
1000 Mbit/s Internet

4
ElasticSearch/Redis
16 core
32 GB RAM
3 TB HDD SATA
1000 Mbit/s Internet

5
Backup
4 core
8 GB RAM
3TB HDD SATA
1000 Mbit/s Internet

Авторизация

- Поддерживается самостоятельная регистрация пользователя через email и заведение пользователей администратором системы через панель Keycloak
- Пользователь самостоятельно может сбросить пароль, через email в случае, если его забудет
- Поддерживается функция «Запомнить меня»



DTFACTORY.TECH

Русский ▾

Вход

E-mail

kalabin@dtfactory.tech

Пароль

.....

Запомнить меня [Забыли пароль?](#)

Вход

[Новый пользователь? Регистрация](#)

Документы

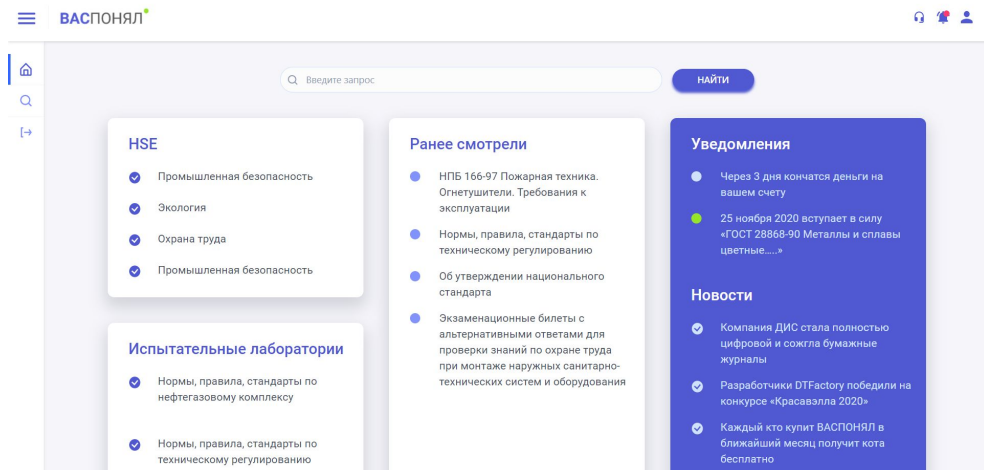
- Поддерживается загрузка документов в любых форматах
- Для индексации поддерживаются документы в формате rtf
- Документ, помимо текста, имеет ряд атрибутов:
 - Тип документа (ГОСТ, СП,...)
 - Статус (действующий, недействующий,...)
 - Дата вступления
 - Дата истечения
 - Принявший орган
 - Авторство
 - Документы, которые ссылаются на этот документ
 - Документы, на которые ссылается документ
 - ...(полный список доступен в приложении)

Группировка документов

- Продукт лицензируется по разделам, которые продаются отдельно. Пример: промышленная безопасность, лаборатории.
- Один документ может относиться сразу к нескольким разделам (например, быть и в промышленной безопасности, и в лабораториях)
- Разделы могут делиться на подразделы (TBD)

Главная страница: компоненты

- Поиск
- Выход из системы
- Связь с поддержкой (в разработке)
- Уведомления (в разработке)
- Личный кабинет (в разработке)
- Подразделы (в разработке)
- Уведомления и новости (в разработке)



Результат поиска

- Результаты поиска выдаются постранично по 10 документов
- Помимо результата поиска выдается его score, источник, статус и дата принятия
- Поддерживается фильтрация по Типу документа, Статусу
- Поддерживается поиск по точному совпадению
- Нажатие на заголовок документа ведет к открытию упрощенной версии
- Нажатие на кнопку «Открыть» ведет к открытию полной версии.

The screenshot shows the search results for 'нпб 166-97' on the VASPONJAL website. The search bar at the top contains the query 'нпб 166-97' and a 'НАЙТИ' button. The results are displayed in a list format with pagination (1-7, 650, далее >). The first result is 'О внесении изменения в НПБ 166-97' with a score of 31.518364, source 'prombez', status 'Действующий', and date '12.01.2000'. The second result is 'НПБ 166-97 Пожарная техника. Огнетушители. Требования к эксплуатации' with a score of 28.658077, source 'prombez', status 'Действующий с 01.03.1998', and date '31.12.1997'. The third result is 'РД 07-166-97 Инструкция по наблюдениям за сдвигами земной поверхности и расположенными на ней объектами при строительстве в Москве подземных сооружений' with a score of 23.650162, source 'prombez', status 'Действующий с 17.09.1997', and date '17.09.1997'. The fourth result is 'НПБ 166-97 могут применяться в части, не противоречащей Техническому регламенту о требованиях пожарной безопасности и в'. On the right side, there are filters for 'Тип документа' (ГОСТ: 1154, НПБ, Постановление, СНиП, СП, Федеральный закон, Стандарт организации) and 'Статус документа' (Действующий: 3721, Недействующий: 1834, Действующий, Отменен в части: 8, Не вступил в силу: 21, Действие прекращено: 90). There is also a 'Настройки' section with a 'Точное совпадение' filter and a 'ПРИМЕНИТЬ' button.

Просмотр документа

Упрощенная версия документа

- Версия, которая изначально хранится в виде html в Elastic, не содержит сложных таблиц и картинок.
- Открывается менее 1 секунды

Полная версия документа

- Версия конвертируется на лету при нажатии кнопки из rtf документа
- Содержит таблицы, рисунки
- Поддерживается увеличение и уменьшение размера шрифта, открытие в новом окне
- Поддерживается функция «Скопировать ссылку» при нажатии ссылка на данный документ копируется в буфер обмена. При нажатии ссылки открывается данный документ в полной версии в новом окне.
- Поддерживается сохранение документа в rtf. При нажатии кнопки происходит скачивание документа из S3

CMS

Разработанная CMS поддерживает следующие виды обновлений документов:

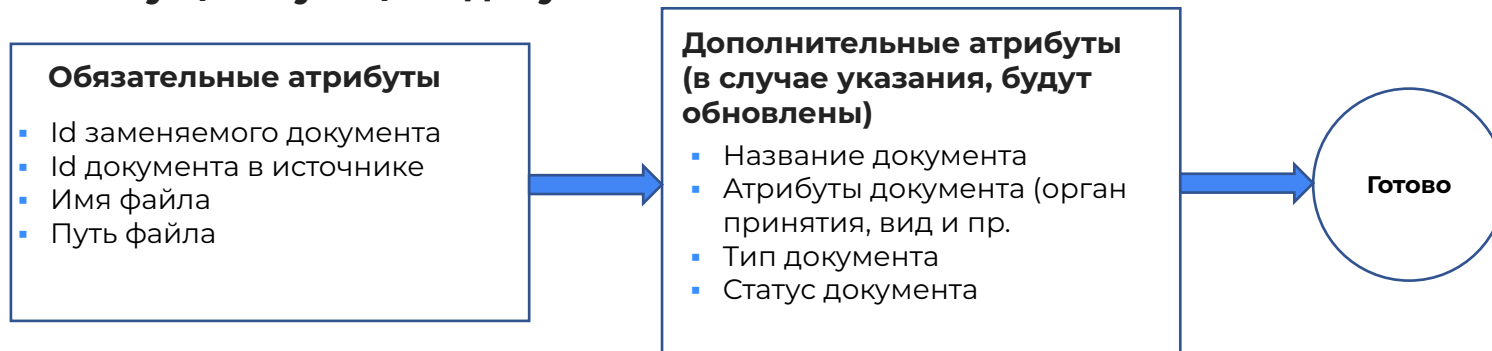
- заменить существующий документ полностью
- обновить часть атрибутов существующего документа
- удалить документ
- загрузить новый документ
- обновить список связанных документов

Протокол:

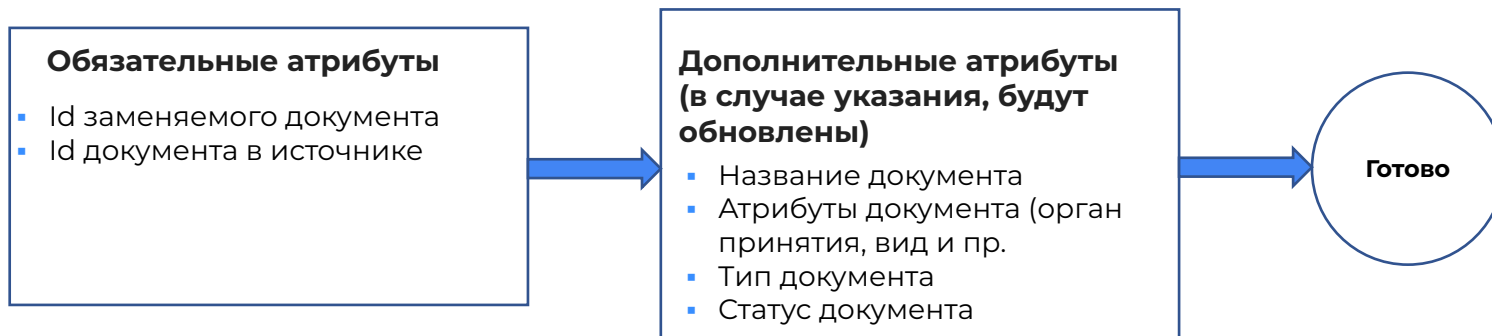
- метод POST
- формат application/json
- кодировке UTF-8
- имена параметров в нижнем регистре
- путь в S3 первая папка это имя бакета
- возвращает 200 в случае успешного завершения

CMS – операции (1/2)

Замена существующего документа полностью:

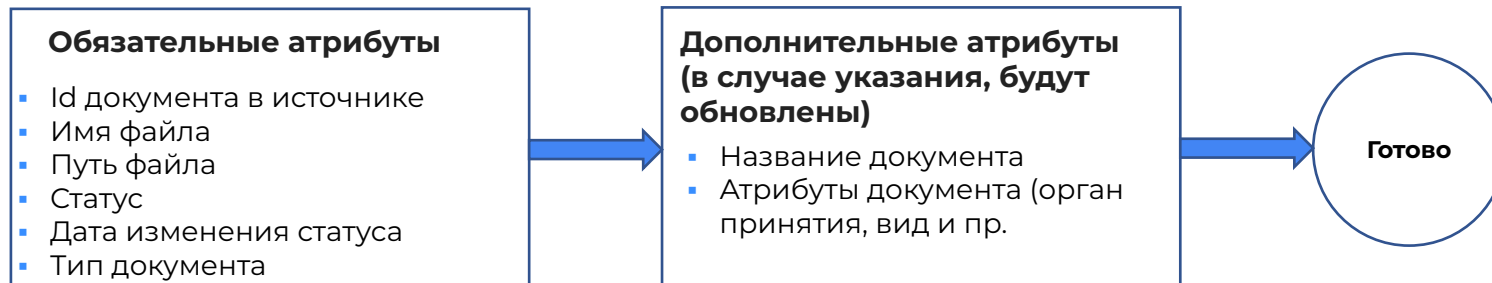


Обновление атрибутов существующего документа:

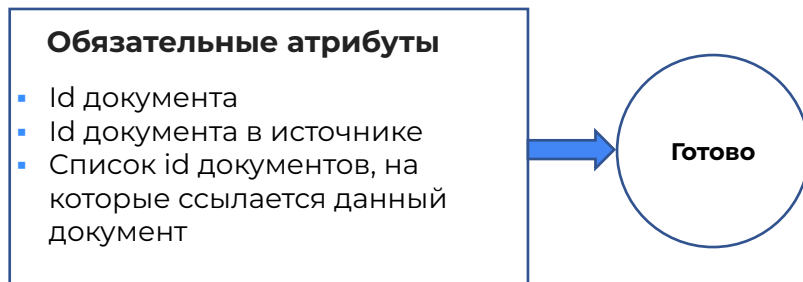


CMS – операции (2/2)

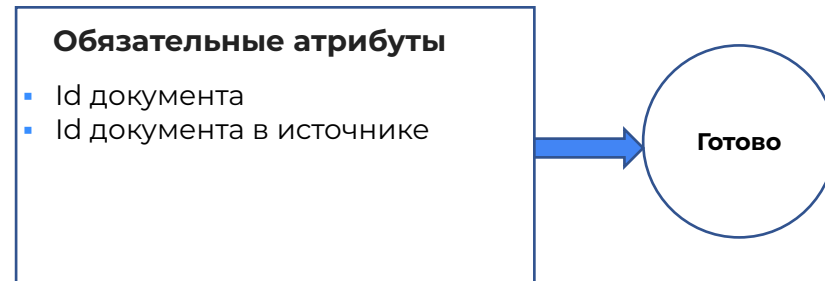
Загрузка нового документа:



Обновление списка связанных документов:



Удаление документа:

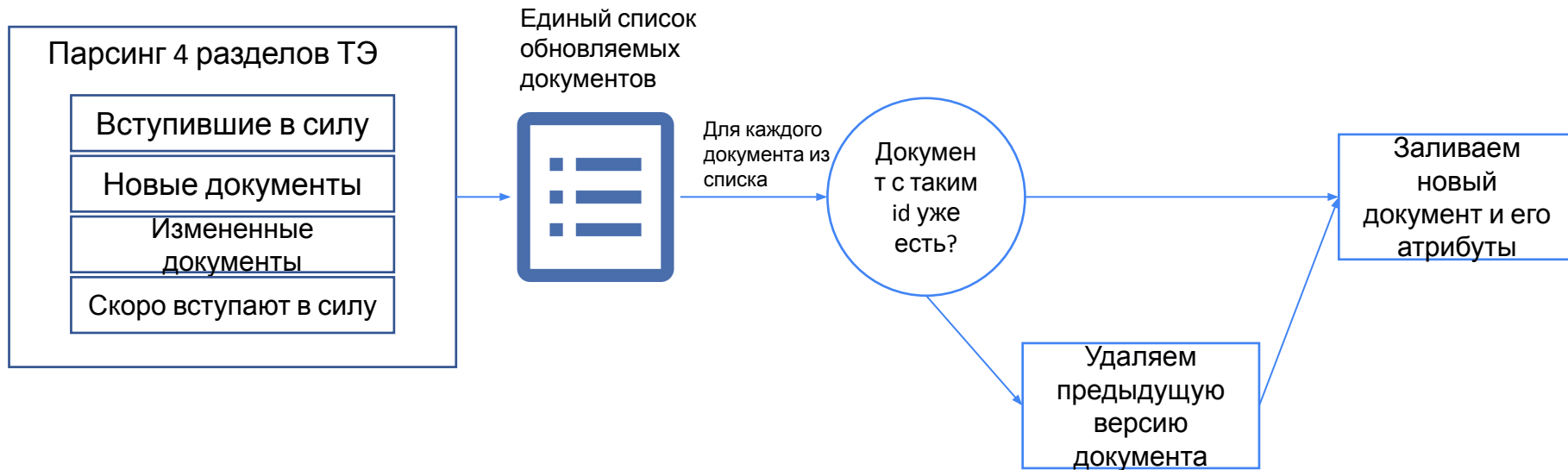


Пакетное обновление документов из ТЭ (1 версия)

Для пакетного обновления используется 4 раздела про обновленные документы на Главной странице ТЭ:

- Производится парсинг 4 подразделов обновленных документов (Новые документы, Измененные документы, Вступившие в силу, Скоро вступают в силу). В список сохраняются id всех документов, которые представлены в этих разделах. Список единый для всех подразделов.
- Из ТЭ выгружаются документы из этого списка и их атрибуты.
- Производится загрузка новых версий всех документов, которые есть в списке.
- Если в системе уже есть такой документ, то он удаляется и прогружается заново со всем атрибутами
- *В дальнейшем возможен более сложный механизм обновления, при котором действие будет зависеть от раздела обновленных документов*

Пакетное обновление документов из ТЭ (схема)



Методы улучшения поиска

Удвоение score за наличие поискового запроса в названии

- Если поисковый запрос содержится в названии документа, данный документ поднимается в выдаче. Это сделано для того, чтобы улучшить поиск по названию.

Словарь синонимов и аббревиатур

- Используется словарь синонимов и аббревиатур. В случае наличия слова-синонима или аббревиатуры, поиск осуществляется также по ним.
- Словарь синонимов и аббревиатур хранится в виде списка вида (слово1, слово2, слово3) и т.д. и может пополняться.

Морфологический плагин

- Используется встроенный морфологический плагин. В поиск попадает не слово целиком, а его основа. Т.е. с точки зрения поиска слова красивый и красивая – равнозначны.