

РЕАЛИЗАЦИЯ КОРПОРАТИВНОГО ИНТЕЛЛЕКТУАЛЬНОГО ПОИСКА В «ГАЗПРОМ НЕФТИ»

РОМАНОВА ОЛЬГА, МЕЙНЦЕР АНТОН



КОНЦЕПЦИЯ ПРОДУКТА

КОРПОРАТИВНЫЙ ПОИСК объединяет источники данных, информационные системы и каталоги, аналитические отчеты, хранилища структурированной и неструктурированной информации, становясь единым окном доступа к данным.

ИНСТРУМЕНТЫ семантического анализа, предоставляемые платформой поиска, позволят не только находить информацию, но и выявлять скрытые связи и знания

ЛЕГЕНДА

□ Приложения по анализу контента, реализованные на базе платформы корпоративного поиска

■ Информационные системы



НЕМНОГО ЦИФР ПРОДУКТА



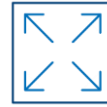
29 135
пользователей

Каждый 3й сотрудник
хотя бы раз заходил в
поиск



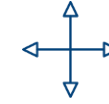
13.6 млн
документов

Различных форматов,
специфики и
назначения



40
источников

Количество
подключений
за 2 года



4
встраивания







В 4 системах встроен
как базовый
инструмент поиска



на 55%
экономит время

По сравнению со
стандартными
способами поиска
информации

ФУНКЦИОНАЛЬНОСТЬ ПОИСКА

					
Полнотекст	Ранжирование	Фильтрация	Аббревиатуры и синонимы	Расширение запроса	Релевантность
Поиск по полному содержимому документов	Инструмент управления поисковой выдачей в зависимости от целей: ранжирование по дате, по релевантности поиска	Удобная система управления выдачей путем выбора наиболее интересных источников, искомых типов документов	Поиск с учетом корпоративного языка	Расширение запроса близкими по смыслу понятиями	Инструментарий управления механизмом формирования поисковой выдачи на основании алгоритмов обработки контента

					
Персонализация	Семантический поиск	Морфология	Исправление запроса	Автоподсказки	Сегментация запроса
Поиск по полному содержимому документов	Поиск с учетом онтологической модели предметной области, расширения понятий и тематически близких документов	Поиск с учетом словоформ	Механизмы корректировки запроса в случае допущения ошибки или набора в неправильной раскладке	Предложения по уточнению запроса на основе наиболее популярных словосочетаний, предложение наиболее подходящих запросу фильтров	Инструментарий разделения запроса на отдельные блоки для повышения качества поиска

ИСТОЧНИКИ ДАННЫХ

Подключенные типы источников

СИСТЕМА
ВНУТРЕННИХ НМД

НОВОСТНЫЕ ПОРТАЛЫ
ДОЧЕРНИХ ОБЩЕСТВ

АНАЛИТИЧЕСКАЯ
ОТЧЕТНОСТЬ

ЭЛЕКТРОННЫЙ
АРХИВ СЭД

КОРПОРАТИВНЫЕ
СЕРВИСЫ

ИНТЕРАКТИВНОЕ
ПОЛОЖЕНИЕ О ЗАКУПКАХ

МОБИЛЬНАЯ ЛЕНТА

СИСТЕМА
РАСПРОСТРАН. ЗНАНИЙ

ПОРТАЛЫ
ИДЕЙ И ИНИЦИАТИВ

БАЗА СОТРУДНИКОВ

КОРПОРАТИВНЫЙ
УНИВЕРСИТЕТ

КОРПОРАТИВНЫЕ
ИЗДАНИЯ

КОРПОРАТИВНАЯ WIKI

БАЗА ГЕОДАННЫХ И
ИССЛЕДОВАНИЙ

ВНЕШНИЕ ИНФ.
РЕСУРСЫ

Типы источников в работе к подключению:

КОРПОРАТИВНЫЙ СЭД

ВНУТРЕННИЙ ИНТРАНЕТ

БИЗНЕС ГЛОССАРИЙ

ПАПКИ DFS

БАЗЫ ТЕХ. ЭКСПЕРТ,
ЮРИДИЧЕСКИЕ БАЗЫ

МЕДИАБАНК

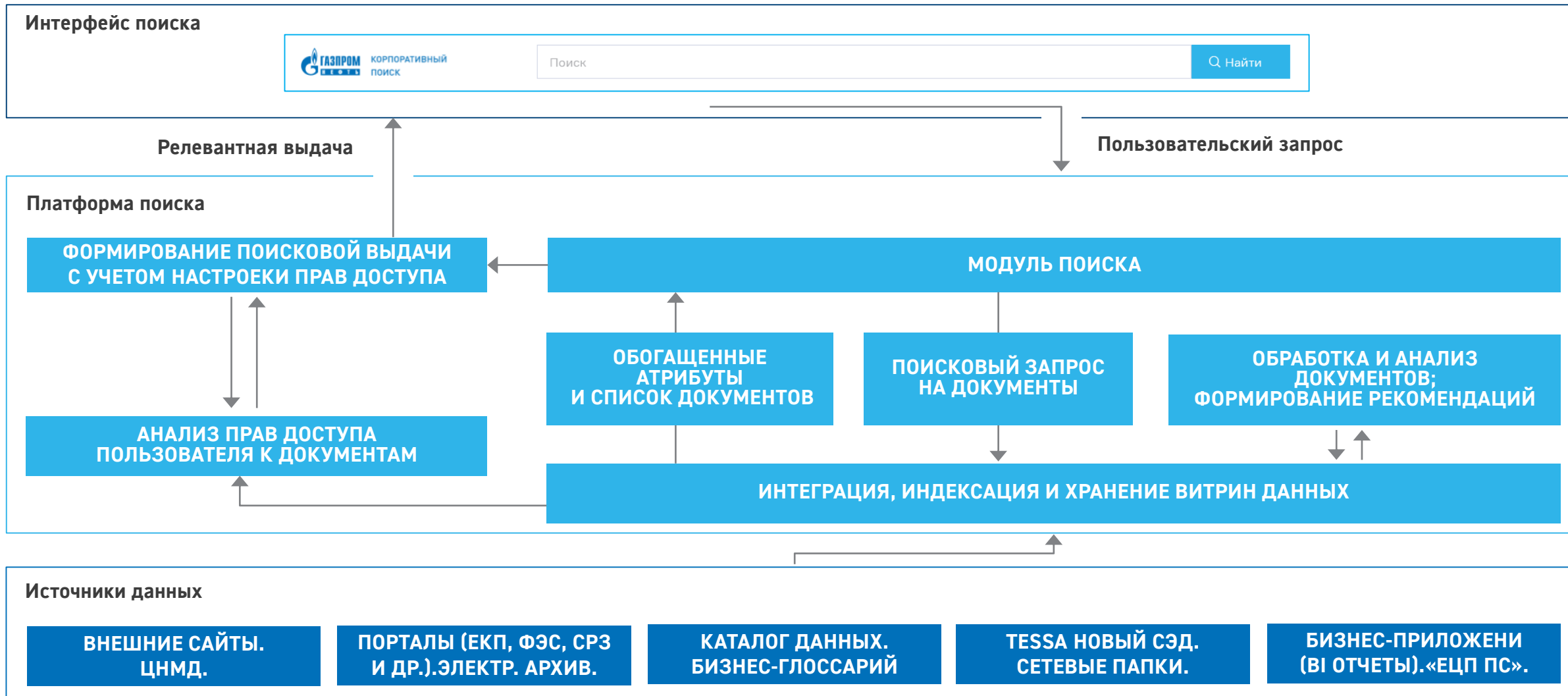
ИС УПРАВЛЕНИЯ
ПРОЕКТАМИ

БАЗА ПЕРЕВОДОВ

ПОИСК ПО НСИ

БАЗЫ КОНТРАГЕНТОВ

БИЗНЕС-АРХИТЕКТУРА



PIPELINE ОБРАБОТКИ ФАЙЛА



РАБОТА С ДАННЫМИ



Интеграции

Инструментарий работы с источниками:

- Использование стандартных коннекторов;
- Разработка веб-сервисов сбора данных;
- Краулеры для автоматизированного сбора информации с порталов;
- Встраивание поиска в информационные системы.



Универсализация

Хранение и представление данных об объектах контента в унифицированном формате единообразного представления в поиске данных различной природы и структуры



Нормализация и очистка

Механизмы нормализации данных при работе с контентом, управление дублями, реализация меппинга данных между источниками (справочники, названия и т.д.)



Каталогизация

Формирование тематического каталога контента – в дальнейшем – в разрезе доменов данных, рабочее место администратора домена для управления контентом

ПОДХОДЫ К ИНТЕГРАЦИИ



Индексация источника Корпоративным поиском

- Индексация данных источника полностью или выбранных разделов.
- Источник появляется в выдаче Корпоративного поиска.
- Пользователи, найдя информацию в Корпоративном поиске, переходят непосредственно в источник, что повышает конверсию и информированность пользователей о полезных материалах источника.



Встраивание поиска в поисковую строку на вашем портале

- Индексация данных источника полностью или выбранных разделов.
- Встраивание переадресации поискового запроса из системы в Корпоративный поиск.
- Поиск возвращает релевантную выдачу по источнику, с фильтрацией, автоподсказками, подсветкой в результатах обработки запроса.



Аналитические инструменты (методы)

- Метод возвращает ключевые слова и аннотацию по документу/объекту источника.
- Метод формирует к документу источника релевантный семантически близкий набор документов.
- Метод возвращает наборы релевантных запросов по ключевым словам документов, с ссылками на источники документов, базовыми параметрами документов.

Варианты реализации:

- Коннектор источника (в случае наличия);
Или
- Реализация источником веб-сервиса по правилам Корпоративного поиска

- Реализация внешнего API корпоративного поиска для встраивания

- Реализация методов API корпоративного поиска для получения специфических данных