



Айкумен – информационные бизнес-системы

ДАнные, СКОРОСТЬ, ИНТЕЛЛЕКТ:
МЫ ПРЕВРАЩАЕМ ДАнные В ЗНАНИЯ

www.iqmen.ru

IQPLATFORM®

v2.0

ОПИСАНИЕ ПРОДУКТА

IQPLATFORM® -

зарегистрированная торговая марка АО «Айкумен – информационные бизнес-системы»



МОСКВА 2017

1. НАЗНАЧЕНИЕ

IQPLATFORM® – информационно-аналитическая платформа для организации безопасной и удобной поисково-аналитической работы с использованием различных типов данных – от множества структурированных баз данных до источников информации на естественных языках.

IQPLATFORM® – BI-инструмент нового поколения, позволяющий пользователям с максимальной полнотой анализировать и применять разнородную информацию для своевременного принятия оптимальных управленческих и бизнес-решений.

Концепция системы заключается в интеграции специалиста-аналитика, аналитического ПО и современных вычислительных средств с наглядным представлением результатов и накоплением аналитического опыта в виде библиотек кейсов, моделей и отчетов.

IQPLATFORM® обеспечивает возможность создания специализированных решений силами предметных аналитиков и экспертов, а большой выбор аналитических средств (графы, BI-отчеты, карты, математические модели) позволяют выполнять полный цикл операций по глубинному анализу собранных и имеющихся данных в следующих целях:

- Выявление скрытых закономерностей и неявных тенденций.
- Сквозной анализ объектов интереса: персон, организаций, бизнес-процессов и технологических операций, финансовых транзакций и банковских счетов, фактов наступления критичных событий и нестандартных ситуаций.
- Сохранение, накопление и использование различных моделей анализа данных.



IQPLATFORM® включена в **Единый реестр российского программного обеспечения** для ЭВМ и баз данных в классе «Поисковые системы. Системы сбора, хранения, обработки, анализа, моделирования и визуализации массивов данных» и официально рекомендована Правительством России к использованию государственными организациями в соответствии с Приказом Министра связи и массовых коммуникаций РФ №653 от 14 декабря 2016 года.



IQPLATFORM® – победитель федерального **конкурса Аналитического центра при Правительстве РФ «Лучшие информационно-аналитические инструменты – 2016»** в номинациях:

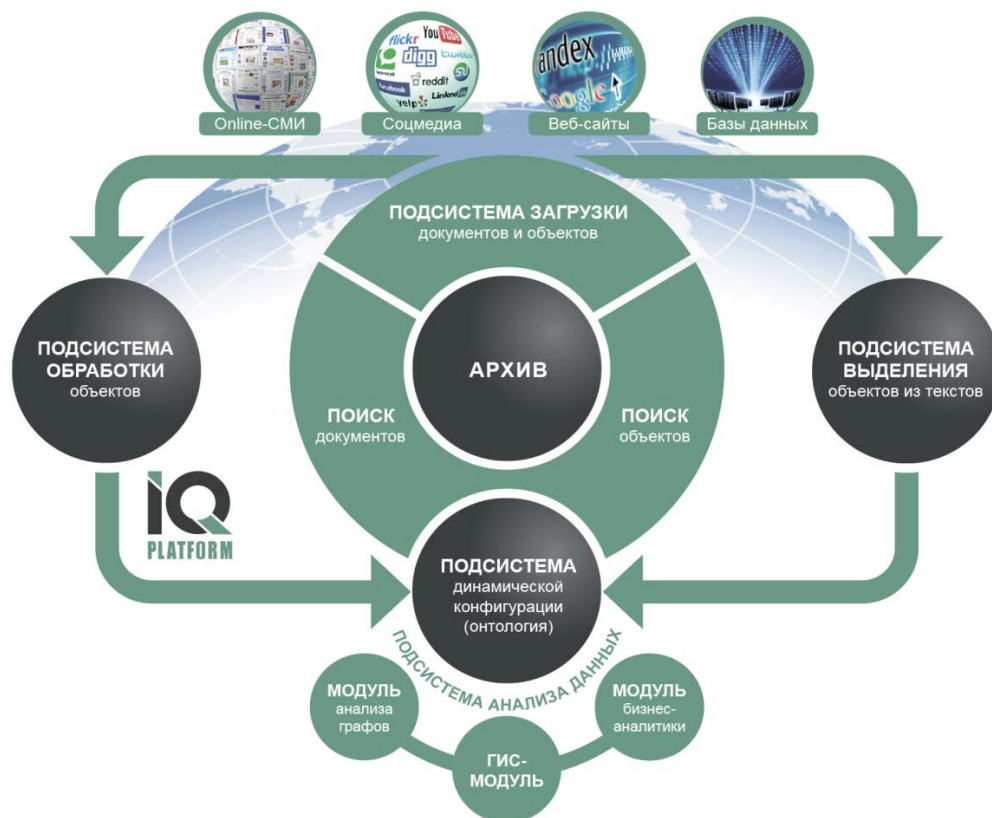
- «Лучшее информационно-аналитическое решение для обработки неструктурированных данных»
- «Лучшее информационно-аналитическое решение для визуализации данных и принятия управленческих решений: АРМ руководителя»

IQPLATFORM® включена в **каталог информационно-аналитических систем**, рекомендованных Аналитическим центром к использованию государственными органами власти Российской Федерации.

Комплексные отраслевые решения на базе информационно-аналитической платформы **IQPLATFORM®** внедрены в крупнейших российских банках, органах государственной власти, промышленных предприятиях и наукоемких производствах.

2. ФУНКЦИОНАЛЬНЫЕ ВОЗМОЖНОСТИ

Модульная архитектура IQPLATFORM® построена на принципах максимальной гибкости и высокой клиентоориентированности:



- **Модульность** – возможность формирования состава функциональных модулей системы для решения предметных задач конкретного Заказчика.
- **Свободная масштабируемость** – увеличение объемов хранения и скорости обработки данных в проекте путем линейного наращивания аппаратных мощностей.
- **Кросс-платформенность** с поддержкой операционных сред Linux/Windows и использованием СУБД различных производителей.

2.1 ПОДСИСТЕМА ДИНАМИЧЕСКОЙ КОНФИГУРАЦИИ (ОНТОЛОГИЯ)

Онтология – центральная подсистема платформы, предназначенная для описания и настройки целевой модели данных, сценариев их обработки и моделей представления. Онтология позволяет развивать индивидуальное решение независимо от разработчиков – только силами аналитика Заказчика прямо в конфигурационном интерфейсе, позволяя встраивать новые источники и расширять область применения системы:

- Создавать новые классы объектов предметной области (персоны, организации, транзакции, банковские счета, телефонные номера и т.п.) с настройкой их отображения в интерфейсе платформы.
- Назначать объектам атрибуты из встроенного списка типов атрибутов (произвольная строка, булев тип, число, справочник, изображение и т.д.).
- Устанавливать правила валидации, идентификации и слияния объектов.
- Создавать новые связи между объектами и правила их формирования для статических (хранящихся в базе данных), динамических (вычисляемых «на лету») и контекстных связей (выделенных из текстов).
- Устанавливать правила автоматического формирования и обнаружения критических или тревожных критериев («алертов») для произвольных объектов.

2.2 ПОДСИСТЕМА ЗАГРУЗКИ ДАННЫХ

Подсистема обеспечивает гибкие возможности загрузки информации в систему и предназначена для сбора структурированных и полуструктурированных данных о заданных объектах из различных источников (внутренние базы данных, корпоративные информационные и учетные системы, сайты с фактической информацией и т.д.).

Гибкие API-интерфейсы загрузки текстовой и числовой информации, а также полнотекстовых документов для формирования собственной документной базы, позволяют Заказчику самостоятельно организовать процесс поступления данных из внутренних и внешних источников. Платформа обладает огромной накопленной библиотекой готовых коннекторов - как однонаправленных с единовременной загрузкой, так и двунаправленных с онлайн-поиском по источнику:

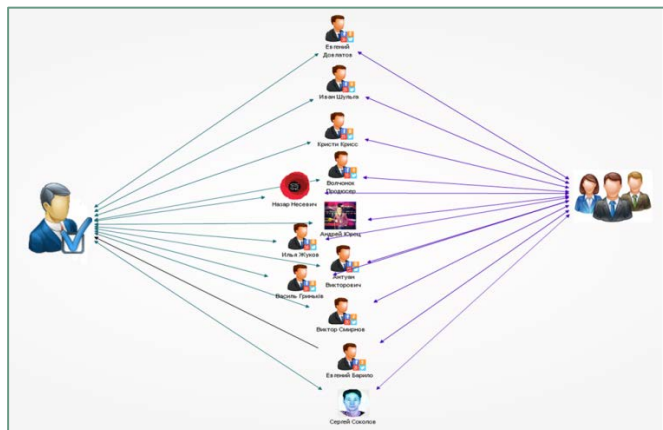
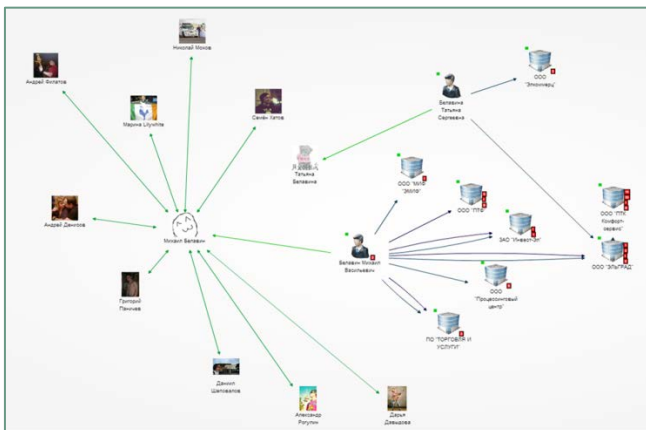
- Online-СМИ: более 5 000 интернет-ресурсов.
- Крупнейшие социальные сети (Facebook, Twitter, Вконтакте, Одноклассники) и наиболее популярные русскоязычные блоги и форумы.
- Основные поисковые системы с возможностью загрузки найденных документов.
- Веб-сайты произвольной структуры.
- Файловые системы.

2.3 ПОДСИСТЕМА АНАЛИЗА ДАННЫХ

Анализ и визуализации данных в едином рабочем пространстве с удобным пользовательским интерфейсом и возможностью интерактивной настройки новых аналитических сценариев:

Анализ графов

- Отображение объектов и связей в виде графа с отображением их атрибутов, группировкой или фильтрацией, определением кратчайшего пути между объектами, а также расчетом статистических показателей по объектам.
- Настройка сценария (правил обхода и глубины) поиска объектов в окружении по заданным характеристикам.
- Построение и анализ кластеров объектов.
- Сбор информации по интересующему объекту и его связям с доступом к внешним источникам за дополнительными данными.
- Поиск похожих объектов с близкими по значению атрибутами или названиями.

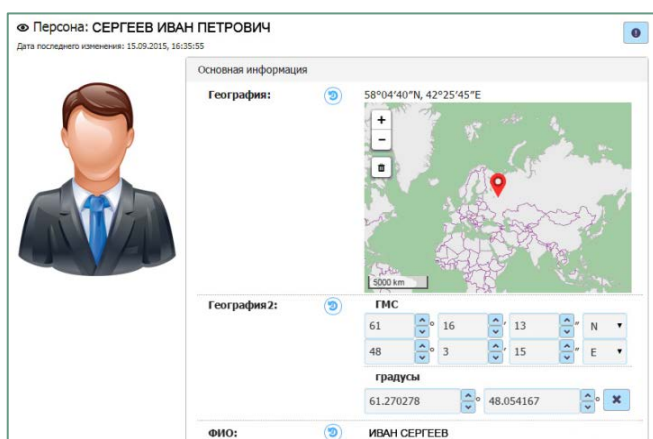


Геоаналитика

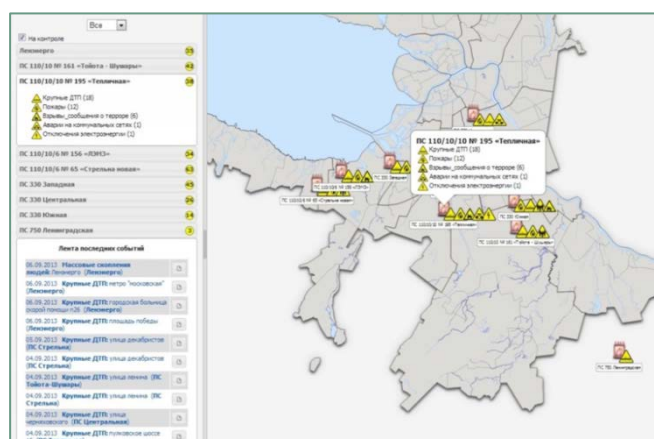
Управление единой информационной средой геопространственных данных с использованием картографического ГИС-сервиса OpenStreetMap для обогащения информационно-аналитических ресурсов платформы геоданными объектами исследования (ГЛОНАСС/GPS/LBS-координаты, геотеги и геометки, почтовые адреса, названия населенных пунктов и улиц) с целью расширения ее аналитических возможностей:

- точная привязка, систематизация и интеграция в едином адресном пространстве всей поступающей и хранимой информации об объекте.
- выявление и динамическое моделирование неочевидных процессов, тенденций и определяющих факторов в географических разрезах.
- графическая визуализация взаимосвязей пространственных отношений и геораспределений объектов.

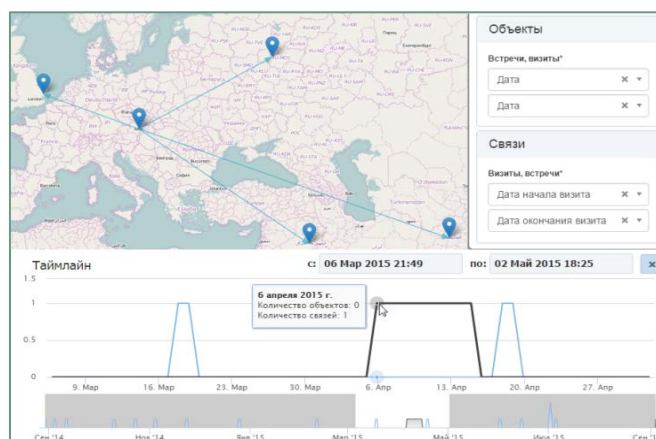
Модуль обладает необходимым набором функций и обеспечивает консолидацию геотрибутов объекта и его окружения для отображения на электронной карте местности вместе с локализованными документами и статистическими отчетами:



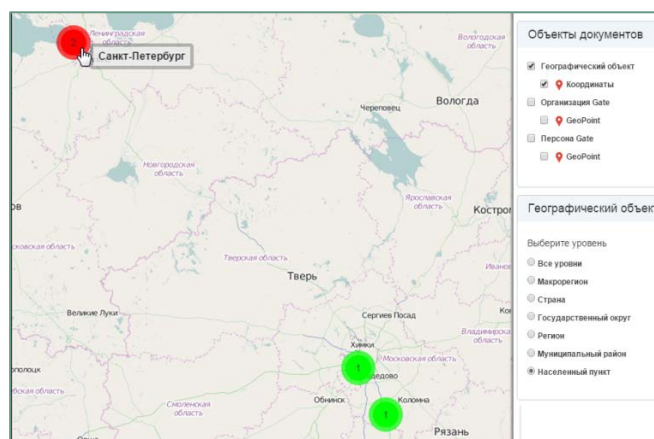
Консолидация геотрибутов в досье объекта



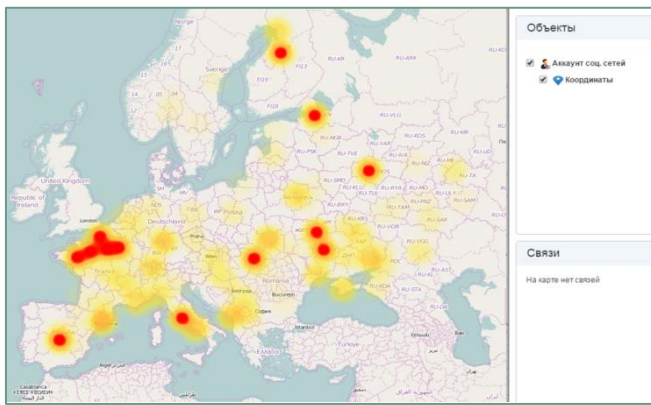
Поиск объектов на карте в выбранных гео зонах и координатах по заданным параметрам.



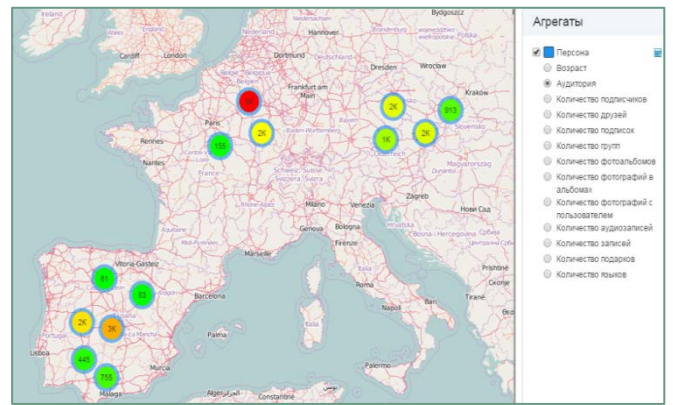
Фильтрация объектов и связей по классам, атрибутам и времени существования



Кластеризация и декластеризация объектов при изменении масштаба карты



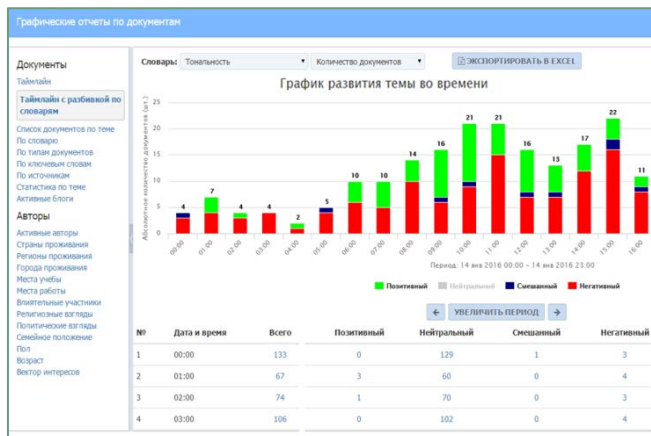
«Тепловая карта» (HeatMap) для визуальной оценки плотности и структуры распределения объектов в зонах контроля



Агрегация объектов и окружения по числовым атрибутам (сумма, средне-арифметическое, мин/макс) с цветовой индикацией значений.

BI-анализ

Построение широкого спектра интерактивных (drill-through) статистических отчетов по объектам, их взаимосвязям и документам, их наглядное представление посредством графиков, диаграмм, гистограмм и таблиц:



2.4 ПОДСИСТЕМА ТЕКСТОВОЙ АНАЛИТИКИ

Обработка, нормализация и структурирование текстовой информации

Анализ больших объемов информации из различных систем Заказчика без предварительной машинной лингвистической обработки обычно крайне затруднен. Система позволяет «сгладить» особенности представления информации в каждом источнике и работать с чистыми данными, уже собранными вокруг единых объектов интереса с помощью набора функциональных модулей:

- Нормализации российских почтовых адресов для приведения их к стандартному формату с указанием кода БД ФИАС.
- Нормализации имен с коррекцией ошибок, учетом транслитерации и уменьшительно-ласкательных форм.
- Определения языка и морфологические модули для русского, английского, португальского, испанского и немецкого языков.
- Выделения объектов и фактов.
- Рубрикации и кластеризации документов, а также выделения сюжетов.
- Извлечения текстовой информации из изображений и растровых документов с применением технологии распознавания текстов ABBYY.

Выделение объектов и фактов

Извлечение объектов, связей и фактической информации из неструктурированного текста с использованием методов текстовой аналитики для обогащения данных и выявления неявных или скрытых закономерностей. Модуль обеспечивает выделение следующих сущностей из текстов на русском и английском языках:

- Географические названия (до населенных пунктов) и полные адреса с возможными сокращениями и привязкой к БД ФИАС.
- Общие сведения: телефоны, e-mail, кредитные карты, банковские счета и т.д.
- Факты о работе персоны в организации с выделением занимаемой должности.
- Названия вузов и выделение факта обучения персоны в вузе.
- Даты с указанием времени.
- Числа, включая процентные соотношения и денежные суммы.
- Цитаты в привязке к персоне.

2.5 ПОДСИСТЕМА ПОИСКА ДАННЫХ

Поисковый модуль с пользовательским конфигурированием обеспечивает распределенный поиск данных по различным критериям для загрузки в систему.

Поиск объектов:

- Полнотекстовый поиск объектов по текстовому содержанию атрибутов с поддержкой всех основных операторов поиска.
- Атрибутивный поиск по объектам.
- Сквозной полнотекстовый поиск по всей базе данных независимо от класса объектов.
- Поиск во внешних источниках данных.

Поиск документов:

- Полнотекстовый поиск документов с поддержкой основных операторов поиска.
- Полнотекстовый вложенный поиск с использованием запросов в запросе.
- Атрибутивный поиск по документам (язык, источник, дата, автор и т.д.).
- Поиск во внешних источниках, например, Google или Yandex.

2.6 АРХИВ

Свободно масштабируемый архив предназначен для хранения неограниченного объема собранных данных с возможностью распределенного поиска по различным критериям для загрузки в систему. Увеличение емкости архива и скорости поиска данных производится путем простого линейного наращивания аппаратных мощностей.

- Архив документов реализован с применением технологии Apache Hadoop.
- Архив объектов с функцией метапоиска по архивному кластеру построен на основе распределенной поисковой платформы Elasticsearch.

2.7 ДОПОЛНИТЕЛЬНЫЕ И СЛУЖЕБНЫЕ ФУНКЦИИ

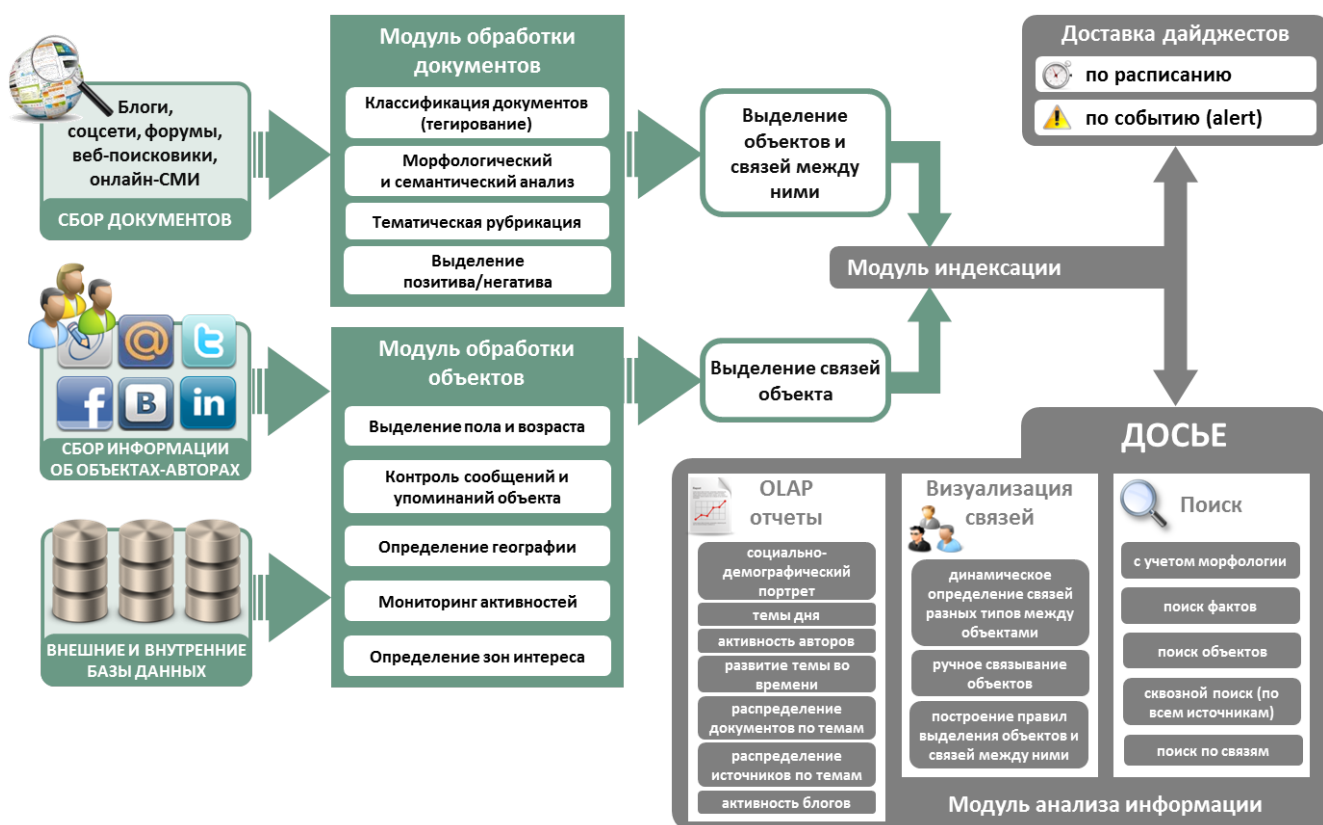
Оперативное оповещение пользователя о системных событиях (новых объектах, связях и изменении их состояния) и регулярных обновлениях отчетов и рубрик с документами с помощью email-/sms-сообщений или интерфейсных push-уведомлений.

Автоматический контроль и диагностика системы с аварийным email/sms-оповещением и сохранением данных путем горячего резервного копирования без остановки работы.

Многоуровневая система безопасности для контроля действий пользователя и надежной защиты информации:

- Разграничение прав доступа пользователей и их групп к функциям, документам, рубрикам и атрибутам.
- Взаимодействие клиента с системой по протоколу HTTPS для снижения риска перехвата трафика через незащищенные каналы связи.
- Журналирование всех действий пользователя, хранение истории изменений объектов с возможностью импорта событий в корпоративные SYSLOG-сервера.

3. ПРИНЦИП РАБОТЫ



4. ТЕХНОЛОГИЧЕСКИЕ ОСОБЕННОСТИ

- **Динамическое конфигурирование** предметных аналитических моделей, сценариев обработки и визуализации данных.
- **Потоковый сбор данных** с обширной библиотекой коннекторов из любых типов внешних и внутренних источников.
- **Высокоточный язык поисковых запросов** с механизмом целевой фильтрации при обработке более 20 млн сообщений в сутки.
- **Встроенный масштабируемый архив** неограниченной ёмкости с распределенным поиском на основе Nadoor.
- **Тематическая рубрикация** для быстрой категоризации документов при сборе и архивном поиске.
- **Автоматическая текстовая обработка** с распознаванием и выделением объектов, матчингом, очисткой и нормализацией неструктурированных данных.

- **Динамическое досье** в виде интерактивной карточки объекта с поиском, привязкой к связям, библиотеками документов и тревожной индикацией.
- **Динамические графы контекстно-объектных связей** со сценариями поиска объектов и выводом документов, атрибутов и алертов.
- **Совместимость с аппаратно-программными комплексами** Oracle Exadata Database Machine / Exalogic Elastic Cloud / Big Data Appliance, IBM Power-8.
- **Интеграционная поддержка СУБД** Oracle, PostgreSQL, TmaxSoft Tiberio.

5. КОНКУРЕНТНЫЕ ПРЕИМУЩЕСТВА

- **Полностью отечественный высокотехнологичный продукт российской разработки**, независимый от зарубежных производителей и соответствующий государственной политике импортозамещения программного обеспечения.
- **Быстрое создание комплексных решений** в конкретной пользовательской специализации с помощью экспертно-аналитической настройки предметных областей и правил идентификации объектов, аналитических сценариев и отчетности.
- **Максимально полный сбор информации** из баз данных любой сложности и наиболее популярных онлайн-источников с возможностью быстрого и сверхточного подключения дополнительных интернет-ресурсов любого типа.
- **Мультиязычный лингвистический и семантический анализ** текстов на основе передовых методов машинной обработки естественного языка.
- **Комплексная бизнес-аналитика в режиме реального времени** с развитой системой BI-отчетности и наглядной визуализации, объединяющая обработку и анализ структурированных и неструктурированных данных.
- **Свободная и незатратная масштабируемость** как односерверной системы, так и многосерверных кластеров с простым увеличением ёмкости архива и скорости поиска во внутренней базе данных.