



ИНТЕЛТЕХ



INTELTECH

АНАЛИТИЧЕСКАЯ ПЛАТФОРМА ОБРАБОТКИ РАЗНОРОДНЫХ ДАННЫХ

1. НАЗНАЧЕНИЕ И РЕШАЕМЫЕ ЗАДАЧИ

Назначение:

Обеспечение сквозного цикла обработки данных, объединяющего различные программные технологии, включая средства разработки сценариев, обработки и хранения данных

Решаемые задачи:

- Импорт данных из разнородных источников;
- Очистка, преобразование, конвертация данных;
- Обработка данных на основе настраиваемых сценариев;
- Визуализация результатов обработки, включая электронные цифровые карты;
- Унификация способа доступа к методам обработки как аналитического ядра системы.

2. ИМПОРТ ДАННЫХ ИЗ РАЗНОРОДНЫХ ИСТОЧНИКОВ

1. Файловые источники

Поддерживаются форматы CSV, XLS, ODT, JSON, XML, GeoJSON.

Для извлечения данных из документов типа ODT, ODS, XLS реализована технология разметки и автоматического извлечения данных из документа как контейнера.

2. Хранилища данных

Реализован алгоритм автоматического анализа источника данных (профилирования). При этом осуществляется оценка (для реляционных объектов):

- количество строк;
- количество уникальных значений атрибутов;
- количество отсутствующих значений атрибутов;
- гистограмма распределения значений атрибутов;
- корреляция значений атрибутов по строкам.

Для извлечения данных из хранилища используется «ленивый» алгоритм. Фактическое извлечение данных выполняется при первом обращении к ним в соответствии со сценарием обработки.

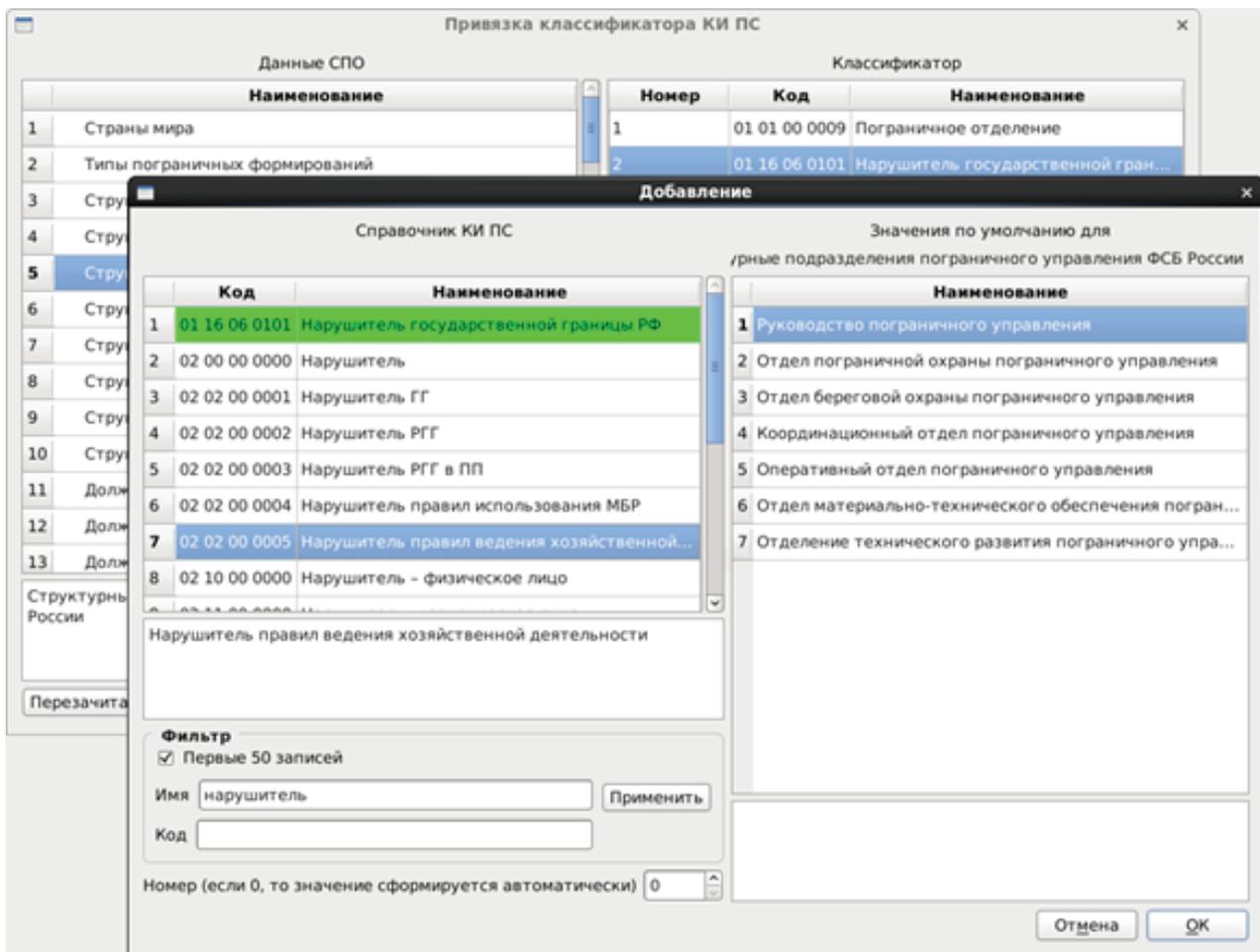
3. Restful API

Поддерживается синхронный алгоритм обращения к сервисам на основе синхронного Restful API, а также асинхронный алгоритм, предусматривающий дополнительное применение брокера сообщений, например, с использованием RabbitMq или универсальной шины взаимодействия.

3. ОЧИСТКА, ПРЕОБРАЗОВАНИЕ, КОНВЕРТАЦИЯ ДАННЫХ

Поддерживаются следующие методы:

- определение допустимых интервалов значений атрибутов данных и последующей фильтрации недопустимых значений;
- привязка (mapping) значений атрибутов и справочников системы;
- преобразование данных на основе декларативных языков, включая XSLT, SQL;
- валидация XML, JSON документов на основе соответствующей схемы;
- преобразование структуры данных, в том числе, из документной в реляционную.



Привязка классификатора КИ ПС

Данные СПО		Классификатор		
Номер	Наименование	Номер	Код	Наименование
1	Страны мира	1	01 01 00 0009	Пограничное отделение
2	Типы пограничных формирований	2	01 16 06 0101	Нарушитель государственной гран...

Добавление

Код	Наименование
1	01 16 06 0101 Нарушитель государственной границы РФ
2	02 00 00 0000 Нарушитель
3	02 02 00 0001 Нарушитель ГГ
4	02 02 00 0002 Нарушитель РГГ
5	02 02 00 0003 Нарушитель РГГ в ПП
6	02 02 00 0004 Нарушитель правил использования МБР
7	02 02 00 0005 Нарушитель правил ведения хозяйственной...
8	02 10 00 0000 Нарушитель - физическое лицо

Значения по умолчанию для
группы подразделений пограничного управления ФСБ России

Наименование
1 Руководство пограничного управления
2 Отдел пограничной охраны пограничного управления
3 Отдел береговой охраны пограничного управления
4 Координационный отдел пограничного управления
5 Оперативный отдел пограничного управления
6 Отдел материально-технического обеспечения погран...
7 Отделение технического развития пограничного упра...

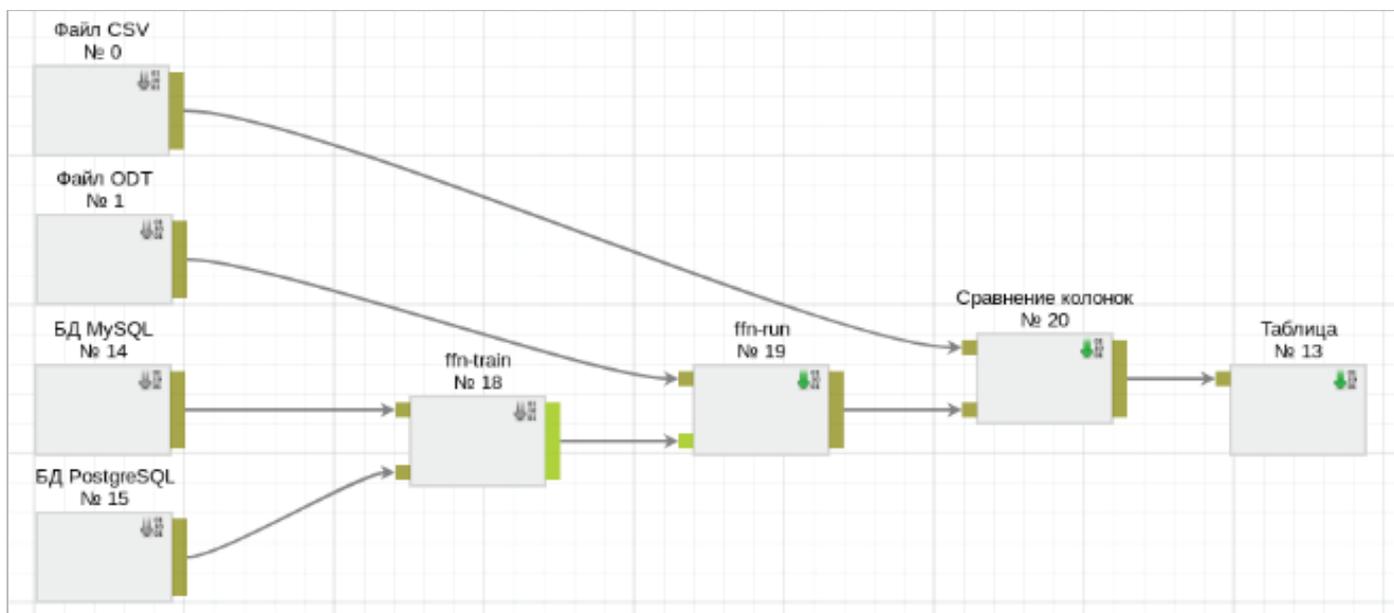
Фильтр
 Первые 50 записей
 Имя Применить
 Код
 Номер (если 0, то значение сформируется автоматически)

Отмена ОК

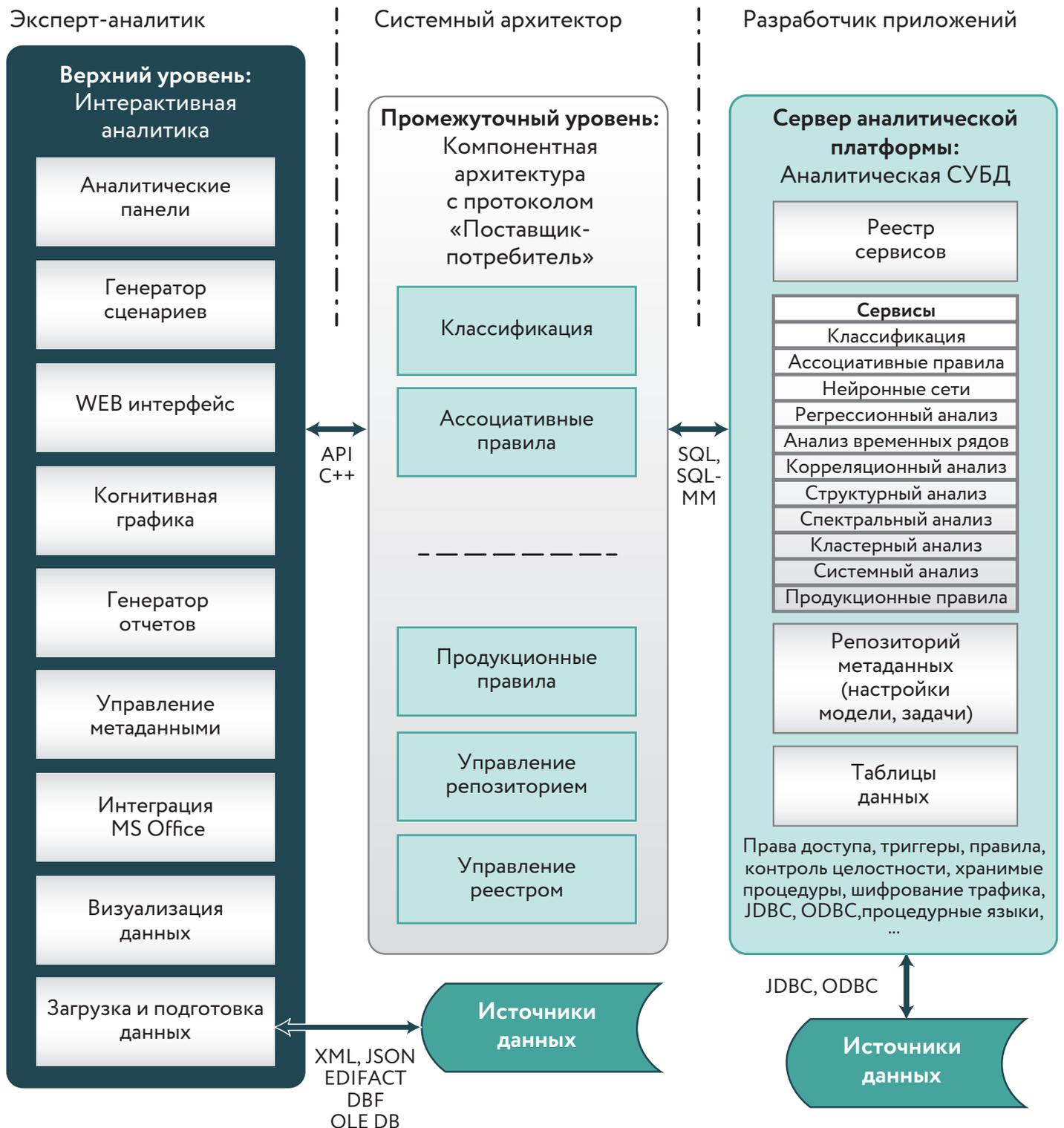
4. ВИДЫ ОБРАБОТКИ

Включает в себя следующие виды обработки данных:

- оперативная (АРМ оператора) и пакетная (сервер расчетов);
- локальная и распределенная (на одном сервере или нескольких взаимодействующих серверах);
- обработка данных с применением стандартизованных библиотек, включая методы искусственного интеллекта и нейросетевые модели вычислений;
- обработка данных с применением пользовательских модулей расширения, реализующих унифицированный программный интерфейс;
- обработка геопространственных данных;
- настраиваемые сценарии обработки с использованием графического редактора сценариев.



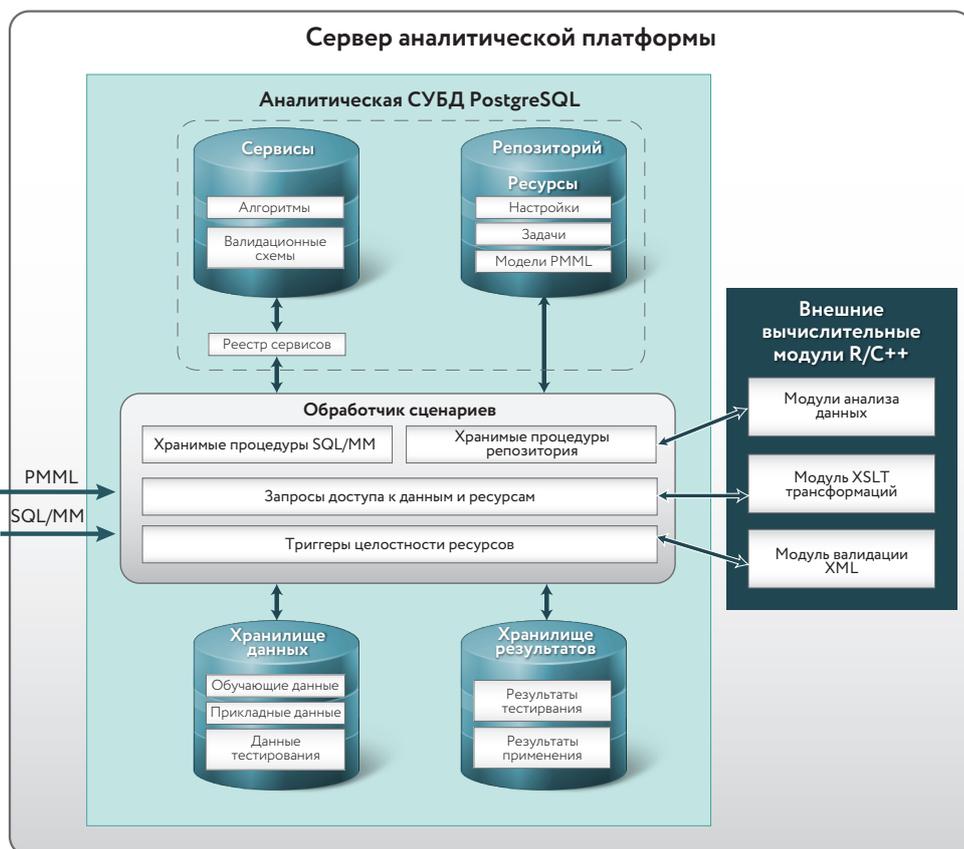
5. АНАЛИТИЧЕСКОЕ ЯДРО – ПЛАТФОРМА SORINTELLA



6. СЕРВЕР АНАЛИТИЧЕСКОЙ ПЛАТФОРМЫ

Тактико-технические характеристики:

- обработка масштабных массивов разнородной информации (до 32 TB);
- расширяемость и масштабируемость аналитики;
- многоплатформенность;
- контроль целостности данных;
- интерфейсы доступа к языкам программирования высокого уровня;
- триггеры и правила для управления процессами;
- система управления правами доступа и авторизации;
- параллельная обработка пользовательских сессий;
- шифрование трафика;
- поддержка стандарта SQL/MM в части обработки пространственных и геоинформационных данных.



Инфраструктура обеспечивает:

- доступ и управление аналитическими ресурсами платформы (задачами, вычислительными моделями и настройками).
- спецификацию о реализованном наборе сервисов (описание алгоритмов).
- спецификации о допустимых настройках для реализованного набора алгоритмов.
- спецификацию по моделям представления знаний.
- спецификации интерфейсных функций стандарта SQL-MM для реализованного набора алгоритмов.
- поддержку целостности ресурсов и сервисов.

7. ТАКТИКО-ТЕХНИЧЕСКИЕ ХАРАКТЕРИСТИКИ АНАЛИТИЧЕСКОЙ ПЛАТФОРМЫ

Сервер аналитической платформы предназначен для выполнения интеллектуального анализа данных в промышленных системах поддержки принятия решений

Методы интеллектуального анализа данных

Назначение	Метод
Классификация данных	Деревья решений
Предсказание событий	Ассоциативные правила
Распознавание состояний	Нейронные сети
Выявление зависимостей	Регрессионный анализ
прогнозирование тренда	Анализ временных рядов
Построение когнитивных карт	Корреляционный анализ
Определение ключевых факторов	Структурный анализ
Фильтрация и сжатие данных	Спектральный анализ
Выявление аномальных состояний	Кластерный анализ
Выявление моделей поведения	Системный анализ
Логический вывод	Продукционные правила

Ключевые факторы производительности

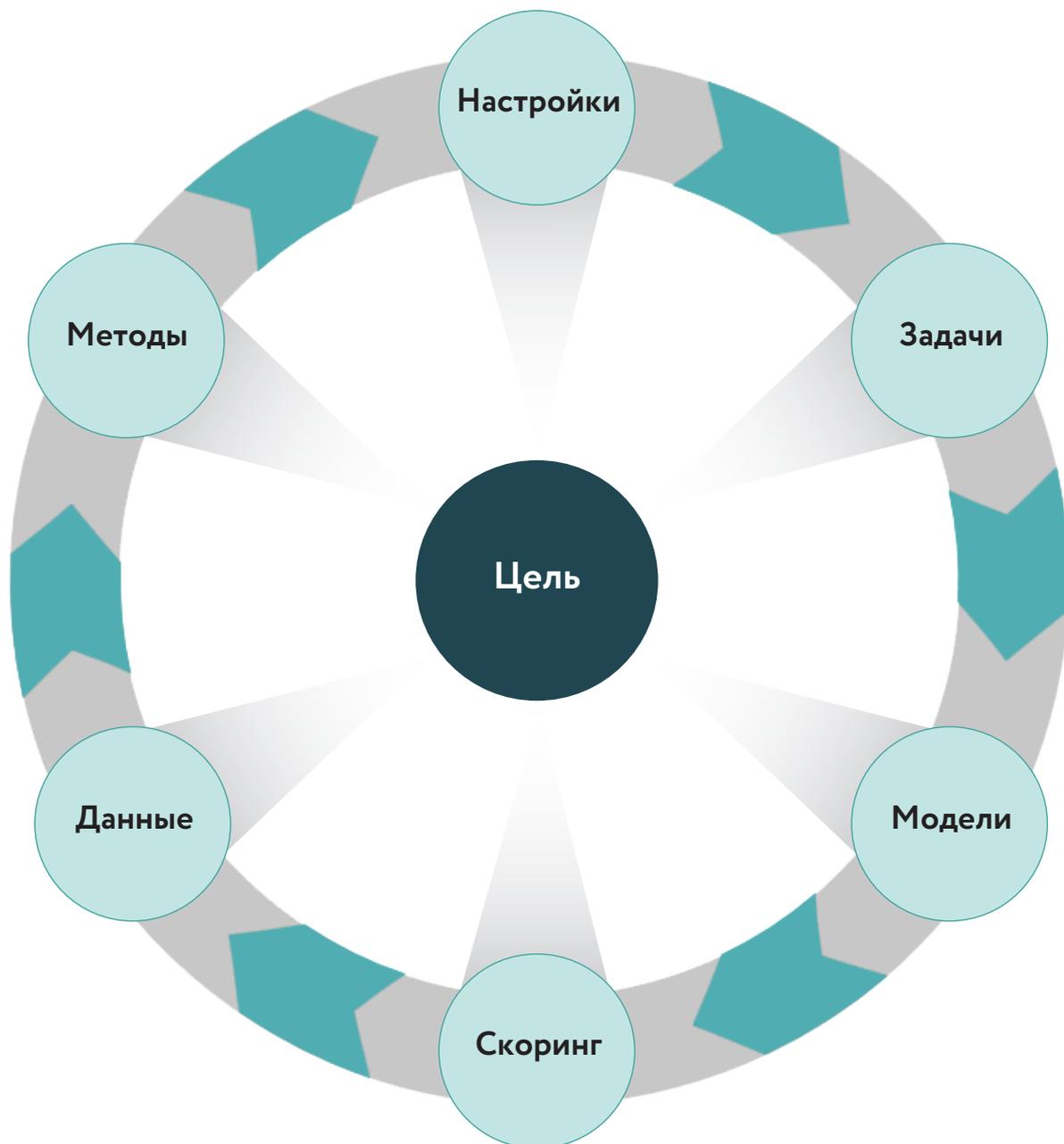
- Серверные (хранимые) процедуры позволяют и позволяют существенно уменьшить трафик между клиентом и сервером
- Перенос функционала на RAM диск увеличивает скорость работы в разы
- Использование ускорителей на DSP-процессорах увеличивает скорость вычислений на порядок

Преимущества

- Обработка масштабных массивов разнородной информации (до 32 TB)
 - Архитектура аналитической подсистемы в соответствии со стандартом SQL/MM Part 6
 - Расширяемость аналитики
 - Поддержка стандарта PMML для моделей аналитики
 - Многоплатформенность
 - Низкое требование к ресурсам
 - Контроль целостности данных
 - Высокая производительность
 - Богатый набор типов данных, включая XML, ГИС, сетевые типы
 - Наследование таблиц
 - Масштабируемость к данным
 - Интерфейсы доступа C, C++, C#, python, perl, ruby, php, Lisp и другие языки
- Методы доступа JDBC, ODBC
- Системные интерфейсы DDS, SOAP
- Процедурные языки PL/pgSQL, pl/Tcl, PL/Perl и pl/Python. Кроме них, существует поддержка PHP, Java, Ruby, R, shell
- Серверные (хранимые) процедуры
- Триггеры и правила для управления процессами
 - Система обмена сообщениями между процессами
 - Развитая система прав доступа
 - Аутентификация пользователя
 - Шифрование трафика
 - Аппаратное ускорение вычислений
 - Возможность использования твердотельных и RAM дисков

8. ТЕХНОЛОГИЧЕСКИЙ ЦИКЛ ПРЕДИКАТИВНОЙ АНАЛИТИКИ

Вершины цикла соответствуют сущностям стандарта SQL/MM. Дуги определяют типовой сценарий взаимодействия сущностей



9. PMML – ЭКСПОРТ-ИМПОРТ АНАЛИТИЧЕСКИХ МОДЕЛЕЙ

История развития

PMML стартовал в 1998 году с версии 0.7. С тех пор стандарт пережил множество релизов вплоть до версии 4.2 в 2014 г.

Проблемы и вызовы

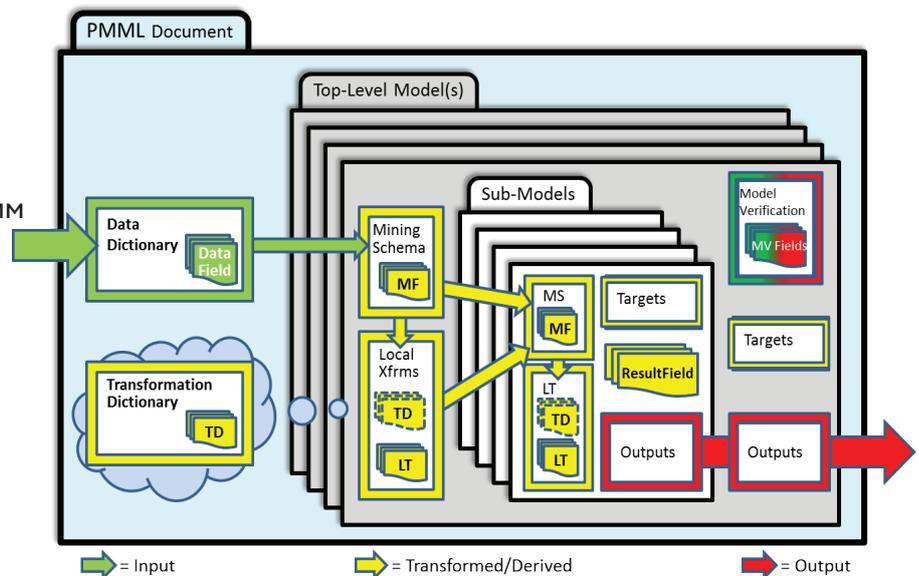
- рыночная адаптация
- эволюция версий
- проблемы алгоритмических расширений
- открытое ПО для восстановления алгоритма

Возможности:

- свободен к коммерческому использованию и имеет открытые коды
- гетерогенное окружение различных производителей
- пре- и пост- обработка встроены в стандарт
- поддержка мультимодельности
- встроенные механизмы расширения

Стандартизованные модели алгоритмов извлечения знаний

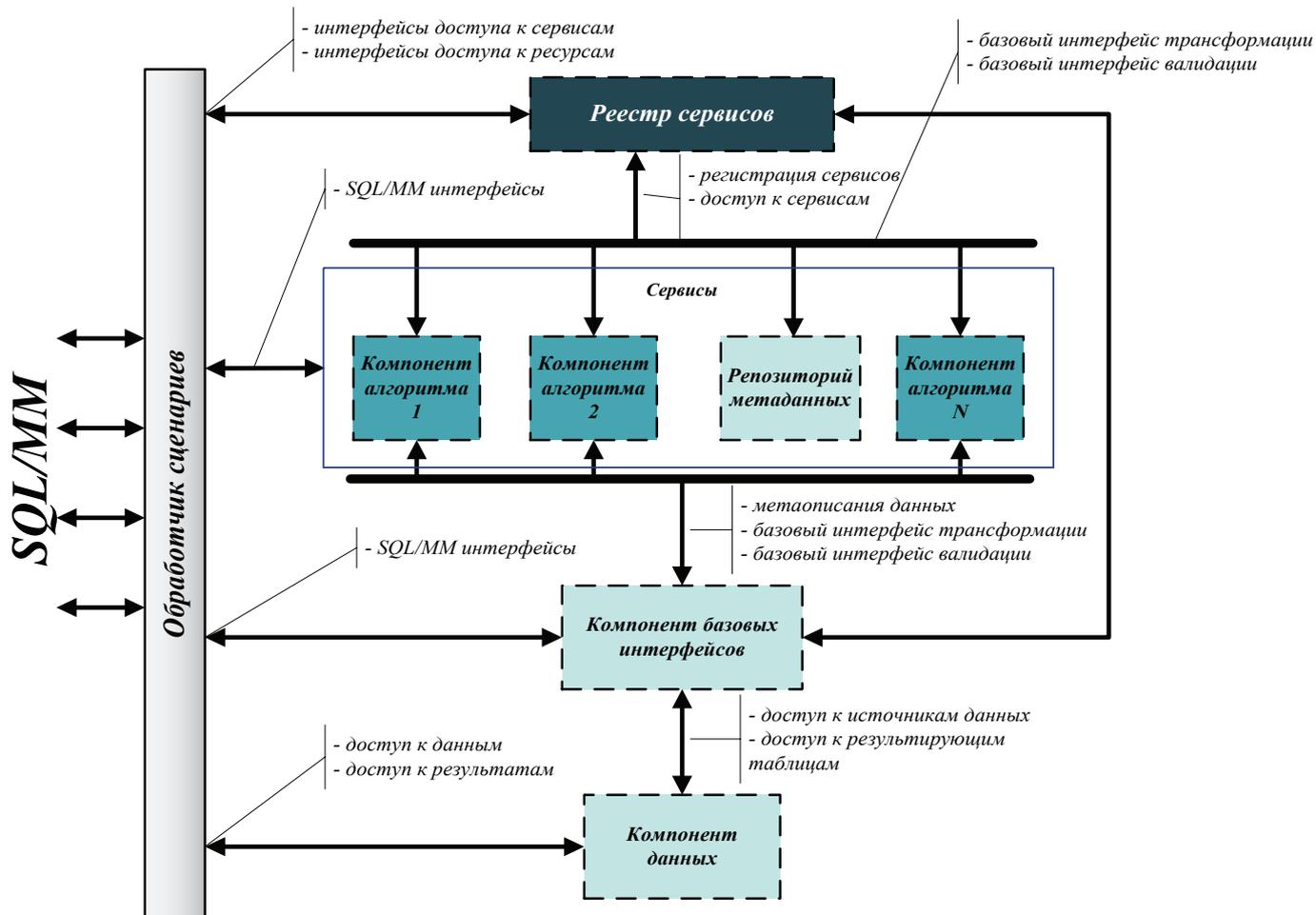
- Ассоциативные модели
- Базовые модели
- Деревья решений
- Кластеризация
- Регрессия
- Обобщенная регрессия
- Классификация по ближайшим соседям
- Нейронные сети
- Наивный байесовский классификатор
- Оценочные карты
- Последовательности
- Обработка текста
- Временные ряды
- Логические правила
- Машина опорных векторов



Поддерживающие компании



10. КОМПОНЕНТНАЯ АРХИТЕКТУРА АНАЛИТИЧЕСКОГО СЕРВЕРА



Масштабируемость обеспечивается компонентной архитектурой программных средств

Состав компонентов поставки определяется задачами пользователя

11. СТРУКТУРА РЕЕСТРА СЕРВИСОВ АНАЛИТИЧЕСКОЙ ПЛАТФОРМЫ

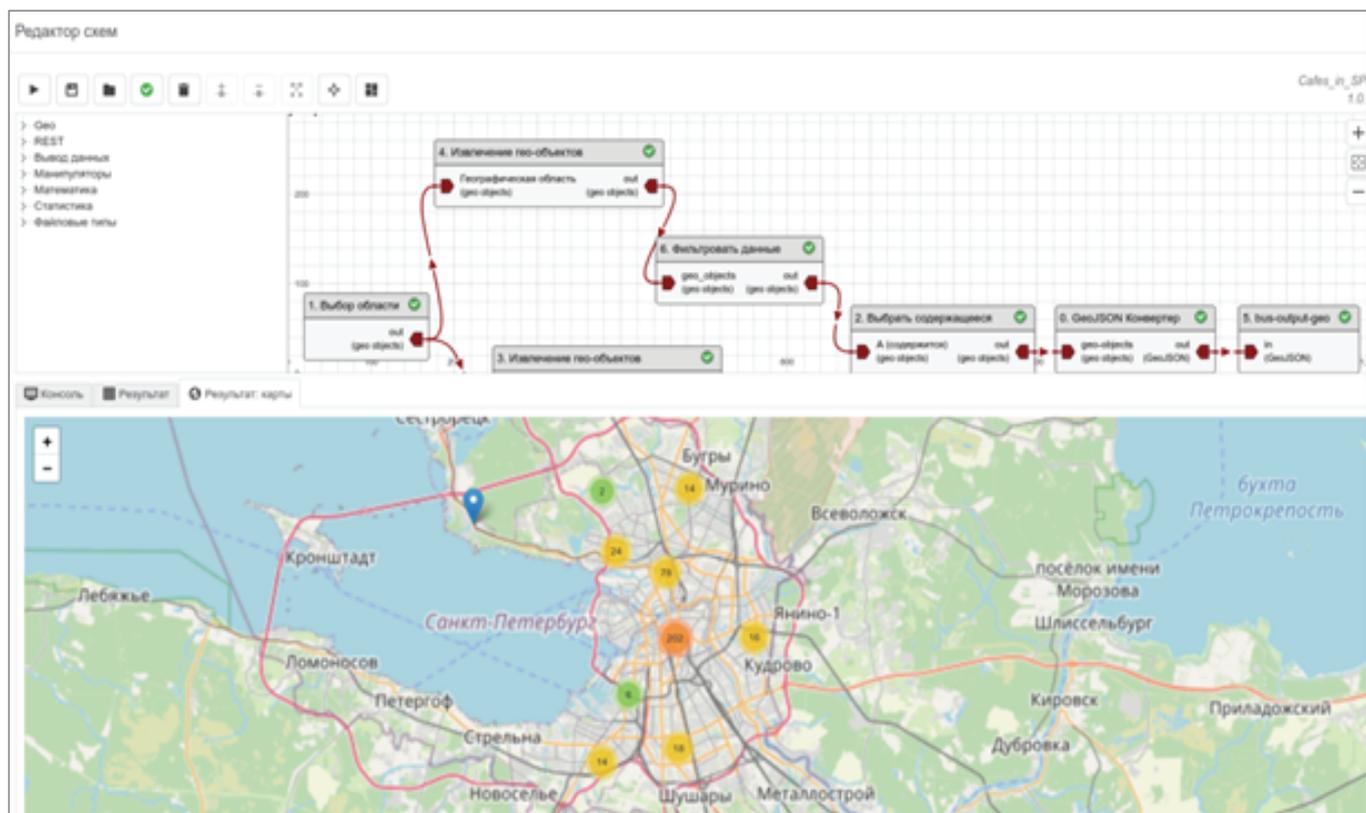
- SR_getServices – получить информацию о сервисах.
- SR_registrateService – регистрация нового сервиса.
- SR_unregistrateService – удаления сервиса из системы.
- SR_getTypes – получить типы сервисов.
- SR_getServiceInfo – получить информацию о сервисе.
- SR_getServiceExeptions – получить описание ошибок (исключений).
- SR_getServiceTypes – получить классы сервиса.
- SR_getServiceTypeInterfaces – получить интерфейсы для класса сервиса.
- SR_getServicesInterfaceParams – получить описание параметров метода.
- SR_getServiceInterfaceReturn – получить описание возвращаемого значения.



12. ГЕОПРОЦЕССИНГ

Основные возможности:

- Автоматическая классификация документов по тематическим, географическим, фактографическим и иным направлениям отбора, включая материалы на иностранных языках;
- Полнотекстовый поиск по загруженным документам;
- Ведение документально-фактографических банков данных с автоматизированными процедурами регистрации и выгрузки данных для различных систем учета;
- Сводный статистический анализ. Подготовка сводных выходных документов.

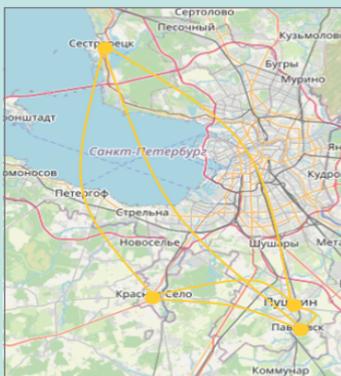


13. ОБРАБОТКА ТЕКСТОВЫХ ДАННЫХ

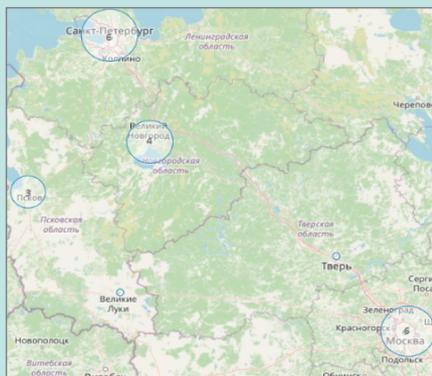
Основные возможности:

- Ускоренный ввод и индексирование документов на основе технологии штрих-кодирования
- Автоматическая классификация документов по тематическим, географическим, фактографическим и иным направлениям отбора, включая материалы на иностранных языках
- Полнотекстовый поиск по загруженным документам
- Ведение документально-фактографических банков данных с автоматизированными процедурами регистрации и выгрузки данных для различных систем учета
- Сводный статистический анализ. Подготовка сводных выходных документов

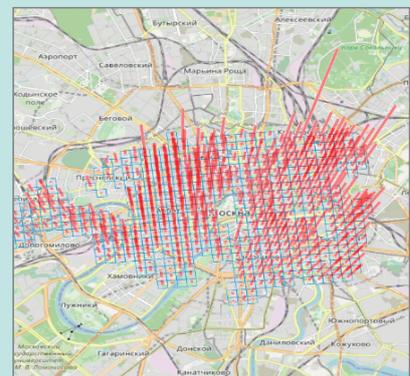
Построение диаграмм событий



Построение связей



Кластеризация



2.5D

Ведение справочников анализируемых сущностей (имена, географические объекты, транспортные средства)

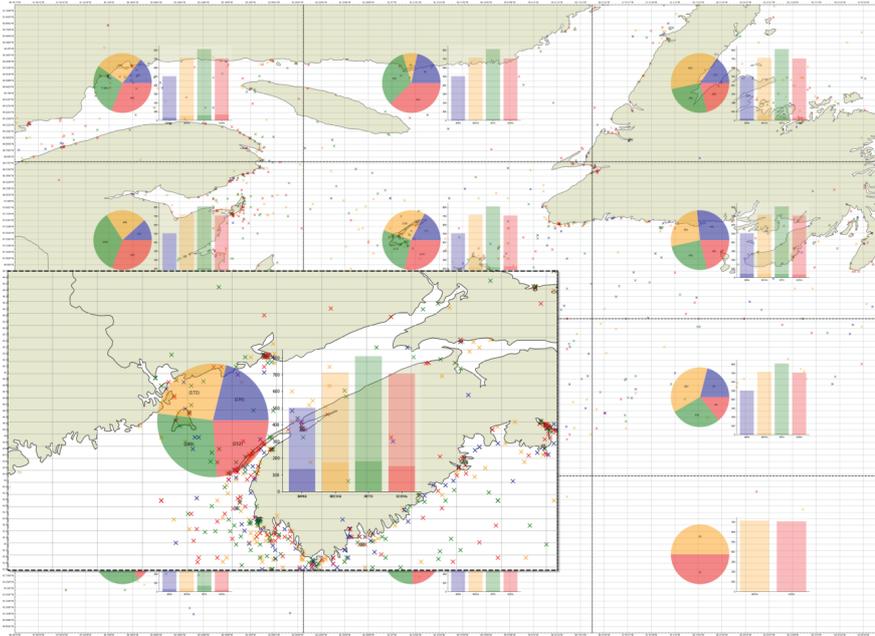
Ведение фактографических банков данных

Сущность	Список статей
Уфа	В Уфе опрокинулась лодка с семьей людей
Тип: Локация	Пропажу в Уфе 17-летнюю девушку нашли мертвой
Название: Уфа	Испорченный праздник
Тип геоточки: Точка	В Уфе ликвидировали открытое горение тарелки кровати
Координаты: [54.7430600, 55.9577900]	
Дополнительная информация: https://ru.wikipedia.org/wiki/%D0%A3%D1%84%D0%80	
Сохранить	Отмена
	Добавить
	Очистить

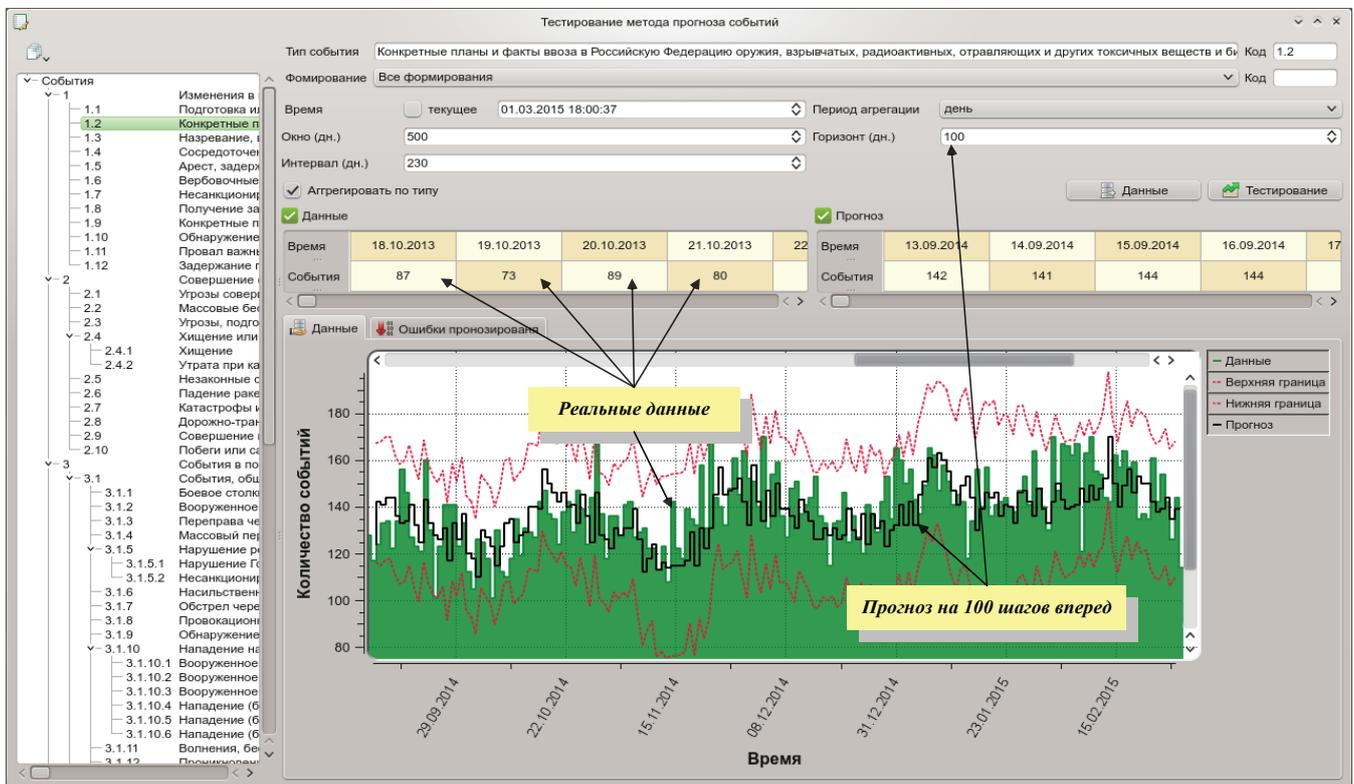
Сущность	Список статей
Уфа	В Уфе опрокинулась лодка с семьей людей
Тип: Локация	Пропажу в Уфе 17-летнюю девушку нашли мертвой
Название: Уфа	Испорченный праздник
Тип геоточки: Точка	В Уфе ликвидировали открытое горение тарелки кровати административного здания
Координаты: [54.74306, 55.95779]	
Дополнительная информация: https://ru.wikipedia.org/wiki/%D0%A3%D1%84%D0%80	
Сохранить	Отмена
	Редактировать

Статья	Атрибуты
В Уфе опрокинулась лодка с семьей людьми	Темы
ТАСС. 9 мая. Лодка, в которой было семь человек, включая троих детей, перевернулась в Уфе, на реке Белая. После оказания медицинской помощи один из находившихся в лодке был госпитализирован, сообщает в воскресенье на сайте управление гражданской защиты Уфы. В 20:05 (18:00 мск), в ИДД г. Уфы поступила информация, что на реке (Белой), в районе Северного автовокзала, перевернулась лодка с людьми. Как выяснилось, трое мужчин и женщина решили покататься на лодке и взяли с собой на водную прогулку троих детей. В какой-то момент произошло опрокидывание плавсредства, и все семеро оказались в воде, - говорится в сообщении. О случившемся в экстренную службу сообщили прожители, отдыхающие на берегу. Спасатели оперативно охватили площадь происшествия, вытаскивая из воды и на берег и парализовали медики. Также в спасении принимали участие очевидцы, подбрасывая в этот момент на лодку. После оказания первой медицинской помощи одного из взрослых, участвовавшего в водной прогулке госпитализировали, остальных отпустили. Обстоятельства происшествия не выяснены.	
	Локации
	Уфа
	река Белая
	Организации
	ИДД г. Уфы
	9.05.2017
Сохранить	Отмена
	Очистить все атрибуты

14. ВИЗУАЛИЗАЦИЯ ДАННЫХ



В составе базовых технологий реализованы методы визуализации деловой и научной графики, а также средства визуализации результатов расчетов на цифровых электронных картах.





АО «Росэлектроника»
121357, г. Москва,
Верейская ул., д. 29, стр. 141
Тел.: +7 (495) 777-42-82
E-mail: info@ruselectronics.ru



ПАО «Интелтех»
197342, Россия, Санкт-Петербург
ул. Кантемировская, д. 8
Тел.: +7 (812) 295-50-69
Email: intelteh@inteltech.ru
Сайт: www.inteltech.ru