GEO-TEXT MINING

АВТОМАТИЧЕСКОЕ ПРЕОБРАЗОВАНИЕ ТЕКСТОВОЙ ЧАСТИ ОТЧЕТОВ ПО ГЕОЛОГОРАЗВЕДКЕ В БАЗЫ ДАННЫХ - ДЛЯ ИХ ИНТЕГРАЦИИ, АНАЛИЗА И СОПОСТАВЛЕНИЯ

ООО «Институт оценки информации», г. Красноярск

Полуян П.В. – генеральный директор,
к.ф.н., доцент СФУ

ieidp.ru

Цифровая трансформация ресурсной отрасли столкнулась с проблемой обработки текстов

- Компьютер не воспринимает тексты на естественном языке
- Быстрый системный поиск по текстам невозможен

К сожалению, мы не можем обработать ваш запрос. Пожалуйста, повторите попытку позже.

Большая часть архивов – это неструктурированные документы

- Отчеты о ГРР, геологоразведочная информация
- Научная аналитика и техническая документация
- Бизнес-переписка, приказы, правила, инструкции

Неструктурированные тексты составляют значительную часть архивов недропользователей

Пример

30 000

единиц хранения геологической информации в архиве агентства NN

до 10 томов

в одном геологоразведочном отчете, объем каждого тома – несколько сот страниц

Тексты оцифрованы, их можно находить и читать с экрана, но обрабатываются они вручную.

Дилемма

- Либо мы тратим на поиск и обработку текста **до 100 человеко-часов на один отчет ГРР**
- Либо отказываемся от анализа и сопоставления

Отчетность по результатам деятельности представляет собой огромный массив данных, уникальный для анализа и сопоставлений.

ПРЕДЛАГАЕМОЕ РЕШЕНИЕ:

АВТОМАТИЧЕСКОЕ ПРЕОБРАЗОВАНИЕ ТЕКСТОВОЙ ДОКУМЕНТАЦИИ (в т.ч. текстовой части отчетов по геологоразведке) В БАЗЫ ДАННЫХ - ДЛЯ ИХ ИНТЕГРАЦИИ, АНАЛИЗА И СОПОСТАВЛЕНИЙ.

Создание системы ИНТЕГРИРОВАННЫХ БАЗ ДАННЫХ позволит вести ускоренный поиск по множеству запросов – для выявления аналогий, закономерностей, анализу изученности и оценке результатов работ в разных условиях с учетом различных параметров.

TEXT MINIG – Интеллектуальный анализ текста - находит применение в науке, бизнесе и спецаналитике.

Инновационный алгоритм TEXT2DB
превращает тексты в базы данных

Предлагаем алгоритм TEXT2DB, который позволит проводить автоматическое структурирование текстовых архивов на основе универсальной онтологии с использованием методов Искусственного Интеллекта.

Эффекты

Дескриптивная аналитика

ускоряется по крайней мере в 10 раз и расширяет сферу охвата в 100 раз

Выгоды

Предиктивная аналитика

за счет исторического охвата становится эффективнее на порядок. Затраты в этой области снижаются в 10 раз.

Новые возможности

Интеллектуальная обработка текстов

обеспечивает поддержку принятия решений. Например: антиплагиат в области ГРР.

Базы данных пригодны для автоматической обработки

Открывается область универсального поиска по всем текстам архива - TEXT MINING

- Кто, что, где, когда?
- Какие компании?
- Какие методики?
- Какое оборудование?
- Какие условия?
- Какие риски?
- Где позитив?
- В чем негатив?

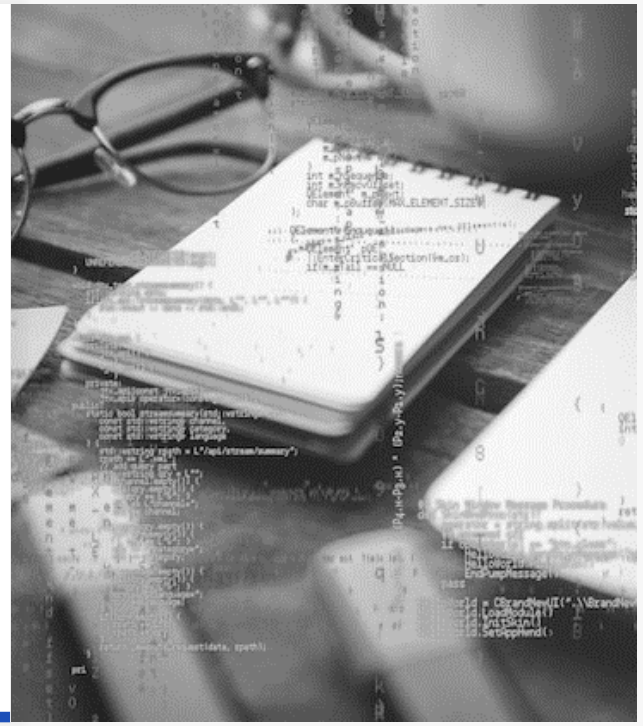
Если текстовые архивы превращаются в реляционную базу данных, - анализ и сопоставление ускоряются в тысячи раз.

Алгоритм TEXT2DB встраивается
в действующие информационные системы

Заказчик получает инструмент для глубинной обработки текстовой информации (TEXT MINING)

Алгоритм дополняет имеющиеся у компании информационные системы.

Сначала мы создаем ПО, нацеленное на интеллектуальную обработку текстов по сейсморазведке и другим видам ГРР, а затем мультиплицируем это решение по всем видам архивной текстовой документации - как «интеллектуальный» помощник для специалистов компании.



Интеллектуальная собственность ООО «ИОИ» защищена и имеет коммерческую оценку

3 млн ₽
Стоимость оценки ИС

«Geosearch (Геологоразведка)» TopViewUser, Crowns Generator Admin"

(свидетельство ФИПС №2020616937 – от 25 июня 2020 г.)

На цифровой платформе представлен **«Geosearch (Геологоразведка) Web-реализация»**
(свидетельство ФИПС №2021610071 от 12 января 2021 г.).

АНАЛОГИ TEXT2DB за рубежом

	ЦП «Институт оценки информации»	Zoho Creator от компании Wrike (Сан-Хосе, Калифорния) https://www.zoho.com/analytics/help/zoho-creator	Amnesthesia — проект на платформе сообщества разработчиков https://github.com/Amnesthesia/txt2db
Алгоритм TEXT2DB для автоматической генерации баз данных на основе текстов.	+	- проектирование баз данных	- заявлен только проект
Используются семантические технологии в комплексе с методами AI	+	- не работает с текстами	+ семантический конвертор
Есть функционал поиска закономерностей	+	+ создает отчеты, но только в виде графиков и формул	- аналитика данных не будет поддерживаться

Суть преимущества перед конкурентами в РФ – гибкое проектирование и адаптация системы под задачи Заказчика

Наши особенности и преимущества

- **Проектирование под задачи Заказчика**
- **Совместимость с архитектурой информационных систем Заказчика**
- **Разработка контролируется Заказчиком, что обеспечивает информационную безопасность**
- **Апробированное на текстах ГРР решение мультиплицируется для интеллектуальной обработки других текстовых архивов**

Прямые конкуренты

Система «АрхиГраф»
разработчик ООО «ТриниДата»

Лицензия на использование системы приобретена в 2021 г. **«Газпром Недра» за 1.8 млн. руб.**

ПЛАН НИОКР ПИЛОТНОГО ПРОЕКТА (6 МЕСЯЦЕВ)

1-2 МЕСЯЦА ЗАПРОС К ПРЕДПРИЯТИЮ-ЗАКАЗЧИКУ:

- Очная встреча-знакомство, определения плана на все 3 этапа - задачи, которые будет решать проект в обработке текстов отчетов ГРР. Заполнение опросника по темам, в продолжение - заочные контакты с 1-2 специалистами (куратор проекта и специалист по сейсморазведке).

3-4 МЕСЯЦА ЗАПРОС К ПРЕДПРИЯТИЮ-ЗАКАЗЧИКУ:

- Заполнение поисковых модулей по каждому из направлений поиска (заочные контакты со специалистами разных подразделений) - с учетом пожеланий БЗ в отношении поисковых запросов.

5-6 МЕСЯЦА ЗАПРОС К ПРЕДПРИЯТИЮ-ЗАКАЗЧИКУ:

- Заочный опрос 5 специалистов разных подразделений по теме интерфейса ПО, участие в тестировании демоверсии ПО специалистов (1-2) от Заказчика
- Подведение итогов и определение плана дальнейшего пилотирования.

Этапы НИОКР определяются на основе утверждения Технического Задания, с оплатой накладных расходов - командировочные, зарплата семи сотрудников команды и оплата консультаций специалистов.

Выполнение конкретных метрик – показатель успешности проекта апробации

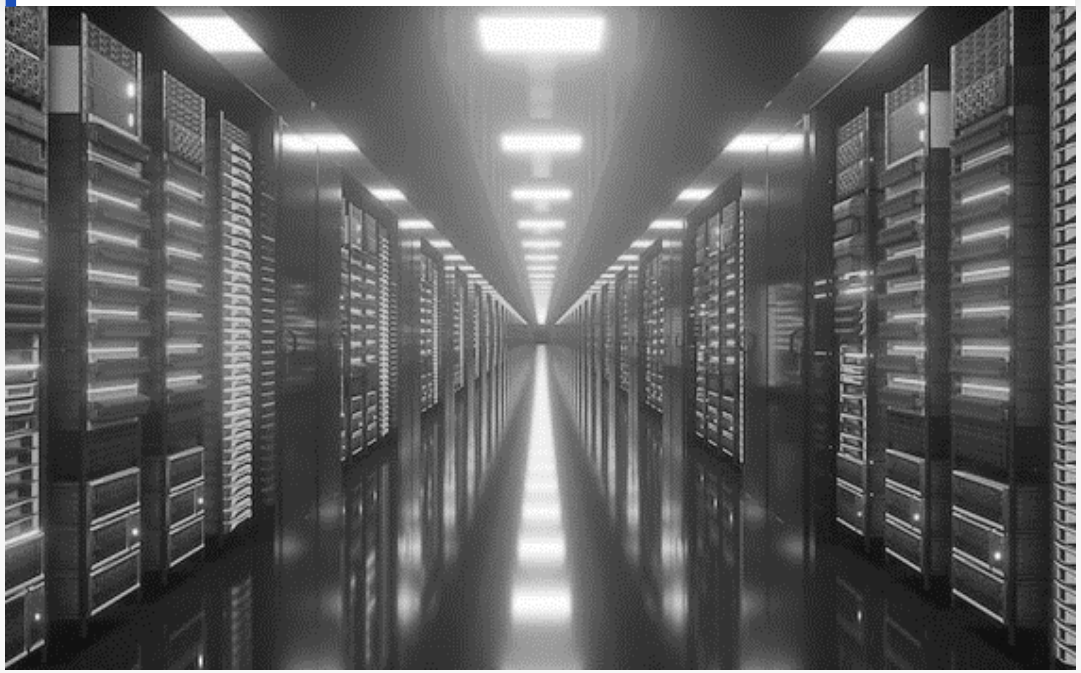
Месяцы	1	2	3	4	5	6
Этап 1	Конфигурация НИОКР, и ТЗ. МЕТРИКИ: не менее 10 проблем-ных опросов, обзор задач.					
Этап 2	Реализация проекта. МЕТРИКИ: классификационная таблица на 1000 терминов, 5 словарей, интеграция 10 текстовых отчетов по сейсморазведке , не менее 100 форм запросов, не менее 3 подходов к алгоритмам поиска.					
Этап 3					Тестирование демоверсии ПО.	

Апробированное решение мультиплицируется для обработки других текстовых архивов

Терабайты текстов получат коммерческую оценку

Базы данных

Это объекты интеллектуальной собственности (согласно ГК РФ ст. 1225).



Что получит Заказчик?

В качестве дополнительного бонуса проводится коммерческая оценка баз данных – создаются нематериальные активы (НМА) на балансе предприятия.

Команда проекта

В офисе ООО «Институт оценки информации»

Полуян Павел Вадимович
учредитель и генеральный директор ООО «ИОИ», кандидат философских наук, СФУ. Отвечает за общее руководство – с первого этапа проекта

Личаргин Дмитрий Викторович
научный руководитель проекта, кандидат технических наук, СФУ

Шикунов Сергей Анатольевич
руководитель AI-направления, кандидат физ-мат. наук

Корболеев Андрей Иванович
программист-разработчик

Бачурина Елена Петровна
концепт-менеджер, СФУ

Поротова Ольга Валерьевна
финансовый директор, кандидат экономических наук, СФУ.

Колодзинский Сергей Геннадьевич
сертифицированный оценщик («МСО» №0387 от 05.02.2008 г.)

Опыт предприятия – гарантия результата

Предприятие ООО «Институт оценки информации» создано в 2019 г. при поддержке Фонда содействия инновациям. Костяк стартапа – специалисты, вышедшие из геологоразведки.

2020 г.

Введена в действие Цифровая платформа открытого доступа – для оказания услуг по интеллектуальной обработке текстов (TEXT MINING), а также для их коммерческой оценки в качестве НМА. Первая продажа услуги

2021 г.

Участвовали в Российском венчурном форуме в Казани. Принимали участие в Московском акселераторе (ENGINEERINGTECH) в московских конференциях InfoSpace-21 и DD StartHub.Moscow.

2022-23 г.

В Тюмени участвовали в конкурсе INNOWEEK-02 — победили на треке FINTECH. Ведем НИОКР за счет частной инвестиции, полученной на INNOWEEK-22 в Тюмени. Успешный кейс – интеллектуальная обработка текстов СМИ «Красноярская газета» .

Письма поддержки от предприятий отрасли

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РФ
Федеральное государственное автономное образовательное учреждение высшего образования
«СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

ИНиГ
Институт нефти и газа
Сибирский федеральный университет

660041 Россия, Красноярск, проспект Свободный, 82 стр. 6
телефон/факс (391) 206-28-61
http://inig.sfu-kras.ru e-mail: inig@sfu-kras.ru

07.07.2021 № 189

на № _____ от _____

В Фонд содействия развитию малых форм предприятий в научно-технической сфере
(Фонд содействия инновациям)

Институт нефти и газа СФУ поддержи
стартапом ООО "Институт оценки инфор
и капитализацией геологоразведочной инфо
у предприятий ресурсной отрасли и может б
актив.

Стартап был организован препод
Полуяном П.В. и Личаргиным Д.В., получи
инновациям и успешно провел НИОКР по
платформы, которая начала работать в 202
Интеллектуальная собственность, созданна
государственной регистрацией, получила ко
2 800 000 рублей и поставлена на баланс пред

ООО "Институт оценки информаци
информации, которые актуальны не только д
других областей экономики. Институт
продолжение НИОКР в данном направлени
Фонда содействия инновациям и
в сотрудничестве.

директор

ИНСТИТУТ НЕФТИ И ГАЗА

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРА

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗО
«УФИМСКИЙ ГОСУДАРСТВЕННЫЙ НЕФТЯНОЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ» (ФГБОУ ВО "УГНТУ")

ул. Космонавтов, 1, г. Уфа, Республика Башкортостан, 450064. Тел.: (347) 242-03-70, факс: (347) 243-14-19. http://www.rusoil.net, E-mail info@ru
77006179, ОГРН 1020203079016, ОКПО 02069450, КПП 027701001

Директору ООО «Институт оценк
информации»
Корболееву А.И.

АБИСС

Общество с ограниченной ответственностью «АБИСС

Исх. № 45 от 30.09.2022 г.

ООО «Институт оценки информации»
Руководителю проекта Полуяну П. В.

Уважаемый Павел Вадимович!

Наша компания занимается оптимизацией процессов бурения, для чего используем биостратиграфические методы изучения образцов с целью выявления дополнительной информации о происхождении и составе горных пород. Мы ознакомились с разработками ООО «Институт оценки информации», касающимися обработки геофизической и геологической информации с целью перевода данных в пригодные для машинной обработки и систематизации. Считаем данное направление работы интересным, готовы принять участие в пилотных разработках Вашей компании. В случае успеха Ваших разработок готовы рассмотреть вариант коммерческого сотрудничества с ООО Институт оценки информации»

Генеральный директор **А.Ф. Галямов**

(подпись)

М.П.

Общество с ограниченной ответственностью
«СОРБ»
50008, Республика Башкортостан, г. Уфа, ул. Кирова, 15, оф. 38А
ИНН 0274937317, ОГРН 1180280028500

2

ООО «Институт оценки информации»

компания ознакомилось с разработками ООО «Институт оценки
касающимися автоматической обработки реестровых данных с
м баз данных, подлежащих оценке и регистрации. Просим предоставить
ую информацию и коммерческое предложение.

директор

А. Д. Бадикова

«СОРБ»

ГАЗПРОМ НЕФТЬ

ОБЩЕСТВО С ОГРАНИЧЕННОЙ ОТВЕТСТВЕННОСТЬЮ
«ГАЗПРОМНЕФТЬ-ТЕХНОЛОГИЧЕСКИЕ ПАРТНЕРСТВА»
(ООО «ГАЗПРОМНЕФТЬ-ТЕХНОЛОГИЧЕСКИЕ ПАРТНЕРСТВА»)

Адрес для корреспонденции:
ул. Якубовича, д.24, литер А, Санкт-Петербург, 190121
ОКПО 31590343, ОГРН 1188617010064
ИНН 8601066868, КПП 860101001
Тел.: +7 812 449-20-83
e-mail: gto_apb@gazprom-neft.ru
www.gazprom-neft.ru

25.05.2023 № ГТП-15/000231

на № _____ от _____

Генеральному директору
ООО «ИОИ»

Полуян П.В.

О подтверждении
заинтересованности в проекте

Уважаемый Павел Вадимович!

Настоящим письмом ООО «Газпромнефть – Технологические партнерства» подтверждает потенциальную заинтересованность в разработках ООО "Институт оценки информации", связанных с интеллектуальной обработкой, оценкой и капитализацией информации, которая накапливается у предприятий ресурсной отрасли, может быть обработана средствами text mining и превращена в финансовый актив.

Решение ООО "Институт оценки информации" предложено к внедрению в периметре группы компаний «Газпром нефть» в рамках акселерационной программы INDUSTRIX, проводимой ООО «Газпромнефть – Технологические партнерства».

В случае успешной реализации проекта в рамках программы INDUSTRIX, ООО «Газпромнефть – Технологические партнерства» подтверждает возможность потенциального сотрудничества и проведения опытно-промышленных испытаний решений ООО «Институт оценки информации».

С уважением,

Директор по развитию открытых инноваций и новых бизнесов **М.Е. Бардин**

ООО «ГАЗПРОМНЕФТЬ-ТЕХНОЛОГИЧЕСКИЕ ПАРТНЕРСТВА»

Мы открыты для сотрудничества

+7 913 839 49 14

DP-IEI@yandex.ru
polyan2002@mail.ru

Poluyan_p

ieidp.ru