

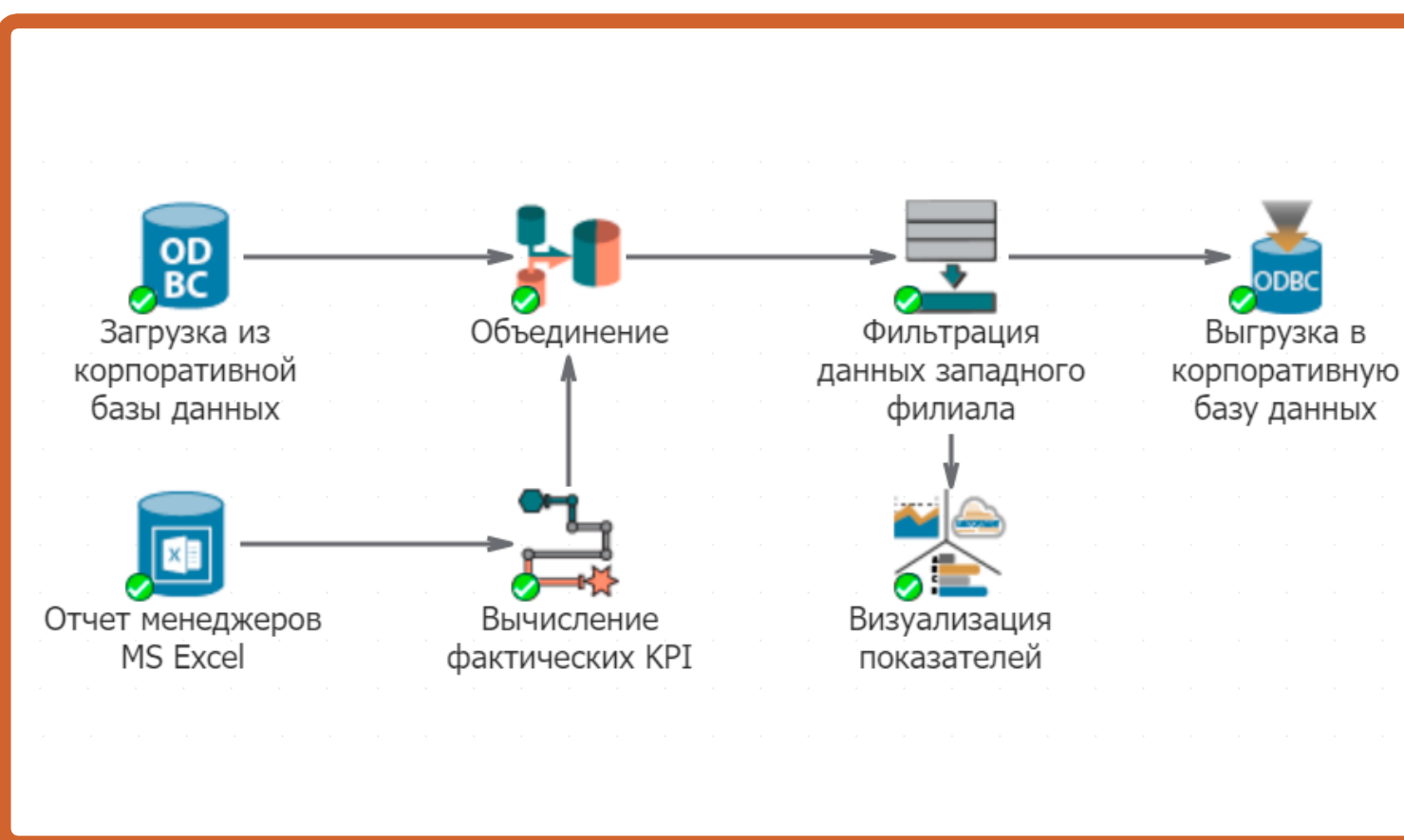


**Примеры
решений для анализа и
автоматизации бизнес-задач
PolyAnalyst 6.5**

ЧТО ТАКОЕ POLYANALYST?

PolyAnalyst- аналитическая платформа и открытая среда разработки.

Это no-code система, где с помощью механизма визуального проектирования аналитического скрипта пользователи, не обладающим специальными математическими знаниями и навыками программирования, самостоятельно создают законченные комплексные **решения для анализа данных и текстов, а также оптимизации и автоматизации бизнес-процессов.**



ПЕРЕДОВАЯ АНАЛИТИКА «ИЗ КОРОБКИ»

В платформу встроены no-code инструменты обработки данных, более 30 готовых алгоритмов машинного обучения и не имеющие конкурентных аналогов инструменты текстового анализа.

АНАЛИТИЧЕСКОЕ ЯДРО

Платформа встраивается в информационную систему и может без контроля со стороны пользователя самостоятельно обрабатывать данные и выдавать результат.

ПОСТРОЕНИЕ ДАШБОРДОВ

PolyAnalyst включает собственные средства для визуализации результатов и построения интерактивных графических веб-отчётов.

Платформа PolyAnalyst включена в реестр Российского ПО. Свидетельство №4414.

Является прямым конкурентом и вариантом импортозамещения для IBM Watson и SAS Viya.

РЕШЕНИЕ БИЗНЕС-ЗАДАЧ НА ОСНОВЕ ТЕХНОЛОГИЙ NLP И ML

Широкий набор бизнес-задач в компаниях различных отраслей решается с применением алгоритмов работы с естественным языком (NLP - Natural Language Processing) и машинного обучения (ML).

Технологии NLP и ML применяются для анализа информации, содержащейся в различных типах документов, возникающих при реализации производственных, управленческих и других функций компаний.

КОРПОРАТИВНАЯ ДОКУМЕНТАЦИЯ И ВНЕШНИЕ ИСТОЧНИКИ

- Договоры
- Счета, накладные
- Финансовые документы
- Жалобы, отзывы
- Внутренние НПА
- Расшифровки телефонных переговоров
- Справочники НСИ
- Технологическая документация, архивы
- Диспетчерские журналы
- Переписка в ЭДО и по электронной почте
- СМИ, Интернет
- Социальные сети
- Другие источники

ФУНКЦИИ

- Юридическая
- Финансовая
- Кадровая
- Продажа услуг и обслуживание клиентов
- Обеспечение ИТ инфраструктурой
- Строительство
- Логистика/организация перевозок
- Разработка и исследования (R&D)
- Эксплуатация инфраструктуры
- Многое другое

ЦИФРОВЫЕ РЕШЕНИЯ МЕГАПЬЮТЕР

Компания Мегапьютер Интеллидженс предлагает к реализации и внедрению решения для **ШИРОКОГО НАБОРА БИЗНЕС-ЗАДАЧ И ЦИФРОВИЗАЦИИ ПРОЦЕССОВ В РАЗЛИЧНЫХ ПОДРАЗДЕЛЕНИЯХ КОМПАНИЙ.**

• Управление персоналом

- I. Разработка карты профессиональных компетенций кандидатов и сотрудников для целей оценки и повышения общего уровня квалификации персонала.
- II. Сокращение доли ручного труда сотрудников путем автоматизации типовых операций (извлечение данных из копий предоставленных документов при оформлении новых работников, перевод документов из бумажного в электронный вид, анализ текстовых баз данных).

• Бухгалтерия/корпоративные финансы

- I. Сокращение ручного труда путем автоматизации типовых операций с текстовыми документами (сверки актов, счетов и накладных).
- II. Повышение эффективности управления товарными запасами путем автоматизированного формирования прогноза спроса в интеллектуальной системе предиктивной аналитики.
- III. Сокращение рисков финансовых операций за счет контроля делового состояния контрагентов через мониторинг как информации от регулирующих органов, так и новостного потока.

Контакт-центр

- I. Автоматизированная классификация и диспетчеризация входящих обращений клиентов по всем каналам коммуникаций (звонки, эл.почта, соц. сети, сайт и т.д.).
- II. Выявление нарождающихся трендов новых тем обращений клиентов.
- III. Интеллектуальный анализ обращений клиентов.

• Корпоративные коммуникации

- I. Контроль и повышение общей удовлетворенности услугами через анализ клиентских отзывов.
- II. Анализ репутационного профиля в медийных средах.
- III. Автоматизация работы с претензиями.

• Правовая функция

- I. Решения Legaltech - повышение эффективности деятельности сотрудников через автоматизацию рутинных операций юридической экспертизы.
- II. Выявление правовых рисков в документации (DueDilligence).

• Капитальное строительство

• Управление корпоративным имуществом

- I. Цифровизация бумажных архивов технологической и нормативной документации, разработка системы индексации документов и системы интеллектуального поиска, включая поиск по атрибутам и другим параметрам.
- II. Повышение скорости выдачи и пользовательской доступности технологической документации с помощью внедрения системы интеллектуального поиска по базе технологической документации и справочникам.
- III. Унификация и автоматизированное управление номенклатурой фонда нормативно-справочной информации.

ПРИМЕРЫ РЕШЕНИЙ НА БАЗЕ ПЛАТФОРМЫ POLYANALYST

PolyAnalyst 6.5

ПЕРЕЧЕНЬ РЕШЕНИЙ

1. Автоматизация присвоения кодов товарам и запасам	стр.7
2. Контроль делового состояния контрагентов через мониторинг новостного потока	стр.8
3. Структурирование и поиск связей в документации	стр.9
4. Автоматизированная классификация и маршрутизация входящих обращений	стр.10
5. Долгосрочный прогноз продаж в ритейл сети посредством data mining	стр.11
6. Анализ интернет-отзывов о продукции Canon	стр.12
7. Извлечение тренда новых тем обращений клиентов	стр.13
8. Агрегирование и анализ исторических и актуальных данных ТОиР в производстве	стр.14
9. Контроль медицинских ресурсов в условиях пандемии инфекции COVID-19	стр.15
10. Автоматическая сверка актов	стр.16
11. Автоматизация анализа дефектов продукции в процессе гарантийного обслуживания	стр.17
12. Управление номенклатурой материально-технических ресурсов	стр.18
13. Роботизированная юридическая экспертиза	стр.19
14. Контроль уровня обслуживания клиентов в онлайн-чатах	стр.20
15. Мониторинг продуктивности бизнеса через анализ отзывов клиентов	стр.21
16. Автоматизация извлечения информации о контрагентах	стр.22
17. Борьба с мошенничеством контрагентов через анализ данных гарантийного обслуживания	стр.23
18. Классификация сообщений фонда целевых капиталов	стр.24
19. Оценка медийного образа	стр.25
20. Оценка мнений и настроений в регионах РФ	стр.26
19. Оцифровка корпоративного архива и система интеллектуального поиска по базе документов	стр.27
20. Анализ цифрового следа обучающихся	стр.28
21. Создание цифровой модели компетенций учёного	стр.29
22. Создание профессионального стандарта	стр.30
23. Интеллектуальный анализ обращений граждан в приемную губернатора	стр.31

АВТОМАТИЗАЦИЯ ПРИСВОЕНИЯ **КОДОВ** ТОВАРАМ И ЗАПАСАМ

ЗАКАЗЧИК: Российская логистическая компания

ПРОБЛЕМА

Логистическая компания, оказывает таможенные услуги, в том числе услуги по классификации товаров по ТН ВЭД. Это десятизначный код-классификатор, который используется для проведения таможенных операций и присваивается товару исходя из его характеристик. Так, например, в коде 62 02 12 000 первые две цифры 62, называемые товарной группой, классифицируют товар как предмет одежды из текстиля, четыре цифры 62 02, что это пальто или накидки, шесть цифр 62 02 12, что сделаны они из хлопка. Сложность классификации в том, что кодов много, одних только товарных групп 96 это первые две цифры. Коды присваиваются вручную исключительно за счет навыка специалиста, который этим занимается и на основе описания товара импортером, то есть человек читает описание и решает под какой код оно подходит. От присвоенного кода зависит ставка ввозной таможенной пошлины, поэтому правильность кода строго проверяется таможенным инспектором. За неправильно указанную информацию, из-за ошибки человека, организации грозит штраф.

ЗАДАЧА

Разработка системы автоматического присвоения кодов товарам.

РЕШЕНИЕ

Мегапьютер получил от заказчика датасет, который включал неструктурированное описание товаров и проставленный заранее код. Имелся некий код и идущее подряд текстовое описание, как с этикетки товара. С помощью лингвистических алгоритмов, это неструктурированное описание разделили на отдельные атрибуты, такие как наименование товара, размер, пол, содержание материалов в составе, тип состава. Далее к этому преобразованному датасету в PolyAnalyst были применены три аналитические методики: чистое машинное обучение, ruled-based подход, то есть классификация на написанных человеком правилах и гибридный подход, который объединяет два предыдущих.

РЕЗУЛЬТАТ

Решение в PolyAnalyst позволяет автоматизировать процедуру присвоения кодов, позволяя частично или полностью избавиться от ручного труда.

Система самостоятельно загружает в себя набор данных о товарах, применяя к ним алгоритмы обработки естественного языка NLP расшифровывает и разбивает на отдельные атрибуты и применяя к уже структурированным данным обученную модель машинного обучения или модель на основе логических правил автоматически без человека кодирует вновь поступающие товары.

- В силу рутинности данных операций кодирование товаров человеком неизбежно приводит к возникновению ошибок и неточностей, что почти не характерно для автоматизированного процесса.
- Система адаптивна к росту количества товаров и исключает необходимость привлечения дополнительного персонала.

КОНТРОЛЬ ДЕЛОВОГО СОСТОЯНИЯ КОНТРАГЕНТОВ ЧЕРЕЗ МОНИТОРИНГ НОВОСТНОГО ПОТОКА

ЗАКАЗЧИК: Крупная банковская организация.

ПРОБЛЕМА

С целью снижения рисков и предотвращения финансовых потерь, компании постоянно оценивают финансовое состояние контрагента. Эффективным методом оценки является мониторинг ковенантных (косвенно демонстрирующих показатели финансовой устойчивости) событий, отмеченных в новостных сводках. Метод требует кропотливого «вычитывания» значительного объема информации и больших кадровых ресурсов, что не гарантирует безошибочность их работы. Для ручного чтения всех новостей, полученных за сутки о крупных компаниях – контрагентах, потребуются сотни часов и полная занятость десятков человек. Однако, «чтение» и анализ новостного потока можно полностью алгоритмизировать.

ЗАДАЧА

Автоматизация анализа новостного потока на предмет упоминания контрагентов и выявления потенциальных рисков контрагентов по заданному перечню.

РЕШЕНИЕ

Алгоритмы текстового анализа PolyAnalyst извлекают упоминания о потенциальных объектах ковенантных событий: компаниях, организациях, именах людей и прочем. Затем извлекается информация о значимых событиях, связанных с объектами, полученными на предыдущем этапе: сделки, поглощения, банкротства, смена руководителей и собственников.

Конечный пользователь получает итоговые результаты в удобном формате (панель), позволяющем легко просматривать информацию о типе ковенантного события, его участниках, связях между ними и дате.

Решение может представлять результаты в виде интерактивных графических отчетов.

РЕЗУЛЬТАТ

Автоматическое решение демонстрирует высокое качество анализа. Полнота результатов составляет от 0.84 до 0.93, точность – от 0.82 до 0.98 по разным типам событий.

Готовое решение за 1 час обрабатывает около 20 миллионов слов. Для обработки информации такого объема команде в 20 человек понадобится не менее 18 рабочих дней.

Решение можно настроить на мониторинг любых интересующих событий.

За счет интуитивного, доступного любому пользователю интерфейса, перестройка поиска и анализа с новостных сообщений, например, о продаже нефтяной компанией иностранных активов государству на новости про определенный штамм гриппа, повлиявший на птицеводческие хозяйства, понадобится 30 минут элементарной работы аналитика.

СТРУКТУРИРОВАНИЕ И ПОИСК СВЯЗЕЙ В ДОКУМЕНТАЦИИ

ЗАКАЗЧИК: Российская железнодорожная компания

ПРОБЛЕМА

Крупная железнодорожная компания столкнулась с проблемой накопления массы неструктурированной, разбросанной по различным директориях или хранилищам, разноформатной (pdf, docx, jpeg) документации. В большинстве случаев связи между документами не указаны явно и присутствуют только в самом тексте в виде отсылок к номеру и названию связанного документа. Информация о дате документа, авторе, подписанте также может быть скрыта только в тексте документа. Очевидно, что чтобы вручную структурировать и установить связи в большом объеме документов потребуется их перечитать и на это уйдет колоссальное количество времени.

ЗАДАЧА

Разработать в PolyAnalyst сценарий для автоматического выявления связей и структурирования документов

РЕШЕНИЕ

В PolyAnalyst загрузили весь корпус разрозненной документации, система ее просмотрела и о каждом документе извлекла необходимую информацию: название, тип, номер, дата подписания. Далее полуавтоматически был составлен словарь, куда была загружена эта извлеченная информация. Таким образом была получена база, содержащую информацию о каждом загруженном документе.

Далее PolyAnalyst настроили на поиск в тексте документов отсылок на другие документы, чьи названия система брала из созданного на предыдущем этапе словаря. В итоговой таблице, которая и является решением поставленной задачи, перечисляются документы и найденные в них ссылки на связанные документы.

РЕЗУЛЬТАТ

В компании-заказчике эта система в первую очередь разрабатывалась для подразделения, которое занимается моделированием бизнес-процессов. Каждый моделируемый процесс содержит определенный набор документации, который его регламентирует. Специалистам необходимо контролировать, не указаны ли лишние документы или наоборот не упущено ли важная документация .

PolyAnalyst структурировал документацию и специалисты компании необходимо просто отобрать относящиеся к моделируемому бизнес-процессу документы и сопоставить их с теми, что прописаны в самой модели. В результате наглядно видно, чего не хватает, а что лишнее.

Также полезно структурировать корпоративную научную документацию. С помощью PolyAnalyst возможно проанализировать выборку научных исследований компании в интересующих областях на предмет их результативности и эффективности расходования средств на такие исследования.

АВТОМАТИЗИРОВАННАЯ КЛАССИФИКАЦИЯ И МАРШРУТИЗАЦИЯ ВХОДЯЩИХ ОБРАЩЕНИЙ

ЗАКАЗЧИК: Крупный дистрибутор ПО и оборудования

ПРОБЛЕМА

В силу динамичного развития дистрибуторского направления и масштабного расширения партнерской сети организация столкнулась с проблемой экспоненциального роста объемов входящей корреспонденции. Подавляющее число поступающих документов составляют запросы о резервировании или заказе продукции от компаний-партнеров. Число таких организаций превышает 6000 тысяч, а объем входящих писем составляет порядка нескольких тысяч в сутки. Ручная обработка, включающая процедуру определения сути обращения и отправки его в соответствующее структурное подразделение, крайне трудоемка и приводит к существенному временному лагу. Это в условиях жесткой конкуренции становится причиной роста упущенной выгоды.

ЗАДАЧА

Внедрение системы автоматизированной классификации и маршрутизации входящих документов.

РЕШЕНИЕ

На базе платформы PolyAnalyst было разработано и интегрировано в информационные системы компании специализированное аналитическое решение. Используя встроенный коннектор, система автоматически подключается к корпоративному хранилищу входящей корреспонденции и с периодичностью раз в минуту загружает себе новые письма. Далее к этим письмам применялись алгоритмы текстового анализа PolyAnalyst: исходя из данных в загруженном в платформу справочнике производителей и перечне типов имеющейся продукции, система определяет и извлекает из письма связку категорий «производитель» + «тип продукции». Таким образом письма классифицировались, формировался отчет и информация направлялась в ответственное за данную категорию письма структурное подразделение.

РЕЗУЛЬТАТ

Разработанное решение за счет продуктивности интеллектуальных инструментов текстового анализа PolyAnalyst продемонстрировало высокую (свыше 90%) точность классификации писем.

Интегрированная аналитическая схема позволила освободить сотрудников от рутинных операций и автоматизировать процессы классификации и маршрутизации. Это в свою очередь привело к росту эффективности ряда бизнес-процессов компании:

- В отличие от ручного труда, система практически в режиме реального времени проводит анализ документации, что существенно сократило временной промежуток между получением письма и отправкой его в ответственное подразделение.
- В силу рутинности данных операций обработка документов человеком неизбежно приводит к возникновению ошибок и неточностей, что почти не характерно для автоматизированного процесса.
- Система адаптивна к росту количества запросов и исключает необходимость привлечения дополнительного персонала.

ДОЛГОСРОЧНЫЙ ПРОГНОЗ ПРОДАЖ В РИТЕЙЛ СЕТИ ПОСРЕДСТВОМ DATA MINING

ЗАКАЗЧИК: Мультиформатная продуктовая розничная компания.

ПРОБЛЕМА

Эффективное составление наиболее точного финансового и закупочного планов, общего бюджета, бюджета продвижения и актуальной системы мотивации персонала ритейл сети обеспечивается при обязательном наличии долгосрочного прогноза продаж. Динамичное расширение сети и ежемесячное открытие нескольких десятков новых точек способствует накоплению внушительного объема показателей продаж. Также дело осложняется крайне большим разбросом объемов продаж по разным магазинам и отдельным товарным группам. Перед компанией возникает непростая задача разработки и поддержки системы анализа и прогнозирования продаж каждой точки сети по всему многотысячному ассортименту, при соблюдении нормы минимального относительного (в %) отклонения прогноза от факта.

ЗАДАЧА

Построение предиктивной модели на основе предоставленных данных, которая позволит предсказывать объем ежедневных продаж в каждом из магазинов сети на глубину в 1.5 года.

РЕШЕНИЕ

Не смотря на то, что данные, из которых строится прогноз представляют собой временные ряды, сотрудники Мегапьютер Интеллидженс выдвинули предположение, что зависимости продаж определяются из заранее известных компонент вроде годичной и недельной сезонности, роста продаж новых магазинов, аномалий в праздничные дни и т.д. Выяснив параметры этих зависимостей, удалось уйти от задачи анализа временных рядов и свести ее к классической задаче data mining.

Далеко не все факторы, влияющие на продажи, отражены в данных (деятельность конкурентов, макроэкономические факторы, долговременные погодные аномалии, и т.д.) Если прогнозная модель систематически ошибается, делается вывод, что действует неизвестный фактор, и требуется коррекция прогноза, что легко и быстро выполняется за счет наличия в PolyAnalyst удобных средств разработки.

РЕЗУЛЬТАТ

В результате, после внедрения системы прогнозирования на базе платформы PolyAnalyst ритейл сеть получила существенное повышение точности прогноза, а конкретно:

- Возможность прогнозировать объем продаж на каждый день в каждом магазине по каждой товарной группе на полтора года вперед со средней точностью 46%;
- Для старых магазинов эта точность повышается до 39.2%;
- Чем больше объемы продаж (по магазину или по товарной группе), тем точность прогноза выше. Например, точность прогноза по игрушкам (на старых магазинах) составляет 10.19%, а общая точность прогноза по самым крупным магазинам доходит до 27%.

Это закономерно приводит к росту эффективности формирования бюджетов, планов продаж и распределению финансовых, временных и трудовых ресурсов, что отражается на КПД операционных процессов всей компании

АНАЛИЗ ИНТЕРНЕТ-ОТЗЫВОВ О ПРОДУКЦИИ CANON

ЗАКАЗЧИК: Canon, подразделение производства МФУ

ПРОБЛЕМА

Отзывы клиентов отражают их жалобы и предложения по улучшению обслуживания и функционалу продукта. Они несут в себе массу полезной информации в разы большую, чем простая, бальная оценка пользователем, но в отличие от нее трудно анализируются, так как представлены в неструктурированном, текстовом формате. Большое количество текстов и неструктурированность данных отзывов делают невозможным проведение ручного анализа и изыскания полезной информации.

ЗАДАЧА

Внедрение автоматического анализа интернет-отзывов клиентов.

РЕШЕНИЕ

В качестве ресурсов для сбора данных использовались крупнейшие форумы с покупательскими отзывами, крупные интернет-магазины, социальные сети, а также сайт компании и данные других каналов обратной связи от клиентов.

Решение для анализа отзывов на базе платформы PolyAnalyst полностью автоматизированно и в кратчайшие сроки выявляет фактические аспекты и сентиментный анализ отзывов покупателей. Решение предоставляет результаты анализа и рекомендации в виде интерактивных графических отчетов для руководства, упрощающих процесс принятия решений, или в табличном виде для интеграции с внешними системами.

РЕЗУЛЬТАТ

Внедренное автоматизированное решение в режиме реального времени с высокой скоростью обрабатывает весь массив отзывов о продукции компании Canon и представляет результаты в виде наглядных веб-отчетов с набором визуализаций.

Таким образом, компания получила возможность получать полезную информацию, скрытую в трудноанализируемых отзывах:

- Определены основные аспекты, влияющие на общую оценку продукции;
- Определены главные конкурентные преимущества;
- Повышены продажи в результате корректировки отрицательных для покупателя аспектов функционала продукта

Canon получили инструмент комплексного и своевременного контроля качества своей продукции, что позволило вести разработку наиболее функционально полезных для пользователя продуктов и услуг и, как следствие, вызвало рост выручки.

ИЗВЛЕЧЕНИЕ ТРЕНДА **НОВЫХ ТЕМ** ОБРАЩЕНИЙ КЛИЕНТОВ

ЗАКАЗЧИК: Западная банковская компания

ПРОБЛЕМА

Динамично меняющаяся обстановка в мире бросает новые вызовы бизнесу. Крупная банковская компания с возрастающей периодичностью сталкивается с проблемой новых по содержанию и проблематике запросов от клиентов, касающихся к примеру изменений в режиме и условиях обслуживания по банковским и сопутствующим продуктам в обстановке пандемии COVID-19. Компании приходится тратить дополнительные средства и время на подробный анализ таких нестандартных вопросов и подготовку подробного, исчерпывающего и способного снять все опасения клиентов ответа. Раннее обнаружение новых тем в запросах способно предоставить компании серьезный выигрыш во времени для подготовки полного объема необходимых аргументов и заявлений.

ЗАДАЧА

Внедрение полуавтоматической системы выделения тренда новых проблемных тем обращений клиентов.

РЕШЕНИЕ

Система автоматически классифицирует и отбирает письма, относящиеся к вопросу готовности банка к работе в условиях пандемии и карантина. Аналитик, работая со словарем синонимов и стоп-слов настраивает сценарий анализа PolyAnalyst на тематику COVID-19 с сопутствующими смысловыми атрибутами. Выявляемые данные о росте числа новых вопросов и интересующих клиентов тем визуализируются в наглядных интерактивных отчетах. Встроенный планировщик настраивается на периодическое получение и обработку новых тем и данных.

РЕЗУЛЬТАТ

Работая с ежедневными визуальными отчетами PolyAnalyst, сотрудники выявляют тренды развития или угасания новых тем клиентских запросов.

Анализ проводится по каждому отделу компании, ведущему переписку с тем или иным типом клиентов, что позволяет категоризировать тематики запросов для каждого отдельного подразделения.

Сформировав пул актуальных и потенциально-актуальных вопросов, компания подготовила ответы на эти вопросы для заблаговременного массового уведомления клиентов.

Это позволило сохранить должный уровень лояльности клиентов, своевременно, не приводя к возможным финансовым оттокам, снять наибольшие их опасения и сократить время и средства, затрачиваемые на подготовку разъяснительных ответов.

АГРЕГИРОВАНИЕ И АНАЛИЗ ИСТОРИЧЕСКИХ И АКТУАЛЬНЫХ ДАННЫХ **ТОИР** В ПРОИЗВОДСТВЕ

ЗАКАЗЧИК: Промышленное предприятие

ПРОБЛЕМА

Существенным отрицательным фактором работы подразделений технического обслуживания и ремонта (ТОиР) промышленных предприятий является разрозненность ресурсов записи и хранения информации о проблемах в работе и поломках оборудования, а также процессах сервисного обслуживания и ремонта. Такими ресурсами являются ИСУ ТОиР, бухгалтерское ПО, производственные системы, рукописные журналы, файлы Excel, диспетчерские программы. Ручной анализ и сопоставление между собой всей массы информации, представленной в этих источниках является непосильной задачей. При этом, встроенные инструменты анализа ИСУ ТОиР обрабатывают только кодированную информация, пропуская важные комментарии, публикуемые в текстовом поле. Анализ производится без учёта производственной загрузки, договорных отношений с поставщиками запасных частей и подрядчиками.

РЕШЕНИЕ

Первоначально проводится экспресс-аудит источников данных, конструирование и согласование расчетных моделей. После внедрения осуществляются сами процессы обработки и интерпретации данных. Аналитическое решение на базе PolyAnalyst для ТОиР агрегирует информацию из разрозненных источников данных о поломках и ремонтах. Собрав данные, из журнала заявок, плана работ, ИСУ ТОиР, журнала ремонтов, бухгалтерской программы, актов остановки проводится сематический анализ и устанавливаются корреляции между интенсивностью отказов, ремонтами, простоями, поставщиками и подрядчиками. Проводя анализ текстовых полей примечаний ИСУ, система изыскивает дополнительную полезную информацию, упускаемую при анализе лишь данных, описываемых в системе установленными кодами. Решение автоматически сопоставляет информацию из разных источников об одних и тех же процессах, проводя сравнение и определяя уровень достоверности.

РЕЗУЛЬТАТ

Комплексный анализ накопленной и вновь поступающей информации о ТОиР с использованием текстового анализа и машинного обучения способствует существенному сокращению потерь в объеме 5-40%. Собранные данные применяются для расчёта основных KPI по сервису и поиска слабых мест на уровне хороших практик использования ИСУ ТОиР.

За счет анализа поломок и их корреляций с оборудованием, ремонтом, запасными частям и их поставщикам, нагрузкой на оборудование и его простоями возможно определение и устранение общих причин поломок и проблем ремонтов. Результатом становится рост ОЕЕ на 0,5-2% и оптимизация стоимости обслуживания на единицу продукции.

По итогам сравнительного анализа однородных записей об истории и текущем состоянии оборудования из различных источников, формируется полная и наиболее достоверная информация, используемая при взаимодействии с сервисными организациями: составление оптимального SLA, обоснованной суммы контракта и дальнейшего контроля соблюдения утвержденного уровня обслуживания.

КОНТРОЛЬ МЕДИЦИНСКИХ РЕСУРСОВ В УСЛОВИЯХ ПАНДЕМИИ ИНФЕКЦИИ COVID-19

ЗАКАЗЧИК: Системы государственного и коммерческого здравоохранения

ПРОБЛЕМА

Важнейшей проблемой бесконтрольного распространением инфекции COVID-19 и пандемий других вирусных заболеваний для системы национального здравоохранения является возможность ее критической перегрузки за счет экспоненциального роста количества госпитализаций больных, и, как следствие, масштабного возрастания запроса на использование ограниченного ресурса коек отделений реанимации и интенсивной терапии (ОРИТ), а также аппаратов искусственной вентиляции легких (ИВЛ). В условиях негативного развития ситуации с распространением этой и других инфекций усиливается значимость принятия наиболее эффективных управленческих решений, а это в первую очередь зависит от полноты и достоверности имеющейся информации и наибольшей точности предсказания развития ситуации.

РЕШЕНИЕ

Сотрудники Мегапьютер Интеллидженс разработали многопараметрические модели прогнозирования развития во времени картины заболеваемости COVID-19, динамики появления случаев критического течения болезни и количества случаев выздоровления. Получаемые данные географически распределены по всем административным округам США. Сложив прогнозируемые цифры с информацией о доступных медицинских ресурсах, мы рассчитали для каждого округа срок, исчерпания мест в отделениях интенсивной терапии и реанимации (ОРИТ) для лечения критически больных с COVID-19. Для наглядного восприятия все данные визуализированы в виде тепловой географической карты.

РЕЗУЛЬТАТ

- Система представляет результаты по текущим данным и результаты прогнозного моделирования в виде наглядных, интерактивных графических карт и дает возможность выгрузки результатов моделирования в различных табличных форматах. Углубленный анализ позволяет пользователю рассматривать картину как в масштабах всей страны, так и "детализировать" прогнозы на уровне региона и муниципального образования при наличии исходных данных для модели.
- Система дает возможность «одновременно» наблюдать и анализировать такие важные параметры, как количество случаев заболеваний, новых случаев за день/неделю/месяц, процент задействованного коечного фонда и прогнозируемый срок наступления "пессимистичного" сценария исчерпания ресурсов данного фонда.

АВТОМАТИЧЕСКАЯ СВЕРКА АКТОВ

ЗАКАЗЧИК: Компания-авиаперевозчик

ПРОБЛЕМА

Масштабный рост бизнеса авиакомпании привел к лавинному увеличению потоков взаимобмена финансовой документацией- разительно выросла масса таких документов, как акты сверки. Их объемы стали насчитывать несколько десятков тысяч в месяц. Регламенты компании обязывают проводить регулярную сверку и обработку этой документации. Данный процесс требовал пропорционального роста бухгалтерского штата, а ручной труд неизбежно приводил к фактическим ошибкам сверки и, как результат, финансовым потерям.

ЗАДАЧА

Автоматизация процесса ручной сверки и обработки актов, счетов, накладных и других финансовых и транспортных документов.

РЕШЕНИЕ

На базе PolyAnalyst был разработан следующий автоматизированный функционал:

1. Предобработка и фиксация акта сверки взаиморасчетов в системе регистрации. Полученные документы подвергаются процедуре автоматического распознавания и интеллектуального поиска ключевых атрибутов/сущностей документов. Далее акты автоматизированно регистрируются в ERP-системе.
2. Проверка акта сверки взаиморасчетов. Документ распознается, автоматически построчно разбирается, проводится интеллектуальный поиск всех ключевых атрибутов (номер документа, наименование, дата, сумма), сверяются суммы. По результатам сверки автоматически переводится в статус «На подпись ГБ», либо выдает информацию для составления протокола разногласий.

РЕЗУЛЬТАТ

По каждому автоматизированному функционалу определены следующие эффекты:

1. Предобработка и фиксация акта сверки взаиморасчетов в системе регистрации: снижение трудозатрат процесса в среднем на 65%, как альтернатива увеличение производительности труда до 450%, что увеличивает объемы проводимой работы при сохранении тех же величин ФОТ.
2. Проверка акта сверки взаиморасчетов: снижение трудозатрат на 60%, или, при той же занятости, увеличение объема обрабатываемых актов сверки на 66%.

Также за счет использования специального дружественного интерфейса, требуется меньшее количество времени на обработку «проблемных» актов сверки, что приводит к еще большему росту производительности труда.

АВТОМАТИЗАЦИЯ АНАЛИЗА **ДЕФЕКТОВ ПРОДУКЦИИ** В ПРОЦЕССЕ ГАРАНТИЙНОГО ОБСЛУЖИВАНИЯ

ЗАКАЗЧИК: Компания-производитель медицинского оборудования

ПРОБЛЕМА

При выявлении проблем в работе оборудования, клиенты обращаются к производителю с гарантийными требованиями. Чаще всего инициируется процедура ремонта, по результатам которой инженеры получают полезную информацию о слабых местах продукта и формируют соответствующий отчет. Анализируя и отчеты и сервисные обращения клиентов, становится возможным предупреждение таких дефектов. Но из-за многокомпонентности оборудования и больших объемов данных, ручной анализ слишком сложен и требует много времени и средств.

ЗАДАЧА

Автоматизация анализа данных о дефектах оборудования из клиентских обращений и тех. отчетов.

РЕШЕНИЕ

Решение на PolyAnalyst способно легко загружать данные из нескольких источников. Автоматизированное извлечение сущностей позволило определить важные даты, термины и имена, относящиеся к теме анализа. Встроенные веб-графики, визуализации взаимосвязей, и облака слов демонстрируют отношения между определенными компонентами продукта и типом возникающих проблем. Решение самостоятельно объединяет информацию из колл-центров и гарантийных заключений для общего анализа проблем. Сравнивая данные из обоих источников, компания смогла выявить наиболее важные проблемы в работе продукта.

Кроме того, PolyAnalyst выявляет начало системного дефекта, определяет подверженный ему компонент, предсказывает тенденцию его поломок и предлагает осуществить редизайн продукта, либо изменить условия гарантийного обслуживания.

РЕЗУЛЬТАТ

По сравнению с ручным анализом данных о поломках, решение PolyAnalyst принесло ряд качественных выгод:

1. Раннее обнаружение возникающих проблем: благодаря регулярному автоматическому обновлению анализируемых данных, компания получает актуальную информацию о возникающих проблемах продукта, и своевременно их решает.
2. Обоснованное проектирование продукта и выбор поставщика: легко определяются проблемные компоненты, их замена или замена поставщика позволяет избежать дорогостоящих неисправностей.
3. Реструктуризация гарантийных соглашений: результаты анализа поломок помогают компании определять наиболее эффективную структуру будущих сервисных соглашений. Специалисты по ремонту обучаются распознавать типичные проблемы в жизненном цикле оборудования.

УПРАВЛЕНИЕ НОМЕНКЛАТУРОЙ МАТЕРИАЛЬНО-ТЕХНИЧЕСКИХ РЕСУРСОВ

ЗАКАЗЧИК: Энергетическая компания федерального уровня.

ПРОБЛЕМА

В производственных, энергетических и добывающих компаниях комплект используемых справочников и классификаторов образует фонд нормативно-справочной информации (НСИ). При правильной организации этих справочников и классификаторов, система НСИ помогает решить такие прикладные задачи, как, например, поиск аналогов и заменяющей продукции, подбор запчастей к оборудованию, поиск поставщиков заданного вида продукции, ценовой анализ по видам продукции. Но информация в НСИ, как правило, пополняется и поддерживается сотрудниками разных подразделений. Как следствие, одни и те же объекты в разных справочниках каталогизируются и описываются по-разному, данные дублируются и противоречат друг другу, из-за чего возникают проблемы взаимодействия и обмена данными между информационными системами предприятия.

ЗАДАЧА

Унифицирование записей, избавление от задвоенностей и каталогизация позиций Единого справочника материалов, услуг и работ.

РЕШЕНИЕ

Для решения данных задач система PolyAnalyst проводит первоначальную обработку данных (удаление опечаток, ошибок и дубликатов), определяет требования к составу и структуре системы свойств и значений классов МТР, разрабатывает единую модель сущностей МТР, формирует классификаторы МТР с учетом требований заказчика и проводит как автоматическую, так и полуавтоматическую, исполняемую под контролем специалиста, классификацию.

РЕЗУЛЬТАТ

PolyAnalyst посредством собственного набора языковых алгоритмов формирует унифицированный фонд нормативно-справочной информации, что приводит к целостности и согласованности корпоративной отчетности, а также баз данных на всех уровнях.

Это приводит к следующим результатам:

- Упрощение учета продукции, автоматизация процесса;
- Экономия времени и средств, затрачиваемых на ручное выполнение параметризации;
- Улучшенное информационное обеспечение закупок;
- Эффективное управление единой базой материально-технических ресурсов;
- Предотвращение потерь компании, связанных с использованием некачественной справочной информации.

РОБОТИЗИРОВАННАЯ ЮРИДИЧЕСКАЯ ЭКСПЕРТИЗА

ЗАКАЗЧИК: Юридический отдел промышленной компании.

ПРОБЛЕМА

Масштабные объемы промышленного производства и товарооборота компании закономерно приводят к процессу активного контрактования. В крупной компании в год может заключаться более 100 000 договоров, включающих миллионы страниц текста. Большая часть договоров запускается по стандартным формам и около трети с оригинальной структурой. По регламенту документы проходят подробную юридическую экспертизу, что «нагружает» квалифицированных сотрудников юридической службы большим объемом рутинной работы и типовых операций, которые не соответствуют их высокой квалификации.

ЗАДАЧА

Автоматизация типовых операций по анализу документов в процессе проведения юридической экспертизы.

РЕШЕНИЕ

Работа над проектом проходит в три этапа:

1. Процесс обучения сотрудников компании специалистами Мегапьютер, для отработки компетенции и технологии на договорах, применяемых в компании;
2. Самостоятельное моделирование и построение архитектуры решения юристами компании;
3. Тестирование разработанного решения по заранее выстроенному алгоритму. Результат тестирования у Заказчика продемонстрировал возможность системы проводить автоматизированную экспертизу $\frac{3}{4}$ всего объема обрабатываемых в компании договоров, отправляя неоднозначные документы на анализ юристам.

РЕЗУЛЬТАТ

Система, построенная юристами компании, показывает высокую производительность, проводя экспертизу более двух сотен документов за один час работы.

Освобождение высококвалифицированных специалистов от рутинной работы ведет к закономерному повышению эффективности юристов и EBITDA компании.

В результате сотрудники компании приобрели навыки создания, тестирования и внедрения собственных аналитических решений на базе платформы PolyAnalyst.

Это позволяет компании самостоятельно модернизировать действующее решение и внедрять его в других подразделениях компании, а в перспективе продавать сторонним компаниям свои решения, монетизируя полученные навыки и компетенции.

КОНТРОЛЬ УРОВНЯ ОБСЛУЖИВАНИЯ КЛИЕНТОВ В ОНЛАЙН-ЧАТАХ

ЗАКАЗЧИК: Крупная финансовая организация.

ПРОБЛЕМА

В колл-центре компании 300 операторов оказывают сервисную поддержку клиентам, использующим финансовые инструменты компании. Контроль качества обслуживания осуществлял сторонний подрядчик, ежемесячно обрабатывая вручную 2500 случайных сообщений из клиентских диалогов и оценивая их на предмет соответствия принятым нормативам. Такое количество анализируемых сообщений составляют только 2,5% от всего объема коммуникаций, что не приводило к полноценному контролю уровня сервиса.

ЗАДАЧА

Замена ручного контроля уровня обслуживания автоматическим решением PolyAnalyst.

РЕШЕНИЕ

Система проводит лингвистический анализ 100% сообщений клиентских чатов. Проверяется их соответствие одиннадцати качественным показателям: уровень эмпатии сотрудника, грамотность структуры его сообщения, общий профессионализм, правильность ответа и принятие должной ответственности за результат. Результаты суммируются в оценочной карте сотрудника, все карты агрегируются в сводные карты, формируя уровни оценки: уровень отдельного сотрудника, рабочей группы, отдела и совокупный уровень всего колл-центра компании.

Результаты представляются руководителям в виде интерактивных дашбордов. В случае возникновения экстремальных отклонений уровня сервиса, ответственный руководитель моментально получает онлайн-оповещение.

РЕЗУЛЬТАТ

Решение позволило обрабатывать с максимальной точностью 100% клиентских коммуникаций.

Фактический экономический эффект от внедрения автоматической системы составил 11,8 млн долларов за счет отказа от ручной работы подрядчика. В 40 раз вырос объем обрабатываемых данных за счет полной автоматизации процесса.

Увеличилась частота представления отчетности с 12 раз в год до 52. Увеличилась глубина и скорость осознания компанией потребностей своих клиентов и уровня их удовлетворенности обслуживанием.

Руководителям доступны все уровни агрегации данных, что облегчает контроль за возникающими проблемами и падением уровня обслуживания.

Количество удовлетворенных и очень удовлетворенных клиентов увеличилось по сравнению с прошлым периодом более чем на 3%. Количество клиентов-промоутеров выросло на 7%.

МОНИТОРИНГ ПРОДУКТИВНОСТИ БИЗНЕСА ЧЕРЕЗ **АНАЛИЗ** **ОТЗЫВОВ КЛИЕНТОВ**

ЗАКАЗЧИК: Сеть ресторанов быстрого питания.

ПРОБЛЕМА

Компания определила, что для поддержки уровня удовлетворенности клиентов и удержания лидерства необходимо своевременно выявлять в ресторанах своей сети негативные происшествия, требующие незамедлительного вмешательства руководства. Поскольку отзывы клиентов отражают жалобы и предложения гостей, для измерения уровня продуктивности бизнеса решено анализировать именно обратную связь гостей. Это требует тщательного изучения и представления результатов по примерно 4000 ежедневно публикуемых отзывов.

ЗАДАЧА

Внедрение системы контроля коммерческой эффективности на платформе PolyAnalyst.

РЕШЕНИЕ

Аналитики Мегапьютер Интеллидженс создали набор изменяемых критериев для извлечения атрибутов из текстовых сообщений. Автоматически и точно извлекается 500 атрибутов: «холодная еда», «пресный вкус», «переваренное блюдо» и «пищевое отравление».

Эти понятия соотносятся с критериями качества компании. Результаты на уровне отдельных блюд и ресторанов агрегируются и оцениваются уже на уровнях выше: франчайзи, города, региона и страны.

Руководство с определенной периодичностью получает детализированные отчеты для управления сетью и своевременного предупреждения возникающих проблемных инцидентов.

РЕЗУЛЬТАТ

Внедренная система представляет из себя многоцелевое масштабируемое решение, обеспечивающее аккумуляцию колоссальной массы данных от франчайзи, региональных ресторанов сети и отзывов гостей.

Ключевые выгоды:

1. Повышение удовлетворенности и лояльности клиентов;
2. Формирование наиболее привлекательного ассортимента меню;
3. Обоснованность решений о рыночной экспансии и управления франчайзи;
4. Своевременные оповещения об инцидентах для эффективного антикризисного управления;
5. Комплексный и своевременный контроль уровня качества обслуживания.

АВТОМАТИЗАЦИЯ ИЗВЛЕЧЕНИЯ ИНФОРМАЦИИ О КОНТРАГЕНТАХ

ЗАКАЗЧИК: Банк, оказывающий финансовые услуги юридическим лицам.

ПРОБЛЕМА

Многие компании заинтересованы в упрощении и автоматизации процесса взаимодействия со своими контрагентами и клиентами. Сбор и обработка первичной документации до сих пор является сильно забюрократизированным процессом. Например, только для простого открытия счета в банке для юр. лиц требуется анализ несколько десятков многостраничных документов. Ручная обработка подаваемых документов требует внушительных затрат на содержание штата специалистов. Человеческий рутинный труд неизбежно приводит к возникновению ошибок во вносимых данных. Это ведет к необходимости своевременного поиска и устранения таких ошибок, а чаще всего приводит к прямым финансовым и, за счет снижения качества работы, репутационным потерям.

ЗАДАЧА

Роботизация и автоматизация извлечения, анализа и верификации первичной информации о компаниях из текстовых документов.

РЕШЕНИЕ

Автоматизированное решение на платформе PolyAnalyst, после загрузки полученных от контрагентов документов на основе встроенных инструментов извлечения «сущностей» (название компании, налоговые и банковские реквизиты, имена учредителей и руководителей, паспортные данные и т.д.), извлекает из текста первичную информацию. Проверяется комплектность и корректность подаваемой документации. Далее система самостоятельно подключается к внешним источникам данных, таких как ЕГРЮЛ, СПАРК и т.д. Автоматически проводится сверка данных и принимается решение об установлении деловых отношений. При положительном решении формируется сводная карточка данных, информацию из которой PolyAnalyst записывает в необходимую внутреннюю информационную систему (ERP, CRM).

РЕЗУЛЬТАТ

Внедрение системы позволило полностью заменить труд 42 сотрудников по работе с первичной документацией.

Обработка ежегодно поступающих в компанию пятисот тысяч документов была ускорена в 30 раз.

Общий экономический эффект измеряется десятками миллионов рублей за счет экономии на ручном труде, ускорении извлечения информации и предотвращении возникающих при человеческой работе ошибок обработки и внесении данных.

БОРЬБА С МОШЕННИЧЕСТВОМ КОНТРАГЕНТОВ ЧЕРЕЗ АНАЛИЗ ДАННЫХ ГАРАНТИЙНОГО ОБСЛУЖИВАНИЯ

ЗАКАЗЧИК: Крупный автопроизводитель.

ПРОБЛЕМА

С целью определения причин возникновения поломок, расчета оптимальной гарантийной компенсации и длительности гарантии, оптимизации затрат на проведение ремонта и борьбы с мошенничествами автодилеров, компания приняла решение начать анализировать объемную массу накопившихся и вновь поступающих гарантийных требований, а также отчетов о проведенных ремонтах реализованных автомобилей.

ЗАДАЧА

Внедрение системы анализа гарантийных требований и ремонтных отчетов на платформе PolyAnalyst.

РЕШЕНИЕ

В качестве предварительной подготовки на основе автоматического анализа жалоб клиентов и отчетов механиков, проводящих ремонт, были сформированы словари и таксономии компонентов автомаобилей и возможных проблем с ними.

Система анализа обнаруживает корреляции ключевых поломок с серией, маркой и годом выпуска автомобиля. Проблемы ранжируются по влиянию на безопасность и затраты на устранение.

Проводя анализ динамики поломок определенных компонентов авто и сопоставляя их с маркой, серией и годом выпуска, система обеспечивает раннее обнаружение пока что малозаметных, но учащающихся поломок, рискующих превратиться в серьезные проблемы в скором будущем. Определяются основные причины проблем и способы их устранения.

Проводится обнаружение подозрительных отклонений от нормальных показателей в отчетах дилеров.

РЕЗУЛЬТАТ

Определяя основные причины возникновения проблем, компания предотвращает их появление у новых моделей еще на этапе проектирования и выбора поставщика.

Контролируя корреляцию типичных поломок и срока их возникновения у определенных моделей и серий, с наибольшей точностью устанавливается стоимости гарантии и период ее предоставления.

Анализируя динамику возникновения поломки предотвращаем появления массовых, системных проблем в работе авто, оптимизируя затраты на проведение ремонта.

Выявляя подозрительные статистические отклонения в отчетах дилеров о поломках и ремонтах, определяются вероятность мошенничества и необходимость глубокого расследования ситуации.

КЛАССИФИКАЦИЯ СООБЩЕНИЙ ФОНДА ЦЕЛЕВЫХ КАПИТАЛОВ

ЗАКАЗЧИК: Фонд Потанина.

ПРОБЛЕМА

В целях привлечения жертвователей (small donors) Фонды целевых капиталов проводят масштабную информационную политику в социальных сетях. Однако, несмотря на это, сообщества Фондов небольшие и активность подписчиков сообществ остается достаточно низкой, а некоторые и вовсе закрываются. Для максимизации активности было решено осуществить оценку эффективности различного рода каналов и сообщений, чтобы определить, какого рода информация является наиболее привлекательной для потребителя

ЗАДАЧА

Классификация авторских сообщений в сообществах Фондов целевых капиталов в социальных сетях на 3 категории (информационные, коммуникативные, репутационные).

РЕШЕНИЕ

Идентифицировано 100 сообществ фондов целевых капиталов в различных социальных сетях и собрано более 43 тыс. сообщений. Для обучения модели подготовлен классификатор на основе 10 тыс. вручную размеченных сообщений. На основе этих сообщений был обучен алгоритм автоматической классификации текстов, а также для каждого сообщений рассчитан индекс онлайн-активности, учитывающий лайки, репосты, комментарии.

РЕЗУЛЬТАТ

1. Разработано автоматическое решение для классификации сообщений социальных сетей, позволяющее разбивать сообщения на категории.
2. Расчет индекса онлайн-активности позволил определить наиболее активные каналы распространения информации.
3. Заказчик получил “портфель медиа активности” каждого Фонда целевых капиталов в социальных сетях с оценкой их активности и понимание того, какой контент наиболее привлекателен.

ОЦЕНКА МЕДИЙНОГО ОБРАЗА

ЗАКАЗЧИК: Томский государственный университет.

ПРОБЛЕМА

Томский государственный университет является одним из ведущих университетов России и входит в программу повышения конкурентоспособности 5-100. Одним из главных источников информации о ВУЗе являются социальные медиа. Из них стейкхолдеры узнают текущее положение университета, а абитуриенты выбирают место учебы по отзывам других студентов. Следовательно, в условиях конкуренции университету очень важно управлять своим имиджем, а также оперативно реагировать на оценки, как положительные, так и отрицательные.

ЗАДАЧА

Определить текущий образ университета, а также его ключевые слабые и сильные стороны.

РЕШЕНИЕ

Были определены основные информационные площадки (сайты, социальные сети, “отзовики”, где публикуется информация об университете. Все данные агрегировались в единую базу, выделялись основные категории (личности, условия и качество обучения), а также осуществлялась оценка тональности (эмоциональный окрас) каждой категории. Также были выделены основные освещаемые события и мероприятия, связанные с университетом, и классифицированы направления деятельности (наука, образование, общество), определены основные “конструкты” медиаобраза университета. Решение автоматизировано, т.е. система сама автоматически подгружает и классифицирует новые инфоповоды, связанные с университетом, “достраивая” таким образом текущий имидж университета, и позволяет осуществлять анализ в динамике.

РЕЗУЛЬТАТ

1. Выделены основные информационные площадки, публикующие информацию об университете. Составлена единая информационная база с выделением сообщений о качестве и условиях обучения.
2. Произведена оценка тональности сообщений для мониторинга мнений об университете и своевременного реагирования на возникающие жалобы и недочеты.
3. При публикации новых сообщений они автоматически попадают в базу и анализируются, что позволяет следить за текущей ситуацией и своевременно принимать необходимое решение.

ОЦЕНКА МНЕНИЙ И НАСТРОЕНИЙ В РЕГИОНАХ РФ

ЗАКАЗЧИК: Администрация Томской области.

ПРОБЛЕМА

Оценка мнений и настроений в регионах связана с проблемой измерения. Социальные медиа стали сегодня одним из основных источников данных о взглядах, настроениях, переживаниях и поведении людей и воспроизводят те же процессы и явления, которые имеют место быть в офлайн. Через социальные сети можно получить оценку людей по условиям своей жизни и существующих на сегодняшний день социальных проблем в регионах.

ЗАДАЧА

Создать модель оценки мнений и настроений населения субъектов РФ посредством анализа высказываний и реакций пользователей в социальных сетях и рассчитать агрегированный показатель условий жизни населения для разработки системы быстрого реагирования на социальные проблемы.

РЕШЕНИЕ

Разработана модель оценки мнений и настроений населения, которая создана на основе достоинств и недостатков, существующих на сегодняшний день методики измерения благополучия и условий жизни населения. Модель состоит из 19 категорий социальной, экономической и политической сфер общественной жизни. Более 60 тыс. публикаций было очищено и размечено по 19 категориям вручную. На основе полученных данных был обучен алгоритм автоматической очистки и классификации текстовых сообщений по категориям, а также по анализу тональности контента региональных сообществ. Разработана система быстрого реагирования на жалобы или проблемы населения.

РЕЗУЛЬТАТ

1. Разработан алгоритм автоматической классификации сообщений по 19 категориям и 3 типам тональности;
2. Классифицировано более 3,2 млн публикаций региональных сообществ для дальнейшего расчета индекса онлайн активности населения национальных и не национальных регионов;
3. Расчет индекса онлайн-активности позволил калькулировать индекс благополучия для каждой категории социальной, экономической и политической сфер общественной жизни по 85 субъектам РФ;
4. С использованием разработанной формулы был рассчитан агрегированный показатель благополучия населения.

ОЦИФРОВКА КОРПОРАТИВНОГО АРХИВА И СИСТЕМА ИНТЕЛЛЕКТУАЛЬНОГО ПОИСКА ПО БАЗЕ ДОКУМЕНТОВ

ПРОБЛЕМА

За период своей деятельности корпорация накопила значительные объемы корпоративной документации (архив). В силу исторических причин и специфики деятельности компании существенную часть данного архива занимают документы в бумажном виде, содержащие техническую информацию (диспетчерские журналы, результаты НИР, НИОКР, эксплуатации оборудования, протоколы испытаний и т.д.). Информация и знания, содержащиеся в бумажных архивах, в настоящее время почти не используются, хотя она часто может быть востребована при проведении новых НИР/НИОКР, формирования моделей прогноза при эксплуатации оборудования, проектировании, решения других производственных задач.

ЗАДАЧА

Оцифровка бумажной документации, разработка решения для анализа электронных текстовых документов на основе механизмов искусственного интеллекта, обработки естественного языка (Natural Language Processing – NLP) и извлечения информации.

РЕШЕНИЕ

На первом этапе специалистами Мегапьютер проводится экспертиза бумажного архива и участие в разработке ТЗ на ЭИР- электронные информационные ресурсы архива. Далее с применением технологии OCR проводится распознавание титулов документов. Распознанные данные анализируются посредством алгоритмов платформы PolyAnalyst и формируются карточки документов (метаданные). Таким образом формируется структура каталога архива, а также проходит первичная классификация и построение онтологий. Далее проходит полная оцифровка бумажного архива через системы OCR и индексирование документов в PolyAnalyst. Результирующая ЭИР и система интеллектуального поиска интегрируется с имеющимися ИТ-системами.

РЕЗУЛЬТАТ

Результатом реализации проекта является формирование электронного цифрового архива и разработка системы интеллектуального поиска по документам архива.

Это – Цифровой актив, который можно (при необходимости) интегрировать в имеющуюся Систему управления знаниями, процессы R&D, проектирования, строительства и т.д.

ЭФФЕКТЫ ВНЕДРЕНИЯ РЕШЕНИЯ:

- Вовлечение «потерянных в бумаге» документов в производственный оборот.
- Повышение скорости доступа к информации, содержащейся в электронных архивах для сотрудников.
- Снижение риска утраты ценных данных, хранящихся в бумажной форме.

АНАЛИЗ ЦИФРОВОГО СЛЕДА

ЗАКАЗЧИК: Университет 2035

ПРОБЛЕМА

Организация Университет 2035 является организатором множества образовательных интенсивов, таких как "Остров 10-21", "Зимний остров" и "Архипелаг". В рамках таких мероприятий Университет 2035 реализует модель обучения с применением аналитики данных. Сложность данного процесса состоит в необходимости ручной обработки большого объема разнородной информации, в том числе данных об освоенных компетенциях обучающихся.

ЗАДАЧА

Разработка и внедрение автоматизированной системы анализа цифрового следа и формирования цифрового компетентностного профиля.

РЕШЕНИЕ

На базе аналитической платформы PolyAnalyst было разработано специализированное интеллектуальное решение, проводящее четырехэтапную обработку данных. С образовательного мероприятия собирается цифровой след каждого участника. Сперва данные, представленные в форме неструктурированных текстов, таблиц, списков, файлов презентаций загружаются из множество разнородных источников и приводятся к единому формату. Далее проходит предобработка данных: проверяется орфография, удаляются дубликаты. На следующем этапе проходит формирование модели. Автоматически формируется и наполняется таксономия образовательных результатов. Данные проходят классификацию, и на завершающем этапе результаты классификации по данным из различных источников сопоставляются.

РЕЗУЛЬТАТ

В рамках одного интеллектуального решения PolyAnalyst используется для формирования таксономии образовательных результатов и ее применения для разметки образовательного контента, анализа цифрового следа, экспорта данных в формы цифровых профилей участников.

ЭФФЕКТЫ ВНЕДРЕНИЯ РЕШЕНИЯ:

- Сокращение объемов рутинной ручной работы и возможность перевести сотрудников на более сложную работу;
- Сокращение временных затрат на обработку данных- решение обрабатывает 4800 документов в минуту;
- Создано решение с возможностью повторного использования для обработки новых документов;
- Наглядное представление результатов работы системы приводит к повышению эффективности принимаемых решений.

СОЗДАНИЕ ЦИФРОВОЙ МОДЕЛИ КОМПЕТЕНЦИЙ УЧЁНОГО

РАЗРАБОТЧИК: Университет 2035

ЗАДАЧА

Коллективом Университета 2035 была поставлена задача по разработке модели компетенций ученого, основанной на анализе больших датасетов и отражающей потребность рынка труда.

РЕШЕНИЕ

Сперва командой проекта было проделана работа по выборке из крупнейших российских и зарубежных рекрутмент-ресурсов вакансий ученых. По собранным названиям вакансий в PolyAnalyst была сформирована таксономия позиций ученых пяти типов: начинающий ученый, исследователь-практик, исследователь-разработчик (R&D), менеджер, академик. Далее по полученным из фреймворков данным о навыках ученых был сформирован обобщенный перечень компетенций для разметки вакансий. Модель компетенций обогащалась за счет формирования таксономий специфических для каждого типа ученых компетенций. Полученные данные обрабатывались, а результаты визуализировались за счет встроенных алгоритмов визуализации PolyAnalyst.

РЕЗУЛЬТАТ

В итоге были получены и визуализированы в виде графовой структуры компетенции четырех типов ученых.

Компетенции были представлены по трем категориям:

- Жесткие навыки (hard skills);
- Гибкие навыки (soft skills);
- Владение инструментами (tools);

С помощью PolyAnalyst решались задачи формирования таксономии профессий и компетенций ученых, а также ее дальнейшего применения для анализа датасетов вакансий с целью демонстрации потребностей рынка труда, а также разметки образовательного контента.

СОЗДАНИЕ ПРОФЕССИОНАЛЬНОГО СТАНДАРТА

РАЗРАБОТЧИК: Университет 2035

ЗАДАЧА

Разработка профстандарта «Специалист по машинному обучению».

РЕШЕНИЕ

На примере профессии «Специалист по машинному обучению» была апробирована методика построения профстандарта на основе данных рынка. Посредством платформы PolyAnalyst был проведен анализ датасетов вакансий с целью поиска вакансий, включающих требование о наличии навыка работы с машинным обучением. Система извлекла необходимые факты и классифицировала их по классам: функция, действие, компетенция.

РЕЗУЛЬТАТ

На основе данных, извлеченных платформой PolyAnalyst, были сформированы трудовые функции и действия, а также необходимые компетенции для специалиста по машинному обучению.

В данный момент разработанный профстандарт проходит процедуру согласования и утверждения в Министерстве труда РФ.

ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ОБРАЩЕНИЙ ГРАЖДАН В ПРИЕМНУЮ ГУБЕРНАТОРА

ЗАКАЗЧИК: Региональное правительство

ПРОБЛЕМА

Региональному Аналитическому центру была поставлена задача за 48 часов провести исследование 1 миллиона текстовых обращений граждан, направленных в приемную правительства из разных источников. Необходимо было 1)определить проблематику обращения 2)определить его эмоциональную окраску 3)провести глубокий анализ обращений, касающихся тематики «мусора». Последняя задача вызвана необходимостью мониторинга отношения граждан к «мусорной реформе». Специалисты центра традиционно используют в своей работе высокоуровневые языки программирования, которые требуют для работы трудоемкого кодирования, размеченного датасета и необходимость многоразового переобучения машинных моделей. Все это стало препятствием в реализации проекта в сжатые сроки.

ЗАДАЧА

Применение системы PolyAnalyst для интеллектуального анализа обращений.

РЕШЕНИЕ

Специалисты Мегапьютер и Аналитического центра совместно передали в PolyAnalyst датасет обращений. Сначала данные предобработали, исправили орфографию, выявили и удалили дубликаты. 1/5 обращений имела разметку в виде принадлежности к определенной тематике (всего 10 тем). На основании этой разметки была обучена и применена к оставшимся данным модель машинного обучения. Это в совокупности с применением алгоритма определения ключевых слов позволило быстро решить задачу 1. Задача 2 была решена с помощью имеющего в PolyAnalyst контекстного алгоритма определения эмоциональной окраски, который в отличие от машинной модели показал высокий результат. Для решения задачи 3 через поисковые запросы были выявлены все сообщения по тематике «мусор». Далее был проведен разведочный анализ через автоматическое выявление ключевых слов и кластеризации. На основании выявленных данных была сформирована таксономия для всех обращений по данной теме.

РЕЗУЛЬТАТ

Применение системы PolyAnalyst позволило в кратчайшие сроки провести глубокий интеллектуальный анализ данных. Наличие в системе собственного языка написания лингвистических правил дало возможность извлекать информацию из данных, не проводя трудоемкую и длительную по времени разметку.

Встроенные машинные алгоритмы позволили провести разведочный анализ данных, результаты которого дали аналитикам виденье того, на какие тематики и объекты в данных необходимо обратить внимание.

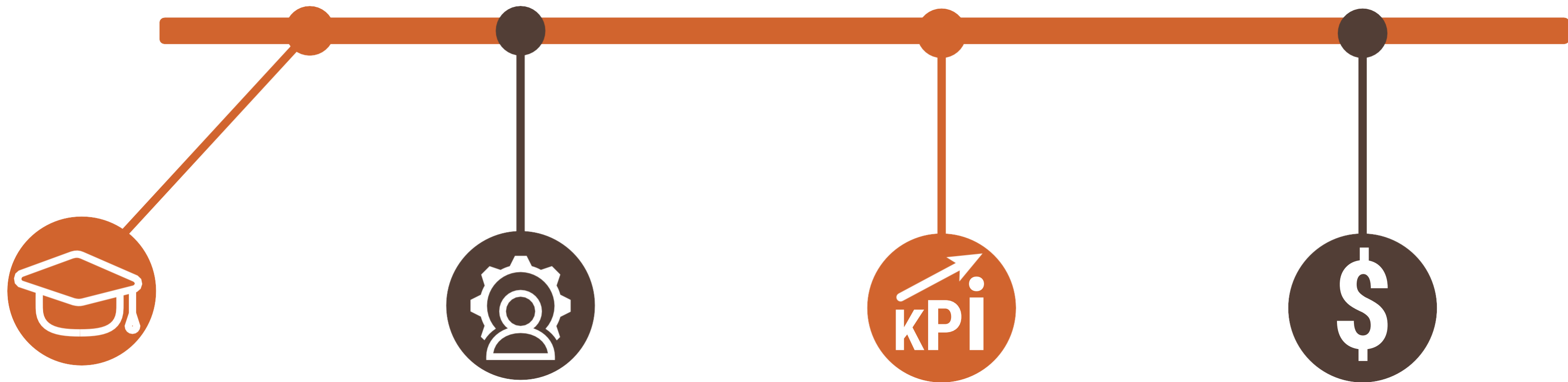
Анализ проводился без программирования в визуальном интерфейсе. Это дало возможность аналитикам центра уже после всего двух дней работы широко ознакомиться с инструментами PolyAnalyst и вести дальнейшие исследования самостоятельно.

Используя встроенные модули визуализации, специалисты сформировали итоговую графическую отчетность без необходимости интеграции данных в сторонние системы.³¹

ЭФФЕКТ КОМПЕТЕНЦИЙ ДЛЯ ЗАКАЗЧИКА

PolyAnalyst - это открытая среда разработки аналитических решений и автоматизации бизнес-процессов. Разработка на платформе **не требует навыков программирования**. Нужны только экспертные знания специалистов в своей предметной области.

Получение **собственных компетенций** по работе с технологиями обработки текстов на естественном языке (Natural Language Processing, NLP) с помощью PolyAnalyst может дать компании – Заказчику дополнительный существенный **экономический эффект при решении различных бизнес задач, в которых возникает необходимость использования NLP**.



ЭФФЕКТ: РАЗРАБОТКА НАБОРА СОБСТВЕННЫХ РЕШЕНИЙ, СОКРАЩАЮЩИХ ТИПОВЫЕ РУТИННЫЕ ОПЕРАЦИИ



- Сотрудники компании получают **навыки** работы с Polyanalyst и разрабатывают **собственное решение по автоматизации**.
- Например, юристы не только автоматизируют экспертизу договоров, но и создают систему поиска по базе юридических документов, прошедших через ЭДО.

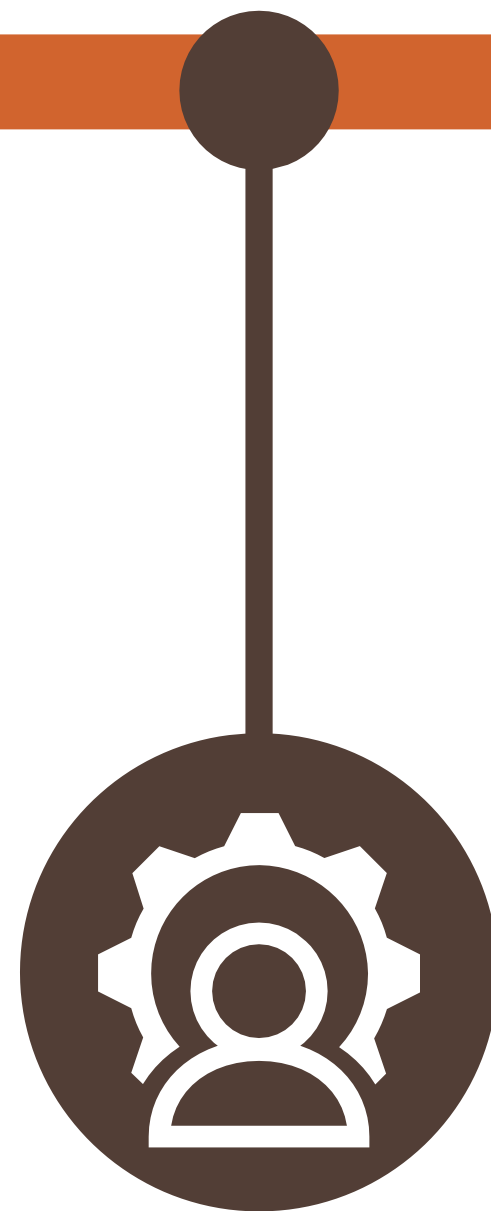
Процесс экспертизы и поиска **ускоряется**, а у юристов и прочих сотрудников, занятых в таких процессах, **высвобождается от рутины дорогостоящее рабочее время**. Это время можно использовать для **повышения собственной квалификации** и решению более **сложных правовых вопросов**, требующих обязательного человеческого участия, таких как разрешения разногласий с контрагентами и участия в судебных мероприятиях.

ЭФФЕКТ: РАЗРАБОТКА КРОСС-ФУНКЦИОНАЛЬНЫХ РЕШЕНИЙ ВНУТРИ КОМПАНИИ (1)

Пример:

Команда **юристов** развивает свое решение, глубже изучает PolyAnalyst, **становясь центром компетенций** своей компании при работе с данными.

Она делится своей практикой с другими подразделениями.



Решения на основе PolyAnalyst можно применять для проведения анализа различных типов документов, не только коммерческих договоров, но и актов, счетов, технологической документации, технико-экономических обоснований.

ЭФФЕКТ: РАЗРАБОТКА КРОСС-ФУНКЦИОНАЛЬНЫХ РЕШЕНИЙ ВНУТРИ КОМПАНИИ (2)

Центр компетенций разрабатывает решения на PolyAnalyst для различных подразделений.

Автоматизация ведет к **снижению затрат** на персонал и **повышению эффективности** его работы, и, как следствие, **росту финансовых показателей компании**.



Снижая объем привлечения к разработке внешних подрядчиков, компания резко сокращает **стоимость** производства ПО, **время** разработки, сокращает **расходы** на эксплуатацию.

ЭФФЕКТ: МОНЕТИЗАЦИЯ/ПРОДАЖА ВНУТРЕННЕЙ КОМПЕТЕНЦИИ НА ВНЕШНЕМ РЫНКЕ

После того, как внутренние решения на базе PolyAnalyst принесли эффект, функционирование систем отлажено и их работоспособность подтверждена **продуктивной эксплуатацией** внутри структур компании, **ЭТИ РЕШЕНИЯ МОГУТ БЫТЬ ПРОДАНЫ** как типовое решение для других компаний отрасли, как собственная разработка на базе платформы PolyAnalyst.

То есть внутренняя экспертиза может быть дополнительно монетизирована. Группа разработки может производить на заказ решения под индивидуальные нужды и специфику клиентов и аналогичных компаний в отрасли.



В результате компания **диверсифицирует** продуктовый портфель, и получает **дополнительный приток** выручки, а это и есть **ЭКОНОМИЧЕСКИЙ ЭФФЕКТ КОМПЕТЕНЦИЙ**.

ПАРТНЕРСКАЯ ПРОГРАММА МЕГАПЬЮТЕР ИНТЕЛЛИДЖЕНС

PolyAnalyst 6.5

ЗАРАБАТЫВАЙТЕ НА POLYANALYST

- **Раньше** MEGAPUTER (и его Дилеры-агенты) продавали лицензии на платформу PolyAnalyst, а MEGAPUTER дополнительно разрабатывал решения под Заказчика на платформе.
- **Теперь** ПАРТНЕРЫ МОГУТ ЗАРАБАТЫВАТЬ НА: (1) продаже лицензий, (2) разработке решений под Заказчика и их интеграции, (3) разработке и **СОВМЕСТНОЙ ПРОДАЖЕ** типовых решений **ПО МОДЕЛИ REVENUE SHARING**

Кто продает
ПЛАТФОРМА POLYANALYST

Кто разрабатывает
РЕШЕНИЯ, СЦЕНАРИИ, ИНТЕРФЕЙСЫ

ВЧЕРА

Мегапьютер/Агент
Продает лицензии на платформу

Заказчик
Сам научился писать сценарии, сам **разрабатывает решение для себя** и никому не показывает

Мегапьютер/Агент
Продает лицензии на платформу

Заказчик/Мегапьютер
Разрабатывают сценарии анализа, делают стыковки с ИТ системами Заказчика через API вместе с партнерами по интеграции

СЕГОДНЯ

Мегапьютер/Агент
Продает лицензии на платформу

Партнер (эксперт в предметной области) разрабатывает решения под Заказчика, делает интеграцию, разрабатывает и продает ТИРАЖИРУЕМЫЕ решения (сценарии и интерфейсы) при поддержке специалистов Мегапьютер

МЕГАПЬЮТЕР ИНТЕЛЛИДЖЕНС

Извлекаем и структурируем факты из текстовых документов

Оцифровываем и роботизируем бизнес-процессы

Строим модели на основе аналитики и Искусственного Интеллекта

107 разработчиков, 16 лингвистов и аналитиков, 9 кандидатов наук

Предоставляем кластерную платформу для анализа Больших Данных

Поддерживаем четверть компаний из списка Fortune 100 и еще более 100 клиентов

Член Ассоциации Разработчиков Программных Продуктов «Отечественный софт»

Платформа PolyAnalyst включена в реестр Российского ПО. Свидетельство №4414

P&G

Canon



СБЕР

Детский мир

Deloitte.

Lilly

hp

НИИ ВОСХОД

Счетная палата Российской Федерации

ПРАВО RU

КОНСИСТ-ОС РОСАТОМ

Roche

EY Building a better working world

фрчч

СИБУР

СВЯЗНОЙ

20.35 УНИВЕРСИТЕТ

КОМАНДА МЕГАПЬЮТЕР



Давид Сазонов

Руководитель направления
текстового анализа
sazonov@megaputer.ru



Алексей Русских

Генеральный директор
+7 (915) 424-23-45
russkikh@megaputer.ru



Сергей Ананян

Исполнительный директор
sananyan@megaputer.com



Гольцов Дмитрий

Заместитель ген. директора
Коммерческое направление
+7 (916) 111-95-49
goltsov@megaputer.ru

Мегапьютер Интеллидженс
Москва, ул. Бауманская 6, офис 723
+7 (499) 753-01-29
www.megaputer.ru