



**«Цифровой эксперт по договорам»
на платформе
PolyAnalyst 6.5**

ПЛАТФОРМА POLYANALYST

АНАЛИТИЧЕСКАЯ ПЛАТФОРМА

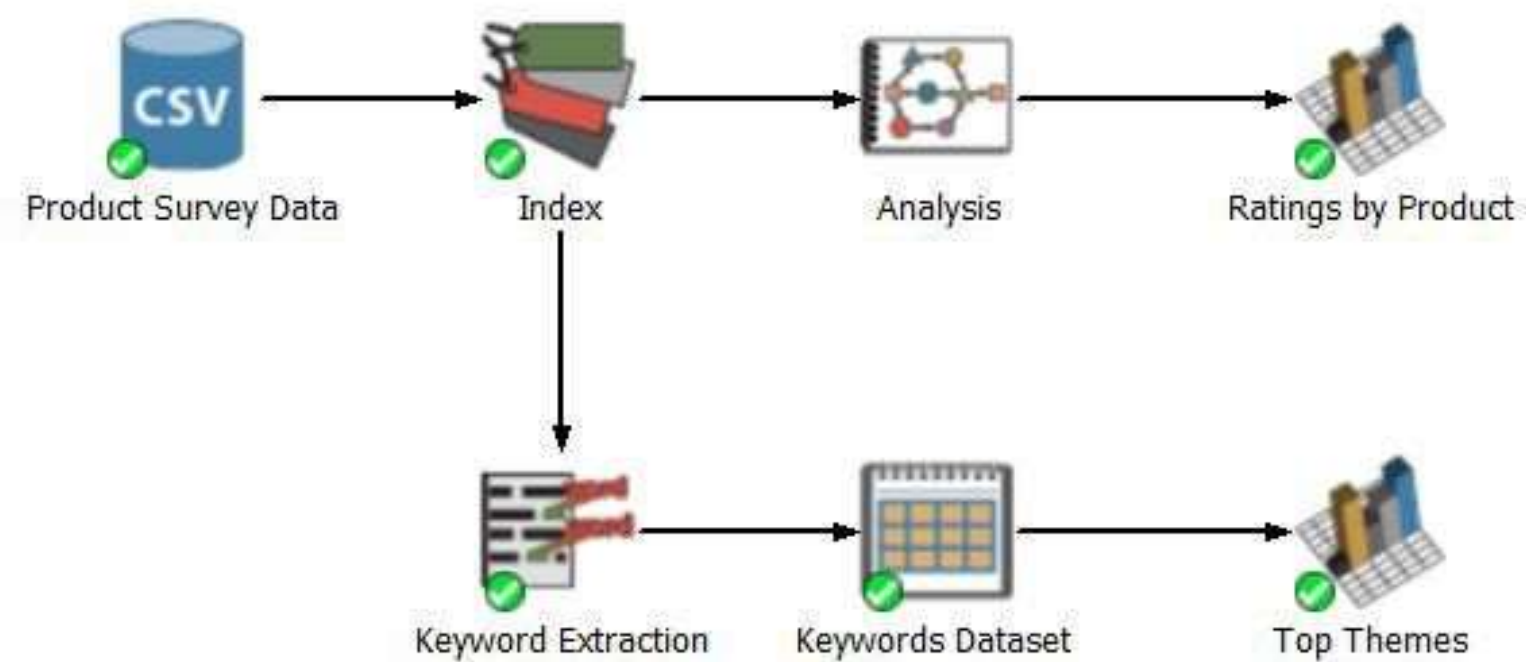
PolyAnalyst – это система для анализа текстовых документов и структурированных данных. Она комбинирует ИИ, лингвистический анализ текста и машинное обучение.

ЛИНГВИСТИЧЕСКИЙ АНАЛИЗ

В системе PolyAnalyst осуществляется лингвистический анализ текста, основанный на заданных правилах и машинном обучении. Анализ включает автоматическое извлечение именованных сущностей, определение тональности, изыскание фактов.

МАШИННОЕ ОБУЧЕНИЕ

В платформу встроены средства, использующие алгоритмы машинного обучения: классификация и кластеризация документов, извлечение условий договора.



PolyAnalyst™ входит в Единый реестр российских программ для ЭВМ и баз данных (реестровый №4414)

МАШИННОЕ ОБУЧЕНИЕ

КЛАССИФИКАЦИЯ ДОКУМЕНТОВ

в том числе иерархическая

ИЗВЛЕЧЕНИЕ ИНФОРМАЦИИ

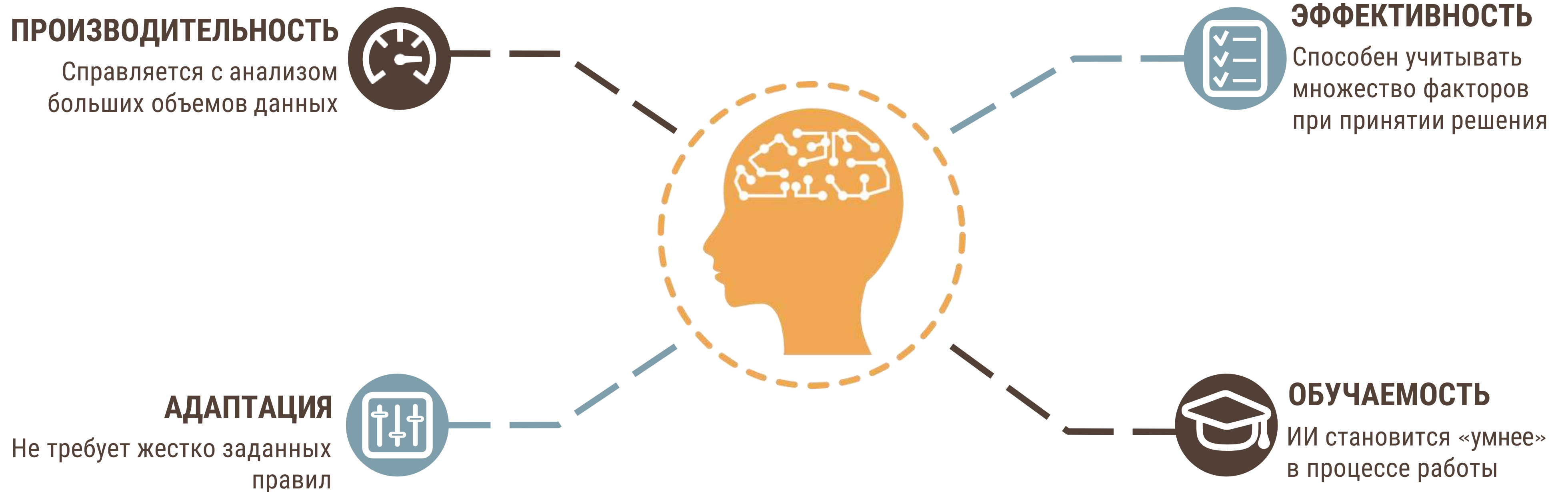
общей (названия компаний, имена людей, даты) и предметно-ориентированной (реквизиты, предмет договора, условия оплаты)

КЛАСТЕРИЗАЦИЯ ДОКУМЕНТОВ

группировка схожих документов

The screenshot displays the POLYANALYST platform interface. On the left, a hierarchical tree structure shows document classification. The root node is 'Виды договоров (5080)', which branches into 'Временное владение и пользование (445)' and 'Аренда (240)'. Under 'Временное владение и пользование', there are sub-nodes for 'оборудование (234)', 'автомобиль (50)', 'земельный участок (53)', and 'Others (181)'. Under 'Аренда', there are sub-nodes for 'Условия расторжения арендатором (204)', 'Письменное уведомление (204)', and 'Уплата арендной платы (208)'. The 'Письменное уведомление' node further branches into 'менее 30 дней (180)' and 'не менее 30 дней (24)'. The 'Уплата арендной платы' node branches into 'не более 5 календарных дней (54)', '5-10 календарных дней (152)', and 'более 10 календарных дней (2)'. On the right, the 'Drill-down' tab is active, showing a text extraction view of a document. The text is structured into sections, with the first section being '1. ОБЩИЕ ПОЛОЖЕНИЯ'. The text describes a lease agreement where the lessor transfers temporary ownership and use of a plot of land to the lessee for a fee. Key terms like 'владение' (ownership) and 'пользование' (use) are highlighted in red. The text also mentions the plot area (200 m2), cadastral number, and address. The document is identified as a lease agreement for a plot of land.

ПРЕИМУЩЕСТВА МАШИННОГО ОБУЧЕНИЯ



ТРУДНОСТИ МАШИННОГО ОБУЧЕНИЯ

Для первичного обучения системы необходимы **размеченные документы**. Разметка делается **вручную**.



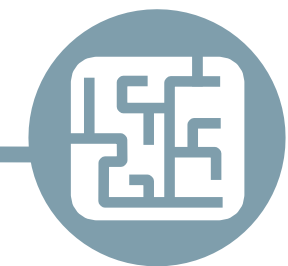
Для качественной модели требуется очень **большое количество** размеченных документов.



Для **повторного обучения** системы **снова** требуется масса проверенных вручную документов.



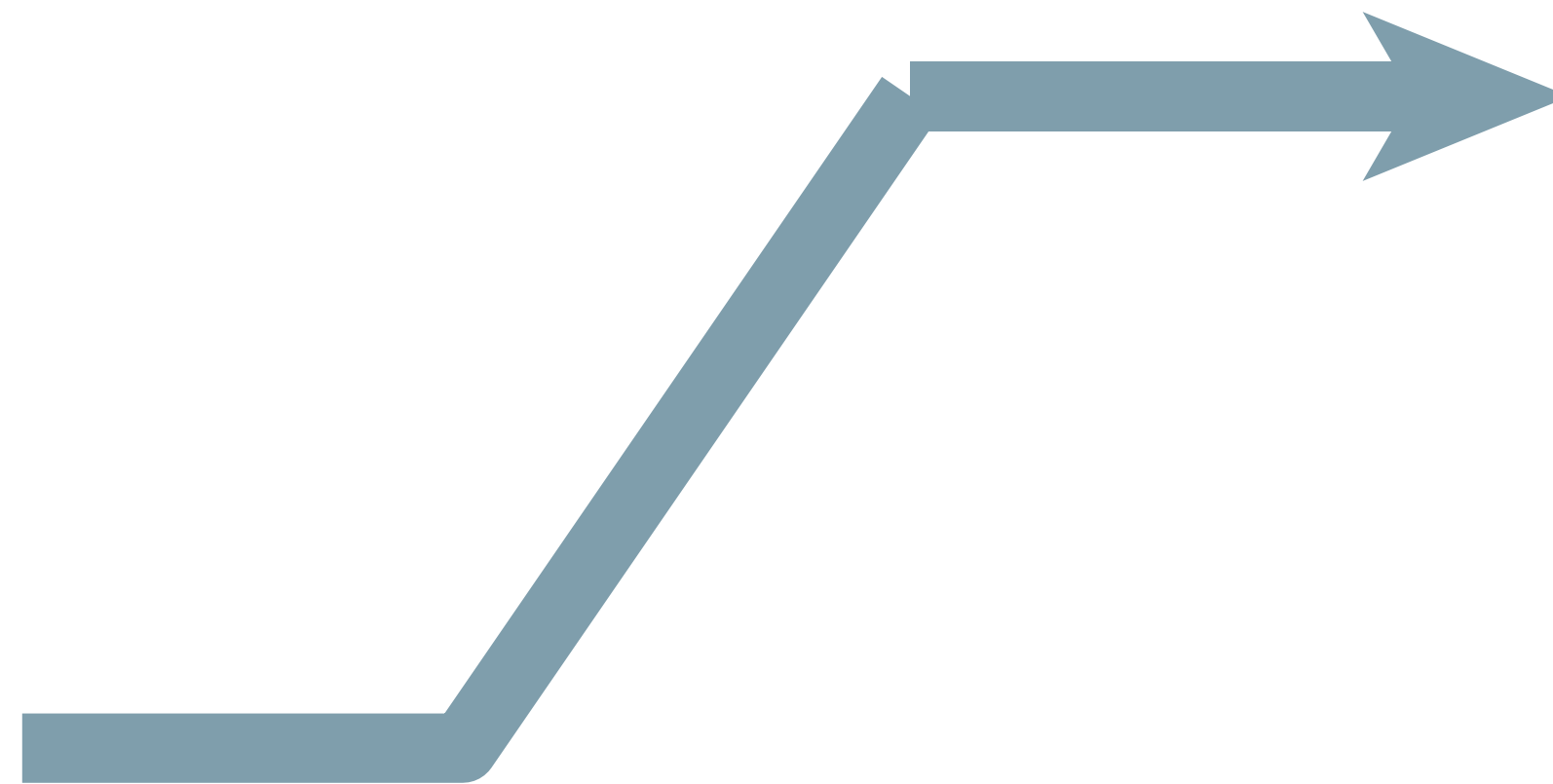
Сложно понять логику принятия решений и исправить ошибки.



УПРОЩЕНИЕ РАЗМЕТКИ ДАННЫХ

ПРОБЛЕМА

Отсутствие/недостаточность
размеченных данных для
машинного обучения



РЕШЕНИЕ

Автоматизация разметки с
помощью **языка запросов**

ПРИМЕРЫ ПРАВИЛ РАЗМЕТКИ

PHRASE

(3, расчет, ежемесячно or ежеквартально)

Расчеты осуществляются **ежемесячно...**

PHRASE

(ежемесячный, синоним(плата))

Ежемесячная оплата услуг
осуществляется...

Размер **ежемесячной платы** установлен
в Приложении № 1 .

PHRASE

(2, синоним(плата), за, каждый, term(период))

...подлежит **уплате** авансом **за каждый
квартал.**

ВОЗМОЖНОСТИ ЯЗЫКА ЗАПРОСОВ

ПОИСК С УЧЕТОМ КОНТЕКСТА

Возможность указать требования к совместной встречаемости слов.

ПАРСИНГ СТРУКТУРЫ ДОКУМЕНТА

Поиск в определенных разделах. Например, найти слово «арендатор» в заголовке.

ИСПОЛЬЗОВАНИЕ СЛОВАРЕЙ

Поиск синонимов и словарей выражений. Например, найти синонимы слова «оплата».

ПОИСК ПО ОНТОЛОГИЯМ

Поиск с учетом семантических связей слов (гиперонимов, гипонимов). Например, найти все виды подзаконных актов.

ПОИСК В КОНТЕКСТЕ ОТРИЦАНИЯ

Возможность исключить из поиска отрицательные и модальные конструкции.

ТАБЛИЦЫ

Работа с таблицами и их содержимым. Например, найти слово «сумма» в ячейке таблицы.

ДОПОЛНИТЕЛЬНЫЕ ПАРАМЕТРЫ ПОИСКА

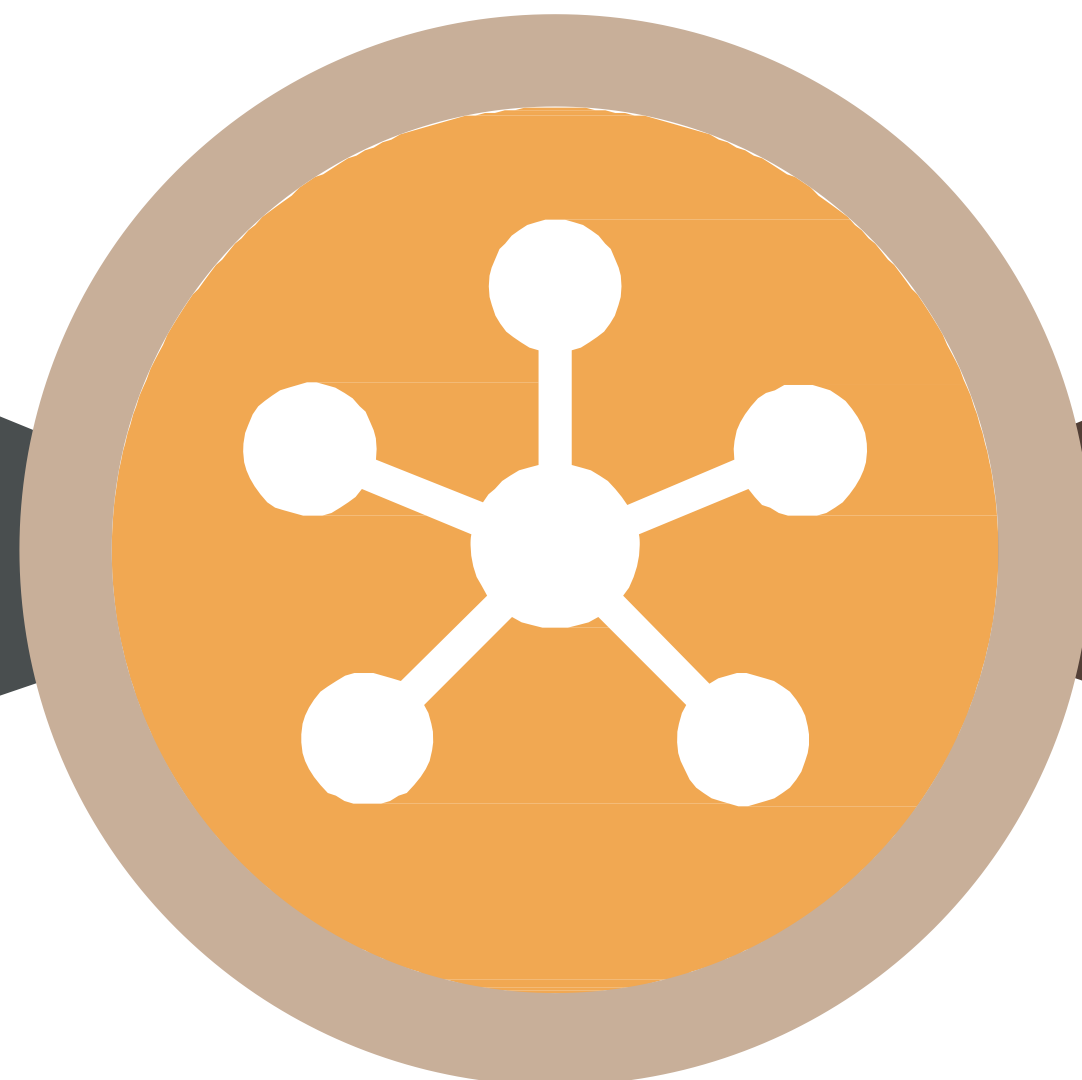
Поиск слов на определенном расстоянии, в определенном регистре. Поиск точных вхождений и с учетом изменения слова.

DATA MINING **ПО ПРАВИЛАМ**

СИСТЕМА ЭКСПЕРТНЫХ ПРАВИЛ

ПРЕИМУЩЕСТВА

Высокая точность,
предсказуемость результатов.



НЕДОСТАТКИ

Необходимость поиска
НОВЫХ
КОНТЕКСТОВ.

DATA MINING ПО ПРАВИЛАМ

ЦЕЛЬ ПРОЕКТА

Автоматизация проверки договоров на предмет их соответствия условиям чек-листов.

РУЧНАЯ ПРОВЕРКА ДОГОВОРОВ



«ЦИФРОВОЙ ЭКСПЕРТ»



ФОРМАЛИЗОВАННЫЕ УСЛОВИЯ ЧЕК-ЛИСТА

[-] Неустойки	<input type="checkbox"/> Уменьшение платы на сумму неустоек
	<input type="checkbox"/> Размер неустойки (штрафа, пени)
[-] НДС	<input type="checkbox"/> Законодательное изменение ставки НДС
[-] Субисполнитель	<input type="checkbox"/> Привлечение субисполнителей
[-] Отказ от исполнения договора	<input type="checkbox"/> Односторонний внесудебный отказ от договора
[-] Политика ИСМ СИБУР	<input type="checkbox"/> Ознакомление с политикой ИСМ и ответственность за ее несоблюдение
[-] Перенос срока оплаты	<input type="checkbox"/> Перенос срока оплаты при непредоставлении оригиналов документов
[-] О не применении ст.823 ГК РФ	<input type="checkbox"/> О не применении ст.823 ГК РФ
	<input type="checkbox"/> Возмещение штрафов контрагентом
[-] Ограничение ответственности СИБУР (10%)	<input type="checkbox"/> Ограничение ответственности СИБУР (10%)
[-] Возмещение имущественных потерь	<input type="checkbox"/> Возмещение имущественных потерь
[-] Доступ контрагента к компьютерной технике СИБУР	<input type="checkbox"/> Доступ контрагента к компьютерной технике СИБУР

ЦИФРОВОЙ ЭКСПЕРТ ПО ДОГОВОРАМ

ПРЕИМУЩЕСТВА СИСТЕМЫ

ДРУЖЕЛЮБНОСТЬ

Специалисты в предметной области легко могут обучиться работе с системой

ИНТЕРПРЕТИРУЕМОСТЬ

Предсказуемость результатов работы системы

АДАПТИРУЕМОСТЬ

Легко исправлять ошибки и вносить изменения в систему

ПРИМЕР ПРАВИЛА, НАПИСАННОГО ЮРИСТОМ

Специалисты в предметной области легко могут обучиться работе с ситемой

```
f(x) Functions ± M Q Check syntax Tools Debug
1  /* О Политике ИСМ и Ключевых правилах безопасности СИБУР*/
2
3  rule: ИСМ_ознакомление
4  {
5      /* Об ознакомлении с Политикой ИСМ и КПВ СИБУР*/
6      query: {sentence(обязан, singleroot(ознакомить),
7
8              персонал or работник,
9
10             phrase(Политика, интегрированной, системы, менеджмента)
11             or
12             phrase(Политика, ИСМ), размещен,
13
14             {entity(Internet Addresses)}:addr)}:m1
15
16     result: Результат = $m1
17           attribute: Интернет_адрес = $addr
18 }
```

ОБРАЗЕЦ ОТЧЕТА

Filter										
Name	Show	Имя документа	Законод...	Односто...	Перенос...	О не при...	Возмеще...	Принятое решение	% соответствия	
<input checked="" type="checkbox"/> Имя документа	<input checked="" type="checkbox"/>	Договор №8	Присутствует	Присутствует	Присутствует	Присутствует	Присутствует	Все условия присутствуют	100	
<input checked="" type="checkbox"/> Законодательное изменение ставки НДС	<input checked="" type="checkbox"/>	Договор №1	Присутствует	Присутствует	Присутствует	Присутствует	Присутствует	Все условия присутствуют	100	
<input checked="" type="checkbox"/> Односторонний внесудебный отказ от договора	<input checked="" type="checkbox"/>	Договор №2	Присутствует	Отсутствует	Присутствует	Присутствует	Присутствует	Требует проверки	70	
<input checked="" type="checkbox"/> Перенос срока оплаты при непредоставлении оригинало...	<input checked="" type="checkbox"/>	Договор №4	Присутствует	Присутствует	Присутствует	Присутствует	Присутствует	Все условия присутствуют	100	
<input checked="" type="checkbox"/> О не применении ст.823 ГК РФ	<input checked="" type="checkbox"/>	Договор №5	Присутствует	Присутствует	Присутствует	Присутствует	Присутствует	Все условия присутствуют	100	
<input checked="" type="checkbox"/> Возмещение имущественных потерь	<input checked="" type="checkbox"/>	Договор №7	Присутствует	Отсутствует	Присутствует	Отсутствует	Присутствует	Требует проверки	50	
<input checked="" type="checkbox"/> Принятое решение	<input checked="" type="checkbox"/>	Договор №6	Присутствует	Отсутствует	Присутствует	Присутствует	Присутствует	Требует проверки	70	
2 % соответствия	<input checked="" type="checkbox"/>	Договор №10	Отсутствует	Отсутствует	Отсутствует	Отсутствует	Отсутствует	Требует проверки	0	
2 #	<input type="checkbox"/>	Договор №11	Отсутствует	Отсутствует	Отсутствует	Отсутствует	Отсутствует	Требует проверки	0	
		Договор №13	Отсутствует	Отсутствует	Отсутствует	Отсутствует	Отсутствует	Требует проверки	0	
		Договор №12	Отсутствует	Отсутствует	Отсутствует	Отсутствует	Отсутствует	Требует проверки	0	
		Договор №14	Отсутствует	Отсутствует	Отсутствует	Отсутствует	Отсутствует	Требует проверки	0	
		Договор №3	Присутствует	Отсутствует	Присутствует	Присутствует	Присутствует	Требует проверки	70	
		Договор №9	Присутствует	Присутствует	Присутствует	Присутствует	Присутствует	Все условия присутствуют	100	

ЭКСПЕРТНЫЕ ПРАВИЛА + МАШИННОЕ ОБУЧЕНИЕ

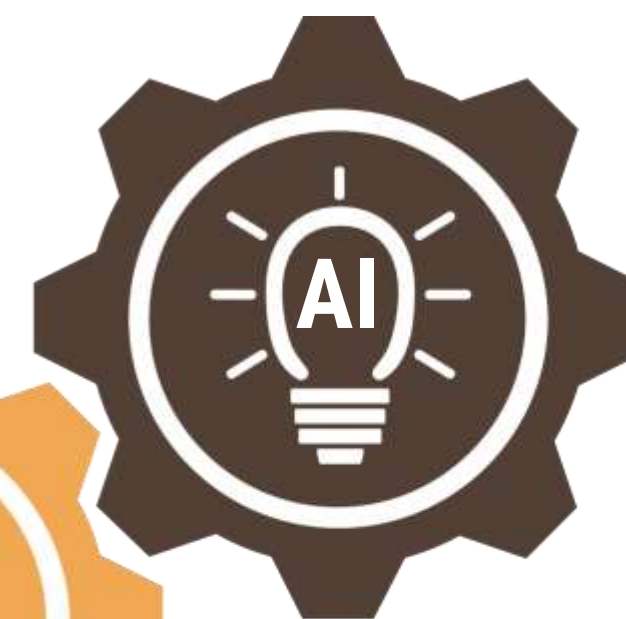
СИСТЕМА ЭКСПЕРТНЫХ ПРАВИЛ



ВАЛИДАЦИЯ РЕЗУЛЬТАТОВ



ОБУЧЕННАЯ МОДЕЛЬ



МАШИННОЕ ОБУЧЕНИЕ



МАШИННОЕ ОБУЧЕНИЕ

Преимущества: высокая полнота

Недостатки: низкая точность

МАШИННОЕ ОБУЧЕНИЕ ДЛЯ ИЗВЛЕЧЕНИЯ ИНФОРМАЦИИ

РЕДАКТОР ПРАВИЛ

для первоначальной разметки

ГРАФИЧЕСКИЙ ИНТЕРФЕЙС

для валидации результатов и повторного обучения

ПОДСВЕТКА

найденной информации в тексте

The screenshot displays the POLYANALYST platform interface. At the top, there are tabs for 'Тексты', 'Сущности', 'Статистика', and 'Настройки'. Below the tabs is a toolbar with icons for 'Выбор колонок', 'Найти', 'Фильтр', 'Настройки', 'Валидация', 'Предыдущий', 'Следующий', 'Настройки', and 'Подсветка'. The main area is divided into three sections:

- Left Panel (Entity List):** Shows a tree view of entities. Under 'Периодичность платежа', there are several epochs: 'Разметка', 'Эпоха 1', 'Эпоха 2', 'Эпоха 3', 'Эпоха 4', and 'Эпоха 5'. 'Эпоха 5' is currently selected.
- Table:** A table with three columns: 'Валидные', 'Невалидные', and 'Результат'. It lists various rules for payment frequency. The row 'Один раз в месяц' is highlighted in blue.
- Right Panel (Text Editor):** Displays a text snippet with a red box highlighting the phrase 'Один раз в месяц'. Below the text, there is a note about document control.

Валидные	Невалидные	Результат
<input checked="" type="checkbox"/>	<input type="checkbox"/>	не позднее числа следующего за отчетным кварталом
<input checked="" type="checkbox"/>	<input type="checkbox"/>	не позднее числа следующего за отчетным месяцем
<input type="checkbox"/>	<input checked="" type="checkbox"/>	один раз в год
<input type="checkbox"/>	<input checked="" type="checkbox"/>	Один раз в месяц
<input checked="" type="checkbox"/>	<input type="checkbox"/>	одного раза в год
<input type="checkbox"/>	<input checked="" type="checkbox"/>	одного раза в месяц
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Оплата производится ежемесячно
<input checked="" type="checkbox"/>	<input type="checkbox"/>	оплата производится ежемесячно
<input checked="" type="checkbox"/>	<input type="checkbox"/>	оплаты производится ежемесячно
<input checked="" type="checkbox"/>	<input type="checkbox"/>	оплаты, ежемесячно
<input type="checkbox"/>	<input type="checkbox"/>	плата оплачивается ежемесячно
<input type="checkbox"/>	<input type="checkbox"/>	плата уплачивается ежемесячно

Соблюдение режима вывоза (выноса), ввоза (вноса) материальных ценностей, возлагается на материально ответственных лиц, главного бухгалтера и руководителей филиала, имеющих право подписи в накладной, а контроль за

АВТОМАТИЗАЦИЯ ЭКСПЕРТИЗЫ ДОГОВОРОВ КОМПАНИИ СИБУР

PolyAnalyst 6.5

ПРОБЛЕМЫ РУЧНОЙ ЭКСПЕРТИЗЫ

Масса рутинной работы

юристы тратят дорогостоящее рабочее время на проведение конвейерных операций

72% стандартной формы



90 000
ДОГОВОРОВ

нестандартной формы **28%**

Нестандартные проверяются на соответствие чек-листам

ГИПОТЕЗЫ РЕШЕНИЯ



1 Применяя семантический и лингвистический анализ, система забирает на себя максимальный объем работы по выявлению документах заложенных юристами сущностей и фактов.

2 Должна поддерживаться возможность написания правил, определяющих проверку договоров на соответствие их содержания чек-листам.

3 Одновременно, за одну часовую итерацию должны проверяться более 500 договоров, всех общеприменяемых форматов (pdf, word, excel).

4 Система обеспечивает доступность и работоспособность 24/7.

РАЗРАБОТАННОЕ РЕШЕНИЕ



ЭФФЕКТ КОМПЕТЕНЦИЙ



Сотрудники компании получают навыки PolyAnalyst и разрабатывают собственное решение, автоматизирующее экспертизу договоров. Процесс экспертизы ускоряется, а у юристов высвобождается от рутины дорогостоящее рабочее время, теперь занимаемое более сложными правовыми вопросами, требуемыми обязательного человеческого участия, таких как разрешения разногласий с контрагентами и участия в судах.

ПЛАТФОРМА POLYANALYST



Команда развивает свое решение, глубже изучает PolyAnalyst, становясь **центром компетенций** своей компании в работе с данными. Система за счет **только собственных человеческих ресурсов** компании теперь способна к проведению анализа широкого множества разносторонней документации, будь то коммерческие договоры, акты, счета, технико-экономические обоснования.



Вы разрабатываете новые решения на PolyAnalyst для коллег из соседних подразделений, Автоматизация ведет к **оптимизированию затрат** на персонал и **повышению эффективности** его работы, и как следствие **росту** показателя **EBITDA** всей компании. Не привлекая к разработке подрядчиков, вы практически **сводите к нулю добавленную стоимость** производства программных систем, и сокращаете расходы организации.



Все продукты обточены и их работоспособность подтверждена **успешной эксплуатацией** внутри структур компании. Решения могут быть проданы, как собственная разработка на базе платформы PolyAnalyst, а имеющаяся экспертиза производит на заказ системы под индивидуальные нужды и специфику клиентов. В результате компания диверсифицирует продуктовый портфель, и получает дополнительный приток выручки.

МЕГАПЬЮТЕР ИНТЕЛЛИДЖЕНС

ЗАРАБАТЫВАЙТЕ НА POLYANALYST

MEGAPUTER РЕАЛИЗУЕТ ПАРТНЕРСКУЮ ПРОГРАММУ **СОВМЕСТНЫХ ПРОДАЖ** ПО МОДЕЛИ **REVENUE SHARING**



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

 ПРАВО RU

 СВЯЗНОЙ


 Детский мир



 РАНХиГС
РОССИЙСКАЯ АКАДЕМИЯ НАРОДНОГО ХОЗЯЙСТВА
И ГОСУДАРСТВЕННОЙ СЛУЖБЫ
ПРИ ПРЕЗИДЕНТЕ РОССИЙСКОЙ ФЕДЕРАЦИИ



 НИИ ВОСХОД

 Счетная палата
Российской Федерации

 СБЕР

 КОНСИСТ-ОС
РОСАТОМ

 МИНПРОМТОРГ
РОССИИ

 фрчч

 СИБУР

 АНАЛИТИЧЕСКИЙ ЦЕНТР
ПРИ ПРАВИТЕЛЬСТВЕ
РОССИЙСКОЙ ФЕДЕРАЦИИ

 20.35
УНИВЕРСИТЕТ

 OCS
DISTRIBUTION

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.ru



Гольцов Дмитрий

Руководитель коммерческого
направления
+7 (916) 111-95-49
goltsov@megaputer.ru

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