

БИШКЕК

КЫРГЫЗСКАЯ РЕСПУБЛИКА

21-23 мая 2025



РЕГИОН СНГ

НА ПУТИ К ДОСТИЖЕНИЮ



ЦУР

Оценка перспектив применения искусственного интеллекта в статистической практике.

Первые результаты

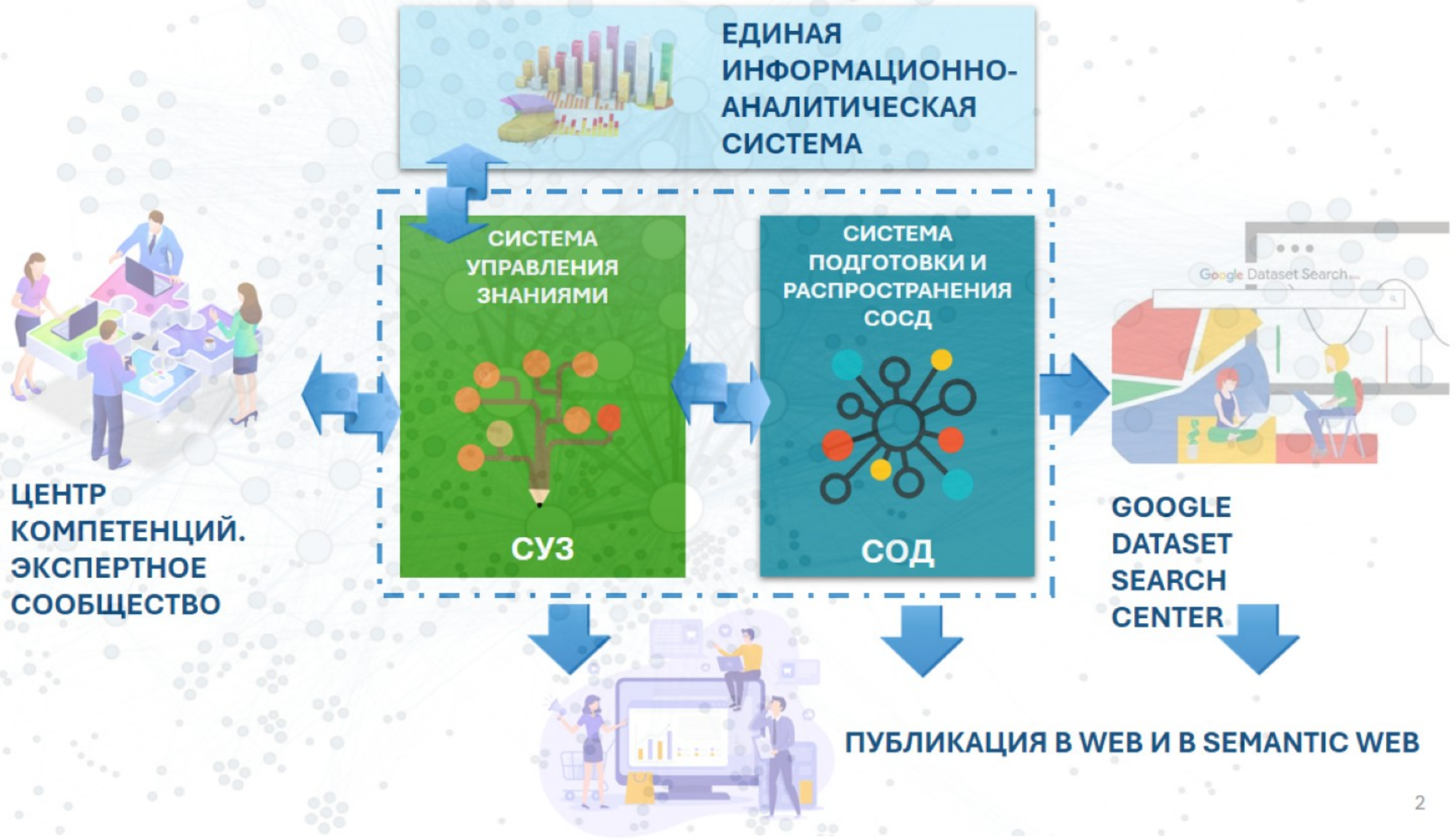
Юрий Акаткин
Научный
руководитель



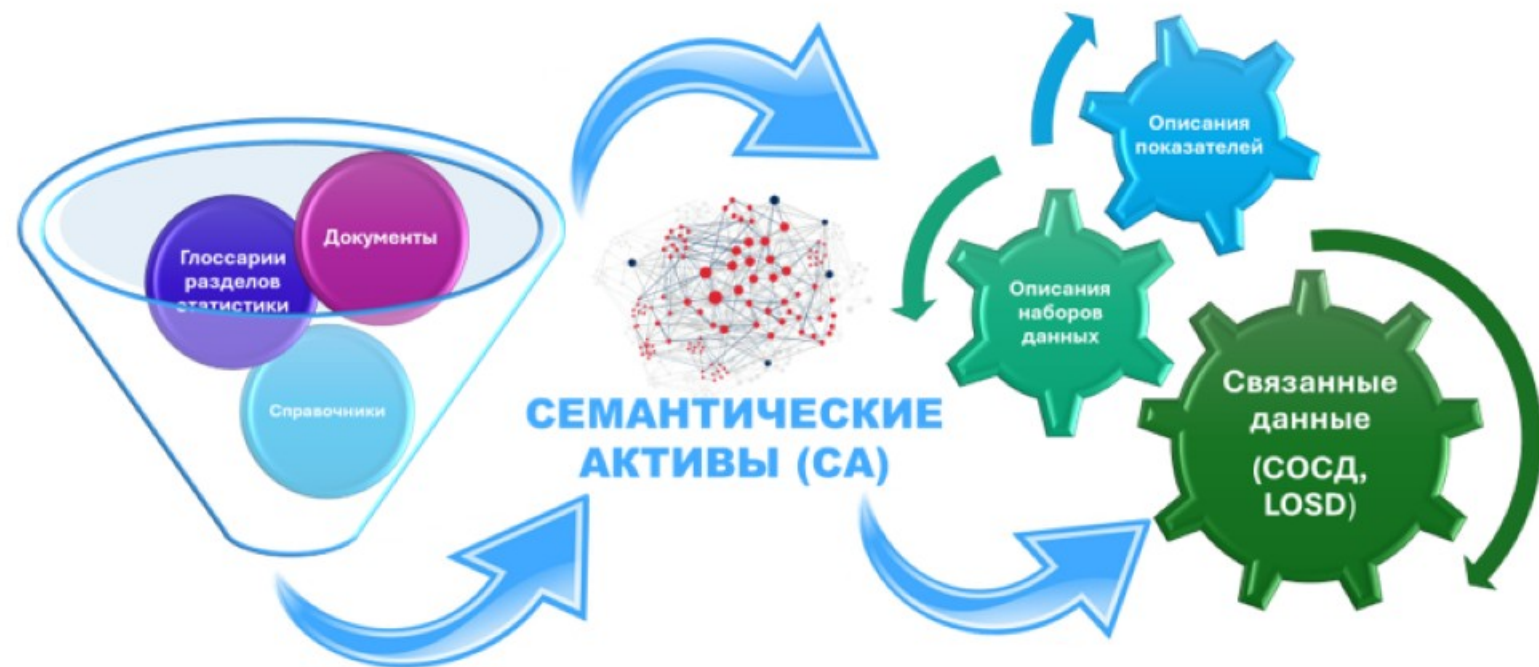
Электронное
Проектирование

Елена Ясиновская
Главный
аналитик

СИСТЕМА УПРАВЛЕНИЯ ЗНАНИЯМИ И СОД В DATAХАБЕ СТАТКОМИТЕТА СНГ



ФОРМИРОВАНИЕ И РАСПРОСТРАНЕНИЕ СОСД И «УМНЫХ» МЕТАДАННЫХ



**СОСД и СА
публикуются в
открытых каталогах
подсистемы СОД**

**СОСД формируются по технологии
Semantic Web в виде наборов данных,
содержащих расширенный состав
метаданных**

ИНТЕРПРЕТАЦИЯ СОСД

СУЗ обеспечивает интерпретацию СОСД для людей, а семантические активы (СА) – для информационных систем

**СЕМАНТИЧЕСКИ
БОГАТАЯ СРЕДА
ИНТЕРПРЕТАЦИИ**



УЛУЧШЕНИЕ КАЧЕСТВА
СТАТИСТИЧЕСКИХ ДАННЫХ
И МЕТАДАННЫХ

ГАРМОНИЗАЦИЯ
СТАТИСТИЧЕСКОЙ
ТЕРМИНОЛОГИИ И
КЛАССИФИКАЦИИ

СООТВЕТСТВИЕ ПРИНЦИПАМ
FAIR

ОБЕСПЕЧЕНИЕ
СЕМАНТИЧЕСКОЙ
ИНТЕРОПЕРАБЕЛЬНОСТИ

ОБЕСПЕЧЕНИЕ ОДНОЗНАЧНОЙ
ИНТЕПРЕТАЦИИ
(МЕТА)ДАННЫХ



СЕМАНТИЧЕСКИЕ
АКТИВЫ



УМНЫЕ МЕТАДАННЫЕ



ОБОГАЩЕННЫЕ
СЕМАНТИКОЙ СОСД



ВИЗУАЛИЗАЦИЯ ДЛЯ
РАБОТЫ ЭКСПЕРТОВ

LLM – МОЩНЫЙ ИНСТРУМЕНТ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

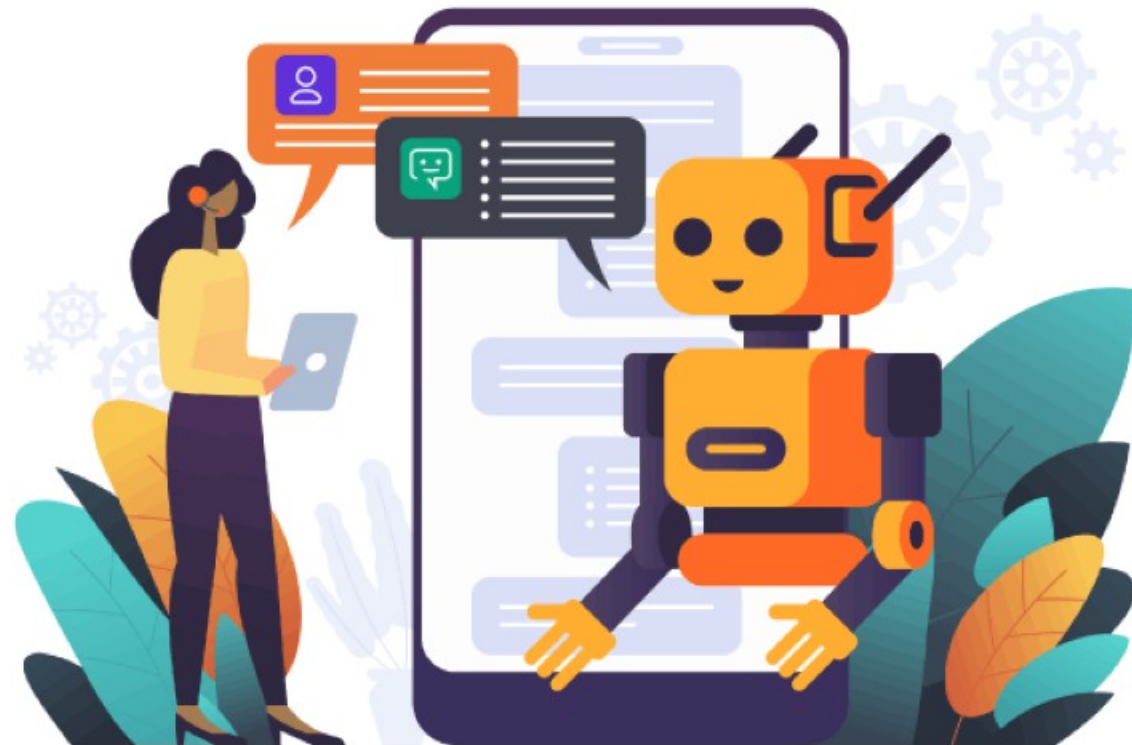
LARGE LANGUAGE MODELS

БОЛЬШИЕ ЯЗЫКОВЫЕ МОДЕЛИ

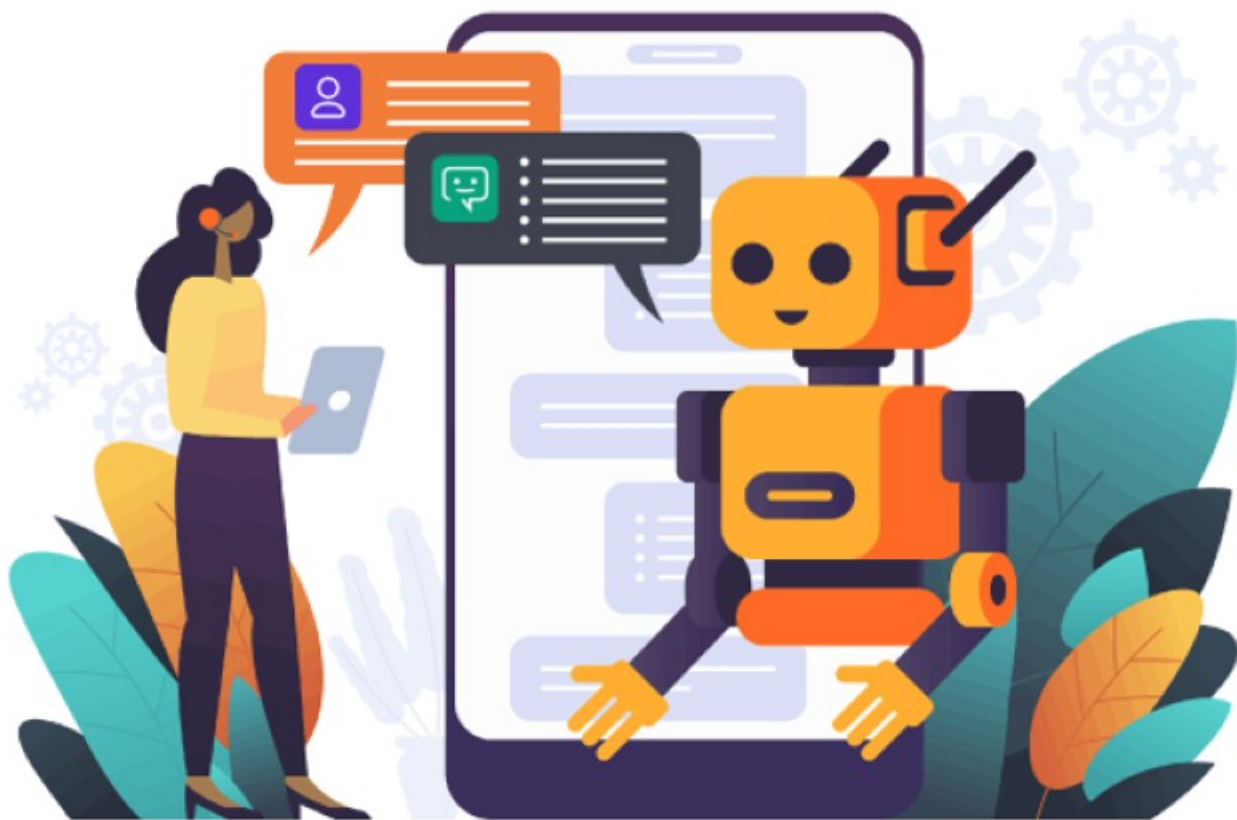
нейросетевые модели, обученные на огромных массивах информации для обработки и генерации текстов на естественном языке

Они способны:

- отвечать на вопросы в заданном контексте
- обобщать, структурировать и анализировать информацию
- генерировать тексты, таблицы и файлы



СЦЕНАРИИ ПРИМЕНЕНИЯ ИИ



ДЛЯ ЭКСПЕРТОВ СУЗ

- Формирование аннотаций документов
- Работа с терминами глоссариев
- Подбор международных эквивалентов

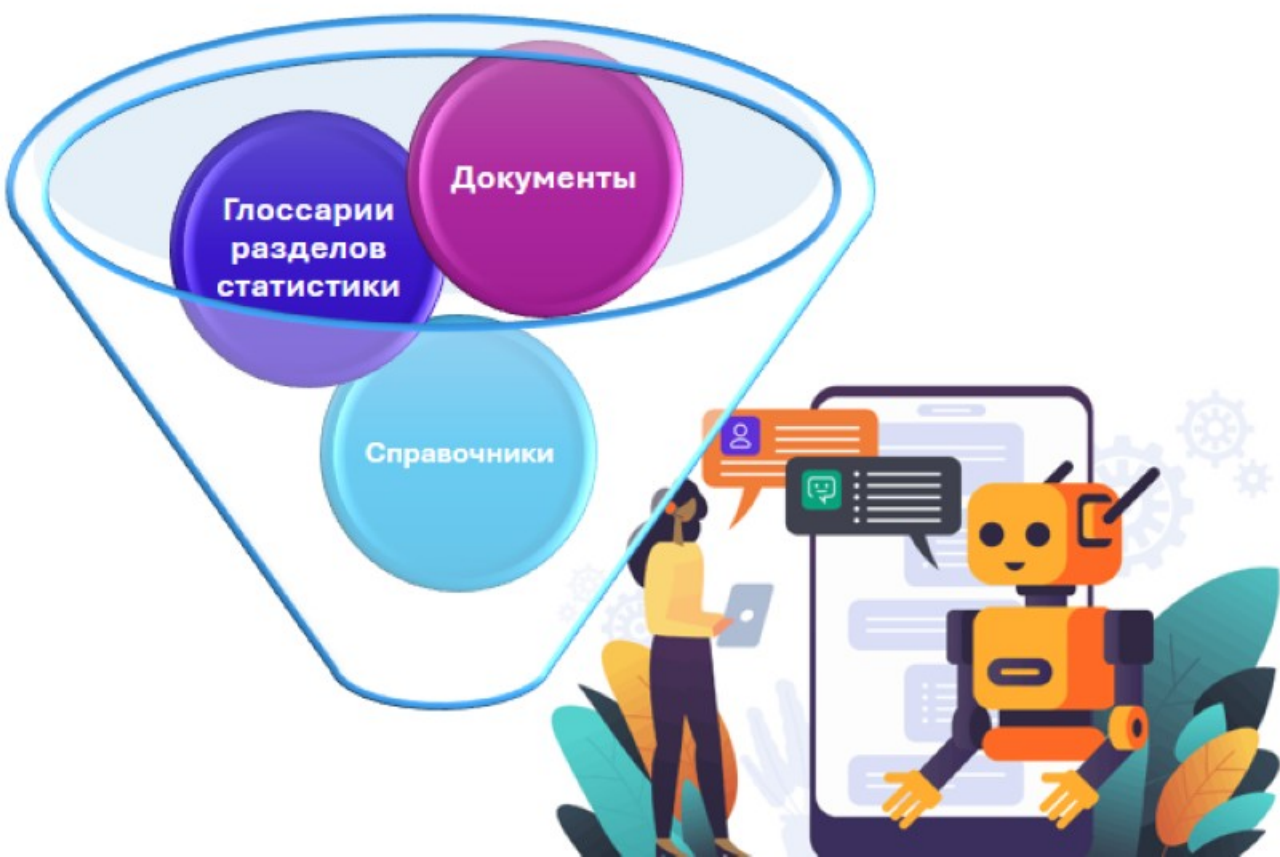
ДЛЯ СТАТИСТИКОВ

- Подготовка методических рекомендаций вопросников в соответствии с международными стандартами
- Актуализация методических рекомендаций и форм сбора
- Контроль соответствия форм сбора

ПРИМЕНЕНИЕ ИИ ДЛЯ РАБОТЫ ЭКСПЕРТОВ

РЕЗУЛЬТАТ

Упрощение и ускорение работы экспертов по наполнению СУЗ, встроенный контроль качества



БИБЛИОТЕКА МАТЕРИАЛОВ

- **Подготовка аннотаций документов**
 - Генерация текста на основе анализа документа с использованием терминов глоссариев СУЗ
 - Проверка аннотации без привлечения редактора и ее итеративная доработка
- **Формирование перечня терминов-кандидатов**

ГЛОССАРИИ

- **Подготовка текстов определений и пояснений**
 - Генерация текстов на основе анализа документов СУЗ и внешних источников
 - Проверка и уточнение определений и пояснений без привлечения редактора и их итеративная доработка
 - Выявление терминов, требующих гармонизации
- **Подбор международных терминов**
 - Международные наименования терминов, в том числе для формирования URI
 - Определения из международных источников для исключения ошибок прямого перевода

ФОРМИРОВАНИЕ ВОПРОСНИКОВ В СУЗ В СООТВЕТСТВИИ С SIMS



ВОПРОСНИКИ

Единообразная структура и быстрая актуализация Вопросников с автоматической генерацией ESMS-отчетов включает:

- **СТРУКТУРИРОВАННЫЕ МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ/ИНСТРУКЦИИ**
специализированные шаблоны СУЗ, соответствующие блокам ESMS, которые заполняются в автоматизированном режиме при взаимодействии статистика с ИИ
- **ФОРМЫ СБОРА ДАННЫХ**
загруженные файлы различных форматов, проверка которых будет осуществляться ИИ

ПРИМЕНЕНИЕ ИИ ДЛЯ ПОДГОТОВКИ ВОПРОСНИКОВ

РЕЗУЛЬТАТ

Автоматизация формирования
Вопросников, встроенный
контроль качества и актуальности,
генерация ESMS-отчетов



МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ

- **Заполнение блоков шаблона**
 - Генерация текстов на основе анализа материалов СУЗ (например, структурированных документов) и внешних источников в соответствии с рекомендациями SIMS
- **Валидация и проверка актуальности**
 - Проверка текстов и получение рекомендаций по улучшению, предложения по обновлению на основе новых данных, терминов, документов

ФОРМЫ СБОРА

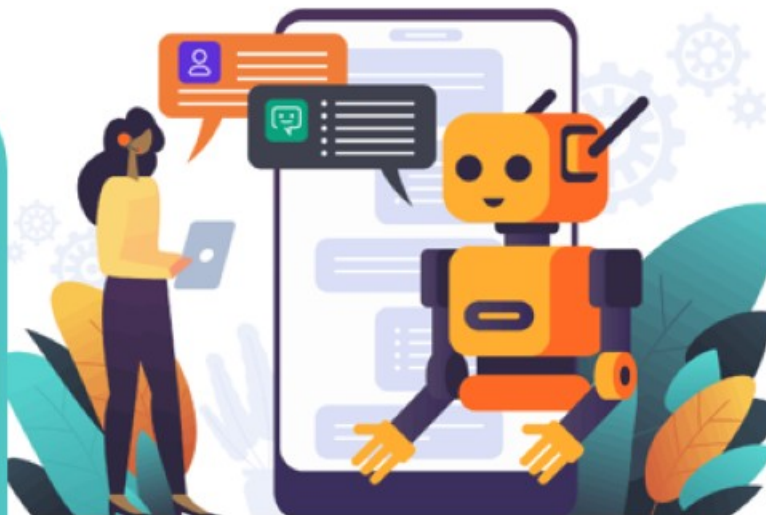
- **Проверка качества форм**
 - Анализ файлов различных форматов, содержащих формы, проверка соответствия методическим рекомендациям/инструкциям
 - Генерация предложений по исправлению форм (например, правильность указанных изменений или периодичность сбора)

ПЕРВЫЕ РЕЗУЛЬТАТЫ. АНАЛИЗ И ИССЛЕДОВАНИЕ

**БАЗОВЫЕ
СЦЕНАРИИ
ПРИМЕНЕНИЯ ИИ**

**ПЕРВИЧНЫЕ
ИССЛЕДОВАНИЯ
LLM**

**ЭКСПЕРИМЕНТЫ
ПО ИНТЕГРАЦИИ
ИИ В ПО СУЗ**



**АНАЛИЗ И ВНЕДРЕНИЕ
МЕЖДУНАРОДНЫХ
СТАНДАРТОВ SIMS/ESMS**

**АДАПТАЦИЯ LLM ДЛЯ
РЕШЕНИЯ
ПРИКЛАДНЫХ ЗАДАЧ**

ВЫБОР LLM

+

○



КРИТЕРИИ ВЫБОРА LLM

- Независимость от внешних факторов
 - Контроль изменений LLM
 - Защита информации
- Разумные требования к оборудованию
- Разумные трудозатраты для тонкой настройки (дообучения) и интеграции в СУЗ-СОД

ОЦЕНКА LLM



КРИТЕРИИ ОЦЕНКИ

- Качество ответов на общетеоретические вопросы (уровень – бакалавр)
- Правильность решения задач (уровень – бакалавр)
- Правильность терминологии
- Возможность анализировать и аннотировать документы в области статистики
- Технологические аспекты (требования к оборудованию, интеграция с СУЗ-СОД)

МЕТОДЫ ОЦЕНКИ



ОЦЕНИВАЕМЫЕ МОДЕЛИ
(~12-17B, ~24-27B)



Контрольные
вопросы: учебник,
практикум,
«старшая» LLM

Ответы на вопросы
и решение задач

Оценка с помощью:
«старшая» LLM,
эксперт

ПЕРВЫЕ РЕЗУЛЬТАТЫ



ХОРОШИЕ РЕЗУЛЬТАТЫ -
ОЦЕНКА ~4 БАЛЛА ИЗ 5

GEMMI3:27B
LLAMA4:17B



Контрольные
вопросы: учебник,
практикум,
«старшая» LLM

Ответы на вопросы
и решение задач

Оценка с помощью:
«старшая» LLM,
эксперт

СОТРУДНИЧЕСТВО И ВЗАИМОДЕЙСТВИЕ

ВОПРОСЫ?

**КОНТАКТЫ ДЛЯ СВЯЗИ
С РАЗРАБОТЧИКАМИ**

elena@semanticpro.org

www.e-projecting.ru