***АИС Деймос является модульной кроссплатформенной системой предиктивной аналитики распределенных открытых источников социальной цифровой среды на основе технологий многопоточной обработки данных.***

Система основана на агрегации данных с целью внедрения цифровых двойников объектов одного из видов открытых источников – социальной сети ВКонтакте. В системе заложена возможность последующего формирования различных структур цифрового двойника этой социальной сети, разработанное решение также позволяет повысить важность анализа квалификаций всех открытых источников. Процедуры системы сводятся к автоматизированному сбору и систематизации информации из открытых источников с помощью скриптов - парсингу в системе предиктивной аналитики.

Цель системы сводится к обеспечению безопасности интернет-пространства через мониторинг связей контента и определение основных объектов открытых источников и представляет собой совокупность локальных систем с подсистемами обратной связи предиктивной аналитики подтипа СППР открытых источников социальной сети с доступной интерпретацией в локальных системах различного назначения.

Подготовка данных для формулирования кроссплаформенных цифровых двойников объектов социальной сети специализируется на ежедневной многопоточной обработке метаданных, для чего предусмотрены модули, реализующие алгоритмы, выбираемые пользователем для данных различного типа. Алгоритмы позволяют привести разрозненное содержание получаемых данных к точным копиям открытых источников, используя свойства эмерджентности и сингуляризации системы.

Технологии управления данными и знаниями в системе созданы с учетом семантики выбираемых пользователем сетов с наборами признаков определенной области исследований на основе:

1) формального описания объектов, их взаимодействия с использованием алгоритмизированных решений различных математических аппаратов, формальных языков и систем, а также проблемной области с учетом жизненного цикла объекта и процессов управления данными и знаниями с использованием инженерии знаний;

2) моделей и методов анализа больших объемов данных, машинного обучения и искусственного интеллекта;

3) интеграции полученных результатов для обеспечения интеллектуального информационного сопровождения в различных исследуемых областях;

4) многоуровневой архитектуры вычислений на основе алгоритмов, комплексная реализация которых делает возможной подготовку данных для дальнейшего анализа структуры связей между объектами исследуемой социальной сети. Исполнение результатов работы системы оценивается на основе свойств достаточности информационного сопровождения в отношении поставленной цели, обеспечивая также такие качества как целостность и доступность данных, а также надежность источника и законные основании их получения.

В качестве прикладных задач, которые решаются в системе, выделяются следующие:

1. Разработка концептуальной модели формализованного наполнения цифрового двойника, которая является отличительной особенностью системы и включает в себя совокупность процессов управления данными и знаниями, интегрированных как как комплекс:

a) структурированных хранилищ данных и знаний

b) формальной модели и методов реализации процессов взаимодействия компонентов информационных объектов

c) системы сбора разнородных данных с процедурами их предварительного категорирования и типологизации по общим особенностям разграничения получаемых сведений

d) формализованной модели на основе процедурно-технологического метода построения и исследования сложных систем

e) подхода к нормализации и интеграции данных для обеспечения информативности

2. Сбор и пополнение коллекций данных различных структур, выявление их особенностей, их классификация, систематизация по предметно-тематическому и форматному принципам, проведение предварительной проверки и анализа данных. Создание архивов для хранения информации.

3. Интерпретация задач информационного сопровождения при построении цифровых двойников в аспекте множества задач для анализа данных и принятия решений с учетом семантики проблемой области.

4. Формирование комплекса моделей и методов для последующего анализа данных, формализации выявленных знаний и получение решения.

5. Возможность интерпретации результатов по управлению данными и знаниями в существующих разработанных сетах для технических классов систем различного назначения с соблюдением свойств кроссплатформенности.

**Основные функциональные характеристики / возможности разрабатываемого ИТ-решения**: Анализ в режиме реального времени потока данных, идентификация цифровых следов объектов исследования, выявление структуры их связей

**Сквозная цифровая технология:**

Перспективные методы и технологии в ИИ

**Приоритетные классы программного обеспечения**:

Средства обработки и визуализации массивов данных

*Функциональные возможности, которые обеспечиваются разработанной системой с учетом требований аналитичности функциональных возможностей:*

1. Вывод массива данных открытых источников в матричном представлении и выделение их эмерджентных свойств и зависимостей, систематизацию и формализацию потоковых данных социальной сети ВКонтакте.
2. Категоризация и формализация наборов данных в рекурсивных вложенностях системы, возможность их качественного и количественного анализа.
3. Вывод исходных данных об объекте с возможностью интерпретации полученных результатов с обеспечением свойств достаточности получаемой пользователем информации в доступном формате для специалистов области разработки.
4. Визуализация метрических данных статистики объекта в пространстве.

### Количественные параметры, определяющие выполнение продуктом своих функций:

**Характеристики программного комплекса:**

* скорость обработки данных (1 мб/с);
* ошибки отклонения (не более 5%);
* способы представления данных (не менее чем в 3 вариациях);
* количество обрабатываемых записей (не менее 100 тыс. объектов);
* размерность облака слов (не менее 10 тыс. записей);
* итерационные вложенности (не менее 3 доступных уровней).

### Входные воздействия, необходимые для выполнения научно-техническим продуктом заданных функций:

Массивы данных открытых источников социальной сети ВКонтакте

### Выходные реакции, обеспечиваемые продуктом в результате выполнения своих функций:

Визуализированные данные в оболочке системы в формате электронных таблиц и их обобщенного пространственного представления

**Профессиональные протоколы и схемы**

Технологические схемы и протоколы процедур сбора системы созданы для полноценной интерпретации данных об объекте исследования и разграничены правилами для возможного сопряженного анализа. Технологии синтеза данных и входящими в него процедурами их автоматизированного сбора, а также последующего базового анализа, описываемого в этом документе, сочетаются с технологиями анализа более высокого порядка, включающем проекции цифровых двойников социальной сети. Обеспечивающие эти цели группы алгоритмов подразделяются на технологические и управляющие. Созданные решения являются универсальными для любых типов клиентских запросов и отвечают требованиям достаточности получаемых от продукта сведений.

Выбор пользователем схем и протоколов является основой для сопутствующих процедур сингуляризации потоков данных. В зависимости от запроса пользователя, формулируются их допустимые варианты, ограниченные правилами, не допускающими обработку данных вне установленного формата, а также с целью исключения образования информационных лопостей в системе и увеличения количества нечетких данных. Схемы предполагают несколько общепринятых в профессиональной среде типов автоматизированного получения данных открытых источников социальной сети, являющихся показателем технологической оснащенности системы – «быстрый парсинг», «средний парсинг» и «глубокий парсинг», в зависимости от задачи пользователя. Также, обработка данных в системе предполагает возможность полноценного вычисления основных характеристик об объекте, - скорости, глубины, частоты, веса и др., что также является показателем эффективности произведенных в системе процедур.

Функциональные возможности технологических протоколов в алгоритмизированных решениях системы:

Технологии пересеченного поиска различных видов и типов данных объектов

Технологии встроенных профессиональных схем поиска, с возможностью адаптации под каждый запрос

Технологии поиска связей любых исследуемых объектов до глубины не менее трех итерационных вложенностей

Технологии наращивания контента поисковых значений об объектах в автоматическом и полуавтоматическом режимах

Описанные возможности адаптированы под управляющие вычисления, в том числе с группированием таких показателей как множество типов одной категории признаков, количество признаков, альтернатив изменения объекта, отношений признаков (межтиповые отношения), состояний каркаса вычислений, правил изменения его состояний, тактов моделирования его структуры и разницу исходных данных и других.