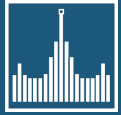


Высшая школа государственного администрирования



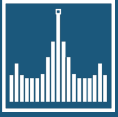
Научная конференция «Ломоносов-2024»
Секция «Авангард цифровой трансформации государственного
администрирования: стратегии, технологии, эффективность»

Подсекция «Слабый и сильный искусственный интеллект в управленческих
практиках»

Массивы данных государственных органов как основа
национальной платформы искусственного интеллекта

Студент 2 курса:
Урмантаев Канат Болатович

Научный руководитель:
Назаренко Сергей Владимирович,
кандидат социологических наук, доцент



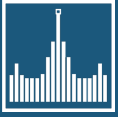
Массивы данных

Основой развития искусственного интеллекта в любом его виде являются массивы (структурированных/неструктурированных) данных. Они используются практически везде, на любых платформах, в любых устройствах и в любых средах разработки. Современный компьютерный мир невозможен без массивов, настолько повсеместно и широко они распространены.

К примеру, массив данных для ChatGPT OpenAI включает в себя миллионы различных текстовых фрагментов, собранных из различных источников интернета. Эти данные представляют собой разнообразные диалоги, текстовые сообщения, статьи, запросы и ответы пользователей, комментарии из социальных сетей, форумов, новостных сайтов и других онлайн-ресурсов.

Благодаря тому, что массив данных для ChatGPT OpenAI является обширным и многообразным, позволяет модели получать широкий контекст и разнообразие информации при генерации ответов на запросы пользователей. Эти данные помогают боту понимать естественный язык, контекст общения, вопросы пользователя и предоставлять соответствующие и информативные ответы.





Массивы данных государственных органов

В свою очередь массивы данных государственных органов играют важную роль в развитии «национального искусственного интеллекта», являются его основой. Государственные органы обладают доступом к разнообразным и объемным данным, таким как например, информация о населении, экономике, здравоохранении, образовании и многом другом.



Эти данные могут быть использованы для обучения искусственной интеллектуальной системы, что позволит «прокачать» ее способности в различных областях, таких как анализ данных, прогнозирование трендов, оптимизация процессов. Это в свою очередь повысит эффективность работы государственных органов, обеспечит открытость и прозрачность их деятельности.



Безопасность, качество данных на этапе внедрения и обработки массивов информации

Для эффективного и безопасного внедрения массивов данных государственных органов на серверы искусственного интеллекта необходимо учитывать несколько важных аспектов:

1. Конфиденциальность данных, так как государственные данные содержат достаточно конфиденциальной информации о гражданах, юридических лицах и т.д. Тут является важным обеспечение высокого уровня защиты данных при их передаче и хранении на серверах искусственного интеллекта. Использование шифрования, механизмов аутентификации и других методов защиты поможет предотвратить несанкционированный доступ к информации. Еще одним эффективным способом защиты информации является ее обезличивание.
2. Качество и актуальность данных также играет очень важную роль. Так, перед тем, как внедрять массивы данных, необходимо убедиться в их качестве и актуальности. Некорректные или устаревшие из них могут привести к неверным выводам и решениям, поэтому крайне важно провести тщательный анализ и очистку данных перед использованием.
3. Регулярный мониторинг и аудит данных обеспечит безопасность и эффективность их использования. Требуется устанавливать системы мониторинга и аудита, которые позволят отслеживать доступ к информации, изменения данных и другие важные параметры, и соответственно вовремя на них реагировать.





Обезличивание исходных данных

Технология обезличивания данных играет ключевую роль в защите персональных данных при их использовании для обучения моделей искусственного интеллекта.

Применение этой технологии важно на этапах:

- ✓ сбора данных необходимо обезличивать персональные данные, удаляя или заменяя идентифицирующие элементы (например, имена, адреса, номера телефонов). Это позволит сохранить конфиденциальность информации и предотвратить возможность идентификации отдельных лиц.
- ✓ обезличенные данные должны храниться в безопасном и защищенном виде, чтобы исключить риск утечки конфиденциальной информации. Использование шифрования и других методов защиты данных поможет обеспечить их безопасность.
- ✓ при обработке данных для обучения моделей искусственного интеллекта крайне важно использовать обезличенные данные. Это поможет предотвратить возможность восстановления персональных данных и в полной мере соблюсти принципы конфиденциальности.

В целом, технология обезличивания данных является важным инструментом для защиты персональных данных при работе с ними в области искусственного интеллекта. Внедрение этой технологии на различных этапах обработки информации поможет обеспечить безопасность и конфиденциальность данных.





Основные направления развития

На сегодняшний день внедрение технологий искусственного интеллекта наиболее актуально в таких значимых для экономики отраслях, как нефтегазовый сектор, энергетика, сельское хозяйство, транспорт и логистика. Бесспорно, использование искусственного интеллекта позволит в значительной мере оптимизировать процессы в данных направлениях.





Smart city и искусственный интеллект

Трансформации городов в полноценные smart city, благоприятные для проживания и работы являются неотъемлемым элементом развития всех крупных мегаполисов мира. Наиболее актуальными направлениями развития являются социально значимые, а также касающиеся общественной безопасности.

Взять к примеру управление пропускной способности дорожного движения в городе, которое может не только сообщать о нарушениях правил дорожного движения, но и в случаях возникновения чрезвычайного происшествия, обеспечивать беспрепятственный проезд аварийных служб и карет скорой помощи.





Интеграция баз данных государственных органов

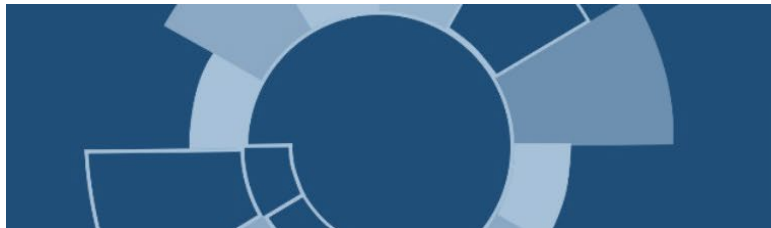
Smart Data Ukimet (SDU) – информационно-аналитическая система, предназначенная для сбора, хранения и комплексного анализа данных, накапливаемых в информационных системах государственных органов Казахстана.

В настоящий момент с ИАС «SDU» интегрировано 86 баз данных информационных систем государственных органов Республики Казахстан, которые ежедневно «стекаются» в «Озеро данных» для создания систем принятия решений на основе больших данных.

Министерством цифрового развития Республики Казахстан реализовано более 30 статистических дашбордов (графических интерфейсов), направленных на обеспечение Правительства объективными аналитическими данными с целью принятия управленческих решений в соответствии с концепцией Data Driven Government.

Очень важным является, что все персональные и чувствительные данные граждан и бизнеса проходят процедуру шифрования и в виде деперсонализированного массива информации попадают в озеро данных.

Наиболее значимым кейсом на платформе SDU является «Цифровая карта семьи», предназначенная для мониторинга социального благосостояния семей в масштабе всей страны, в целях выявления наиболее социально уязвимых слоев населения и таргетирования на них мер господдержки.



Smart Data Ukimet

ИНФОРМАЦИОННО-АНАЛИТИЧЕСКАЯ СИСТЕМА



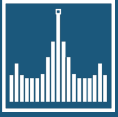
Искусственный интеллект в сфере оказания государственных услуг

Большое значение также имеет широкое использование возможностей искусственного интеллекта при оказании государственных услуг.

На примере Казахстана – это реализация проекта «Цифровая карта семьи». Данный проект оптимизирует усилия государства для оценки уязвимости семей без участия самих граждан. Таким образом, реализация мер поддержки осуществляется в автоматическом режиме, даже в случае отсутствия у граждан доступа к интернет-платформам государственных служб.

В рамках Всемирного Правительственного саммита в Дубае, инновационный проект Казахстана был признан лучшим в категории «Инклюзивная цифровая трансформация», получив престижную награду «GovTech Prize 2024».





Заключение

Исходя из перечисленного можно сделать следующие выводы:

- ✓ Создание единого пространства больших данных (bigdata), является основой для применения искусственного интеллекта.
- ✓ Единое хранилище данных необходимо для автоматизации процессов сбора и хранения информации, предоставляемой государственными органами, коммерческими организациями и академическими учреждениями. В дальнейшем это позволит эффективно обмениваться этими данными для дальнейшего анализа, прогнозирования или принятия важных решений.
- ✓ Качество проведенной обработки данных влияют на результаты дальнейшего мониторинг состояния и управление путем прогнозирования развития различных ситуаций.
- ✓ Важным преимуществом данного подхода является возможность моделирование последствий управленческих решений, на базе использования информационно-аналитического модуля.

Несомненно, информационно-аналитическая система существенно улучшает возможности государственных органов в управлении данными, что в свою очередь способствует более эффективному принятию решений и более точному анализу текущей ситуации.

При этом не менее важными остаются вопросы безопасности данных, их защиты, что в свою очередь делает обязательным качественное законодательное регулирование указанных систем, которое может защитить общественность от потенциальных угроз, связанных с использованием этих технологий.

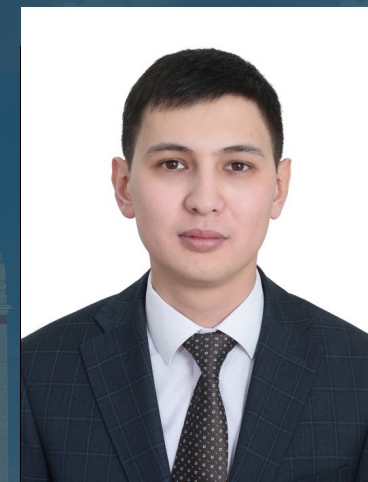




**Научная конференция «Ломоносов-2024»
Секция «Авангард цифровой трансформации
государственного администрирования: стратегии,
технологии, эффективность»
Подсекция «Слабый и сильный искусственный
интеллект в управленческих практиках»**

**Массивы данных государственных
органов как основа национальной
платформы искусственного интеллекта**

**Научный руководитель:
Назаренко Сергей Владимирович,
кандидат социологических наук, доцент**



**Студент 2 курса:
Урмантаев Канат
Болатович**