

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И ЦИФРОВАЯ ТРАНСФОРМАЦИЯ ИСТОРИЧЕСКОГО ИСТОЧНИКА

Опыт работы со стенограммой по проблемам этногенеза кыргызов

Нурзада Гарибовна Жантелиева

кандидат исторических наук, доцент

Кыргызский национальный университет им. Ж. Баласагына

1. Актуальность исследования

- Историческая наука входит в этап цифровой трансформации.
- Архивные документы становятся объектами цифровой обработки и анализа.
- Большие массивы текстов требуют новых инструментов поиска, систематизации и сопоставления.
- Искусственный интеллект расширяет технические возможности историка.
- При этом научная интерпретация и источниковедческая критика остаются ответственностью исследователя.

2. Объект исследования

Стенограмма научной сессии 1956 года, посвящённой проблемам происхождения и этногенеза кыргызского народа.

Особенность источника:

- коллективный характер;
- многоголосая научная дискуссия;
- участие специалистов разных дисциплин;
- фиксация вопросов, ответов, аргументов и полемики;
- отражение состояния исторической науки середины XX века.

3. Почему стенограмма — особый исторический источник?

Стенограмма отличается от монографии или статьи.

Она фиксирует не только итоговое положение автора, но и процесс научного обсуждения.

В ней можно изучать:

- научные позиции;
- аргументацию;
- полемику;
- терминологию;
- структуру научного сообщества;
- взаимодействие различных дисциплин.

4. Научная сессия 1956 года

В материалах подчёркивается участие специалистов из Москвы и Ленинграда, работавших в связи с исследовательскими планами Института истории АН Киргизской ССР.

Сессия была ориентирована на дальнейшую разработку истории кыргызского народа и проблемы его происхождения.

5. Междисциплинарность источника

Проблема этногенеза рассматривалась с привлечением разных групп знаний:

ИСТОРИЯ → АРХЕОЛОГИЯ → ЭТНОГРАФИЯ → АНТРОПОЛОГИЯ → ЛИНГВИСТИКА

Такой комплексный характер обсуждения делает стенограмму особенно перспективной для цифрового анализа.

6. А. Н. Бернштам в стенограмме

Важное место занимает выступление А. Н. Бернштама «Сложение тюрко-язычного населения Средней Азии и происхождение киргизского народа».

Бернштам подчёркивал необходимость рассматривать происхождение кыргызского народа в контексте истории тюркоязычного населения Средней Азии.

7. Ключевая идея Бернштама

Бернштам говорил о длительном процессе вступления на территорию Средней Азии тюркоязычных племён в разные исторические периоды.

Следовательно, этногенез рассматривается как процесс, а не как одномоментное событие.

Важен не поиск единственной причины, а реконструкция сложного исторического взаимодействия.

8. От текста к цифровому источнику

Традиционный путь:

АРХИВ → ДОКУМЕНТ → ЧТЕНИЕ → ИНТЕРПРЕТАЦИЯ

Цифровой путь:

**АРХИВ → ЦИФРОВАЯ КОПИЯ → OCR → ПРОВЕРКА → СТРУКТУРИРОВАНИЕ → АНАЛИЗ →
ИНТЕРПРЕТАЦИЯ**

9. Что означает «цифровая трансформация»?

Это не простое превращение бумажного документа в электронный файл.

Цифровая трансформация включает:

- создание цифрового представления;
- распознавание текста;
- нормализацию и техническую проверку;
- структурирование данных;
- создание метаданных;
- полнотекстовый поиск;
- тематическое и сетевое сопоставление.

10. Где применяется ИИ?

ИИ может помочь в:

- распознавании и обработке текста;
- поиске имён и терминов;
- выявлении тематических блоков;
- классификации фрагментов;
- построении рабочих указателей;
- сопоставлении повторяющихся сюжетов;
- извлечении метаданных;
- подготовке цифрового корпуса источников.

11. Модель работы с источником

1. Архивный документ



2. Цифровая копия



3. Распознавание текста



4. Редакторская проверка



5. Источниковедческая разметка



6. ИИ-анализ



7. Проверка историком



12. Первый уровень: сохранение

Цель — обеспечить цифровую сохранность источника и воспроизводимость работы.

Фиксируются:

- происхождение документа;
- архивный контекст;
- название;
- датировка;
- объём;
- структура;
- сведения об авторах и участниках.

13. Второй уровень: распознавание

Большие архивные тексты требуют OCR и последующей проверки.

Проблемы:

- ошибки распознавания;
- старые орфографические формы;
- имена собственные;
- географические названия;
- специальные термины;
- дефекты исходной копии.

ИИ может ускорять обработку, но не отменяет верификацию.

14. Третий уровень: структурирование

Стенограмма должна быть представлена как структурированный объект.

Возможные единицы разметки:

- выступление;
- доклад;
- вопрос;
- ответ;
- реплика;
- имя исследователя;
- научный термин;
- географическое название;
- упомянутый источник.

15. Четвёртый уровень: семантический анализ

ИИ позволяет перейти от поиска отдельных слов к поиску смысловых блоков.

Например:

«этногенез» → происхождение → миграции → тюркоязычное население → этнические контакты → формирование общностей.

16. Пятый уровень: карта научной дискуссии

Стенограмма позволяет реконструировать:

КТО выступает?



КАКУЮ ПОЗИЦИЮ ФОРМУЛИРУЕТ?



КАКИЕ ИСТОЧНИКИ ИСПОЛЬЗУЕТ?



КТО КРИТИКУЕТ?



КАКИЕ ВОПРОСЫ ОСТАЮТСЯ СПОРНЫМИ?

17. Бернштам и научная полемика

Материалы позволяют рассматривать концепцию Бернштама не изолированно, а внутри научной дискуссии.

В последующих материалах отражена критика отдельных положений его концепции, в частности со стороны Л. П. Кызласова.

Для цифрового анализа это означает необходимость связывать не только имена, но и аргументы.

18. Что именно может анализировать ИИ?

Уровень 1 — поиск: имена, термины, даты.

Уровень 2 — классификация: темы и типы фрагментов.

Уровень 3 — сопоставление: повторяющиеся аргументы и сюжеты.

Уровень 4 — визуализация: связи между участниками, темами и источниками.

Уровень 5 — исследовательская навигация по большому корпусу.

19. Принцип «человек в контуре»

ИИ не должен получать право самостоятельно устанавливать историческую истину.

Историк:

- проверяет распознавание;
- устанавливает контекст;
- различает факт и интерпретацию;
- проверяет атрибуцию;
- сопоставляет источники;
- принимает научное решение.

20. Основные риски ИИ

- галлюцинации и вымышленные сведения;
- ошибочная атрибуция высказывания;
- потеря контекста;
- смешение разных научных позиций;
- искажение терминологии;
- ошибочное распознавание имён;
- автоматическая интерпретация спорного материала.

21. Источниковедческая критика в цифровой среде

Цифровой инструмент не отменяет классические вопросы:

КТО создал источник?

КОГДА?

ГДЕ?

В КАКОМ КОНТЕКСТЕ?

С КАКОЙ ЦЕЛЬЮ?

КАК ДОКУМЕНТ СОХРАНИЛСЯ?

КАК ОН БЫЛ ОЦИФРОВАН?

КАКИЕ ИЗМЕНЕНИЯ ВНЕСЕНЫ ПРИ ОБРАБОТКЕ?

22. Стенограмма как «многослойный» источник

Первичный слой — содержание выступлений.

Второй слой — структура научной дискуссии.

Третий слой — научная терминология эпохи.

Четвёртый слой — отношения между исследовательскими позициями.

Пятый слой — исторический контекст советской науки.

Цифровые методы позволяют исследовать эти слои одновременно.

23. От отдельного документа к цифровому корпусу

Перспектива исследования:

СТЕНОГРАММА 1956

+

АРХИВНЫЕ МАТЕРИАЛЫ А. Н. БЕРНШТАМА

+

ПУБЛИКАЦИИ УЧАСТНИКОВ ДИСКУССИИ

+

АРХЕОЛОГИЧЕСКИЕ И ЭТНОГРАФИЧЕСКИЕ МАТЕРИАЛЫ

=

ЦИФРОВОЙ КОРПУС ПО ИСТОРИИ И ЭТНОГЕНЕЗУ КЫРГЫЗОВ

24. Возможности цифрового корпуса

- единый полнотекстовый поиск;
- поиск по персоналиям;
- поиск по географическим названиям;
- тематическая навигация;
- сопоставление научных концепций;
- хронологическая навигация;
- связь документов с архивными единицами хранения;
- создание исследовательских визуализаций.

25. Почему это важно для истории Кыргызстана?

Значительная часть исторического наследия сохраняется в архивных фондах и не всегда доступна исследователю в удобной форме.

Цифровая трансформация позволяет:

- сохранять документы;
- повышать доступность источников;
- облегчать научный поиск;
- объединять разрозненные материалы;
- создавать новые формы исторического исследования.

26. Методологический результат

Опыт работы со стенограммой показывает:

1. Цифровизация меняет способы доступа к источнику.
2. ИИ ускоряет первичную обработку больших массивов.
3. Структурирование делает источник аналитически доступным.
4. Междисциплинарность источника открывает возможности для цифрового анализа.
5. Источниковедческая критика остаётся обязательной.

27. Научная новизна подхода

Новизна заключается не только в применении ИИ к историческому тексту.

Она состоит в соединении:

АРХИВИСТИКИ

+

ИСТОЧНИКОВЕДЕНИЯ

+

ИСТОРИИ ЭТНОГЕНЕЗА

+

ЦИФРОВЫХ ГУМАНИТАРНЫХ МЕТОДОВ

+

ИИ

28. Практический результат работы

Работа со стенограммой позволяет подготовить источник к современному научному использованию:

- научное редактирование;
- структурирование;
- система имён и терминов;
- указатели;
- цифровой поиск;
- тематическая разметка;
- дальнейшее включение в цифровой корпус.

29. От архивного документа к новым исследованиям

Архивный документ



Цифровая обработка



Структурированные данные



Новые вопросы



Сопоставление источников



Новые исторические исследования

Цифровизация здесь становится не конечным результатом, а исследовательским методом.

30. Основной вывод

Искусственный интеллект не заменяет историка.

Он создаёт новые возможности для работы историка с источником.

Главная задача цифровой трансформации — не автоматизировать историческое мышление, а сделать сложные исторические материалы более доступными для системного, проверяемого и междисциплинарного исследования.

31. Перспективы дальнейшей работы

- завершение научной подготовки стенограммы;
- создание структурированной цифровой версии;
- разработка системы метаданных;
- формирование именового и тематического указателей;
- создание цифрового корпуса материалов по этногенезу кыргызов;
- применение ИИ для навигации и сопоставления источников;
- разработка принципов верификации результатов ИИ.

32. Заключение

Стенограмма 1956 года — это не только документ истории науки.

Это источник, позволяющий увидеть процесс формирования научного знания о происхождении и этногенезе кыргызского народа.

Её цифровая трансформация создаёт возможность соединить архивное наследие, источниковедение и современные технологии искусственного интеллекта.

Именно в этом я вижу перспективу цифровой гуманитаристики в исследовании истории Кыргызстана.

СПАСИБО ЗА ВНИМАНИЕ

Нурзада Гарибовна Жантелиева

кандидат исторических наук, доцент

Кыргызский национальный университет им. Ж. Баласагына

«От сохранения источника — к новым возможностям его научного прочтения»