

Технология преобразования лекционного материала в учебные пособия на основе нейросетевого распознавания речи и больших языковых моделей

С.Н. Коренская, М.Г. Жабицкий, П.В. Анни. Р.И. Хазеев, Е.В. Маркидонов, П.С. Эрнст

Аннотация- В статье рассматривается задача автоматизированного преобразования лекционного материала в структурированные учебные документы на основе нейросетевого распознавания речи и больших языковых моделей. Предложена архитектура человеко-машинной интеллектуальной системы, основанная на функциональном распределении задач между преподавателем, интеллектуальным конвейером обработки и контекстными источниками знаний. Разработана формальная постановка задачи, определены исходные и целевые информационные объекты, требования к процессу преобразования и последовательность специализированных стадий обработки, включающих автоматическую транскрибацию, контекстную коррекцию, нормализацию письменной речи и семантическое структурирование материала. Архитектура реализована в виде локального программного прототипа и верифицирована в ходе двухэтапных исследований, включавших экспериментальную апробацию на университетских лекциях и независимую функциональную проверку после переноса в иную вычислительную среду. Полученные результаты подтвердили работоспособность предложенного архитектурного подхода, возможность формирования структурированных учебных документов, существенное сокращение трудоёмкости подготовки учебных материалов, а также воспроизводимость и переносимость программной реализации. Определены ограничения исследовательского прототипа и основные направления дальнейшего развития предложенной архитектуры.

Ключевые слова - автоматизированное формирование учебных пособий, большие языковые модели, нейросетевое распознавание речи, человеко-машинная интеллектуальная система, интеллектуальный конвейер, контекстное обогащение.

1. ВВЕДЕНИЕ

Одной из ключевых задач цифровой трансформации образования является кардинальное повышение производительности интеллектуального труда преподавателя [1-6]. Несмотря на широкое распространение цифровых образовательных технологий, значительная часть времени преподавателя по-прежнему затрачивается не на создание нового знания

и не на непосредственную работу со студентами, а на трудоёмкую переработку уже подготовленного учебного материала в различные формы его представления.

За последние годы в образовательных организациях накоплены большие объёмы аудио- и видеозаписей лекционных занятий высокого качества. Однако сами по себе видеозаписи и автоматически сформированные стенограммы обладают ограниченной пригодностью для самостоятельного изучения. Полноценное учебное пособие требует структурирования материала, устранения особенностей устной речи, восстановления терминологии, логического деления на разделы, формирования иллюстративных элементов и единого оформления.

Традиционно подобная переработка выполняется преподавателем вручную и представляет собой одну из наиболее трудоёмких составляющих подготовки учебно-методических материалов. В результате значительная часть прочитанных лекционных курсов так и не получает завершённого текстового представления в виде учебных пособий.

Современные средства нейросетевого распознавания речи и большие языковые модели впервые позволяют передать значительную часть этих рутинных когнитивных операций программной системе. Однако существующие сервисы, как правило, решают лишь отдельные задачи [7-11]: распознавание речи, исправление текста или генерацию отдельных фрагментов - и не обеспечивают сквозного автоматизированного преобразования лекционного материала в целостный структурированный учебный документ.

В связи с этим актуальной является разработка методики и цифровой технологии, объединяющих нейросетевое распознавание речи, большие языковые модели и специализированные алгоритмы формирования документов в единый процесс автоматизированного преобразования лекционного материала. Рассматриваемая задача не ограничивается сферой образования. Аналогичная потребность возникает при обработке материалов совещаний, экспертных обсуждений, мозговых штурмов, интервью и иных форм

Статья получена 17 июля 2026

Коренская С.Н. – магистрант ВИШ НИЯУ МИФИ, (e-mail: korenskayas@mail.ru);

Жабицкий М.Г. - заместитель директора ВИШ НИЯУ МИФИ (e-mail: jabitsky@mail.ru);

Анни П.В. – специалист компании Red Hat (e-mail: pavel.anni@gmail.com)

Хазеев Р.И. - директор по цифровым технологиям ПАО «Ростелеком» (e-mail: Roman.khazeev@yandex.ru)

Маркидонов Е.В - студент ВИШ НИЯУ МИФИ; (e-mail: egorm987@gmail.com)

Эрнст П.С. - студент ВИШ НИЯУ МИФИ (e-mail: ernst7591@mail.ru)

содержательной устной коммуникации, результаты которых должны быть преобразованы в структурированный документ: протокол, отчёт, аналитическую записку, перечень решений или план действий. Во всех подобных случаях требуется не только распознавание речи, но и контролируемое сохранение существенного смыслового содержания, восстановление логической структуры обсуждения, выделение ключевых положений и представление результата в форме, пригодной для дальнейшего использования. Тем самым исследуемая задача представляет собой частный случай более широкого класса задач автоматизированного преобразования речевого материала в структурированные текстовые документы.

2. ПОСТАНОВКА ЗАДАЧИ АВТОМАТИЗИРОВАННОГО ФОРМИРОВАНИЯ УЧЕБНОГО ПОСОБИЯ

Автоматизированное формирование учебного пособия рассматривается как задача интеллектуального преобразования авторского лекционного материала в структурированный образовательный документ. В отличие от генерации текста по заданной теме, в рассматриваемой постановке основным источником содержания является лекция преподавателя, отражающая образовательные цели, последовательность раскрытия понятий, смысловые акценты, определения, пояснения и примеры. Требуется сохранить данный педагогический замысел, одновременно сократив трудоёмкость подготовки учебного материала и повысив качество его текстового и структурного представления.

Решение задачи основывается на распределении интеллектуальных функций между преподавателем, программной системой и контекстными источниками знаний. При этом автоматизации подлежат преимущественно повторяющиеся когнитивные операции, связанные с распознаванием, коррекцией, нормализацией, структурированием и оформлением материала. Целеполагание, формирование образовательного сценария и итоговая оценка качества результата сохраняются за преподавателем.

Для системной постановки задачи необходимо определить функциональные роли участников процесса, исходные и целевые информационные объекты, требования к преобразованию и последовательность изменения состояния обрабатываемого материала.

2.1. КОНЦЕПТУАЛЬНАЯ МОДЕЛЬ ПРОЦЕССА

Предлагаемая методика основана на распределении интеллектуальных функций между участниками процесса в соответствии с их компетенциями. Множество функциональных ролей представим в виде

$$R = \{R_P, R_S, R_K\},$$

где R_P - преподаватель; R_S - интеллектуальная программная система; R_K - контекстные источники знаний. Ролевая модель исследуемого бизнес-процесса достаточно подробно анализировалась [2-5].

Преподаватель выполняет целеполагающую и педагогическую функции. Он определяет образовательные результаты, состав и последовательность раскрытия материала, уровень подробности, смысловые акценты, способы объяснения и набор примеров. Тем самым преподаватель формирует

авторский образовательный сценарий, который должен сохраняться при последующих преобразованиях лекционного материала. После формирования проекта учебного документа преподаватель также выполняет функцию его педагогической верификации, оценивая полноту, корректность, последовательность изложения и соответствие материала целевой аудитории.

Интеллектуальная система выполняет функции преобразования формы представления информации. К ним относятся распознавание речи, устранение ошибок автоматической транскрипции, преобразование устной речи в письменную, выделение смысловых фрагментов, формирование структуры, согласование терминологии и подготовка проекта учебного документа. Передача данных операций системе позволяет заменить последовательную ручную переработку всего лекционного материала экспертным контролем промежуточных версий и итогового результата.

Контекстные источники знаний обеспечивают предметную поддержку процесса преобразования. В качестве таких источников могут использоваться учебники, учебно-методические материалы, справочники, предметные словари, нормативные документы и специализированные терминологические материалы. Их назначением являются уточнение терминов, устранение неоднозначностей, восстановление неявных смысловых связей и обеспечение согласованности понятийного аппарата. Контекстные источники не должны подменять лекционный материал самостоятельным изложением темы или изменять заданный преподавателем образовательный сценарий.

Таким образом, преподаватель определяет цели, содержание и критерии качества, интеллектуальная система выполняет трудоёмкие операции преобразования, а контекстные источники обеспечивают предметную поддержку. Такое распределение функций направлено на повышение производительности интеллектуального труда преподавателя при сохранении его ведущей роли в формировании и верификации учебного материала.

2.2. ИСХОДНЫЕ И ЦЕЛЕВЫЕ ИНФОРМАЦИОННЫЕ ОБЪЕКТЫ

Множество входных информационных объектов представим в виде

$$I_{in} = \{M_{lec}, C_{disc}, K_{ctx}, R_{edu}\},$$

где M_{lec} - исходный лекционный материал; C_{disc} - сведения о дисциплине, теме и месте лекции в структуре курса; K_{ctx} - контекстные источники знаний и предметная терминология; R_{edu} - требования преподавателя к целевому учебному документу.

Основным входным объектом M_{lec} является аудио- или видеозапись лекции. Она содержит не только предметную информацию, но и последовательность её раскрытия, авторские пояснения, смысловые акценты, примеры и переходы между темами. При наличии презентация преподавателя может использоваться как дополнительный информационный объект, обеспечивающий уточнение терминологии, структуры и содержания отдельных фрагментов. В текущей реализации основным обрабатываемым каналом является речевая составляющая аудио- или видеозаписи,

а совместная обработка речи и визуальных материалов рассматривается как направление дальнейшего развития. Объект C_disc задаёт образовательный контекст преобразования: наименование дисциплины и темы, место материала в структуре курса, характеристики целевой аудитории и предполагаемый уровень подготовки обучающихся.

Объект K_stx включает материалы, используемые для контекстной поддержки распознавания и последующего преобразования текста. Их применение должно быть ограничено задачами уточнения, согласования и контролируемого раскрытия предметного содержания. Контекстное обогащение исходных материалов обсуждалось в ряде работ, например [1, 3, 12].

Объект R_edu содержит требования преподавателя к структуре, стилю, уровню подробности, терминологии и составу целевого документа. Данные требования могут уточняться по результатам экспертной оценки по мере формирования промежуточных версий и итогового артефакта работы.

Целевым информационным объектом является структурированный учебный документ

$$D_edu = \{H, A, S, Def, Exp, Ex, Con\},$$

где H - название документа; A - аннотация; S - система логически связанных разделов и подразделов; Def - определения понятий; Exp - пояснения; Ex - примеры; Con - промежуточные и итоговые выводы.

В дальнейшем состав D_edu может быть расширен связанными изображениями, схемами, таблицами, формулами и другими визуальными материалами. В границах рассматриваемого прототипа основным результатом является структурированный текстовый документ.

Следовательно, целевой результат принципиально отличается от автоматической стенограммы. Стенограмма воспроизводит последовательность речевых фрагментов, тогда как учебный документ должен представлять материал в письменной, логически организованной и методически пригодной форме. Вместе с тем преобразование не должно превращать лекцию в сокращённый конспект или изменять её авторскую содержательную логику.

2.3. ТРЕБОВАНИЯ К ПРОЦЕССУ ПРЕОБРАЗОВАНИЯ

Требования к автоматизированному преобразованию лекционного материала могут быть объединены в четыре группы: требования к сохранению содержания, требования к преобразованию формы, требования к педагогическому качеству и эксплуатационные требования.

К требованиям сохранения содержания относятся:

- сохранение существенного смыслового содержания лекции;
- сохранение фактов, числовых данных, терминов, определений, причинно-следственных связей, пояснений и примеров;
- сохранение авторской последовательности раскрытия материала и смысловых акцентов;
- недопущение внесения в учебный документ неподтверждённых положений, отсутствующих в лекции и разрешённых контекстных источниках;

- возможность установления смыслового соответствия между исходным лекционным материалом и сформированным документом.

К требованиям преобразования формы относятся:

- устранение ошибок автоматического распознавания речи;
- восстановление специальной терминологии, имён собственных, обозначений и сокращений;
- удаление повторов, речевых пауз, незавершённых конструкций и других особенностей устного изложения, не несущих самостоятельного смыслового содержания;
- преобразование устной речи в связный письменный текст;
- выделение смысловых блоков и формирование логической структуры документа;
- согласование терминологии, определений и обозначений в пределах всего документа;
- формирование заголовков, аннотации, пояснений, примеров и выводов в соответствии с требованиями R_edu.

К требованиям педагогического качества относятся:

- соответствие документа образовательным целям и целевой аудитории;
- полнота раскрытия основных положений лекции;
- логическая связность разделов;
- достаточность пояснений и примеров;
- сохранение авторского педагогического сценария;
- возможность экспертной оценки и последующего уточнения результата преподавателем.

К эксплуатационным требованиям относятся:

- обработка продолжительных аудио- и видеоматериалов;
- локальное выполнение основных операций без обязательной передачи лекционных материалов внешним сервисам;
- сохранение промежуточных результатов;
- возможность контроля отдельных стадий преобразования;
- повторяемость обработки при одинаковых исходных данных и настройках;
- возможность изменения требований и повторного формирования документа без ручной переработки всего текста.

Ключевым ограничением является недопустимость повышения формальной структурированности документа за счёт необоснованного сокращения содержания лекции. Удалению подлежат прежде всего речевые дефекты, повторы и фрагменты, не несущие самостоятельного значения. Факты, определения, аргументы, примеры и пояснения должны сохраняться либо преобразовываться без утраты выполняемой ими образовательной функции.

Тем самым качество результата определяется не минимизацией объёма текста, а достижением рационального соотношения между полнотой содержания, логической структурированностью, письменной связностью и соответствием педагогическому замыслу преподавателя [7, 12].

2.4. СТРУКТУРНО-ФУНКЦИОНАЛЬНАЯ ПОСТАНОВКА ЗАДАЧИ

В общем виде задача автоматизированного формирования учебного пособия может быть

представлена как преобразование множества входных информационных объектов в целевой документ:

$$D_{edu} = F(M_{lec}, C_{disc}, K_{ctx}, R_{edu}),$$

где F - совокупность взаимосвязанных операций интеллектуального преобразования лекционного материала. Функция F должна обеспечивать формирование документа D_{edu} , удовлетворяющего заданным требованиям к содержанию, структуре, стилю и педагогическому качеству. При этом лекционный материал M_{lec} остаётся основным источником содержания, C_{disc} определяет образовательный контекст, K_{ctx} обеспечивает предметную поддержку, а R_{edu} задаёт требования к результату.

Внутренняя структура преобразования может быть представлена последовательностью состояний информационного объекта:

$$M_{lec} \rightarrow T_{raw} \rightarrow T_{corr} \rightarrow T_{norm} \rightarrow D_{edu},$$

где T_{raw} - автоматическая стенограмма лекционного материала; T_{corr} - текст после контекстной коррекции ошибок распознавания; T_{norm} - нормализованный письменный текст; D_{edu} - структурированный учебный документ.

На стадии формирования T_{raw} выполняется преобразование речевого сигнала в последовательность текстовых фрагментов. Основным критерием результата является получение стенограммы, охватывающей весь исходный лекционный материал и пригодной для дальнейшей обработки.

На стадии формирования T_{corr} устраняются ошибки распознавания, уточняются специальные термины, обозначения, имена собственные и смысловые связи. Коррекция выполняется с использованием локального контекста лекции, сведений о дисциплине и разрешённых контекстных источников. Критерием результата является терминологическая и фактическая согласованность текста с исходным материалом.

На стадии формирования T_{norm} осуществляется преобразование особенностей устной речи в письменное изложение: устраняются повторы и речевые дефекты, восстанавливаются завершённые предложения, уточняются связи между фрагментами и формируются логически целостные текстовые блоки. Критерием результата является связность письменного текста при сохранении существенного содержания исходной лекции.

На стадии формирования D_{edu} выполняются семантическое структурирование материала, выделение разделов и подразделов, формирование заголовков, определений, пояснений, примеров и выводов, а также оформление документа в соответствии с R_{edu} . Критерием результата является соответствие сформированного документа установленным содержательным, структурным и педагогическим требованиям.

Полученный документ передаётся преподавателю для педагогической верификации. Результат экспертной оценки может быть представлен множеством замечаний и уточнений

$$E_P = \{e_1, e_2, \dots, e_n\}.$$

На основании E_P корректируются требования R_{edu} и при необходимости итерационно повторяется преобразование:

$$\begin{aligned} R_{edu}^{(i+1)} &= G(R_{edu}^{(i)}, E_P^{(i)}), \\ D_{edu}^{(i+1)} &= F(M_{lec}, \\ &C_{disc}, K_{ctx}, R_{edu}^{(i+1)}) \end{aligned}$$

где i - номер итерации; G - функция уточнения требований по результатам экспертной оценки. Итерационный характер процесса позволяет сочетать скорость автоматизированной обработки с предметной и педагогической ответственностью преподавателя. При этом преподаватель переходит от ручного переписывания и последовательного редактирования всего лекционного материала к постановке требований, оценке результата и целевому уточнению отдельных элементов документа.

Таким образом, задача автоматизированного формирования учебного пособия сводится к последовательности контролируемых преобразований информационного объекта при заданном распределении функций между преподавателем, интеллектуальной системой и контекстными источниками знаний. Целевым эффектом является повышение производительности интеллектуального труда преподавателя при одновременном обеспечении полноты, связности и педагогического качества формируемого учебного документа. Реализация функции F в виде программного конвейера рассматривается в следующем разделе.

3. АРХИТЕКТУРА ЧЕЛОВЕКО-МАШИННОЙ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ ФОРМИРОВАНИЯ УЧЕБНЫХ ДОКУМЕНТОВ

В отличие от традиционных средств автоматизации подготовки учебных материалов, разработанная система рассматривает преподавателя не как внешнего пользователя программного обеспечения, а как полноценный функциональный элемент процесса формирования учебного документа. Автоматизированная обработка лекционного материала осуществляется интеллектуальным конвейером, тогда как преподаватель определяет педагогические требования, выполняет экспертную оценку результатов и принимает решение о готовности документа. Такое распределение интеллектуальных функций позволяет существенно сократить объем рутинной когнитивной работы без потери авторского педагогического замысла. Формально архитектура системы может быть представлена кортежем

$$S_{HMI} = \langle H, P, K, D, C \rangle,$$

где H - преподаватель; P - интеллектуальный конвейер обработки лекционного материала; K - контекстные источники знаний; D - информационные объекты, формируемые в процессе обработки; C - контур экспертного управления качеством.

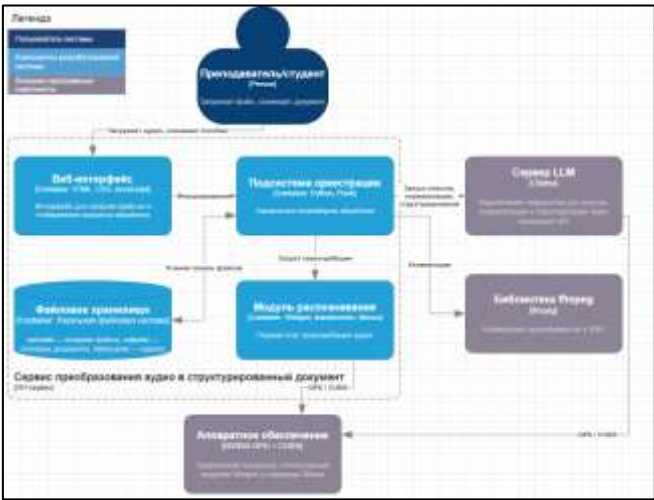


Рисунок 1 - Архитектура человеко-машинной интеллектуальной системы формирования учебных документов.

Предлагаемая архитектура является модульной. Это позволяет заменять отдельные программные компоненты без изменения общей организации процесса. В частности, средства автоматического распознавания речи, большие языковые модели и сервисы контекстной коррекции могут изменяться по мере развития технологий при сохранении общей архитектуры человеко-машинной системы. Более того, цифровой инструмент изначально имеет возможность подключения различных облачных или внутриконтурных LLM и обогащения их предметным контекстом в различных вариантах. Исходный материал может порождать несколько версий конечного продукта для различных образовательных целей и уровней зрелости аудитории.

3.1. ИНТЕЛЛЕКТУАЛЬНЫЙ КОНВЕЙЕР ОБРАБОТКИ ЛЕКЦИОННОГО МАТЕРИАЛА

Интеллектуальный конвейер реализует последовательность специализированных преобразований исходного лекционного материала в структурированный учебный документ.

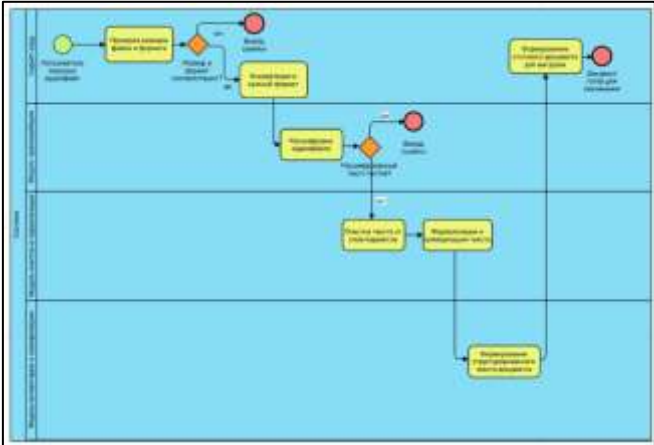


Рисунок 2 - Интеллектуальный конвейер обработки лекционного материала.

Архитектурно конвейер представляет собой композицию функциональных операторов

$$P = F_S \circ F_N \circ F_C \circ F_P \circ F_R \circ F_A,$$

где F_A - подготовка аудиоматериала; F_R - автоматическое распознавание речи; F_P - первичная очистка текста; F_C - контекстная коррекция терминологии; F_N - нормализация

письменной речи; F_S - семантическое структурирование учебного документа.

Функциональные компоненты конвейера представлены в таблице 1.

Таблица 1 - Функциональные компоненты интеллектуального конвейера

Компонент	Входной объект	Выходной объект	Основная функция
Подготовка материала	Видеолекция	Аудиопоток	Выделение и нормализация аудио
Распознавание речи	Аудиопоток	Первичный текст	Автоматическая транскрипция
Очистка текста	Первичный текст	Очищенный текст	Удаление речевых артефактов
Контекстная коррекция	Очищенный текст	Скорректированный текст	Исправление терминологии с использованием базы знаний
Нормализация	Скорректированный текст	Нормализованный текст	Преобразование устной речи в письменную
Семантическое структурирование	Нормализованный текст	Проект учебного документа	Формирование разделов, определений, примеров и выводов

Каждый компонент выполняет строго определенную функцию и взаимодействует с соседними компонентами через формализованные информационные объекты. Такое разделение обеспечивает независимую модернизацию отдельных этапов обработки и упрощает адаптацию системы к новым инструментам искусственного интеллекта.

3.2. ОБРАБОТКА ПРОДОЛЖИТЕЛЬНЫХ ЛЕКЦИЙ

Продолжительные лекционные материалы не могут обрабатываться как единый информационный объект из-за ограничений вычислительных ресурсов, контекстного окна языковых моделей и необходимости сохранения управляемости промежуточных результатов [7-10]. Поэтому исходный материал предварительно разбивается на последовательность частично перекрывающихся фрагментов

$$M_{lec} = \{T_1, T_2, \dots, T_n\},$$

где T_i - отдельный фрагмент лекции, а пересечение соседних фрагментов используется для сохранения смысловых связей на границах разбиения.

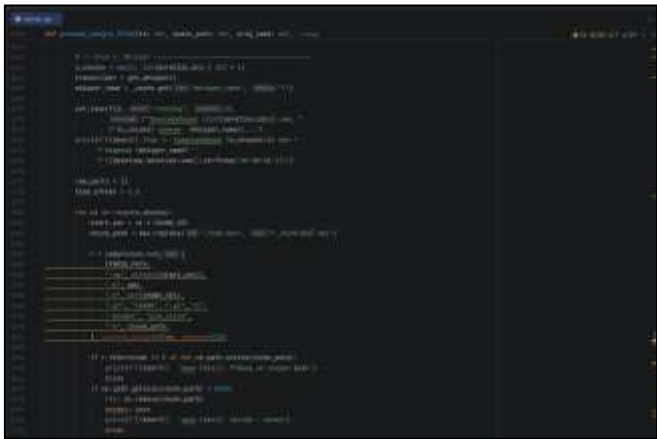


Рисунок 3 - Схема обработки продолжительной лекции



Обработка каждого фрагмента выполняется тем же интеллектуальным конвейером, что и для короткого материала. При этом результат формируется не простым объединением локально обработанных текстов, а посредством дополнительного оператора глобального согласования:

$$D_{edu} = G(F(T_1) \cdot F(T_2) \dots F(T_n)),$$

где $F(T_i)$ - результат обработки i -го фрагмента, а G - оператор согласования терминологии, структуры и логических связей между частями документа. Для сохранения целостности материала применяются следующие механизмы:

- перекрытие соседних фрагментов;
- использование единого терминологического словаря;
- передача накопленного контекста между этапами обработки;
- сохранение промежуточных результатов;
- финальная унификация структуры и стиля документа.

Простое последовательное объединение локальных результатов может приводить к повторению одних и тех же положений, расхождению терминологии и нарушению логики изложения. Поэтому после объединения выполняется дополнительная проверка связности разделов, устранение дублирования и выравнивание уровня детализации. В результате

продолжительная лекция преобразуется в единый учебный документ, а не в механическую совокупность независимо обработанных фрагментов.

3.3. РАСПРЕДЕЛЕНИЕ ИНТЕЛЛЕКТУАЛЬНЫХ ФУНКЦИЙ В ЧЕЛОВЕКО-МАШИННОЙ СИСТЕМЕ

Ключевой особенностью разработанной архитектуры является изменение принципов взаимодействия преподавателя и программной системы. В традиционной технологии автоматизированной подготовки документов преподаватель непосредственно редактирует автоматически сформированный текст, постепенно доводя его до требуемого качества. При таком подходе большая часть трудозатрат приходится на многократное выполнение рутинных операций анализа, исправления и структурирования текста [2-5].

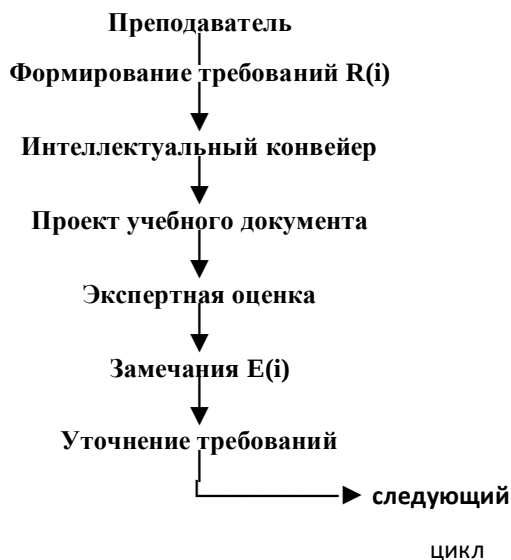
В предлагаемой архитектуре реализован иной принцип взаимодействия. Преподаватель не выполняет последовательное редактирование документа, а осуществляет управление процессом его формирования посредством постановки требований и экспертной оценки результата. Все операции преобразования текста выполняются интеллектуальным конвейером автоматически. Таким образом, объектом управления становится не сам документ, а процесс его формирования.

Такое распределение функций позволяет разделить деятельность человека и программной системы в соответствии с их сильными сторонами. Распределение интеллектуальных функций представлено в таблице 2.

Таблица 2 - Распределение интеллектуальных функций в человеко-машинной системе

Функции преподавателя	Функции интеллектуальной системы
Определение структуры и педагогических требований	Автоматическая обработка лекционного материала
Формирование критериев качества документа	Распознавание речи и преобразование аудио в текст
Экспертная оценка содержания и структуры	Контекстная коррекция терминологии
Анализ замечаний и принятие решения о необходимости повторной обработки	Нормализация письменной речи
Уточнение требований к следующему циклу формирования	Семантическое структурирование документа
Подтверждение готовности итогового документа	Автоматическое формирование новой версии документа

Механизм взаимодействия преподавателя и интеллектуального конвейера реализован в виде замкнутого контура управления качеством формирования документа.



Пусть после выполнения очередного цикла сформирован документ $D^{(i)}$. По результатам его анализа преподавателем формируется множество экспертных замечаний

$$E^{(i)} = \{e_1, e_2, \dots, e_m\},$$

на основании которого уточняется множество требований

$$R^{(i+1)} = G(R^{(i)}, E^{(i)}).$$

Интеллектуальный конвейер повторно выполняет обработку исходного лекционного материала уже с учетом обновленных требований

$$D^{(i+1)} = F(M_{\text{lec}}, C_{\text{disc}}, K_{\text{ctx}}, R^{(i+1)}).$$

Принципиальным отличием предложенной архитектуры является то, что преподаватель не редактирует промежуточный текст вручную. Вместо этого корректируются требования к процессу формирования документа, после чего система автоматически генерирует новую согласованную версию. Такой подход исключает накопление локальных исправлений, обеспечивает воспроизводимость процесса формирования документа и позволяет повторно использовать накопленный опыт экспертной оценки.

Таким образом, научная новизна предложенной архитектуры заключается не в использовании отдельных технологий искусственного интеллекта, а в реализации человеко-машинной системы, основанной на функциональном распределении интеллектуальных задач между преподавателем и программными компонентами. Автоматизированная система принимает на себя выполнение трудоемких операций обработки информации, тогда как преподаватель сохраняет функции постановки целей, экспертной оценки и управления качеством результата. Это обеспечивает существенное снижение трудоемкости подготовки учебных документов при сохранении авторского педагогического замысла.

4. РЕАЛИЗАЦИЯ АРХИТЕКТУРЫ В ПРОГРАММНОМ ПРОТОТИПЕ

Разработанная архитектура была реализована в виде исследовательского программного прототипа, предназначенного для проверки работоспособности

предложенного человеко-машинного подхода к автоматизированному формированию учебных документов. Целью реализации являлось не создание законченного коммерческого продукта, а практическая верификация архитектурных решений, предложенных в предыдущем разделе. Поэтому выбор конкретных программных средств определялся возможностью быстрого построения экспериментальной системы при сохранении всех ключевых архитектурных принципов: модульности, локального выполнения, наблюдаемости процесса обработки и поддержки итерационного взаимодействия с преподавателем [13-16].

4.1. РЕАЛИЗАЦИЯ АРХИТЕКТУРНЫХ КОМПОНЕНТОВ

Программный прототип реализует все основные компоненты предложенной человеко-машинной системы (рис. 5).

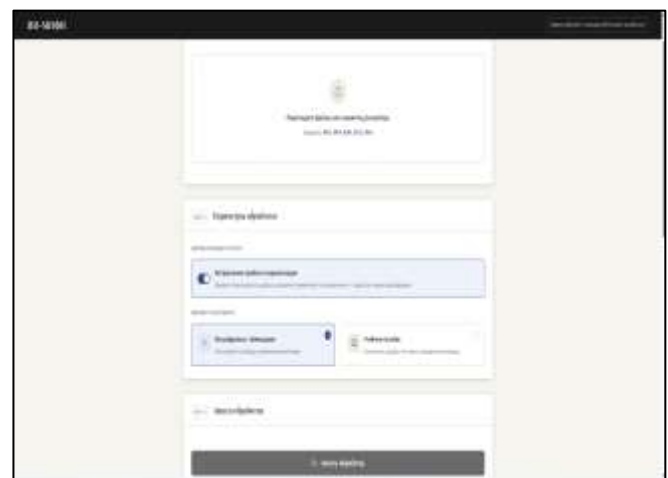
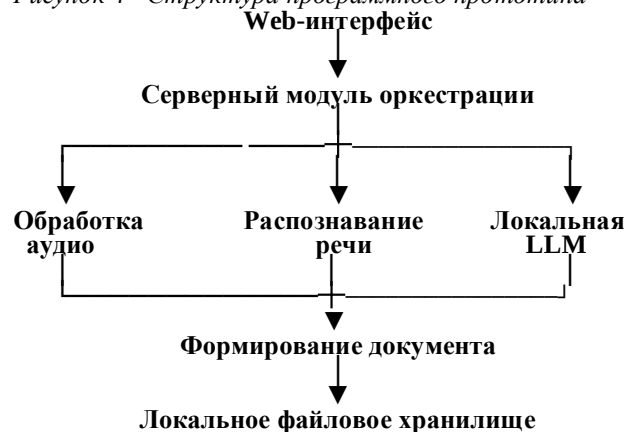


Рисунок 4 - Структура программного прототипа



Веб-интерфейс обеспечивает взаимодействие преподавателя с системой, загрузку исходных материалов, выбор режима обработки и получение сформированного документа. Серверный модуль выполняет оркестрацию всех стадий интеллектуального конвейера, контролирует последовательность их выполнения и организует обмен данными между функциональными компонентами.

Отдельные программные модули реализуют подготовку аудиоматериала, автоматическое распознавание речи, взаимодействие с локальной языковой моделью и формирование итогового учебного документа. Все промежуточные результаты сохраняются в локальном файловом хранилище, что обеспечивает возможность контроля процесса обработки, повторного

выполнения отдельных стадий и накопления истории выполненных задач.

Подобная организация непосредственно отражает архитектурную декомпозицию, рассмотренную в разделе 3, и обеспечивает независимое развитие отдельных компонентов без изменения общей структуры системы.

4.2. ВЫБОР ПРОГРАММНЫХ СРЕДСТВ РЕАЛИЗАЦИИ

При разработке прототипа использовались свободно распространяемые программные средства, позволяющие реализовать архитектуру локального интеллектуального конвейера без зависимости от внешних облачных сервисов.

Таблица 3 Используемые программные компоненты

Архитектурная функция	Программная реализация
Подготовка аудио	FFmpeg
Автоматическое распознавание речи	Whisper
Контекстная коррекция и нормализация	Qwen2.5 / Llama 3 (Ollama)
Оркестрация обработки	Python
Пользовательский интерфейс	Flask, HTML, JavaScript
Формирование результата	Markdown

Следует отметить, что приведённый набор программных средств относится исключительно к реализации экспериментального прототипа. Предложенная архитектура не зависит от использования конкретных моделей искусственного интеллекта или программных библиотек. При модернизации системы отдельные компоненты могут быть заменены более совершенными аналогами при сохранении общей архитектуры человеко-машинной системы и интерфейсов взаимодействия функциональных блоков.

4.3. РЕАЛИЗАЦИЯ АРХИТЕКТУРНЫХ ТРЕБОВАНИЙ

Одним из требований, сформулированных при проектировании архитектуры, являлось обеспечение устойчивости работы интеллектуального конвейера и возможности воспроизводимого выполнения отдельных стадий обработки. Для реализации данных требований в программном прототипе использован ряд инженерных решений.

Во-первых, предусмотрена возможность автоматического переключения между различными языковыми моделями, включая использование резервной модели при возникновении ошибок основной.

Во-вторых, вычислительные компоненты поддерживают выполнение как на графическом процессоре, так и на центральном процессоре, что обеспечивает работоспособность системы на вычислительных платформах различной конфигурации.

В-третьих, при обнаружении аномального результата генерации выполняется возврат к предыдущему корректному состоянию обработки с сохранением

исходного текста соответствующей стадии конвейера. Для повышения воспроизводимости используются ограничения длины генерируемого ответа и пониженные значения параметра температуры языковой модели.

Наконец, архитектура предусматривает обязательное сохранение всех промежуточных результатов (стенограмма, скорректированный текст, нормализованный текст и итоговый документ), что позволяет повторно выполнять отдельные стадии интеллектуального конвейера без необходимости повторной обработки исходной аудиозаписи.

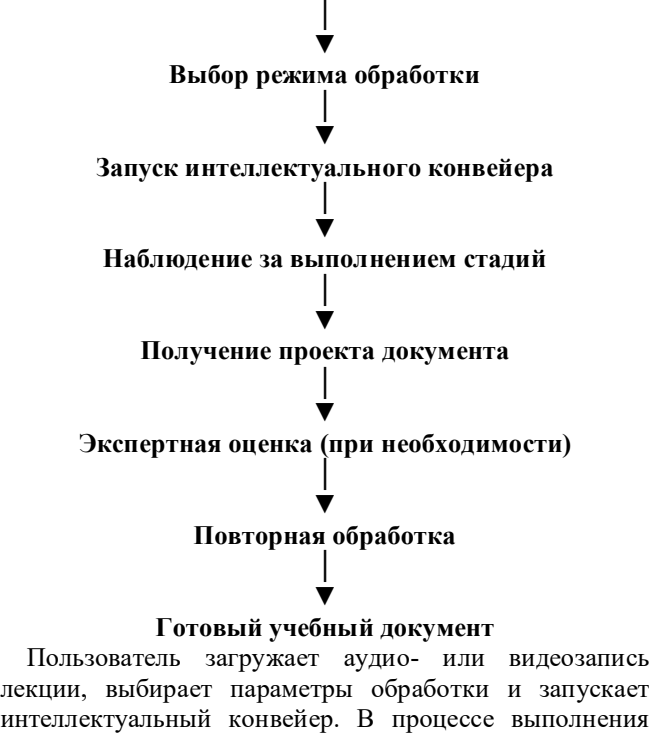
Таким образом, программная реализация обеспечивает не только выполнение последовательности преобразований, но и наблюдаемость процесса обработки, возможность локализации ошибок и поддержку итерационного совершенствования результата.

4.4. РЕАЛИЗАЦИЯ ПОЛЬЗОВАТЕЛЬСКОГО СЦЕНАРИЯ

Работа пользователя с программным прототипом соответствует общей логике человеко-машинной системы, предложенной в предыдущем разделе.



Рисунок 6 - Основной сценарий работы с системой



отображается текущее состояние обработки и сохраняются промежуточные результаты отдельных стадий. После завершения автоматического цикла пользователь получает проект учебного документа, который может быть либо принят без изменений, либо направлен на повторную обработку после уточнения требований.

Следует отметить, что повторный цикл не предполагает ручного редактирования автоматически сформированного текста. Преподаватель осуществляет экспертную оценку результата и формирует замечания, которые используются системой для повторного выполнения необходимых стадий интеллектуального конвейера. Тем самым в программном прототипе реализована предложенная в разделе 3 архитектура человеко-машинной интеллектуальной системы, основанная на распределении интеллектуальных функций между преподавателем и программными компонентами.

5. МЕТОДИКА ВЕРИФИКАЦИИ АРХИТЕКТУРЫ И ПРОГРАММНОГО ПРОТОТИПА

Верификация предложенной архитектуры человеко-машинной интеллектуальной системы выполнялась в два последовательных этапа. На первом этапе проводилась экспериментальная апробация разработанного программного прототипа автором выпускной квалификационной работы. Ее целью являлась проверка реализуемости предложенной архитектуры при обработке реальных лекционных материалов, а также получение экспериментальных данных для оценки качества отдельных стадий интеллектуального конвейера. На втором этапе выполнялась независимая функциональная верификация программного прототипа после его переноса в иную вычислительную среду. Данный этап был направлен на подтверждение воспроизводимости предложенной архитектуры и проверку возможности ее автономного локального развертывания. Подходы к испытаниям основаны на.

Разделение исследований на два самостоятельных этапа позволило независимо оценить архитектурную состоятельность предложенного подхода и корректность его программной реализации.

5.1. ПРОГРАММА ЭКСПЕРИМЕНТАЛЬНОЙ АПРОБАЦИИ

Экспериментальная апробация проводилась в процессе разработки программного прототипа и охватывала полный цикл автоматизированного формирования учебного документа по материалам университетских лекций.

В качестве экспериментальных материалов использовались три лекции общей продолжительностью около четырех часов. Исходные записи отличались качеством звука, темпом речи преподавателя, объемом предметной терминологии и наличием реплик слушателей, что позволило проверить работоспособность интеллектуального конвейера при обработке разнородных лекционных материалов.

В ходе исследований последовательно оценивались все стадии разработанного конвейера: подготовка аудиоматериала, автоматическая транскрибация, контекстная коррекция, нормализация письменной речи,

семантическое структурирование и формирование итогового учебного документа.

Таблица 4 Программа экспериментальной апробации

Проверяемый этап	Метод проверки
Подготовка аудиоматериала	Контроль корректности преобразования входных данных
Автоматическая транскрибация	Сравнение с исходной записью и расчет доли корректно распознанных слов
Контекстная коррекция	Экспертная оценка исправления ошибок распознавания
Нормализация текста	Экспертная оценка преобразования устной речи в письменную форму
Семантическое структурирование	Анализ структуры и полноты итогового документа
Производительность обработки	Измерение времени выполнения стадий и сравнение с ручной технологией подготовки учебных материалов

Для оценки качества автоматически сформированных учебных документов использовалась экспертная проверка, включавшая анализ сохранности содержания лекции, логической структуры документа, корректности терминологии, полноты представления определений, примеров и выводов. Дополнительно оценивались трудозатраты преподавателя при традиционной подготовке учебного пособия и при использовании разработанного интеллектуального конвейера.

Полученные в ходе экспериментальной апробации количественные и качественные показатели послужили исходными данными для анализа эффективности предложенной архитектуры.

5.2. ПРОГРАММА НЕЗАВИСИМОЙ ФУНКЦИОНАЛЬНОЙ ВЕРИФИКАЦИИ

После завершения разработки программного прототипа была выполнена его независимая функциональная проверка другим составом участников исследования.

Испытания проводились после переноса программного комплекса на отдельную виртуальную машину под управлением Ubuntu Server 24.04 LTS с конфигурацией 12 виртуальных процессорных ядер и 24 ГБ оперативной памяти. Проверка выполнялась без использования графического ускорителя, при локальном запуске системы автоматического распознавания речи Whisper и языковой модели Ollama.

В качестве входного материала использовалась дополнительная лекция продолжительностью 18 минут, ранее не применявшаяся при разработке программного прототипа.

Целью проверки являлось подтверждение возможности локального развертывания системы, воспроизводимости полного пользовательского сценария и корректного функционирования всех стадий интеллектуального конвейера в новой вычислительной среде.

В ходе испытаний контролировалось выполнение полного пользовательского сценария, включавшего загрузку лекционного материала, запуск обработки, последовательное выполнение всех стадий интеллектуального конвейера, формирование итогового документа, сохранение промежуточных результатов и завершение обработки без участия разработчика программной системы.

Дополнительно выполнялась функциональная проверка отдельных компонентов программного комплекса, включая работу веб-интерфейса, обработку аудиоматериала, автоматическую транскрибацию, контекстную коррекцию, нормализацию текста, структурирование документа, ведение истории задач и обработку нештатных ситуаций.

5.3. КРИТЕРИИ ВЕРИФИКАЦИИ

Для обеспечения сопоставимости результатов обеих серий исследований использовалась единая система критериев оценки, включающая функциональные, содержательные и эксплуатационные характеристики разработанной системы.

Таблица 5 - Критерии верификации

Группа критериев	Проверяемые характеристики
Функциональные	успешная загрузка материалов, запуск обработки, последовательное выполнение всех стадий интеллектуального конвейера, формирование и сохранение итогового документа
Содержательные	логическая связность текста, наличие аннотации, структуры документа, определений, примеров и выводов, отсутствие критических смысловых дефектов
Эксплуатационные	возможность локального развертывания, автономная работа без внешних сервисов, функционирование в CPU-only конфигурации, отображение хода обработки, сохранение промежуточных результатов, возможность повторного выполнения отдельных стадий

В рамках независимой функциональной проверки дополнительно контролировалось выполнение 21 функции программного комплекса, охватывающей полный жизненный цикл обработки лекционного материала - от загрузки исходного файла до формирования структурированного Markdown-документа и сохранения результатов обработки.

Представленная программа верификации позволила оценить как корректность функционирования отдельных компонентов интеллектуального конвейера, так и работоспособность предложенной архитектуры в целом. Результаты выполненных экспериментальных исследований и их интерпретация приводятся в следующем разделе.

6. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

6.1. РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТАЛЬНОЙ АПРОБАЦИИ

Экспериментальная апробация, выполненная разработчиком, подтвердила возможность практической реализации предложенной архитектуры интеллектуального конвейера автоматизированного формирования учебных документов. В ходе исследований были обработаны три университетские лекции общей продолжительностью около четырех часов, различавшиеся тематикой, качеством исходной записи и насыщенностью предметной терминологией.

На этапе автоматической транскрибации была достигнута высокая корректность распознавания речи: доля правильно распознанных слов превысила 99,8 %. Полученный результат подтвердил возможность использования локальной модели Whisper в качестве первого этапа интеллектуального конвейера при обработке длительных образовательных материалов.

Последующая контекстная коррекция позволила устранить значительную часть ошибок автоматического распознавания. Для лекций с умеренным использованием специальной терминологии эффективность исправления ошибок достигала практически полного уровня, тогда как при высокой терминологической насыщенности качество коррекции определялось полнотой контекстной информации и используемых промптов. При этом ложные исправления, изменяющие исходный смысл лекционного материала, в ходе апробации не выявлены.

Этап нормализации подтвердил возможность автоматического преобразования устной речи в письменный учебный текст. Из итоговых документов были устранены речевые повторы, слова-паразиты и особенности устного изложения, восстановлены пунктуация и абзацная структура при сохранении смысловой последовательности материала.

Особый интерес представляли результаты семантического структурирования. Во всех трех экспериментах система автоматически сформировала аннотацию, тематические разделы, определения и итоговые выводы. Экспертная оценка показала высокое соответствие структуры содержанию лекций: качество выделения разделов составило 4–5 баллов, качество автоматически сформированных аннотаций - 4–5 баллов, качество итоговых выводов - 4–5 баллов.

Таблица 6 Основные результаты экспериментальной апробации

Лекция	Длительность, мин	Качество транскрибации	Время обработки	Итоговый результат
1	69,9	> 99,8 %	около 35 мин	структурированный учебный документ

Лекция	Длительность, мин	Качество транскрибации	Время обработки	Итоговый результат
2	118,0	> 99,8 %	около 60 мин	структурированный учебный документ
3	70,2	> 99,8 %	около 35 мин	структурированный учебный документ

Одним из наиболее значимых результатов экспериментов стало подтверждение существенного сокращения трудоемкости подготовки учебных материалов. Сопоставление с традиционной технологией показало снижение трудозатрат приблизительно на 82 % при подготовке текстовой расшифровки и на 77 % при формировании структурированного учебного пособия. Часовая лекция, ручная обработка которой обычно требует около трех часов, автоматически преобразовывалась системой в течение 30–40 минут.

Таким образом, экспериментальная апробация подтвердила, что предложенный интеллектуальный конвейер обеспечивает автоматизированное преобразование лекционного материала не только в текстовую расшифровку, но и в структурированный учебный документ, существенно сокращая объем рутинной интеллектуальной работы преподавателя.

6.2. РЕЗУЛЬТАТЫ НЕЗАВИСИМОЙ ФУНКЦИОНАЛЬНОЙ ПРОВЕРКИ

Второй этап исследований был направлен на подтверждение воспроизводимости предложенной архитектуры и независимую проверку работоспособности программного прототипа после его переноса в новую вычислительную среду.

Программный комплекс был успешно развернут другим составом участников исследования на виртуальном сервере под управлением Ubuntu Server 24.04 LTS с использованием только центрального процессора без графического ускорителя. Все основные компоненты системы - модуль предварительной обработки, автоматическая транскрибация, локальная языковая модель, веб-интерфейс и модуль формирования документа - функционировали в штатном режиме.

В ходе проверки был полностью выполнен пользовательский сценарий обработки лекционного материала: загрузка исходного файла, запуск интеллектуального конвейера, последовательное выполнение всех стадий обработки, формирование структурированного Markdown-документа и сохранение результатов обработки. Дополнительно подтверждена корректная работа подсистем хранения промежуточных результатов, истории задач и повторного выполнения отдельных стадий обработки.

Независимая функциональная проверка охватила 21 функцию программного комплекса, включая загрузку файлов, предварительную обработку аудиоматериала, автоматическую транскрибацию, контекстную коррекцию, нормализацию текста, семантическое структурирование, формирование итогового документа и обработку нештатных ситуаций. Критических отказов, препятствующих завершению полного

пользовательского сценария, в ходе испытаний не выявлено.

Одновременно были зафиксированы отдельные особенности функционирования программного прототипа. При работе в CPU-only конфигурации продолжительность обработки существенно возрастала по сравнению с использованием графического ускорителя, однако это не влияло на полноту выполнения интеллектуального конвейера. Кроме того, при переключении вкладок браузера наблюдались отдельные недостатки отображения текущего состояния обработки, относящиеся к реализации пользовательского интерфейса и не оказывающие влияния на формирование итогового документа.

Полученные результаты подтвердили переносимость предложенной архитектуры и воспроизводимость её программной реализации. Независимое развертывание системы без участия первоначального разработчика показало, что предложенные архитектурные решения не зависят от конкретной вычислительной среды и могут использоваться в автономном локальном режиме без обращения к внешним облачным сервисам искусственного интеллекта.

6.3. ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Проведённые исследования позволяют сделать несколько выводов, имеющих значение не только для рассматриваемой задачи автоматизированного формирования учебных пособий, но и для проектирования человеко-машинных интеллектуальных систем в целом.

Прежде всего, экспериментальная апробация и независимая функциональная проверка подтвердили работоспособность предложенной архитектуры интеллектуального конвейера. Последовательное выполнение специализированных стадий обработки обеспечивает преобразование лекционного материала не только в автоматическую стенограмму, но и в структурированный учебный документ, включающий аннотацию, логически связанные разделы, определения, пояснения и выводы. Тем самым подтверждена основная гипотеза исследования о возможности построения специализированного интеллектуального конвейера, ориентированного на автоматизацию подготовки учебных материалов.

Вторым важным результатом является подтверждение переносимости архитектурного решения. Независимое развертывание программного прототипа другим составом участников показало, что предложенная архитектура не зависит от среды первоначальной разработки и допускает воспроизводимую реализацию на иной вычислительной платформе. При этом замена аппаратной конфигурации и повторное развертывание системы не потребовали изменения архитектурных принципов или переработки программных компонентов.

Проведённые исследования также подтвердили возможность полностью локального выполнения интеллектуального конвейера. Все стадии обработки - от автоматического распознавания речи до формирования итогового документа - выполнялись без использования внешних облачных сервисов искусственного интеллекта. Это обстоятельство представляется существенным для практического применения системы при обработке материалов ограниченного доступа, включая

корпоративные образовательные курсы, внутреннюю техническую документацию и материалы, содержащие конфиденциальную информацию. Одновременно результаты испытаний показали, что локальное выполнение на CPU-only конфигурации сопровождается увеличением времени обработки, однако не приводит к ухудшению функциональной полноты интеллектуального конвейера.

Полученные результаты позволяют сделать ещё один принципиальный вывод. Качество автоматически формируемого учебного документа не может оцениваться исключительно точностью автоматической транскрибации [7-12.]. Даже при высокой корректности распознавания речи итоговый результат определяется эффективностью последующих стадий интеллектуального конвейера - контекстной коррекции, нормализации письменной речи и семантического структурирования материала. Следовательно, критериями качества подобных систем должны выступать не только показатели распознавания речи, но и сохранение содержания лекции, логическая связность документа, полнота представления материала и соответствие авторскому педагогическому замыслу.

Вместе с тем проведённые исследования выявили характерный риск использования больших языковых моделей при подготовке учебных документов. Стремясь улучшить стиль изложения и устранить речевые повторы, языковая модель может одновременно сокращать отдельные пояснения, примеры и комментарии преподавателя. С точки зрения языковой обработки подобные преобразования повышают компактность текста, однако с педагогической точки зрения могут приводить к снижению полноты раскрытия темы. Следовательно, интеллектуальный конвейер должен ориентироваться не на максимальное сокращение текста, а на сохранение образовательной ценности исходной лекции.

Полученные результаты подтверждают и предложенную в работе концепцию распределения интеллектуальных функций между преподавателем и программной системой. Основной эффект автоматизации достигается не вследствие полного исключения преподавателя из процесса подготовки учебного пособия, а благодаря изменению характера его деятельности. Рутинные операции преобразования информации выполняются интеллектуальным конвейером, тогда как преподаватель сохраняет функции определения педагогических целей, экспертной оценки качества и принятия решения о готовности итогового документа. Именно такое распределение функций обеспечивает одновременно повышение производительности интеллектуального труда и сохранение авторского содержания учебного материала.

Наконец, следует отметить, что разработанный программный прототип представляет собой исследовательскую реализацию предложенной архитектуры и предназначен прежде всего для экспериментальной проверки её основных принципов. Для промышленного применения необходимы развитие подсистем хранения данных, поддержка многопользовательской работы, управление очередями обработки, версионность документов, встроенный

редактор результатов и интеграция с образовательными информационными системами.

6.4. ОГРАНИЧЕНИЯ ИССЛЕДОВАНИЯ

Несмотря на положительные результаты проведённых исследований, выполненная работа имеет ряд ограничений, которые следует учитывать при интерпретации полученных выводов.

Во-первых, экспериментальная апробация выполнялась на ограниченном количестве лекционных материалов. Несмотря на различие их продолжительности, качества записи и предметной специфики, полученная выборка не позволяет делать статистически обоснованные выводы относительно всех возможных образовательных дисциплин и сценариев использования системы.

Во-вторых, оценка качества автоматически формируемых учебных документов основывалась преимущественно на экспертном анализе преподавателей и разработчиков системы. В рамках настоящего исследования не проводился педагогический эксперимент с участием обучающихся, направленный на сравнение эффективности обучения по автоматически сформированным и традиционно подготовленным учебным материалам. Поэтому влияние предложенного подхода непосредственно на образовательные результаты требует отдельного исследования.

В-третьих, разработанный программный прототип ориентирован главным образом на обработку устной речи. Презентации, схемы, рисунки и другие визуальные материалы, сопровождающие лекцию, в текущей реализации практически не участвуют в формировании итогового документа. Между тем именно они нередко содержат существенную часть образовательной информации. Расширение архитектуры средствами мультимодальной обработки представляется одним из наиболее перспективных направлений дальнейших исследований.

Кроме того, качество автоматически формируемого документа в определённой степени зависит от характеристик исходной записи, особенностей предметной области и степени насыщенности материала специальной терминологией. Несмотря на использование контекстной коррекции, итоговый документ должен рассматриваться как проект учебного пособия, требующий обязательной экспертной проверки преподавателем перед публикацией или использованием в учебном процессе.

Наконец, в настоящем исследовании не проводилось сравнительного анализа нескольких альтернативных архитектур интеллектуальных конвейеров. Выполненная работа подтверждает реализуемость и практическую эффективность предложенной архитектуры, однако не ставила целью доказательство её превосходства над другими возможными подходами. Проведение сравнительных исследований различных архитектурных решений представляет самостоятельное направление дальнейших исследований.

В целом полученные результаты позволяют сделать вывод, что предложенная архитектура человеко-машинной интеллектуальной системы обеспечивает практическую основу для автоматизированного формирования учебных документов и подтверждает перспективность архитектурного подхода, основанного

на распределении интеллектуальных функций между преподавателем и программными компонентами системы. Вместе с тем дальнейшее развитие данного направления связано с расширением мультимодальных возможностей, совершенствованием механизмов контекстного обогащения, развитием средств экспертного взаимодействия и проведением полномасштабных педагогических исследований эффективности создаваемых учебных материалов.

7. ЗАКЛЮЧЕНИЕ И НАПРАВЛЕНИЯ ДАЛЬНЕЙШИХ ИССЛЕДОВАНИЙ

В работе рассмотрена задача повышения производительности и качества интеллектуального труда преподавателя при подготовке учебных пособий на основе лекционного материала. Показано, что традиционный процесс последовательной ручной переработки аудио- и видеозаписей лекций обладает высокой трудоёмкостью и плохо масштабируется, тогда как существующие средства автоматической транскрипции не обеспечивают получения структурированного учебного документа, пригодного для непосредственного использования в образовательном процессе.

Основным результатом исследования является предложенная архитектура человеко-машинной интеллектуальной системы, основанная на рациональном распределении интеллектуальных функций между преподавателем и программными компонентами. В разработанной архитектуре преподаватель определяет педагогические цели, формирует требования к структуре и содержанию учебного пособия, выполняет экспертную оценку результата и принимает решение о его готовности, тогда как интеллектуальный конвейер автоматически осуществляет транскрипцию, контекстную коррекцию, нормализацию письменной речи, семантическое структурирование и формирование проекта учебного документа. Такой подход позволяет заменить последовательную ручную переработку материала итерационным процессом экспертного управления автоматизированным формированием документа.

Для практической проверки предложенной архитектуры разработан исследовательский программный прототип, реализующий локальный интеллектуальный конвейер обработки лекционного материала. Архитектура построена на принципах модульности и независимости отдельных функциональных компонентов, благодаря чему используемые средства автоматического распознавания речи и большие языковые модели являются заменяемыми элементами реализации и не определяют архитектурную структуру системы.

Экспериментальная апробация программного прототипа на реальных университетских лекциях подтвердила работоспособность предложенного подхода и возможность существенного сокращения трудоёмкости подготовки учебных материалов при сохранении их содержательной целостности. Последующая независимая функциональная верификация после переноса системы в новую вычислительную среду подтвердила воспроизводимость программной реализации, возможность локального развертывания без использования внешних облачных сервисов и

корректное функционирование полного интеллектуального конвейера на иной аппаратно-программной платформе.

Проведённые исследования одновременно позволили определить существующие ограничения разработанного прототипа. К ним относятся отсутствие полноценной обработки мультимодального лекционного материала, зависимость качества результата от характеристик исходной записи и предметной области, необходимость обязательной экспертной проверки автоматически сформированного учебного документа перед публикацией, а также ограниченная функциональность исследовательского программного прототипа по сравнению с промышленными информационными системами.

Дальнейшее развитие исследования целесообразно сосредоточить по четырём взаимосвязанным направлениям. Первое связано с развитием мультимодальной обработки лекционного материала и совместным использованием аудиозаписи, презентаций, видеоряда и временных меток. Второе направление предполагает повышение содержательной достоверности автоматически формируемых документов за счёт использования предметных словарей, специализированных баз знаний и механизмов контроля сохранения терминов, числовых данных, примеров и иных существенных элементов исходной лекции. Третье направление связано с развитием средств человеко-машинного взаимодействия, включая поддержку версионности документов, формализацию экспертных замечаний преподавателя и автоматизированную повторную генерацию отдельных разделов учебного пособия. Наконец, четвёртое направление предполагает развитие исследовательского прототипа до уровня промышленной информационной системы, поддерживающей различные форматы публикации, коллективную работу пользователей и интеграцию с образовательными платформами, а также проведение педагогических исследований эффективности использования автоматически формируемых учебных материалов в реальном образовательном процессе.

Полученные результаты позволяют сделать вывод, что предложенный архитектурный подход обеспечивает практическую основу для построения человеко-машинных интеллектуальных систем, ориентированных на повышение производительности экспертной деятельности при сохранении ведущей роли человека в постановке целей, экспертной оценке и принятии решений. При этом задача автоматизированного формирования учебных пособий представляет собой одну из возможных областей применения предложенного архитектурного подхода, а оценка его применимости к другим видам интеллектуальной деятельности требует проведения самостоятельных исследований.

БИБЛИОГРАФИЯ

- [1] Dickey E., Bejarano A. GAIDE: A Framework for Using Generative AI to Assist in Course Content Development // Proceedings of the IEEE Frontiers in Education Conference (FIE). - 2024. - DOI: 10.1109/FIE61694.2024.10893132.
- [2] Dennison D. V., Ahtisham B., Chourasia K., Arora N., Singh R., Kizilcec R. F., Nambi A., Ganu T., Vashistha A. Shiksha Copilot: Teacher–AI Collaboration for Curating and Customizing Lesson Plans in Low-Resource Schools // Proceedings of the ACM on Human-Computer Interaction. - 2026. - DOI: 10.1145/3788074.

- [3] Baek I., Kim Y., Kwak N. PedaCo-Gen: Scaffolding Pedagogical Agency in Human-AI Collaborative Video Authoring // Extended Abstracts of the CHI Conference on Human Factors in Computing Systems. - 2026. - DOI: 10.1145/3772363.3798741.
- [4] UNESCO. Guidance for Generative AI in Education and Research. - Paris: UNESCO, 2023.
- [5] UNESCO. AI Competency Framework for Teachers. - Paris: UNESCO, 2024.
- [6] Трофимов В. В., Трофимова Е. В. Цифровая трансформация и этика высшего образования // Высшее образование в России. - 2021.
- [7] Retkowski F., Züfle M., Sudmann A., Pfau D., Watanabe S., Niehues J., Waibel A. Summarizing Speech: A Comprehensive Survey // Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing (EMNLP). - 2025. - P. 27275–27306. - DOI: 10.18653/v1/2025.emnlp-main.1388.
- [8] Koay J. J., Roustai A., Dai X., Liu F. A Sliding-Window Approach to Automatic Creation of Meeting Minutes // Proceedings of the NAACL Student Research Workshop. - 2021. - P. 68–75. - DOI: 10.18653/v1/2021.naacl-srw.10.
- [9] Retkowski F., Waibel A. From Text Segmentation to Smart Chaptering: A Novel Benchmark for Structuring Video Transcriptions // Proceedings of EACL 2024. - 2024. - P. 406–419. - DOI: 10.18653/v1/2024.eacl-long.25.
- [10] Kirstein F., Kumar S., Ruas T., Gipp B. Re-FRAME the Meeting Summarization SCOPE: Fact-Based Summarization and Personalization via Questions // Findings of EMNLP. - 2025. - DOI: 10.18653/v1/2025.findings-emnlp.1094.
- [11] Gupta A., Singh D., Cowan G. A., Kadhiresan N., Srivastava S., Sriraja Y., Mantri Y. K. AUTOSUMM: A Comprehensive Framework for LLM-Based Conversation Summarization // Proceedings of ACL Industry Track. - 2025. - DOI: 10.18653/v1/2025.acl-industry.35.
- [12] Mondal I. From Fluent to Useful: Generative AI That Models Purpose, Audience, and Presenter for Scientific Communication // Proceedings of ACL Student Research Workshop. - 2026. - DOI: 10.18653/v1/2026.acl-srw.87.
- [13] OpenAI. Whisper: Robust Speech Recognition via Large-Scale Weak Supervision. - 2022. ГКД <https://github.com/openai/whisper>
- [14] Ollama. Run Large Language Models Locally. - 2024.
- [15] Mihalcea R., Tarau P. TextRank: Bringing Order into Texts // Proceedings of EMNLP. - 2004. - P. 404–411.
- [16] ГОСТ Р ИСО/МЭК 25010–2015. Системная и программная инженерия. Требования и оценка качества систем и программного обеспечения (SQuaRE). Модели качества систем и программных продуктов.

Technology for Automated Transformation of Lecture Materials into Educational Textbooks Using Neural Speech Recognition and Large Language Models

S.N. Korenskaya, M.G. Zhabitskii, P.V. Anni, R.I. Khazeev, E.V. Markidonov, P.S. Ernst

Abstract -The paper addresses the automated transformation of lecture materials into structured educational documents using neural speech recognition and large language models. An architecture of a human-machine intelligent system is proposed, based on the functional distribution of tasks among the instructor, an intelligent processing pipeline, and contextual knowledge sources. A formal problem statement is developed; the input and target information objects, transformation requirements, and a sequence of specialized processing stages are defined, including automatic transcription, contextual correction, written-speech normalization, and semantic structuring of the material. The architecture is implemented as a local software prototype and verified through a two-stage study comprising experimental evaluation on university lectures and independent functional testing after deployment in a different computing environment. The results confirmed the operability of the proposed architectural approach, the feasibility of generating structured educational documents, a substantial reduction in the labor required to prepare teaching materials, and the reproducibility and portability of the software implementation. The limitations of the research prototype and the main directions for further development of the proposed architecture are identified.

Keywords: automated educational textbook generation, large language models, neural speech recognition, human-machine intelligent system, intelligent processing pipeline, context enrichment.

REFERENCES

- [1] Dickey E., Bejarano A. GAIDE: A Framework for Using Generative AI to Assist in Course Content Development // Proceedings of the IEEE Frontiers in Education Conference (FIE). - 2024. - DOI: 10.1109/FIE61694.2024.10893132.
- [2] Dennison D. V., Ahtisham B., Chourasia K., Arora N., Singh R., Kizilcec R. F., Nambi A., Ganu T., Vashistha A. Shiksha Copilot: Teacher-AI Collaboration for Curating and Customizing Lesson Plans in Low-Resource Schools // Proceedings of the ACM on Human-Computer Interaction. - 2026. - DOI: 10.1145/3788074.
- [3] Baek I., Kim Y., Kwak N. PedaCo-Gen: Scaffolding Pedagogical Agency in Human-AI Collaborative Video Authoring // Extended Abstracts of the CHI Conference on Human Factors in Computing Systems. - 2026. - DOI: 10.1145/3772363.3798741.
- [4] UNESCO. Guidance for Generative AI in Education and Research. - Paris: UNESCO, 2023.
- [5] UNESCO. AI Competency Framework for Teachers. - Paris: UNESCO, 2024.
- [6] Trofimov V. V., Trofimova E. V. Digital Transformation and Ethics of Higher Education // Higher Education in Russia. — 2021. — Vol. 30, No. 8–9. — P. 9–22. — DOI: 10.31992/0869-3617-2021-30-8-9-9-22.
- [7] Retkowski F., Züfle M., Sudmann A., Pfau D., Watanabe S., Niehues J., Waibel A. Summarizing Speech: A Comprehensive Survey // Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing (EMNLP). - 2025. - P. 27275–27306. - DOI: 10.18653/v1/2025.emnlp-main.1388.
- [8] Koay J. J., Roustai A., Dai X., Liu F. A Sliding-Window Approach to Automatic Creation of Meeting Minutes // Proceedings of the NAACL Student Research Workshop. - 2021. - P. 68–75. - DOI: 10.18653/v1/2021.naacl-srw.10.
- [9] Retkowski F., Waibel A. From Text Segmentation to Smart Chaptering: A Novel Benchmark for Structuring Video Transcriptions // Proceedings of EACL 2024. - 2024. - P. 406–419. - DOI: 10.18653/v1/2024.eacl-long.25.
- [10] Kirstein F., Kumar S., Ruas T., Gipp B. Re-FRAME the Meeting Summarization SCOPE: Fact-Based Summarization and Personalization via Questions // Findings of EMNLP. - 2025. - DOI: 10.18653/v1/2025.findings-emnlp.1094.
- [11] Gupta A., Singh D., Cowan G. A., Kadhiresan N., Srivastava S., Sriraja Y., Mantri Y. K. AUTOSUMM: A Comprehensive Framework for LLM-Based Conversation Summarization // Proceedings of ACL Industry Track. - 2025. - DOI: 10.18653/v1/2025.acl-industry.35.
- [12] Mondal I. From Fluent to Useful: Generative AI That Models Purpose, Audience, and Presenter for Scientific Communication // Proceedings of ACL Student Research Workshop. - 2026. - DOI: 10.18653/v1/2026.acl-srw.87.
- [13] OpenAI. Whisper: Robust Speech Recognition via Large-Scale Weak Supervision. - 2022. TKD <https://github.com/openai/whisper>
- [14] Ollama. Run Large Language Models Locally. - 2024.
- [15] Mihalcea R., Tarau P. TextRank: Bringing Order into Texts // Proceedings of EMNLP. - 2004. - P. 404–411.
- [16] GOST R ISO/IEC 25010–2015. Systems and Software Engineering. Systems and Software Quality Requirements and Evaluation (SQuaRE). System and Software Quality Models. Moscow: Standartinform, 2015. (In Russian).