

УДК 81'322.2

DOI 10.22213/2618-9763-2025-3-92-103

Е. В. Исаева, кандидат филологических наук, доцент

Пермский государственный национальный исследовательский университет, Пермь, Россия

П. С. Ермакова, студент

Национальный исследовательский технологический университет МИСИС, Москва, Россия

МЕТОДЫ И ТЕХНОЛОГИИ КОГНИТИВНОЙ АДАПТАЦИИ СПЕЦИАЛИЗИРОВАННЫХ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ ТЕКСТОВ ДЛЯ ЦЕЛЕВОЙ АУДИТОРИИ: ОБЗОР И ПЕРСПЕКТИВЫ РАЗРАБОТКИ ИНФОРМАЦИОННЫХ СИСТЕМ

В статье рассматривается проблема когнитивной адаптации специализированных социально-экономических текстов для целевой аудитории. Для решения проблемы выбран междисциплинарный подход на стыке когнитивно-дискурсивной лингвистики, терминоведения, информационных технологий, компьютерной лингвистики и обработки естественного языка. Приводится системное описание методов для автоматизации процесса адаптации текстов и создания комплексных информационных систем, генерирующих вторичные тексты, оптимизированные для восприятия конкретной группой реципиентов. Особое внимание уделяется вопросам когнитивной нагрузки и концептуальной плотности текста, а также семантической инвариантности вторичного текста. Дается сравнение методов с обоснованием эффективности их использования для различных этапов проектирования информационной системы, обеспечивающей автоматизацию данного процесса. Рассматриваются инструменты статистической оценки сложности текста, классификации целевой аудитории, общего упрощения текста и адаптации терминологии, реализованные на языке программирования Python. В качестве иллюстрации предполагаемого решения приводится пример полуавтоматического (с использованием нейронной сети с последующей экспертной корректировкой) преобразования фрагмента специализированного социально-экономического текста из научной статьи. В результатах предлагается блок-схема прототипа разрабатываемой информационной системы.

Представленное исследование может послужить основой для разработки информационной системы, которая может быть полезной государственным и муниципальным органам для улучшения коммуникации с гражданами, средствам массовой информации для повышения читаемости публикаций и вовлеченности аудитории, образовательным учреждениям для преподавания сложных социально-экономических дисциплин, бизнесу для объяснения финансовых, юридических и социальных аспектов своей деятельности и др. Предложенный подход является универсальным и может быть масштабирован на другие предметные области.

Ключевые слова: адаптация социально-экономических текстов; обработка естественного языка; когнитивная адаптация текста; восприятие информации; автоматизация обработки текста; генерация текста; симплификация текста; BART; T5.

Введение

В условиях цифровизации экономики, развития электронных сервисов и повышения прозрачности информации особое значение приобретает доступность и эффективность представления специализированных текстов – в частности социально-экономического содержания. Такие тексты зачастую обладают высокой терминологической насыщенностью, сложной логико-семантической структурой, ориентированной на профессиональную аудиторию. В результате этого возникают серьезные барьеры для восприятия текстов у людей, не обладающих профильной подготовкой. Это снижает вовлеченность граждан в социальные процессы, ограничивает доступ к важной информации, приводит к неспособности адресата реагировать на содержа-

ние текста и препятствует принятию взвешенных решений [1, 2].

Когнитивная адаптация – это процесс трансформации текста с учетом особенностей восприятия и уровня подготовки пользователя, способствующих «передаче основного содержания текста, его смысла в новых дискурсивных условиях» [3, с. 51]. В данном случае мы имеем дело с интердискурсивной адаптацией текста, которая определяется А. А. Дьяковой как «функциональное изменение параметров текста», включающей в себя: «изменение объема текста и количества текстообразующих концептов, повышение/понижение категориального статуса ключевых слов адаптируемого текста, изменение способа концептуализации содержания, изменение процентного соотношения свер-

нутых и развернутых структур»¹. Данный процесс становится необходимым инструментом в ответ на растущие требования к персонализации и доступности информации. Применение когнитивной адаптации особенно актуально в контексте трендов, связанных с развитием систем электронного управления, открытых данных, инклюзивного образования и цифровой грамотности. Современные вызовы, такие как повышение финансовой и правовой осведомленности населения, требуют, чтобы тексты были не только формально корректными, логически не содержащими противоречия, но и естественными в реальных обстоятельствах его восприятия, должны соответствовать фоновым знаниям целевой аудитории [4], а значит, быть когнитивно-доступными для реципиента.

В последние годы наблюдается активное развитие технологий обработки естественного языка (*NLP*), направленных на упрощение и перефразирование текста, оценку его сложности, автоматический семантический анализ и классификацию пользователей. Среди таких решений можно выделить языковые модели *T5*, *BART*, *PEGASUS*, инструменты оценки читаемости (*TextStat*, *Readability Transformers*), семантические модели (*Word2Vec*, *ConceptNet*)² [5] и классификаторы целевой аудитории (*FastText*, *BERT*) [6]. Однако несмотря на наличие этих компонентов большинство решений остаются разрозненными, не обеспечивают достаточную интердискурсивную адаптацию текста и не интегрированы в полноценные когнитивно-адаптивные системы.

При проектировании подобных систем особое внимание должно уделяться вопросам надежности, защищенности и интерпретируемости. Как подчеркивают Н. Т. Юнусов, С. В. Смирнов и С. А. Сакули: «...все большее количество государственных и коммерческих систем с высокими требованиями безопасности начинают использовать глубокие нейронные сети», а значит, устойчивость таких моделей к некорректным данным или внешним вмешательствам становится важным условием их применения в реальных условиях³.

В данной статье предпринимается попытка системного описания методов для автоматизации когнитивной адаптации текстов и определения перспектив создания комплексных информаци-

онных систем, способных адаптировать социально-экономические тексты под конкретные группы пользователей. Новизна подхода заключается в анализе и обосновании гибридной архитектуры, объединяющей различные технологии – от оценки сложности до семантической переформулировки, реинжиниринга текста и персонализированной подачи информации.

Целью статьи является обзор современных методов и технологий когнитивной адаптации специализированных текстов с акцентом на их применение в социально-экономической сфере и возможности построения информационных систем нового поколения. Для достижения этой цели решаются следующие задачи:

- 1) раскрытие теоретических основ когнитивной адаптации текста и ее значимости;
- 2) формализация задач, стоящих перед системами адаптации социально-экономических материалов;
- 3) анализ существующих моделей и инструментов, применяемых для упрощения, оценки сложности и семантической адаптации текстов;
- 4) обоснование гибридного подхода к когнитивной адаптации и описание архитектуры возможной информационной системы;
- 5) выявление проблем и ограничений, мешающих эффективной реализации таких систем, а также определение направлений их дальнейшего развития.

Реализация этих задач позволит не только сформулировать требования к когнитивно-адаптивным системам, но и обозначить путь к их практическому применению в образовании, медиа, управлении и цифровых сервисах.

Существующие подходы к когнитивной адаптации текста

Когнитивная адаптация текста представляет собой методику преобразования способа подачи информации, направленную на оптимизацию ее восприятия конкретной целевой аудиторией, с учетом ее когнитивных особенностей, уровня подготовки, фоновых знаний и предпочтений. Основная цель когнитивной адаптации заключается в облегчении восприятия, интерпретации и интериоризации информации без искажения ее первоначального смысла. Это особенно важно в контексте специализированных текстов, где высокая степень терминологической

¹ Дьякова А. А. Интердискурсивная адаптация текста : автореферат диссертации на соискание ученой степени кандидата филологических наук. Волгоград : Волгоградский государственный педагогический университет, 2009. 20 с.

² Sequence-to-Sequence Models for Automated Text Simplification / R. M. Botarleanu [et al.] // Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer, 2020. Vol. 12164 LNAI. Pp. 31–36.

³ Юнусов Н. Т., Смирнов С. В., Сакулин С. А. Состязательные примеры в задаче классификации изображений // Социально-экономическое управление: теория и практика. 2019. № 4 (39). С. 74–77.

насыщенности и сложность синтаксических конструкций могут существенно затруднить понимание даже для профессиональных читателей.

Восприятие текста различными группами пользователей варьируется в зависимости от их способности осмыслить, постичь содержание и смысл возникающей степени субъективности при интерпретации смысла, когнитивных особенностей, профессионального фона и уровня абстрактного мышления. Например, один и тот же текст в области макроэкономики будет восприниматься по-разному студентом экономического факультета, экспертом-финансистом и широкой аудиторией без профильного экономического образования. Следовательно, эффективная адаптация должна учитывать такие параметры, как уровень владения темой, наличие предварительных знаний, мотивацию и цели восприятия текста. Когнитивные факторы играют ключевую роль в процессе обработки текста. К ним относятся: внимание, память, восприятие, языковая компетенция, способность к логическому анализу и интерпретации значений в контексте. Поскольку содержание есть проекция текста (внешняя форма, представленная языковыми средствами, и внутренняя структура) на сознание адресата, а смысл – обратная проекция сознания реципиента на содержание текста^{1,2}, вполне закономерно, что один и тот же текст имеет тем большее количество смыслов, чем большее количество проекций сознания на текст он способен спроецировать. Отметим важность концептуальной плотности текста. По мнению А. И. Новикова, концепты кодируют знание и опыт человека, его восприятие действительности сквозь призму языка³, соответственно, чем она выше, тем выше когнитивная нагрузка на текст и ниже эффективность усвоения информации, особенно в случае ограниченного времени на ознакомление с материалом. Поэтому адаптация текста должна стремиться к снижению когнитивной нагрузки за счет упрощения структуры или рассредоточению плотности когнитивной нагрузки на отрезок текста через разъяснение терминов, использование иллюстраций и примеров, переформулировки сложных конструкций. При этом, как отмечается в работе «Принципы интердискурсивной адаптации текстов» (см. выше), когнитивная адаптация текста может повлечь его усложнение, например, на синтаксическом уровне.

При адаптации текста следует учитывать и «культурно-специфичные модели и сценарии, определяющие поведение и мировосприятие представителей разных лингвосообществ» [7, с. 266].

Специализированные социально-экономические тексты обладают рядом особенностей, затрудняющих их восприятие. *Во-первых*, они насыщены профессиональной терминологией, часто междисциплинарного характера. *Во-вторых*, в подобных текстах нередко содержатся абстрактные или обобщенные идеи, понимание которых требует определенного уровня подготовки и компетентности. *В-третьих*, логическая организация материала может быть довольно усложненной: текст состоит из объемных абзацев, включает многослойные рассуждения и обращения к дополнительным источникам. Все это предъявляет высокие требования к когнитивным способностям читателя. Согласно Н. И. Власенко и И. А. Толмачевой: «...трудность данной работы заключается в том, что большое количество экономических терминов появилось в русском языке не так давно, либо они были заимствованы из английского языка. Понятия, к которым относятся эти термины, отсутствуют в русском языке или формируются под влиянием российских экономических реалий» [8, с. 126].

Следовательно, адаптация текстов в области социально-экономических знаний с учетом когнитивных особенностей становится важной задачей на пути к обеспечению равного доступа к информации. Особенно это актуально в условиях цифровизации, когда важны оперативность восприятия и точность понимания.

При этом переработка специализированных материалов – это не просто упрощение, а многоэтапный процесс трансформации, направленный на сохранение смысловой точности при повышении степени понятности для более широкой аудитории.

В условиях стремительного увеличения объемов данных и усложнения профессиональной терминологии (особенно в социально-экономической сфере) адаптация является ключевым фактором успешной коммуникации между экспертами и широкой аудиторией. Когнитивная адаптация направлена на расширение целевой аудитории узкоспециализированных текстов без потери смысловой глубины. Профессиональные материалы изначально рассчитаны на узкий

¹ Первоисточник: Новиков А. И. Семантика текста и ее формализация. Москва : Наука, 1983. 213 с.

² Кирсанова И. В., Шияпова А. А. Проблема определения «субъективности» в процессах восприятия и понимания речевых произведений // Современные научные разработки. Инновационный аспект : сб. ст. Междунар. науч. конф. Санкт-Петербург, 2023. С. 41–47. DOI: 10.58351/230330.2023.48.48.002. EDN: YCZGNQ

³ Первоисточник: Краткий словарь когнитивных терминов / Е. С. Кубрякова, В. З. Демьянков, Ю. Г. Панкрац, Л. Г. Лузина. Москва : Филол. ф-т, МГУ им. Ломоносова, 1997. 245 с. (с. 90).

круг специалистов, однако современные тенденции открытого доступа к информации требуют их адаптации для студентов и начинающих специалистов, представителей смежных профессиональных областей, журналистов и популяризаторов науки, а также широкой аудитории, интересующейся данной темой. Ключевая задача когнитивной адаптации заключается в том, чтобы сохранить научную или профессиональную ценность контента, сделав его доступным для понимания без специальной подготовки [9].

Когнитивная адаптация направлена на снижение когнитивной нагрузки на читателя. Высокая когнитивная нагрузка – один из главных барьеров в восприятии специализированной информации. Длинные сложные предложения, большое количество новых терминов, абстрактные концепции – все это затрудняет усвоение текста. Адаптация направлена на структурирование материала, логическое упрощение, использование более ясных языковых конструкций и пояснений, что способствует снижению ментального напряжения при чтении.

При адаптации текстов важно сохранить семантическую точность. Одна из наиболее сложных задач при упрощении текста – не исказить исходный смысл. Особенно это актуально в экономике, праве, медицине и других областях, где значение терминов строго определено и любое упрощение может привести к потере нюансов или возникновению двусмысленности. Поэтому важным принципом адаптации является сохранение семантической точности – адаптированный текст должен передавать те же ключевые идеи, что и оригинал, но в более понятной форме¹.

Более того, необходимо учитывать персонализацию подачи информации в зависимости от уровня подготовки пользователя. Разные пользователи воспринимают информацию по-разному, и один и тот же текст может быть как слишком сложным, так и слишком упрощенным – в зависимости от читателя. Персонализация позволяет адаптировать тексты в соответствии с когнитивным профилем пользователя: его профессиональной сферой, уровнем образования, возрастом, опытом работы с конкретной темой. Это может быть реализовано через автоматическую кластеризацию и классификацию аудитории и генерацию разных версий одного и того же материала, что особенно актуально при разработке адаптивных информационных систем.

Для наглядной демонстрации описанных выше принципов когнитивной адаптации рассмотрим преобразование фрагмента специализированного социально-экономического текста из научной статьи «Новые социально-экономические тренды современного мира», сложного для восприятия широкой аудиторией без профильной подготовки.

Исходный текст (из аналитического экономического отчета):

Прежде всего, это старая методологическая проблема, идущая ещё от Вебера и Дюркгейма: дискуссия о роли эмпирии, теории и методов. И если проблема методологии в целом решена, и мы стоим на почве тектологии (синергетики, системности) и применяем это в практике эмпирических исследований, то проблема методов только обострилась. Так, современный ракурс проблемы метода Дюркгейма показал, что в чистом виде естественнонаучный метод не может быть применен, так как субъект-исследователь, желаем мы того или нет, влияет на объект социологического исследования, поэтому Дюркгеймом были сформулированы принципы непредвзятости. В дальнейшем в 30-х годах основатель политической эмпирической социологии доктор Джордж Гэллуп разработал принцип репрезентативности. А в нашей стране знаменитый профессор Л. Н. Коган, лидер уральской социологической школы, в своих лекциях ещё в 70–80-е годы предложил структуру научного (в противовес обыденному) знания. Структуру, которая нами, его учениками, была доработана.²

Данный фрагмент характеризуется высокой абстрактностью, насыщенностью терминами и сложной синтаксической организацией, что затрудняет его понимание непрофессиональной аудиторией. Ниже приведен адаптированный вариант, созданный с использованием трансформерной модели и последующей «ручной» корректировки.

Адаптированный вариант:

Учёные давно спорят о том, как лучше изучать общество: через теорию, практические наблюдения или специальные методы. Сегодня считается, что теоретическая основа у нас есть, но выбор конкретных методов остаётся сложной задачей. Например, ещё Дюркгейм показал, что исследователь не может быть полностью нейтральным – он всё равно влияет на результат. Позже Гэллуп предложил делать

¹ Overview of the Transformer-based Models for NLP Tasks / A. Gillioz [et al.] // Federated Conference on Computer Science and Information Systems. 2020. Pp. 179–183. DOI: 10.15439/2020F20

² Захарова Н. Л., Половникова Т. В. Новые социально-экономические тренды современного мира // Социально-экономическое управление: теория и практика. 2023. Т. 19, № 2. С. 43–48. DOI: 10.22213/2618-9763-2023-2-43-48. EDN: PQLFNA

выборки, чтобы мнение людей отражало реальную картину. В СССР профессор Коган учил, как отличать научное знание от обыденного, и его подход до сих пор используется.

В данной версии реализованы основные приемы когнитивной адаптации:

1) упрощение синтаксиса – сложные конструкции разбиты на более короткие и понятные предложения;

2) снижение терминологической нагрузки – специализированные термины («тектология», «репрезентативность», «естественнонаучный метод») заменены или объяснены простыми словами;

3) сохранение смысла – ключевые идеи оригинала сохранены и переданы в доступной форме;

4) расширение аудитории – адаптированный текст становится понятным для студентов, заинтересованных читателей и неспециалистов.

Данный пример иллюстрирует, как когнитивная адаптация может эффективно преобразовывать сложный академический материал, сделав его доступным широкой аудитории без искажения исходного содержания.

Ключевое преимущество адаптированных текстов в том, что они позволяют сохранить глубину содержания, значительно расширив при этом целевую аудиторию материала.

Современные технологии когнитивной адаптации: обзор методов и решений

Современные достижения в области обработки естественного языка (NLP) предоставляют эффективный инструментарий для реализации когнитивной адаптации специализированных текстов, позволяющий автоматизировать процессы упрощения, персонализации и повышения доступности информации, особенно в социально-экономической сфере, где уровень абстракции и терминологическая насыщенность затрудняют восприятие текстов широкой аудиторией. В рамках представленного далее обзора прорисованы ключевые направления, составляющие основу NLP-решений для когнитивной адаптации специализированных текстов.

Одним из часто упоминаемых в научной литературе направлений в области упрощения текстов является применение моделей перефразирования. Современные нейросетевые архитектуры на базе трансформеров, такие как T5, BART и PEGASUS, активно используются для преобразования сложных текстов в более понятные и доступные варианты¹. Например, мо-

дель T5, основанная на подходе Text2text (текст-в-текст), отличается высокой адаптивностью и точностью при работе с разнообразным содержанием. В статье «Использование языковых моделей T5 для задачи упрощения текста» приводится описание эксперимента, сравнивающего эффективность симплификации текста с различными архитектурами на базе T5, «в которых использовались разные значения параметра скорости обучения для каждого набора данных, размеры батча, а также исключение из обучения дополнительного набора данных» [10, с. 228]: *Absum*, *Paraphraser*, *Reference*, *SumGazeta* и *Base*². Авторы констатируют наилучший результат, достигнутый моделью *Reference* после ее дообучения, и делают выводы о важности подготовки качественных данных репрезентативного объема.

Модель BART, сочетающая черты автоэнкодера и генератора текста, обученная на двух задачах по удалению шума из текста: заполнение текста (восстановление замаскированных последовательностей токенов) и перестановка предложений (восстановление исходного порядка предложений), показала высокое значение вторичной генерации текста в условиях ограниченных ресурсов [11]. Она особенно эффективна в задачах перефразирования и составления кратких резюме, что делает ее подходящей для медицинских, юридических и аналитических документов. Модель была успешно адаптирована для решения задачи генерации описаний хода лечения в больнице для неврологических пациентов на основе выписных эпикризов. По результатам экспертной оценки 62 % автоматизированных резюме были отмечены как соответствующие стандартам здравоохранения, на основании чего авторы полагают, что этот метод может быть полезен в клинической практике [12].

PEGASUS, ориентированная на генерацию аннотаций и кратких описаний на основе абстрактной суммаризации (т. е. генерации нового текста с сохранением смысла), оказывается особенно полезной при обработке обширных экономических отчетов и аналитики. Как отмечается в ряде исследований, системы, построенные на основе предварительно обученной архитектуры кодирующе-декодирующего трансформера, такого как PEGASUS, благодаря алгоритму маскировки значимых предложений текста для достижения высокой степени абстракции

¹ Overview of the Transformer-based Models for NLP / Gillioz A. [et al.] // TasksProceedings of the Federated Conference on Computer Science and Information Systems. С. 179–183. DOI: 10.15439/2020F20

² URL: <https://huggingface.co/cointegrated/rut5-base-absum> (дата обращения: 04.08.2025) ; URL: <https://huggingface.co/cointegrated/rut5-base-paraphraser> (дата обращения: 04.08.2025) ; URL: https://huggingface.co/IlyaGusev/rut5_base_sum_gazeta (дата обращения: 04.08.2025) ; URL: <https://huggingface.co/sberbank-ai/ruT5-base> (дата обращения: 04.08.2025).

и дообучению для конкретной предметной области, хорошо справляются с решением проблем переноса и обобщения в узкоспециализированных текстах^{1, 2}. Исследования абстрактной суммаризации научных текстов, написанных на разных языках Индии (в т. ч. хинди и гуджарати), показывает кросс-языковой потенциал архитектуры PEGASUS [13]. Эти технологии могут быть интегрированы в цифровые системы для автоматической генерации доступных версий нормативных документов, методических рекомендаций и экспертных материалов.

Одним из обязательных этапов когнитивной адаптации является анализ читаемости оригинального текста. Это позволяет выявить трудные для восприятия фрагменты и определить, какие из них требуют трансформации. Среди инструментов, предназначенных для такой оценки, стоит выделить библиотеку *TextStat*³ (<https://textstat.org/>) на языке *Python*, которая рассчитывает классические показатели сложности, включая индекс Флеша, формулы *GunningFog* и *SMOG*⁴. В работе [14] отмечается, что *TextStat* позволяет рассчитывать статистические показатели поверхностных характеристик текста и на основе статистических показателей определять читабельность, сложность и академичность текста. Пакет *TextStat* стандартно используется для анализа доступности информации на веб-сайтах, как, например, приведено в исследовании доступности информации о COVID-19 на правительственных веб-сайтах [15].

Более продвинутый подход к когнитивной оценке текста реализован в проектах типа *Readability-Transformers* (<https://github.com/One-Theta/readability-transformers>), использующих трансформеры для анализа синтаксических и семантических характеристик текста с целью прогнозирования его воспринимаемости. Например, в работе⁵ предлагается модель на осно-

ве *RoBERTa* с взвешенным пулингом, который эффективно использует информацию о нескольких скрытых состояниях, благодаря чему превосходит традиционные формулы оценки читабельности текста. Не менее интересен проект *SimpleText*⁶, специализирующийся на автоматическом упрощении текстов и оценке их доступности – его концепции можно эффективно адаптировать под русскоязычный контекст.

Еще одним значимым направлением адаптации выступает семантическая переработка терминологии. Это особенно важно для документов, насыщенных профессиональными терминами. Такие модели, как *Word2Vec* (<https://www.tensorflow.org/text/tutorials/word2vec/>)⁷, *FastText* [16] и *GloVe* (<https://nlp.stanford.edu/projects/glove/>)⁸, создают векторные представления слов, что позволяет находить синонимичные и семантически близкие выражения, более понятные широкой аудитории. Согласно статье Роберта-Михай Ботарлеану, М. Даскалу, С. Кроссли, Д. Макнамара: «Лексическая база данных WordNet содержит синсеты (т. е. наборы синонимов), которые могут быть использованы для создания пар синонимов путем пересечения синсетов различных словарных терминов» [17, с. 32]. В свою очередь, *ConceptNet*⁹ – «многоязычная база знаний, предназначенная для представления слов, фраз и их общих взаимосвязей» [18], что особенно ценно при объяснении или замене абстрактных терминов. Такие подходы позволяют сохранять точность при переводе сложных формулировок на более доступный язык без искажения сути.

Персонализация – еще один значительный элемент когнитивной адаптации. Для эффективной настройки текстов под конкретного пользователя важно точно определить уровень его подготовки. Здесь применимы методы классификации, основанные на таких моделях, как

¹ PEGASUS: Pre-training with Extracted Gap-sentences for Abstractive Summarization / J. Zhang [et al.] // 37th International Conference on Machine Learning, ICML 2020. International Machine Learning Society (IMLS), 2019. Pp. 11265–11276.

² Также см.: Dai S. et al. BDKG at MEDIQA 2021: System Report for the Radiology Report Summarization Task // Proceedings of the 20th Workshop on Biomedical Language Processing. Stroudsburg, PA, USA: Association for Computational Linguistics, 2021. Pp. 103–111 ; Kim S. Using Pre-Trained Transformer for Better Lay Summarization // Proceedings of the First Workshop on Scholarly Document Processing. Stroudsburg, PA, USA: Association for Computational Linguistics, 2020. Pp. 328–335 ; Jo S.G. et al. Learning Cluster Patterns for Abstractive Summarization // IEEE Access. Institute of Electrical and Electronics Engineers Inc., 2023. Vol. 11. Pp. 146065–146075.

³ URL: <https://textstat.org/> (дата обращения: 04.08.2025)

⁴ См. Hüning M., Hüning M. TextStat Simple text analysis tool // Dutch Linguistics, Free University of Berlin. 2005.

⁵ Rating Ease of Readability using Transformers / V. S. Alaparthi [et al.] // 2022 14th International Conference on Computer and Automation Engineering (ICCAE). IEEE, 2022. Pp. 117–121. DOI: 10.1109/ICCAE55086.2022.9762413

⁶ URL: <https://simpletext.app/> (дата обращения: 04.08.2025).

⁷ Также см. Mustafa G. et al. Multi-label classification of research articles using Word2Vec and identification of similarity threshold // Sci Rep. 2021. Т. 11, № 1. С. 21900.

⁸ Также см. Brown K.S. et al. Investigating the Extent to which Distributional Semantic Models Capture a Broad Range of Semantic Relations // Cogn Sci. 2023. Т. 47, № 5.

⁹ URL: <https://conceptnet.io/> (дата обращения: 04.08.2025).

*FastText*¹, *BERT*² и *RoBERTa* [19]. *FastText* обеспечивает быструю и точную категоризацию на основе текстовых данных, что удобно при массовой обработке пользовательских запросов. *BERT* и *RoBERTa*, благодаря своей глубокой архитектуре, позволяют обучить систему тонкому различию в уровнях профессиональной компетентности, что открывает возможность генерации адаптированных версий одного и того же текста для разных категорий читателей.

Когнитивная адаптация специализированных текстов базируется на интеграции современных *NLP*-технологий, каждая из которых выполняет свою функцию: от анализа сложности до генерации персонализированных текстов. Системы, сочетающие эти технологии, способны существенно повысить доступность информации, сохранив ее точность и релевантность для разных категорий пользователей.

Результаты. Проектирование архитектуры решения для когнитивной адаптации экономических текстов

Наиболее перспективным направлением в развитии когнитивной адаптации специализированных текстов сегодня выступает использование гибридных подходов, объединяющих в рамках одной информационной системы сразу несколько технологических решений. Такой интегративный формат позволяет не только учитывать множество аспектов восприятия текста, но и значительно повысить точность, гибкость и персонализацию адаптации по сравнению с применением отдельных, изолированных инструментов. Гибридные подходы к адаптации текста базируются на поэтапной координации нескольких функциональных модулей. Далее рассмотрим модель такой гибридной системы и набор методов, которые могут найти применение на разных этапах обработки экономического текста.

Первая фаза предполагает автоматическое определение уровня сложности исходного материала. С использованием таких инструментов, как *TextStat* и *Readability-Transformers*, осуществляется оценка читаемости и выделение фрагментов, затрудняющих восприятие. Это позволяет установить исходный уровень сложности текста, который служит отправной точкой для последующих преобразований [20].

На следующем этапе система проводит анализ предполагаемой аудитории. С помощью мо-

делей машинного обучения, таких как *FastText*, *BERT* или *RoBERTa*, происходит классификация пользователей по степени подготовки. Эти данные позволяют подобрать наиболее эффективную стратегию адаптации, соотнося ее с когнитивными характеристиками конкретной целевой группы.

После этого активируется модуль текстового упрощения. Современные языковые трансформеры, включая *T5*, *BART* и *PEGASUS*, производят переработку текста, улучшая его читаемость за счет упрощения синтаксических конструкций, устранения сложных оборотов и увеличения прозрачности изложения. При этом сохраняется логика повествования и передается основное содержание без искажений.

Дополнительный уровень адаптации обеспечивается через семантическую трансформацию терминов. В процессе генерации упрощенной версии система может использовать такие модели, как *Word2Vec*, *FastText*, *GloVe* или семантические сети, такие как *ConceptNet*, для автоматической замены сложных, узкоспециализированных терминов на более доступные аналоги. Это позволяет избежать потери смысла и при этом сделать текст значительно понятнее для неспециалистов.

Использование гибридного подхода обеспечивает целостную обработку текста – от анализа сложности и понимания аудитории до глубокой переформулировки и семантической адаптации. Такой подход не только технологически эффективен, но и открывает реальные возможности для создания масштабируемых и интеллектуальных систем адаптации текстов, особенно востребованных в образовании, цифровом госуправлении, медиа и других сферах с высокой социальной значимостью.

Информационная система, предназначенная для когнитивной адаптации специализированных материалов, представляет собой многослойную архитектуру, направленную на индивидуализированное упрощение и преобразование сложных текстов в более доступные для восприятия формы. В ее основе лежит интегративный подход, объединяющий функциональность современных языковых моделей, методы семантического анализа и технологии пользовательской классификации.

Процесс начинается с работы модуля первичной обработки, который получает исходный текст, проводит автоматическую оценку его

¹ Bag of Tricks for Efficient Text Classification / Joulin A. [et al.] // 15th Conference of the European Chapter of the Association for Computational Linguistics, EACL 2017 – Proceedings of Conference. Association for Computational Linguistics (ACL), 2016. Vol. 2. Pp. 427–431.

² URL: <https://github.com/google-research/bert> (дата обращения: 04.08.2025).

сложности с применением показателей читаемости и анализирует его лексико-семантическую структуру. На этом этапе определяется потенциальная аудитория, что позволяет сформировать стартовые параметры адаптационного процесса. Затем активируется модуль пользовательской классификации, задача которого – оценить уровень подготовки конкретного пользователя или целевой группы на основе соответствующих характеристик. С помощью моделей машинного обучения (например, *FastText* или *BERT*), определяется цифровой когнитивный профиль (цифровой двойник) пользователя и подбирается соответствующая стратегия для адаптации текста.

На основе этих данных активируется модуль генерации адаптированной версии, который формирует упрощенный и когнитивно доступный вариант текста. Он использует модели типа *T5*, *BART* или *PEGASUS* для перефразирования, а также подключает семантические ресурсы (*Word2Vec*, *ConceptNet*) для замены сложных терминов на более понятные эквиваленты. При этом соблюдается баланс между упрощением и сохранением точности содержания. За это отвечает модуль сравнения и верификации, контролирующий соответствие адаптированной версии исходному смыслу и обеспечивающий корректность терминологических замен.

Важным элементом системы выступает пользовательский интерфейс, обеспечивающий интерактивное взаимодействие с контентом. Он предоставляет возможность выбора желаемого уровня сложности текста, предлагает альтернативные формулировки, поясняет значение профессиональных терминов и позволяет легко переключаться между оригинальной и адаптированной версиями. Благодаря такому функционалу система становится универсальной и применимой в разнообразных ситуациях – от образовательной среды до профессионального консультирования.

Для наглядного представления структуры, компонентов и используемых методов далее приведем блок-схему системы, включающей ключевые модули обработки текста, модели машинного обучения и механизмы взаимодействия с пользователем (рисунок).



Источник: выполнен авторами.

Блок-схема прототипа информационной системы, способной адаптировать специализированные тексты с учетом особенностей целевой аудитории

Flow chart of the information system prototype for adapting specialized texts with regard to the target readership characteristics

Как указано в работе *A hybrid model of complexity estimation: Evidence from Russian legal texts* (см. выше), ключевое преимущество гибридной архитектуры заключается в ее способности точно подстраивать процессы адаптации под конкретного пользователя. Такая настройка достигается за счет комбинирования различных языковых моделей и аналитических алгоритмов, что позволяет одновременно учитывать характеристики исходного текста и особенности восприятия целевой аудитории. Такая гибридная система обладает высокой масштабируемостью: ее можно адаптировать для разных языков, тематических областей и уровней пользовательской подготовки.

Кроме того, способность работать в режиме реального времени делает данное решение особенно ценным для онлайн-обучающих платформ, интеллектуальных справочных систем и цифровых сервисов с широкой пользовательской аудиторией, где критичны скорость и точность подачи информации. Особенно эффективны такие решения в сферах, где требуется донести одну и ту же информацию до разных категорий пользователей – от школьников и студентов до профессиональных аналитиков, юристов или госслужащих. В контексте публичной политики они способствуют обеспечению прозрачности экономических и правовых документов, делая их доступными для широкой аудитории и повышая общий уровень информированности общества.

Обсуждение предполагаемых решений

Несмотря на значительный прогресс в области когнитивной адаптации текстов и стремительное развитие технологий обработки естественного языка, существующие решения все еще далеки от универсальности и безошибочности. Разработка и внедрение систем когнитивной адаптации сталкиваются с рядом важных технических и методологических ограничений.

Наиболее существенным препятствием остается ограниченная универсальность современных языковых моделей. Такие модели, как *T5*, *BART* и *PEGASUS*, демонстрируют высокую эффективность при работе с обобщенными корпусами текстов. Однако результаты свидетельствуют о снижении их точности и надежности при адаптации узкоспециализированных материалов, таких как экономические отчеты, юридические заключения или научные статьи. Это наблюдение подтверждает предположение, что большинство моделей обучаются на общих данных и не учитывают контекстуальные особен-

ности терминологии, что увеличивает риск искажения смысла при упрощении или перефразировании.

Еще одной яркой проблемой является нарушение семантической точности. При автоматическом упрощении текста модели часто выбирают формально корректные синонимы, которые в конкретном контексте оказываются неуместными или двусмысленными. Это особенно критично в сферах, где точность формулировок имеет юридическое или финансовое значение [21].

Важное наблюдение, подтвержденное в предыдущих исследованиях, касается нехватки качественных русскоязычных обучающих данных. Большинство передовых моделей разрабатывались для англоязычного пространства, в то время как существующие русскоязычные решения, особенно в таких тематиках, как социальная политика, налогообложение и макроэкономика, существенно уступают по полноте и зрелости. Это требует дополнительных ресурсов для адаптации моделей под лингвистические особенности русского языка и профессиональную терминологию.

Неожиданным и важным препятствием является сложность классификации целевой аудитории. Для точного подбора уровня упрощения требуются размеченные обучающие выборки, которые часто отсутствуют на практике. Более того, уровень подготовки пользователя может быть скрытым или ситуативным, что делает персонализацию затруднительной. При этом использование пользовательских данных порождает вопросы этики, конфиденциальности и информационной безопасности.

Следует особо отметить отсутствие устоявшихся стандартов когнитивной адаптации. На данный момент не существует общепринятой методологии, определяющей, каким должен быть адаптированный текст для различных типов аудитории. Это затрудняет оценку качества адаптации, сравнение результатов различных систем и разработку унифицированных решений. Более того, восприятие текста всегда субъективно: то, что одному пользователю кажется понятным, другому может показаться сложным.

Технические ограничения также заслуживают внимания. Интеграция современных языковых моделей, особенно в условиях обработки в реальном времени, требует значительных вычислительных ресурсов. Это снижает масштабируемость решений и ограничивает их применение в организациях с ограниченной инфраструктурой.

Таким образом, представленные результаты подчеркивают, что несмотря на достигнутые успехи, технологии когнитивной адаптации находятся на стадии становления. Для преодоления существующих барьеров необходим комплексный междисциплинарный подход, включающий сотрудничество специалистов в области вычислительной лингвистики, когнитивной психологии, машинного обучения и экспертов по предметной области. Только при тесной интеграции этих направлений возможно создание эффективных, точных и этически безопасных систем когнитивной адаптации специализированных текстов.

В заключение следует подчеркнуть, что адаптация социально-экономических текстов представляет собой актуальное направление на стыке лингвистики, когнитивных наук и искусственного интеллекта. В условиях роста информационной нагрузки, усложнения профессионального языка и стремления к более широкой коммуникации, когнитивная адаптация становится ключевым инструментом обеспечения доступности знаний и прозрачности информации.

Выводы

Рассмотренные в статье методы – от оценки читаемости и перефразирования до семантической подстройки терминологии и классификации аудитории – демонстрируют высокий потенциал для автоматизации процесса адаптации. Современные языковые модели, такие как *T5*, *BART*, *PEGASUS*, а также средства семантического анализа и классификации, уже сегодня позволяют создавать интеллектуальные системы, способные трансформировать сложные тексты без потери содержания и смысловой точности.

Наиболее эффективным подходом является интеграция этих решений в виде гибридных систем, способных учитывать особенности как текста, так и читателя. Такие системы могут применяться в образовании, науке, медиа, госуправлении и других сферах, где важно обеспечить понятную подачу сложной информации.

Однако на пути широкого внедрения остаются значимые вызовы: недостаток русскоязычных специализированных данных, риски семантических искажений, необходимость персонализации и этические аспекты обработки пользовательской информации. Преодоление этих ограничений требует дальнейших исследований и разработки устойчивых методологических подходов к когнитивной адаптации.

В перспективе возможно создание полноценных адаптивных информационных систем

нового поколения, которые будут автоматически подстраивать тексты под индивидуальные когнитивные профили пользователей, обеспечивая тем самым равный доступ к информации и способствуя росту цифровой и профессиональной грамотности в обществе.

Библиографические ссылки

1. Хафизова Э. И. Адаптация сложности текстов и текстовых заданий // Молодой ученый. 2021. № 354. С. 266–268.
2. Демьянков В. З. Приемлемость, уместность и адаптация текста // Вопросы когнитивной лингвистики. 2019. № 4. С. 4–19. DOI:10.20916/1812-3228-2019-4-9-19. EDN: TXHKWL
3. Ионов С. В. Принципы интердискурсивной адаптации текстов // Научный вестник Воронежского государственного архитектурно-строительного университета. Серия: Современные лингвистические и методико-дидактические исследования. 2006. № 5. С. 50–59.
4. Демьянков В. З. Приемлемость, уместность и адаптация текста // Вопросы когнитивной лингвистики. 2019. № 4. С. 4–19. DOI: 10.20916/1812-3228-2019-4-9-19. EDN: TXHKWL
5. Speer R., Havasi C. ConceptNet 5: A Large Semantic Network for Relational Knowledge. Springer, Berlin, Heidelberg, 2013. Pp. 161–176. DOI: 10.1007/978-3-642-35085-6_6
6. Imperial J. M. BERT Embeddings for Automatic Readability Assessment // ACL Anthology. 2021. Pp. 611–618. URL: <https://aclanthology.org/2021.ranlp-1.69/> (дата обращения: 11.07.2025).
7. Хунцзюнь Д. Роль когнитивной лингвистики в понимании процессов восприятия и интерпретации текста при переводе // Управление образованием: теория и практика. 2024. Т. 14, № 1–2. С. 265–273. DOI: 10.25726/c3852-9347-4601-ево
8. Vlasenko N. I., Tolmacheva I. A. Translation features of technical texts (examples of economic texts) // Russian Linguistic Bulletin. 2020. No 4 (24). Pp. 126–128. DOI: 10.18454/RULB.2020.24.4.20. EDN: WKYBSJ
9. Anggawijaya M. H., Adika D. Enhancing Target Text Comprehension for Lay Audience through Paraphrasing // JurnalHumaya: Jurnal Hukum, Humaniora, Masyarakat, dan Budaya. Universitas Terbuka, 2023. Vol. 3, no. 1. Pp. 1–14. DOI: 10.33830/humaya.v3i1.4282
10. Умарова Д. З. Принципы отбора и адаптации текстов по специальности // Бюллетень науки и практики. 2023. Т. 9, № 11. С. 371–375. DOI: 10.33619/2414-2948/96/48/. EDN: CWVNCC
11. Васильев Д. Д., Пятаева А. В. Использование языковых моделей T5 для задачи упрощения текста // Программные продукты и системы. 2023. Т. 20. С. 228–236. DOI: 10.15827/0236-235X.142.228-236. EDN: AYXLBW
12. Chernyshev D., Dobrov B. Investigating the Pre-Training Bias in Low-Resource Abstractive Summarization // IEEE Access. 2024. Vol. 12. Pp. 47219–47230. DOI: 10.1109/ACCESS.2024.3379139

13. A Method to Automate the Discharge Summary Hospital Course for Neurology Patients / V. C. Hartman [et al.] // Journal of the American Medical Informatics Association. Oxford University Press, 2023. Vol. 30, no. 12. Pp. 1995–2003. DOI: <https://doi.org/10.1093/jamia/ocad177>

14. Implementing Deep Learning-Based Approaches for Article Summarization in Indian Languages / R. Tangsali [et al.] // CEUR Workshop Proc. CEUR-WS, 2022. Vol. 3395. Pp. 449–453. URL: <https://ceur-ws.org/Vol-3395/T6-10.pdf> (дата обращения: 11.07.2025).

15. Kostadimas D., Kermanidis K. L., Andronikos T. Exploring the Effectiveness of Shallow and L2 Learner-Suitable Textual Features for Supervised and Unsupervised Sentence-Based Readability Assessment // Applied Sciences. 2024. Vol. 14, no. 17. Pp. 7997. DOI: 10.3390/app14177997

16. Improving access to COVID-19 information by ensuring the readability of government websites / T. Serry [et al.] // Health Promotion Journal of Australia. John Wiley and Sons Inc. 2023. Vol. 34, no. 2. Pp. 595–602. DOI: 10.1002/hpja.610

17. FastText.zip: Compressing text classification models / A. Joulin [et al.]. 2016. DOI: <https://doi.org/10.48550/arXiv.1612.03651>

18. Sequence-to-Sequence Models for Automated Text Simplification / R. M. Botarleanu [et al.] // Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer, 2020. Vol. 12164 LNAI. Pp. 31–36. DOI: 10.1007/978-3-030-52240-7_6

19. Rajabi Z., Valavi M. R., Hourali M. A Context-Based Disambiguation Model for Sentiment Concepts Using a Bag-of-Concepts Approach // Cognit Comput. 2020. Vol. 12, no. 6. Pp. 1299–1312. DOI: 10.48550/arXiv.2008.03020

20. RoBERTa: A Robustly Optimized BERT Pre-training Approach / Y. Liu [et al.]. 2019. URL: <https://arxiv.org/pdf/1907.11692> (дата обращения: 11.07.2025).

21. Blinova O., Tarasov N. A hybrid model of complexity estimation: Evidence from Russian legal texts // Front ArtifIntell. 2022. Vol. 5. DOI: 10.3389/frai.2022.1008530

References

1. Khafizova E.I. [Adaptation of text complexity and text tasks]. *Molodoy uchenyy*, 2021, pp. 266–268. (in Russ.).

2. Demyankov V.Z. [Acceptability, appropriateness, and adaptation of text]. *Issues of Cognitive Linguistics*, 2019, pp. 4–19. (in Russ.). DOI: 10.20916/1812-3228-2019-4-9-19. EDN: TXHKWL

3. Ionova S.V. [The Principles of Interdiscourse Texts Adaptation]. *Modern Linguistic and Methodical-and-Didactic Researches*, 2006, pp. 50–59. (in Russ.).

4. Demyankov V.Z. [Acceptability, appropriateness, and adaptation of text]. *Issues of Cognitive Linguistics*, 2019, pp. 4–19. (in Russ.). DOI: 10.20916/1812-3228-2019-4-9-19. EDN: TXHKWL

5. Speer R., Havasi C. ConceptNet 5: A Large Semantic Network for Relational Knowledge. Springer, Berlin, Heidelberg, 2013, pp. 161–176. (in Engl.). DOI: 10.1007/978-3-642-35085-6_6

6. Imperial J.M. BERT Embeddings for Automatic Readability Assessment. ACL Anthology, 2021, pp. 611–618. (in Engl.). Available at: <https://aclanthology.org/2021.ranlp-1.69/> (accessed 11.07.2025).

7. Hongjun D. [The role of cognitive linguistics in understanding the processes of perception and interpretation of text in translation]. *Education management review*, 2024, vol. 14, pp. 265–273 (in Russ.). DOI: 10.25726/C3852-9347-4601-E

8. Vlasenko N.I., Tolmacheva I.A. [Translation features of technical texts (examples of economic texts)]. *Russian Linguistic Bulletin*, 2020, no. 4 (24), pp. 126–128. (in Engl.). DOI: 10.18454/RULB.2020.24.4.20. EDN: WKYBSJ

9. Anggawijaya M.H., Adika D. Enhancing Target Text Comprehension for Lay Audience through Paraphrasing. *JurnalHumaya: Jurnal Hukum, Humaniora, Masyarakat, dan Budaya*. Universitas Terbuka, 2023, vol. 3, no. 1, pp. 1–14. (in Engl.). DOI: 10.33830/humaya.v3i1.4282

10. Umarova D.Z. [Principles for selection and adaptation of texts by specialty]. *Byulleten of science and practice*, 2023, vol. 9, pp. 371–375 (in Russ.). DOI: 10.33619/2414-2948/96/48/. EDN: CWNCC

11. Vasiliev D.D., Pyataeva A.V. [T5 language models for text simplification]. *Software & Systems*, 2023, vol. 20, pp. 228–236. (in Russ.). DOI: 10.15827/0236-235X.142.228-236. EDN: AYXLBW

12. Chernyshev D., Dobrov B. Investigating the Pre-Training Bias in Low-Resource Abstractive Summarization. IEEE Access, 2024, vol. 12, pp. 47219–47230. (in Engl.). DOI: 10.1109/ACCESS.2024.3379139

13. Hartman V.C. et al. A Method to Automate the Discharge Summary Hospital Course for Neurology Patients. Journal of the American Medical Informatics Association. Oxford University Press, 2023, vol. 30, no. 12, pp. 1995–2003. (in Engl.). DOI: <https://doi.org/10.1093/jamia/ocad177>

14. Tangsali R. et al. Implementing Deep Learning-Based Approaches for Article Summarization in Indian Languages. CEUR Workshop Proc. CEUR-WS, 2022, vol. 3395, pp. 449–453. (in Engl.). Available at: <https://ceur-ws.org/Vol-3395/T6-10.pdf> (accessed 11.07.2025).

15. Kostadimas D., Kermanidis K.L., Andronikos T. Exploring the Effectiveness of Shallow and L2 Learner-Suitable Textual Features for Supervised and Unsupervised Sentence-Based Readability Assessment. Applied Sciences, 2024, vol. 14, no. 17, p. 7997. (in Engl.). DOI: 10.3390/app14177997

16. Serry T. et al. Improving access to COVID-19 information by ensuring the readability of government websites. Health Promotion Journal of Australia. John Wiley and Sons Inc, 2023, vol. 34, no. 2, pp. 595–602. (in Engl.). DOI: 10.1002/hpja.610

17. Joulin A. et al. FastText.zip: Compressing text classification models. 2016. (in Engl.). DOI: <https://doi.org/10.48550/arXiv.1612.03651>

18. Botarleanu R.M. et al. Sequence-to-Sequence Models for Automated Text Simplification. Lecture

Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer, 2020, vol. 12164 LNAI, pp. 31-36. (in Engl.). DOI: 10.1007/978-3-030-52240-7_6

19. Rajabi Z., Valavi M.R., Hourali M. A Context-Based Disambiguation Model for Sentiment Concepts Using a Bag-of-Concepts Approach. *Cognit Comput*, 2020, vol. 12, no. 6, pp. 1299-1312. (in Engl.). DOI: 10.48550/arXiv.2008.03020

20. Liu Y. et al. RoBERTa: A Robustly Optimized BERT Pretraining Approach, 2019. (in Engl.). Available at: <https://arxiv.org/pdf/1907.11692> (accessed 11.07.2025).

21. Blinova O., Tarasov N. A hybrid model of complexity estimation: Evidence from Russian legal texts. *Front ArtifIntell*, 2022, vol. 5. (in Engl.). DOI: 10.3389/frai.2022.1008530

E. V. Isaeva, Candidate of Philological Science, Associate Professor
Perm State University, Perm, Russia

P. S. Ermakova, Student

National University of Science and Technology MISIS, Moscow, Russia

METHODS AND TECHNIQUES FOR COGNITIVE ADAPTATION OF SPECIALIZED SOCIO-ECONOMIC TEXTS FOR TARGET AUDIENCES: REVIEW AND PROSPECTS FOR THE DEVELOPMENT OF INFORMATION SYSTEMS

The article discusses the problem of cognitive adaptation of specialised socio-economic texts for the target audience. To solve the problem, an interdisciplinary approach was chosen at the intersection of cognitive-discursive linguistics, terminology, information technology, computational linguistics, and natural language processing. We provide a systematic description of methods for automating the process of text adaptation and creating complex information systems that generate secondary texts optimised for perception by a specific group of recipients. Particular attention is paid to issues of cognitive load and conceptual density of the text, as well as the semantic invariance of the secondary text. We compare methods and justify their effectiveness for different stages of designing an information system that automates this process. We examine tools for statistical assessment of text complexity, target audience classification, general text simplification, and terminology adaptation, implemented in the Python programming language. As an illustration of the proposed solution, an example is given of the semi-automatic (using a neural network with subsequent expert correction) conversion of a fragment of a specialized socio-economic text from a scientific article. The results propose a flowchart of the prototype of the information system under development.

The presented study can serve as a basis for the development of an information system that can be useful to state and municipal authorities for improving communication with citizens, to the media for increasing the readability of publications and audience engagement, to educational institutions for teaching complex socio-economic disciplines, businesses to explain the financial, legal, and social aspects of their activities, etc. The proposed approach is universal and can be scaled to other subject areas.

Keywords: adaptation of socio-economic texts; natural language processing; cognitive adaptation of text; information perception; automated text processing; text generation; text simplification; BART; T5.

Получена: 07.07.2025

ГРНТИ 16.31.21

Образец цитирования

Исаева Е. В., Ермакова П. С. Методы и технологии когнитивной адаптации специализированных социально-экономических текстов для целевой аудитории: обзор и перспективы разработки информационных систем // Социально-экономическое управление: теория и практика. 2025. Т. 21, № 3. С. 92–103. DOI: 10.22213/2618-9763-2025-3-92-103

For Citation

Isaeva E.V., Ermakova P.S. [Methods and techniques for cognitive adaptation of specialized socio-economic texts for target audiences: review and prospects for the development of information systems]. *Social'no-ekonomicheskoe upravlenie: teoria i praktika*, 2025, vol. 21, no. 3, pp. 92-103 (in Russ.). DOI: 10.22213/2618-9763-2025-3-92-103