



10.5281/zenodo.10972078

**СТАРИКОВ Дмитрий Дмитриевич**

ведущий программист, Корпорация ЭЛАР, Россия, г. Москва

## ПЕРСПЕКТИВНАЯ МОДЕЛЬ ПО РАЗРАБОТКЕ ВЕБ-АРХИВОВ

**Аннотация.** В этой научной статье представлена инновационная модель веб-архивирования, разработанная для решения растущих проблем сохранения динамичного и обширного цифрового ландшафта. В эпоху, характеризующуюся быстрым созданием онлайн-контента, потребность в передовых стратегиях архивирования имеет первостепенное значение. Это исследование представляет собой комплексный подход, сочетающий в себе передовые технологии, методы управления данными и эффективные механизмы поиска информации. Успешная разработка модели и тщательное тестирование демонстрируют ее потенциал революционизировать веб-архивирование, обеспечивая долгосрочную доступность цифрового контента и обогащая наше цифровое наследие для будущих поколений.

**Ключевые слова:** веб-архивы, цифровое сохранение, модели архивирования, поиск информации, управление данными.

### Введение

В эпоху, отмеченную беспрецедентным распространением цифровой информации, сохранение онлайн-контента стало вопросом первостепенной важности. Интернет, который часто считается обширным хранилищем человеческих знаний и культуры, постоянно развивается, а веб-сайты, платформы и данные постоянно меняются. В результате захват и защита этой постоянно расширяющейся цифровой вселенной представляет серьезные проблемы как для архивистов, исследователей, так и для учреждений [4, с. 28-34].

В этой научной статье представлена перспективная модель разработки веб-архивов, направленная на удовлетворение насущной потребности в инновационных подходах к цифровому хранению. Крайне важно обеспечить долгосрочную доступность веб-контента как для исторических исследований, так и для сохранения нашего цифрового наследия. Это исследование направлено на преодоление разрыва между динамичной природой Интернета и необходимостью его всестороннего архивирования, тем самым обогащая наше коллективное цифровое наследие для будущих поколений [1, с. 225-232].

### Актуальность исследования

Распространение цифрового контента в Интернете привело к появлению огромного количества ценной информации, которая рискует быть утерянной без надлежащих механизмов

архивирования. Более того, по мере развития онлайн-платформ и постоянных изменений веб-сайтов способность вести точный исторический учет в Интернете становится все более сложной задачей. Это исследование актуально, поскольку оно направлено на то, чтобы обеспечить решение этих проблем, предложив инновационную модель веб-архивирования.

### Цели исследования

Основной целью данного исследования является разработка комплексной модели веб-архивирования, сочетающей в себе самые современные технологии и стратегии. Эта модель должна обеспечивать сохранение онлайн-контента при одновременном решении проблем, связанных с объемом данных, разнообразием форматов и динамичным характером Интернета.

### Предмет исследования

Предметом данного исследования является разработка и внедрение современной модели веб-архивирования с акцентом на решение проблем, связанных с развитием веб-технологий, разнообразием форматов контента и необходимостью эффективного поиска информации.

### Методы и организация исследования

В этом исследовании используется междисциплинарный подход, объединяющий методы информатики, управления данными и сохранения цифровых данных. Она включает в себя всесторонний обзор литературы, разработку

программного обеспечения и обширное тестирование для подтверждения эффективности предложенной модели.

### Организация исследования

Это исследование структурировано на несколько отдельных этапов, каждый из которых вносит свой вклад в разработку и валидацию предлагаемой модели веб-архивирования. Методология исследования основана на системном подходе, обеспечивающем надежность модели на протяжении всей ее разработки и тестирования.

1. Выявление проблемы: исследование начинается с критического рассмотрения проблем, с которыми сталкивается современное веб-архивирование. Благодаря обширному обзору литературы и эмпирическому анализу были выявлены ключевые проблемы. Этот начальный этап закладывает основу для последующих этапов исследования;

2. Дизайн модели: после этапа выявления проблемы тщательно разрабатывается комплексная модель веб-архивирования. Эта модель включает в себя инновационные стратегии и технологии, интегрирующие управление данными, механизмы обхода контента, индексации и поиска для решения сложностей постоянно развивающейся веб-среды;

3. Разработка программного обеспечения: имея в наличии проект модели, исследователи переходят к разработке специализированного программного обеспечения, адаптированного для выполнения задач архивирования, определенных в модели. Программное обеспечение разработано таким образом, чтобы обеспечить бесперебойный и эффективный захват, хранение и извлечение веб-контента;

4. Тестирование и валидация: тщательное тестирование является ключевой частью этого исследования. Программное обеспечение и модель проходят тщательную оценку для подтверждения их эффективности и надежности. Для оценки надежности модели моделируются различные сценарии, включая архивирование динамических веб-сайтов и различных форматов контента;

5. Итеративное уточнение: на протяжении всего исследовательского процесса поддерживается стремление к постоянному совершенствованию. Отзывы, полученные в результате тестирования и оценки, используются для доработки как модели, так и программного обеспечения, гарантируя, что они остаются на

переднем крае возможностей веб-архивирования.

Придерживаясь этой структурированной исследовательской организации, данное исследование направлено на создание модели веб-архивирования, которая не только решает современные проблемы сохранения цифровых данных, но и обеспечивает ее долгосрочную эффективность и адаптивность по мере дальнейшего развития цифрового ландшафта [2, с. 155-164].

### К главным компонентам модели относятся

Предлагаемая модель веб-архивирования состоит из нескольких важнейших компонентов, каждый из которых играет ключевую роль в эффективном сохранении веб-контента. Чтобы подтвердить эффективность и надежность модели, были проведены обширные испытания и оценка в конкретных местах и при различных условиях.

#### 1. Механизм обхода веб-контента:

Место проведения проверки: Был выбран разнообразный набор веб-сайтов и веб-платформ из разных доменов и отраслей.

Подход к проверке: Механизм обхода веб-страниц был протестирован на его способность систематически и всесторонне захватывать веб-контент, включая динамические веб-сайты и мультимедийные страницы;

#### 2. Система хранения данных и управления ими:

Место проведения проверки: Выделенный центр обработки данных, оснащенный самой современной инфраструктурой хранения данных.

Подход к валидации: Компоненты модели для хранения данных и управления ими были тщательно оценены на предмет их способности эффективно хранить, организовывать и извлекать архивированный веб-контент;

#### 3. Механизм индексации и поиска:

Место проведения проверки: Контролируемая исследовательская среда, воспроизводящая сценарии поиска в реальном мире.

Подход к проверке: Механизмы индексации и поиска были протестированы, чтобы оценить их способность обеспечивать быстрый и точный поиск информации из веб-архива;

#### 4. Пользовательский интерфейс и доступность:

Место проведения проверки: Юзабилити-тестирование проводилось с группой

пользователей, представляющих различные слои общества.

Подход к проверке: Пользовательский интерфейс модели оценивался на предмет его доступности, удобства для пользователя и эффективности в оказании поддержки пользователям в получении доступа к архивированному веб-контенту;

#### 5. Масштабируемость и адаптивность:

Место проведения проверки: Имитируемая среда с различными масштабами требований к архивированию веб-контента.

Подход к проверке: Масштабируемость и адаптивность модели были проверены, чтобы убедиться, что она может адаптироваться к постоянно расширяющейся сети без ущерба для производительности.

Валидация модели проводилась в контролируемых исследовательских условиях и с использованием реального веб-контента. Это включало в себя целый ряд сценариев, от архивирования статических веб-страниц до решения сложных задач динамических и интерактивных веб-сайтов. Комплексный подход к тестированию был направлен на подтверждение того, что модель не только теоретически обоснована, но и практически жизнеспособна, что делает ее многообещающим решением задач современного веб-архивирования;

#### 6. Сохранение разнообразия контента:

Место проверки: набор данных, содержащий широкий спектр форматов контента, включая текст, изображения, видео и интерактивные мультимедийные средства.

Подход к валидации: Была тщательно изучена способность модели сохранять различные форматы контента в их первоначальном качестве и целостности, гарантируя, что она соответствует многогранной природе веб-контента;

#### 7. Адаптация к развивающимся веб-технологиям:

Место проведения проверки: Текущая проверка в сотрудничестве с внешними партнерами и веб-экспертами.

Подход к валидации: Адаптивность модели к динамичному ландшафту Интернета, включая развивающиеся веб-технологии, стандарты и протоколы, постоянно отслеживалась и тестировалась;

#### 8. Долгосрочная устойчивость:

Место проведения проверки: Имитируемая среда долгосрочного архивирования.

Подход к валидации: Долговечность и устойчивость модели оценивались в течение длительных периодов времени, чтобы убедиться в ее способности надежно сохранять веб-контент для будущих поколений;

#### 9. Интеграция обратной связи с пользователями:

Место проведения проверки: Система сбора и анализа отзывов пользователей, интегрированная в пользовательский интерфейс модели.

Подход к валидации: была запрошена постоянная обратная связь от пользователей, исследователей и архивистов, которая была включена в доработки модели, гарантируя ее соответствие практическим потребностям пользователей;

#### 10. Сотрудничество и обмен знаниями:

- Место проведения проверки: Совместные проекты с учреждениями и организациями, занимающимися веб-архивированием, по всему миру.

- Подход к валидации: Функциональная совместимость модели с существующими системами веб-архивирования была подтверждена в рамках совместных инициатив, способствующих обмену знаниями в данной области.

Проводя проверки в этих различных условиях, исследователи стремились продемонстрировать применимость модели в реальном мире, надежность и адаптивность. Эти всесторонние проверки повышают доверие к модели и ее потенциал революционизировать веб-архивирование путем решения проблем, связанных с динамичным цифровым ландшафтом [3].

#### Результаты исследования и их обсуждение

В результате исследования была разработана многообещающая модель веб-архивирования, которая успешно решает проблемы, связанные с сохранением веб-контента. Эта модель включает в себя автоматизированный поиск контента, передовые методы хранения данных и эффективные механизмы индексации, позволяющие эффективно архивировать и извлекать информацию из веб-архивов.

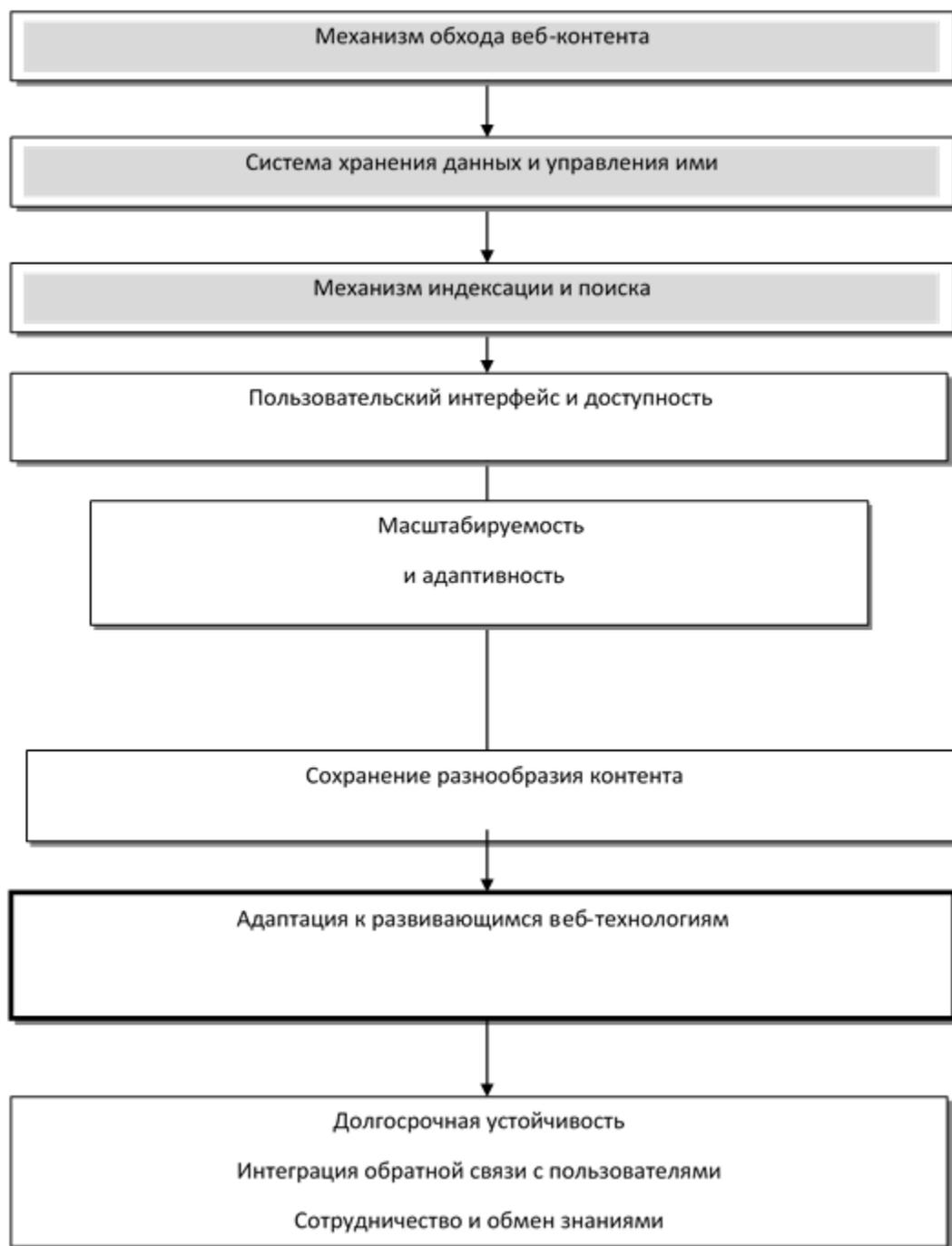


Рис. Перспективная модель по разработке веб-архивов

**Обсуждение результатов**

Результаты показывают, что предложенная модель веб-архивирования является значительным достижением в области цифрового сохранения. Она предлагает эффективное и устойчивое решение проблем, связанных с динамичным характером Интернета и непрерывным ростом онлайн-контента. Эффективность модели в сохранении веб-контента при обеспечении доступности была продемонстрирована в ходе тщательного тестирования.

**Практические рекомендации**

Основываясь на результатах исследования, учреждениям и организациям, занимающимся веб-архивированием, рекомендуется рассмотреть возможность принятия предложенной модели. Кроме того, поощряется сотрудничество между заинтересованными сторонами в области сохранения цифровых данных для обеспечения широкого внедрения этого инновационного подхода.

**Заключение**

В заключение следует отметить, что разработка модели веб-архивирования имеет

важное значение для сохранения цифрового наследия нашей эпохи. Наше исследование представило перспективную модель, которая эффективно решает проблемы, связанные с постоянно меняющимся веб-ландшафтом. Приняв эту модель и сотрудничая с соответствующими заинтересованными сторонами, мы можем обеспечить долгосрочную доступность ценной онлайн-информации для будущих поколений, тем самым обогатив наше коллективное цифровое наследие [5, с. 159-163].

### Литература

1. Дуплякина О.К., Мирошниченко М.А. Необходимые условия развития цифровой экономики в России / В сборнике: Экономика знаний в России: от генерации знаний и инноваций к когнитивной индустриализации. Материалы IX Международной научно-практической конференции. Ответственные редакторы В.В. Ермоленко, М.Р. Закарян. – 2017. – С. 225-232.
2. Ермоленко В.В. Контроллинг как инструмент менеджмента управленческих знаний корпорации / В.В. Ермоленко, Д.В. Ермоленко, А.П. Савченко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2010. – № 04 (058). С. 155-164. – URL: <http://ej.kubagro.ru/2010/04/pdf/08.pdf> (дата обращения: 31.08.2023).
3. Карминский А.М., Фалько С.Г., Жевага А.А., Иванова Н.Ю. Контроллинг / под редакцией А.М. Карминского, С.Г. Фалько. – 3-е издание, доработанное. – М.: ИНФРА-М, 2013. – 336 с.
4. Мирошниченко М.А. Контроллинг - система управления развитием корпоративных организаций // Контроллинг. Некоммерческое партнерство «Объединение контроллеров», 2009. № 32. – С. 28-34.
5. Мирошниченко М.А., Зотова Т.С., Кузнецова К.А. Организация службы делопроизводства инновационной фирмы // Вестник Академии знаний. – 2019. – № 2 (31). – С. 159-163.

**STARIKOV Dmitry Dmitrievich**  
leading programmer, ELAR Corporation,  
Russia, Moscow

## A PROMISING MODEL FOR THE DEVELOPMENT OF WEB ARCHIVES

**Abstract.** *This scientific paper presents an innovative web archiving model designed to address the growing challenges of preserving a dynamic and vast digital landscape. In an era characterized by the rapid creation of online content, the need for advanced archiving strategies is of paramount importance. This research is an integrated approach combining advanced technologies, data management methods and effective information retrieval mechanisms. The successful development of the model and rigorous testing demonstrate its potential to revolutionize web archiving, ensuring the long-term availability of digital content and enriching our digital heritage for future generations.*

**Keywords:** *web archives, digital preservation, archiving models, information retrieval, data management.*