

ОТЗЫВ

**официального оппонента Т.В. Майстрович
на диссертацию Руслана Сулеймановича Сулейманова, представленную на
соискание ученой степени кандидата технических наук по специальности
05.25.05 – Информационные системы и процессы, по теме «Интеграция
цифровых информационных ресурсов в электронные библиотеки»**

Представленное к защите Русланом Сулеймановичем Сулеймановым диссертационное исследование несомненно актуально. Диссертант верно отметил, что «в эпоху нарастающей цифровизации необходим принципиально новый подход к разработке информационного обеспечения образования и науки» (стр. 3). Обратим внимание на то, что, будучи технологической по организации, электронная среда по своему содержанию является площадкой, где реализуется большинство социальных практик, в том числе информационно-библиотечная деятельность. Поэтому совершенствование инструментария работы с электронными ресурсами, включая электронные библиотеки, - весомый вклад в повышение эффективности информационной поддержки науки и образования. В качестве ключевой проблемы при создании электронных библиотек автором выделена интеграция данных, чему и посвящено проведенное диссертационное исследование.

Определив проблемную ситуацию, Р.С. Сулейманов сформулировал цель своей работы: «улучшение качества интеграции цифровых информационных ресурсов из разных источников с помощью модели и методики, учитывающих разные структуры данных» (стр. 6). Понятно, что диссертация сама по себе не улучшает качество чего-либо, а предлагает механизмы и направления действий для этого. Несмотря на данную стилистическую погрешность, поставленные задачи вполне адекватно отражают основную целевую установку исследования.

Результаты, полученные в рамках решения поставленных задач, определили научную новизну, которую содержит данное диссертационное исследование: обоснование и разработка новой методики интеграции цифровых данных из разнородных и распределённых источников для электронных библиотек, включая обоснование моделей конструктора правил интеграции данных из распределенных источников и технологии извлечения метаданных из текстовых документов; создание грамматик и словарей для анализа полнотекстовых оцифрованных документов. На мой взгляд, вклад Р.С. Сулейманова в науку можно уточнить по двум аспектам. Во-первых, разработана не просто методика, а получил обоснование инструментарий для решения задач, связанных с организацией электронных ресурсов.

Соответственно (во-вторых) значение этого инструментария выходит за рамки только электронных библиотек.

Теоретическая значимость представленной к защите работы соответствует уровню диссертации на соискание ученой степени кандидата наук и заключается в расширении представлений о механизмах формирования электронных библиотек в том числе посредством создания модели конструктора правил интеграции электронных документов из распределенных источников.

Положения диссертации, выносимые на защиту, базируются на общей концепции автора и коррелируются с поставленными задачами. Несмотря на некоторую стилистическую погрешность в формулировках, они отражают суть проведенного исследования и раскрывают полученные результаты.

Структура диссертации Р.С. Сулейманова логична и позволяет последовательно решать поставленные задачи.

Первая глава («Современные подходы к интеграции данных в электронных библиотеках») содержит изложение теоретических основ формирования и функционирования целевого объекта (электронной библиотеки), для совершенствование которого предназначается разрабатываемый инструментарий. В первом параграфе представлен генезис электронных библиотек, определяется авторское понимание этого объекта и доказывается, что «электронные библиотеки заняли свою нишу в общей структуре передачи знаний. Дальнейшее их развитие связано с переводом все большего количества библиотечных фондов в электронный формат, а также с интеграцией библиотек» (стр.18). Поскольку разработка электронной библиотеки имеет свою специфику, в частности, необходимость описания файлов и хранения их метаданных, то во втором параграфе произведено сравнение двух базовых стандартов - MARC и Dublin Core. Полученные при сравнении результаты определили проблемное поле второй главы.

В начале второй главы («Обоснование принципов конструктора полей интеграции информации и модели извлечения метаданных из полнотекстовых документов») Р.С. Сулейманов обозначил основные пользовательские требования к электронной библиотеке, которые были учтены при разработке ER-схемы электронной библиотеки, основными элементами которой являются материалы, источники и авторы. К сожалению, в тексте диссертации не приводится раскрытие термина «материалы» и не дается ссылка на использованный в данном случае стандарт либо другой терминологический документ. В содержании первого параграфа хотелось бы обратить внимание на два момента. Во-первых, и в теоретическом осмыслении, и в построении методик важно, что они применимы для электронных библиотек

различного вида - от универсальных до специализированных. Во-вторых, инструментарий в целом позволяет адаптировать как ресурсы, производимые для конкретной электронной библиотеки (путем оцифровки), так и обеспечивать интеграцию с другими базами данных, в том числе исключительно электронного происхождения. И этот момент весьма существенен в свете таких масштабных проектов, как создание единых электронных сред - науки, образования, экономики и т.д.

Дальнейшее исследование вопросов интеграции электронных ресурсов (параграф 2) привело диссертанта к заключению, что существующие методики не всегда эффективны и требуют иного подхода, в качестве которого автором предлагается конструктор полей интеграции данных (параграф 3). О научной добросовестности Р.С. Сулейманова говорит то, что его доказательства базируются на проведенном эксперименте по проверке эффективности разработанной методики (параграф 4), последующего анализа его результатов (параграф 5) и расширения эмпирической базы до 10 000 документов. Комплексная обработка полученных данных дала возможность диссертанту выявить наиболее трудно вписывающиеся в автоматизированный технологический процесс атрибуты электронного документа: наименований материалов, сведения об издателях и месте издания, поскольку они не имеют формализованной структуры, однозначных правил могут содержать любое количество символов и знаков препинания и написания, для которых можно разработать универсальные грамматики (параграф 6). Для преодоления обозначенных трудностей предложен конструктор интеграции данных для электронной библиотеки, позволяющий объединять данные из других библиотек вне зависимости от используемого формата хранения метаданных, а также разработана модель извлечения метаданных из полных текстов материалов.

В приложениях представлены все основные разработки, обоснованию которых посвящено диссертационное исследование: грамматики для извлечения ISBN, информации об издателе, кодах рубрикаторов, дате и месте публикации, авторах и наименовании, а так же приводятся фрагменты исходного кода электронной библиотеки. Таким образом диссертационное исследование приобрело заверченный вид, что адекватно отражено в заключении к работе.

Использованный для исследования набор методов показал свою адекватность поставленным задачам и способствовал достижению заявленной цели.

Итоги диссертационного исследования Р.С. Сулейманова уже получили практическое применение. Предложенный им инструментарий используется в управлении библиотечным фондом электронной библиотеки Московского педагогического государственного университета. По результатам

диссертационного исследования были зарегистрированы две программы для ЭВМ, одна в соавторстве, другая (№ 2019661660 - «Конструктор правил интеграции данных для электронных библиотек») личная разработка автора.

Основные положения диссертации нашли отражение в семи публикациях автора и его выступлениях на научных мероприятиях. Исходный код конструктора правил интеграции информации из распределенных источников выложен в открытый репозиторий по лицензии GNU General Public License (универсальная общественная лицензия GNU).

Естественно, данная работа не свободна от ряда недостатков. На некоторые уже было обращено внимание выше. Замечание о неудачных формулировках (помимо цели исследования), относится и к вынесенным на защиту положениям. Например, первое положение: «Анализ имеющихся способов интеграции информации из распределённых источников выявил проблемы, возникающие в основном из-за разных форматов хранения метаданных в электронных библиотеках» (стр. 9) - звучит как констатация факта, а не как теоретический конструкт, нуждающийся в доказательстве.

Недостатки вычитанности текста особенно заметны на стр.13, где автором приводятся прогностические данные на 2020 и начало 2021 года, но не указывается, оказались ли они верными.

В списке использованной литературы преобладают публикации конца 1990-х - первого десятилетия 2020-х годов. Если с тех пор по теме диссертации не вышло работ, достойных внимания, это надо было указать при обзоре степени разработанности тематики.

Для данного диссертационного исследования (особенно в пятом параграфе второй главы) было бы уместно обратиться к международным и современным национальным стандартам по библиографическому описанию: ГОСТ Р 7.0.3 СИБИД "Издания. Основные элементы. Термины и определения", ГОСТ Р 7.0.100 СИБИД "Библиографическая запись. Библиографическое описание. Общие требования и правила составления" и другим.

Однако данные замечания касаются периферийных аспектов диссертации и не оказывают влияния на ее позитивную оценку. Можно констатировать, что цель исследования достигнута, поставленные задачи решены, положения, выносимые на защиту, нашли доказательное и развернутое обоснование. Диссертация содержит решение актуальной научной проблемы, имеет важное научное и практическое значение. Автореферат соответствует содержанию текста диссертации.

Диссертационное исследование Р.С. Сулейманова «Интеграция цифровых информационных ресурсов в электронные библиотеки» является

завершенным и нашедшим внедрение научным исследованием. Представленная к защите работа соответствует паспорту специальности 05.25.05 «Информационные системы и процессы» и предъявляемым к кандидатским диссертациям требованиям пп. 9-11, 13, 14 «Положения о присуждении ученых степеней», утвержденного постановлением Правительства Российской Федерации от 24 сентября 2013 г. № 842 (ред. от 01.10.2018, с изм. от 26.05.2020). Следовательно, Руслан Сулейманович Сулейманов заслуживает присуждения искомой ученой степени кандидата технических наук по специальности 05.25.05 «Информационные системы и процессы».

Т.В. Майстрович
доктор педагогических наук, доцент

Майстрович Татьяна Викторовна, доктор педагогических наук по специальности 05.25.03 «Библиотечное дело, библиографоведение и книговедение», ведущий научный сотрудник Научно-исследовательского отдела библиотечного дела Фундаментальной библиотеки Федерального государственного бюджетного учреждения науки «Институт научной информации по общественным наукам Российской академии наук» (юридический адрес ИНИОН РАН: 117418, Москва, Нахимовский проспект, д. 51/21; электронная почта: office-inion@mail.ru; телефон: (499) 124-3076).

Телефон: +7 (903) 585-32-10

Адрес электронной почты: t-maistr@yandex.ru

25.02.2021

Подпись Т.В. Майстрович заверяю

Ученый секретарь ИНИОН РАН Я.Г. Гудкова

