

Цифровой профиль пользователя библиотек: структурно-функциональный подход. (Часть 1)

А. И. Каптерев^{1,2,3}

¹ГПНТБ России, Москва, Российская Федерация

²Российская государственная библиотека, Москва, Российская Федерация

³Московский городской педагогический университет, Москва,
Российская Федерация

kapterev@narod.ru, <https://orcid.org/0000-0002-2556-8028>

Аннотация. Статья является первой из двух, посвящённых данной теме. Теоретическая значимость работы заключается в демонстрации возможностей цифрового профиля (ЦП) пользователя библиотек. Рассмотрены следующие составляющие проблемы: а) структура ЦП пользователя с выделением слоёв – субъектного, объектного, предметного, коммуникативного и б) функции каждого слоя – идентификационная, кластеризационная, фильтрационная, коллаборационная. На каждом слое выделены доминирующие факторы, анализ которых позволяет моделировать, формировать и использовать профили для проактивного удовлетворения информационных потребностей пользователей библиотек. Практическая значимость исследования – разработка и применение ЦП пользователя библиотечными услугами для персонализации обслуживания: проактивное информирование о новых поступлениях в фонд, цифровое сопровождение пользователя (курирование), поддержка исследовательских проектов; приведение фонда в большее соответствие запросам основного контингента пользователей (оптимизация состава фонда, изменение политики комплектования, каталогизации и рекламы библиотечных услуг); лучшее понимание взаимосвязи запросов и потребностей пользователей в результате использования информации из ЦП, что позволит более обоснованно решать проблемы в области комплектования и предоставления услуг; согласование деятельности библиотеки с планами развития местного сообщества и его ожиданиями.

Статья выполнена в рамках государственного задания ГПНТБ России на 2022-2024 гг. № 1021062311368-2-5.8.3 «Развитие электронного библиотечного знания как научной и учебной дисциплины в условиях трансформации библиотечных фондов, справочно-библиографического и документного обслуживания в цифровой среде (FNEG-2022-0004)».

Ключевые слова: пользователь библиотек, цифровой профиль, модель, слои, формирование, использование

Для цитирования: Каптерев А. И. Цифровой профиль пользователя библиотек: структурно-функциональный подход. (Часть 1) // Научные и технические библиотеки. 2024. № 7. С. 82–102. <https://doi.org/10.33186/1027-3689-2024-7-82-102>

MODERN INFORMATION TECHNOLOGIES

UDC 004.65:024 + 004:02

<https://doi.org/10.33186/1027-3689-2024-7-82-102>

Library user digital profiling: The structural and functional approach. (Part I)

Andrey I. Kapterev^{1, 2, 3}

¹*Russian National Public Library for Science and Technology,
Moscow, Russian Federation*

²*Russian State Library, Moscow, Russian Federation*

³*Moscow City Pedagogical University, Moscow, Russian Federation
kapterev@narod.ru, <https://orcid.org/0000-0002-2556-8028>*

Abstract. This is the first of two articles on the problem. The theoretical significance of the article lies in the demonstration of the potential of library user digital profiles (DP). The author examines the following problem components: a) user DP structure with specifying the layers – subject, object, domain, communications, and b) layer function – identification, clusterization, filtration, and collaboration. For each layer, the dominant factors are identified that enable to model, structure and use DPs for proactive meeting of library users' information demands. The practical value of the study lies in the DP development and application for services personalization: proactive building awareness of new acquisitions, digital user support and curation, support of research projects; reconciling library collection and the demands of core user groups (collection structure enhancement, changing of the policy of acquisitions, cataloguing, and library services promotion); need-demand interdependence through using DP information, which makes

acquisitions and services more efficient; and accommodation of library activities with local community's plans and expectations.

The article was carried out within the framework of the Government order for the Russian National Public Library for Science and Technology for 2022-2024. No. 1021062311368-2-5.8.3 "Development of e-librarianship as a scientific and academic discipline in the context of the transformation of library collections, reference, bibliographic and document services in the digital environment (FNEG-2022-0004)".

Keywords: library user, digital profile, model, strata, development, usage

Cite: Kapterev A. I. Library user digital profiling: The structural and functional approach. (Part I) // Scientific and technical libraries. 2024. No. 7, pp. 82–102. <https://doi.org/10.33186/1027-3689-2024-7-82-102>

Модель ЦП пользователя библиотек

Задачи проведённого исследования: а) контент-анализ литературы для выявления сущности, структуры и определений понятия «цифровой профиль», формулирование авторского определения, моделирование ЦП пользователя, определение этапов, методов и средств его формирования и использования в библиотечно-библиографической деятельности.

В процессе работы применялись структурно-функциональный подход и следующие методы: анализ процессов цифровизации и цифровой трансформации социокультурных объектов, концептуальное моделирование, анализ отраслевого информационного потока.

В «Стратегическом направлении в области цифровой трансформации отрасли культуры Российской Федерации до 2030 года» отмечено, что «не обеспечены полнота сбора и хранение первичных данных из учреждений культуры о посетителях и клиентах для наполнения их цифровых профилей» [1. С. 26]. Определим **ЦП пользователя библиотечными услугами и продуктами** как набор данных о пользователе, собранный библиотекой с помощью электронных средств. По содержанию он может включать разнообразную информацию о пользователе и его цифровом следе: интересы, историю чтения и опыт использования библиотечных продуктов и услуг. Форма ЦП зависит от того, какие средства его формирования использовала библиотека: от простой БД запросов пользователя (как удовлетворённых, так и отказанных или перенаправленных) до проактивной рекомендательной системы, включающей

взаимосвязанные факторы, принадлежащие разным слоям и доступные количественному измерению и интеллектуальному анализу данных.

К этой проблеме в разное время обращались: один из крупнейших отечественных специалистов по информационным ресурсам и цифровой гуманитаристике А. Б. Антопольский [2], специалисты ГПНТБ России М. В. Гончаров и К. А. Колосов, обратившие внимание на важность альтметрик [3], директор ГПНТБ России А. С. Карауш, изучающий такие перспективные направления, как контентная и коллаборативная фильтрация информационных запросов [4]. О. А. Лаврёнова и А. А. Винберг изучают пользователей библиотек в связи с проблематикой открытых и связанных данных [5], моя коллега М. Ю. Нещерет анализирует опыт создания библиотеками личных кабинетов пользователей [6], Н. А. Сергеева продемонстрировала возможности работы с удалёнными пользователями на примере РНБ [7], Я. Л. Шрайберг, крупнейший организатор библиотечной деятельности в России и учёный-библиотековед, на протяжении многих лет активно анализирует тренды цифровизации библиотек [8]. Из зарубежных авторов назову лишь некоторых, чьи исследования ЦП пользователей библиотечными услугами созвучны нашему пониманию проблемы. Например, Бриони Бирди изучал характеристики читателей художественной литературы в связи с их жанровыми предпочтениями с помощью метода репертуарных решёток [9]; китайские специалисты Джао Линь, Венджан Гуан, Джинг Джанг исследовали применение технологии интеллектуального анализа данных с улучшенным алгоритмом кластеризации в библиотечной системе персонализированных рекомендаций книг [10]; учёные из Бангладеш Джиман Сарма, Танни Митра и Мохаммад Хусейн предложили эффективную систему рекомендаций книг для онлайн-пользователей на базе метода кластеризации [11]; специалисты из Индонезии Хафид Ахмад Аджатма и З. К. Байзал разработали систему рекомендаций книг с использованием матричной факторизации и коллаборативной фильтрации (CF) методом чередующихся наименьших квадратов, которая сравнивается с методом разложения по сингулярным значениям для получения точной системы рекомендаций [12]; испанские специалисты К. Порсел, Дж. Морено и И. Херрера-Видма представили модель нечёткой лингвистической рекомендательной системы, которая помогает пользователям университетских электронных библиотек получить доступ к исследовательским ресурсам, расширяет возможности сотрудничества в рамках университета и улучшает услуги

университетской цифровой библиотеки [13]. Мы видим, что проблема использования ЦП изучается в разных странах и в разных аспектах. Однако нам не удалось выявить целостный анализ ЦП пользователя библиотечными продуктами и услугами в существующих библиотечно-ведческих публикациях.

Мы предлагаем в структуре ЦП пользователя выделить несколько слоев: **а) субъектный, б) объектный, в) предметный, г) коммуникативный.**

На субъектном слое, основной функцией которого является идентификация личности, отметим персональные данные, охраняемые многими законодательными актами: ФИО, паспортные данные (возраст, место регистрации) и др.

На объектном слое ЦП менеджеров библиотеки могут интересоваться такие данные пользователя, как: образование, специальность, квалификация, профессиональный стаж, принадлежность к организации, занимаемая должность, уровень дохода в семье (зависимость от родителей или опекунов), профессиональная роль. Эти факторы должны интересовать библиотекаря не сами по себе, а в соотношении с моделями предметных областей, в том числе применяемых для систематизации в фондах библиотек и в используемых АБИС.

На предметном слое ЦП пользователя следует детально анализировать форму и содержание информационных запросов: а) очная, дистанционная форма; б) отраслевой характер; в) однократность, повторяемость, частота запроса; г) тип запрашиваемой информации; д) вид документов; е) число запрошенных документов.

На коммуникативном слое выделим факторы обратной связи: прямые (отзывы пользователей, соотношение выданных и повторно заказанных документов, косвенные (реакции пользователей в соцсетях, участие в профессиональных мероприятиях, запросы из референтных групп и др.). Совместив слои, функции, доминирующие факторы на каждом слое, методы и средства, применяемые в процессе формирования и использования ЦП, построим концептуальную модель ЦП пользователя библиотек (см. табл.).

Концептуальная модель ЦП пользователя библиотек

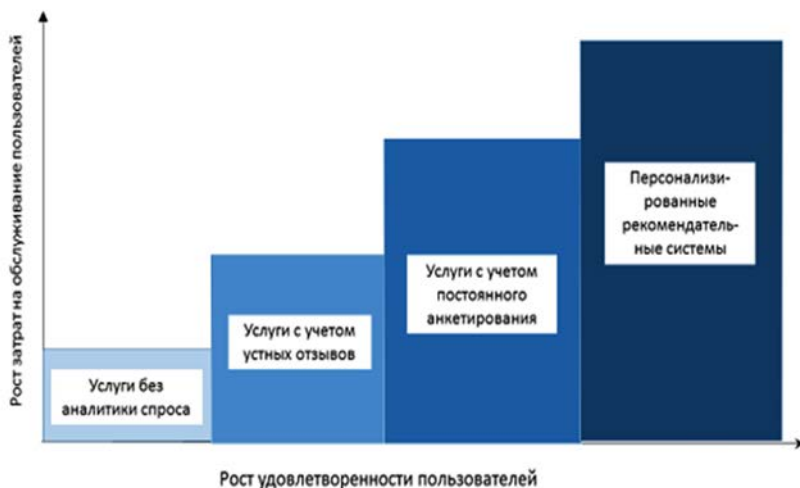
Слои	Целевая функция	Факторы	Методы	Средства
Субъектный	Идентификация	Персональные данные	Формирование	СУБД + CRM
Объектный	Кластеризация	Социологические данные	Обработка	SPSS
Предметный	Фильтрация	Запросы	Анализ	АБИС
Коммуникативный	Коллаборация	Потребности	Прогноз	Рекомендательные системы (CXDP)

Адаптируя идеи классика современного маркетинга Роджера Дж. Беста [14. Р. 164], можем показать различные CRM¹-стратегии библиотечного обслуживания, демонстрирующие зависимость удовлетворённости пользователя от финансирования (рис. 1). Очевидно, что персонализация обслуживания – дело затратное, но, если мы провозглашаем курс на клиентоориентированность библиотечной деятельности, то следует учитывать и траты на интеллектуальный анализ данных, без которого достичь удовлетворённости пользователей невозможно.

Сегодня цифровая трансформация рассматривается как стратегическое направление развития многих отраслей, в том числе и библиотечной [1]. Виртуальные коммуникации и образ жизни современного человека, предполагающий дистанционное электронное взаимодействие в режиме 24/7, ставят проблему цифровой идентификации личности как минимум в социальном, психологическом, экономическом, юридическом и техническом аспектах. В науке изучение идентификации личности включает различные подходы и методы, применяемые в разных странах.

Социологи изучают личность как субъект, который стремится найти жизненный путь, соответствующий его индивидуальности, используя все имеющиеся в распоряжении способности и личностные свойства в качестве ресурса для достижения целей.

¹ CRM – Customer Relationship Management (управление отношениями с потребителями-клиентами).



**Рис. 1. CRM-стратегии библиотечного обслуживания.
Субъектный слой ЦП**

Психологи рассматривают проблему идентификации с точки зрения индивидуальных особенностей и характеристик, таких как личностные ценности, стереотипы и механизмы самоактуализации. Когнитивная психология изучает процессы получения субъектом информации, её представления, хранения и преобразования в знания, а также результаты влияния полученной информации на сознание и поведение человека. У истоков когнитивной психологии стояли американские психологи Джордж А. Миллер, Джером Брунер, Ульрик Найссер. Их коллега Джордж Келли предложил модель личности, основанную на образе человека как исследователя. Для этого ему пришлось разработать особый метод – метод репертуарных решёток. «Интерпретация человека как исследователя, постоянно занимающегося построением своего собственного “образа мира” при помощи конструктов, предложенная Келли, сегодня весьма актуальна» [15]. В настоящее время многие организации для анализа поведения используют компетентностный подход. В этом случае критерием является демонстрация человеком определённых компетенций – моделей поведения, обеспечивающих успешное выполнение задач и описанных в поведенческих

индикаторах. Оценивая способности и черты личности, мы можем оценить предрасположенность человека к демонстрации того или иного профессионального поведения и компетенций, то есть успехов в решении профессиональных задач. Этот подход мигрировал из психологии и в педагогику [16, 17]. Интересный опыт описан Татьяной Крупа: «На каждом шаге образовательной траектории [модель – А. К.] xMANN принимает конкретное задание и выдаёт вероятность его успешного выполнения, а затем обновляет значение в памяти для конкретной связки “задание – результат”». Модель xMANN отслеживает состояния учащихся путём чтения и записи в матрицу значений с применением корреляционных весов, вычисленных на входном задании и связке “задание – результат» [18. С. 133]. «Благодаря использованию более богатой модели данных точность предсказания увеличивается на 15–20%. Это позволяет повысить уровень автоматизации оказания цифровых информационных услуг за счёт использования алгоритма предсказания оптимальных образовательных траекторий пользователей, а, следовательно, их индивидуализации». [Там же. С. 134].

В исторических исследованиях личность, как правило, анализируется с точки зрения её роли в истории, что предполагает анализ влияния исторических деятелей и их личностных качеств на события и процессы в различные периоды. Цифровизация подобных исследований даёт результаты визуальных лент времени (таймлайнов), широко применяемых в образовательных процессах.

ЦП используется также в экономике: индивидуальный ЦП кандидата на должность сравнивается с модельным, отчёт формируется в табличной и/или графической форме [19]. Работу по составлению ЦП потребителя маркетологи рекомендуют начать со статистической информации из таких источников, как Google Analytics, Google AdWords, Яндекс.Метрика, Яндекс.Директ. Данные систем помогают аналитикам провести анализ и разделить свою целевую аудиторию на сегменты, выделив общие характеристики в каждом из них.

В юриспруденции становится актуальной проблема защиты ЦП личности. Однако пока в законодательстве Российской Федерации понятие «цифровой профиль личности» отсутствует, его правовое положение и механизм защиты не определены [20]. Юристы предложили такое определение: «Цифровая личность – наиболее полная совокупность различного рода достоверных сведений о субъекте, выраженных

в цифровой форме посредством размещения их в аккаунтах и/или профилях цифровых платформ с помощью программ для ЭВМ и позволяющих осуществить идентификацию этого субъекта среди других в реальном мире» [21]. При создании ЦП пользователя необходимо соблюдать конфиденциальность и обеспечить безопасность информации. Информация должна использоваться только в целях, для которых она была собрана, и должна быть защищена от несанкционированного доступа, использования и раскрытия. Для извлечения индивидуальной информации, как уже было показано многими исследователями, иногда достаточно иметь доступ к агрегированным данным, например, средним или суммарным значениям за период или географическим регионам. Дифференциальная приватность основана на внесении в данные случайного шума по определённым, достаточно сложным алгоритмам. Разумеется, часть информации при этом теряется, но статистические параметры даже по зашумлённым данным удаётся подсчитать с удовлетворительной точностью. «Алгоритмы разрабатываются так, чтобы можно было менять уровень вносимого шума в зависимости от того, насколько важно скрыть индивидуальные данные по сравнению с тем, насколько важно получить более точные результаты при анализе. Усовершенствованный алгоритм позволяет защитить данные с той же надёжностью, но с меньшими вычислительными затратами на добавление шума, что важно при работе с большими массивами данных» [22].

В последние десятилетия во всём мире активно развиваются биометрические технологии идентификации личности. Они позволяют установить личность человека по его физиологическим или поведенческим характеристикам. В целом научное изучение идентификации личности включает в себя многообразные подходы и методы, позволяющие исследователям проникать в сложные процессы и явления, связанные с личностью и её развитием.

Технологии четвёртой промышленной революции, стремительно ворвавшиеся в повседневную жизнь большинства людей, высвечивают новые ракурсы проблемы формирования идентичности личности, которые, являясь недостаточно изученными с научной точки зрения, тем не менее стали повседневной реальностью каждого человека, полем для практических экспериментов [23]. Так, например, в связи с развитием технологий ИИ появилась возможность моделирования ЦП с использованием нейронных сетей. Нейронная сеть обладает многослойностью

(сегментацией факторов), токенизацией (разбиением каждого фактора на ряд параметров с конкретными количественными весами), векторизацией (отображением значений весов в n-мерном пространстве). Самообучаемость на сравниваемых выборках и уточнение весов позволяют отбирать документы, наиболее релевантные либо по контенту (контентно-зависимая фильтрация), либо по запросам пользователей, чьи ЦП наиболее близки (коллаборативная фильтрация). В последнее время применяются и гибридные методы фильтрации, которые детально рассмотрим в следующей статье.

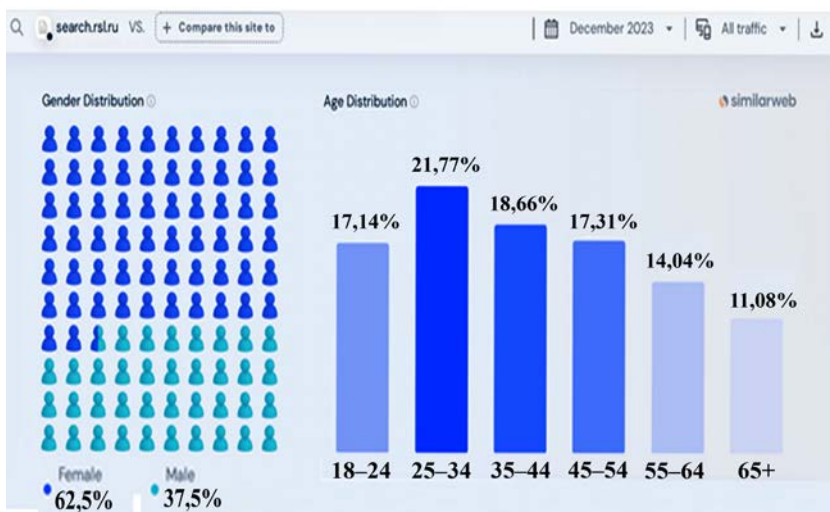


Рис. 2. Половозрастной состав посетителей сайта. Объектный слой

Первоначально необходимо собрать базовую информацию о пользователе (имя, фамилия, возраст, пол, контактные данные, адрес, информация для связи, дата рождения). Сбор информации для ЦП начинается с заполнения читательского формуляра при регистрации либо сотрудником, либо самим пользователем (при удалённой регистрации). На субъектном слое библиотека собирает в основном анкетные, демографические (возраст, пол, семейное положение, уровень образования, профессия, стаж и местоположение) и психографические данные: (информационные цели, жизненный этап, языковые предпочтения и уровень владения языком). Эта информация даёт представле-

ние о том, кто ваши читатели. Например, с помощью сервиса Similarweb можно получить информацию о половозрастном составе пользователей, обращавшихся к portalу удалённых ресурсов РГБ в определённый временной период (например, в декабре 2023 г.) (рис. 2).

В качестве объектов ЦП пользователя мы рассматриваем документы, попадающие в зону его взаимодействия с библиотекой, как реально запрошенные, так и потенциально доступные. Библиотеки, архивы и музеи сталкиваются с проблемой обеспечения доступа к быстро-растущим коллекциям при ограниченном бюджете. Такие характеристики пользователя, как направление и уровень его образования, стаж работы по выбранной специальности и другие данные о целях, интересах и мотивах обращения в библиотеку становятся теми фильтрами, через которые необходимо «просеивать» имеющиеся в фонде и особенно новые поступления при выполнении запросов и последующей кластеризации потенциально пригодных для пользователя (пертинентных) документов. Ключом к этому являются создание, связывание и публикация метаданных хорошего качества, которые позволяют обнаруживать релевантные документы, получать к ним доступ и распространять. Проблема усугубляется неоднозначностью отнесения отдельных документов, проблемных кластеров на стыках наук, да и целых предметных областей к тем или иным научным разделам в разных классификационных системах. Поскольку вопрос о единой, признаваемой всеми классификационной системе документов, пригодной для использования в библиотеках, не решён, паллиативным, но всё же распространённым механизмом стало использование метаданных.

Спектр метаданных, необходимых для управления цифровой библиотекой и сохранения её коллекций в долгосрочной перспективе, гораздо обширнее и сложнее, чем что-либо в традиционной библиотеке. Он включает в себя ту же описательную информацию, которая направляет пользователей к необходимым им ресурсам, но должен дополнять её всеобъемлющими метаданными: техническими деталями файлов, составляющих его коллекции, документацией о правах интеллектуальной собственности и обширным набором, необходимым для поддержки процессов хранения и поиска в долгосрочной перспективе. Для учёта всего этого требуется использование множества стандартов метаданных, все из которых в перспективе гипотетически должны быть сведены в единое интегрированное целое.

Библиотечные специалисты давно осознали этот факт и используют различные метаданные документов. В России широко применяется машиночитаемый формат MARC, существует ГОСТ «Набор элементов метаданных “Дублинское ядро”» [24]. В Европейском Союзе в настоящее время развивается проект OpenAIRE под названием OpenAIRE-Advance [25], краткую аннотацию сервисов которого любезно разместил в своей статье А. Б. Антопольский [26]. Он отметил, что «финансируемый ЕС проект EINFACENTRAL предпринял шаги для исправления разрозненных форматов метаданных путём разработки общего каталога европейских электронных инфраструктур. Команда проекта собрала и опубликовала информацию о более чем 150 цифровых услугах и ресурсах на портале Европейского открытого научного облака ЕС (EOSC). Перечисленные услуги постоянно добавляются, и каталог доступен для бесплатного использования» [Там же. С. 9]. Александр Борисович дал и интересный обзор исследований связанных открытых данных в цифровой гуманитаристике [27].

Два тематических исследования, проведённых Институтом Варбурга в Лондоне, показывают, как стратегия работы с метаданными может быть реализована в библиотечной практике. Описанная стратегия, изложенная в одной из книг, гарантирует, что метаданные цифровой библиотеки будут поддерживать все её операции, будут полностью совместимы с другими и обеспечат их долгосрочное сохранение. В работе содержится как введение в передовые методы работы с метаданными цифровой библиотеки, так и руководство по их практическому внедрению [28].

В другой работе, посвящённой использованию связанных данных в цифровой гуманитаристике, эксперты по метаданным Сет ван Хуланд и Рубен Верборг знакомят с основными концепциями стандартов метаданных и связанных данных и показывают, как их можно применять к существующим метаданным. Работа с метаданными предполагает: а) создание метаданных, включая архитектуру, модели данных и стандарты; б) очистку метаданных; в) согласование метаданных; г) обогащение метаданных за счёт связанных данных и распознавания именованных объектов; д) экспорт метаданных [29].

При создании метаданных необходимо отслеживать логику связанных данных с помощью основных парадигм их моделирования. Проблема несоответствия документных классификаций уходит глубоко

в различия тех классификаций наук, которые стали фундаментом для создания конкретных документных классификаций. Специалисты отмечают сходства и различия между табличными данными, реляционными базами данных, XML и RDF, рассматривают влияние моделирования данных на метаданные, анализируют построение запросов в графах знаний, знакомят со SPARQL через DBpedia, а также с проприетарным языком запросов Metaweb.

На этапе очистки метаданных обычно анализируют качество данных и пытаются понять, как устранить распространённые проблемы с их качеством. Как известно, «абсолютного качества» не существует, оно всегда относительно. Возникает вопрос, как метаданные эволюционируют с течением времени. Для профилирования и очистки метаданных используют общие функции OpenRefine – настольного приложения с открытым исходным кодом для очистки данных и преобразования их в другие форматы.

На этапе согласования метаданных рассматривают возможности и ограничения повторного использования контролируемых словарей и тезаурусов. Распознавание именованных сущностей (NER) – это одна из функций, предлагаемых языком ИИ Microsoft Azure, набором алгоритмов машинного обучения и ИИ в облаке для разработки интеллектуальных приложений, использующих письменный язык. Функция NER может обнаруживать и классифицировать сущности в неструктурированном тексте – такие, как люди, места, организации и количества.

На этапе обогащения изучают возможности и ограничения применения NER к метаданным, анализируют большие данные и оценивают возможности применения методов «дистанционного чтения» к метаданным культурного наследия. Так, например, отдельные специалисты дают сравнительный анализ различных сервисов NER (Zemanta, Alchemy и DBpedia Spotlight) и тестируют их на описательных полях метаданных, предоставленных Британской библиотекой через Europeana [Там же].

На этапе экспорта метаданных рассматривают, как публиковать свою коллекцию. Здесь интересен опыт практических преимуществ стиля передачи репрезентативного состояния (REST). REST – это архитектурный стиль программного обеспечения, который определяет набор правил, используемых для создания веб-сервисов. Веб-службы, следующие архитектурному стилю REST, известны как веб-службы RESTful. Это позволяет запрашивающим системам получать доступ к

веб-ресурсам и манипулировать ими с помощью единообразного и предопределённого набора правил. Взаимодействие в системах на базе REST происходит через интернет-протокол передачи гипертекста (HTTP).

Принципы связанных данных часто неправильно понимаются и должны быть реализованы хорошо продуманным образом. Связанные данные создают огромные проблемы с точки зрения качества метаданных, и поэтому крайне важно выработать критический взгляд и провести различие между тем, что осуществимо, а что нет [30].

Когда же речь идёт об использовании связанных данных (LOD) в библиотеках, то предполагают: работу по созданию, идентификации и представлению семантически насыщенных связанных данных; выбор концепции онтологий; сравнительный анализ проектов LOD (Open Graph, DBpedia, BIBFRAME и расширение Bib от Schema.org) [31]; оценку бюджетов электронных ресурсов, рабочих процессов, поставщиков и совместных инициатив; использование наборов данных для оценки коллекций по тематике, форматам или другим факторам; работу с данными с использованием распространённых приложений, таких как Excel, Access или MySQL; ключевые особенности создания хранилища данных и витрин данных; методы сравнительного анализа процессов обработки метаданных [32]; механизмы обмена и повторного использования данных с расширенными разделами, посвящёнными открытому доступу, открытым данным и инициативам по открытым стандартам; процессы, гарантирующие сохранность данных и их пригодность к использованию с течением времени.

Следует отметить, что в последнее время происходит переключение на контролируемые словари, краудсорсинг, использование API, профилирование данных. Часто в библиотеках эти задачи и связанные с ними процессы по-прежнему выполняются отдельными сотрудниками, расположенными на разных ветвях организационного дерева, а иногда даже в разных зданиях. Это физическое и психологическое разделение часто способствует излишнему дублированию функций и снижает эффективность работы организации. Последовательное внедрение технологий на протяжении десятилетий разбросало метаданные наших библиотек по множеству БД, электронных таблиц и даже неструктурированных текстовых документов.

Мы не должны забывать и ещё об одной характеристике метаданных: они постоянно расширяемы, всегда можно добавить ещё один

слой для описания метаданных поверх другого. Сообщество компьютерных наук за последние несколько лет выпустило специальные руководства по связанным данным [33, 34]. Отметим также ставшее классическим руководство по семантическому вебу Дина Аллеманга и Джеймса Хендлера [35]. Хотя эта тема актуальна в контексте связанных данных, стандарты метаданных за последние годы широко рассматривались и в других источниках.

Основываясь на работах по связанным данным, можно предположить, что мы откажемся от наших реляционных БД в пользу хранилищ триплетов – наборов данных, представленных моделью «субъект – предикат – объект» («Сергей – отец Елены» или «берёза является деревом»). Хранилище триплетов, или RDF-хранилище – специальным образом устроенная БД, в которой можно хранить и искать триплеты, используя семантические запросы. Схожесть таких хранилищ с реляционными БД в том, что информация извлекается через язык запросов, но, в отличие от них, такое хранилище оптимизировано для хранения именно триплетов. Помимо запросов, обмен информацией может производиться через RDF или другой формат. Реальность такова, что мы всё-таки продолжим использовать реляционные БД в течение следующих лет (и, возможно, десятилетий), поскольку они превосходно справляются со структурированными данными. Однако выбор программного обеспечения как для разных типов библиотек, так и для отдельных процессов всё ещё является предметом дискуссий.

Предметный и коммуникативный слои ЦП пользователя мы анализируем в следующей статье.

Подводя промежуточный итог, констатируем, что периодический анализ новых поступлений, а позднее и всего фонда с использованием описанных методов позволит оптимизировать фонды и повысить эффективность обслуживания читателей в библиотеках различных типов.

В качестве перспектив практического использования ЦП пользователя библиотечными услугами мы рассматриваем следующие направления:

персонализация обслуживания пользователей, в том числе проактивное информирование о новых поступлениях в фонд, цифровое сопровождение интересов пользователя (курирование), поддержка исследовательских проектов;

приведение фонда в большее соответствие запросам основного контингента пользователей (оптимизация состава фонда, изменение

политики комплектования, каталогизации и рекламы библиотечных услуг);

эффективное изучение взаимосвязи запросов и потребностей пользователей библиотеки для принятия более обоснованных решений в комплектовании и предоставлении услуг;

приведение в соответствие деятельности библиотеки с планами развития местного сообщества.

Список источников

1. **Стратегическое** направление в области цифровой трансформации отрасли культуры Российской Федерации до 2030 года. Утв. распоряжением Правительства Российской Федерации от 11 декабря 2023 г. № 3550-р.
2. **Антопольский А. Б.** Информационные объекты цифровой гуманитаристики: вопросы типологии // Информационные ресурсы России. 2023. № 1(190). С. 56–69. DOI 10.52815/0204-3653_2023_1190_56. EDN SSCEVU.
3. **Гончаров М. В., Колосов К. А.** Вопросы расчёта альтметрик на основе данных, формируемых при обращении пользователей к электронным каталогам библиотек // Научные и технические библиотеки. 2020. № 11. С. 73–88.
4. **Карауш А. С.** Перспективные направления развития информационных технологий в публичных библиотеках // Информационные технологии, компьютерные системы и издательская продукция для библиотек : материалы конференции «LIBCOM-2010» : электронный ресурс, Ершово, 15–19 ноября 2010 г. Ершово : Государственная публичная научно-техническая библиотека России, 2010. С. 28–32. EDN ZBAECH.
5. **Лаврёнова О. А., Винберг А. А.** Современные пользователи библиотек и пространство связанных открытых данных // Библиотековедение. 2020. Т. 69, № 3. С. 243–260. DOI 10.25281/0869-608X-2020-69-3-243-260.
6. **Нещерет М. Ю.** Библиотечно-информационные услуги в личном кабинете пользователя библиотеки // Библиотековедение. 2023. Т. 72, № 3. С. 213–223. DOI 10.25281/0869-608X-2023-72-3-213-223.
7. **Сергеева Н. А.** Библиотека online/offline: на примере Российской национальной библиотеки (РНБ) // Евразийский Союз Учёных. Серия: педагогические, психологические и философские науки. # 8 (101), 2022. С. 33–37.
8. **Шрайберг Я. Л.** Особенности компоненты цифровой трансформации общества, активно влияющие на технологические и поведенческие модели деятельности современных библиотек (Ежегодный доклад Седьмого международного профессионального форума «Со-чи-2023») // Научные и технические библиотеки. 2023. № 8. С. 13–84. DOI 10.33186/1027-3689-2023-8-13-84. EDN FTIDSG.

9. **Birdi B.** (2011) Investigating fiction reader characteristics using personal construct theory. *Aslib Proceedings*, 63 (2–3). 275–294. ISSN 0001-253X.
<http://dx.doi.org/10.1108/00012531111135709> (дата обращения: 25.01.2024).
10. **Xiao Lin, Wenjuan Guan, Ying Zhang.** Application of Data Mining Technology with Improved Clustering Algorithm in Library Personalized Book Recommendation System // *International Journal of Advanced Computer Science and Applications*, Vol. 14, No. 11. 2023. P. 494–504.
11. **Sarma D., Mittra T., Hossain M. S.** Personalized book recommendation system using machine learning algorithm. *International Journal of Advanced Computer Science and Applications*, 2021, 12 (1):212–219.
12. **Hafid Ahmad Adyatma, Z. K. A. Baizal.** Book Recommender System Using Matrix Factorization with Alternating Least Square Method // *Journal of Information System Research (JOSH)*. Volume 4, No. 4, Juli 2023, pp. 1286–1292. ISSN 2686-228X (media online).
<https://ejurnal.seminar-id.com/index.php/josh/> DOI 10.47065/josh.v4i4.3816.
13. **Porcel C., Moreno J. M., Herrera-Viedma E.** A multi-disciplinar recommender system to advice research resources in University Digital Libraries // *Expert Systems with Applications* 36 (2009). P. 12520–12528.
14. **Best Roger J.** *Market-Based Management: Strategies for Growing Customer Value and Profitability.* Pearson College Div; 5th edition. 2008. 515 p.
15. **Левицкая И. А.** Теория личностных конструкторов Дж. Келли: на пути к когнитивной философии образования // *Вестник Нижегородского университета им. Н. И. Лобачевского. Сер. Социальные науки.* 2013. № 2 (30). С. 114–119.
16. **Kapterev A.** Professional Environment as an Information Analysis Object // *Knowledge Organization*. 1993. Vol. 20, No. 4. P. 24–29. EDN DSGPZI.
17. **Свидетельство** о государственной регистрации программы для ЭВМ № 2017617742 Российская Федерация. Система дистанционного профориентационного тестирования старших школьников «ПРОФСИЛА» : № 2017614461 : заявл. 16.05.2017 : опубл. 11.07.2017 / А. И. Каптерев; заявитель Государственное автономное образовательное учреждение высшего образования города Москвы «Московский городской педагогический университет». EDN EFEIYY.
18. **Крупa Т. В.** Новый подход к процессу автоматизации обучения на основе данных о поведении пользователей в цифровых библиотеках // *Научные и технические библиотеки.* 2022. № 4. С. 126–136. <https://doi.org/10.33186/1027-3689-2022-4-126-136>.
19. **Онлайн-оценка** кандидатов и персонала // ONTARGET. URL: <https://ontarget.ru/> (дата обращения: 09.01.2024).
20. **Виноградова Е. В., Полякова Т. А., Минбалеев А. В.** Цифровой профиль: понятие, механизмы регулирования и проблемы реализации // *Правоприменение.* 2021. № 4. URL: <https://clck.ru/hFN47i> (дата обращения: 09.01.2024).
21. **Попану Д. Г.** Правовое понятие цифровой личности человека // *Трибуна учёного.* 2023. № 10. С. 96–103.
22. **Внесение** шума в данные позволяет использовать их, не раскрывая личную информацию // *Открытые системы.* Официальный сайт. URL: <https://www.osp.ru/articles/2023/0131/13058040> (дата обращения: 11.01.2024).

23. **Соловьёва Л. Н.** Цифровая идентичность как новый вид идентичности человека информационной эпохи // Общество: философия, история, культура. 2018. № 12. С. 40–43.
24. **ГОСТ Р 7.0.10-2019 (ИСО 15836-1:2017).** Набор элементов метаданных «Дублинское ядро». Основные (ядерные) элементы // Электронный фонд правовой и научно-технической информации / информ. компания «Кодекс». URL: <https://docs.cntd.ru/document/1200163563?ysclid=lrhy3k1jvd51627693> (дата обращения: 11.01.2024).
25. **OpenAIRE-Advance** helps move Europe further down the open science path. URL: <https://ec.europa.eu/research-andinnovation/en/projects/success-stories/all/openaire-advance-helps-move-europefurther-down-open-science-path> (available: 17.01.2024).
26. **Антопольский А. Б.** Эволюция сервисов научной информации в условиях открытой цифровой науки // Информация и инновации. 2023. Т. 18, № 1. С. 5–20. DOI 10.31432/1994-2443-2023-18-1-5-20. EDN SQZKDE.
27. **Антопольский А. Б.** Связанные открытые данные в цифровой гуманитаристике (обзор публикаций) // Научно-техническая информация. Серия 1: Организация и методика информационной работы. 2022. № 5. С. 23–30. DOI 10.36535/0548-0019-2022-05-4. EDN IJGUQW.
28. **Richard Gartner.** Metadata in the Digital Library: Building an Integrated Strategy with XML: Facet Publishing; 1st edition (November 2, 2021). London. 200 p. 978–1783304851
29. **Seth van Hooland and Ruben Verborgh.** Linked Data for Libraries, Archives and Museums. How to Clean, Link and Publish Your Metadata. Chicago, 2014. 254 p. 978-0-8389-1251-5
30. **Verborgh R. and De Wilde M.** (2013) Using OpenRefine, Packt Publishing. Birmingham-Mumbai. 114 p.
31. **Linked Data for the Perplexed Librarian (An ALCTS Monograph)** / American Library Association. https://ebook-hunter.org/linked-data-for-the-perplexed-librarian-by-unknown_5f306b1bbdfb9f0b1d269091/ (дата обращения: 13.01.2024).
32. **Assessment Strategies in Technical Services (An ALCTS Monograph)** / Kimberley A. EdwardsMichelle Leonard // <https://www.alastore.ala.org/content/assessment-strategies-technical-services-alcts-monograph%E2%80%94printe-book-bundle> (дата обращения: 13.01.2024).
33. **Heath T. and Bizer C.** (2011) Linked Data: evolving the web into a global data space, Morgan & Claypool.
34. **Wood D., Zaidman M. and Ruth L.** (2013) Linked Data: structured data on the web, Manning.
35. **Allemang D. and Hendler J.** (2011) Semantic Web for the Working Ontologist, Morgan Kaufmann.

References

1. **Strategicheskoe napravlenie v oblasti tsifrovoy` transformatsii otrasli kul`turny` Rossijskoi` Federatsii do 2030 goda.** Utv. rasporyazheniem Pravitel'stva Rossii'skoi` Federatsii ot 11 dekabria 2023 g. № 3550-r.

2. **Antopol'skii A. B.** Informatcionny'e ob`ekty` tcifrovoi` gumanitaristiki: voprosy` ti-pologii // Informatcionny'e resursy` Rossii. 2023. № 1(190). S. 56–69. DOI 10.52815/0204-3653_2023_1190_56. EDN SSCEVU.
3. **Goncharov M. V., Kolosov K. A.** Voprosy` raschyota al'tmetrik na osnove danny`kh, formiruemy`kh pri obrashchenii pol'zovatelei` k e`lektronny`m katalogam bibliotek // Nauchny'e i tekhnicheskie biblioteki. 2020. № 11. S. 73–88.
4. **Karaush A. S.** Perspektivny'e napravleniia razvitiia informatcionny`kh tekhnologii` v publichny`kh bibliotekakh // Informatcionny'e tekhnologii, komp'iuterny'e sistemy` i izdatel'skaia produktsiia dlia bibliotek : materialy` konferentsii «LIBCOM-2010» : e`lektronny`i resurs, Ershovo, 15–19 noiabria 2010 g. Ershovo : Gosudarstvennaia publichnaia nauchno-tekhnicheskaia biblioteka Rossii, 2010. S. 28–32. EDN ZBAECH.
5. **Lavryonova O. A., Vinberg A. A.** Sovremenny'e pol'zovateli bibliotek i prostranstvo svyazanny`kh otkry`ty`kh danny`kh // Bibliotekovedenie. 2020. T. 69, № 3. S. 243–260. DOI 10.25281/0869-608X-2020-69-3-243-260.
6. **Neshcheret M. Iu.** Bibliotechno-informatcionny'e uslugi v lichnom kabinete pol'zovatelja biblioteki // Bibliotekovedenie. 2023. T. 72, № 3. S. 213–223. DOI 10.25281/0869-608X-2023-72-3-213-223.
7. **Sergeeva N. A.** Biblioteka online/offline: na primere Rossii`skoi` natsional'noi` biblioteki (RNB) // Evrazijskii` Soiuz Uchyony`kh. Seriya: pedagogicheskie, psihologicheskie i filosofskie nauki. # 8 (101), 2022. S. 33–37.
8. **Shrai'berg Ia. L.** Osobenny'e komponenty` tcifrovoi` transformatsii obshchestva, aktivno vliiaiuschie na tekhnologicheskie i povedencheskie modeli deiatel'nosti sovremenny`kh bibliotek (Ezhgodny`i` doclad Sed'mogo mezhdunarodnogo professional'nogo foruma «Sochi-2023») // Nauchny'e i tekhnicheskie biblioteki. 2023. № 8. S. 13–84. DOI 10.33186/1027-3689-2023-8-13-84. EDN FTIDSG.
9. **Birdi B.** (2011) Investigating fiction reader characteristics using personal construct theory. *Aslib Proceedings*, 63 (2–3). 275–294. ISSN 0001-253X. <http://dx.doi.org/10.1108/00012531111135709> (data obrashcheniia: 25.01.2024).
10. **Xiao Lin, Wenjuan Guan, Ying Zhang.** Application of Data Mining Technology with Improved Clustering Algorithm in Library Personalized Book Recommendation System // *International Journal of Advanced Computer Science and Applications*, Vol. 14, No. 11. 2023. P. 494–504.
11. **Sarma D., Mittra T., Hossain M. S.** Personalized book recommendation system using machine learning algorithm. *International Journal of Advanced Computer Science and Applications*, 2021, 12 (1):212–219.
12. **Hafid Ahmad Adyatma, Z. K. A. Baizal.** Book Recommender System Using Matrix Factorization with Alternating Least Square Method // *Journal of Information System Research (JOSH)*. Volume 4, No. 4, Juli 2023, pp. 1286–1292. ISSN 2686-228X (media online). <https://ejournal.seminar-id.com/index.php/josh/> DOI 10.47065/josh.v4i4.3816.
13. **Porcel C., Moreno J. M., Herrera-Viedma E.** A multi-disciplinar recommender system to advice research resources in University Digital Libraries // *Expert Systems with Applications* 36 (2009). P. 12520–12528.

14. **Best Roger J.** Market-Based Management: Strategies for Growing Customer Value and Profitability. Pearson College Div; 5th edition. 2008. 515 p.
15. **Levitckaia I. A.** Teoriia lichnostny'kh konstruktov Dzh. Kelli: na puti k kognitivnoi' filosofii obrazovaniia // Vestneyk Nizhegorodskogo universiteta im. N. I. Lobachevskogo. Ser. Sotcial'ny'e nauki. 2013. № 2 (30). S. 114–119.
16. **Kapterev A.** Professional Environment as an Information Analysis Object // Knowledge Organization. 1993. Vol. 20, No. 4. P. 24–29. EDN DSGPZl.
17. **Svidetel'stvo** o gosudarstvennoi' registratsii programmy` dlia E`VM № 2017617742 Rossii'skaia Federatsiia. Sistema distantsionnogo proforientatsionnogo testirovaniia starshikh shkol'nikov «PROFSILA» : № 2017614461 : zaiavl. 16.05.2017 : opubl. 11.07.2017 / A. I. Kapterev; zaiavitel' Gosudarstvennoe avtonomnoe obrazovatel'noe uchrezhdenie vy'sshego obrazovaniia goroda Moskvy' «Moskovskii' gorodskoi' pedagogicheskii' universitet». EDN EFEIYY.
18. **Krupa T. V.** Novy'i' podhod k protsessu avtomatizatscii obucheniia na osnove danny'kh o povedenii pol'zovatelei' v tcifrov'y'kh bibliotekakh // Nauchny'e i tekhnicheskie biblioteki. 2022. № 4. S. 126–136. <https://doi.org/10.33186/1027-3689-2022-4-126-136>.
19. **Onlai'n-ocenka** kandidatov i personala // ONTARGET. URL: <https://ontarget.ru/> (data obrashcheniia: 09.01.2024).
20. **Vinogradova E. V., Poliakova T. A., Minbaleev A. V.** Tcifrovoi' profil': poniatie, mehanizmy' regulirovaniia i problemy' realizatscii // Pravoprimenenie. 2021. № 4. URL: <https://clck.ru/hFN47i> (data obrashcheniia: 09.01.2024).
21. **Popanu D. G.** Pravovoe poniatie tcifrovoi' lichnosti cheloveka // Tribuna uchyonogo. 2023. № 10. S. 96–103.
22. **Vnesenie** shuma v danny'e pozvoliaet ispol'zovat' ikh, ne raskry'vaia lichnuiu informatsiiu // Otkry'ty'e sistemy'. Ofitsial'ny'i' sai't. URL: <https://www.osp.ru/articles/2023/0131/13058040> (data obrashcheniia: 11.01.2024).
23. **Solov'yova L. N.** Tcifrovaia identichnost' kak novy'i' vid identichnosti cheloveka informatsionnoi' e'pohi // Obshchestvo: filosofii, istoriia, kul'tura. 2018. № 12. S. 40–43.
24. **GOST R 7.0.10-2019** (ISO 15836-1:2017). Nabor e'lementov metadanny'kh «Dublinskoe iadro». Osnovny'e (iaderny'e) e'lementy' // E'lektronny'i' fond pravovoi' i nauchno-tekhnicheskoi' informatsii / inform. kompaniia «Kodeks». URL: <https://docs.cntd.ru/document/1200163563?ysclid=lrhy3k1jvd51627693> (data obrashcheniia: 11.01.2024).
25. **OpenAIRE-Advance** helps move Europe further down the open science path. URL: <https://ec.europa.eu/research-andinnovation/en/projects/success-stories/all/openaire-advance-helps-move-europefurther-down-open-science-path> (available: 17.01.2024).
26. **Antopol'skii' A. B.** E'voliutsiia servisov nauchnoi' informatsii v usloviakh otkry'toi' tcifrovoi' nauki // Informatciia i innovatsii. 2023. T. 18, № 1. S. 5–20. DOI 10.31432/1994-2443-2023-18-1-5-20. EDN SQZKDE.
27. **Antopol'skii' A. B.** Sviazanny'e otkry'ty'e danny'e v tcifrovoi' gumanitaristike (obzor publikatsii) // Nauchno-tekhnicheskaiia informatciia. Seriia 1: Organizatsiia i metodika informatsionnoi' raboty'. 2022. № 5. S. 23–30. DOI 10.36535/0548-0019-2022-05-4. EDN IUGUQW.

28. **Richard Gartner.** Metadata in the Digital Library: Building an Integrated Strategy with XML: Facet Publishing; 1st edition (November 2, 2021). London. 200 p. 978–1783304851
29. **Seth van Hooland and Ruben Verborgh.** Linked Data for Libraries, Archives and Museums. How to Clean, Link and Publish Your Metadata. Chicago, 2014. 254 p. 978-0-8389-1251-5
30. **Verborgh R. and De Wilde M.** (2013) Using OpenRefine, Packt Publishing. Birmingham-Mumbai. 114 p.
31. **Linked Data for the Perplexed Librarian** (An ALCTS Monograph) / American Library Association. https://ebook-hunter.org/linked-data-for-the-perplexed-librarian-by-unknown_5f306b1bbdfb9f0b1d269091/ (data obrashcheniia: 13.01.2024).
32. **Assessment Strategies in Technical Services** (An ALCTS Monograph) / Kimberley A. Edwards-Michelle Leonard // <https://www.alastore.ala.org/content/assessment-strategies-technical-services-alcts-monograph%E2%80%94printe-book-bundle> (data obrashcheniia: 13.01.2024).
33. **Heath T. and Bizer C.** (2011) Linked Data: evolving the web into a global data space, Morgan & Claypool.
34. **Wood D., Zaidman M. and Ruth L.** (2013) Linked Data: structured data on the web, Manning.
35. **Allemang D. and Hendler J.** (2011) Semantic Web for the Working Ontologist, Morgan Kaufmann.

Информация об авторе / Author

Каптерев Андрей Игоревич – доктор социол. наук, доктор пед. наук, профессор, главный научный сотрудник ГПНТБ России, главный научный сотрудник Российской государственной библиотеки, профессор Московского городского педагогического университета, Москва, Российская Федерация
kapterev@narod.ru

Andrey I. Kapterev – Dr. Sc. (Sociology), Dr. Sc. (Pedagogy), Professor, Chief Researcher of Russian National Public Library for Science and Technology, Chief Researcher of Russian State Library; Professor of Moscow City Pedagogical University, Moscow, Russian Federation
kapterev@narod.ru