

О роли и месте институционального репозитория в цифровой экосистеме научной организации

И. А. Филозова, Т. Н. Заикина

Аннотация—В любой деятельности человек постоянно сталкивается с необходимостью поиска, обработки, хранения и передачи информации. В глобальном социальном процессе информатизации общества информация является одним из основных ресурсов его развития, а информационные системы и технологии используются во всех областях человеческой деятельности. В науке и образовании важнейшая роль отводится научным информационным системам. В широком смысле такие системы предназначены для автоматизации научной, научно-образовательной и научно-организационной деятельности научных коллективов. Они весьма разнообразны, среди них есть системы общего назначения, используемые во всех областях науки, есть узкоспециализированные системы, имеющие специфическое применение в определенной предметной области. К этому классу информационных систем относятся, в частности, наукометрические базы данных, автоматизированные библиотечно-информационные системы, информационные системы для организации и проведения научных мероприятий, информационные системы для совместной работы исследователей, репозитории открытого доступа (OAR), информационные системы поддержки текущих исследований (Current Research Information System, CRIS-системы), библиографические менеджеры, системы идентификации авторов. В данной статье рассматривается особый тип научных информационных систем — институциональные репозитории публикаций — и их место и роль в цифровой экосистеме крупной научной организации, ключевые требования и характеристики, которым они должны удовлетворять.

Ключевые слова—институциональный репозиторий, информационные системы, научная коммуникация, цифровая экосистема.

Статья получена 13 августа 2024.

И. А. Филозова — начальник Группы развития и сопровождения информационных систем общеинститутского назначения, Лаборатория информационных технологий, Объединенный институт ядерных исследований (141980 г. Дубна, Московская обл., ул. Жолио-Кюри, 6); старший преподаватель, Государственный университет «Дубна», Институт системного анализа и управления (141980, Россия, Московская обл., г. Дубна, ул. Университетская, 19), ORCID: <https://orcid.org/0000-0003-3441-7093>, fia@jinr.ru.

Т. Н. Заикина — инженер-программист 2 категории Лаборатория информационных технологий, Объединенный институт ядерных исследований (141980 г. Дубна, Московская обл., ул. Жолио-Кюри, 6), ORCID: <https://orcid.org/0000-0003-0805-7995>, ztanya@jinr.ru.

I. ВВЕДЕНИЕ

Необходимость в системной обработке научно-технической информации возникла, когда наука стала доступным занятием для широких масс населения, т.е. примерно с начала XX в. Активная автоматизация началась в 1970-1980-х гг., как и в других областях, например, в это время появились первые автоматизированные библиотечные информационные системы (АБИС), пришедшие на смену традиционным библиотечным технологиям. Глобальный процесс информатизации уже охватил все сферы человеческой деятельности, вследствие чего в современном обществе накоплены огромные массивы информации в цифровой форме, в т.ч. и научно-технической. Согласно источнику [1] к началу 2014 г. в мире издавалось 34 274 научных рецензируемых журналов, тогда как их история началась всего с двух европейских изданий в 1665 г. Это были французский журнал «Le Journal des Scavans» и английский «Philosophical Transactions». Охват и изучение такого объема научных работ в традиционных печатных форматах представляется трудно выполнимой задачей для современного исследователя. Рост объемов научно-технической информации и перевод ее в цифровую форму определил насущную потребность в автоматизации целого ряда информационных процессов. На стыке XX-го и XXI вв. произошла глубокая трансформация научной коммуникации в связи с появлением и бурным развитием сетевых и цифровых технологий, формированием и эволюцией открытых движений (в т.ч. открытой науки), информатизации общества. Появились новые формы и средства, без которых развитие современной науки уже трудно представить [2].

II. НАУЧНЫЕ ИНФОРМАЦИОННЫЕ СИСТЕМЫ

Научные информационные системы представляют собой достаточно широкий класс систем, предназначенных для обеспечения информационной работы научных и исследовательских коллективов. Они очень разнообразны и имеют широкий спектр применения. К таким системам относят: наукометрические базы данных, автоматизированные библиотечные системы, информационные системы для организации и проведения научных конференций, информационные системы для совместной работы исследователей,

репозитории Открытого доступа (Open Access Repository), информационные системы поддержки текущих исследований (Current Research Information System, CRIS-системы).

А. Наукометрические базы данных

Наукометрическая база данных (НМБД) — библиографическая и реферативная база данных с инструментами для отслеживания цитируемости статей, опубликованных в научных изданиях [3]. В такие базы данных вносятся метаданные публикаций и их полные тексты (при доступности), изданные в журналах, которые индексируются данной НМБД. Наукометрические базы данных выполняют расчет основных наукометрических показателей, к которым относят индекс цитирования, индекс Хирша и его модификации, а также импакт-фактор.

Индекс цитирования — мера значимости научной работы ученого или научного коллектива. Значение индекса цитирования определяется количеством ссылок на публикацию или фамилию автора в других источниках.

Индекс Хирша (h-индекс) — количественная характеристика ученого, основанная на количестве его публикаций и количестве цитирований этих публикаций. Например, ученый имеет индекс Хирша 5, если 5 из его статей цитируются как минимум 5 раз каждая.

Импакт-фактор показывает, сколько раз в среднем цитируется каждая опубликованная в журнале статья в течение двух последующих лет после выхода.

Наиболее популярными и значимыми НМБД являются Российский индекс научного цитирования (РИНЦ), Web of Science и Scopus. Также широкую известность получили платформы Publons и Google Scholar.

Наукометрические базы данных являются общепринятым инструментом работы с научно-технической информацией в разных областях научной деятельности.

В. Информационные системы для организации и проведения научных мероприятий

Деятельность любой научной или образовательной организации не ограничивается основным видом деятельности и включает подготовку и проведение научных мероприятий — конференций, школ, симпозиумов, рабочих совещаний и т.п. различного масштаба — от локальных до международных. Эти мероприятия проводятся как правило на регулярной основе и требуют от оргкомитета значительных усилий. Автоматизация ряда процессов подготовки и проведения конференции значимо облегчает работу оргкомитета и делает ее более эффективной. Существует значительное количество программного обеспечения для управления жизненным циклом научных мероприятий. Среди наиболее известных ИС поддержки конференции можно назвать EasyChair Conference System (<http://www.easychair.org>) российский портал Agora (<http://agora.guru.ru>), Интернет-портал «Конференции Уральского отделения РАН» (<http://confer.uran.ru/>),

Indico (©CERN) (<http://indico-software.org/>), Microsoft's Academic Conference Management Service (CMT) [4]. Типовыми бизнес-процессами конференции являются (<https://cmt3.research.microsoft.com/>):

- создание и поддержка сайта мероприятия;
- подача и регистрация заявок;
- организация рецензирования заявок;
- информационная поддержка: публикация списка участников, списка принятых докладов;
- загрузка тезисов и презентаций участников;
- рассылка сообщений и уведомлений;
- формирование отчетов;
- составление и коррекция программы конференции;
- учет оплаты регистрационных взносов;
- размещение иногородних участников.

В то же время каждое научное мероприятие имеет свою специфику, соответственно, бизнес-процессы и требования к информационному сопровождению будут отличаться. Наиболее полный охват перечисленных выше функций обеспечивает ПО Indico, что подтверждается его применением по всему миру. Кроме ЦЕРНа многие научные коллаборации, международные организации используют данный продукт в качестве инструмента для организации встреч и обмена информацией. Среди пользователей Indico такие всемирно известные научные центры как DESY (Deutsches Elektronen-Synchrotron, Германия), ОИЯИ, INFN (Istituto Nazionale di Fisica Nucleare, Италия) и многие другие. Пользователем Indico является ООН.

С. CRIS системы

Особую актуальность для научно-исследовательских организаций имеют Current Research Information Systems — информационные ресурсы для хранения, управления и обмена метаданными о проводимых в научных организациях разработках [5], [6]. Данные системы также известны как RIMS (Research Information Management Systems) [7]. Ядро CRIS систем опирается на модель данных Общеευропейского формата исследовательской информации CERIF, разработанную и поддерживаемую некоммерческой организацией euroCRIS. Публикации являются важнейшим элементом научной деятельности, но в ней также присутствуют и другие элементы, нуждающиеся в описании. Поэтому, кроме публикаций, в модель также включены такие сущности как профили ученых, оборудование, проекты, менеджеры проектов, подразделения, финансирующие организации, результаты исследований и связи между ними [8].

Среди известных CRIS-систем есть как собственные разработки, так и решения на базе программного обеспечения с открытым исходным кодом (например, DSpace-CRIS). Примерами таких ресурсов являются информационные системы Research Portal Уральского федерального университета им. первого Президента России Б. Н. Ельцина (<https://sciencedata.urfu.ru/portal/ru/>), ИСТИНА Московского государственного университета (<https://istina.msu.ru/>), Астраханского государственного

университета (<https://science.asu.edu.ru/>). В разных странах мира существует десятки ресурсов на основе программной платформы DSpace-CRIS (<https://wiki.lyrasis.org/display/DSPACECRIS/DSpace-CRIS+Users>).

Важным аспектом является совместимость CRIS систем и институциональных репозиториев (Institutional repository, IR) — архивов для сбора, хранения и распространения цифровых копий интеллектуальной продукции организации. Т.е. должны быть обеспечены рабочие процессы обмена информацией между CRIS и IRs. Ранее эти два типа систем считались конкурирующими друг с другом, но в настоящее время совершенствование механизмов обмена информацией о результатах исследований и связанных с ними метаданных позволило интегрировать их в единую систему, которая получила название CRIS&OAR (Open Access Repository).

D. Репозитории открытого доступа

Репозитории открытого доступа представляют собой общедоступные архивы цифровых объектов различной природы с развитым инструментарием для поиска. В научных и образовательных организациях в таких архивах представлена интеллектуальная продукция сотрудников (как правило, научные публикации). Создание и развитие репозиториев важно по двум основным причинам: они являются неотъемлемой частью инфраструктуры открытой науки и обеспечивают доступность результатов научных исследований. Основное преимущество репозиториев в оперативности распространения научных результатов, что позитивно влияет на цитируемость и на показатели альтернативных метрик публикации [9]. Кроме того, наукометрические базы данных и подобные ресурсы по отдельному взятому автору в общем случае не могут обеспечить полное покрытие первоисточников его публикационной активности.

Традиционно репозитории классифицируют по охвату контента на институциональные (архивы организаций) и тематические (предметные) — хранилища публикаций исследователей из разных организаций по одной или смежным дисциплинам, а также по масштабу — национальные и международные [10], [11]. Они могут существовать и вне модели Открытого доступа [12].

Институциональный репозиторий (ИР) / Электронный архив организации — информационная система для накопления, долговременного хранения информации и обеспечения надежного доступа к цифровым объектам, представляющим собой результат интеллектуальной деятельности научного или образовательного учреждения. Цель институционального репозитория — представление научного потенциала организации в глобальной сети. ИР обеспечивает доступность научных, научно-организационных и образовательных материалов аффилированных с организацией авторов. Это могут быть статьи, препринты, диссертации, технические отчеты, рабочие заметки и документы, труды конференций, выпускные квалификационные работы

студентов и другие цифровые публикации и/или документы. Основными задачами институционального репозитория являются:

- 1) Обеспечение систематического сбора и надежного долговременного хранения в цифровом виде метаданных и полных текстов публикаций.
- 2) Предоставление набора сервисов для доступа и управления информационными ресурсами сотрудникам организации и неавторизованным пользователям.

Организация-держатель репозитория несет полную ответственность за размещаемый в нем контент.

Многие научные и образовательные учреждения создают и поддерживают репозитории публикаций своих сотрудников и используют предметные архивы в разных областях науки [5], [13], [14], [15]. В качестве примера предметного репозитория можно привести крупнейшую информационную платформу в области физики высоких энергий INSPIRE (<https://inspirehep.net/>). Поддержка и развитие INSPIRE осуществляется в сотрудничестве с CERN (Европейская организация по ядерным исследованиям, Швейцария), DESY (Немецкий Электронный Синхротрон, Германия), Fermilab (Национальная ускорительная лаборатория им. Энрико Ферми, США), IHEP (Институт ядерной физики, Китай), IN2P3 (Национальный институт ядерной физики и физики элементарных частиц, Франция) и SLAC (Центр Стэнфордского линейного ускорителя, США). На данный момент платформа содержит сведения о более 1,600 млн. публикаций, более 700 тыс. авторов, около 12 тыс. научных организаций, более 3,8 тыс. научных экспериментах, около 4 тыс. научных журналов. Институциональный репозиторий ЦЕРНа CERN Document Server (<https://cds.cern.ch>) обеспечивает доступ к более, чем 0,5 млн. публикаций. Аналогичный архив создан и поддерживается в DESY (<https://bib-pubdb1.desy.de/>). Опыт создания и эксплуатации институциональных репозиториев есть и в России, например, электронный научный архив УрФУ (<https://elar.urfu.ru/>), Цифровой репозиторий Южного федерального университета (<https://hub.sfedu.ru/repository/>), открытый репозиторий Института биологии южных морей имени А. О. Ковалевского РАН (<https://repository.marine-research.ru/home>) и др.

E. Другое ПО

Рассмотренные выше виды информационных систем и сервисов, применяемых в научной деятельности, далеко не полностью покрывают потребности и информационные нужды коллективов исследователей. В зависимости от области научной деятельности, организации внедряют и эксплуатируют самое разнообразное программное обеспечение, начиная от систем бухгалтерского и кадрового учета до специализированных продуктов, автоматизирующих специфические бизнес-процессы предприятия.

В частности, в организациях, где ведутся ИТ-разработки, активно используются инструменты для

совместной работы исследователей, сервисы для хостинга IT-проектов и их совместной разработки, такие как GitHub [16], [17] и GitLab [18].

Коллективную разработку программного обеспечения трудно эффективно организовать без применения методов и инструментов управления проектами [19], [20], [21]. В этой связи широко распространена практика использования таких программных продуктов как Trello, Bitrix, Yougile, Яндекс Трекер, Planigo, Jira и др. [22].

В физических экспериментах (численность крупных коллабораций достигает нескольких сотен и тысяч человек) стандартной практикой является использование сервисов для управления и обмена научно-технической документацией с ведением версионности. Это не только официальные документы, но и рабочая документация (отчеты, заметки и т.п.), требующая специальной маркировки о внесенных в данную версию изменениях. Очень популярным в физическом экспериментальном сообществе был сервер документов для совместной работы DocDB (<https://github.com/ericvaandering/DocDB>), разработанный для коллаборации ВТeV в Fermilab (<https://www.fnal.gov/>) [23]. Несмотря на то, что сервис морально устарел, он продолжает использоваться по сей день.

У научных организаций в области экспериментальной физики также есть потребности в сфере автоматизации. Сегодня невозможно представить ни один современный крупный физический эксперимент без автоматизации сбора, хранения, обработки и анализа экспериментальных данных. Информационные системы для поддержки обработки и анализа данных и других видов коллаборационной деятельности являются насущной необходимостью и обеспечивают успешность работы коллаборации. Кроме того, каждая физическая установка крупного эксперимента уникальна, что порождает потребность в специализированных информационных системах для онлайн и офлайн обработки данных в современных физических экспериментах [24].

III. РОЛЬ И МЕСТО ИНСТИТУЦИОНАЛЬНЫХ РЕПОЗИТОРИЕВ В ЦИФРОВОЙ ЭКОСИСТЕМЕ НАУЧНОЙ ОРГАНИЗАЦИИ

В научной организации, особенно крупной, протекает масса информационных процессов по поиску, сбору и обработке научно-технической информации как из внешних источников, так и из локальных информационных систем. В идеале все многообразие эксплуатируемых в организации информационных систем должно быть интегрировано в единую корпоративную информационную систему (КИС), охватывающую основные бизнес-процессы организации на общем информационном пространстве.

Типичный состав КИС в широком смысле может включать такие информационные системы как система управления ресурсами предприятия, система управления персоналом, система управления проектами, система автоматизированного проектирования (CAD/CAM/CAE), система электронного документооборота (Electronic Document Management System), система организации

рабочего пространства, среда Internet / Intranet, система управления информационными ресурсами — управление корпоративным контентом (Enterprise content management, ECM), система управленческой отчетности (MIS), специализированные рабочие места пользователей, системы моделирования и представления бизнес-процессов, системы математического и имитационного моделирования процессов, системы математического (в т.ч. статистического) анализа данных, специализированные продукты или системы для реализации частных задач, корпоративные порталы (обработка и доставка информации, доступ к разнообразным сервисам), мобильные приложения (доступ к корпоративным информационным каналам в любое время в любом месте), системы совместной работы сотрудников и многие другие. Каждое из перечисленных IT-решений в свою очередь также является корпоративной информационной системой в узком смысле.

Структура КИС хорошо иллюстрируется в виде двухуровневой управленческой пирамиды [25], в которой нижний уровень соответствует оперативному слою, а верхний — стратегическому (Рис.1). На вход пирамиды поступает информация об основных управляемых ресурсах (финансовых, материальных, кадровых, информационных), на выходе — результат основной деятельности организации. По мере движения информационного потока вверх по пирамиде происходит структурирование первичной информации, ее обезличивание, свертка и фильтрация. На вершине управленческой пирамиды формируются сведения для высшего руководства организации, генерируются отчеты, отражающие показатели, необходимые для выработки стратегических решений по управлению и развитию.

Научные информационные системы (системы класса CRIS&OAR) располагаются на стыке операционного и стратегического слоев пирамиды КИС (Рис.1). Ядро данных составляют библиографическая информация, данные об аффилиации авторов и информация о темах исследований (научных программах, проектах). Компонент CRIS характеризуется как административная, комплексная, интегративная, персоно-центричная, аналитическая система для управления кадровой информацией, данными об исследованиях и разработках (НИОКР), организационных подразделений, экспериментальных установках и оборудовании. Компонент OAR представляет собой документо-центричную систему открытого доступа для сбора и долговременного хранения результатов исследований и разработок (публикаций), предоставление услуг членам организации по управлению и распространению цифровых ресурсов [26]. С помощью CRIS&OAR систем научные организации представляют в реальном времени данные о публикациях своих сотрудников, о научных исследованиях и проектах, их результатах; ведут учет публикационной активности аффилированных с ними авторов; производят

оперативный мониторинг научной деятельности и анализируют наукометрические показатели [27]. CRIS&OAR система в рамках КИС должна взаимодействовать с другими ИС, в частности, с системой управления персоналом для импорта

актуальных данных о сотрудниках и структурных подразделениях, с системами управления корпоративным контентом, с системой архивирования и управления адресами электронной почты сотрудников (Email Archiving & Management) и др. Эти данные необходимы для формирования и актуализации профиля автора [28].

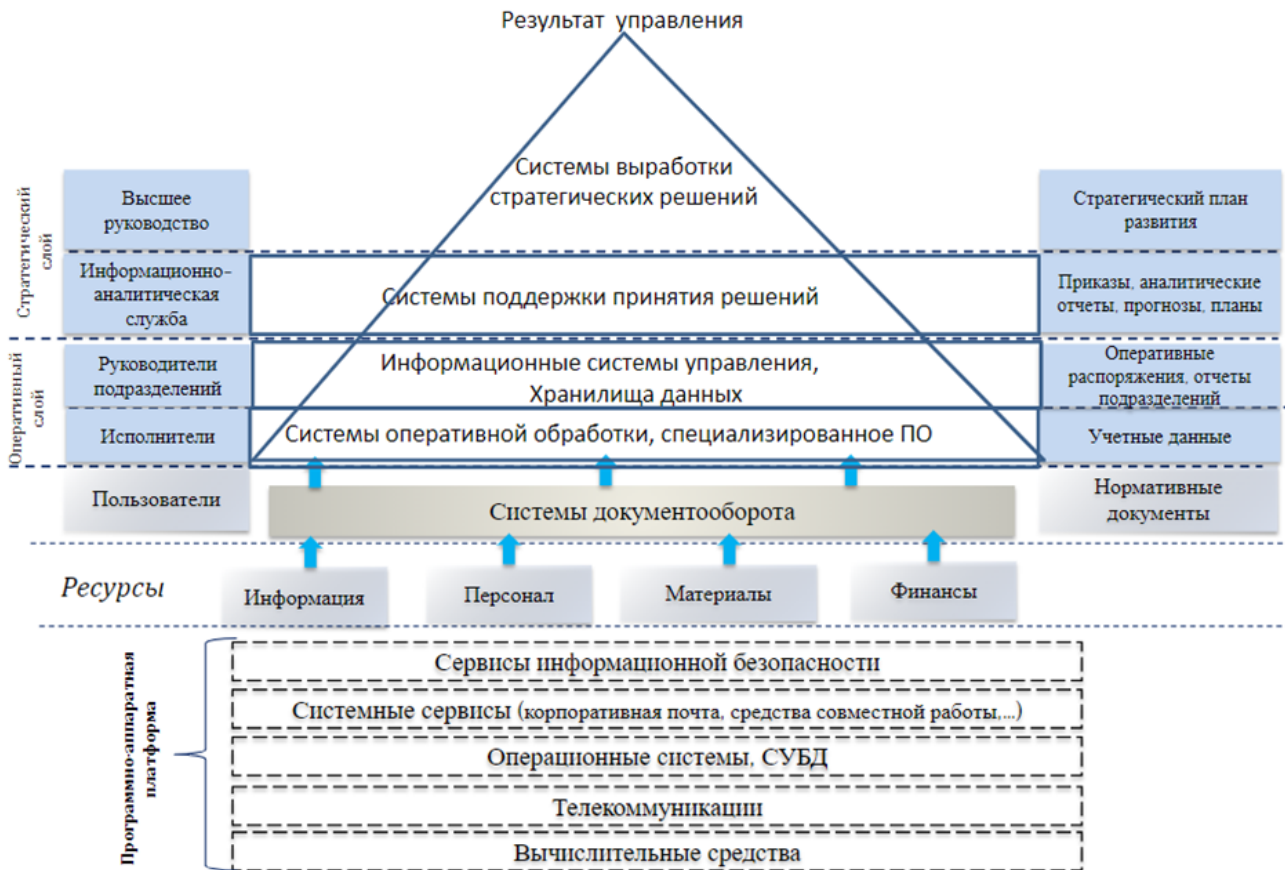


Рис.1 Структура корпоративной информационной системы

Основная функция CRIS&OAR — это обработка, хранение и обмен метаданными, характеризующими профессиональную активность и эффективность научной организации [6].

Происходящие в современной экономике процессы стимулируют бизнес к активному применению цифровых решений. Организациям приходится модернизировать свою инфраструктуру и внедрять цифровые платформы (ЦП) [27], [29]. Цифровая трансформация затрагивает не только бизнес, но и другие сферы человеческой деятельности — образование, здравоохранение, науку, сельское хозяйство, юриспруденцию, туризм и пр. Функционирование организаций как цифровых экосистем (ЦЭС) стало предметом активного интереса и обсуждения [30]. Большое количество публикаций на данную тему демонстрирует ее актуальность и интерес к ней специалистов в разных областях [31]–[35]. Построение и развитие экосистем в различных сферах деятельности является одним из основных трендов текущего момента. Традиционные бизнес-модели

становятся неконкурентоспособными и все чаще основной формой взаимодействия участников цифровой среды становятся экосистемы. Следуя данной тенденции, современные организации будут вынуждены либо создавать свои собственные экосистемы, либо интегрироваться в уже имеющиеся. Рассматривая структуру, границы и основные элементы цифровой экосистемы для бизнеса в работе [36] автор описывает несколько уровней ЦЭС. Проецируя данный подход на научную сферу, можно дать следующую графическую интерпретацию (Рис.2).

На первом уровне цифровой экосистемы находятся непосредственно пользователи (сотрудники организации-владельца ЦЭС и сотрудники организаций-партнеров) и цифровые платформы с собственными сервисами (институциональные цифровые платформы). Этот уровень как раз соответствует пирамиде КИС (Рис.1) как интегрированной информационной системе организации. На втором уровне находятся внешние цифровые сервисы, взаимодействующие с сервисами цифровых платформ организации-владельца ЦЭС, но не

принадлежащих ей. На третьем уровне располагаются продукты и услуги поставщиков разнообразных IT-решений, которые взаимодействуют с цифровыми сервисами второго уровня. Организации, интегрированные в цифровую экосистему, могут входить и в состав других экосистем или сами являться

экосистемами. Отличительной чертой ЦЭС является бесшовная цифровая среда, в которой функционируют сервисы организации-владельца ЦЭС и сервисы организаций-партнеров. Пользователи получают доступ к сервисам ЦЭС в режиме «одного окна».



Рис.2. Элементы ЦЭС

IV. ОСОБЕННОСТИ ИНСТИТУЦИОНАЛЬНОГО РЕПОЗИТОРИЯ ОТКРЫТОГО ДОСТУПА КРУПНОЙ НАУЧНОЙ ОРГАНИЗАЦИИ НА ПРИМЕРЕ ОИЯИ

Опыт построения и эксплуатации репозитория открытого доступа разных типов обобщен и представлен в работе [9]. Согласно данному исследованию, кроме основной функции репозитория — обеспечение долговременного хранения произведений, — возможности, предоставляемые в рассмотренных проектах, включают депонирование — размещение авторами метаданных своих публикаций и их полных текстов с учетом эмбарго, премодерацию (проверку метаданных на формальные признаки научной публикации), оценку качества метаданных и их чистоты, регистрацию пользователей, навигацию и поиск по контенту репозитория. Такие возможности как комментирование, доска обсуждений, коллективная работа над рукописью, рецензирование не являются встроенными функциями рассмотренных проектов. Репозитории ОД реализуют принципы FAIR [37]. А именно, любой объект, размещенный на платформе репозитория, должен быть легко обнаруживаемым, легкодоступным, интероперабельным и доступным для повторного использования. Таким образом, репозитории отвечают всем фундаментальным принципам открытой науки.

Объединенный институт ядерных исследований (ОИЯИ) — крупная международная межправительственная организация, членами-участниками которой сегодня являются 16 государств (<https://www.jinr.ru/about/>).

В состав ОИЯИ входит семь научных лабораторий: Лаборатория теоретической физики

им. Н. Н. Боголюбова, Лаборатория ядерных проблем им. В. П. Джелепова, Лаборатория ядерных реакций им. Г. Н. Флерова, Лаборатория физики высоких энергий им. В. И. Векслера и А. М. Балдина, Лаборатория нейтронной физики им. И. М. Франка, Лаборатория информационных технологий им. М. Г. Мещерякова, Лаборатория радиационной биологии. Каждая из лабораторий по масштабам исследований сопоставима с большим академическим институтом. Кроме научных лабораторий в структуру ОИЯИ также входят научно-организационные службы, административно-хозяйственные и производственные подразделения.

Штат ОИЯИ насчитывает около 4500 человек, из них более 1200 — научные сотрудники, в том числе действительные члены и члены-корреспонденты национальных академий наук, более 260 докторов и 570 кандидатов наук, около 2000 — инженерно-технический персонал.

Результаты научно-исследовательской деятельности коллективов подразделений ОИЯИ представляются на научных мероприятиях и в публикациях. Ежегодно в редакции журналов и оргкомитеты конференций ОИЯИ направляет более 1500 научных статей и докладов, которые представляют около 3000 авторов. Таким образом, ОИЯИ является крупным производителем интеллектуальной продукции, сбор и учет которой является нетривиальной задачей.

А. Цифровая экосистема ОИЯИ

В рамках реализации проекта Цифровой экосистемы ОИЯИ как комплексной цифровой среды, объединяющей большое число информационных сервисов и бизнес-процессов, в опытной эксплуатации

уже находятся несколько сотен различных сервисов, сгруппированных по пяти основным категориям: научные сервисы, административные сервисы, сетевые сервисы, информационные сервисы, другие. К группе научных сервисов относятся цифровые сервисы для поддержки научной деятельности, такие как вычислительные и программные сервисы, научное программное обеспечение (библиотеки программ, операционные системы, репозитории открытого программного обеспечения), Проблемно-тематический план ОИЯИ, ИС мониторинга показателей работы ОИЯИ, сервис организации научных мероприятий, сервисы научно-технической библиотеки ОИЯИ и др. В группу административных сервисов входят сервисы поддержки административной работы, такие как система электронного документооборота, сервис дистанционного заказа и получения справок из бухгалтерии и отдела кадров, Информационно-справочная система ОИЯИ, сервис работы с исходящими платежными поручениями по данным ИС, архив нормативных документов, входящая и исходящая корреспонденция, шаблоны приказов, документы основного делопроизводства, сервис обеспечения закупочной деятельности и др. В категорию сетевых сервисов включены: сервис Учетная запись SSO (общая информация, смена пароля, правила работы в сети), сервис управления подпиской на списки рассылки, перечень зарегистрированных на пользователе сетевых элементов (IP, MAC-адреса), почтовые ящики пользователя и т.п. В группу информационных сервисов входят: административная информационно-справочная система по кадрам, интерактивная карта ОИЯИ, телефонный справочник ОИЯИ, сервис планирования и учета экскурсий в ОИЯИ, Персональная ИНформация о сотрудниках (ПИН) и пр.

Цифровая экосистема обеспечивает доступ ко всем функционирующим на текущий момент сервисам через единую учетную запись пользователя в соответствии с разграничением прав доступа, оперативный и удобный поиск информации, бесшовное переключение пользователя между различными сервисами, входящими в цифровую экосистему.

В настоящее время исторически сложившиеся бизнес-процессы по учету сведений о публикационной активности сотрудников организованы таким образом, что эти данные хранятся в нескольких информационных системах: в АБИС научно-технической библиотеки, в электронных версиях ежегодно выпускаемых библиографических указателей работ сотрудников ОИЯИ, в ИС Персональная ИНформация о сотрудниках. Каждая из перечисленных ИС отражает особенности организации бизнес-процесса по сбору и учету сведений о публикациях сотрудников. Так, АБИС научно-технической библиотеки ориентирована прежде всего на технологические библиотечные процессы, такие как комплектование, каталогизация, создание читательских карточек, формуляров, выпуск бюллетеней и т.п. ИС

Персональная ИНформация о сотрудниках предназначена для сбора и анализа информации о результатах научной деятельности сотрудников. В ней хранятся сведения о публикациях, научной деятельности (участие в научных мероприятиях, их подготовке и проведении, научное руководство и консультирование аспирантов, дипломников). В обеих системах предусмотрен ручной ввод информации, что не позволяет исключить ошибки из-за потенциальной многовариантности представления данных о публикации. Несмотря на то, что существует методология каталогизации для специалиста, анализирующего содержание документа, не всегда результаты этого анализа совместимы с возможностями формата хранения и представления библиографических данных в ИС.

В. Ключевые требования к институциональному репозиторию публикаций ОИЯИ

Основу контента институционального репозитория составляют метаданные публикаций и их полные тексты (при наличии) аффилированных с данной научной организацией авторов. Для тематического репозитория при формировании контента важно обеспечить максимально полный охват предметной области и/или смежных областей.

К основным задачам современного институционального репозитория можно отнести:

- Систематический сбор метаданных публикаций сотрудников и их полных текстов (при наличии).
- Обеспечение надежного и долговременного хранения контента.
- Обеспечение доступности результатов научно-исследовательской деятельности для широкой общественности в режиме поиска (в т.ч. полнотекстового) и навигации с учетом разграничения прав.

Для успешного решения перечисленных задач, репозиторий должен удовлетворять ряду требований.

1) Требования к структуре и контенту репозитория
Обеспечение и управление структурированностью информационного пространства репозитория публикаций является одним из ключевых требований. Репозиторий должен обеспечивать хранение публикаций основных традиционных научных жанров — журнальных статей, препринтов, диссертаций и авторефератов, монографий и их глав, авторами которых являются (или являлись в прошлом) сотрудники ОИЯИ. Кроме того, в репозитории должны храниться профили автора (учетные карточки пользователей) и сведения об организационной структуре (подразделениях ОИЯИ). Контент должен быть структурирован по коллекциям, которые входят в тематические разделы. Описание разделов, коллекций и их элементов представлено в Таблице 1.

Структура репозитория должна быть масштабируемой, т.е. при включении в структуру новых разделов ее реструктуризация не должна приводить к потере

контента. Раздел Иное предназначен для обозначения потенциальных разделов, которые могут быть введены в будущем.

2) Требования к модели данных

Наиболее известная модель, применяемая для унификации и стандартизации данных научных исследований в CRIS-системах, является модель CERIF (Common European Research Information Format) —

общеевропейский формат для исследовательской информации [38]. В основе формата лежит гибкая модель данных для представления исследовательской информации на основе модели Entity-Relationship. Формат разработан и поддерживается Международной профессиональной ассоциацией разработчиков научных информационных систем euroCRIS (www.eurocris.org) [8].

Таблица 1. Структура ИР ОИЯИ

Разделы	Коллекции	Элементы коллекции
Персоны	Архив сотрудников ОИЯИ	Бывшие сотрудники (профили)
	Сотрудники	Сотрудники (профили)
Подразделения ОИЯИ	Административно-хозяйственные подразделения	Подразделение
	Лаборатории и научные подразделения	Подразделение
Публикации ОИЯИ	База данных публикаций сотрудников	Библиографические метаданные + полный текст (при наличии)
	Публикации, вносимые пользователем	
Иное	Тематические коллекции	Публикации по тематике коллекции

В модели данных CERIF определены сущности, используемые для описания научных исследований и взаимосвязи между ними — Персона, Организация, Публикация, Проект, Событие, Оборудование и т.д. (<https://eurocris.org/Uploads/Web%20pages/CERIF-1.6/documentation/MInfo.html>). Модель поддерживает многоязычность и является расширяемой, предназначена для обмена данными в гетерогенных распределенных средах. Гибкость и богатство семантики модели позволяет обеспечить интероперабельность CRIS-систем с репозиториями открытого доступа и с хранилищами открытых научных данных.

Таким образом, эта модель данных может быть основой для интеграции CRIS-систем, институциональных репозиториев и хранилищ открытых данных, а также обеспечивать надежное управление интеллектуальной собственностью учреждения (Рис. 3) [39].

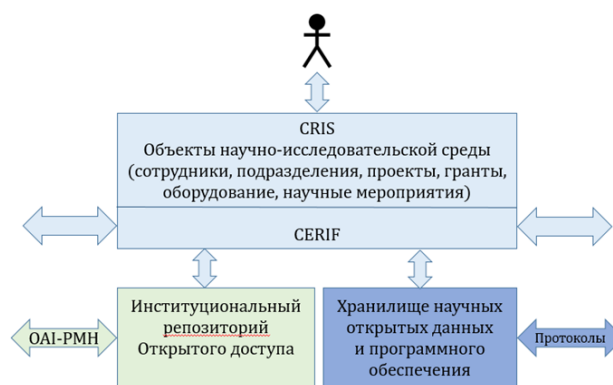


Рис. 3. Институциональные научные системы
OAI-PMH - The Open Archives Initiative Protocol for Metadata Harvesting¹

Также в CRIS&OAR системах широко применяется профессиональный библиотечный формат marc21 для записи библиографической информации на машиночитаемые носители и XML-версия структуры MARC — формат marcxml [40], [41].

3) Ключевая функциональность институционального репозитория

Основные потенциальные пользователи ИР принадлежат к трем категориям: неавторизованный пользователь (гость), автор, модератор (редактор) и администратор. В Таблице 2 представлено описание основных

¹ Протокол для обмена метаданными, позволяет создавать и автоматически обновлять контент репозитория, импортируя метаданные из различных источников одновременно.

пользователей и соответствующие им базовые функции.

Таблица 2. Основные пользователи и возможности ИР ОИЯИ

Роль	Описание доступных возможностей
Гость (неавторизованный пользователь)	Поиск и просмотр по общедоступным разделам репозитория (Подразделения ОИЯИ; Публикации ОИЯИ; Иное.)
Автор (зарегистрированный пользователь)	<ul style="list-style-type: none"> ✓ Авторизация по единой учетной записи SSO; ✓ Поиск и просмотр по всем разделам репозитория: Подразделения ОИЯИ; Публикации ОИЯИ; Персоны; Иное. ✓ Внесение публикаций (метаданных и полных текстов) в режиме самоархивирования.
Администратор ИР	Пользователь, наделенный максимальными правами, для поддержки репозитория в работоспособном состоянии.
Редактор	Пользователь системы, наделенный правами на просмотр и редактирование внесенных авторами данных. Выполняет премодерацию контента: проверку библиографической информации по формальным признакам. При необходимости вносит изменения в библиографические метаданные.

Кроме вышеуказанных в таблице 2 акторов, в ИР должны присутствовать функциональные архитектурные

элементы, представленные в Таблице 3.

Таблица 3. Ключевые элементы и функции ИР

Элемент	Функции
Подсистема хранения	Обеспечение надежного и долговременного хранения цифровых объектов (метаданных публикаций и их полных текстов).
Поисковая машина	Навигация и поиск по разделам и коллекциям репозитория с настройками поисковой выдачи и фильтрацией.
Подсистема курирования метаданных	<p>Курирование метаданных о сотрудниках и подразделениях. Актуализирует записи коллекций Персоны и Подразделения — справочников сотрудников ОИЯИ (потенциальных авторов научных публикаций) и структурных подразделений ОИЯИ всех уровней иерархии соответственно:</p> <ul style="list-style-type: none"> • Обновление данных о работающих сотрудниках: изменение электронной почты, добавление идентификаторов различных внешних систем идентификации авторов, таких как Inspire, ORCID, Scopus (при наличии). • Перенос метаданных неработающих (увольняемых) сотрудников в архивную коллекцию. • Внесение метаданных о новых (принятых на работу) сотрудниках. • Перевод сотрудника из одного подразделения в другое. • Обновление данных о действующих структурных подразделениях (переименование, перевод на другой уровень иерархии и т.п.). • Удаление метаданных ликвидируемых структурных подразделений. • Внесение метаданных о новых (вновь созданных) структурных подразделениях.
Подсистема пакетной загрузки	Формирование и преобразование контента в поддерживаемые форматы хранения и загрузка его в репозиторий в пакетном режиме.

4) Другие функциональные, нефункциональные требования и характеристики ИР

При работе с научными документами (публикациями типичных научных жанров) важно учитывать следующие артефакты.

Полный охват информационных источников подразумевает 1) извлечение метаданных публикаций аффилированных с организацией (ОИЯИ) авторов в

полном объеме из данного источника; 2) включение в систему новых источников по мере необходимости. Для реализации данного требования необходимо обеспечить корректное функционирование процедур ввода данных. Основные стратегии наполнения контента: ввод пользователем (внесение вручную), сбор данных посредством открытых и партнерских API из доверенных внешних источников и регистрация их в

системе (автоматический ввод).

Актуальность, полнота и достоверность происхождения документов. Обеспечение авторизации пользователей по единой учетной записи. Верификация вводимой информации (премодерация контента).

Наличие интеллектуальных служб обслуживания запросов пользователя. Обеспечение комплексного и точного информационного поиска с настройками и фильтрацией, включая комбинированный поиск по метаданным, полному тексту и ссылкам; удобная навигация.

Поддержка уникальных идентификаторов (URI, Uniform Resource Identifier) цифровых объектов для точной адресации ресурса. Возможность такой идентификации необходима для обеспечения интеграции объектов репозитория в глобальное информационное пространство.

Интероперабельность. Поддержка принятых стандартов метаданных для хранения, экспорта и импорта данных и протоколов обмена информации с другими информационными системами. К таким стандартам, в частности, относятся формат машиночитаемой каталогизационной записи marc (Machine-Readable Cataloging), marxml (xml-версия marc), набор элементов данных для описания ресурсов Dublin Core (dc).

Наличие графического пользовательского интерфейса. Пользователь работает с репозиторием посредством web-браузера без установки какого-либо дополнительного ПО.

Поддержка версииности хранимого элемента. Управление версиями позволяет хранить несколько версий одного элемента хранения.

Поддержка многоязычности. Возможность локализации как минимум на английский и русский языки, как максимум — на языки стран-участниц ОИЯИ.

Поддержка профилей авторов (пользователей) с набором характеристик: идентификационные номера автора различных систем идентификации, таких как ORCID, ResearcherId, InspireId и др.; варианты имени автора, указываемые в публикациях (библиографическое имя).

Поддержка разнообразия форматов размещаемых файлов: pdf; Open Office (ODS, ODP, ODT), Microsoft Office (DOC, DOCX, XLS, XLSX, PPT, PPTX), JPEG, PNG, GIF, TIFF.

Доступность и надежность. Репозиторий должен выполнять задачи в условиях реальной работы пользователей. К таким задачам относится поиск (простой и многокритериальный), внесение новых метаданных, редактирование метаданных, загрузка полных текстов. Необходимы процедуры резервного копирования и восстановления после сбоев.

Масштабируемость. Репозиторий должен хранить неограниченное количество записей (не менее 1 млн.) с годовым приращением 1000 ÷ 2000, обрабатывать файлы произвольных размеров.

Производительность. Время внесения записи в

репозиторий не должно превышать 5-6 сек. Должна быть обеспечена параллельная работа не менее 10 пользователей на внесение информации и не менее 100 пользователей на доступ (чтение).

Требования к организационному и методическому обеспечению. Поскольку потенциальными пользователями ИР являются все сотрудники организации, участвующие в подготовке и публикации научной продукции, требуется описать и задокументировать рабочие процессы внесения публикаций пользователем в репозиторий, модерации внесенных записей и нормативные документы, регламентирующие взаимодействие всех категорий пользователей с системой. Необходима встроенная система онлайн-помощи, контекстные подсказки без перенасыщения и загромождения интерфейса.

Также полезными и привлекательными для пользователя будут такие сервисы как:

Создание личных и групповых книжных полок. Книжная полка (корзина) позволяет формировать списки публикаций для их последующего анализа, добавляя их из контента репозитория или из поисковой выдачи.

Создание групп пользователей. Объединение пользователей в группы для каких-либо совместных действий, например, просмотра книжных полок, модерации.

Расылка оповещений. Пользователи могут настроить и получать на регулярной основе (ежедневно, еженедельно и т.д.) автоматические уведомления по электронной почте, например, о поступлении новых публикаций.

Обмен сообщениями позволяет пользователям отправлять друг другу и просматривать сообщения внутри системы.

Доска обсуждения позволяет пользователям оставлять комментарии к публикациям.

Расчет и визуализация статистических показателей, таких как количество публикаций по подразделениям, сотрудникам, типам документов; количество просмотров и т.п.

5) Ограничения и функциональные границы

Для построения ИР должны использоваться программные платформы с открытым исходным кодом и успешным опытом применения в научной и академической среде.

За рамками автоматизации остается оценка значимости результатов исследований, т.к. данная функция практически не актуальна для репозитория вне зависимости от их типа (институциональный или тематический). Интеграция с метрическими сервисами внешних систем является отдельной сложной задачей.

V. ТРЕБОВАНИЯ В ПРОЦЕССЕ РАЗРАБОТКИ ИНСТИТУЦИОНАЛЬНОГО РЕПОЗИТОРИЯ

A. Приоритетность требований

Институциональный репозиторий, как и любая современная информационная система, должен обладать

богатой функциональностью и соответствовать широкому набору характеристик. Как правило, реализация системы требований к подобным IT-разработкам за одну итерацию невозможна. Соответственно, управление процессом разработки зависит от степени значимости (важности) требований. Для оценки приоритетности реализации требований и соответствия основным бизнес-задачам производят ранжирование выявленных требований. Для этого может быть применен широко известный метод MoSCoW [42], [43]. Данный метод классифицирует требования по следующим признакам:

M (Must) — критически важное требование, должно быть реализовано в обязательном порядке.

S (Should) — важное требование, реализация которого должна быть выполнена после реализации требований группы **Must**.

C (Could) — желаемое требование, которое можно реализовать при наличии необходимых ресурсов.

W (Would) — желаемое требование, которое хотелось бы реализовать в будущем.

К группе **M** (требования 1-й очереди) следует отнести функциональные требования, покрывающие базовую функциональность репозитория и обеспечивающие надежность и доступность системы. Большая часть этих требований описана в таблицах 2 и 3:

- Пакетная загрузка цифровых объектов в репозиторий;
- Обеспечение надежного и долговременного хранения метаданных публикаций и их полных текстов;
- Обеспечение процедур резервного копирования;
- Поиск (простой и многокритериальный) и навигация по контенту репозитория;
- Просмотр и фильтрация поисковой выдачи;
- Депонирование (внесение) публикаций в репозиторий через веб-интерфейс;
- Просмотр профиля автора;
- Осуществление системной поддержки;
- Управление правами доступа пользователей;
- Премодерация контента;
- Редактирование контента репозитория (коррекция профилей автора и метаданных публикаций).

В группу **S** (требования 2-й очереди) входит:

- Проверка хранимых цифровых объектов на дублетность;
- Обогащение метаданных публикаций;
- Обеспечение документацией;
- Встроенная система онлайн-помощи посредством контекстных подсказок без перенасыщения и загромождения интерфейса;
- Поддержка версионности хранимого элемента;
- Поддержка многоязычности пользовательского интерфейса;
- Расширение набора характеристик профилей авторов (пользователей).

Группу **C** (требования 3-й очереди) составляют требования, относящиеся к улучшению характеристик надежности и доступности репозитория в реальных условиях, такие как:

- Расширение разнообразия форматов размещаемых файлов;
- Достижение приемлемых значений показателей времени восстановления после сбоев и допустимого времени простоя.

К группе **W** (требования 4-й очереди) относятся возможности и условия, направленные на усовершенствование функционала репозитория и повышение комфортности работы пользователей с ним:

- Управление поисковой выдачей (сохранение результатов поиска, создание личных и групповых книжных полок);
- Создание групп пользователей;
- Рассылка оповещений;
- Обмен сообщениями внутри системы;
- Доска обсуждения;
- Расчет и визуализация статистических показателей: количество публикаций подразделениям, по сотрудникам, по типам документов; количеству просмотров цифрового объекта.

В. Изменение требований

Изменение требований в IT-проектах — хорошо известный факт в программной инженерии. К большому сожалению разработчиков, требования практически никогда не остаются неизменными в течение жизненного цикла информационной системы. Одной из причин тому является пользовательский опыт использования системы, формирующий новое осознание и ожидания от системы, что приводит к запросам на ее модернизацию.

Одна из реализаций прототипа институционального репозитория публикаций ОИЯИ была выполнена на программной платформе Invenio-JOIN. О данном этапе разработки подробнее можно узнать в [28], [44]. Данный продукт обеспечивал следующие ключевые возможности:

- Размещение цифровых материалов и хранение цифровых активов ОИЯИ, и обеспечение долговременного хранения;
- Управление политикой доступа;
- Автоматизация ввода библиографических метаданных путем импорта библиографических метаданных с использованием номеров популярных систем идентификации: DOI, ISBN, идентификаторов arXiv, WoS, Medline, PubMed, INSPIRE;
- Многокритериальный поиск и просмотр цифровых ресурсов репозитория по метаданным и полному тексту;
- Автоматическое отслеживание записей-дубликатов при вводе;
- Верификация введенной информации (редакторами и экспертами);

- Генерация статистические отчетов по разным критериям (по подразделениям; по сотрудникам; по грантам; по периодическим изданиям; по темам и проектам Проблемно-тематического плана; по индексируемым реферативным БД (Web of Science, Scopus, Pubmed); по значениям импакт-фактора журнала (Clarivate); по типам документов и др.);
- Поддержка коммуникации групп пользователей с сервисами создания и совместного использования книжных полок, рассылки оповещений, обмена сообщениями, доски обсуждения.

Для приемочного тестирования прототипа JOIN² была организована группа из потенциальных пользователей системы. Перед участниками группы ставилась задача внести от 5 до 10 публикаций в репозиторий и выполнить поисковые запросы по разным критериям. Все участники группы успешно справились, не испытывая серьезных затруднений. По результатам тестирования была получена и обработана обратная связь. Наиболее часто пользователи указывали на следующие неудобства и/или недостатки:

- Устаревший дизайн web-интерфейса;
- Обязательная привязка публикации к темам/проектам темплана вызывала неудобства.

Если авторский коллектив из нескольких подразделений, нужно точно знать по какой теме работают соавторы;

- Большой список наименований типов публикаций на первом шаге ввода иногда вызывал затруднения с выбором;

- Система выдает предупреждения и рекомендации при сохранении ввода, например, о роли автора (ответственный автор, редактор, иллюстратор, рецензент и т.п.). Эта информация полезна для статистики, но требует дополнительных усилий от пользователя при вводе.

Кроме того, пользователи высказали пожелание свести ручной ввод новых публикаций к минимуму. Большинство работ сотрудников уже проиндексированы наукометрическими базами данных и размещены в тематических репозиториях. Чтобы удовлетворить данную потребность в институциональном репозитории должна присутствовать подсистема сбора данных о публикациях, задачей которой является автоматизированный регламентированный сбор библиографических метаданных публикаций и полных текстов (при наличии) из внешних доверенных источников с помощью открытых и партнерских API с сохранением их в промежуточные базы данных для каждого источника соответственно.

Таблица 4. Новые требования к ИР, выявленные в процессе тестирования прототипа

Элемент	Функции
Подсистема сбора данных о публикациях	Регламентированный сбор библиографических метаданных публикаций и полных текстов (при наличии) из внешних источников с помощью открытых и партнерских API, сохранение результатов в промежуточные базы данных для каждого источника соответственно.
Подсистема курирования метаданных	<p><i>Курирование метаданных публикаций</i></p> <p>Осуществляет предварительную обработку собранных метаданных:</p> <ul style="list-style-type: none"> • Проверка на дубли (выявление и идентификация повторных метаданных); • Обогащение метаданных (дополнение полей); • Запись обработанных данных в единую промежуточную базу данных. <p>Производит коррекцию метаданных по учету изменения типа публикации (например, перевод препринта в статью).</p>
	<p><i>Курирование профилей авторов</i></p> <p>Установление связей авторства (связывания профиля автора с его публикациями)</p> <p>Обогащение профилей авторов метаданными (добавление вариантов имени автора, идентификаторов автора разных систем).</p>

Также необходимо обеспечить автоматическую загрузку собранных метаданных и полных текстов в репозиторий в пакетном режиме и установление связей с профилями авторов. Таким образом, требуется реализация двух основных стратегий наполнения контента: ручной ввод и автоматический сбор с последующей загрузкой. Влияние этого изменения на систему требований состоит в том, что подсистема курирования метаданных должна включать компонент

актуализации метаданных публикаций. Таблица 4 содержит требования к институциональному репозиторию, выявленные в процессе тестирования его опытного прототипа. Данные требования были внесены в реестр требований к ИР и приняты к реализации на следующем этапе разработки.

VI. ЗАКЛЮЧЕНИЕ

Для крупной научной организации актуальность задачи

анализа публикационной активности ее сотрудников не вызывает сомнений. Современным подходом к ее решению является построение, поддержка и развитие репозитория публикаций аффилированных с данным научным учреждением авторов. Это новый вид информационных систем, которые возникли в результате эволюционирования движений открытой науки и служат для накопления, долговременного хранения и обеспечения надежного доступа к цифровым объектам — результатам интеллектуальной деятельности научных организаций. Построение таких систем возможно как на базе собственных разработок, так и готовых решений, в т.ч. на программном обеспечении с открытым исходным кодом, что наиболее соответствует концепции открытой науки. При создании и развитии информационных ресурсов такого типа следует учитывать современные тенденции и процессы в сфере научной коммуникации, построения и развития цифровой среды в научных организациях, а также особенности бизнес-процессов и возможности их интеграции в цифровую экосистему организации-владельца репозитория. Кроме того, такие системы являются открытыми, характеризуются наличием сложных взаимосвязей их структурных элементов с постоянным усложнением структуры и появлением у них новых свойств и функций. Таким образом, они должны развиваться по принципу коэволюции с другими сервисами цифровой экосистемы.

БИБЛИОГРАФИЯ

- [1] Научные журналы: количество, темпы роста / Т. Н. Домнина, О. А. Хачко // Информационное обеспечение науки: новые технологии: Сборник научных трудов. - Москва: БЕН РАН, 2015. - С. 83-96. - [14] с. : [табл., графики]
- [2] Широкаянова А. А. Новая роль и формы научной коммуникации в информационную эпоху // Социология. - 2013. - № 1. - С. 103-116.
- [3] Гонашвили А.С. Наукометрические базы данных и работа с ними: научно-методическое пособие / А.С. Гонашвили. – СПб.: Университет при МПА ЕвразЭС, 2020. – 57 с.
- [4] Гарфутдинова А.Р., Макаровских Т.А. АВТОМАТИЗИРОВАННАЯ СИСТЕМА ДЛЯ ОРГАНИЗАЦИИ И ПРОВЕДЕНИЯ КОНФЕРЕНЦИИ / Современные информационные технологии и ИТ-образование, Том 2 № 11 (2015). - с.85-91.
- [5] Ударцева О.М. Научные процессы в вузах «Приоритета–2030»: системы управления и мониторинга данных. Научные и технические библиотеки. 2022;(9):33-53. <https://doi.org/10.33186/1027-3689-2022-9-33-53>
- [6] Альперин Б. Л., Зибарева И. В., Ведягин А. А. Роль CRIS систем в администрировании научно-исследовательской организации // Управление наукой: теория и практика. 2022. Том 4. № 1. С. 143-156. DOI: <https://doi.org/10.19181/sntp.2022.4.1.8>
- [7] Bryant R. Practices and patterns in research information management: findings from a global survey / R. Bryant, A. Clemens, P. de Castro [et al.] // OCLC Online Computer Library. 2018. DOI: 10.25333/BGFG-D241.
- [8] Паринов С. И. Международная профессиональная ассоциация разработчиков научных информационных систем euroCRIS и её главный продукт CERIF // Электронные библиотеки: перспективные методы и технологии, электронные коллекции: материалы XVI Всероссийской научной конференции R CD L-2014 (Дубна, 13-16 октября 2014 г.). Дубна : Объединенный институт ядерных исследований, 2014. С. 26-29.
- [9] Засурский И.И., Соколова Д.В., Трищенко Н.Д. Репозитории открытого доступа: функции и тенденции развития. Научные и технические библиотеки. 2020;1(9):121-142. <https://doi.org/10.33186/1027-3689-2020-9-121-142>
- [10] Трищенко Н. Д. Открытый доступ к науке: анализ преимуществ и пути перехода к новой модели обмена знаниями. – Москва : Ассоциация интернет-издателей; Екатеринбург : Кабинетный учёный, 2017. – 198 с.
- [11] Рождественская М. Ю. Репозиторий как реализация идей открытого доступа к научным публикациям: подход к классификации // Библиосфера. – 2015. – № 2. – С. 86–94.
- [12] Шрайберг Я. Л., Земсков А. И.. Модели открытого доступа: история, виды, особенности, терминология // Научные и технические библиотеки. – 2008. – №5
- [13] Дудникова О. В., Богомолов А. А. Цифровой репозиторий Южного федерального университета в научном и учебном пространстве вуза //Наука и научная информация. – 2021. – Т. 4. – №. 3. – С. 82-93
- [14] Юдина И. Г., Федотова О. А. Репозитории научных публикаций открытого доступа: история и перспективы развития //Информационное общество. – 2020. – №. 6. – С. 67-79.
- [15] Патук М. И., Наумова В. В., Ерёмченко В. С. Цифровой репозиторий " geologyscience.ru": открытый доступ к научным публикациям по геологии России //Электронные библиотеки. – 2020. – Т. 23. – №. 6. – С. 1324-1338.
- [16] Valerio Cosentino, PictureJavier Luis and PictureJordi Cabot. Findings from GitHub: methods, datasets and limitations / MSR'16: Proceedings of the 13th International Conference on Mining Software Repositories, May 2016, pp.137–141.- URL: <https://doi.org/10.1145/2901739.2901776>
- [17] Laura Dabbish a.o. Social coding in GitHub: transparency and collaboration in an open software repository / Laura Dabbish, Colleen Stuart, Jason TsayJim Herbsleb // CSCW '12: Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work/February 2012, pp. 1277–1286, URL: <https://doi.org/10.1145/2145204.2145396>
- [18] J. C. Cortés Ríos et al., "A Methodology for Using GitLab for Software Engineering Learning Analytics," 2019 IEEE/ACM 12th International Workshop on Cooperative and Human Aspects of Software Engineering (CHASE), Montreal, QC, Canada, 2019, pp. 3-6, doi: 10.1109/CHASE.2019.00009.
- [19] САВИНА А.Г., МАЛЯВКИНА Л.И., САВИН Д.А. ПРОБЛЕМЫ ВЫБОРА И ОБЗОР РЫНКА СИСТЕМ АВТОМАТИЗАЦИИ ПРОЕКТИНГО УПРАВЛЕНИЯ / ВЕСТНИК ОРЕЛГИЭТ, 2020 г., №3 (53).-с.131-139
- [20] Сергеева Н. В. ПРИМЕНЕНИЕ ЦИФРОВЫХ ТЕХНОЛОГИЙ В УПРАВЛЕНИИ ПРОЕКТАМИ В СОВРЕМЕННЫХ УСЛОВИЯХ //Трансформация национальной социально-экономической системы России. – 2023. – С. 175-180.
- [21] Полянин А. В., Жданов А. А. ИСПОЛЬЗОВАНИЕ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ В УПРАВЛЕНИИ ПРЕДПРИЯТИЕМ В СОВРЕМЕННЫХ ЭКОНОМИЧЕСКИХ УСЛОВИЯХ //МЕЖДИСЦИПЛИНАРНОСТЬ НАУЧНЫХ ИССЛЕДОВАНИЙ КАК ФАКТОР. – 2023. – С. 171.
- [22] Чистякова К.А., Юдин В.В. Практические методы управления реализацией инновационных проектов на основе использования программного обеспечения “Jira”. Наука и искусство управления. 2023; (1):80-93. <https://doi.org/10.28995/2782-2222-2023-1-80-93>
- [23] The BTeV program at Fermilab/ BTeV Collaboration, A. Kulyavtsev (Carnegie Mellon U.) et al./AIP Conf.Proc. 444 (1998) 1, 479-502. - DOI: 10.1063/1.56609]
- [24] ALEXANDROV, Evgeny Igorevich et al. Информационные системы для онлайн и офлайн обработки данных в современных экспериментах физики высоких энергий. Современные информационные технологии и ИТ-образование, [S.l.], v. 15, n. 3, p. 654-671, sep. 2019. ISSN 2411-1473. Доступно на: <<http://sitito.cs.msu.ru/index.php/SITITO/article/view/568>>. Дата доступа: 02 apr. 2024 doi: <https://doi.org/10.25559/SITITO.15.201903.654-671>.
- [25] Король И.А. Структура корпоративной информационной системы / Информатизация образования. 2009. № 1 (54). С. 69-82.
- [26] Filozova, I. Concept of JINR Corporate Information System / I.Filozova, V.Korenkov, T.Strizh // XXV International Symposium on Nuclear Electronics & Computing (NEC'2015), Budva, Becici, Montenegro, Sept.28 - Oct.2, 2015 = XXV Международный симпозиум по ядерной электронике и компьютерингу : Book of Abstracts / International Symposium on Nuclear Electronics & Computing (25; 2015; Budva). – Dubna : JINR, 2015. – p.39. – (JINR ; E10,11-2015-65) .

- [27] Суворова С. Д., Куликова О. М. Цифровая трансформация бизнеса // Инновационная экономика: перспективы развития и совершенствования. – 2022. – №. 2 (60). – С. 54-59.
- [28] Filozova I. et al. JINR Open Access Repository based on the JOIN² Platform /Irina Filozova, Tatiana Zaikina, Galina Shestakova, Roman Semenov, Martin Köhler, Alexander Wagner, Laura Baracci on behalf of the JOIN² project //CEUR workshop proceedings, 2020. – Т. 2790. – С. 142-155. - <https://ceur-ws.org/Vol-2790/paper14.pdf>
- [29] Тульчинский Г. Л. Цифровая трансформация образования: вызовы высшей школе //Философские науки. – 2017. – №. 6. – С. 121-136.
- [30] Володина Н. Л. Преимущества создания цифровой экосистемы //Организатор производства. – 2021. – Т. 29. – №. 4. – С. 104-111.
- [31] Иванов А. Л., Шустова И. С. Исследование цифровых экосистем как фундаментального элемента цифровой экономики
- [32] Тамбиева Д. А. Цифровые экосистемы в сельском хозяйстве //Исследование проблем экономики и финансов. – 2021. – №. 1. – С. 6.
- [33] Гаврилов С. Н., Володина С. И. Информационная (цифровая) экосистема адвокатуры в контексте экосистемы цифровой экономики России //Актуальные проблемы российского права. – 2019. – №. 6 (103). – С. 156-166.
- [34] Морозов М. А., Морозова Н. С. Концепция цифровой экосистемы индустрии туризма и гостеприимства //Современные проблемы сервиса и туризма. – 2020. – Т. 14. – №. 4. – С. 27-36.
- [35] Фролов А. А., Соколов А. В., Егоров Д. В. Ключевые характеристики цифровых экосистем в политике //Управленческое консультирование. – 2023. – №. 2 (170). – С. 46-55.
- [36] Каленов О.Е. Цифровые экосистемы организаций. Вестник Российского экономического университета имени Г. В. Плеханова. 2022;(1):139-147. <https://doi.org/10.21686/2413-2829-2022-1-139-147>
- [37] Земсков А. И. и др. Развитие систем открытого доступа к научным данным (Research Data). Принципы FAIR //ГПНТБ, 2022. – С. 65.
- [38] Тютюнов И. А. Использование стандарта CERIF для описания научных исследований // Материалы V Всероссийской научно-практической конференции «Современные проблемы физико-математических наук». – 2019. – С. 247-250.
- [39] Asserson A., Jeffery K. CRIS and institutional repositories // Data Science Journal. – 2010. – Т. 9. – С. CRIS14-CRIS23
- [40] Fritz D. A., Fritz R. J. MARC 21 for everyone: a practical guide. – American Library Association, 2003.
- [41] Das D. Marc 21: The standard exchange format for the 21 st century. – 2004.
- [42] A Guide to the Business Analysis Body of Knowledge® (BABOK® Guide). Version 3. Toronto: International Institute of Business Analysis, 2015.
- [43] Agile extension to the BABOK Guide. Toronto: International Institute of Business Analysis, 2013.
- [44] Filozova I. et al. Curation of bibliographic metadata of the institutional repository on the Invenio-JOIN² platform. – 2023.

On the role and place of the institutional repository in the digital ecosystem of a scientific organization

Irina Filozova, Tatiana Zaikina

Abstract— In any activity, a person is constantly faced with the need to search, process, store and transmit information. In the global social process of informatization of society, information is one of the main resources of its development, and information systems and technologies are used in all areas of human activity. Scientific information systems play an important role in science and education. In a broad sense, such systems are designed to automate scientific, scientific-educational and scientific-organizational activities of research teams. Systems alike are very diverse: there are general-purpose — used in all areas of science, and highly specialized — that have specific application in a particular subject area. This class of information systems includes, in particular, scientometric databases, automated library information systems, information systems for organizing and holding scientific conferences, information systems for collaborative work of researchers, Open Access Repositories (OAR), information systems for supporting current research (Current Research Information System, CRIS-systems), bibliographic managers, author identification systems. This paper examines a special type of scientific information systems — institutional publication repositories — and their place and role in the digital ecosystem of a large scientific organization.

Keywords— digital ecosystem, information systems, institutional repository, scientific communication.

Filozova Irina A., Head of Group, Laboratory of Information Technologies, Joint Institute for Nuclear Research (6 Joliot-Curie St Dubna Moscow Region Russia 141980); Senior Lecturer, Dubna State University, Institute of System Analysis and Management (19 Universitetskaya St Dubna Moscow region Russia 141982), ORCID: <https://orcid.org/0000-0003-3441-7093>, fia@jinr.ru

Zaikina Tatyana N., software engineer 2nd category of the Laboratory of Information Technologies, Joint Institute for Nuclear Research (6 Joliot-Curie St Dubna Moscow Region Russia 141980), ORCID: <https://orcid.org/0000-0003-0805-7995>, ztanya@jinr.ru

REFERENCES

- [1] Scientific journals: quantity, growth rates / T. N. Domnina, O. A. Khachko // Information support of science: new technologies: Collection of scientific papers. - Moscow: BEN RAS, 2015. - P. 83-96. - [14] p.: [tables, graphs]. - in russian
- [2] Litvinova N.N., Razumova I.K. Attitude to Open Access in Russian Scholarly Community 2020: Two Years Later. Scholarly Research and Information. 2020;3(4):226-260. (In Russ.) <https://doi.org/10.24108/2658-3143-2020-3-4-243-277>. - in russian
- [3] Gonashvili A.S. Scientometric databases and working with them: scientific and methodological manual / A.S. Gonashvili. - St. Petersburg: University at the IPA EurAsEC, 2020. - 57 p.- in russian
- [4] Garfutdinova A.R., Makarovskikh T.A. AUTOMATED SYSTEM FOR ORGANIZING AND HOLDING A CONFERENCE / Modern information technologies and IT education, Vol. 2 No. 11 (2015). - pp.85-91. - in russian
- [5] Udartseva O.M. Research processes at “Priority–2030” universities: The systems of control and monitoring. Scientific and Technical Libraries. 2022;(9):33-53. (In Russ.) <https://doi.org/10.33186/1027-3689-2022-9-33-53>. - in russian
- [6] Alperin B. L., Zibareva I. V., Vedyagin A. A. The role of CRIS systems in the administration of a research organization // Science Management: Theory and Practice. 2022. Vol. 4. No. 1. P. 143-156. DOI: <https://doi.org/10.19181/smtp.2022.4.1.8>. - in russian
- [7] Bryant R. Practices and patterns in research information management: findings from a global survey / R. Bryant, A. Clemens, P. de Castro [et al.] // OCLC Online Computer Library. 2018. DOI: 10.25333/BGFG-D241.
- [8] Parinov S. I. International professional association of developers of scientific information systems euroCRIS and its main product CERIF // Electronic libraries: promising methods and technologies, electronic collections: materials of the XVI All-Russian scientific conference RCDL-2014 (Dubna, October 13-16, 2014). Dubna: Joint Institute for Nuclear Research, 2014. P. 26-29. - in russian
- [9] Zassoursky I.I., Sokolova D.V., Trishchenko N.D. Open access repositories: Functions and trends. Scientific and Technical Libraries. 2020;1(9):121-142. (In Russ.) <https://doi.org/10.33186/1027-3689-2020-9-121-142>. - in russian
- [10] Trishchenko N. D. Open access to science: analysis of advantages and ways of transition to a new model of knowledge exchange. - Moscow: Association of Internet Publishers; Yekaterinburg: Cabinet scientist, 2017. - 198 p.- in russian
- [11] Rozhdestvenskaya M. Yu. Repository as an implementation of ideas of open access to scientific publications: an approach to classification // Bibliosphere. - 2015. - No. 2. - P. 86-94.- in russian
- [12] Shraiberg Ya. L, Zemskov A. I. Open access models: history, types, features, terminology // Scientific and technical libraries. - 2008. - No. 5.- in russian
- [13] Dudnikova O. V., Bogomolov A. A. Digital repository of the Southern Federal University in the scientific and educational space of the university // Science and scientific information. - 2021. - Vol. 4. - No. 3. – P. 82-93.- in russian
- [14] Yudina I. G., Fedotova O. A. Open access repositories of scientific publications: history and development prospects // Information Society. – 2020. – No. 6. – P. 67-79.- in russian
- [15] Patuk M. I., Naumova V. V., Eremenko V. S. Digital repository "geologyscience.ru": open access to scientific publications on the geology of Russia // Electronic libraries. – 2020. – Vol. 23. – No. 6. – P. 1324-1338.- in russian
- [16] Valerio Cosentino, PictureJavier Luis and PictureJordi Cabot. Findings from GitHub: methods, datasets and limitations / MSR'16: Proceedings of the 13th International Conference on Mining Software Repositories, May 2016, pp.137–141.- URL: <https://doi.org/10.1145/2901739.2901776>

- [17] Laura Dabbish a.o. Social coding in GitHub: transparency and collaboration in an open software repository / Laura Dabbish, Colleen Stuart, Jason TsayJim Herbsleb // CSCW '12: Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work February 2012, pp. 1277-1286, URL: <https://doi.org/10.1145/2145204.2145396>
- [18] J. C. Cortés Ríos et al., "A Methodology for Using GitLab for Software Engineering Learning Analytics," 2019 IEEE/ACM 12th International Workshop on Cooperative and Human Aspects of Software Engineering (CHASE), Montreal, QC, Canada, 2019, pp. 3-6, doi: 10.1109/CHASE.2019.00009.
- [19] SAVINA A.G., MALYAVKINA L.I., SAVIN D.A. PROBLEMS OF CHOICE AND MARKET REVIEW OF PROJECT MANAGEMENT AUTOMATION SYSTEMS / BULLETIN OF ORELGIET, 2020, No. 3 (53).-pp. 131-139.- in russian
- [20] Sergeeva N.V. APPLICATION OF DIGITAL TECHNOLOGIES IN PROJECT MANAGEMENT IN MODERN CONDITIONS // Transformation of the national socio-economic system of Russia. - 2023. - P. 175-180.- in russian
- [21] Polyanin A. V., Zhdanov A. A. USE OF INFORMATION TECHNOLOGIES IN ENTERPRISE MANAGEMENT IN MODERN ECONOMIC CONDITIONS // INTERDISCIPLINARITY OF SCIENTIFIC RESEARCH AS A FACTOR. - 2023. - P. 171.- in russian
- [22] Chistyakova K.A., Yudin V.V. Practical methods of managing the implementation of innovative projects based on the use of the software "Jira". Science and art of management. 2023; (1):80-93. <https://doi.org/10.28995/2782-2222-2023-1-80-93>.- in russian
- [23] The BTeV program at Fermilab/ BTeV Collaboration, A. Kulyavtsev (Carnegie Mellon U.) et al./AIP Conf.Proc. 444 (1998) 1, 479-502. - DOI: 10.1063/1.56609]
- [24] ALEXANDROV, Evgeny Igorevich et al. Информационные системы для онлайн и офлайн обработки данных в современных экспериментах физики высоких энергий. Современные информационные технологии и ИТ-образование, [S.l.], v. 15, n. 3, p. 654-671, sep. 2019. ISSN 2411-1473. Доступно на: <<http://sitito.cs.msu.ru/index.php/SITITO/article/view/568>>. Дата доступа: 02 apr. 2024 doi: <https://doi.org/10.25559/SITITO.15.201903.654-671>.- in russian
- [25] Korol I.A. Structure of the corporate information system / Informatization of education. 2009. No. 1 (54). P. 69-82.- in russian
- [26] Filozova, I. Concept of JINR Corporate Information System / I.Filozova, V.Korenkov, T.Strizh // XXV International Symposium on Nuclear Electronics & Computing (NEC'2015), Budva, Vecici, Montenegro, Sept.28 - Oct.2, 2015 = XXV Международный симпозиум по ядерной электронике и компьютерингу : Book of Abstracts / International Symposium on Nuclear Electronics & Computing (25; 2015; Budva). - Dubna : JINR, 2015. - p.39. - (JINR ; E10,11-2015-65) .
- [27] Suvorova S. D., Kulikova O. M. Digital transformation of business // Innovative economy: prospects for development and improvement. - 2022. - No. 2 (60). - P. 54-59.- in Russian.
- [28] Filozova I. et al. JINR Open Access Repository based on the JOIN² Platform /Irina Filozova, Tatiana Zaikina, Galina Shestakova, Roman Semenov, Martin Köhler, Alexander Wagner, Laura Baracci on behalf of the JOIN² project //CEUR workshop proceedings, 2020. – T. 2790. – С. 142-155. - <https://ceur-ws.org/Vol-2790/paper14.pdf>
- [29] Tulchinsky G. L. Digital transformation of education: challenges to higher education // Philosophical sciences. - 2017. - No. 6. - P. 121-136.- in russian
- [30] Volodina N. L. Benefits of creating a digital ecosystem // Production organizer. - 2021. - Vol. 29. - No. 4. - P. 104-111.- in russian
- [31] Ivanov A. L., Shustova I. S. Study of digital ecosystems as a fundamental element of the digital economy // Creative economy. - 2020. - Vol. 14. - No. 5. – P. 655-670.- in russian
- [32] Tambieva D. A. Digital ecosystems in agriculture // Study of problems of economics and finance. – 2021. – No. 1. – P. 6.- in russian
- [33] Gavrilov S. N., Volodina S. I. Information (digital) ecosystem of advocacy in the context of the ecosystem of the digital economy of Russia // Actual problems of Russian law. – 2019. – No. 6 (103). – P. 156-166.- in russian
- [34] Morozov M. A., Morozova N. S. Concept of the digital ecosystem of the tourism and hospitality industry // Modern problems of service and tourism. – 2020. – Vol. 14. – No. 4. – P. 27-36.- in russian
- [35] Frolov A. A., Sokolov A. V., Egorov D. V. Key characteristics of digital ecosystems in politics // Management Consulting. - 2023. - No. 2 (170). - P. 46-55.- in russian
- [36] Kalenov O.E. Digital Ecosystems of Organizations. Vestnik of the Plekhanov Russian University of Economics. 2022;(1):139-147. (In Russ.) <https://doi.org/10.21686/2413-2829-2022-1-139-147>.- in russian
- [37] Zemskov A. I. et al. Development of open access systems to scientific data (Research Data). FAIR principles // Responsible for the issue. - 2022. - P. 65.- in russian
- [38] Tyutyunov I. A. Using the CERIF standard to describe scientific research // Proceedings of the V All-Russian scientific and practical conference "Modern problems of physical and mathematical sciences". - 2019. - P. 247-250.
- [39] Asserson A., Jeffery K. CRIS and institutional repositories // Data Science Journal. – 2010. – T. 9. – С. CRIS14-CRIS23
- [40] Fritz D. A., Fritz R. J. MARC 21 for everyone: a practical guide. – American Library Association, 2003.
- [41] Das D. Marc 21: The standard exchange format for the 21 st century. – 2004.
- [42] A Guide to the Business Analysis Body of Knowledge® (BABOK® Guide). Version 3. Toronto: International Institute of Business Analysis, 2015.
- [43] Agile extension to the BABOK Guide. Toronto: International Institute of Business Analysis, 2013.
- [44] Filozova I. et al. Curation of bibliographic metadata of the institutional repository on the Invenio-JOIN2 platform. – 2023.