

Требования к прототипу системы управления информационными ресурсами в распределенных информационных системах поддержки научных исследований

А. М. ФЕДОТОВ*, Ю. В. ЛЕОНОВА

Институт вычислительных технологий СО РАН, Новосибирск, Россия

*Контактный e-mail: fedotov@sbras.ru

Описаны технологические подходы, применяемые для создания распределенных информационных систем поддержки научных исследований. Рассмотрены архитектурные решения, разработанные для создания таких систем и принципы интеграции этих систем с внешними источниками. Определены функциональные требования к модели такой информационной системы, которые обусловлены, во-первых, информационными потребностями исследователей, а во-вторых, необходимостью обеспечения надежного и долговременного хранения информации. Рассмотрены правила представления и преобразования метаданных. Описан и реализован прототип такой системы.

Ключевые слова: информационная система, электронная библиотека, словарь-справочник, распределенные информационные ресурсы, интеграция данных, цифровой репозиторий, поиск информации, метаданные, функциональные требования, библиографические базы данных.

Библиографическая ссылка: Федотов А.М., Леонова Ю.В. Требования к прототипу системы управления информационными ресурсами в распределенных информационных системах поддержки научных исследований // Вычислительные технологии. 2018. Т. 23, № 5. С. 82–109. DOI: 10.25743/ICT.2018.23.5.008.

Введение

В современном обществе быстро возрастают объемы научной информации. Для повышения эффективности научных исследований ученым необходим доступ к информации о результатах исследований, выполняемых в интересующей области. Поэтому любое научное исследование обычно начинается с поиска необходимой информации об исследованиях в данной области, но ее поиск в постоянно возрастающем объеме статей, книг, монографий, отчетов, патентов становится все сложнее. По подсчетам американских специалистов, от 10 до 20 % научных исследований можно было бы не проводить, если правильно подобрать информацию по изучаемой проблеме. Вполне закономерно, что большинство ученых около трети своего рабочего времени тратит на сбор и обработку научной информации. И этот показатель имеет тенденцию к росту. Ученым приходится

много времени уделять работе по унификации данных, систематизации информационных ресурсов, поиску и обработке информации, позволяющим быстро ознакомиться с результатами других исследований и исключить их дублирование.

Основой коммуникации в научном сообществе является научная публикация. «Публикация выступает как первичный источник сведений о научном знании, отношениях между учеными, строении и динамике научных объединений и т. п. Для науковеда, философа, логика, методолога, специалиста по информатике, социолога науки той конечной реальностью, из которой исследователь черпает свои представления о науке, выступают публикации. Отличающиеся друг от друга изображения науки в различных исследовательских традициях ... становятся объектами изучения лишь постольку, поскольку сведения о них имеются в научной публикации» [1]. Главной задачей создания информационной системы поддержки научных исследований является повышение качества информационного обмена в научном сообществе. Однако, чтобы эффективно пользоваться накопленной ранее информацией, необходимы специальные инструменты и технологии, при помощи которых могут быть реализованы специальные приемы работы с информацией [2].

Стремительное развитие глобальных информационных и вычислительных сетей ведет к изменению фундаментальных парадигм обработки данных, которое можно охарактеризовать как переход к поддержке и развитию распределенных информационных ресурсов (ИР) [3, 4]. Поэтому важнейшей задачей, связанной с технологией работы с информацией, является исследование способов интеграции распределенных источников данных.

Под интеграцией информационных ресурсов понимается их объединение с целью использования (с помощью удобных и унифицированных пользовательских интерфейсов) разнородной информации с сохранением ее свойств, особенностей представления и пользовательских возможностей манипулирования с ней. При этом объединение ресурсов не обязательно должно осуществляться физически, оно может быть виртуальным, главное — оно должно обеспечивать пользователю восприятие доступной информации как единого информационного пространства. В частности, такие системы позволяют работать с гетерогенными наборами и базами данных или системами баз данных, обеспечивая пользователю эффективность информационных поисков независимо от особенностей конкретных систем хранения ресурсов, к которым осуществляется доступ [5].

Одна из главных задач интеграции ресурсов состоит в необходимости установить связи между конкретными научными фактами (например, «что означает термин *кибернетика*» или «кто автор данной статьи») и сущностями информационной системы (персонами, фактами, данными, документами, публикациями, ключевыми терминами и т. п.). Поэтому исследование принципов автоматизации научно-информационного процесса и разработка алгоритмов, включая алгоритмы связывания и идентификации данных, является весьма актуальной проблемой.

Научные информационные системы (НИС) — информационные системы поддержки научных исследований оперируют с различного рода информацией. Это могут быть публикации, электронные документы, электронные коллекции, онтологические описания, массивы данных, логические описания и др. Как правило, эти ресурсы разнородны и распределены. Как показано в [6], создание единой централизованной НИС невозможно. Развитие НИС демонстрирует, что во многих случаях научные данные хранятся в различных информационных системах. Тем не менее эти данные интересны для широкого круга научной общественности, и пользователь НИС должен получать ответ на запро-

сы независимо от физического расположения данных. Эти ресурсы, востребованные разными группами исследователей, могут оказаться недоступными из-за проблем с их поиском и идентификацией. Семантические связи между информационными ресурсами повышают их ценность и предоставляют дополнительные возможности для информационного поиска и идентификации. Работы по унификации данных и систематизации информационных ресурсов актуальны для многих предметных областей.

Сложность использования ресурсов обусловлена проблемами обеспечения актуальности версий ресурсов, а также отсутствием единых механизмов преобразования информации при разнородности представления и хранения ресурсов. Интеграция информационных ресурсов в единую информационную среду и организация доступа к ним позволяют унифицировать процесс обмена результатами научных исследований и повышают эффективность взаимодействия отдельных групп исследователей. Актуальной задачей является создание модели информационной системы поддержки научных исследований, позволяющей:

- унифицировать процесс обмена результатами научных исследований;
- оперировать с данными и документами, интегрированными в открытое семантическое пространство;
- предоставлять сервисы преобразования разнородных ресурсов, реализующих средства описания, представления, автоматического связывания ресурсов, а также взаимодействовать с поисковыми и классификационными механизмами в соответствии с потребностями пользователей.

Модель должна обеспечивать такие функции, как:

- публикация ресурсов, включающая процедуры регистрации, именованя, аннотирования и определения формата;
- аналитическая обработка ресурсов;
- доступ к опубликованным ресурсам, включая функции динамического формирования;
- для автоматизированного функционирования мониторинг ресурсов и актуализация их метаописаний, функции уведомления пользователей о появлении новых ресурсов и обновлении существующих, функция диспетчеризации.

1. Анализ текущего состояния исследований в данной области

Работы в направлении создания систем, интегрирующих информационные ресурсы, в ИВТ СО РАН начались в конце 1990-х гг. В результате работ сложилось понимание того, что информационная система для поддержки научных исследований НИС должна основываться на использовании концепции электронных (цифровых) библиотек [7, 8]. В рамках такого подхода цифровые библиотеки рассматриваются как отдельная конкретная технология работы с цифровой информацией, образующая класс информационных систем (ИС), предназначенных для управления информационными ресурсами [9, 10].

Под термином “электронная библиотека” (ЭБ) в данной работе будем понимать систему управления структурированными каталогизированными коллекциями разнородных электронных (цифровых) объектов (ресурсов). Система управления ИР не только обеспечивает поиск, навигацию по рубрикам (по словарям-классификаторам, управляемым словарям), но и непосредственно предоставляет пользователю конкретный ресурс (публикацию, документ, фотографию, описание факта и др.), а также

дополнительные сведения о нем, например географическую привязку, информацию об авторах, фактах, библиографию, перечень организаций и т. д.

Многие технологические разработки информационных систем так или иначе ориентированы на поддержку научных исследований и направлены на удовлетворение информационных потребностей пользователей. В монографии [11], изданной ВИНТИ в 1976 г. и содержащей подробный обзор теоретических проблем информационного обслуживания пользователей, выделены два типа информационных потребностей исследователей — потребности в сведениях об источниках необходимой научной информации и потребности в самой научной информации. Для удовлетворения информационных потребностей первого типа предназначены информационные системы, получившие название документальных, второго типа — фактографических.

В настоящее время наиболее востребованным средством информационного обеспечения научно-образовательной деятельности становятся интеллектуальные информационные системы (ИИС), сочетающие возможности информационных систем обоих названных типов и позволяющие удовлетворять информационные потребности квалифицированного пользователя в соответствии со схемой “документ \mapsto факт \mapsto знание” [2, 12].

1.1. Концептуальные модели

Концептуальная модель ИС должна описывать то, какие сущности могут быть представлены в ИС, а также фиксировать правила и связи, что предполагает классификацию сущностей, абстрагирование, обобщение. Основной целью разработки концептуальной модели научно-образовательной информационной системы является создание модели с конкретной архитектурой для последующей ее реализации в виде программной системы.

В мире разработан целый ряд моделей, описывающих ИР в научной сфере, в области электронных библиотек, культурного и научного наследия. Наибольший интерес представляют *эталонные модели*¹, поскольку они дают абстрактное представление сущностей и отношений, составляющих информационное наполнение системы. Перечислим наиболее популярные из них (обзор таких моделей представлен в работе [13]).

DELOS DLRM (DELOS Digital Library Reference Model). Эталонная модель электронной библиотеки разработана группой специалистов ассоциации в сфере электронных библиотек (DELOS) в 2006–2007 гг. на основе анализа имеющихся библиотечных систем. Большое внимание в модели уделено функциональным возможностям электронной библиотеки. Последнее описание датируется 2011 г. [14, 15].

CIDOC CRM (Committee on Documentation. Conceptual Reference Model) представляет собой формальную онтологию, предназначенную для описания информации в области культурного наследия. Разработка модели началась в 1996 г., а в 2006 г. она стала стандартом (новая версия стандарта ISO 21127:2014²) “Эталонная онтология для обмена информацией культурного наследия” (A reference ontology for the interchange of cultural heritage information) [16].

¹Эталонная модель (reference model) — это абстрактное представление понятий и отношений между ними в некоторой проблемной области. На основе эталонной модели строятся более конкретные и детально описанные модели, в итоге воплощенные в реально существующие объекты и механизмы.

²http://www.iso.org/iso/catalogue_detail?csnumber=57832

OAIS RM (Open Archival Information System) — стандарт ISO-14721:2012³. Эта модель была использована многими организациями для разработки наборов метаданных и организации крупных хранилищ цифровых объектов. На основе данной модели создана концепция “институционального репозитория” как системы долговременного хранения, накопления информации и обеспечения надежного доступа к цифровым объектам, представляющим собой результат интеллектуальной деятельности научного или образовательного учреждения [17].

FRBR (Functional Requirements for Bibliographic Records) — “Функциональные требования к библиографическим записям”. Разработаны международной федерацией библиотечных ассоциаций и учреждений (International Federation of Library Associations and Institutions, IFLA). ER — модель документальной системы, работающей с библиографической информацией [18].

CERIF (Common European Research Information Format) — общеевропейский исследовательский информационный формат [19]. Модель CERIF предназначена для поддержки управления научно-исследовательской информацией, а также настройки и взаимодействия между исследовательскими информационными системами и научными электронными библиотеками.

ЕНИИП (ИСИП) РАН — Единое научное информационное пространство (Интегрированная система информационных ресурсов) представляет модель документально-фактографической ИС [20], реализованной в виде системы управления электронными библиотеками [21]. Выделено четыре основных группы сущностей: участники научной деятельности, научная деятельность, результаты научной деятельности, документы и публикации.

ИРИС СО РАН — Интегрированная распределенная информационная система [4, 7, 22–24]. Представляет модель документально-фактографической ИС, реализованной в виде системы управления электронными библиотеками. Основные сущности: документ, персона, организация, событие, факт.

1.2. Прикладные разработки

Многие технологические разработки ИС в той или иной степени используют эти модели, так или иначе ориентированы на поддержку научно-образовательной деятельности и направлены на удовлетворение информационных потребностей пользователей. Разработки можно разбить на два класса. Большая часть из них представлена библиотечными системами (в том числе электронными каталогами), удовлетворяющими информационные потребности лишь частично. Однако среди них имеются чисто документальные системы, разработанные без учета моделей и стандартов, играющие большую роль в научно-образовательной деятельности. К ним относятся:

eLIBRARY — крупнейшая в России электронная библиотека научных публикаций⁴, обладающая возможностями поиска и получения информации. Система имеет собственную модель данных. Аналогами системы являются: Google Scholar (Google), Scopus (Elsevier), Web of Science (Thompson Reuters).

*Информика (единое окно)*⁵ — информационная система “Единое окно доступа к образовательным ресурсам”, созданная для предоставления свободного доступа и обеспе-

³http://www.iso.org/iso/catalogue_detail.htm?csnumber=57284

⁴<http://elibrary.ru/>

⁵<http://window.edu.ru/>

чения эффективной навигации и поиска учебно-методических и информационно-справочных ресурсов для всех уровней общего и профессионального образования [25]. Имеет свой “стандарт” на представление информации, рубрикаторы и метаописания образовательных ресурсов [26].

Ко второму классу относятся системы, разработанные в рамках концептуальных моделей. В качестве примеров приведем следующие:

ИСТИНА — интеллектуальная система тематического исследования научно-технической информации. Основной целью проекта является сбор, систематизация, хранение, анализ и выдача по запросу информации, характеризующей результаты деятельности научных и образовательных организаций [27]. Система ориентирована на удовлетворение потребностей пользователей (научных сотрудников) в процессе организации научной работы и подготовки отчетных материалов.

MathNET — общероссийский математический портал⁶, созданный Математическим институтом им. В.А. Стеклова РАН, является документальной системой. В ней содержатся информация о российских математиках, база данных публикаций журналов Отделения математики РАН, а также информация об организациях математического профиля [28]. Базовой платформой системы является универсальная информационная система ИСИР (Интегрированная система информационных ресурсов). В качестве основного тематического рубрикатора в системе использованы рубрикатор MSC (Mathematics Subject Classification), созданный Американским математическим обществом (AMS), а также УДК — Универсальная десятичная классификация. Портал содержит следующие разделы: журналы, видеотека, библиотека, персоналии, организации, конференции.

euroCRIS занимается развитием модели современной научной информационной системы⁷, включая создание комплекса необходимых мер (одним из важнейших здесь является формат научных данных CERIF [19]) для взаимодействия и поддержки сотрудничества между участниками, принадлежащими разным организациям независимо от их национальной принадлежности.

СОЦИОНЕТ — система информационной поддержки научно-образовательной деятельности (первоначально в области общественных наук) [29]. Ориентирована на ученых, преподавателей и студентов, собирающих информацию по своей тематике и отслеживающих появление новых публикаций, которые соответствуют их научным интересам. Основная функция системы — накопление метаданных, описывающих научные публикации и другие научные информационные ресурсы (научные отчеты, рабочие записки, тезисы докладов, статистические таблицы, базы данных, сведения о научных организациях и персонах, материалы научных проектов в социально-экономической области и т. д.). Используемая схема метаданных в системе своя, однако поддерживается совместимость с форматом научных данных CERIF.

Научное наследие России — распределенная цифровая библиотека⁸, реализована на платформе ЕНИП [30]. Это документальная ИС с элементами управления информационными ресурсами. Основные элементы функциональности:

- доступ к ресурсам — запрос, определение местоположения, извлечение, трансформация и сохранение ресурса; поиск может осуществляться как по атрибутам ресурса, так и по полным текстам;

⁶<http://www.mathnet.ru/>

⁷<http://www.eurocris.org/>

⁸<http://e-heritage.ru/index.html>

- управление ресурсом — создание нового ресурса, внесение его в цифровую библиотеку, удаление старого ресурса и изменение существующего;
- управление метаданными — их создание, обработка и преобразование; состав метаданных определяется соглашениями;
- управление словарями — их создание, обработка и преобразование; состав словарей определяется соглашениями;
- управление участниками — их регистрация, подписка, права доступа и персональная информация;
- управление цифровой библиотекой — управление коллекциями, группами пользователей, членством, так же как общее управление политикой, качеством или функциональностью;
- системное администрирование — установка, конфигурирование и т. п.

АБИС — Автоматизированная библиотечная информационная система. Появление АБИС в библиотеках сделало переворот в создании ЭБ. Теперь каждая уважающая себя библиотека, имеющая электронный каталог, выставляет его в сети Интернет, называя его при этом ЭБ. Исключение составляет сайт Библиотеки Конгресса США⁹, где действительно присутствует содержательный контент. Мировая цифровая библиотека¹⁰ (World Digital Library) — проект Библиотеки Конгресса, она предоставляет бесплатный доступ в сети Интернет в многоязычном формате к большому количеству материалов по культуре разных стран мира. Самым продвинутым с точки зрения нашего проекта можно назвать проект “Европеана” (Europeana), но, как и все библиотечные проекты, он является чисто документальным.

Все ИС обладают как достоинствами, так и недостатками. Среди них нет системы, которая в полном объеме удовлетворяла бы информационным запросам пользователей. Главными их недостатками являются ограниченность возможностей проведения аналитической работы с ресурсами, обеспечения интеграции ресурсов как внутри каждой из систем, так и с внешними системами (низкая интероперабельность), плохая идентификация ресурсов и слабо развитые сервисы таксономии. Это крайне неудобно для научных работников, которым в процессе исследований необходимо установить связи между конкретными научными понятиями и сущностями информационной системы (персонами, фактами, документами, публикациями и т. п.). Наиболее близко к нашей задаче находятся информационная система СОЦИОНЕТ (но в ней нет поддержки аналитической работы с данными) и разрабатываемый пакет “Либмета” [31].

1.3. Сервисы таксономии

Существует большое количество разнообразных средств для решения узкоспециализированных задач, имеются обобщенные методологии для создания открытых и распределенных систем. Однако они не позволяют полностью решить сформулированную задачу. Поэтому необходимы интеграция этих средств в единое целое и разработка механизмов, которые могли бы адекватно связать между собой данные и модели.

Исследование возможностей разных поисковых систем и разработка новых методик поиска информации в больших массивах данных является актуальной областью научных исследований во многих странах. Например, ежегодно проводимая Национальным институтом стандартов и технологии (National Institute of Standards and Technology)

⁹<https://www.loc.gov/>

¹⁰<https://www.wdl.org/ru/>

США международная конференция TREC (Text REtrieval Conference) посвящена разработке критериев эффективности информационного поиска в разных системах, способам достижения наилучшего соотношения полноты и точности информационного поиска [<http://trec.nist.gov/>; <http://tac.nist.gov/>].

Несмотря на появление новых способов создания, обработки, хранения и поиска информации, проблема доступа к информационным ресурсам и получения релевантных документов для потребителя обострилась. Это связано с неограниченным ростом информационных массивов и распространением полнотекстовых баз данных. Качественно составленный тезаурус является источником семантической информации, применяемой в различных технических приложениях, в особенности в области информационного поиска. Пользователь, осуществляющий поиск информации, может по тезаурусу найти интересующие его понятия в данной предметной области и все его термины, в том числе и являющиеся синонимами. Использование связей тезауруса позволяет расширить границы поискового запроса, улучшить качество информационного поиска в выбранной предметной области.

Основная проблема состоит в высокой трудоемкости и стоимости ручного составления тезауруса, малой гибкости процесса его построения [32]. В тезаурусах для ручного индексирования совокупности близких понятий сводятся к одному, наиболее представительному понятию для уменьшения субъективности индексирования. В автоматизированных тезаурусах семантически близкие понятия представлены в виде отдельных единиц, что позволяет использовать при поиске синонимические ряды. Трудность построения тезауруса, соответствующего всему тематическому многообразию индексируемой информации, является основной причиной его непопулярности в современных информационных системах. Но если рассматривать вопрос об эффективности информационных систем в определенных областях знаний, то создание и использование специализированных тезаурусов представляет несомненный интерес и переводит систему в совершенно другой качественный класс.

2. Предметная область НИС

Научная информационная система автоматизирует процессы научно-информационного обмена. В связи с этим можно сформулировать основные цели, стоящие перед НИС как системами управления информационными ресурсами [33]:

- управление ИР;
- обеспечение и управление доступом к информации;
- долговременное хранение информации;
- сохранение научного и культурного наследия;
- поддержка аналитической работы с информацией;
- повышение эффективности научных исследований и обучения.

Часть реального мира, которая моделируется НИС, называется ее предметной областью. Поскольку модель предметной области, поддерживаемая НИС, материализуется в форме организованных необходимым образом информационных объектов, она называется информационной моделью (рис. 1). Информационные объекты характеризуются метаданными, описывающими реальный объект, и могут быть снабжены аннотациями. Информационные объекты могут иметь информационное содержание (контент).

Ввиду того что информация в НИС отображает некоторые сущности реального мира (физические объекты: предметы, процессы, явления, персоны, публикации, документы,

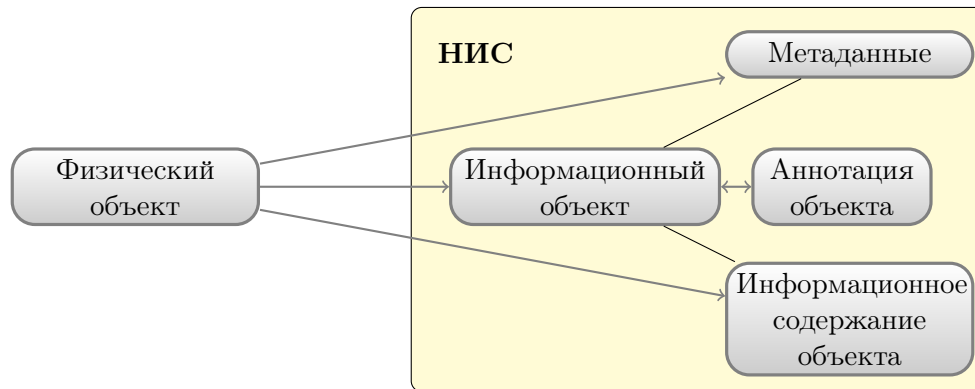


Рис. 1. Структура сущностей НИС

алгоритмы, программы, файлы, факты, ключевые термины и т. д.), следует рассматривать НИС как множество информационных объектов — наборов данных, представляющих (описывающих) эти сущности. Отметим, что разработка модели НИС должна использовать онтологические описания и концептуальные модели, обобщающие накопленный опыт в сфере создания и использования ЭБ [12].

Информационный ресурс — это понятие, включающее любую сущность НИС. В информационном пространстве все сущности (документы, публикации, персоны, события, факты, программы и любые другие сущности реального или виртуального мира) представлены только в форме некоторых информационных ресурсов (объектов). Информационный ресурс — это абстрактное понятие, выражаемое экземплярами одной из своих специализаций. В частности, экземплярами понятия информационного ресурса являются экземпляры информационного объекта любого типа (например, документы, базы данных, коллекции, функции и т. п.).

Наиболее значимы для НИС информационные ресурсы, выражающие научное произведение [18]. Научное произведение — это определенная система понятий. Научное произведение может быть выражено в форме публикации: учебника, монографии, статьи и т. д. Существуют и другие формы воплощения научных произведений: чертежи, планы, эскизы, модели, компьютерные программы, различного рода карты и т. п. Научное произведение отличается от художественного тем, что оно закрепляет результаты познания закономерностей развития внешнего мира (природы и общества) и не имеет персонажей, повествователей, рассказчиков [34].

Научное произведение формально является основным результатом работы ученого. Публикуя материал, ученый знакомит научную общественность с результатами своих исследований, их анализом и выводами. Но помимо донесения информации о проведенной работе у публикации в НИС есть еще одна полезная функция — обеспечение поиска публикаций по аналогии, что существенно облегчает отбор научной информации об исследованиях в данной области.

2.1. Информационный ресурс

Реализация информационного ресурса (информационного объекта) — это единица информации, представляющая собой уникально именованный набор данных, структурированных в виде присущих ему именованных атрибутов и методов, характеризующих

его свойства и связи (отношения) с другими ресурсами. Онтологическая модель НИС основана на концептуальных моделях FRBR [18], RM OAIIS [17] и DELOS DLRM [15]. В соответствии с концептуальной моделью каждый ИР должен:

- иметь идентификатор;
- быть организованным в соответствии с описанием ресурса (ресурсы могут быть сложными и структурированными, а с организационной точки зрения они могут группироваться в наборы ресурсов, которые рассматриваются как единая сущность);
- может регулироваться функциями, управляющими его жизненным циклом, имеет набор присущих ему атрибутов и методов, характеризующих его свойства и связи с другими ресурсами;
- выражается через информационный объект;
- должен быть описан метаданными, а также может быть расширен дополнительными метаданными и аннотациями.

В информационной системе каждому ресурсу соответствует информационный объект, который является традиционным вторичным информационным объектом, содержа-

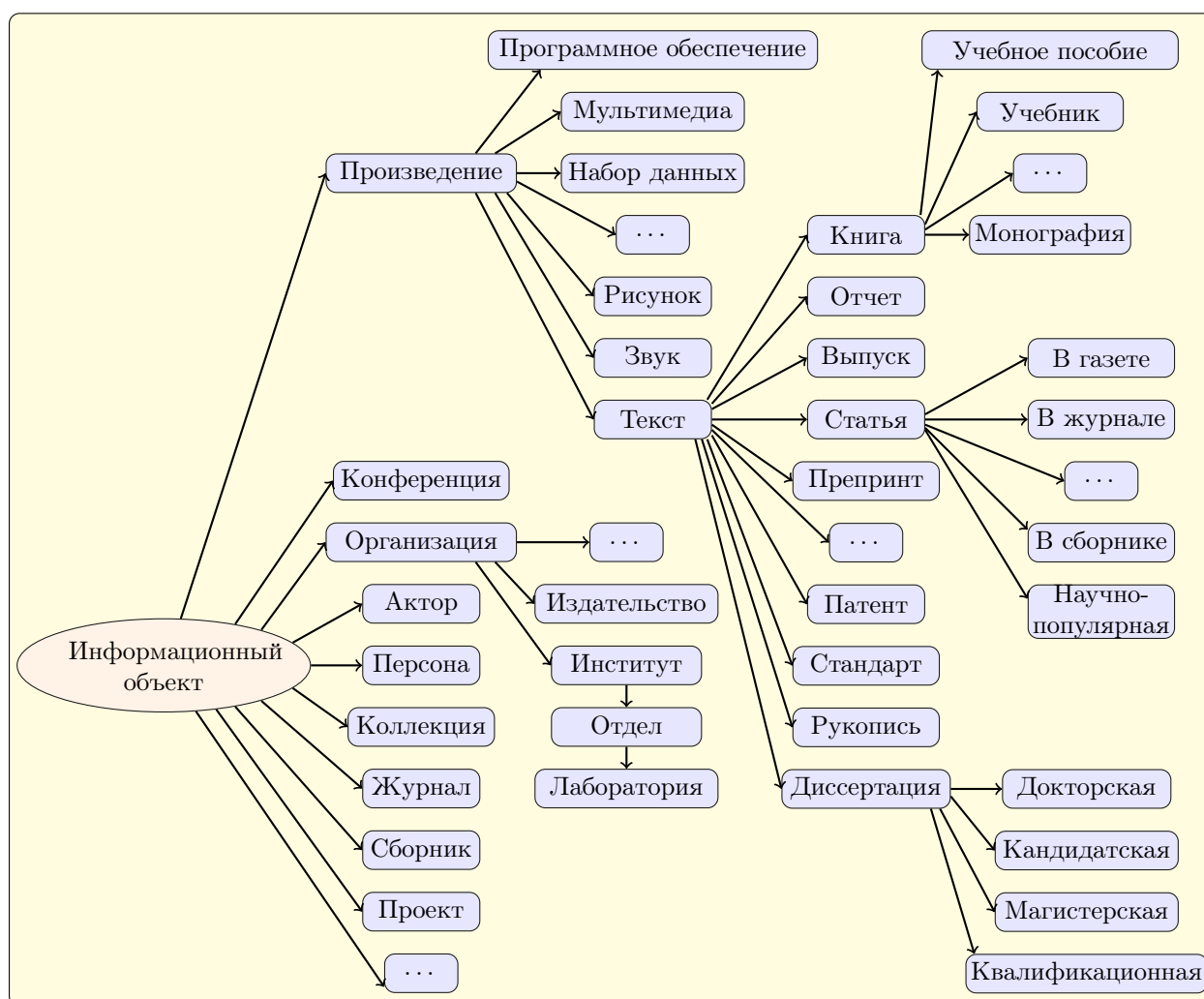


Рис. 2. Иерархия классов НИС (фрагмент)

щим описание первичного ресурса, т. е. информационный объект — это объект, который хранит информацию об объектах НИС (физических объектах, ресурсах, информационных объектах).

На рис. 2 приведена часть иерархии классов информационных объектов, представленных в НИС. Основу содержания НИС составляют информационные объекты, представляющие основные типы сущностей:

- субъекты (актор¹¹, персона, организация, действующие лица, приложение и т. п.);
- объекты (публикация, журнал, документ, факт, научный результат, мероприятие, проект, фотография и др.);
- отношения (понятие, ключевой термин, событие, время, место и т. п.).

Информационный объект — наиболее общее понятие в системе, он представляет произвольную единицу информации в НИС. Информационные объекты описывают все классы сущностей научного информационного пространства, такие как публикация, персона, ключевой термин или понятие, словарная статья, факт, функция, организация, пользователь и т. д., а также связи между ними [24].

Информационный объект является корневым объектом в представляемой модели, он охватывает все объекты, информация о которых хранится в НИС. Каждый информационный объект в НИС состоит из таких объектов, как (см. рис. 1):

- информационное содержание объекта (первичный информационный объект включает, например, изображение, полный текст и т. д.), который может использоваться самостоятельно или может отсутствовать;
- метаданные — объект, главная цель которого состоит в том, чтобы дать информацию об ИР (как правило, о первичном информационном объекте);
- аннотация — объект, главная цель которого состоит в том, чтобы аннотировать ИР или его часть. Примеры таких аннотаций включают примечания, структурированные комментарии и связи. Аннотации помогают интерпретировать ИР, содержат либо поддержку, либо детальные объяснения, либо информацию о том, как можно использовать ИР.

2.2. Профили метаданных

Эффективное средство описания информационных объектов — метаданные, которые являются неотъемлемой частью информационного объекта и описывают реальный объект или группу объектов.

Важное свойство метаданных — специфичность относительно области применения описываемых объектов (ресурсов). Метаданные могут характеризовать сущности, которые относятся как к виртуальному (информационному) пространству, так и к реальному миру (персоны, организации, события). Метаданные могут быть частью ИР, а могут храниться отдельно от ИР.

Метаданные необходимы для решения следующих задач [2, 22, 35]:

- предоставление сведений о документах (их содержании, структуре, способах использования и т. д.);
- систематизация информации о документах;
- выбор из множества документов определенного подмножества по формальным признакам и сопоставление документов по формальным признакам;

¹¹Актор — лицо, составляющее акты

- внутрисистемные технологические задачи, связанные с обеспечением подготовки документов, размещением документов в информационной среде и т. д.;
- внешние технологические задачи, связанные с обменом данными с внешними информационными системами.

Очевидно, что для успешного обмена и интерпретации описаний необходимо стандартизировать формат элементов метаописаний, синтаксис и семантику его элементов. Основная функция стандартов состоит в упрощении обмена и автоматизированной обработки информации из различных источников. Для автоматического размещения в хранилище информации о ресурсе без привлечения дополнительного персонала описание ресурса должно быть оформлено в соответствии со стандартом, а с источниками, предоставляющими ресурсы, должно быть заключено соглашение о взаимной поддержке стандарта и качества описаний. Описания, оформленные не по стандарту или некачественные (например, недостоверные или неполные), затруднительно использовать для автоматического размещения без дополнительной обработки.

Основные требования к системе метаданных [35, 36]:

- универсальность в рамках установленного понимания ИР как объекта систематизации;
- структурированность и формализованность метаданных, необходимые для их автоматической обработки;
- достаточная выразительность для обеспечения решения задач, требующих наличия метаданных;
- совместимость с международными стандартами и протоколами в области метаданных и информационного поиска (создание условий для интероперабельности);
- возможность задания ограничений целостности, отражающих взаимосвязи полей описания ИР;
- обеспечение возможности хранения метаданных как совместно с ИР, так и отдельно от него;
- возможность представления в метаданных сведений о создателях, правообладателях, распространителях ИР и отношениях между ИР.

Для определенной группы функциональных задач или пользователей создают профили применения стандартов. Это позволяет облегчить создание систем, которые работают с метаданными. Профиль можно определить как “один или сочетание нескольких базовых стандартов с идентификацией выбранных классов, подмножеств, факультативных возможностей и параметров этих базовых стандартов, необходимых для выполнения конкретной функции” [37, 38]. В области метаданных ресурсов для публикаций профиль должен содержать список обязательных элементов, присутствующих в описании ресурса, задавать словари для описания значений элементов, которые дополняют или расширяют определенное в стандарте допустимое множество значений. Кроме того, могут быть предложены дополнительные элементы описания. Таким образом, основу разработки НИС составляют стандарты и международные рекомендации, формирующие профиль НИС, под которым понимается один или набор нескольких базовых нормативно-технических документов (стандартов и спецификаций), ориентированных на решение определенной задачи (реализацию заданной функции либо группы функций приложения или среды), с указанием, если нужно, выбранных классов, подмножеств, опций базовых стандартов, необходимых для выполнения конкретной функции [38]. Наиболее важен профиль метаданных информации, циркулирующей в системе. Профиль метаданных должен соответствовать следующим требованиям [4, 12, 20, 33, 39]:

- включать описания основных типов информации, необходимой для поддержки научно-образовательной деятельности;
- быть открытым, т. е. обеспечивать доступ к информации в соответствии с ее описанием (метаданными);
- быть расширяемым, т. е. обеспечивать возможность детализации описаний;
- обеспечивать возможность интеграции информации и уникальной идентификации информации;
- обеспечивать отбор, систематизацию и классификацию информации;
- обеспечивать возможности размещения и поиска информации в распределенной среде и интероперабельности с другими системами;
- быть ориентированным на современные технологии описания и использования информации.

Отметим, что серьезной проблемой является идентификация ИР [40, 41], позволяющая получать библиографические сведения, а также устанавливать связи определенного ресурса с другими фактами и объектами. Проблема связывания данных выходит за рамки этой статьи.

В основе взаимодействия распределенных информационных систем лежит интеграция метаданных, которая обеспечивается наличие стандартов для форматов их представления, и унификация нормативно-справочной информации (профиля ИС). Под интеграцией данных, с точки зрения пользователя, следует понимать:

- возможность свободно группировать любые имеющиеся разнородные ресурсы или их части по любому признаку в произвольные реальные и/или виртуальные коллекции;
- возможность организовывать по всем массивам данных прозрачный для конечного потребителя сквозной поиск информации.

Реализация механизмов интеграции данных немислима без их стандартизации — данные одного типа должны описываться и предоставляться единым образом в соответствии с нормативными документами. В частности, в стандартизованном виде должны предоставляться следующие типы информационных ресурсов:

- географические информационные ресурсы (картографические материалы, спутниковые снимки, данные полевых наблюдений и т. п.), а также соответствующие базы метаданных;
- фактографические базы данных и метаданных;
- библиографические базы данных и электронные каталоги;
- полнотекстовые базы данных и цифровые репозитории;
- авторитетные базы данных (описывающие субъекты информационного взаимодействия: персоны, организации и т. п.);
- другие ресурсы (аудио- и видеозаписи, электронные презентации и др.), снабженные стандартизованными метаданными.

По назначению выделяют три основных вида метаданных:

- описательные (библиографические описания ИР) — метаданные в стандартном понимании, описывающие контент ресурса в соответствии с выбранной схемой данных (например, Dublin Core [42] или МЕКОФ [43]), библиографические данные (если ресурс — публикация), аннотацию, идентификаторы ресурса (например, URI или DOI [44]) и т. п.;
- структурные (формат, объем и структура ИР) — характеризуют общую структуру ресурса и ее компоненты, структурные правила определения ассоциативных

связей между ресурсами, объем, используемый формат и другие свойства описываемого ресурса;

- административные или служебные (правообладатели, права на доступ и коррекцию ИР и т. д.) — метаданные, несущие исключительно служебную информацию. К ним относятся, например, дата создания или модификации ресурса, владелец прав на ресурс (не путать с автором), информация о существующих версиях, копиях ресурса и другие сведения, необходимые для решения технологических задач системы управления ресурсами, например правила представления документов пользователю, правила преобразования схем данных структурных метаданных.

2.3. Протокол Z39.50

Протокол Z39.50 (ISO-23950) [45] — удобная основа для создания профилированных НИС. Главной отличительной особенностью Z39.50 является стандартизация метаданных, схем данных, без чего невозможно построение НИС из разнородных источников информации. Z39.50 — это мировой стандарт, которого придерживаются многие информационные организации и объединения во всем мире.

Таким образом, для решения задач интероперабельности данных необходимо создать схемы метаданных, описывающих типы ИР, их свойства. Каждая предметная область оперирует собственными терминами, отношениями между информационными ресурсами, вследствие чего требуется создавать схему метаданных для каждой предметной области. В то же время ряд схем и определений метаданных могут быть использованы, возможно, с переопределениями во многих предметных областях, поэтому требуется разработать методы использования существующих схем метаданных для новых предметных областей.

Схемы метаданных для конкретных областей разрабатываются на основе имеющихся схем метаданных в области предметных профилей — схем метаданных, собранных из элементов различных схем (или пространств имен) и оптимизированных для использования в конкретной предметной области [47]. Под предметными профилями, оптимизированными для предметной области, понимают следующее. Значения терминов словарей метаданных (элементов схемы или пространств имен) предметной области и используемых в профиле схем могут различаться. В предметной области могут использоваться отличные от принятых в исходной схеме словари для классификации значений свойств данных. Разработка предметного профиля для предметной области может включать переопределение терминов словарей метаданных, формальные или неформальные инструкции по использованию метаданных в предметной области, пополнение исходных схем новыми словарями значений свойств и их изменение.

Для совместного использования данных при интеграции различных схем необходимо согласовать схемы или форматы метаданных — определить семантические отношения между терминами словарей метаданных и словарей классификации значений данных, описать конфликты и правила их разрешения.

3. Требования, предъявляемые к НИС

К большинству ИС предъявляются повышенные требования: с точки зрения пользователей — удобство в использовании и простота в изучении; с технической точки зрения —

взаимодействие с другими ИС и обеспечение для взаимодействия стандартизованных сервисов и протоколов.

Рассмотрим требования к НИС в контексте “Интегрированной распределенной информационной системы” (ИРИС) СО РАН. В работах [12, 33], исходя из пользовательских предпочтений, определен профиль НИС как набор базовых нормативно-технических документов (стандартов и спецификаций), реализующих пользовательские и функциональные требования, требования к модели и управляемым словарям данных, производительности и удобству использования, а также требования к безопасности.

В [6, 12, 33, 46] сформулирован ряд требований для систем, работающих с научными документами. Суммируя сказанное, можно определить набор наиболее общих функциональных требований к НИС.

1. Сбор информационных ресурсов. Для сбора информации необходимо использовать различные варианты ввода данных: ввод данных пользователями, сбор данных в сети Интернет посредством специальных программных агентов (“пауков”), обмен данными с другими ИС. Таким образом, НИС должна поддерживать работу с внешними источниками (например, каталогами библиотек и журналов, цифровыми репозиториями информационных ресурсов и т. п.).

2. Релевантность документов. При автоматическом сборе информации в сети Интернет может накапливаться и нерелевантная или малорелевантная информация для данной НИС. Решение проблемы возможно следующими способами. Создание подробных форматов представления метаданных о ресурсах и структурированных справочников для тематической классификации ресурсов. НИС должна вкладывать описания в метаданные на веб-страницы и предоставлять интерактивные средства пользователям для создания метаданных определенного формата при размещении ресурсов. Разделение информационных ресурсов в зависимости от варианта поступления в систему (размещенные экспертами/пользователями и “пауком”), а также указание степени достоверности информации с учетом ее источника. Указание поисковым средствам пространства поиска и классификации информации, а также задание критериев оценки качества введенной информации. Использование схем классификации ресурсов согласно потребностям пользователей и классификация ресурсов в соответствии с этими схемами.

3. Актуальность, полнота, достоверность происхождения документов. Способы решения проблем актуальности и полноты аналогичны способам решения проблемы охвата ресурсов. Способы определения достоверности происхождения информации следующие: для интерактивного ввода — информация вводится только аутентифицированными пользователями; для автоматизированных систем сбора — накладывание ограничений на область действия агента, выполняющего сбор информации; для обмена информацией с другими ИС — задание фильтров на импортируемые информационные ресурсы; для всех способов ввода — должны выполняться проверка и классификация всей введенной информации.

4. Использование интеллектуальных служб обработки запросов пользователя. Службы обработки запросов пользователей должны обеспечивать поиск по атрибутам, полнотекстовый поиск, просмотр ресурсов по категориям, семантический поиск (необязателен).

5. Извлечение знаний. (Использование частичной автоматизации извлечения знаний.) В основе подхода лежит представление смысла текста в виде семантической сети, принцип построения которой основывается на использовании частоты совместной встречаемости понятий в тексте. Пользователю сеть представляется в виде тематиче-

ского дерева (дерева ключевых терминов и связанных с ними понятий), что позволяет выполнять навигацию и существенно облегчает процесс исследования текста и поиска требуемой информации. Данный подход используется также для решения таких задач, как автоматическое реферирование, тематическая классификация и кластеризация текстов, семантический поиск и т. д.

Кроме того, к НИС, работающим с разными типами информационных ресурсов, предъявляются следующие требования.

6. Поддержка нецентрализованных архитектур информационных систем. Это необходимое условие для полноты, аутентичности и актуальности информации. Опыт эксплуатации НИС показал сложность создания централизованных научных систем, охватывающих научную информацию в какой-то области науки или в какой-то стране.

7. Структурированность информационного пространства. Структурированность является важнейшим свойством информационного пространства. Это означает, что выделены его элементы, установлены связи между ними, элементы и связи упорядочены. Свойство структурированности в разных видах информационных пространств может быть выражено в разной степени. Высокий уровень структурированности обеспечивает возможность представления информации в виде документов и манипулирования данными с помощью программно-технических средств информационных систем. Таким образом, НИС должна обеспечивать поддержку взаимосвязей и идентификации информационных ресурсов. Слабоструктурированные информационные объекты содержат в себе структурированную и неструктурированную части. Структурированная часть информационного объекта может быть представлена в виде информационной модели, неструктурированная — в виде совокупности терминов некоторой онтологии, характерной для предметной области. Для представления неструктурированных информационных объектов с помощью онтологии необходимо определить термины предметной области. Термины можно выделить путем индексирования всех информационных ресурсов.

8. Использование классификации информации при информационном поиске. Для поддержки сложных функций поиска и классификации информации помимо хранения полнотекстового описания необходимо реализовывать поиск по атрибутам, полнотекстовый поиск, просмотр ресурсов по категориям и словарям-классификаторам. Выбор классификаторов определяется степенью специализации системы. Для реализации этих функций, а также обеспечения идентификации и классификации ресурсов НИС должна содержать словари-классификаторы.

9. Адаптивное представление информации. Для повышения скорости поиска и точности подбора информации без потери качества поиска НИС должна учитывать запросы пользователей, их компетентность при работе с НИС, ограничения по времени. НИС должна обеспечивать возможность пользователю получать различные уровни абстракции при представлении информации от кратких описаний для максимального быстрого поиска до очень подробных описаний информационных объектов.

10. Историчность информации. Спецификой научной информации является ее быстрое устаревание и потеря актуальности. Для многих типов информационных ресурсов важно хранить всю информацию о всех изменениях и иметь возможность восстановить состояние ресурса на любой момент времени. Например, информация об авторах может меняться со временем при смене персоной фамилии, места работы. Также необходимо учитывать переформирование и переименование организаций, наименования гео-

графических объектов также могут меняться. Поэтому следует учитывать временной фактор и использовать актуальную для сущностей информацию, связанную с промежутками времени. При распознавании сущностей необходимо обеспечивать выполнение запросов на какой-либо момент времени в прошлом, т. е. создание среза истинности информации о сущностях на произвольную дату.

11. Архив. Как было отмечено выше, большая часть научной информации быстро устаревает. Но существуют информационные ресурсы, к которым необходимо обеспечивать доступ длительное время. К таковым, например, относятся документы, имеющие длительную юридическую силу, патенты или мультимедийная информация об исторических событиях, которая может быть востребована через любой период времени. Кроме того, научные отчеты институтов, речи ученых могут также иметь огромную историческую ценность, становясь только еще ценнее со временем. Поэтому система должна поддерживать возможность длительного хранения информационных ресурсов и их восстановления.

12. Поддержка распределенности. В современном мире наблюдается бурный рост в направлении разработки распределенных ИС, что обусловлено [23]:

- увеличением пропускной способности каналов связи и, как следствие, скоростью обмена по ним, которая приближается к скоростям внутренних шин компьютеров;
- ростом производительности компьютеров как по скорости, так и по объемам памяти, и оперативной, и внешней;
- широким проникновением компьютеров и компьютерных технологий в повседневную деятельность как большинства организаций и учреждений, так и граждан;
- развитием Интернета, обеспечивающего простой и надежный доступ к невероятному числу информационных ресурсов;
- развитием самих информационных технологий и “сетевых” программирования.

Тенденция к увеличению доли распределенных систем существует и в СО РАН, которое объединяет большое число научно-исследовательских организаций и коллективов в различных городах России и имеет разветвленную структуру. Эти организации обладают разнообразными научными информационными ресурсами, которые представляют значительный интерес для мирового научного сообщества, работников промышленных предприятий и бизнеса, заинтересованных во внедрении результатов научных исследований. Таким образом, при создании современной НИС следует уделять особое внимание разработке средств взаимодействия между НИС этих организаций и средств поддержки распределенного хранения ресурсов НИС в различных хранилищах. При этом необходима поддержка средств поиска и каталогизации, работающих над всеми разнородными ресурсами распределенной НИС. В условиях работы в распределенной среде к НИС предъявляются требования обеспечить поддержку:

- принятых стандартов метаданных для экспорта и импорта данных;
- протоколов обмена информацией с другими информационными системами;
- создания ссылок на внутренние ресурсы как в интерфейсах пользователей, так и на системном уровне.

4. Информационная модель НИС

Как уже отмечалось, НИС — это каталогизированные распределенные ИС, позволяющие хранить, обрабатывать, распространять, анализировать, а также организовывать

поиск в разнообразных коллекциях электронных (цифровых) документов (электронных библиотек). Основная задача, решаемая НИС, — это управление и интеграция ИР, включая поддержку унифицированного доступа к ним, а также эффективная навигация в них [7].

Под интеграцией ИР понимается их объединение с целью использования (с помощью удобных и унифицированных пользовательских интерфейсов) разнородной информации с сохранением ее свойств, особенностей представления и пользовательских возможностей манипулирования с ней. При этом объединение ресурсов не обязательно должно осуществляться физически, оно может быть виртуальным, главное — оно должно обеспечивать пользователю восприятие доступной информации как единого информационного пространства. В частности, такие системы позволяют работать с гетерогенными наборами и базами данных или системами баз данных, обеспечивая пользователю эффективность информационных поисков независимо от особенностей конкретных систем хранения ресурсов, к которым осуществляется доступ.

Под эффективной навигацией в информационной системе понимается возможность для пользователя находить интересующую его информацию с наибольшей полнотой и точностью при наименьших затратах усилий во всем доступном информационном пространстве. При таком подходе хорошо известные информационно-поисковые системы, используемые в ИС и базах данных, являются частными случаями навигационных средств [4, 7].

Наиболее важным выводом из вышесказанного является то, что информационная модель НИС должна быть многоуровневой и состоять как минимум из таких компонентов, как хранилище данных (репозиторий), сервер метаданных, сервер приложений (диспетчер), словари-справочники (рис. 3) [4, 7, 12]:.

Репозиторий — это независимая система долговременного хранения и доступа к разнородным цифровым объектам. Цифровой репозиторий является одним из важнейших компонентов распределенной системы и предназначен только для обеспечения “функции” долговременного хранения информационных ресурсов. Может изменить-

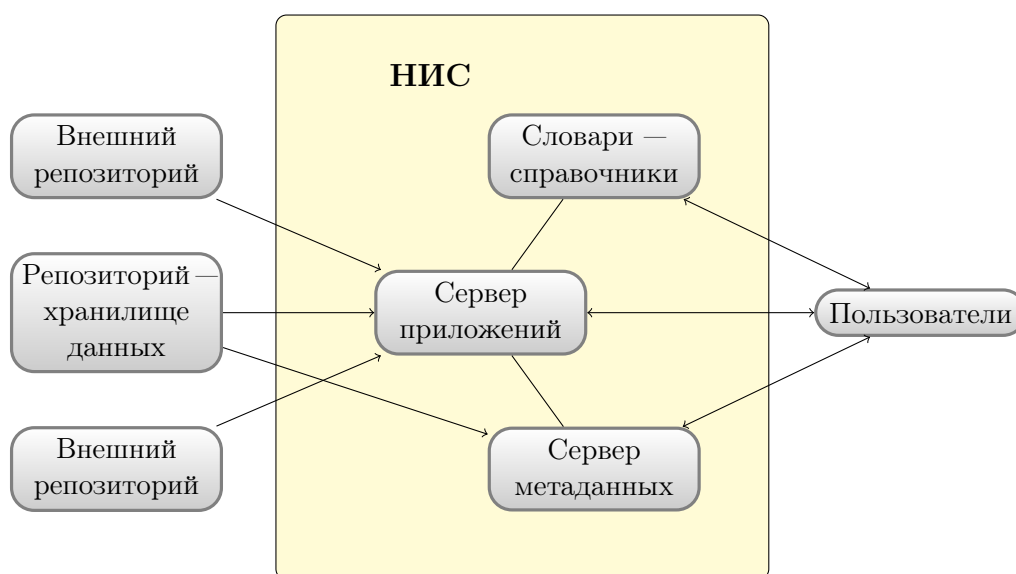


Рис. 3. Информационная модель НИС

ся система, могут поменяться интерфейсы и сервисы, но ИР, несущий информацию, не изменяется, поэтому должен храниться вечно и независимо. Таким образом, функция хранения данных отделена и не зависит от других функций и сервисов системы.

Сервер метаданных должен обеспечивать работу с метаданными — каталогизацию всех информационных ресурсов в соответствии с общепринятыми международными стандартами.

Сервер приложений должен обеспечивать работу основных сервисов НИС, которые необходимы для формирования информационных ресурсов с использованием и без использования диалоговых пользовательских интерфейсов. Сервисы позволяют использовать метаданные других информационных систем в диалоговом и пакетных режимах. Их функциональность должна обеспечивать поиск и извлечение метаданных из других систем, конвертирование полученных метаданных в схемы и структуры локальной системы.

Справочники — управляемые словари (ключевые признаки, ключевые термины) — это особый вид метаданных, отражающих наиболее существенные свойства информационного объекта и имеющих наиболее важное значение с точки зрения НИС. Специфика словарей определяется терминологией конкретной предметной области, которой посвящена НИС. Необходимо рассматривать различные типы ключевых терминов (ключевые термины в стандартном понимании; ключевые термины, описывающие персону, организации, временные периоды, географические понятия). Это набор баз данных (нормативных словарей), содержащих информацию об авторах и других персонах (авторитетные записи), географических пунктах, городах, издательствах, имеющих отношение к конкретной теме или разделу НИС (например, к научной школе), тематические словари-классификаторы, тезаурусы, рубрикаторы, описания предметной области и классификаторы документов.

Используемый профиль определяет список элементов данных (полей), необходимых для создания записи соответствующего типа, и раскрывает содержание элементов данных. Для эффективной работы сервера приложений используется набор словарей-классификаторов, содержащих как классификационные признаки, так и наборы ключевых терминов (с отношениями порядка), по которым производятся систематизация и классификация материала.

Особым типом объекта является объект *коллекция*. Коллекция может состоять из любой совокупности (группировки, агрегации) любого типа информационных ресурсов, т. е. коллекциями могут быть совокупности как субъектов, так и объектов. Критерии отбора для таких совокупностей могут определяться, например, общностью местоположения, общностью авторов, хронологией, тематикой, происхождением или принадлежностью и т. д. Коллекции могут содержать любое число объектов, и критерии отбора этих объектов со временем могут изменяться.

5. Практическая реализация

Рассмотренная модель НИС реализована в виде прототипа Системы управления электронными библиотеками (СУЭБ ИРИС). СУЭБ ИРИС оперирует электронными коллекциями. Электронная коллекция — это совокупность информационных объектов, объединенных по смысловому признаку и имеющих одинаковую структуру (схему данных). На рис. 4 изображены базовые технологии, на которых строится работа системы.

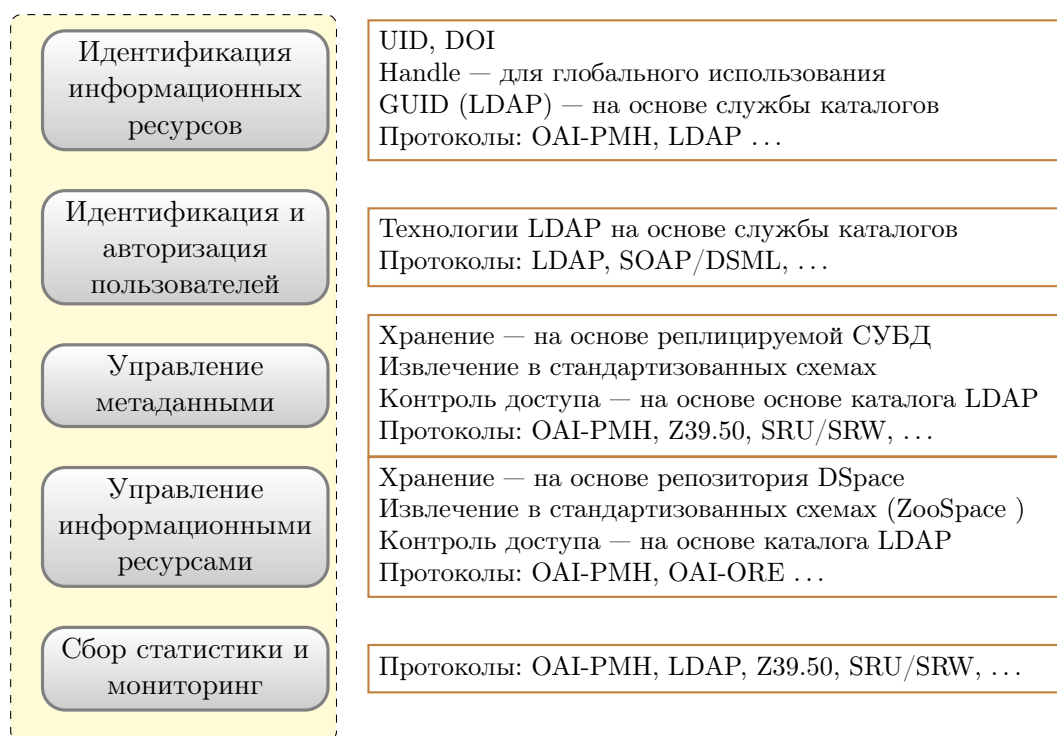


Рис. 4. Используемые технологии реализации ИИС

СУЭБ позволяет работать с двумя видами коллекций: каталогами и тезаурусами. Принципиальное отличие каталогов от тезаурусов состоит в том, что в тезаурусах можно организовывать иерархические зависимости между записями (родитель — потомок, часть — целое и т. п.). Коллекции-каталоги предназначены для хранения и обработки метаданных о документах различной природы: публикациях, ключевых терминах, персонах, организациях, фотографиях и т. д. Коллекции-тезаурусы предназначены для работы с управляемыми словарями и со словарями-классификаторами.

Сервер метаданных СУЭБ содержит служебную коллекцию “Основной каталог метаданных”, которая включает документы, описывающие все метаданные, которые можно использовать в системе. Документы “Основного каталога” содержат описания схемы метаданных QDC [42], расширенной метаданными для соответствия МЕКОФ [43], и служебных метаданных, которые описывают структуру объектов, пользовательские интерфейсы, ассоциативные связи между документами, права доступа к документам и т. д. (при желании он может быть расширен новыми метаданными). Априори каждая коллекция (в зависимости от вида) имеет минимальный обязательный набор метаданных. Администратор коллекции имеет возможность доопределить схему метаданных коллекции исходя из имеющихся метаданных из “Основного каталога”.

В СУЭБ представлено два вида ассоциативных связей между документами (записями): жесткие и мягкие. Жесткие связи реализованы средствами СУБД путем ссылок на первичные ключи записи. К сожалению, такой тип связи не защищен от нарушения целостности (в случае неправильного изменения или удаления записи). Мягкие связи реализуются через процедуру поиска соответствий. Такой способ установления связей защищен от любых нарушений целостности БД и достаточно удобен пользователям, поскольку для указания на необходимость связи используются наглядные мнемонические определения. Соответствия устанавливаются такими двумя способами, как:

— ссылка на идентификатор записи — уникальный (в пределах одной коллекции) текстовый код, формируемый в рамках конкретной коллекции по определенным правилам. Например, для коллекции, содержащей описания персон, идентификатор формируется (на русском языке) последовательно из фамилии, инициалов, года рождения;

— ссылка на ключевой термин — особый вид метаданных, выбираемый из словаря ключевых терминов, по существу представляющий собой тезаурус предметной области коллекции. Ссылка определяет запись, в которой данный ключевой термин присутствует в метаданных.

С целью организации обмена метаданными между репозиториями и сервером метаданных (а также с другими системами, имеющими расширенный профиль) создан специальный сервис, выполняющий преобразование метаданных из внутренней схемы в другие схемы метаданных. Реализован OAI-PMH-сервис, который в пакетном режиме периодически, в соответствии с расписанием, проводит синхронизацию метаданных репозитория и сервера метаданных. Для заполнения “Основного каталога метаданных” в соответствии с созданными схемами метаданных используются управляемые словари из справочного блока сопровождения. Для обеспечения интероперабельности данных также задействован сервер приложений на основе ZooPARK-ZS [5], реализующий доступ к метаданным системы по протоколам Z39.50 [45] и SRW/SRU [48].

Разработанная модель может быть использована как типовая для работы с документами в сфере научно-образовательной деятельности, поскольку решает основные задачи, стоящие перед этими системами: обеспечение надежного долговременного хранения цифровых (электронных) документов с сохранением всех смысловых и функциональных характеристик исходных документов; обеспечение “прозрачного” поиска и доступа пользователей к документам как для ознакомления, так и для анализа содержащихся в них фактов; организация сбора информации по удаленным цифровым репозиториям, поддерживающим протоколы OAI-PMH, SRW/SRU, Z39.50.

Рассмотренная технология создания и поддержки информационных ресурсов кроме работы с научными коллекциями с успехом была реализована в научно-образовательной сфере на примере научной школы Алексея Андреевича Ляпунова — основателя теоретического программирования и российской кибернетики, а также в виде электронных учебных пособий по курсам “Современные проблемы информатики и вычислительной техники”, “Вычислительные системы”, “Информатика” и “Экология” и др.

Благодарности. Исследования выполнены при частичной поддержке РФФИ (грант № 18-07-01457), Интеграционного проекта СО РАН АААА-А18-118022190008-8 (№ 0316-2018-0002) и темы г.з. РФ АААА-А17-117120670141-7 (№ 0316-2018-0009).

Список литературы / References

- [1] **Мирский Э.М.** Массив публикаций и система научной дисциплины “Системные исследования”. 1977. Адрес доступа: <http://courier-edu.ru/pril/posobie/mirdis.htm>
Mirskiy, E.M. The array of publications and the system of scientific discipline “System Studies”. 1977. Available at: <http://courier-edu.ru/pril/posobie/mirdis.htm> (In Russ.)
- [2] **Шокин Ю.И., Федотов А.М., Баракнин В.Б.** Проблемы поиска информации. Новосибирск: Наука, 2010. 198 с.
Shokin, Yu.I., Fedotov, A.M., Barakhnin, V.B. Problems of information retrieval. Novosibirsk: Nauka, 2010. 198 p. (In Russ.)

- [3] **Жижимов О.Л., Федотов А.М., Чубаров Л.Б., Шокин Ю.И.** Технология создания распределенных информационно-вычислительных ресурсов СО РАН // Тр. Первой Междунар. конф. “Системный анализ и информационные технологии” САИТ-2005. 12–16 сентября 2005 г., Переславль-Залесский. Т. 2. М., 2005. С. 161–165.
Zhizhimov O.L., Fedotov A.M., Chubarov L.B., Shokin Yu.I. Technology for creating distributed information and computing resources of the SB RAS // Proc. the First Intern. Conf. “System analysis and information technology” SAIT-2005. September 12–16, 2005, Pereslavl-Zalesskiy. T. 2. Moscow, 2005. P. 161–165. (In Russ.)
- [4] **Федотов А.М.** Методологии построения распределенных систем // Вычисл. технологии. 2006. Т. 11. Спецвыпуск: Избр. докл. X Рос. конф. “Распределенные информационно-вычислительные ресурсы” (DICR-2005), Новосибирск 6–8 октября 2005 г. С. 3–16.
Fedotov, A.M. Methodologies of the distributed systems buildup // Comput. Technologies. 2006. Vol. 11. Special issue: Selected Contributions of the X Russ. Conf. “Distributed Information and Computing Resources” (DICR-2005), Novosibirsk October 6–8. 2005. P. 3–16. (In Russ.)
- [5] **Жижимов О.Л., Федотов А.М., Шокин Ю.И.** Технологическая платформа массовой интеграции гетерогенных данных // Вестн. Новосиб. гос. ун-та. Сер. Информ. технологии. 2013. Т. 11, вып. 1. С. 24–41.
Zhizhimov, O.L., Fedotov, A.M., Shokin, Yu.I. Technology platform for the mass integration of heterogeneous data // Novosibirsk State Univ. J. of Inform. Technologies. 2013. Vol. 11, iss. 1. P. 24–41. (In Russ.)
- [6] **Кулагин М.В., Лопатенко А.С.** Научные информационные системы и электронные библиотеки. Потребность в интеграции // Сб. тр. Третьей Всерос. конф. по электронным библиотекам (RCDL’2001). Петрозаводск, 11–13 сентября 2001 г. КарНЦ, 2001. С. 14–19.
Kulagin, M.V., Lopatenko, A.S. Scientific information systems and electronic libraries. The need for integration // Proc. of the Third All-Rus. Conf. on Electronic Libraries (RCDL’2001). Petrozavodsk, 11–13 September 2001. KarNTs, 2001. P. 14–19. (In Russ.)
- [7] **Шокин Ю.И., Федотов А.М., Гуськов А.Е. и др.** Электронные библиотеки — путь интеграции информационных ресурсов Сибирского отделения РАН // Вестн. КазНУ. Сер. Математика, механика, информатика. 2005. № 2. С. 115–127. (ISSN 1563-0285).
Shokin, Yu.I., Fedotov, A.M., Guskov, A.E. et al. Electronic libraries the way of integration of information resources of the Siberian Branch of the RAS // Vestn. KazNU. Ser. Matematika, Mechanika, Informatika. 2005. No. 2. P. 115–127. (In Russ.)
- [8] **Шокин Ю.И., Федотов А.М.** Электронная библиотека Сибирского отделения РАН // Информационное общество. 2000. № 2. С. 22–31.
Shokin, Yu.I., Fedotov, A.M. Electronic library of the Siberian Branch of the RAS // Information Society. 2000. No. 2. P. 22–31. (In Russ.)
- [9] **Жижимов О.Л., Мазов Н.А., Федотов А.М.** Некоторые заметки об эволюции цифровых репозиториях традиционных библиотек к полнофункциональным электронным библиотекам // Вестн. Владивосток. гос. ун-та экономики и сервиса. Территория новых возможностей. 2010. Т. 7, № 3. С. 55–63.
Zhizhimov, O.L., Mazov, N.A., Fedotov, A.M. Some notes on the evolution of digital repositories of traditional libraries towards full-featured electronic libraries // Vladivostok State Univ. of Economics and Service. 2010. Vol. 7, No. 3. P. 55–63. (In Russ.)
- [10] **Антопольский А.Б., Вигурский К.В.** Концепция электронных библиотек // Электронные библиотеки: Рос. науч. электронный журн. 1999. Т. 2, вып. 2. Адрес доступа: http://elib.ict.nsc.ru/jspui/bitstream/ICT/1222/1/DL_antapol_1999.pdf

- Antopolsky, A.V., Vigursky, K.V.** The concept of electronic libraries // Electronic Libraries: Russ. Sci. Electronic J. 1999. T. 2, iss. 2. Available at: http://elib.ict.nsc.ru/jspui/bitstream/ICT/1222/1/DL_antapol_1999.pdf (In Russ.)
- [11] **Михайлов А.И., Черный А.И., Гиляревский Р.С.** Научные коммуникации и информатика. М.: Наука, 1976. 82 с.
Mikhaylov, A.I., Chernyy, A.I., Gilyarevskiy, R.S. Scientific communications and informatics. Moscow: Nauka, 1976. 82 p. (In Russ.)
- [12] **Жижимов О.Л., Федотов А.М., Федотова О.А.** Построение типовой модели информационной системы для работы с документами по научному наследию // Вестн. Новосиб. гос. ун-та. Сер. Информ. технологии. 2012. Т. 10, № 3. С. 5–14.
Zhizhimov, O.L., Fedotov, A.M., Fedotova, O.A. Building a generic model of information system for working with documents on the scientific heritage // Novosibirsk State Univ. J. of Inform. Technologies. 2012. Vol. 10, No. 3. P. 5–14. (In Russ.)
- [13] **Резниченко В.А., Проскудина Г.Ю., Кудим К.А.** Концептуальная модель электронной библиотеки [Электронный текст] // Тр. XI Всерос. науч. конф. RCDL'2009, Петрозаводск (Карелия), 17–21 сентября 2009 г. С. 23–31.
Reznichenko, V.A., Proskudina, G.Yu., Kudim, K.A. The conceptual model of the electronic library [Electronic text] // Proc. of the 11th All-Russ. Sci. Conf. RCDL'2009, Petrozavodsk (Karelia), September 17–21, 2009. P. 23–31. (In Russ.)
- [14] D3.2b The Digital Library Reference Model. Funded under the Seventh Framework Programme, ICTProgramme “Cultural Heritage and Technology Enhanced Learning”. Project Number: 231551, April 2011. Available at: http://nmis.isti.cnr.it/casarsa/FUB-DL/readings/DL_Conceptual_Model_v1.0.pdf
- [15] **Candela, L., Castelli, D., Dobрева, M. et al.** The DELOS Digital Library Reference Model Foundations for Digital Libraries. IST-2002-2.3.1.12. Technology-enhanced Learning and Access to Cultural Heritage. Version 0.98, Dec. 2007. 215 p. Available at: <https://pureportal.strath.ac.uk/en/publications/the-delos-digital-library-reference-model-foundations-for-digital>
- [16] Definition of the CIDOC Conceptual Reference Model. Produced by the ICOM/CIDOC Documentation Standards Group, Continued by the CIDOC CRM Special Interest Group. Version 5.0.4, Nov. 2011. Available at: <http://www.cidoc-crm.org/html/5.0.4/cidoc-crm.html>
- [17] ISO-14721 Reference Model for an Open Archival Information System (OAIS), Draft Recommended Standard, CCSDS 650.0-P-1.1 (Pink Book). June 2012. Available at: <https://public.ccsds.org/pubs/650x0m2.pdf>
- [18] Функциональные требования к библиографическим записям: концептуальная модель: окончательный отчет: Пер. с англ. В.В. Арефьев. М.: Рос. гос. библиотека, 2006. 150 с.
Functional requirements for bibliographic records: final report / IFLA Study Group on the Functional Requirements for Bibliographic Records. Munchen: K.G. Saur, 1998. viii, 136 p. (UBCIM publ., new series, vol. 19). Available at: <https://www.ifla.org/publications/functional-requirements-for-bibliographic-records>
- [19] CERIF 2008-1.2 Full Data Model (FDM). Introduction and Specification. Available at: http://www.eurocris.org/Uploads/Web%20pages/CERIF2008/Release_1.2/CERIF2008_1.2_FDM.pdf
- [20] **Бездушный А.Н., Бездушный А.А., Серебряков В.А., Филиппов В.И.** Интеграция метаданных Единого научного информационного пространства РАН. М.: Вычисл. центр РАН им. А.А. Дородницына, 2006. 258 с.

- Bezdushny A.N., Bezdushny A.A., Serebryakov, V.A., Filippov, V.I.** Integration of the metadata of the Unified Scientific Information Space of the RAS. Moscow: Vychisl. Tsentr RAN im. A.A. Dorodnitsyna, 2006. 258 p. (In Russ.)
- [21] **Захаров А.А., Серебряков В.А.** Система управления электронными библиотеками LibMeta // Тр. 12-й Всерос. науч. конф. “Электронные библиотеки: перспективные методы и технологии, электронные коллекции” RCDL’2010, Казань. Казан. ун-т, 2010. С. 28–37.
Zakharov, A.A., Serebryakov, V.A. Electronic library management system LibMeta // Proc. of the 12th All-Russ. Sci. Conf. “Digital Libraries: Advanced Methods and Technologies, Digital Collections” RCDL’2010, Kazan. Kazanskiy un-t, 2010. P. 28–37. (In Russ.)
- [22] **Жижимов О.Л., Пестунов И.А., Федотов А.М.** Структура сервисов управления метаданными для разнородных информационных систем // Электронные библиотеки. 2012. Т. 15, № 6. Адрес доступа: <http://elbib.kpfu.ru/ru/article/328>
Zhizhimov, O.L., Pestunov, I.A., Fedotov, A.M. Structure of metadata services management for heterogeneous information systems // Digital Library. 2012. Vol. 15, No. 6. Available at: <http://elbib.kpfu.ru/ru/article/328> (In Russ.)
- [23] **Шокин Ю.И., Федотов А.М., Жижимов О.Л.** Технология распределенных информационных систем // Матер. конф. “Современные информационные технологии для научных исследований”. Магадан, 2008. Магадан: СВНЦ ДВО РАН, 2008. С. 18–21.
Shokin Yu. I., Fedotov, A.M., Zhizhimov, O.L. Technology of distributed information systems // Proc. of the Conference “Modern Information Technologies for Scientific Research”. Magadan, 2008. Magadan: SVNTs DVO RAN, 2008. P. 18–21. (In Russ.)
- [24] **Шокин Ю.И., Федотов А.М., Жижимов О.Л., Федотова О.А.** Эволюция информационных систем: от Web-сайтов до систем управления информационными ресурсами // Вестн. Новосиб. гос. ун-та. Сер. Информ. технологии. 2015. Т. 13, № 1. С. 117–134.
Shokin, Yu.I., Fedotov, A.M., Zhizhimov, O.L., Fedotova, O.A. The evolution of information systems: from websites to information resource management systems // Novosibirsk State Univ. J. of Inform. Technologies. 2015. Vol. 13, No. 1. P. 117–134. (In Russ.)
- [25] **Абрамов А.Г., Булакина М.Б., Сигалов А.В.** Единое окно доступа к образовательным ресурсам: информационное наполнение и использование в учебном процессе // Дистанц. и виртуал. обучение. 2011. № 8. С. 65–80.
Abramov, A.G., Bulakina, M.B., Sigalov, A.V. Single window of access to educational resources: content and use in the educational process // Distance and Virtual Training. 2011. No. 8. P. 65–80. (In Russ.)
- [26] Стандарт ГНИИ ИТТ “Информика”. Метаданные информационных образовательных ресурсов для Интернет-каталогов. М.: ГНИИ ИТТ “Информика”, 2004. 89 с.
Standard GNII ITT “Informika”. Metadata informational educational resources for Internet-catalogs. Moscow: GNII ITT “Informika”, 2004. 89 p. (In Russ.)
- [27] **Афонин С.А. и др.** Интеллектуальная система тематического исследования научно-технической информации (ИСТИНА) / Под ред. В.А. Садовниченко. М.: МГУ, 2014. 262 с.
Afonin, S.A. et al. Intellectual system of case study of scientific and technical information (ISTINA) / Ed. V.A. Sadovnichiy. Moscow: MGU, 2014. 262 p. (In Russ.)
- [28] **Аджиев А.С., Бездушный А.Н., Серебряков В.А.** О реализации веб-системы математической информации // Рос. науч. электронный журн. Электронные библиотеки. 2004. Т. 7. Вып. 1. Адрес доступа: <http://www.elbib.ru/index.phtml?page=elbib/rus/journal/2004/part1/ABS>
Adzhiev, A.S., Bezdushny, A.N., Serebryakov, V.A. About implementation of Web-system of mathematical information // Russ. Digital Libraries J. 2004. Vol. 7, iss. 1. Available at: <http://www.elbib.ru/index.phtml?page=elbib/eng/journal/2004/part1/ABS> (In Russ.)

- [29] **Паринов С.И., Ляпунов В.М., Пузырев Р.Л.** Система Соционет как платформа для разработки научных информационных ресурсов и онлайн-сервисов // Рос. науч. электронный журн. Электронные библиотеки. 2003. Т. 6, вып. 1. Адрес доступа: <http://www.elbib.ru/index.phtml?page=elbib/rus/journal/2003/part1/PLP>
Parinov, S.I., Lyapunov, V.M., Puzyrev, R.L. The Socionet as a platform for creation of information resources and online services for a community of researchers // Russ. Digital Libraries J. 2003. Vol. 6, iss. 1. Available at: <http://www.elbib.ru/index.phtml?page=elbib/rus/journal/2003/part1/PLP> (In Russ.)
- [30] **Калёнов Н.Е., Савин Г.И., Серебряков В.А., Сотников А.Н.** Принципы построения и формирования электронной библиотеки “Научное наследие России” // Программные продукты и системы. 2012. № 4. С. 30–40.
Kalenov, N.E., Savin, G.I., Serebryakov, V.A., Sotnikov, A.N. Scientific heritage of Russia digital library: Construction and sources aggregation philosophy // Programmnye Produkty i Sistemy. (Software & Systems). 2012. No. 4. P. 30–40. (In Russ.)
- [31] **Атаева О.М., Серебряков В.А.** Онтология цифровой семантической библиотеки LibMeta // Информатика и ее применения. 2018. Том 12, № 1. С. 2–10.
Ataeva, O.M., Serebryakov, V.A. Ontology of the digital semantic library LibMeta // Informatics and Applications. 2018. Vol. 12, No. 1. P. 2–10. (In Russ.)
- [32] Онтологии и тезаурусы: Учеб. пособие / В.Д. Соловьев, Б.В. Добров, В.В. Иванов, Н.В. Лукашевич. М.: Интернет-ун-т информ. технологий; БИНОМ. Лаб. знаний, 2009. 173 с.
Ontologies and thesauri: models, tools, applications: a tutorial / V.V. Dobrov, V.V. Ivanov, N.V. Lukashevich, V.D. Solovyov. Moscow: Intern. Univ. Inform. Technologi; BINOMI. Lab. Znaniy, 2009. 173 p. (In Russ.)
- [33] **Федотов А.М., Барахнин В.Б., Жижимов О.Л., Федотова О.А.** Технология создания корпоративных информационных систем учета трудов научных работников // Вестн. Новосиб. гос. ун-та. Сер. Информ. технологии. 2011. Т. 9, вып. 2. С. 31–41.
Fedotov, A.M., Barakhnin, V.B., Zhizhimov, O.L., Fedotova, O.A. The technology of creation of corporate information systems for accounting resources, created by researchers // Novosibirsk State Univ. J. of Inform. Technologies. 2011. Vol. 9, No. 2. P. 31–41. (In Russ.)
- [34] **Колесникова Н.И.** От конспекта к диссертации: Учеб. пособие по развитию навыков письменной речи. М.: Флинта: Наука, 2002. 288 с.
Kolesnikova, N.I. From the abstract to the thesis: A manual on the development of writing skills. Moscow: Flint: Nauka, 2002. 288 p. (In Russ.)
- [35] **Когаловский М.Р.** Метаданные, их свойства, функции, классификация и средства представления // Тр. 14-й Всерос. науч. конф. “Электронные библиотеки: перспективные методы и технологии, электронные коллекции” RCDL-2012, 15–18 октября 2012 г. Переславль-Залесский, 2012. С. 25–36. Адрес доступа: <http://rcdl.ru/doc/2012/paper3.pdf>
Kogalovsky, M.R. Metadata, their properties, functions, classification and means of representation // Proc. of the 14th All-Rus. Sci. Conf. “Electronic Libraries: Advanced Methods and Technologies, Digital Collections” RCDL-2012, 15–18 October 2012. Pereslavl-Zalessky, 2012. P. 25–36. Available at: <http://rcdl.ru/doc/2012/paper3.pdf> (In Russ.)
- [36] **Башмаков А.И., Старых В.А.** Систематизация информационных ресурсов для сферы образования: классификация и метаданные. М.: Фонд “Европейский центр по качеству”, 2003. 384 с.
Bashmakov, A.I., Staryh, V.A. Systematization of information resources for education: classification and metadata. Moscow: Foundation “European Center for Quality”, 2003. 384 p. (In Russ.)

- [37] ISO/IEC TR 10000-1:1998. Information technology — Framework and taxonomy of International Standardized Profiles. Pt 1: General principles and documentation framework. Available at: <https://www.iso.org/ru/standard/30726.html>
- [38] ГОСТ Р ИСО / МЭК ТО 10000-2-99 Информационная технология (ИТ). Основы и таксономия функциональных стандартов. Ч. 2. Принципы и таксономия профилей ВОС. Адрес доступа: <http://docs.cntd.ru/document/1200027469>
GOST R ISO / IEC TO 10000-2-99 Information technology (IT). Framework and taxonomy of International Standardized Profiles. Pt 2. Principles and taxonomy for OSI profiles. Available at: <http://docs.cntd.ru/document/1200027469> (In Russ.)
- [39] **Федотов А.М., Жижимов О.Л., Федотова О.А., Барахнин В.Б.** Модель информационной системы для поддержки научно-педагогической деятельности // Вестн. НГУ. Сер. Информ. технологии. 2014. Т. 12, № 1. С. 89–101.
Fedotov, A.M., Zhizhimov, O.L., Fedotova O.A., Barakhnin, V.B. A model of information system to support scientific and educational activities // Novosibirsk State Univ. J. of Inform. Technologies. 2014. Vol. 12, No. 1. P. 89–101. (In Russ.)
- [40] **Жижимов О.Л., Федотов А.М., Шокин Ю.И.** Технология интеграции разнородных информационных ресурсов // Тр. V Междунар. конф. “Системный анализ и информационные технологии” САИТ-2013. Красноярск: ИВМ СО РАН, 2013. Т. 2. С. 129–136. Адрес доступа: <http://elib.sbras.ru:8080/jspui/bitstream/SBRAS/9212/1/ZooSpace.pdf>
Zhizhimov, O.L., Fedotov, A.M., Shokin, Yu.I. The technology of integration of heterogeneous information resources // Proc. of the V Intern. Conf. “System Analysis and Information Technologies” SAIT-2013. Krasnoyarsk: ICM SB RAS, 2013. T. 2. P. 129–136. Available at: <http://elib.sbras.ru:8080/jspui/bitstream/SBRAS/9212/1/ZooSpace.pdf> (In Russ.)
- [41] **Федотов А.М., Жижимов О.Л., Князева А.А. и др.** Проблемы авторитетного контроля для распределенных электронных библиотек и библиографических баз данных // Вестн. НГУ. Сер. Информ. технологии. 2011. Т. 9, № 1. С. 89–101.
Fedotov, A.M., Zhizhimov, O.L., Knyazeva A.A. et al. Problems of authority control for distributed digital libraries and bibliographic databases // Novosibirsk State Univ. J. of Inform. Technologies. 2011. Vol. 9, No. 1. P. 89–101.
- [42] DCMI — Dublin Core Metadata Initiative. Available at: <http://www.dublincore.org/>
- [43] ГОСТ 7.19-2001. Система стандартов по информации, библиотечному и издательскому делу. Формат для обмена данными. Содержание записи. Адрес доступа: <http://docs.cntd.ru/document/1200025970>
GOST 7.19-2001. System of standards on information, librarianship and publishing. Format for data exchange. Content of the record. Available at: <http://docs.cntd.ru/document/1200025970> (In Russ.)
- [44] International DOI Foundation. Available at: <http://www.doi.org/>
- [45] ANSI/NISO Z39.50-2003. Information retrieval (Z39.50): Application service definition and protocol specification. NISO Press, Bethesda, Maryland, U.S.A. Nov. 2002. Available at: <https://www.loc.gov/z3950/agency/Z39-50-2003.pdf>
- [46] Формирование открытых баз данных информационных ресурсов в области образования, науки и культуры: Отчет о НИР. ВИНТИ РАН, 2005 г. Адрес доступа: <http://science.viniti.ru/>
Formation of open databases for information resources in education, science and culture: Otchet o NIR. VINITI RAN, 2005. Available at: <http://science.viniti.ru/> (In Russ.)

- [47] Heery, R., Patel, M. Application profiles: mixing and matching metadata schemas. Ariadne, 25 Sept. 2000. Available at: <http://www.ariadne.ac.uk/issue25/app-profiles/intro.html>
- [48] SRU (Search / Retrieve via URL). Available at: <http://www.loc.gov/standards/sru/>

Поступила в редакцию 22 августа 2018 г.

Requirements for the prototype of the information resources management system in distributed information systems for the support of scientific research

FEDOTOV, ANATOLIY M.* , LEONOVA, YULIYA V.

Institute of Computational Technologies SB RAS, Novosibirsk, 630090, Russia

*Corresponding author: Fedotov, Anatoliy M., e-mail: fedotov@sbras.ru

The article addresses the description of technological approaches used to create distributed information systems for supporting scientific research. We describe the developed architectural solutions which are intended to create such systems and the principles for integrating of these systems with the external sources. The functional requirements for the model of such an information system are determined. These requirements are determined, firstly, by the information needs of researchers, and secondly, by ensuring a reliable and long-term storage of information.

The considered model of SIS is implemented in the form of a prototype of the Electronic Library Management System (ELMS IRIS). ELMS IRIS operates with electronic collections. The electronic collection is considered as a set of information objects, united on a semantic basis and having the same structure (data schema). An ELMS allows to work with two types of collections which are the catalogs and the thesauri. The rules of representation and transformation of metadata are discussed. In order to organize metadata exchange between repositories and a metadata server (as well as with other systems with an extended profile), a special service is created that converts metadata from the internal schema into other metadata schemes.

The OAI-PMH service is implemented, which periodically synchronizes the metadata of the repository and the metadata server in accordance with the schedule in batch mode. To fill the “Main metadata catalog”, in accordance with the metadata schemas created, managed dictionaries are used from the reference support block. To ensure data interoperability, the application server based on ZooPARK-ZS is also implemented, which provides access to system metadata via Z39.50 and SRW/SRU protocols.

The developed model can be used as a model of the system for working with documents related to scientific and educational activities, since it solves the main tasks for these systems which are to provide a system for reliable long-term storage of digital (electronic) documents while preserving all the semantic and functional characteristics of source documents as well as ensuring “transparent” search and access of users to documents. These actions are necessary both for familiarization and for the analysis of the facts contained therein; and organization of collection of information on remote digital repositories supporting protocols OAI-PMH, SRW/SRU, Z39.50.

Keywords: information system, electronic library, dictionary-directory, distributed information resources, data integration, digital repository, information retrieval, meta-data, functional requirements, bibliographic databases.

Cite: Fedotov, A.M., Leonova, Yu.V. Requirements for the prototype of the information resources management system in distributed information systems for the support of scientific research // Computational Technologies. 2018. Vol. 23, No. 5. P. 82–109. (In Russ.) DOI: 10.25743/ICT.2018.23.5.008.

Acknowledgements. This research was partly supported by RFBR (grant No. 18-07-01457), Integration Project of SB RAS AAAA-A18-118022190008-8 (No. 0316-2018-0002) and subjects of state order AAAA-A17-117120670141-7 (No. 0316-2018-0009).

Received 22 August 2018