

Учреждение образования
«Белорусский государственный университет культуры и искусств»
Институт повышения квалификации и переподготовки кадров

УТВЕРЖДАЮ

Ректор университета

Н.В.Карчевская

09 .2022



**МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ
ПО УЧЕБНОЙ ДИСЦИПЛИНЕ
«МЕТАДААННЫЕ»
специальности переподготовки
2-23 02 71 Библиотекведение и библиография**

Составители:

Федорина А.И., доцент кафедры информационных ресурсов и коммуникаций учреждения образования «Белорусский государственный университет культуры и искусств»;

Предеина А.В., доцент кафедры культурологии и психолого-педагогических дисциплин Института повышения квалификации и переподготовки кадров учреждения образования «Белорусский государственный университет культуры и искусств», кандидат педагогических наук

Рассмотрен и рекомендован к утверждению:

*Кафедрой культурологии и психолого-педагогических дисциплин
(протокол от 07.09.2022 № 36);*

*Советом Института повышения квалификации и переподготовки кадров
(протокол от 22.09.2022 № 8).*

СОДЕРЖАНИЕ

| | | |
|-----|--|----|
| 1 | ПОЯСНИТЕЛЬНАЯ ЗАПИСКА..... | 4 |
| 2 | ТЕОРЕТИЧЕСКИЙ РАЗДЕЛ..... | 6 |
| | Конспект лекций..... | 6 |
| 3 | ПРАКТИЧЕСКИЙ РАЗДЕЛ..... | 47 |
| | Тематика и вопросы круглых столов | 47 |
| 4 | РАЗДЕЛ КОНТРОЛЯ ЗНАНИЙ..... | 48 |
| 4.1 | Материалы для самоконтроля по дисциплине «Метаданные»..... | 48 |
| 4.2 | Задания для самостоятельной работы..... | 50 |
| 5 | ВСПОМОГАТЕЛЬНЫЙ РАЗДЕЛ..... | 52 |
| 5.1 | Учебная программа..... | 52 |
| 5.2 | Основная литература..... | 68 |
| 5.3 | Дополнительная литература..... | 68 |

1 ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

Представленные методические рекомендации по дисциплине «Метаданные» предусматривают реализацию требований, установленных Образовательным стандартом переподготовки руководящих работников и специалистов по специальности 1-23 02 71 «Библиотекведение и библиография», утвержденным постановлением Министерства образования Республики Беларусь от 28.09.2020 № 259.

Дисциплина рассчитана на 42 учебных часа, которые включают 14 учебных часов аудиторных занятий (в том числе: 6 учебных часов лекций, 6 учебных часов практических занятий, 2 учебных часа круглых столов); 28 учебных часов самостоятельной работы.

Дисциплина «Метаданные» относится к компоненту учреждения образования типового учебного плана и изучается во взаимосвязи с дисциплинами «Каталогизация документов», «Основы информационных технологий» и другими.

Цель методических рекомендаций: оказание методической помощи в изучении современных подходов к каталогизации электронных документов сетевого распространения.

Задачи:

освоение теоретических знаний о метаданных;
сформировать представление о форматах метаданных;
познакомиться с метаданными цифровых ресурсов;
познакомиться с технологией формирования метаданных цифровых ресурсов;
овладеть методикой каталогизации национальных электронных документов сетевого распространения Беларуси.

Методы обучения: традиционные (методы организации и осуществления учебно-познавательной деятельности; методы стимулирования и мотивации учебно-познавательной деятельности; методы контроля и самоконтроля); активный; интерактивный.

Средства обучения: печатные издания; презентации, видео- и аудиозаписи.

Основные требования к результатам учебной деятельности слушателей.

В результате освоения дисциплины «Метаданные» у слушателя должны быть сформированы социально-личностные, академические и профессиональные компетенции.

Требования к формированию социально-личностных компетенций.

Слушатель должен:

знать идеологию белорусского государства, нравственные и правовые нормы, уметь учитывать их в своей профессиональной деятельности;

уметь ориентироваться в процессах и тенденциях развития современного общества.

Требования к формированию академических компетенций.

Слушатель должен:

владеть базовыми теоретическими знаниями в области документоведения, библиотековедения и библиографоведения, уметь применять их для решения практических и научных задач в библиотечной сфере;

уметь проводить сравнительный анализ и оценку теоретических концепций библиотечного дела и библиографии с целью использования библиотечно-библиографических знаний в профессиональной деятельности;

владеть теоретическими знаниями о сущности библиографии;

уметь использовать информационные технологии в библиотечной сфере.

Требования к формированию профессиональных компетенций.

Слушатель должен:

использовать теоретические знания о метаданных;

анализировать и использовать структуру коммуникативных форматов в автоматизированной обработке документов;

знать стандарты метаданных;

знать технологию формирования метаданных цифровых ресурсов;

уметь представлять метаданные цифровых ресурсов;

уметь осуществлять каталогизацию национальных электронных документов сетевого распространения Беларуси;

уметь работать в АРМ «Каталогизатор» АБИС.

2 ТЕОРЕТИЧЕСКИЙ РАЗДЕЛ

Конспект лекций

Раздел 1 Основа теории метаданных

Тема 1.1 Метаданные: терминосистема, свойства, функции, классификация

Хранение и обработка данных в компьютерных системах, обмен данными между ними и доступ к ним пользователей невозможны без явно представленных описаний свойств этих данных. Эти описания необходимы программным средствам, выполняющим указанные функции, а также пользователям для формулировки запросов, анализа данных и интерпретации их содержания. Описания такого рода называются метаданными и являются особым видом информационных ресурсов. Их создание часто требует значительных усилий и существенных затрат. Однако они существенно повышают ценность данных, обеспечивают более широкие возможности их использования. Метаданные начали использоваться задолго до появления компьютерных систем и до введения этого термина в научно-техническую лексику. Библиографические описания использованных источников в публикациях, библиотечные каталоги, различные классификаторы, аннотации статей – все это примеры метаданных. Рассмотрим метаданные, используемые в компьютерных системах. В этой сфере метаданные начали использоваться уже на ранней стадии ее развития. Когда начали создаваться языки и технологии программирования, термин «метаданные» еще не существовал. Однако в исходном коде программы на языке программирования должно было содержаться описание типов данных, которыми эта программа оперирует. На их основе компилятор обеспечивает требуемое для этих данных распределение оперативной памяти, оценивает правомерность выполнения над ними предусмотренных в программе операций. Если программа оперирует данными во внешней памяти, то в ее исходном коде, а для некоторых операционных систем и в языках управления заданиями, должны описываться свойства файлов, в которых хранятся эти данные. Указанные описания, несомненно, являются примерами видов метаданных. Рождение технологий баз данных и информационно-поисковых систем также потребовало использования метаданных. При проектировании конкретной базы данных необходимо сформировать концептуальную схему предметной области, представляя ее средствами какого-либо языка концептуального моделирования. На ее основе нужно создать описание структуры базы данных, ограничений целостности, полномочий пользователей и т.д. Для этой цели используется язык описания данных СУБД, выбранной для реализации системы базы данных. Указанное описание представляется в виде схемы создаваемой базы данных, которая, также как и концептуальная схема предметной области, является метаданными. В ранних документальных информационно-поисковых

системах каждый документ представлялся его поисковым образом, который идентифицировал этот документ, характеризовал его содержание набором дескрипторов и заменял этот документ в процессах обработки пользовательских запросов. Поисковый образ исходного документа является его метаданными. Метаданными текстовых публикаций в электронных библиотечных каталогах служат также их библиографические описания и аннотации, рубрики тематических классификаторов, к которым относится содержание данных публикаций, и другие их характеристики. В 1960-1970 гг. были довольно широко распространены генераторы отчетов с разработанным и впервые реализованным компанией IBM входным языком PRG (Report Program Generator). Этот язык позволяет описывать формат генерируемого отчета. Такие описания также, несомненно, являются метаданными. Появление гипертекстовых технологий, а затем и Всемирной паутины, предусматривает использование описаний гипертекстовой разметки исходного текста в гипертекстовых издательских системах и в веб-страницах. Совокупность тегов разметки также представляет собой метаданные таких ресурсов. Создание систем, основанных на знаниях, и Семантического Веба, позволило явным образом представлять и использовать для поиска и логического вывода семантику данных, а также онтологию предметной области. Для их описания в последние годы активно используются, в частности, язык RDF, языки описания онтологий RDFS, OWL и OWL2, а также профили языка OWL2. Такие описания также являются примерами метаданных. В научных информационных системах необходимо описывать свойства используемых в них данных, связанные с особенностью предметной области исследований. Вместе с тем, могут быть необходимы описания характеристик научных приборов, с помощью которых они получены, место и время регистрации этих данных и т.п.

Метаданные необходимы в современных электронных библиотеках. Они описывают контент библиотеки в целом, составляющие его коллекции информационных ресурсов, отдельные ресурсы, содержащиеся в коллекциях, их классификаторы, организации, владеющие информационными ресурсами библиотеки, ее пользователей (их профили), ряд других информационных объектов и сервисов таких систем.

Осознание необходимости метаданных привело к созданию инструментария управления этим специфическим видом информационных ресурсов. Впервые серьезное внимание проблемам управления метаданными начало уделяться в 1970-е годы в контексте информационных систем, оперирующих структурированными данными. Была предложена концепция системы словаря-справочника данных, которая получила развитие. Несколько позднее была реализована концепция интегрированных словарей-справочников данных – систем управления базами данных, выполняющих традиционные функции СУБД и одновременно функции системы словаря-справочника данных. Идея интеграции системы словаря-справочника данных и СУБД, хотя и в ограниченной форме, позднее нашла воплощение в стандарте языка SQL. В нем предусмотрено представление схемы базы данных (на «логическом» уровне) в форме набора системных таблиц базы

данных, которыми можно оперировать с помощью обычных средств языка. В результате базы данных SQL стали самоописываемыми – они содержат пользовательские данные вместе с описывающими их метаданными. Развитие информационных технологий и сферы их применения привело к существенному расширению функций метаданных и их многообразия.

Средства представления метаданных и управления ими созданы и развиваются как для информационных, так и для других компьютерных систем. Содержание метаданных, их функции и средства их представления зависят от используемых информационных технологий, функциональных возможностей и предметной области использующих их систем, природы описываемых ресурсов, контекста и характера их использования, а также от многих других факторов. В последние два десятилетия метаданные стали объектом для многих исследований, главным образом, в связи с развитием технологий Семантического Веба, электронных библиотек и ряда других новых пластов информационных технологий. Кроме того, стала актуальной необходимость обеспечения обмена метаданными между различными системами, обеспечения интероперабельности и повторного использования информационных ресурсов. Все это вызвало активную деятельность по стандартизации метаданных, которой занимаются международные органы стандартизации, промышленные консорциумы, научные и другие сообщества. В результате были созданы многочисленные стандарты описания метаданных горизонтальной и вертикальной сферы. Активно используются на практике стандарты платформы XML, Дублинское ядро, дескриптивное подмножество языка SQL, большое число схем метаданных для различных областей применения, стандарты языков концептуального моделирования, многочисленные стандарты метаданных в науке – геоданных, астрономических данных, а также научных данных в других областях исследований. Использование термина «метаданные» существенно активизировалось в последние годы. Однако, к сожалению, не сложилось однозначное понимание этого термина. Широко распространенная абстрактная формула «метаданные – это данные о данных» не раскрывает многообразия свойств и функций этого вида информационных ресурсов. Хотя метаданным в последние годы посвящено огромное количество публикаций, в них мало внимания уделяется систематическому обсуждению их общих свойств и функций. Чаще всего обсуждаются проблемы, связанные с созданием систем метаданных для конкретных областей научных исследований, бизнеса и других сфер деятельности, разработок электронных библиотек, разнообразных репозиториев цифровых информационных объектов, информационных систем для конкретных сфер применения. При этом во многих работах, посвященных электронным библиотекам или приложениям, основанным на стандартах Семантического Веба, доминирует весьма ограниченный взгляд на функции метаданных. Метаданные рассматриваются в них лишь как средство описания контента (содержания) информационных ресурсов, хотя это только одна из многих возможных функций метаданных. Формирование адекватного и достаточно полного представления о независимых от конкретной области применения свойствах

и функциях метаданных является, на наш взгляд, актуальной проблемой. В данной работе предлагается систематический взгляд на метаданные как на информационный ресурс особого рода, анализируются распространенные трактовки этого термина, приводятся примеры метаданных, используемых в различных видах компьютерных систем, рассматриваются общие (независимые от сферы применения) свойства метаданных и их функции, выразительные средства для их представления, обсуждаются вопросы их обобщенной классификации.

Существуют различные точки зрения на содержание термина «метаданные». Неоднозначность в понимании смысла и функций метаданных связана с доминирующими в настоящее время двумя подходами. Первый из них свойственен библиотечному сообществу. Он связан с созданием и использованием документальных ИПС, электронных каталогов библиотечных фондов, а в последние годы – с разработками репозиториев цифровых объектов, представляющими собой основную проблематику электронных библиотек. При этом представители библиотечного сообщества имеют в виду, главным образом, текстовые информационные объекты. Истоки этого подхода некоторые авторы видят в библиотечной науке (Library Science). В отличие от него, истоки второго подхода, по их мнению, относятся к «компьютерным наукам» (Computer Sciences). К этой сфере, относятся развивающиеся уже около полувека технологии баз данных и другие области информатики, связанные с управлением данными и знаниями. Нужно заметить, что на достигнутой стадии развития электронных библиотек, технологии их разработки широко используют самые разнообразные достижения «компьютерных наук», которые далеко выходят за рамки предлагаемых «библиотечной наукой». Это в первую очередь связано с расширением разнообразия видов информационных ресурсов, составляющих контент современных электронных библиотек, и с расширением объема требований пользователей к функциональности этих информационных систем. Сосуществование двух указанных подходов, а также огромное многообразие видов метаданных и сфер их использования, в значительной мере определяют разнообразие имеющихся трактовок термина «метаданные».

Рассмотрим некоторые наиболее распространенные из них. Определение 1. Метаданные – это данные о данных. Определение 2. Метаданные – это информация, которая делает данные полезными. Определение 3. Метаданные – это машинно-обрабатываемые данные, которые описывают некоторые ресурсы, цифровые и нецифровые. Определение 4. Метаданными называется допускающая компьютерную обработку и интерпретацию человеком информация о цифровых и нецифровых объектах. Определение 5. Метаданные – это структурированные данные, которые описывают характеристики некоторого ресурса. Определение 6. Метаданными называется структурированная информация, которая описывает, поясняет, указывает местоположение и иным образом облегчает поиск, использование информационного ресурса, а также управление им. Определение 7. Метаданные – это структурированные

данные, представляющие собой характеристики описываемых сущностей для целей их идентификации, поиска, оценки, управления ими. Определение 8. Метаданные – это описание не только состава данных, их структуры (формата) представления, места хранения и других их признаков, но и поддерживающих их информационных систем, технологий, пользователей, методов доступа. Определение 9. Во Всемирной паутине метаданные – это слабоструктурированные данные, как правило, соответствующие согласованным моделям, обеспечивающим операционную интероперабельность в неоднородной среде. Определение 10. Метаданными называется любая дескриптивная информация о других источниках данных, которая способствует организации, идентификации, представлению, определению местоположения, обеспечению интероперабельности, управлению и использованию этих данных. Определение 11. Метаданные – это данные, относящиеся к некоторому элементу данных, т.е. метаданные – это данные о данных. Для элемента данных спецификация, определяющая, каким образом он может создаваться, в каких контекстах он может использоваться, как его трансформировали или как он может интерпретироваться либо обрабатываться. Определение 12. Метаданные обычно определяются как «данные о данных», которые предназначены для выражения семантики информации, следовательно, для улучшения ее поиска и выборки, понимания и использования. Определение 13. Метаданные – это данные из более общей формальной системы, описывающей заданную систему данных. Проанализируем кратко этот довольно представительный набор определений. Определения 1 и 2 являются чрезмерно общими, и не дают конструктивного представления о существе метаданных в определении 3 декларируется универсальный характер (цифровой и нецифровой) информационных ресурсов, которые могут описываться метаданными. Отмечается важное свойство метаданных – возможность их компьютерной обработки. Однако при этом не указывается предназначение метаданных – ориентированы ли они на поддержку системных функций или на конечного пользователя. Напротив, в определении 4 явно указывается, что метаданные могут описывать как цифровые, так и нецифровые информационные ресурсы, и должны быть воспринимаемыми человеком. Определение 5 имеет такой же неконструктивный характер, как и первые два, но его авторы дополнительно отмечают, что метаданные представляют собой структурированные данные. Такая точка зрения высказывается во многих источниках. Однако метаданные могут быть не только структурированными данными. В определениях 6 и 7 метаданные также трактуются как структурированные данные. Но эти определения более конструктивны по сравнению с предыдущими, поскольку в них предлагается некоторая конкретизация функций метаданных. В определении 8 делается акцент на функции описания представления данных. Кроме того, важно, что в нем расширяется «область действия» метаданных. В соответствии с этим определением, описываемым ресурсом могут быть не только собственно данные, но и поддерживающие их системы, используемые в них технологии, пользователи рассматриваемых данных и др. Авторы определения 9

полагают, что метаданные в Вебе – это слабоструктурированные данные. На наш взгляд, однако, эти метаданные могут быть и неструктурированными, например, текстами на естественных языках. В этом определении акцентируется важная функция метаданных – обеспечение интероперабельности в неоднородных средах информационных ресурсов, описываемых метаданными. В определении 10 также отмечается эта функция наряду с другими существенными функциями метаданных – обеспечение организации, идентификации, поиска (определения местоположения) и, вообще, управления описываемыми данными. При этом авторы не ограничиваются представлением о метаданных как о структурированных данных. В определении 11 описываемым ресурсом является элемент данных. Поскольку элемент данных – категория, свойственная структурированным данным, данное определение относится лишь к метаданным, описывающим структурированные данные. Нужно заметить при этом, что метаданные элементов данных – лишь только составная часть метаданных структурированных данных. В рассматриваемом определении не охватываются такие, например, аспекты их описания, как структурные характеристики данных, ограничения целостности, полномочия доступа, семантика данных и другие их свойства. В определении 12 указывается еще одна важная функция метаданных – описание семантики информационных ресурсов, которое является основой для реализации ряда существенных функций поддерживающих их систем. Наконец, важное свойство метаданных отмечается в весьма лаконичном определении 13. Описание свойств данных и других ресурсов любой природы может осуществляться лишь в терминах более абстрактной системы понятий. Так, в системе базы данных описание хранимых данных осуществляется не в терминах значений этих данных, а в таких терминах, как хранимая запись, элемент данных, ключ записи, указатель, тип хранимой записи, ключ индексирования, область памяти, и др., т.е. в системе терминов модели данных среды хранения базы данных. Завершая анализ приведенных определений термина «метаданные», можно констатировать, что они характеризуют разные точки зрения на метаданные, акцентируют отдельные их свойства и/или функции. Некоторые из них являются слишком абстрактными и в силу этого малопродуктивными. Другие относятся к частному виду метаданных и поэтому не носят общего характера. Хотя такие определения предлагаются авторами в конкретном контексте, в ряде случаев их специфика явным образом в определении не оговаривается. В результате приписываемые метаданным свойства или функции либо категория описываемых ресурсов неправомерно обобщаются.

3 Ресурсы, описываемые метаданными при рождении термина «метаданные» имелось в виду, что они предназначены для описания свойств данных. По прошествии времени разнообразие ресурсов, описание свойств которых стало необходимым в современных компьютерных системах, существенно расширилось. Помимо собственно данных, требуются описания систем, обеспечивающих их хранение, доступ к ним и их обработку, описания характеристик пользователей (их полномочий доступа, их профиля, определяющего, в частности, характер их информационных потребностей и

т.п.). Возникли также потребности описания концептуальных схем и онтологий, разнообразных сервисов, средств вычислительной техники, научных приборов и других технических устройств, потоков работ, бизнес-процессов, а также организаций – владельцев, генераторов и пользователей информационных ресурсов. Нужны описания не только структурированных, но и неструктурированных данных. С помощью метаданных должны описываться различного рода ограничения, которым должны удовлетворять данные, доступ к ним и их использование. Метаданные в бизнесе описывают специфические для этой сферы деятельности ограничения, называемые бизнес-правилами. Бизнес-правила характеризуют «правила игры» в бизнесе, своего рода руководство для бизнес-деятельности. Следует упомянуть также использование метаданных для описания разрабатываемых программных систем и процессов их функционирования. Некоторые такие метаданные формируются CASE-инструментами и представляются диаграммами на языке UML. Также указывают авторы об использовании метаданных для описания программ, их фрагментов, различного рода процессов и их компонентов, сред функционирования систем, совокупностей событий и отдельных событий, уже отмечавшихся ранее ограничений, людей и их ролей в ИТ-системах, организаций, их подразделений, индивидуумов или индивидуумов в определенной роли. Перечисленными примерами отнюдь не исчерпывается множество видов ресурсов, описываемых метаданными. Поэтому наряду с метаданными информационных ресурсов будем обсуждать далее и метаданные ресурсов произвольного вида.

4 Свойства метаданных

Конкретный состав функций метаданных зависит от особенностей использующей их системы, от характера описываемых метаданными ее ресурсов, от базовых информационных технологий системы, потребностей ее пользователей и от многих других факторов. В этом разделе рассматриваются общие свойства метаданных, не специфичные для какого-либо конкретного случая их применения или для какого-либо вида описываемых ресурсов. Относительный характер метаданных. Разделение информационных ресурсов на данные и метаданные является относительным. Информационные ресурсы, играющие роль метаданных в одних случаях, являются данными в других, и наоборот. Действительно, совокупность тегов гипертекстовой разметки веб-страницы, название публикации или ее аннотация, содержащаяся в ее тексте, могут использоваться в качестве метаданных, описывающих перечисленные ресурсы. Вместе с тем, все они являются составными частями соответствующих ресурсов и тем самым являются данными. Многоуровневость метаданных. Как уже отмечалось, что описание свойств данных и любого другого ресурса может осуществляться только в терминах более абстрактной системы понятий. Такие системы понятий и их связей воплощают модели более абстрактного уровня (метамодели) по отношению к моделям, в терминах которых представлены описываемые ресурсы. При этом метаданные, как и данные, являются информационными ресурсами. Для них также могут существовать метаданные, представляющие собой метаметаданные по отношению к исходным описываемым ресурсам. Такая

иерархия метаданных может включать любое число уровней. Так, стандарт Meta Object Facility (MOF) [33] консорциума OMG предусматривает три уровня, спецификация стандартов Дублинского ядра – два уровня: предназначенный непосредственно для практических целей и обеспечивающий его описание. Спецификации более высоких уровней иерархии метаданных обозначаются терминами «метаметаданные», «мета-мета-метаданные» и т.д. Ресурсы, описываемые метаданными. Как уже отмечалось, метаданные могут использоваться для описания свойств не только данных, но и многих других ресурсов различных видов.

Автономные/встроенные метаданные. Метаданные могут быть автономными (отчужденными) от описываемого ресурса или встроенными в него. Примеры автономных метаданных: схемы баз данных (они отчуждены от описываемого ими наполнения базы данных); DTD, описывающие тип XML-документов и хранимые в Вебе вне конкретных экземпляров XML-документов; репозиторий метаданных открытого архива в технологии открытых архивов OAI; полнотекстовый индекс коллекции текстовых документов электронной библиотеки. Примеры встроенных метаданных: HTML-разметка веб-страницы; разметка веб-страницы средствами микроформатов; DTD, специфицированное в теле XML-документа; аннотация статьи, содержащаяся в ее тексте; глоссарий, содержащийся в тексте спецификации официального стандарта. Если неструктурированные метаданные встроены в текстовый документ, то для компьютерной обработки необходимо извлечь их из текста. Методам решения этой сложной задачи посвящен в последнее время ряд исследований. Информационные ресурсы со встроенными описывающими их метаданными естественно называть самоописываемыми. Метаданные независимые и зависимые от контента описываемых информационных ресурсов. Различаются метаданные, независимые от контента описываемого информационного ресурса и зависимые от контента. Примерами метаданных, независимых от контента, могут служить: дата создания и место хранения текстового документа; тип файла, в котором он содержится; URL веб-страницы. Примеры метаданных, зависимых от контента описываемого ресурса: аннотация текста; его статистические характеристики (частотные характеристики вхождений слов словаря, длина текста в литерках и т.п.). Некоторые из таких метаданных могут продуцироваться автоматически из текста данного документа, другие формируются вручную. Метаданные независимые и зависимые от предметной области. Метаданные первой из этих категорий обычно описываются выразительными средствами стандартов горизонтальной сферы (стандартов, не ориентированных на какую-либо конкретную область применения). Метаданные второй категории характеризуются большим разнообразием. Существуют собственные системы метаданных в различных областях научных исследований, в музейном деле, в образовательной сфере и т.д. Системные и пользовательские метаданные. Метаданные могут быть предназначены для компьютерной обработки, связанной с исполнением системных функций (системные метаданные). Они могут быть также предназначены для пользователей информационной системы. Метаданные,

предназначенные для пользователей, могут служить, напри мер, для информирования пользователей о свойствах информационных ресурсов, свойствах системных сервисов, контекста использования системы или предоставляют пользователю инструкцию о порядке выполнения требуемых ему операций. Существуют системы, служащие для поддержки всей совокупности метаданных информационной системы, называемые системами словарей-справочников данных. Их компонент справочник данных управляет метаданными, предназначенными для системы, а словарь данных – метаданными, предназначенными для пользователей. Структурированные и неструктурированные метаданные. Для поддержки системных функций, как правило, используются структурированные метаданные. Метаданные, предназначенные для пользователей, могут быть как структурированными, так и неструктурированными, например, они могут быть обычными текстами. В отличие от структурированных метаданных, для неструктурированных метаданных не существует каких-либо общепринятых стандартов их представления. Структурированные метаданные привлекают большое внимание. Им посвящена масса публикаций. Структурированные метаданные, так же как и неструктурированные, могут описывать и структурированные, и не структурированные данные. Уровень гранулярности описания ресурсов. Метаданные могут описывать информационные и другие ресурсы на любом требуемом уровне их гранулярности. Например, в электронных библиотеках используются метаданные, описывающие коллекции информационных ресурсов в целом, а также отдельные входящие в них информационные объекты. В системах баз данных схема базы данных описывает не только структуру базы данных в целом, но и свойства ее структурных компонентов – в реляционных базах данных отдельных таблиц, данных, содержащихся в их столбцах и т.д. Статические и динамические метаданные. Метаданные могут быть как статическими, так и динамическими. Так, в системах баз данных относительно неизменна схема базы данных. В то же время каталоги коллекций в электронной библиотеке изменяются при включении в них новых информационных объектов. Степень формализованности метаданных. Для представления метаданных могут служить в различной степени формализованные выразительные средства – естественные языки, полуформальные языки, например, набор элементов метаданных Дублинского ядра, а также формальные языки в дедуктивных базах данных интенционал описывается средствами логики. Язык OWL для описания онтологий Веба основан на дескриптивной логике. Зависимость метаданных от информационной архитектуры системы. Состав и содержание метаданных зависит от того уровня информационной архитектуры информационной системы, которому принадлежит ресурс, описываемый метаданными. Действительно, метаданные, описывающие веб-страницы на уровне хранения, представляют собой совокупность тегов гипертекстовой разметки, которую можно видеть в исходном коде страницы. В то же время, на «логическом» уровне, в том виде, как браузер воспроизводит эту страницу на экране компьютера, с этой страницей ассоциируются совсем иные

метаданные, например, ее URL и описание ее контента в различного рода каталогах для пользователей Веба. В системах баз данных метаданные среды хранения описывают структуру хранимых данных и ее свойства, индексы, характеристики пространства памяти и другие элементы среды хранения. В то же время, метаданные на «логическом» уровне представления данных описывают их видение пользователем. В реляционной системе базы данных метаданные описывают состав таблиц базы данных, свойства каждой таблицы, связи между таблицами, полномочия пользователей и т.п., но не организацию хранения базы данных. Явное и неявное представление метаданных. Многие свойства различных ресурсов информационных и других систем описываются явным образом представленными метаданными. Эти метаданные поддерживаются в форме текстовых документов, иллюстраций или изображений, сообщений по электронной почте, в электронных таблицах, в базах данных и в различных ИТ-инструментах, например, в репозиториях метаданных CASE-инструментов или открытых архивов. Вместе с тем, некоторые их свойства могут быть не представлены в системах явными описаниями. Такая ситуация имеет место, например, в научных публикациях. Для содержащихся в них ссылок на использованные источники не указывается семантика этих связей.

Для представления метаданных могут использоваться разнообразные выразительные *средства*: естественные и искусственные языки, наборы элементов метаданных (схемы метаданных), графические схемы и диаграммы, а также другие средства. Рассмотрены только основные свойства метаданных, которые не зависят от предметной области, в которой они применяются. Как уже отмечалось, разнообразие метаданных, используемых в конкретных областях применения, чрезвычайно велико. Во многих направлениях научных исследований существуют собственные, специфичные для них системы метаданных. Составляющие их виды метаданных позволяют описывать многообразие свойств исследуемых объектов, явлений, процессов. Например, это характеристики измерительных приборов, обеспечивающих получение научных данных, разнообразных компьютерных и иных моделей, в соответствии с которыми вычислены или оценены значения используемых данных, характеристики публикаций, источников информационных ресурсов и многое другое.

Метаданные выполняют разнообразные *функции* в системах, их использующих. Конкретные их функции и состав существенным образом зависят от информационных технологий, на которых базируется система, от ее функциональности, свойств поддерживаемых в ней информационных ресурсов, способов их организации в системе, особенностей задач их обработки и от многих других факторов. Рассмотрим кратко основные функции, выполняемые метаданными, не исчерпывающие всех их возможностей. Приведенные ниже примеры выразительных средств представления метаданных, выполняющих обсуждаемые функции, относятся к числу «универсальных», не специфичных для какой-либо конкретной предметной области или деятельности какого-либо профессионального сообщества. Описание моделей предметной области. Важнейшая функция

метаданных – описание абстрактных моделей предметной области. С разработки этих моделей начинается проектирование информационных систем. Это – концептуальная схема предметной области, дополняемая в системах семантического уровня онтологией. Такие модели не зависят от технологий реализации разрабатываемой информационной системы. Для спецификации концептуальной схемы используются языки концептуального моделирования различного уровня формальности, например, диаграммы UML или формальные языки представления знаний, основанные на математической логике. Тем самым концептуальная схема может в различной степени удерживать семантику предметной области. Онтологии также могут представляться различными средствами – от классификаторов до формальных языков описания онтологий. В последние годы языки описания онтологий OWL и OWL2 консорциума W3C стали иногда использоваться для описания концептуальной схемы предметной области.

Описание источников данных. Эти функции метаданных выполняют, в частности, схемы баз данных, определения типа XML-документов, XML-схемы для репозиториях информационных объектов, онтологии источников информации. Метаданные этого вида представляются, соответственно, языками описания данных (например, дескриптивный подязык языка SQL, язык описания объектов ODL объектных баз данных), подязык языка XML и язык XML-схем, языки описания онтологий (например, RDFS, OWL2).

Описание информационных объектов. Эта функция необходима в электронных библиотеках, системах открытых архивов и других репозиториях информационных объектов с гранулярностью доступа к их информационным ресурсам на уровне отдельного объекта. В последнее время для представления таких метаданных часто используются описатели, представленные в форме XML-документов, библиотечных каталожных записей различных диалектов стандарта MARC. Эти метаданные описывают структурные свойства описываемых информационных объектов, различные другие их характеристики. В «интеллектуальных» системах описывается семантика информационных объектов (см. далее). Описание семантики информационных источников, информационных объектов или их фрагментов. Метаданные, используемые для этой цели, называются семантическими. В простейших случаях семантика информационных источников или отдельных информационных объектов описывается с помощью рубрикаторов, например, российских ГРНТИ или ББК, либо международных УДК или JEL (Journal of Economic Literature Classification System). Другим более многоаспектным, но неформальным средством описания семантики является набор элементов метаданных Дублинского ядра (Dublin Core, DC). Для текстовых документов в электронных библиотеках и коллекций информационных ресурсов электронных библиотек, для структурных элементов баз данных, для веб-страниц, а также для электронных библиотек или других информационных репозиториях в целом используют также семантическое аннотирование, заключающееся в ассоциировании с описываемым ресурсом некоторой информации (семантической аннотации), характеризующей его семантику. Семантическая

аннотация может быть встроена в описываемый ресурс или автономна от него. Одним из простейших средств встроеного семантического аннотирования, используемого для описания семантики фрагментов веб-страниц, являются так называемые микроформаты. Некоторые из них поддерживаются популярными поисковыми машинами Веба. Например, Яндекс поддерживает микроформаты: hCard – для разметки контактной информации (адресов, телефонов и т. д.); hRecipe – для описания кулинарных рецептов; hReview – для разметки рецензий, отзывов; hProduct – для разметки описаний товаров. Для семантического аннотирования могут использоваться различные средства – от естественных до формальных искусственных языков. Могут использоваться, например, управляемые словари и тезаурусы, таксономии и формальные онтологии. Частным случаем семантического аннотирования является онтологическое аннотирование. В этом случае описание семантики ресурса осуществляется в терминах некоторой онтологии. Онтологии, как и другие средства представления знаний, используются также для автономного формального описания семантики информационных источников. В настоящее время для этого чаще всего служат разработанные консорциумом W3C языки RDF, RDFS, OWL, OWL2 и его профили. Следует отметить, что вопреки мнению, высказанному в ряде других публикаций, метаданные могут описывать семантику не только структурированных, но и неструктурированных данных. В указанной работе утверждается: «Структурированные данные обладают метаданными, которые позволяют нам понять, что они означают. Для неструктурированных данных не существует какого-либо эквивалента». Это утверждение ошибочно хотя бы потому, что для любых информационных ресурсов, структурированных и неструктурированных, могут использоваться семантические аннотации. Описание представления данных на разных уровнях информационной архитектуры. В системах баз данных метаданные описывают представление хранимых данных (внутренняя схема базы данных), «логическое» представление полной базы данных (концептуальная схема базы данных) и представления для пользователей или приложений необходимых им фрагментов базы данных или полной базы данных (внешние схемы). Такие метаданные определяются на языках описания данных СУБД. В среде Веба разметка хранимых HTML-файлов описывает, каким образом веб-браузер должен формировать представление веб-страниц на экране пользовательского компьютера и какие другие файлы из Интернет нужно для этого использовать. Идентификация описываемых ресурсов. Метаданные описывают, какие встроенные в экземпляры ресурсов их элементы или ассоциированные с ними атрибуты обеспечивают их идентификацию. Примерами встроенных идентифицирующих элементов являются первичные и вторичные ключи строк таблиц в системах баз данных. Их значения уникальным образом идентифицируют отдельные строки или, соответственно, некоторые их множества. XML-документы при наличии DTD или XML-схемы идентифицируются их встроенными идентификаторами. В электронных библиотеках текстовых документов могут использоваться их уникальные идентификаторы. Часто они генерируются

системой, недоступны в явном виде пользователям, и служат для ссылок на эти документы в генерируемых системой каталогах. Для идентификации текстовых документов могут также использоваться наборы значений дескрипторов или ключевых слов. В Вебе в качестве уникальных идентификаторов веб-страниц используются указатели URL или URI. Первый из них идентифицирует страницу по ее местоположению в Интернет. В ГИС точки земной поверхности идентифицируются их широтой и долготой. Наконец, нужно упомянуть индексы УДК и ISBN, идентификаторы цифровых объектов DOI, элементы метаданных Дублинского ядра. Существуют и другие примеры видов метаданных, используемых для идентификации информационных ресурсов. Управление структурированными данными. Это комплекс задач, связанных с обеспечением эффективного хранения и обработки структурированных данных, а также доступа к ним. Чаще всего для решения этих задач используют технологии баз данных. Метаданные, представленные в виде схемы баз данных, обеспечивают механизмы СУБД в процессе ее функционирования необходимой информацией об организации базы данных, ограничениях целостности данных и управления доступом. Верификация данных. Метаданные структурированных данных (например, схема базы данных) позволяют контролировать их форматы и типы данных, проверять соблюдение ограничений целостности. Эти задачи решают механизмы СУБД. Для XML-документов с заданным определением типа документов (DTD) используемый XML-процессор может проверить корректность их структуры. Онтология SQL-системы базы данных в системах доступа к данным, основанная на онтологиях, позволяет обнаруживать неполноту и противоречивость данных в базе данных. Предоставление пользователям сведений об описываемых ресурсах. Метаданные могут быть предназначены не только для поддержки системных функций, но и для информирования пользователей о свойствах информационных ресурсов, сервисов, которые обеспечивают доступ к ним и о других описываемых ресурсах. Ориентированные на пользователя метаданные могут быть как структурированными, так и неструктурированными данными. Последние могут представляться в виде текстов на естественных языках, схем, чертежей, фрагментов карт местности и др. Метаданные рассматриваемого назначения могут описывать генезис описываемых ресурсов, связанные с ними предположения, способы доступа и условия их использования, их семантику, используемую терминологию, а также иные сведения. Примерами таких метаданных являются инструкции для пользователей по работе с системой, оглавление или аннотация книги. Управление доступом. Метаданные позволяют описывать ограничения доступа и полномочия пользователей на доступ к информационным ресурсам или на выполнение иных операций используемой системы. Эти полномочия и ограничения проверяются системными механизмами при попытке пользователя выполнить защищенную операцию. В системах баз данных ограничения доступа описываются в схеме базы данных. Описание характеристик пользователей и их информационных потребностей. Во многих

информационных системах имеется возможность поддерживать сведения о пользователях и их информационных потребностях, необходимые для учета их работы и сбора статистических данных об использовании системных ресурсов. Совокупность таких метаданных называется профилем пользователя. Обеспечение обнаружения и поиска информационных ресурсов. Это — одна из главных функций метаданных, предусматривающая использование их в критериях поиска. При этом могут использоваться не только идентифицирующие, но и другие виды метаданных. Обнаружение требуемых информационных ресурсов может также осуществляться путем навигации в каталогах, предметных указателях, рубрикаторах. Использование семантических метаданных в процессе поиска позволяет существенно снизить уровень информационного шума. Распространение информационных ресурсов. Поддерживаемые в информационной системе метаданные (например, в профилях пользователей), характеризующие информационные потребности пользователей, позволяют обеспечить избирательное распространение информации средствами системы. Тематическая систематизация информационных ресурсов. Эта функция выполняется путем формирования каталогов и предметных указателей содержащихся в системе информационных ресурсов, индексирования ресурсов с помощью рубрикаторов. Обеспечение анализа данных. Эта функция характерна для систем многомерных или реляционных баз данных и хранилищ данных и связана с использованием методов OLAP. Используемые метаданные содержатся в схемах баз данных. Обеспечение интероперабельности и повторного использования информационных ресурсов. Техническая и/или семантическая интероперабельность и на этой основе повторное использование информационных ресурсов может обеспечиваться путем унификации представления метаданных средствами открытых стандартов. Например, база данных, описанная схемой, представленной на дескриптивном подязыке языка SQL или на языке описания объектов ODL стандарта ODMG, может повторно использоваться другой СУБД, поддерживающей стандарт SQL или, соответственно, ODMG. Интеграция данных из многих источников. Рассмотрим эту функцию на примере создания и функционирования системы виртуальной интеграции информационных ресурсов из множества источников. Здесь необходимы метаданные, которые описывают: локальные схемы интегрируемых источников, глобальную схему, отображения между локальными схемами и глобальной схемой. В системах семантической интеграции данных необходимы также метаданные, включающие онтологии локальных источников, общую онтологию системы интеграции, а также отображения онтологий. В системах виртуальной интеграции данных с архитектурой адаптеров-посредника необходимы также метаданные, которые описывают регистрационные данные источников в посреднике — центральном компоненте архитектуры систем интеграции данных. Описание авторских прав на интеллектуальную собственность. Эту функцию поддерживают метаданные, предоставляющие информацию об авторских правах на описываемые ресурсы. Наукометрия в научных электронных библиотеках.

Идентифицирующие метаданные (автор, организация, название работы), описания рубрик классификаторов и ссылок цитирования в описаниях научных публикаций в электронных библиотеках позволяют оценивать цитируемость работ отдельных авторов или организаций, в которых они работают, анализировать тематическую структуру контента библиотеки, формировать рейтинги востребованности и цитируемости работ по авторам и по отдельным публикациям. Благодаря классификации связей между публикациями, авторами и публикациями, организациями и авторами, и осуществляемой на ее основе семантической структуризации контента научных электронных библиотек с поддержкой метаданных, описывающих сформированную семантическую структуру, можно получать многоаспектную наукометрическую и науковедческую информацию. Рассмотренные в этом разделе функции метаданных используются наиболее часто. Однако в некоторых компьютерных системах они могут выполнять и иные специфические функции, обсуждение которых не входит в задачу этой работы.

Возможны различные подходы к *классификации* метаданных. Имеют право на существование классификации метаданных по их функциям, по уровням семантической абстрактности, по их свойствам, по уровням информационной архитектуры, к которым относятся описываемые ресурсы, а также по ряду других критериев. Ранее были рассмотрены детальные классификации метаданных, независимых от предметной области применения, по их функциям, а также по их свойствам. В литературе чаще всего предлагаются агрегированные функциональные классификации. Наиболее популярна классификация, в которой различаются описательные, структурные и административные метаданные. Описательные метаданные описывают контент ресурса (например, это – набор значений элементов метаданных Дублинского ядра), библиографические его данные (если это – публикация), аннотацию, идентификаторы ресурса (например, URI или DOI) и т. п. Структурные метаданные характеризуют общую структуру ресурса и ее компоненты, объем и другие подобные свойства описываемого ресурса. Административные метаданные описывают даты создания и обновления ресурса, кем он создан или модифицирован, кто владелец прав на этот ресурс, полномочия доступа пользователей, сведения об имеющихся версиях и/или копиях ресурса, где они хранятся и другие сведения, необходимые для администрирования и управления ресурсом. При этом полагают, что административные метаданные состоят из нескольких подмножеств, два из которых иногда рассматриваются как самостоятельные классы метаданных. Это метаданные управления правами на интеллектуальную собственность, а также метаданные, описывающие различные аспекты деятельности, связанной с сохранением ресурсов. В материалах, посвященных созданию цифровых архивов, предлагается та же классификация. Однако в состав административных метаданных авторы включают технические метаданные (описание ИТ-аспектов информационных ресурсов и используемых информационных технологий), метаданные управления правами (декларации прав на ресурс и ограничения доступа к нему), а также метаданные

цифрового происхождения (сведения о создании и последующей обработке цифрового ресурса, в том числе, подробности об ответственности за каждое событие на протяжении его жизни). Предлагают и несколько более расширенную классификация, предусматривающую следующие классы метаданных.

Административные метаданные. Служат для управления и администрирования электронными коллекциями и другими информационными ресурсами. Предоставляют информацию о сборе данных, их местоположении, правах и способах репродуцирования ресурсов, условиях легального доступа.

Описательные метаданные. Используются для идентификации и описания свойств коллекций и связанных с ними информационных ресурсов. К таким метаданным относятся каталожные записи, информация для помощи при поиске, описания версий ресурсов, специализированные индексы, кураторская информация, гипертекстовые связи между ресурсами, аннотации создателей и пользователей.

Метаданные сохранения. Предназначены для управления сохранением коллекций и других информационных ресурсов. К этому классу метаданных относятся документация о физических условиях сохранения ресурсов, о предпринятых действиях для физического сохранения, о цифровых версиях ресурсов, об изменениях, которые были произведены во время их оцифровки и сохранения.

Технические метаданные. Включают, в частности, описание функционирования систем, документацию по оборудованию и программному обеспечению, технологическую информацию о проведенной оцифровке, данные аутентикации и обеспечения информационной безопасности (ключи шифрования, пароли).

Метаданные об использовании. Описывают характер использования ресурсов, в том числе, повторного использования, версии ресурсов, авторские права на ресурсы, включают журнал регистрации поиска и др.

Предпринимались также попытки создания обобщенной классификации *научных метаданных*. Например, классификацию, которая включает следующие классы научных метаданных:

Описание трансформаций данных (описание научных моделей, компьютерных программ и др.)

Описание семантики данных

Описание структуры данных

Описание атрибутов данных

Описание значений данных.

Эта классификация, к сожалению, является довольно узкой. Она не включает метаданные, описывающие научные приборы, процессы исследовательской деятельности, ее инфраструктурные и другие ресурсы, связанные с наукой и научной деятельностью.

Для представления метаданных могут использоваться *средства представления метаданных*: естественные языки, а также различные искусственные языки, такие как языки разметки, схемы метаданных,

графические средства. Рассмотрим кратко указанные категории выразительных средств.

Естественные языки. Такие языки обладают наиболее богатыми выразительными возможностями по сравнению с другими средствами представления метаданных. Однако они не обеспечивают однозначности и строгости интерпретации метаданных, представленных их средствами. Такие неструктурированные метаданные в малой степени пригодны для компьютерной обработки и предназначены, главным образом, для пользователей. На естественных языках представляются, например, аннотации публикаций, различные сведения об описываемых ресурсах и их авторах.

Искусственные языки. Это большой пласт разнообразных языков. К их числу относятся языки описания данных СУБД, концептуального моделирования, описания онтологий, бизнес-процессов, потоков работ. К этой категории относятся, например, дескриптивный подязык языка SQL, язык описания объектов ODL, язык описания интерфейсов IDL стандарта CORBA, языки консорциума W3C: OWL, OWL2, RDF, RDFS, язык XML-схем. Сюда же следует отнести языки разметки, схемы метаданных, визуальные языки, которые рассматриваются ниже, и многие другие языки в различных областях информационных технологий. Языки разметки. Среди этих искусственных языков наиболее популярны XML, HTML, XHTML, Tex (язык разметки математических текстов) и его оболочка LaTeX. Во многих системах метаданные представляются в виде XML-документов с определенной XML-схемой.

Схемы метаданных. Средства этой категории представляют собой особый вид искусственных языков, который стал в последние годы чрезвычайно популярным. Прототипом для их разработок в значительной мере стал набор элементов метаданных Дублинского ядра [12, 17, 24] – стандарта структурированных метаданных с нетипизированными значениями элементов. Попытки типизации значений элементов метаданных Дублинского ядра привели к дополнению в его спецификацию описателей типов для значений некоторых его элементов, называемых квалификаторами. Термин «схема метаданных» широко используется в литературе и, по сути, является синонимом термина «набор элементов метаданных». Схема метаданных – это набор элементов метаданных, каждый из которых обладает некоторым именем и семантикой, принимает значения с установленной семантикой, иногда – значения из управляемого словаря. Существуют схемы метаданных, созданные различными профессиональными сообществами для многих сфер применения. Метаданные описываемых ресурсов, выраженные средствами схем метаданных, часто кодируются в виде XML-документов. Визуальные языки. Визуальные языки давно начали использоваться в CASE-технологиях. Наиболее популярным в этой области в настоящее время стал язык объектного анализа и проектирования UML [42]. Основным средством представления метаданных в этом языке являются графические диаграммы нескольких типов, которые на стадии проектирования позволяют описать различные аспекты создаваемой сложной программной системы.

Для обеспечения интероперабельности и повторного использования метаданных важное значение имеет *стандартизация* средств их представления. Деятельность в этой области активно проводится с давнего времени официальными органами стандартизации, промышленными консорциумами, различными профессиональными сообществами. Разработано большое число стандартов метаданных как независимых от сферы применения, так и предназначенных для специфических применений. К первой группе относятся, например, дескриптивный подязык языка SQL, язык описания объектов ODL консорциума ODMG, стандарты консорциума OMG: язык UML, язык описания интерфейсов CORBA IDL, Common Warehouse Model (CMW); стандарты консорциума W3C: XML Schema, RDF, RDFS, OWL, OWL2 с его профилями и язык описания интерфейсов веб-сервисов WSDL; набор элементов Дублинского ядра (Dublin Core, DC), поддерживаемый Директоратом Дублинского ядра; язык моделирования бизнес-процессов BPMN и другие. Значительное место занимают многочисленные стандарты научных метаданных, созданные во многих областях исследований.

Метаданные являются важнейшим видом информационных ресурсов. Сущность метаданных, их свойства и возможные функции пока еще не осознаны многими специалистами в достаточной мере. Множество определений термина «метаданные» существует не случайно. Причина в том, что довольно сложно дать качественное и лаконичное определение этого термина. Проблематика метаданных в компьютерных системах является непреходящей. В последние годы все более активно используются семантические метаданные. Разработки новых технологий и новых видов приложений по необходимости потребуют создания и использования новых видов метаданных. В связи с этим будут создаваться и новые выразительные средства для их представления. Однако рассмотренные свойства и функции метаданных будут присущи и вновь создаваемым их видам.

В «Техническом руководстве по программам создания цифрового контента по культуре» (Technical Guidelines for Digital Cultural Content Creation Programmes) проекта MINERVA говорится, что метаданные можно определить буквально как «данные о данных», но под этим термином понимают, в основном, структурированные данные о ресурсах, которые можно использовать для поддержки разнообразных действий с этими ресурсами. Если создавать веб-приложение с соблюдением соответствующих стандартов, то можно облегчить его взаимодействие и совместимость с другими веб-сайтами, поисковыми машинами и порталами. Ресурсом может быть что угодно, то, что имеет идентичность, и ресурс может быть как цифровым, так и не цифровым. Операции с цифровыми ресурсами могут включать такие действия как раскрытие, обнаружение, управление (включая управление правами) и долгосрочное хранение. Единичный ресурс может потребовать различных метаданных для поддержки этих функций. Если же информация, публикуемая в Сети, будет систематизирована и представлена в более или менее структурированной форме, то поиск может быть более эффективным.

Метаданные можно использовать для решения следующих задач:

- облегчения поиска контента по большому числу агентов (порталов, агрегаторов, поисковых машин);
- продвижения материалов с помощью агрегаторов;
- повышения визуализации и понимания имеющихся в организации ресурсов;
- предоставления потенциальным пользователям возможности определять релевантность ресурсов до получения доступа к ним;
- стимулирования создания совместимых услуг;
- повышения визуализации контента в поисковых машинах, таких как Google, Google Scholar, Yahoo и др.;
- увеличения трафика и объема операций веб-сайта;
- представления ресурсов новым рынкам, чтобы интеллектуальные агенты пользователей могли анализировать информацию и с большей отдачей использовать ресурсы.

Преимущества использования метаданных равно действительны как для поставщиков контента, предоставляющих ресурсы на бесплатной основе, так и для поставщиков коммерческой информации и информации ограниченного доступа.

Берестова Т.Ф. также отмечает, что информационная природа метаданных не вызывает сомнения. Они возникли как явление информатизации и функционируют чаще всего как машинная информация, которая имеет много общего с семантической. Наиболее удачно понятие информация дефинировал А. В. Соколов: «Семантическая информация – это амбивалентное понятие, обозначающее средство для выражения смысла через коммуникативные знаки». Общность семантической и машинной информации основана на том, что любая социальная информация существует только в социуме (человеческом обществе), и она выражает смыслы. Семантическая информация – это источник знаний, смыслы, представленные в виде человекочитаемой информации. Машинная информация также выражает смыслы, но через машиночитаемые знаки. Смыслы машиночитаемой информации связаны с командами, а передаются они в виде электрических импульсов по алгоритмическим программам и воспринимаются компьютером, если он снабжен соответствующим программным обеспечением. «Структура семантической и машинной коммуникации изоморфны (подобны по форме): в обоих случаях передаваемая информация амбивалентна и включает смысловое (идеальное) содержание (допустим человекочитаемые указания или машиночитаемые команды) и материальную форму (текст, изображение, закодированное сообщение)». Большинство исследователей метаданных признает их разновидностью информационных ресурсов. Определение последних уже предложено нами, и в данной статье мы только повторим его: «Информационный ресурс – это целенаправленно созданная информация, обладающая инструментальной функцией и используемая в качестве средства определенного вида деятельности, она обладает ярко выраженными потребительскими ценностями, которые зачастую позволяют

позиционировать этот продукт в качестве товара на информационном рынке». Феномен метаданные возникает как явление и понятие, производное от явления информационные ресурсы, а потому обладает всеми характеристиками, перечисленными в дефиниции информационных ресурсов. К сожалению, вопрос об объеме понятия информационные ресурсы не имеет общепринятого решения, и это обуславливает подходы к широкому или узкому пониманию термина метаданные. Если информационные ресурсы – это общее явление и понятие, включающее в себя и традиционные (нецифровые), и машинные (цифровые) ресурсы, то метаданные – термин для обозначения вторичной информации о любых информационных ресурсах. Если признается, что информационные ресурсы существуют исключительно в электронной среде, то метаданные – разновидность вторичной информации, которая может быть только машиночитаема.

Снятие противоречия, отражающего разное понимание явления метаданные, связано с разрешением вопроса об объеме понятия информационные ресурсы. При дальнейшем изучении метаданных мы будем опираться на узкое понимание данного термина, обозначающего феномен машиночитаемых информационных ресурсов.

В работах о метаданных нередко представлены позиции, которые определяют этот феномен беспредельно широко, и часто наряду с электронными феноменами в число метаданных включаются явления, зародившиеся и функционирующие в рамках традиционной (печатной) коммуникации. Так, М. Р. Когаловский считает, что метаданные существовали задолго до появления компьютерных систем, и в качестве примера таковых он называет библиографические описания, различные каталоги, классификации и аннотации. Он относит к метаданным также: информационные потребности пользователей, характеристики владельцев генераторов и ИПС, слабоструктурированные данные Веба и вообще неструктурированные данные, в частности тексты на естественных языках, совокупность разметки тегов, необходимых в гипертекстовых системах, поисковый образ документа, описания схемы баз и банков данных, концептуальных схем и онтологий, научных приборов, технических сервисов, устройств программных систем и их элементов, процессов и сред их функционирования, бизнес-процессов и бизнес-правил (полномочия и ограничения в доступе и использовании). Исходя из такого подхода, к метаданным можно отнести и нецифровые объекты, а значительная часть подобной информации, как известно, воспринимается человеком. При этом упомянутый автор отмечает: если метаданные присутствуют в тексте, изображениях, сообщениях электронной почты и других объектах в скрытом, неявном виде, то их приходится извлекать вручную, но при создании особых программ даже в неструктурированных ресурсах метаданные изымаются из текста автоматически.

Столь широкий подход к использованию термина метаданные кажется нам неправомерным. Ведь явление метаданные зародилось только после возникновения информационно-коммуникационных технологий, только во времена информатизации, термин метаданные сформировался

применительно к машиночитаемой информации – это обстоятельство, которое надо учитывать при проведении познавательных процедур по отношению к феномену метаданные. Признав метаданные вторично-информационным ресурсом, нет резона при этом включать в их число уже давно известные и неплохо изученные явления вторичной информации, уже обозначенные специальными терминами и закрепленные в профессиональной лексике у очень широкого круга специалистов и ученых документально-коммуникационной сферы. Разновидности вторичной информации давно выявлены и названы, это – библиографическая и вторично-семантическая (аналитико-синтетически переработанная) информация, а значит, нет смысла для их обозначения использовать термин метаданные, появившийся значительно позднее и обозначающий явление, порожденное автоматизацией информационных процессов. Если же метаданные – еще один вид вторичной информации, то надо выявить его общность с другими вторично-информационными объектами и указать на различия с ними.

Очень часто представление метаданных осуществляют через перечисление технологических операций, которые производят с их помощью. При подобной перечислительной конкретизации метаданными называют любую информацию, необходимую для анализа, проектирования, построения, внедрения и применения компьютерной системы при использовании информационных технологий. В ряде публикаций говорится, что метаданные упрощают и рационализируют управление ресурсами, могут быть инструментом генерирования и управления записями, дают правила взаимосвязи внутри информационных систем, они обеспечивают полноценное использование и понимание данных и создание запросов, метаданные выступают как технические стандарты, информируют об условиях долговременной сохранности отражаемых объектов и включают различные их характеристики: идентификацию, программно-аппаратное обеспечение, описание физических свойств объекта, формат файла, номер версии, информацию о миграции и возможных потерях во время ее процесса, данные на аутентичность, информацию об авторских правах, стоимости доставки.

В приведенных примерах под термином метаданные, в одном ряду со вторичными информационными явлениями, широко представлены явления первичной информации, например, информация об авторских правах, условиях сохранности, правила взаимосвязи внутри информационных систем, а также данные о технических устройствах, обеспечивающих функционирование метаданных и других информационных ресурсов. Иногда даже стандарты метаданных фактически нельзя отнести к вторично-информационным явлениям, например, стандарт для идентификации организаций (ISO 6523) в ряд ли можно назвать вторичным информационным ресурсом.

Одним из способов познания разнообразия тех или иных феноменов является классифицирование. В настоящее время есть ряд классификаций метаданных на основе использования таких признаков, как уровни

семантической абстракции и архитектуры, свойств и выполняемых функций и т. д. Для данной публикации наиболее продуктивно использовать разделение метаданных на описательные (контентные), структурные и административные. В административные метаданные включены несколько разновидностей: технические метаданные, данные, используемые в управлении сохранением и обработкой ресурса, а также в управлении правами интеллектуальных и иных собственников.

Структурные метаданные включают описание объемов, структуры ресурса и его компонентов.

Структурные, и особенно административные, метаданные пока слабо дифференцированы, а группы их разновидностей не имеют собственного «словесного» обозначения, именно в них нет четкого разделения на объекты вторично информационной природы и объекты, обладающие признаками первичной (фактографической) информации. Третья группа, так называемые описательные метаданные, характеризует информационный ресурс (т. е. контент) по структуре или по предметным областям, целям и другим качествам, с помощью этих метаданных ресурс может быть идентифицирован и аутентифицирован.

Название описательные метаданные нам кажется неудачным, потому что и в других группах метаданных также используется описание. Описание – это технологическая процедура, общая и для административных, и для структурных метаданных. Вопрос о характеристиках структурных и административных метаданных требует особого изучения, но самые предварительные суждения можно высказать уже в данной статье. Считаем, что описывать информационный ресурс можно и по признакам, связанным с его структурой, а также по возможностям защиты, сохранения, движения в инфосистеме и т. д., но необходимо разобраться, каков характер у такого описания: является ли полученный результат вторичной информацией (информацией об информации) или он несет первичную (фактографическую) информацию. Безусловно, вторичным является описание контента, которое осуществляется с помощью описательных метаданных в рамках определенного формата (стандарта) представления той или иной категории ресурсов. Другие метаданные не всегда правомерно называть информацией об информации.

В результате детального изучения явления метаданные каждая из групп должна быть четко терминологически обозначена, их совокупность – очерчена и более детально дифференцирована, а затем и классифицирована на основе разных признаков (как общих с другими информационными объектами, так и специфичных, проявляющихся только у метаданных). Решение этих задач позволит создать дефиницию метаданных как вторично-информационного явления и выявить особенности феноменов, называемых сегодня метаданные, но не обладающих признаком вторичности информации. Все сказанное приводит нас к выводу: явление метаданные – разнородно, а термин метаданные – многозначен (полисемичен). Многозначность термина не только тормозит научное осмысление феномена метаданные, но и затрудняет научное познание, профессиональное общение.

Тема 1.2 Форматы метаданных

Используют следующие форматы метаданных. Стандарты метаданных, созданных для описания различных видов документированной информации: архивной (AIF, CWM, DDI и др.); музейной (CDWA, MIDAS, Object ID и др.) «Ядро Ассоциации Визуальных Ресурсов» – Visual Resources Association Core, или VRA Core – для описания музейных объектов.; библиографической (MARC, RUSMARC); научной (CERIF, GRIB, VICAR и др.) и др.

Стандарты метаданных в области культурного наследия. Библиотеки: RUSMARC, UNIMARC, Концептуальные модели IFLA (FRBR, FRAD, FRSAD, LRM), METS (Metadata Encoding and Transmission Standard), MODS (Metadata Object Description Schema), BIBFRAME.

Архивы: EAD (Encoded Archival Description), ISAD (G) (General International Standard Archival Description), ISAAR (CPF) (International Standard Archival Authority Record for Corporate Bodies, Persons and Families).

Музеи: CDWA (Categories for the Description of Works of Art), CDWA Lite, LIDO (Lightweight Information Describing Objects).

BIBFRAME (BF – Bibliographic Framework Initiative) можно перевести как «Основы библиографической структуры» или «Начала библиографической конструкции. Разработчики считают, что BIBFRAME является инициативой преобразования библиографических стандартов по описанию в модель связанных данных на более высоком уровне, чтобы сделать библиографическую информацию более полезной как внутри, так и вне библиотечного сообщества.

Перспектива превращения конкретных, привычных, осязаемых стандартов, которые десятилетиями служили библиографическому отражению ресурсов, в абстрактную модель связанных данных настораживает. Но авторы проекта успокаивают, обещают, что BIBFRAME принесёт новые способы и чёткого различия между содержанием произведения и его физическим/цифровым воплощением, и однозначного определения информационных объектов, и управления связями между объектами.

BIBFRAME не только заменит формат MARC, но и включит в себя различные модели размещения содержания произведения и правил каталогизации, новые методы ввода данных.

Отвечая на вопрос о различиях между MARC и BIBFRAME, разработчики подчёркивают, что формат MARC фокусируется на каталожных записях, которые независимо друг от друга непонятны. Странная аргументация, ведь 50 лет назад MARC – машиночитаемая каталогизация – была задумана именно как противовес карточной каталогизации. Или предполагается, что записей даже в электронном каталоге совсем не будет?

Другое отличие в том, что модель BIBFRAME во многом зависит от отношений между ресурсами (связи «произведение – произведение» и т.д.), она управляет этими отношениями с помощью идентификаторов для объектов. MARC уже использует некоторые из этих идей (географические

коды, коды языков), но BIBFRAME стремится сделать эти аспекты нормой, а не исключением.

Модель BIBFRAME является формальной точкой входа библиотечного сообщества в более крупную сеть данных, где связи между объектами имеют первостепенное значение.

Структура и модель BIBFRAME

Основные компоненты BIBFRAME:

модель,

словарь,

инструменты.

Модель BIBFRAME 2.0 организует информацию на трех основных уровнях абстракции:

произведение,

воплощение,

физическая единица.

Здесь видна переключку с объектами FRBR (Функциональные требования к библиографическим записям) и RDF (Структура описания ресурса). Но в наборе компонентов, их количестве и названиях терминов есть отличия. Сопоставление объектов FRBR, RDF и BIBFRAME – тема для отдельного исследования.

По содержанию объекты FRBR, RDF и BIBFRAME совпадают.

Произведение. Самый высокий уровень абстракции. В модели BIBFRAME отражает концептуальную сущность ресурса: авторы, языки, о чем книга (предметы).

Воплощение. Произведение может иметь одну или несколько материальных реализаций, например конкретную форму опубликования. Воплощение – совокупность экземпляров произведения – отражает такую информацию, как издатель, место и дата публикации, формат.

Физическая единица. Является экземпляром воплощения (физическим или электронным), отражает такую информацию, как ее местоположение (физическое или виртуальное), полочный индекс и штрих-код.

BIBFRAME 2.0 также дает определения дополнительным ключевым понятиям, которые имеют отношение к основным классам. Например, агенты – это лица, организации, юрисдикции и т.д., связанные с произведением или воплощением через роли, такие как автор, редактор, художник, фотограф, композитор, иллюстратор и т.д.

Словарь BIBFRAME. Вместо набора областей и элементов ISBD, библиотечных стандартов и правил, в частности RDA (Описание ресурса и доступ), BIBFRAME предлагает использовать терминологические словари для ввода в запись элементов данных о содержании произведения и его носителе.

В BIBFRAME 2.0 словарь основательно переработан и назван Словарь BIBFRAME 1.0. Для него разрабатываются соответствующие инструменты и другие вспомогательные компоненты.

Словарь BIBFRAME состоит из трёх основных классов RDF и свойств. Дополнительные классы являются подклассами базовых классов.

Свойства описывают характеристики ресурса, а также отношения между ресурсами.

В RDA обе структурные части (правила и словарь) равны по объёму. В BIBFRAME от первой части осталась только упрощённая модель RDF в качестве введения, концепции, идеологической установки, а вторая часть – Словарь – стала основной, базовой.

Инструменты BIBFRAME – инструменты и демонстрации еще не разработаны. Инструментарий, на примере предшественника BIBRAFME – онлайн-ового RDA Toolkit, – это единый ресурс для реализации RDA, который включает в себя:

инструкции RDA, доступные для поиска и просмотра;

структура – по объектам FRBR;

набор элементов RDA, расположенный сначала по объектам FRBR, а затем в алфавитном порядке названий элементов, по сути это словарь;

инструменты для настройки рабочих процессов;

связи с другими каталогизационными ресурсами и т.д.

Пошаговое внедрение. Салли МакКаллум назвала восемь шагов на пути к внедрению BIBFRAME.

Шаг 1. Пересмотр модели и словаря.

Шаг 2. Особенности конверсии MARC в BIBFRAME, главная из которых – RDF-технологии. Модель RDF – это «грамматика» Семантического веба с триплетной структурой.

Раздел 2 Формирование метаданных

Тема 2.1 Метаданные цифровых ресурсов (8 часов, в том числе: 4 часа – практические занятия, 4 часа – самостоятельная работа)

В сфере метаданных главную роль играет стандарт, определенный Инициативой метаданных Дублинского ядра (DCMI). Эта организация разрабатывает и продвигает принятие стандартов определения метаданных для описания цифровых ресурсов. В частности, DCMI разработала стандартный словарь с указанием основных характеристик наиболее распространенных онлайн-овых ресурсов. Этот словарь, изначально принятый для описания библиографических ссылок, был генерализирован и адаптирован к описанию огромного числа ресурсов.

Под руководством Национальной организации по информационным стандартам (NISO) стандарт DCM был принят как стандарт ISO 15836:2003. Публикация стандартов ISO является официальным признанием использования набора, определяемого Дублинским ядром (ДЯ). ДЯ было переведено на более чем 20 языков и используется во всем мире для интеграции разных типов информации.

Первоначальная версия Дублинского ядра, которая включала 13 элементов, была предложена на состоявшемся в 1995 году в Дублине (США) симпозиуме, организованном «законодателем мод» в области библиотечного дела Online Computer Library Center (OCLC) и National Center for

Supercomputing Applications (NCSA), для описания информационных ресурсов библиотечных систем, информационных ресурсов Веб и т.п. Развитие Дублинского ядра поддерживается специально учрежденной для этой цели организацией – Инициативой по метаданным Дублинского ядра (Dublin Core Metadata Initiative, DCMИ).

Действующая версия спецификаций Дублинского ядра DC 1.1 была принята в качестве стандарта DCMИ в 1999 г., а впоследствии приобрела статус официальных стандартов Организации национальных стандартов информационных технологий США – NISO (стандарт NISO Z39.85-2001) и Международной организации стандартизации – ISO (стандарт ISO: 15836-2003).

В РФ действует ГОСТ Р 7.0.10-2019 ИСО 15836-1:2017).

Стандарт ДЯ состоит из группы элементов, используемых для описания ресурсов. Первая группа элементов, изначально задуманная для описаний, создаваемых авторами веб-ресурсов, была впоследствии принята разными сообществами, включая музеи, другие государственные организации и коммерческие предприятия. Это помогло придти к необходимому консенсусу в вопросах стандартизации на всех уровнях.

Успех ДЯ ядра объясняется простотой понимания составляющих его элементов, его всемирно признанной семантики и использования на разных языках.

ДЯ поддается расширению за счет использования уточняющих элементов:

квалификаторов и схем кодирования.

Простейший уровень ДЯ состоит из 15 базовых элементов:

Название элемента Описание элемента
DC.title (Название) Имя, данное ресурсу

DC.creator
(Создатель)

Автор (создатель) ресурса

DC.subject
(Предмет)

Тема содержания ресурса

DC.description
(Описание)

Сообщение о содержании ресурса: описание может включать, но не только, аннотацию, оглавление, графическую форму или свободный текст ресурса

DC.publisher
(Издатель)

Лицо или организация, ответственные за ввод ресурса в обращение

DC.contributor
(Соисполнитель)

Лицо или организация, внесшие вклад в создание содержания ресурса

DC.date (Дата) Дата или период времени, связанные с событием в жизни ресурса

DC.type (Тип) Природа или жанр содержания ресурса

DC.format

(Формат)

Физическое или цифровое представление ресурса

DC.identifier

(Идентификатор)

Однозначная ссылка на ресурс в пределах данного контекста

(URL, ISBN и др.)

DC.source

(Источник)

Ссылка на ресурс, из которого извлечен данный ресурс

Используя эти элементы, можно описывать ресурс как группу элементных/ценностных пар.

Так, для описания данной главы, мы могли бы использовать следующую последовательность элементных/ценностных пар:

DC.title="Значение использования метаданных"

DC.creator="MINERVA WP5 Quality Accessibility Usability"

DC.publisher="MINERVA Project"

DC.type="текст"

DC.format="html"

DC.language="ru"

Все элементы языка – не обязательные и могут, при необходимости, ПОВТОРЯТЬСЯ.

Например: если имеется не один, а несколько авторов, элемент DC.creator может повторяться. Квалифицированный (компетентный) уровень ДЯ добавляет элементы и вводит группу квалификаторов, позволяющих уточнить семантику базовых элементов.

Название элемента Описание элемента

DC.audience

(Аудитория)

Группа лиц, на которых ориентирован ресурс или которым он может быть полезен

DC.provenance

(Происхождение)

Информация о возможных изменениях, внесенных с момента создания ресурса.

DC.rightsHolder

(Правообладатель)

Определяет владельца прав на использование ресурса.

DC.instructionalMethod

(Метод обучения)

Учебный процесс, в котором может использоваться данный ресурс.

DC.accrualMethod

(Метод сбора)

Метод, с помощью которого ресурс добавляется в коллекцию.

DC.accrualPeriodicity

(Периодичность сбора)

Частота, с которой ресурс добавляется в коллекцию.

DC.accrualPolicy

(Политика сбора)

Политика, управляющая добавлением ресурса в коллекцию.

Квалификаторы позволяют более подробно описывать информацию, предоставляемую базовыми элементами. Они выражены в уточняющих элементах – элементах, которые конкретизируют определенные характеристики, и схемах кодирования – группах ссылок, служащих для интерпретации ценности элементов. Например: мы можем переделать описание данной главы и представить ее в следующей форме с большей детализацией:

DC.title=“Значение использования метаданных”

DC.creator=“MINERVA WP5 Quality Accessibility Usability”

DC.publisher=“MINERVA Project”

DC.type=“текст”

DC.format=“text/html”

объем “62 kb”

DC.language=“ru”

В приведенном выше описании мы указали, что тип ресурса – «текст» на основании стандартной схемы кодирования словаря типов DCMI, что его формат – text/html на основании стандарта MIME Type, и что язык – английский на основании стандарта. Мы также использовали элемент DC.extent (Протяженность), который позволяет нам определить размер ресурса. Структурирование расширяющего элемента говорит о том, что мы имеем дело со вспомогательным определением (уточнением) элемента формата.

Язык, определенный в соответствии со стандартом ДЯ, не зависит от каких-либо специальных стандартов, и мы можем выбирать терминологию, которая больше всего ему соответствует. К примеру, мы можем сделать так, чтобы элементы языка соответствовали полям таблицы БД или свойствам объекта на языке программирования. В любом случае DCMI дает нам некоторые указания, как использовать обычный синтаксис для описания онлайн-ресурсов, таких как веб-страницы. Их можно описать в терминах элементов ДЯ, выраженных <мета>метками в документе XHTML того же самого ресурса, цитируя источник/UTL принятой схемы. Например: данную главу можно описать с помощью следующего XHTML-кода, собрав метаданные в разделе <head> (заголовок) исходного кода XHTML каждой страницы и используя элементы <meta> и <link>:

```
head>
```

```
<link rel=“schema.DC” href= “http://purl.org/dc/elements/1.1/” />
```

```
<link rel=“schema.DCTERMS” href= “http://purl.org/dc/terms/” />
```

```
<meta name=“DC.title” lang=“ru” content=“Значение использования метаданных” />
```

```

    <meta name="DC.creator" content="MINERVA WP5 Quality Accessibility
Usability" />
    <meta name="DC.publisher" content="MINERVA Project" />
    <meta name="DC.type" scheme="DCTERMS.DCMIType" content="Text"
/>
    <meta          name="DC.format"          scheme="DCTERMS.IMT"
content="text/html"/>
    <meta name="DC.format" scheme="DCTERMS.extent" content="62kb" />
    <meta  name="DC.language"  scheme="DCTERMS.  ISO639-2"
content="eng" />
</head>

```

В данном примере элементы `<link>` указывают на схемы, используемые для элементов простого и квалифицированного ДЯ, а элемент `<meta>` соответствует элементным/ценностным парам, использованным в первых примерах. Там, где это возможно, данные отдельных элементов следует выбирать из нормализованного словаря – совокупности тщательно определенных терминов. Это позволит нам значительно улучшить результаты поиска, поскольку компьютеры могут идентифицировать отдельные слова, но им сложно понимать концепции, ассоциации, синонимы и пр.

Без определенного уровня терминологического контроля неполные или неправильные метаданные могут стать причиной низкого качества результатов поиска. Нормализованные словари и тезаурусы можно использовать с помощью схем кодирования. Тим Бернес-Ли, Джеймс Хендлер и Ора Лассила определили семантическую паутину как «новую форму содержания Всемирной паутины, понятную компьютерам» (The semantic web, Scientific American, May 2001).

Сущность семантической паутины – обнаружение информации и услуг, которые будут распознаваться по Унифицированному идентификатору ресурсов (URI), используя концепции, а не ключевые слова и открывая новые возможности для автоматизации услуг. Для описания имеющихся в Сети данных машиной (в этом смысле исследователи используют слово «семантический») следует определить модели представления знаний. Семантическая паутинa объединяет под общим названием «W3C Semantic Web Activity» различные принципы проектирования, рабочие группы и комплекс развивающихся технологий. Основные сложности внедрения семантической паутины заключаются в определении и универсальном распространении стандартных форматов, обеспечивающих совместимость приложений и внедрение полностью автоматизированного способа дедуктивных заключений, экспортируя в Сеть правила из любой базы знаний. Семантическая паутинa – это декларативная среда, которая использует стандарты и средства, основанные на XML Namespace и XML Schema. Этот комплект стандартов W3C предоставляет элементарный синтаксис для структурирования контента внутри документов, но не ассоциирует напрямую семантику со значением контента. Технология W3C для кодирования, обмена и повторного использования метаданных

структурированной сети называется «Модель для описания ресурсов» (RDF). Чтобы выразить ограничения по ассоциациям и избежать кодирования синтаксически правильных операторов без всякого смысла, необходим механизм представления классов объектов. Для этого и был разработан язык описания словарей RDF или модель RDF.

После того, как мы выразим данные и правила данных, мы сталкиваемся с необходимостью иметь язык, способный экспортировать наши знания (онтологии) и предоставить их в доступ любому приложению: язык веб-онтологий (OWL).

Все эти компоненты организованы, как правило, в т.н. «Стек понятий семантической паутины»: над XML (придающего ресурсам вид структуры) и RDF (для выражения значений или, правильнее сказать, для обозначения того, что некоторые элементы обладают некоторыми свойствами) представлен онтологический уровень, т.е. уровень, на котором происходит формальное определение отношений между терминами. Верхний уровень – это логический уровень, на котором утверждения, представленные в паутине, могут использоваться для порождения нового знания без применения уникальной, универсальной системы рассуждений, но с применением логики объединения для представления всех достоверных результатов.

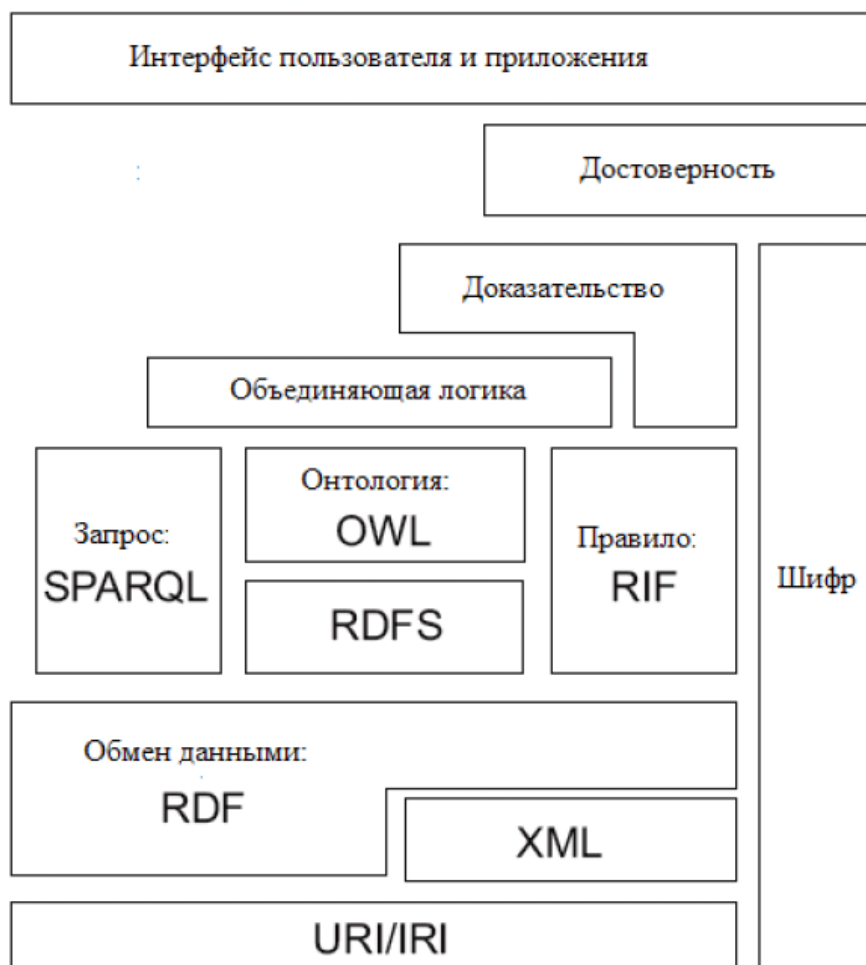


Схема понятий семантической паутины

Модель для описания ресурсов (RDF) – это универсальная, базовая система для кодификации, обмена и повторного использования структурированных метаданных. Она поддерживает совместимость веб-приложений, осуществляющих обмен машинопонимаемой информацией.

Модель данных RDF, представляющая операторы RDF в синтаксически нейтральной форме, очень проста и базируется на 3-х типах объектов: ресурсах, свойствах и операторах.

Первые два однозначно идентифицируются по URI.

Ресурсы – это все, что можно описать выражением RDF: веб-страница или ее часть, элемент XML внутри исходного документа и даже целая совокупность веб-страниц или объект, к которому нет прямого доступа по Сети.

Свойства – это конкретный аспект, характеристика, определение или отношение, используемое для описания ресурса.

Каждое свойство имеет определенный смысл. Оно определяет допустимые значения, типы ресурсов, которые может описать, и связи с другими свойствами.

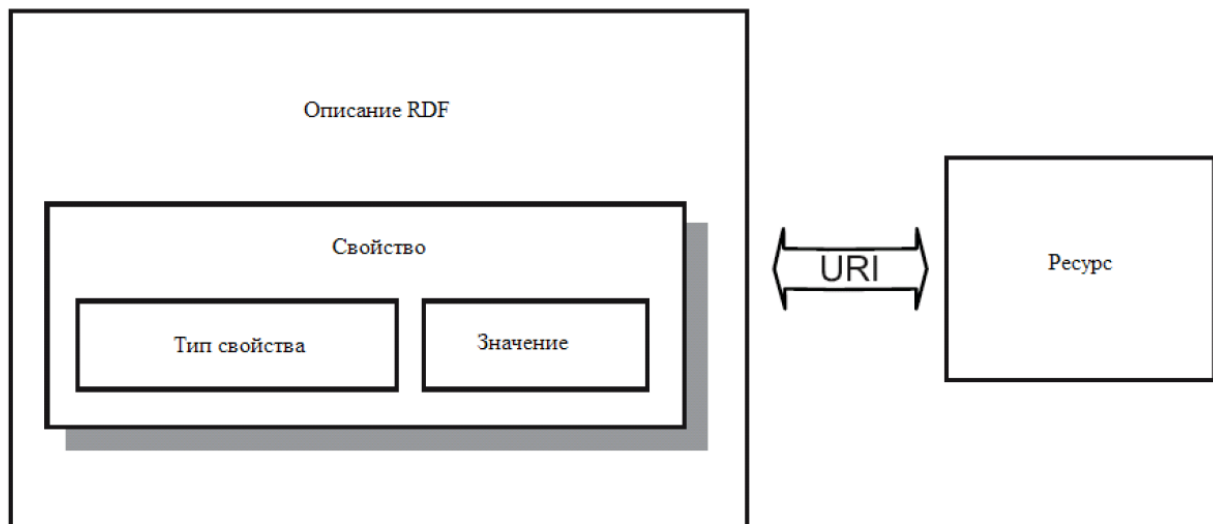
Свойства, ассоциируемые с ресурсом, идентифицируются по имени и обладают значениями.

Операторы – ресурс с определенным свойством, определяемый именем, и значение этого свойства для определенного ресурса, формируют оператор RDF.

Таким образом, оператор – это триплет, состоящее из субъекта (ресурса), предиката (свойства) и объекта (значения).

Объект оператора (значение) может быть выражением (последовательностью символов или каким-либо другим примитивным типом, описываемым XML) или другим ресурсом.

Совокупность свойств, описывающих определенный ресурс, называется «описание».



Модель данных RDF

Лит

Дублинское ядро используется при формировании метаданных документов репозитория, организованных на системе DSpace. А. М. Федотов, А. Т. Байдавлетов, О. Л. Жижимов, М. А. Самбетбаева, О. А. Федотова представили формирование цифрового репозитория в научно-образовательной информационной системе.

Показать сокращенную информацию

| | | |
|-------------------------|---|----|
| dc.contributor.author | Предеина, Анжелика Вадимовна | ru |
| dc.contributor.author | Федорина, Алла Ивановна | ru |
| dc.date.accessioned | 2021-07-02T09:13:53Z | |
| dc.date.available | 2021-07-02T09:13:53Z | |
| dc.date.issued | 2021 | |
| dc.identifier.citation | Предеина, А. В. Специфика отражения краеведческого компонента в региональных сводных электронных каталогах библиотек Беларуси / А. В. Предеина, А. И. Федорина // Культура. Наука. Творчество : XV Международная научно-практическая конференция (Минск, 13 мая 2021 г.) : сборник научных статей / Белорусский государственный университет культуры и искусств [и др.] - Минск, 2021. - [Вып. 15]. - С. 215-220. - Библиогр.: с. 219-220 (14 назв.). | ru |
| dc.identifier.uri | http://repository.buk.by/123456789/23891 | |
| dc.description.abstract | Сводные каталоги (СвК) | ru |
| dc.language.iso | other | ru |
| dc.publisher | БГУКИ | ru |
| dc.subject | сводные каталоги | ru |
| dc.subject | краеведческий аспект | ru |
| dc.subject | региональные сводные электронные каталоги | ru |
| dc.subject | корпоративные электронные ресурсы | ru |
| dc.title | Специфика отражения краеведческого компонента в региональных сводных электронных каталогах библиотек Беларуси | ru |
| dc.type | Article | ru |

Поиск в DSpace

Поиск в DSpace
 В этой коллекции

Просмотр

Весь DSpace
[Сообщества и коллекции](#)
[Дата публикации](#)
[Авторы](#)
[Названия](#)
[Тематика](#)
 Эта коллекция
[Дата публикации](#)
[Авторы](#)
[Названия](#)
[Тематика](#)

Моя учетная запись

[Войти](#)
[Регистрация](#)

Файлы в этом документе



Имя: SPECIFIKA OTRAZHE...
 Размер: 182,2Кб
 Формат: PDF

[Открыть](#)

Акти
 Чтобы

Для организации системы долговременного хранения информационных ресурсов международной организацией по стандартизации (ISO) предложен стандарт ISO-14721:2012 1 (OAIS – Open Archive Information System) [6]. Эталонная модель OAIS 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. К основным особенностям институционального репозитория относятся:

обеспечение разграниченного доступа к разнородным цифровым объектам (публикациям, изображениям и т. д.);

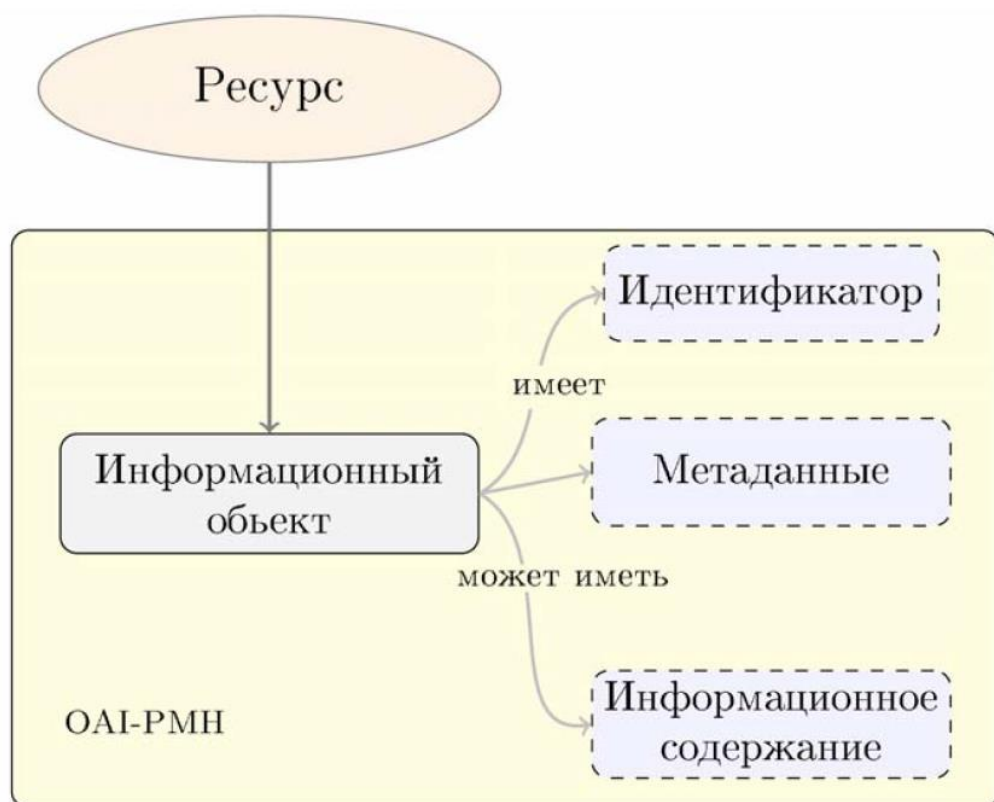
организация доступа к информационным ресурсам для мирового сообщества (в том числе с помощью полнотекстового индексирования мировыми поисковыми системами);

унифицированный доступ к метаданным по стандартным протоколам (поддержка интероперабельности);

возможность организации единой точки доступа к информационным ресурсам;

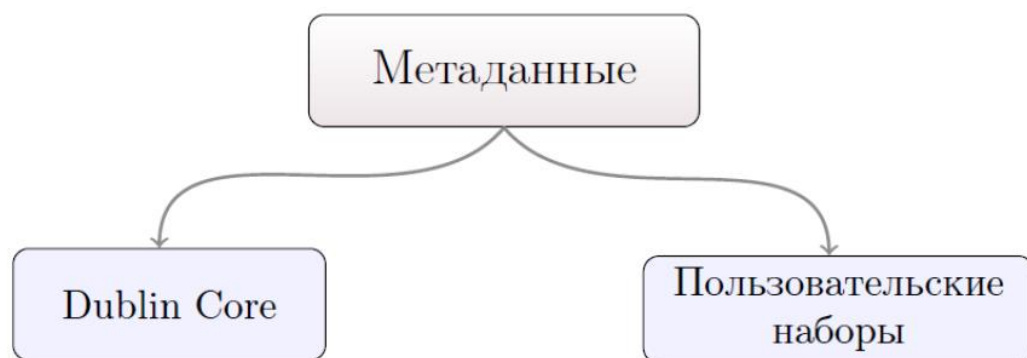
Инициатива «Открытые архивы» (OAI) разрабатывает и продвигает стандарты интероперабельности с целью эффективного распространения

электронных ресурсов, а также повышения доступности обмена научной информацией. Протокол OAI-PMH имеет простую и гибкую информационную модель представленную на рисунке:



Основная сущность, с которой оперирует модель, – это Ресурс (Resource). В общем случае Ресурс интерпретируется в терминах библиографических данных, описывающих ресурс, хотя возможны и другие интерпретации. В информационной системе Ресурсу соответствует Информационный объект (Item – элемент), который является традиционным вторичным информационным объектом, содержащим описание первичного ресурса (в нашем случае публикации или документа).

Информационный объект концептуально является контейнером, который хранит или динамически генерирует метаданные об отдельном ресурсе в нескольких форматах, каждый из которых может быть собран в виде записи OAI-PMH. Каждый информационный объект имеет идентификатор записи или уникальный идентификатор. Уникальный идентификатор однозначно определяет Информационный объект в репозитории, он используется в запросах OAI-PMH для извлечения метаданных. Протокол OAI-PMH требует обязательного включения в описание ресурса набора метаданных Дублинское ядро (Dublin Core – DC). Также желательно включать в описание пользовательские наборы метаданных (например, MARC 3), расширяющие DC, что представлено на рисунке:



Тема 2.2 Метаданные объектов национальной библиографии

В соответствии с руководством ИФЛА «Национальная библиография в электронную эру» рассмотрены следующие положения. Национальные библиографические агентства, выделяя средства и персонал для описания электронных ресурсов, должны учитывать и необходимость обеспечения долговременного доступа к ним.

Отправной точкой оценки требований к метаданным является изучение пользователей и путей использования национальной библиографии. Использование и функционирование системы возможно только при условии наличия необходимых метаданных. Отдельные пользователи и их группы будут предъявлять к метаданным различные требования при реализации стандартных функций, таких как поиск, а кроме того, могут иметь особые потребности, отсутствующие у других групп или у отдельных пользователей. При изучении потребностей принимается допущение, что национальные библиографические агентства не будут специально создавать метаданные для индивидуальных пользователей (хотя они могут предложить пользователям сервис, позволяющий им присваивать свои собственные метки (tags) и создавать аннотации). Таким образом, в ходе анализа индивидуальные пользователи рассматриваются как члены определенных групп пользователей.

В связи с появлением Интернета для всех видов национальных публикаций стал необходим единый стандарт.

Необходимо учитывать, что уровень каталогизации, применяемый для обработки соответствующих типов ресурсов, зависит:

- от уровня метаданных, присвоенных ресурсу;
- от значимости материала для национальной библиографии;
- от содержания материала, но не его носителя.

Первая из этих рекомендаций обуславливает значительные изменения в подходе к созданию национальной библиографии. Внедрение разных уровней каталогизации оказывает прямое воздействие на пользователей национальной библиографии, следовательно, национальные библиографические агентства должны принимать во внимание значимость электронного ресурса и степень его воздействия на пользователей.

В дополнение к этому процессы создания национальной библиографии должны быть упрощены и автоматизированы там, где это возможно. Дублирование работ должно быть исключено посредством организации взаимодействия между разными секторами и посредством упразднения барьеров, препятствующих обмену данными.

Национальным библиографическим агентствам рекомендовано использовать готовые метаданные, относящиеся к электронному ресурсу.

Метаданные, используемые повторно, могут создаваться не для целей поиска и использования ресурсов. Было бы правильнее рассматривать процесс повторного использования метаданных как смену их предназначения. Метаданные, предназначение которых было изменено, вряд ли будут настолько же эффективны, как и метаданные, специально созданные для поддержки функции поиска и использования ресурсов.

Функциональным недостатком метаданных, предназначение которых было изменено, является их несоответствие содержанию контент-стандартов, специально разработанных для поддержки функции поиска и использования ресурсов. Это оказывает влияние на возможность метаданных поддерживать полный спектр задач по поиску и использованию ресурсов. Это особенно важно для более сложных задач, таких как идентификация и отбор ресурсов, и для представления ресурса в более широком библиографическом и тематическом контексте.

Идентификация и описание связей между ресурсами в значительной степени сохраняет свою зависимость от человеческого фактора. В Библиотеке Конгресса США проводились эксперименты с так называемыми записями «уровня доступа» к электронным ресурсам, которые показали, что при вмешательстве человека фокус смещается ближе к авторитетному контролю, чем к описанию ресурса. В результате в Библиотеке Конгресса был отмечен потенциал значительного снижения затрат.

Метаданные имеются не у всех электронных ресурсов. Электронные страницы Интернета или ресурсы, полученные по обязательному экземпляру и предоставленные на добровольной основе, могут не иметь метаданных. Значение этого факта зачастую преуменьшается, так как полнотекстовые электронные ресурсы могут рассматриваться как описывающие сами себя. Это справедливо только в случае, если полнотекстовые ресурсы могут быть проиндексированы и включены в систему поиска. Индексирование ключевых словами, отражающими содержание электронного ресурса, значительно отличается от использования контролируемых словарей, таких как предметные системы индексирования или классификационные схемы. Различие проявится в точности поиска и возможности повторного использования ресурса. Поиск по ключевым словам будет, скорее всего, содержать много неверных результатов поиска и не отразит материалы на других языках или использующие другую терминологию.

Многие электронные ресурсы не являются текстовыми. Они могут содержать изображения или звуковые записи, комбинации нетекстовых материалов. Поиск ресурсов этого типа зависит от наличия метаданных для организации этого процесса. Например, поиск необходимых иллюстраций

для включения в презентацию – это долгая и сложная процедура; если нет словесной характеристики, позволяющей описать иллюстрации тематически, это может оказаться практически невозможным.

Рабочая группа выделила четыре уровня соответствующих метаданных.

Авторитетный уровень – это самый высокий уровень метаданных, присвоенных ресурсу. Определение в качестве авторитетного означает, что (специфические) точки доступа контролируются авторитетными записями, а описание ресурса подчиняется правилам, подробно разработанным контент-стандартами.

Полный уровень каталогизации – это формирование максимально возможного объема метаданных при описании ресурсов и обеспечении доступа к ним. В этом случае описание ресурса полностью соответствует опубликованным контент-стандартам, но не предполагает обязательств обеспечивать любую точку доступа авторитетными данными. Примерами полного уровня метаданных могут служить записи в форматах MARC, профессиональном Dublin Core, ONIX и в других книготорговых форматах.

Расширенный уровень показывает, что исходный ресурс был дополнен некоторыми метаданными. Расширенные метаданные могут не соответствовать в точности используемым контент-стандартам, а также не иметь никаких обязательств по обеспечению точек доступа авторитетного контроля. Примерами расширенных метаданных, которые могут быть использованы национальными библиографическими агентствами, являются: непрофессиональный Dublin Core и заголовки TEI.

Базовый уровень не предполагает наличия метаданных, присваиваемых ресурсу национальным библиографическим агентством или создавшей его организацией. Как правило, эта категория включает ресурсы, напрямую получаемые из Интернета без метаданных или с метаданными, сформированными посредством автоматизированной обработки.

Уровень метаданных, требуемый для поиска электронных ресурсов с целью их включения в систему национальной библиографии, должен быть определен национальным библиографическим агентством. При решении этого вопроса должна быть принята во внимание значимость ресурса по отношению ко всему объему национальных публикаций. Ответственность за решение о значимости ресурсов обычно возлагается на тех, кто отвечает за установление критериев отбора для национальной библиографии и коллекций национальных документов. Те, кто формирует национальные каталоги, будут ответственны за определение наиболее подходящего уровня метаданных. Таблица, приведенная ниже, иллюстрирует, каким образом те или иные уровни метаданных могут соответствовать определенным категориям ресурсов.

| Значимость | Определение | Рекомендуемый уровень |
|------------|---|-------------------------|
| Высокая | Официально опубликованные; аналог или заменитель печатного источника; могут содержать стандартные идентификаторы. Ресурсы, ценные для проведения научных исследований | Авторитетный |
| Средняя | Материалы высокой культурной ценности или инновационные сайты, которые официально не опубликованы | Полный |
| Низкая | Персональная или недолговечная информация, способная вызывать незначительный интерес у современной или будущей аудитории | Базовый или Расширенный |

Целесообразным представлено взаимодействие библиотек, издательств, книжной торговли по представлению метаданных, содержащим идентификаторы и о сближении стандартов метаданных.

По мере того как системы национальной библиографии становятся все более и более доступными он-лайн, проблема использования постоянных идентификаторов цитирования и ссылок становится решающей на международном уровне.

Постоянные идентификаторы – это такие идентификаторы, которые позволяют извлекать информацию на долгосрочной основе. Даже если ресурс перемещается в каталоге или электронном хранилище, идентификатор остается тем же, так как система проверяет соответствие между идентификатором, известным общественности, и действующим адресом

Для определения местонахождения ресурса в Интернете используется следующий идентификатор: URL (унифицированный указатель ресурса), который и идентифицирует, и указывает местонахождение ресурса путем применения синтаксической структуры протокола (доменное имя и фактическое имя файла, в котором в настоящее время находится ресурс).

Однако один URL не может обеспечить связь с ресурсом, особенно если тот перемещен в архив. Постоянный и уникальный идентификатор сохранил бы доступ к такому ресурсу независимо от его фактического местонахождения. Подобная форма картирования осуществляется с помощью базы данных, содержащей преобразователь адресов.

Постоянный унифицированный указатель ресурса (PURL) был создан и внедрен OCLC в качестве сервиса, позволяющего присваивать имена и определять местонахождение основных ресурсов Интернета. Он предполагался в качестве промежуточной системы, которая должна была использоваться до окончательного внедрения системы URN. PURL аналогичен URL, за исключением того, что выводит пользователя на

интерпретационный сервис вместо указания фактического местонахождения цифрового ресурса. Интерпретационный сервис затем перенаправляет пользователя на соответствующий URL.

Система «Код архивированного ресурса» (ARK), которая не зависит от местонахождения ресурса и используемого протокола, – это новый подход к постоянной идентификации. Она была создана в 2001 г. Джоном Кунзе для хранителей архивированных цифровых объектов. В ней реализуется принцип управления ресурсами и схемами их наименования, действующий в течение длительного времени.

Идентификатор ARK – это схема наименования цифровых объектов для постоянного доступа к ним (в том числе к изображениям, текстам, наборам данных и средствам поиска).

Идентификатор – это ассоциация между строкой (последовательность символов) и информационным ресурсом. Такая ассоциация отражается записью, которая связывает идентифицирующую строку с набором характеристик идентифицируемого ресурса. Идентификатор ARK – это специально созданный, уникальный, действующий URL.

Схема основана на трех требованиях:

ссылка от объекта на службу управления ресурсами;

ссылка от объекта к метаданным, которые его описывают;

ссылка на сам объект (или соответствующий заменитель).

Создание системы идентификаторов цифрового объекта (DOI) было начато Ассоциацией американских издателей, а сейчас продолжается Международным фондом DOI (International DOI Foundation): <http://www.doi.org>. Система уникальных идентификаторов DOI опирается на систему идентификационных дескрипторов (Handle System) и позволяет присваивать уникальный цифровой идентификатор коммерческим электронным публикациям. Преимущества DOI – быстрый поиск документа в интернете, переход к его электронному тексту, гарантия актуального местонахождения документа в интернете. Примеры структуры DOI и отражения в информационно-поисковых системах.

Структура кода DOI



Научные и технические библиотеки

Расширенный поиск

ГЛАВНАЯ

О ЖУРНАЛЕ

ТЕКУЩИЙ ВЫПУСК

[Главная](#) > [№ 3 \(2018\)](#) > [Бахтурина](#)

От MARC 21 к модели BIBFRAME: эволюция машиночитаемых форматов Библиотеки Конгресса США

Т.А. Бахтурина

<https://doi.org/10.33186/1027-3689-2018-3-62-70>



ПОЛНЫЙ ТЕКСТ:

[PDF \(RUS\)](#)

Аннотация

Об авторе

КЛЮЧЕВЫЕ СЛОВА

[машиночитаемые форматы](#), [Библиотека Конгресса США](#), [модель BIBFRAME](#), [терминология](#)

3 ПРАКТИЧЕСКИЙ РАЗДЕЛ

Тематика и вопросы круглых столов

Круглый стол

Тема 1.1 Метаданные: терминосистема, свойства, функции, классификация (2 часа)

Метаданные: общая характеристика.

Цель: освоение терминосистемы, свойств, функций, классификаций метаданных.

Задача: формирование системных представлений о метаданных.

Порядок взаимодействия:

1. Открытие круглого стола.
2. Презентация темы, целей и задач круглого стола.
3. Презентация вопросов круглого стола.
4. Обсуждение вопросов круглого стола.
5. Подведение итогов круглого стола.

Вопросы для обсуждения:

1. Определение понятия «метаданные», их свойства, классификация, функции.

2. Ресурсы, описываемые метаданными.

Средства представления, стандартизация метаданных

4 РАЗДЕЛ КОНТРОЛЯ ЗНАНИЙ

4.1 Материалы для самоконтроля по дисциплине «Метаданные»

1. Определение понятия «метаданные».
2. Ресурсы, описываемые метаданными.
3. Свойства метаданных.
4. Универсальные функции метаданных.
5. Специфические функции метаданных.
6. Виды метаданных.
7. Средства представления, стандартизация метаданных.
8. «Электронные информационные ресурсы»: общая характеристика.
9. ГОСТ 7.82–2001. «Библиографическая запись. Библиографическое описание электронных ресурсов. Общие требования и правила составления»: общая характеристика.
10. ГОСТ Р ИСО 489-1–2019. «Информация и документация. Управление документами. Часть 1. Понятия и принципы»: понятие, принципы, требования к метаданным документам.
11. СТБ 7.38–2017. «Электронные издания. Основные виды и выходные сведения»: общая характеристика.
12. ГОСТ 7.70–2003. «Описание баз данных и машиночитаемых информационных массивов. Состав и обозначение характеристик»: система метаданных для описания электронных информационных ресурсов.
13. Цифровые коллекции как объект формирования метаданных.
14. ГОСТ Р 7.0.96–2016. «Электронные библиотеки. Основные виды. Структура. Технология формирования»: характеристика метаданных.
15. Метаданные для описания веб-сайтов.
16. Универсальные стандарты метаданных.
17. Стандарты метаданных для описания архивных документов.
18. Стандарты метаданных для описания музейных предметов.
19. Стандарты метаданных для описания информационных ресурсов библиотек.
20. Стандарты метаданных в области культурного наследия.
21. BIBFRAME: общая характеристика.
22. Дублинское ядро (DC, Dublin Core): определение, структура, стандартизация.
23. Дублинское ядро (DC, Dublin Core): набор элементов метаданных в соответствии со стандартом ГОСТ Р 7.0.10–2019 (ИСО 15836–1:2017). Набор элементов метаданных «Дублинское ядро». Основные (ядерные) элементы (ISO 15836-1:2017, Information and documentation – The Dublin Core metadata element set – Part 1: Core elements, MOD).
24. DSpace как система формирования депозитария. Виды метаданных.
25. Авторские профили в наукометрических базах данных.
26. Идентификаторы научной коммуникации: DOI, ORCID, ROR. Значение, структура.

27. Роль ИФЛА в организации учета электронных ресурсов национальной библиографией.

28. Библиографические стандарты и семантика Интернета (Семантический веб).

29. Виртуальный международный авторитетный файл (VIAF).

30. Уровни метаданных национальных электронных ресурсов: авторитетный, библиографический.

31. Методика каталогизации национальных электронных документов сетевого распространения Беларуси.

4.2 Задания для самостоятельной работы

Раздел 1 Основа теории метаданных

Тема 1.1 Метаданные: терминосистема, свойства, функции, классификация

Задание (2 часа)

Подготовить материалы, необходимые для обсуждения на круглом столе, по следующим вопросам:

1. Метаданные, их свойства, классификация, функции.
2. Ресурсы, описываемые метаданными.
3. Средства представления, стандартизация метаданных.

Литература для самостоятельного изучения: [2, с. 19–20; 7, с. 7–11; 25, с. 59–77; 26, с. 39–51].

Требования к проверке результатов самостоятельной работы. Проверка результатов самостоятельной работы осуществляется при проведении самоконтроля (вопросы № 1–7).

Тема 1.2 Форматы метаданных

Задание (4 часа)

Изучить с использованием рекомендуемой литературы следующие вопросы:

1. Стандарты метаданных для описания архивных документов.
2. Стандарты метаданных для описания музейных предметов.
3. Стандарты метаданных для описания информационных ресурсов библиотек.
4. Стандарты метаданных в области культурного наследия.

Литература для самостоятельного изучения: [3, с. 12–17; 4, с. 63–71; 15, с. 37–53].

Требования к проверке результатов самостоятельной работы. Проверка результатов самостоятельной работы осуществляется при проведении при проведении самоконтроля (вопросы № 16–21).

Раздел 2 Формирование метаданных

Тема 2.1 Метаданные цифровых ресурсов

Задание (4 часа)

Изучить с использованием рекомендуемой литературы следующие вопросы:

1. Универсальные стандарты метаданных: «Дублинское ядро» (Dublin Core), CIDOC-CRM (CIDOC Conceptual Reference Model), VRA Core (Visual Resource Association Core Categories). «Схема метаданных для описания объектов» (Metadata Object Description Scheme, или MODS). Стандарт METS (Metadata Encoding and Transmission Standard – Стандарт кодирования и передачи метаданных).

2. Семантическая сеть и роль метаданных в ней.

3. Электронные информационные ресурсы: определение понятия, классификация. Электронный документ. Цифровые коллекции. Веб-сайты. Система метаданных для описания электронных информационных ресурсов. ГОСТ Р ИСО 489-1–2019. Информация и документация. Управление документами. Часть 1. Понятия и принципы. (ISO 15489-1:2016, Information and documentation – Records management – Part 1: Concepts and principles, IDT).

4. Проанализировать и привести пример метаданных платформы DSpace репозитория учреждения образования Беларуси (по выбору).

5. Авторские профили в наукометрических базах данных. Идентификаторы современной научной коммуникации: DOI (Digital Object Identifier) – цифровой идентификатор объекта, ORCID (Open Researcher and Contributor ID) – открытый идентификатор автора и исследователя и ROR (Research Organization Registry) – реестр исследовательских организаций.

Литература для самостоятельного изучения: [8, с. 3–16; 20, с. 26–27; 30, с. 24–38; 32, с. 8].

Требования к проверке результатов самостоятельной работы. Проверка результатов самостоятельной работы осуществляется при проведении при проведении самоконтроля (вопросы № 8–16; 22–26).

Тема 2.2 Метаданные объектов национальной библиографии

Задание (6 часов)

Изучить с использованием рекомендуемой литературы следующие вопросы:

1. Электронные ресурсы как объект национальной библиографии.
2. Роль ИФЛА в организации учета электронных ресурсов национальной библиографией.
3. Стандарты национальной библиографии представления метаданных электронных ресурсов.
4. Авторитетный контроль. Виртуальный международный авторитетный файл (VIAF).
5. Библиографические стандарты и семантика Интернета (Семантический Web).
6. Представить по титульному листу на разные виды национальных электронных документов сетевого распространения Беларуси и расписать библиографические сведения метками полей и подполей в соответствии со структурой BELMARC-формата.

Литература для самостоятельного изучения: [1, с. 135–152; 17, с. 37–43; 28, с. 168–176].

Требования к проверке результатов самостоятельной работы. Проверка результатов самостоятельной работы осуществляется при проведении при проведении самоконтроля (вопросы № 27–31).

5 ВСПОМОГАТЕЛЬНЫЙ РАЗДЕЛ УЧЕБНАЯ ПРОГРАММА

Учреждение образования
«Белорусский государственный университет культуры и искусств»
Институт повышения квалификации и переподготовки кадров

УТВЕРЖДАЮ

Ректор университета

Н.В.Карчевская

. .2022

Рег. № УД-___ Пп/эуч.

УЧЕБНАЯ ПРОГРАММА ПО ДИСЦИПЛИНЕ

МЕТАДАННЫЕ

специальности переподготовки
2-23 02 71 Библиотекведение и библиография,
квалификация: библиотекарь-библиограф,
в соответствии с типовым учебным планом переподготовки,
утвержденным постановлением Министерства образования Республики Беларусь
от 28.09.2020 № 260, рег. № 25-13/126

Минск, 2022

Разработчики программы:

А.И. Федорина, доцент кафедры информационных ресурсов и коммуникаций учреждения образования «Белорусский государственный университет культуры и искусств»;

А.В. Предеина, доцент кафедры и психолого-педагогических дисциплин Института повышения квалификации и переподготовки кадров учреждения образования «Белорусский государственный университет культуры и искусств», кандидат педагогических наук

Рекомендована к утверждению:

кафедрой культурологии и психолого-педагогических дисциплин Института повышения квалификации и переподготовки кадров учреждения образования «Белорусский государственный университет культуры и искусств»

Протокол заседания от 07.06.2022 № 33

Советом Института повышения квалификации и переподготовки кадров учреждения образования «Белорусский государственный университет культуры и искусств»

Протокол заседания от 14.06.2022 № 6

Научно-методическим советом учреждения образования «Белорусский государственный университет культуры и искусств»

Протокол заседания от 15.06.2022 № 5

ВВЕДЕНИЕ

Учебная программа по дисциплине «Метаданные» предусматривает реализацию требований, установленных Образовательным стандартом переподготовки руководящих работников и специалистов по специальности 2-23 02 71 «Библиотечное дело и библиография», утвержденным постановлением Министерства образования Республики Беларусь от 28.09.2020 № 259.

Дисциплина рассчитана на 30 учебных часов, которые включают 14 учебных часов аудиторных занятий (в том числе, 4 учебных часа лекций, 8 учебных часов практических занятий, 2 учебных часа – круглый стол); 16 учебных часов самостоятельной работы.

Дисциплина «Метаданные» относится к компоненту учреждения образования типового учебного плана и изучается во взаимосвязи с дисциплинами «Каталогизация документов», «Основы информационных технологий» и другими.

Актуальность учебной программы обусловлена необходимостью приобретения компетенций по каталогизации электронных документов сетевого распространения с целью развития информационной грамотности, овладения методиками, которые являются неотъемлемой частью библиотечной деятельности.

Цель: освоить современные подходы к каталогизации электронных документов сетевого распространения.

Задачи:

освоение теоретических знаний о метаданных;
сформировать представление о форматах метаданных;
познакомиться с метаданными цифровых ресурсов;
познакомиться с технологией формирования метаданных цифровых ресурсов;

овладеть методикой каталогизации национальных электронных документов сетевого распространения Беларуси.

Методы обучения: традиционные (методы организации и осуществления учебно-познавательной деятельности; методы стимулирования и мотивации учебно-познавательной деятельности; методы контроля и самоконтроля); активный; интерактивный.

Средства обучения: печатные издания; презентации, видео- и аудиозаписи.

Формы учебных занятий: лекции, практические занятия, круглый стол.

Форма получения образования: заочная.

Основные требования к результатам учебной деятельности слушателей.

В результате освоения дисциплины «Метаданные» у слушателя должны быть сформированы социально-личностные, академические и профессиональные компетенции.

Требования к формированию социально-личностных компетенций.

Слушатель должен:

знать идеологию белорусского государства, нравственные и правовые нормы, уметь учитывать их в своей профессиональной деятельности;

уметь ориентироваться в процессах и тенденциях развития современного общества.

Требования к формированию академических компетенций.

Слушатель должен:

владеть базовыми теоретическими знаниями в области документоведения, библиотековедения и библиографоведения, уметь применять их для решения практических и научных задач в библиотечной сфере;

уметь использовать информационные технологии в библиотечной сфере.

Требования к формированию профессиональных компетенций.

Слушатель должен:

использовать теоретические знания о метаданных;

анализировать и использовать структуру коммуникативных форматов в автоматизированной обработке документов;

освоение теоретических знаний о метаданных;

знать форматы метаданных;

знать технологию формирования метаданных цифровых ресурсов;

уметь представлять метаданные цифровых ресурсов;

уметь осуществлять каталогизацию национальных электронных документов сетевого распространения Беларуси;

уметь работать в АРМ «Каталогизатор» АБИС.

В содержание учебной программы включены материалы для аудиторной работы с использованием активных форм обучения и задания для самостоятельной работы.

Контроль самостоятельной работы слушателей осуществляется во время учебных занятий и при самоконтроле. Вопросы и задания для самостоятельной работы включены в материалы для самоконтроля.

СОДЕРЖАНИЕ ПРОГРАММЫ

Раздел 1 Основа теории метаданных

Тема 1.1 Метаданные: терминосистема, свойства, функции, классификация (6 часов, в том числе: 2 часа – лекция, 2 часа – круглый стол, 2 часа – самостоятельная работа)

Определение понятия «метаданные». Ресурсы, описываемые метаданными. Свойства метаданных: относительный характер метаданных; многоуровневость метаданных; ресурсы, описываемые метаданными; автономные/встроенные метаданные; метаданные независимые и зависимые от контента описываемых информационных ресурсов; метаданные независимые и зависимые от предметной области; системные и пользовательские метаданные; структурированные и неструктурированные метаданные; уровень гранулярности описания ресурсов; статические и динамические метаданные; степень формализованности метаданных; зависимость метаданных от информационной архитектуры системы; явное и неявное представление метаданных; средства представления метаданных.

Функции метаданных. Универсальные: описание моделей предметной области; описание источников данных; описание информационных объектов; описание семантики информационных источников, информационных объектов или их фрагментов; описание представления данных на разных уровнях информационной архитектуры; идентификация описываемых ресурсов; управление структурированными данными; верификация данных; предоставление пользователям сведений об описываемых ресурсах; управление доступом; описание характеристик пользователей и их информационных потребностей; обеспечение обнаружения и поиска информационных ресурсов; распространение информационных ресурсов; тематическая систематизация информационных ресурсов; обеспечение анализа данных; обеспечение интероперабельности и повторного использования информационных ресурсов; интеграция данных из многих источников; описание авторских прав на интеллектуальную собственность; наукометрия в научных электронных библиотеках. Специфические функции.

Классификации метаданных. Средства представления, стандартизация метаданных.

Тема 1.2 Форматы метаданных (8 часов, в том числе: 2 часа – лекция, 2 часа – практические занятия, 4 часа – самостоятельная работа)

Стандарты метаданных, созданных для описания различных видов документированной информации: архивной (AIF, CWM, DDI и др.); музейной (CDWA, MIDAS, Object ID и др.) «Ядро Ассоциации Визуальных Ресурсов» – Visual Resources Association Core, или VRA Core – для описания музейных объектов.; библиографической (MARC, RUSMARC); научной (CERIF, GRIB, VICAR и др.) и др.

Стандарты метаданных в области культурного наследия. Библиотеки: RUSMARC, UNIMARC, Концептуальные модели IFLA (FRBR, FRAD, FRASD, LRM), METS (Metadata Encoding and Transmission Standard), MODS (Metadata Object Description Schema), BIBFRAME.

Архивы: EAD (Encoded Archival Description), ISAD (G) (General International Standard Archival Description), ISAAR (CPF) (International Standard Archival Authority Record for Corporate Bodies, Persons and Families).

Музеи: CDWA (Categories for the Description of Works of Art), CDWA Lite, LIDO (Lightweight Information Describing Objects).

Раздел 2 Формирование метаданных

Тема 2.1 Метаданные цифровых ресурсов (8 часов, в том числе: 4 часа – практические занятия, 4 часа – самостоятельная работа)

Стандарты метаданных. Универсальные схемы метаданных: «Дублинское ядро» (Dublin Core), CIDOC-CRM (CIDOC Conceptual Reference Model), VRA Core (Visual Resource Association Core Categories). «Схема метаданных для описания объектов» (Metadata Object Description Scheme, или MODS). Стандарт METS (Metadata Encoding and Transmission Standard – Стандарт кодирования и передачи метаданных).

Семантическая сеть и роль метаданных в ней.

Электронные информационные ресурсы: определение понятия, классификация. Электронный документ. Цифровые коллекции. Веб-сайты. Система метаданных для описания электронных информационных ресурсов. ГОСТ Р ИСО 489-1–2019. Информация и документация. Управление документами. Часть 1. Понятия и принципы. (ISO 15489-1:2016, Information and documentation – Records management – Part 1: Concepts and principles, IDT).

Система управления электронным контентом DSpace. Репозитории учреждений.

Авторские профили в наукометрических базах данных. Идентификаторы современной научной коммуникации: DOI (Digital Object Identifier) – цифровой идентификатор объекта, ORCID (Open Researcher and Contributor ID) – открытый идентификатор автора и исследователя и ROR (Research Organization Registry) – реестр исследовательских организаций.

Тема 2.2 Метаданные объектов национальной библиографии (8 часов, в том числе: 2 часа – практические занятия, 6 часов – самостоятельная работа)

Электронные ресурсы как объект национальной библиографии. Критерии отбора электронных ресурсов.

Роль ИФЛА в организации учета электронных ресурсов национальной библиографией. «Функциональных требований к библиографическим записям» (FRBR). «Функциональные требования к авторитетным данным» (FRAD). «Международный стандарт библиографического описания» (ISBD). Национальные правила каталогизации. RDA – «Описание ресурса и доступ к

нему». Стандарты идентификации. Международный стандартный книжный номер (ISBN) (<http://www.isbn-international.org/>). Международный стандартный номер сериального издания (ISSN) (<http://www.issn.org/>). Международный стандартный номер издания музыкального произведения (ISMN) (<http://ismninternational.org/>). Связывающий ISSN (ISSN-L). ISTC или Международный стандартный код текста. ISAN или Международный стандартный номер аудиовизуального документа. ISWC или Международный стандартный код музыкального произведения. ISRC или Международный стандартный код аудио/видео записи. ISNI/ISPI или Международный стандартный идентификатор имени лица / организации. NBN или Регистрационный номер в системе национальной библиографии. Регистр NSDL (Национальной научной электронной библиотеки).

Авторитетный контроль. Виртуальный международный авторитетный файл (VIAF).

Библиографические стандарты и семантика Интернета (Семантический Web). RDF(Инфраструктура описания ресурса). OWL (Онтологический язык Web).

Обеспечение доступа к электронным ресурсам национальной библиографии. Пользователи и использование национальной библиографии: элементы метаданных. Уровень метаданных, описывающих электронный ресурс: авторитетный, полный, расширенный, базовый. Значимость электронного ресурса для национальной библиографии. Совместимость библиографического формата. Форматы MARC (Стандарты MARC, Библиотека Конгресса США UNIMARC, MARC 21). Форматы XML. Библиотеки MARCXML, MarcXchange. В издательском деле и книготорговле – ONIX (Online Information eXchange) (Формат электронного обмена информацией в книжной отрасли). Стандарт ISO (ISO 15836) Dublin Core («Дублинское ядро»). Совместимость протокола: Z39.50 (ISO 23950), SRU/SRW – SRU (Поиск/Извлечение с помощью URL), SRW (Веб-сервис извлечения при поиске), OAI-PMH (Протокол Проекта открытого архива для сбора метаданных, постоянные идентификаторы, унифицированное наименование ресурса (URN), The Handle System (Система идентификационных дескрипторов), идентификатор цифрового объекта (DOI), постоянный URL (PURL), код архивированного ресурса (ARK). Организация и управление системами национальной библиографии. Взаимодействие с издателями: интеграция и совместное использование метаданных.

Сводный электронный информационный ресурс (ЭИР) национальной библиографии Беларуси. Методика каталогизации национальных электронных документов сетевого распространения Беларуси.

МАТЕРИАЛЫ ДЛЯ АУДИТОРНОЙ РАБОТЫ С ПРИМЕНЕНИЕМ АКТИВНЫХ ФОРМ ОБУЧЕНИЯ

Круглый стол

Тема 1.1 Метаданные: терминосистема, свойства, функции, классификация (2 часа)

Метаданные: общая характеристика.

Цель: освоение терминосистемы, свойств, функций, классификаций метаданных.

Задача: формирование системных представлений о метаданных.

Порядок взаимодействия:

1. Открытие круглого стола.
2. Презентация темы, целей и задач круглого стола.
3. Презентация вопросов круглого стола.
4. Обсуждение вопросов круглого стола.
5. Подведение итогов круглого стола.

Вопросы для обсуждения:

3. Определение понятия “метаданные”, их свойства, классификация, функции.
4. Ресурсы, описываемые метаданными.
5. Средства представления, стандартизация метаданных.

МАТЕРИАЛЫ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ

Раздел 1 Основа теории метаданных

Тема 1.1 Метаданные: терминосистема, свойства, функции, классификация

Задание (2 часа)

Подготовить материалы, необходимые для обсуждения на круглом столе, по следующим вопросам:

4. Метаданные, их свойства, классификация, функции.
5. Ресурсы, описываемые метаданными.
6. Средства представления, стандартизация метаданных.

Литература для самостоятельного изучения: [2, с. 19–20; 7, с. 7–11; 25, с. 59–77; 26, с. 39–51].

Требования к проверке результатов самостоятельной работы. Проверка результатов самостоятельной работы осуществляется при проведении самоконтроля (вопросы № 1–7).

Тема 1.2 Форматы метаданных

Задание (4 часа)

Изучить с использованием рекомендуемой литературы следующие вопросы:

5. Стандарты метаданных для описания архивных документов.
6. Стандарты метаданных для описания музейных предметов.
7. Стандарты метаданных для описания информационных ресурсов библиотек.
8. Стандарты метаданных в области культурного наследия.

Литература для самостоятельного изучения: [3, с. 12–17; 4, с. 63–71; 15, с. 37–53].

Требования к проверке результатов самостоятельной работы. Проверка результатов самостоятельной работы осуществляется при проведении при проведении самоконтроля (вопросы № 16–21).

Раздел 2 Формирование метаданных

Тема 2.1 Метаданные цифровых ресурсов

Задание (4 часа)

Изучить с использованием рекомендуемой литературы следующие вопросы:

1. Универсальные стандарты метаданных: «Дублинское ядро» (Dublin Core), CIDOC-CRM (CIDOC Conceptual Reference Model), VRA Core (Visual Resource Association Core Categories). «Схема метаданных для описания объектов» (Metadata Object Description Scheme, или MODS). Стандарт METS (Metadata Encoding and Transmission Standard – Стандарт кодирования и передачи метаданных).

2. Семантическая сеть и роль метаданных в ней.

3. Электронные информационные ресурсы: определение понятия, классификация. Электронный документ. Цифровые коллекции. Веб-сайты. Система метаданных для описания электронных информационных ресурсов. ГОСТ Р ИСО 489-1–2019. Информация и документация. Управление документами. Часть 1. Понятия и принципы. (ISO 15489-1:2016, Information and documentation – Records management – Part 1: Concepts and principles, IDT).

4. Проанализировать и привести пример метаданных платформы DSpace репозитория учреждения образования Беларуси (по выбору).

5. Авторские профили в наукометрических базах данных. Идентификаторы современной научной коммуникации: DOI (Digital Object Identifier) – цифровой идентификатор объекта, ORCID (Open Researcher and Contributor ID) – открытый идентификатор автора и исследователя и ROR (Research Organization Registry) – реестр исследовательских организаций.

Литература для самостоятельного изучения: [8, с. 3–16; 20, с. 26–27; 30, с. 24–38; 32, с. 8].

Требования к проверке результатов самостоятельной работы. Проверка результатов самостоятельной работы осуществляется при проведении при проведении самоконтроля (вопросы № 8–16; 22–26).

Тема 2.2 Метаданные объектов национальной библиографии

Задание (6 часов)

Изучить с использованием рекомендуемой литературы следующие вопросы:

7. Электронные ресурсы как объект национальной библиографии.

8. Роль ИФЛА в организации учета электронных ресурсов национальной библиографией.

9. Стандарты национальной библиографии представления метаданных электронных ресурсов.

10. Авторитетный контроль. Виртуальный международный авторитетный файл (VIAF).

11. Библиографические стандарты и семантика Интернета (Семантический Web).

12. Представить по титульному листу на разные виды национальных электронных документов сетевого распространения Беларуси и расписать библиографические сведения метками полей и подполей в соответствии со структурой BELMARC-формата.

Литература для самостоятельного изучения: [1, с. 135–152; 17, с. 37–43; 28, с. 168–176].

Требования к проверке результатов самостоятельной работы. Проверка результатов самостоятельной работы осуществляется при проведении при проведении самоконтроля (вопросы № 27–31).

Вопросы для самоконтроля по дисциплине «Метаданные»

32. Определение понятия «метаданные».
33. Ресурсы, описываемые метаданными.
34. Свойства метаданных.
35. Универсальные функции метаданных.
36. Специфические функции метаданных.
37. Классификации метаданных.
38. Средства представления, стандартизация метаданных.
39. «Электронные информационные ресурсы»: общая характеристика.
40. ГОСТ 7.82–2001. «Библиографическая запись. Библиографическое описание электронных ресурсов. Общие требования и правила составления»: общая характеристика.
41. ГОСТ Р ИСО 489-1–2019. «Информация и документация. Управление документами. Часть 1. Понятия и принципы»: понятие, принципы, требования к метаданным документам.
42. СТБ 7.38–2017. «Электронные издания. Основные виды и выходные сведения»: общая характеристика.
43. ГОСТ 7.70–2003. «Описание баз данных и машиночитаемых информационных массивов. Состав и обозначение характеристик»: система метаданных для описания электронных информационных ресурсов.
44. Цифровые коллекции как объект формирования метаданных.
45. ГОСТ Р 7.0.96–2016. «Электронные библиотеки. Основные виды. Структура. Технология формирования»: характеристика метаданных.
46. Метаданные для описания веб-сайтов.
47. Универсальные стандарты метаданных.
48. Стандарты метаданных для описания архивных документов.
49. Стандарты метаданных для описания музейных предметов.
50. Стандарты метаданных для описания информационных ресурсов библиотек.
51. Стандарты метаданных в области культурного наследия.
52. BIBFRAME: общая характеристика.
53. Дублинское ядро (DC, Dublin Core): определение, структура, стандартизация.
54. Дублинское ядро (DC, Dublin Core): набор элементов метаданных в соответствии со стандартом ГОСТ Р 7.0.10–2019 (ИСО 15836–1:2017). Набор элементов метаданных «Дублинское ядро». Основные (ядерные) элементы (ISO 15836-1:2017, Information and documentation – The Dublin Core metadata element set – Part 1: Core elements, MOD).
55. DSpace как система формирования депозитария. Виды метаданных.
56. Авторские профили в наукометрических базах данных.
57. Идентификаторы научной коммуникации: DOI, ORCID, ROR. Значение, структура.

58. Роль ИФЛА в организации учета электронных ресурсов национальной библиографией.

59. Библиографические стандарты и семантика Интернета (Семантический веб).

60. Виртуальный международный авторитетный файл (VIAF).

61. Уровни метаданных национальных электронных ресурсов: авторитетный, библиографический.

62. Методика каталогизации национальных электронных документов сетевого распространения Беларуси.

СПИСОК РЕКОМЕНДУЕМОЙ ЛИТЕРАТУРЫ

Основная

1. Нацыянальныя дакументы Беларусі: методыка і тэхналогія сучаснай бібліятэчнай працы [Электронны рэсурс] : навукова-метадычнае выданне / Нацыянальная бібліятэка Беларусі. – Мінск : Нацыянальная бібліятэка Беларусі, 2019. – С. 135–152. – (Серыя «БібліяКансультант» ; вып. 7). – Режим доступа: <https://e-catalog.nlb.by/Record/BY-NLB-br0001533423>. – Дата доступа: 03.06.2022.

2. Штратникова, А. В. Библиография библиографии [Электронный ресурс] : учебник / А. В. Штратникова. – 2-е изд., испр. и доп. – М. : Юрайт, 2022. – С. 19–20 ; 74–79, 99–120. – Режим доступа: <https://e-catalog.nlb.by/Record/BY-RLST-kn-859508>. – Дата доступа: 03.06.2022.

Дополнительная

3. Балацкая, Н. М. Веб-архивирование как задача национальной и краеведческой библиографии [Электронный ресурс] / Н. М. Балацкая, М. Б. Мартиросова // Библиосфера. – 2021. – № 3. – С. 12–17. – Режим доступа: <https://doi.org/10.20913/1815-3186-2021-3-12-17>. – Дата доступа: 03.06.2022.

4. Бахтурина, Т. А. От MARC 21 к модели VIBFRAME: эволюция машиночитаемых форматов Библиотеки Конгресса США [Электронный ресурс] / Т. А. Бахтурина // Науч. и техн. б-ки. – 2018. – № 3. – С. 63–71. – Режим доступа: <https://ntb.gpntb.ru/jour/article/view/289/290>. – Дата доступа: 03.06.2022.

5. Белорусский коммуникативный формат представления авторитетных/нормативных записей в машиночитаемой форме (BELMARC/Authorities), 2-я редакция (2013) [Электронный ресурс]. – 2-я редакция. – [Мінск : Нацыянальная бібліятэка Беларусі, 2013]. – Режим доступа: <https://nlb.by/content/bibliotekaryam/belmarc/belmarc-authorities/>. – Дата доступа: 03.06.2022.

6. Белорусский коммуникативный формат представления библиографических записей в машиночитаемой форме (BELMARC) [Электронный ресурс]. – 2-я редакция. – [Мінск : Нацыянальная бібліятэка Беларусі, 2012]. – Режим доступа: <https://nlb.by/content/bibliotekaryam/belmarc/>. – Дата доступа: 03.06.2022.

7. Берестова, Т. Ф. Что скрывается за термином «метаданные»? [Электронный ресурс] / Т. Ф. Берестова // Вестник культуры и искусств. – 2017. – № 1 (49). – С. 7–11. – Режим доступа: <https://cyberleninka.ru/article/n/chto-skryvaetsya-za-terminom-metadannye>. – Дата доступа: 03.06.2022.

8. Библиографическая информация в цифровой культуре: о дискуссионных проблемах и итогах III Международного библиографического конгресса [Электронный ресурс] / А. Е. Гуськов, О. Н. Жлобинская, Д. В. Косяков, Н. К. Леликова, О. Л. Лаврик //

Библиосфера. – 2021. – № 2. – С. 3–16. – Режим доступа: <https://doi.org/10.20913/1815-3186-2021-2-59-64>. – Дата доступа: 03.06.2022.

9. Взаимодействие веб-сайтов по культуре с пользователем. Рекомендации [Электронный ресурс] / под редакцией рабочей группы проекта MINERVA ЕС «Качество, доступность и удобство работы». – Режим доступа: <https://ifap.ru/library/book469.pdf>. – Дата доступа: 03.06.2022.

10. ГОСТ 7.1–2003. Библиографическая запись. Библиографическое описание. Общие требования и правила составления = Библиографічний запис. Бібліографічне описання. Агульні патрабаванні і правілы састаўлення. – Взамен ГОСТ 7.1–84, ГОСТ 7.16–79, ГОСТ 7.18–79, ГОСТ 7.34–81, ГОСТ 7.40–82 ; введ. 2004–11–01. – Минск, 2004. – 48 с.

11. ГОСТ 7.70–2003. Описание баз данных и машиночитаемых информационных массивов. Состав и обозначение характеристик. – Взамен ГОСТ 7.70–96 ; введ. 2004–07–01. – М. : Стандартинформ, 2008. – 11 с.

12. ГОСТ 7.82–2001. Библиографическая запись. Библиографическое описание электронных ресурсов. Общие требования и правила составления = Библиографічний запис. Бібліографічне описання электронных рэсурсаў. Агульні патрабаванні і правілы складання. – Введ. 2002–07–01. – Минск, 2002. – 23 с.

13. ГОСТ Р 7.0.10–2019 (ИСО 15836–1:2017) [Электронный ресурс]. Набор элементов метаданных «Дублинское ядро». Основные (ядерные) элементы (ISO 15836-1:2017, Information and documentation – The Dublin Core metadata element set – Part 1: Core elements, MOD) [Электронный ресурс]. – Взамен ГОСТ Р 7.0.10-2010 (ИСО 15836:2003) ; введ. 2020–01–01. – Режим доступа: <https://www.ifap.ru/library/gost/70102019.pdf>. – Дата доступа: 03.06.2022.

14. ГОСТ Р 7.0.96–2016. Электронные библиотеки. Основные виды. Структура. Технология формирования [Электронный ресурс]. – Введ. 2017–07–01. – Режим доступа: <https://ifap.ru/library/gost/70962016.pdf>. – Дата доступа: 03.06.2022.

15. Грузова, А. А. Библиографическая деятельность в сетевой информационной среде [Электронный ресурс] / А. А. Грузова // Научные и технические библиотеки. – 2019. – № 7. – С. 37–53. – Режим доступа: <https://ntb.gpntb.ru/jour/article/view/453/443>. – Дата доступа: 03.06.2022.

16. Когаловский, М. Р. Метаданные, их свойства, функции, классификация и средства представления [Электронный ресурс] / М. Р. Когаловский. – Режим доступа: <http://ceur-ws.org/Vol-934/paper3.pdf> <https://doi.org/10.20913/1815-3186-2021-3-12-17>. – Дата доступа: 03.06.2022.

17. Леликова, Н. К. Национальная библиография в цифровую эпоху: международная регламентация и отечественный опыт [Электронный ресурс] / Н. К. Леликова // Библиосфера. – 2021. – № 2. – С. 37–43. – Режим доступа: <https://doi.org/10.20913/13/1815-3186-2021-2-37-43>. – Дата доступа: 03.06.2022.

18. Национальная библиография в электронную эру: руководство и новые направления развития / рабочая группа ИФЛА по созданию руководств, регламентирующих работу систем национальной библиографии ;

редактор Майя Жумер [Электронный ресурс]. – Мюнхен : К. Г. Заур, 2009. – (Серия публикаций ИФЛА по библиографическому учету ; том 39). – 132 с. – Режим доступа: <https://repository.ifla.org/bitstream/123456789/903/1/ifla-guidelines-national-bibliographies-electronic-age-ru.pdf>. – Дата доступа: 03.06.2022.

19. Нормативные документы и методические материалы [Электронный ресурс] // Президентская библиотека имени Б. Н. Ельцина. – Режим доступа: https://www.prilib.ru/contentprocessing#systema_SIBID. – Дата доступа: 03.06.2022.

20. Петрашкевич, А. Региональный сводный каталог: содержание и технология формирования / А. Петрашкевич, Н. Шамко // *Бібліятэчны свет*. – 2017. – № 1. – С. 26–27.

21. Положение об обязательном бесплатном экземпляре документов [Электронный ресурс] : от 03.09.2008 г. № 1284 (в ред. от 10.08.2009 г. № 1052). – Режим доступа: <https://natbook.org.by/index.php?id=25>. – Дата доступа: 03.06.2022.

22. Принципы качества веб-сайтов по культуре. Руководство [Электронный ресурс] / под ред. пятой рабочей группы проекта Minerva «Определение потребностей пользователей, содержания и критериев качества веб-сайтов по культуре». – М., 2006. – 62 с. – Режим доступа: http://www.ifarcom.ru/files/publications/Kachestvo_websaitov_po_kulture.pdf. – Дата доступа: 03.06.2022.

23. Предеина, А. В. Базы данных библиотек: актуальные вопросы создания и использования [Электронный ресурс] / А. В. Предеина // *Бібліяпанарама*. Вып. 22 / Брэсцкая абласная бібліятэка імя М. Горкага, Адзел бібліятэказнаўства ; [складанне і рэдакцыя: Л. К. Кароль, А. С. Кузьміна]. – Мінск : БудМедыяПраект, 2019. – С. 102–106. – Режим доступа: http://www.brl.by/resursy/pechatnye-izdaniya/bibliopanorama_22.pdf. – Дата доступа: 03.06.2022.

24. Предеина, А. В. Региональные сводные электронные каталоги как корпоративный информационный ресурс библиотек Беларуси: вопросы отражения краеведческого компонента [Электронный ресурс] / А. В. Предеина, А. И. Федорина // *Культура: теория и практика* / Московский государственный институт культуры. – 2021. – № 3. – Режим доступа: <https://cyberleninka.ru/article/n/regionalnye-svodnye-elektronnye-katalogi-kak-korporativnyu-informatsionnyu-resurs-bibliotek-belarusi-voprosy-otrazheniya/viewer>. – Дата доступа: 03.06.2022.

25. Редькина, Н. С. Компетенции библиографа: непрерывность развития и точки разрыва [Электронный ресурс] / Н. С. Редькина // *Библиосфера*. – 2021. – № 4. – С. 59–77. – Режим доступа: <https://doi.org/10.20913/1815-3186-2021-4-59-77>. – Дата доступа: 03.06.2022.

26. Романовски, С. Сравнительный анализ эволюции процессов каталогизации и информационных технологий в направлении создания следующего поколения библиотечных систем автоматизации / С. Романовски // *Научные и технические библиотеки*. – 2017. – № 6. – С. 39–51.

27. Справочник библиотекаря. Книга 1. Интерактивные и мультимедийные технологии / авторы: Н. П. Опарина, С. В. Савкина, Е. В. Роот ; под ред. И. С. Пилко. – 5-е издание. – Санкт-Петербург : Профессия, 2021. – С. 43–104.

28. Справочник библиотекаря. Книга 2. Грузова, А. А. Управление проектами в библиотечно-информационной деятельности / А. А. Грузова. – Пятое издание. – Санкт-Петербург : Профессия, 2021. – С. 168–176.

29. Стандарты метаданных в области культурного наследия [Электронный ресурс] // Президентская библиотека имени Б. Н. Ельцина. – Режим доступа: https://www.prilib.ru/metadata_standards. – Дата доступа: 03.06.2022.

30. Стегаева, М. В. Каталогизация в цифровую эпоху: практика Президентской библиотеки им. Б. Н. Ельцина [Электронный ресурс] / М. В. Стегаева // Научные и технические библиотеки. – 2020. – № 2. – С. 24–38. – Режим доступа: <https://doi.org/10.33186/1027-3689-2020-2-24-38>. – Дата доступа: 03.06.2022.

31. СТБ 7.38–2017. Электронные издания. Основные виды и выходные сведения = Электронныя выданні. Асноўныя віды і выхадныя звесткі. – Введ. 2017–10–01.– Минск, 2017. – 16 с.

32. СТБ 7.53–2019. Издания. Международный стандартный книжный номер. Использование и издательское оформление = Выданні. Міжнародны стандартны кніжны нумар. Выкарыстанне і выдавецкае афармленне. – Взамен ГОСТ 7.53–2008; введ. 2019–11–01. – Минск, 2019. – 8 с.

Нарматыўныя прававыя акты

33. Кодэкс Рэспублікі Беларусь аб культуры [Электронны рэсурс] : 20 ліпеня 2016 г., № 413-3 : прыняты Палатай прадстаўнікоў 24 чэрвеня 2016 г. : адобр. Саветам Рэсп. 30 чэрвеня 2016 г. – Режим доступа: <https://pravo.by/document/?guid=12551&p0=Hk1600413>. – Дата доступа: 03.06.2022.

34. Об авторском праве и смежных правах [Электронный ресурс] : Закон Респ. Беларусь, 17 мая 2011 г., № 262-З. – Режим доступа: <https://etalonline.by/document/?regnum=h11100262>. – Дата доступа: 03.06.2022.

35. Об информации, информатизации и защите информации [Электронный ресурс] : Закон Респ. Беларусь, 10 нояб. 2008 г., № 455-З. – Режим доступа: <https://pravo.by/document/?guid=3871&p0=h10800455>. – Дата доступа: 03.06.2022.

36. Государственная программа «Культура Беларуси» на 2021–2025 гг. [Электронный ресурс]. – Режим доступа: https://pravo.by/upload/docs/op/C22100053_1612558800.pdf. – Дата доступа: 03.06.2022.

5.2 Основная литература

1. Нацыянальныя дакументы Беларусі: методыка і тэхналогія сучаснай бібліятэчнай працы [Электронны рэсурс] : навукова-метадычнае выданне / Нацыянальная бібліятэка Беларусі. – Мінск : Нацыянальная бібліятэка Беларусі, 2019. – С. 135–152. – (Серыя «БібліяКансультант» ; вып. 7). – Режим доступа: <https://e-catalog.nlb.by/Record/BY-NLB-br0001533423>. – Дата доступа: 03.06.2022.

2. Штратникова, А. В. Библиография библиографии [Электронный ресурс] : учебник / А. В. Штратникова. – 2-е изд., испр. и доп. – М. : Юрайт, 2022. – С. 19–20 ; 74–79, 99–120. – Режим доступа: <https://e-catalog.nlb.by/Record/BY-RLST-kn-859508>. – Дата доступа: 03.06.2022.

5.3 Дополнительная литература

3. Балацкая, Н. М. Веб-архивирование как задача национальной и краеведческой библиографии [Электронный ресурс] / Н. М. Балацкая, М. Б. Мартыросова // Библиосфера. – 2021. – № 3. – С. 12–17. – Режим доступа: <https://doi.org/10.20913/1815-3186-2021-3-12-17>. – Дата доступа: 03.06.2022.

4. Бахтурина, Т. А. От MARC 21 к модели BIBFRAME: эволюция машиночитаемых форматов Библиотеки Конгресса США [Электронный ресурс] / Т. А. Бахтурина // Науч. и техн. б-ки. – 2018. – № 3. – С. 63–71. – Режим доступа: <https://ntb.gpntb.ru/jour/article/view/289/290>. – Дата доступа: 03.06.2022.

5. Белорусский коммуникативный формат представления авторитетных/нормативных записей в машиночитаемой форме (BELMARC/Authorities), 2-я редакция (2013) [Электронный ресурс]. – 2-я редакция. – [Мінск : Нацыянальная бібліятэка Беларусі, 2013]. – Режим доступа: <https://nlb.by/content/bibliotekaryam/belmarc/belmarc-authorities/>. – Дата доступа: 03.06.2022.

6. Белорусский коммуникативный формат представления библиографических записей в машиночитаемой форме (BELMARC) [Электронный ресурс]. – 2-я редакция. – [Мінск : Нацыянальная бібліятэка Беларусі, 2012]. – Режим доступа: <https://nlb.by/content/bibliotekaryam/belmarc/>. – Дата доступа: 03.06.2022.

7. Берестова, Т. Ф. Что скрывается за термином «метаданные»? [Электронный ресурс] / Т. Ф. Берестова // Вестник культуры и искусств. – 2017. – № 1 (49). – С. 7–11. – Режим доступа: <https://cyberleninka.ru/article/n/chto-skryvaetsya-za-terminom-metadannye>. – Дата доступа: 03.06.2022.

8. Библиографическая информация в цифровой культуре: о дискуссионных проблемах и итогах III Международного библиографического конгресса [Электронный ресурс] / А. Е. Гуськов, О. Н. Жлобинская, Д. В. Косяков, Н. К. Леликова, О. Л. Лаврик //

Библиосфера. – 2021. – № 2. – С. 3–16. – Режим доступа: <https://doi.org/10.20913/1815-3186-2021-2-59-64>. – Дата доступа: 03.06.2022.

9. Взаимодействие веб-сайтов по культуре с пользователем. Рекомендации [Электронный ресурс] / под редакцией рабочей группы проекта MINERVA ЕС «Качество, доступность и удобство работы». – Режим доступа: <https://ifap.ru/library/book469.pdf>. – Дата доступа: 03.06.2022.

10. ГОСТ 7.1–2003. Библиографическая запись. Библиографическое описание. Общие требования и правила составления = Библиографічний запис. Бібліографічне апісанне. Агульні патрабаванні і правілы састаўлення. – Взамен ГОСТ 7.1–84, ГОСТ 7.16–79, ГОСТ 7.18–79, ГОСТ 7.34–81, ГОСТ 7.40–82 ; введ. 2004–11–01. – Минск, 2004. – 48 с.

11. ГОСТ 7.70–2003. Описание баз данных и машиночитаемых информационных массивов. Состав и обозначение характеристик. – Взамен ГОСТ 7.70–96 ; введ. 2004–07–01. – М. : Стандартинформ, 2008. – 11 с.

12. ГОСТ 7.82–2001. Библиографическая запись. Библиографическое описание электронных ресурсов. Общие требования и правила составления = Библиографічний запис. Бібліографічне апісанне электронных рэсурсаў. Агульні патрабаванні і правілы складання. – Введ. 2002–07–01. – Минск, 2002. – 23 с.

13. ГОСТ Р 7.0.10–2019 (ИСО 15836–1:2017) [Электронный ресурс]. Набор элементов метаданных «Дублинское ядро». Основные (ядерные) элементы (ISO 15836-1:2017, Information and documentation – The Dublin Core metadata element set – Part 1: Core elements, MOD) [Электронный ресурс]. – Взамен ГОСТ Р 7.0.10-2010 (ИСО 15836:2003) ; введ. 2020–01–01. – Режим доступа: <https://www.ifap.ru/library/gost/70102019.pdf>. – Дата доступа: 03.06.2022.

14. ГОСТ Р 7.0.96–2016. Электронные библиотеки. Основные виды. Структура. Технология формирования [Электронный ресурс]. – Введ. 2017–07–01. – Режим доступа: <https://ifap.ru/library/gost/70962016.pdf>. – Дата доступа: 03.06.2022.

15. Грузова, А. А. Библиографическая деятельность в сетевой информационной среде [Электронный ресурс] / А. А. Грузова // Научные и технические библиотеки. – 2019. – № 7. – С. 37–53. – Режим доступа: <https://ntb.gpntb.ru/jour/article/view/453/443>. – Дата доступа: 03.06.2022.

16. Когаловский, М. Р. Метаданные, их свойства, функции, классификация и средства представления [Электронный ресурс] / М. Р. Когаловский. – Режим доступа: <http://ceur-ws.org/Vol-934/paper3.pdf> <https://doi.org/10.20913/1815-3186-2021-3-12-17>. – Дата доступа: 03.06.2022.

17. Леликова, Н. К. Национальная библиография в цифровую эпоху: международная регламентация и отечественный опыт [Электронный ресурс] / Н. К. Леликова // Библиосфера. – 2021. – № 2. – С. 37–43. – Режим доступа: <https://doi.org/10.20913/13/1815-3186-2021-2-37-43>. – Дата доступа: 03.06.2022.

18. Национальная библиография в электронную эру: руководство и новые направления развития / рабочая группа ИФЛА по созданию руководств, регламентирующих работу систем национальной библиографии ;

редактор Майя Жумер [Электронный ресурс]. – Мюнхен : К. Г. Заур, 2009. – (Серия публикаций ИФЛА по библиографическому учету ; том 39). – 132 с. – Режим доступа: <https://repository.ifla.org/bitstream/123456789/903/1/ifla-guidelines-national-bibliographies-electronic-age-ru.pdf>. – Дата доступа: 03.06.2022.

19. Нормативные документы и методические материалы [Электронный ресурс] // Президентская библиотека имени Б. Н. Ельцина. – Режим доступа: https://www.prilib.ru/contentprocessing#systema_SIBID. – Дата доступа: 03.06.2022.

20. Петрашкевич, А. Региональный сводный каталог: содержание и технология формирования / А. Петрашкевич, Н. Шамко // *Бібліятэчны свет*. – 2017. – № 1. – С. 26–27.

21. Положение об обязательном бесплатном экземпляре документов [Электронный ресурс] : от 03.09.2008 г. № 1284 (в ред. от 10.08.2009 г. № 1052). – Режим доступа: <https://natbook.org.by/index.php?id=25>. – Дата доступа: 03.06.2022.

22. Принципы качества веб-сайтов по культуре. Руководство [Электронный ресурс] / под ред. пятой рабочей группы проекта Minerva «Определение потребностей пользователей, содержания и критериев качества веб-сайтов по культуре». – М., 2006. – 62 с. – Режим доступа: http://www.ifarcom.ru/files/publications/Kachestvo_websaitov_po_kulture.pdf. – Дата доступа: 03.06.2022.

23. Предеина, А. В. Базы данных библиотек: актуальные вопросы создания и использования [Электронный ресурс] / А. В. Предеина // *Бібліяпанарама*. Вып. 22 / Брэсцкая абласная бібліятэка імя М. Горкага, Адзел бібліятэказнаўства ; [складанне і рэдакцыя: Л. К. Кароль, А. С. Кузьміна]. – Мінск : БудМедыяПраект, 2019. – С. 102–106. – Режим доступа: http://www.brl.by/resursy/pechatnye-izdaniya/bibliopanorama_22.pdf. – Дата доступа: 03.06.2022.

24. Предеина, А. В. Региональные сводные электронные каталоги как корпоративный информационный ресурс библиотек Беларуси: вопросы отражения краеведческого компонента [Электронный ресурс] / А. В. Предеина, А. И. Федорина // *Культура: теория и практика* / Московский государственный институт культуры. – 2021. – № 3. – Режим доступа: <https://cyberleninka.ru/article/n/regionalnye-svodnye-elektronnye-katalogi-kak-korporativnyu-informatsionnyu-resurs-bibliotek-belarusi-voprosy-otrazheniya/viewer>. – Дата доступа: 03.06.2022.

25. Редькина, Н. С. Компетенции библиографа: непрерывность развития и точки разрыва [Электронный ресурс] / Н. С. Редькина // *Библиосфера*. – 2021. – № 4. – С. 59–77. – Режим доступа: <https://doi.org/10.20913/1815-3186-2021-4-59-77>. – Дата доступа: 03.06.2022.

26. Романовски, С. Сравнительный анализ эволюции процессов каталогизации и информационных технологий в направлении создания следующего поколения библиотечных систем автоматизации / С. Романовски // *Научные и технические библиотеки*. – 2017. – № 6. – С. 39–51.

27. Справочник библиотекаря. Книга 1. Интерактивные и мультимедийные технологии / авторы: Н. П. Опарина, С. В. Савкина, Е. В. Роот ; под ред. И. С. Пилко. – 5-е издание. – Санкт-Петербург : Профессия, 2021. – С. 43–104.

28. Справочник библиотекаря. Книга 2. Грузова, А. А. Управление проектами в библиотечно-информационной деятельности / А. А. Грузова. – Пятое издание. – Санкт-Петербург : Профессия, 2021. – С. 168–176.

29. Стандарты метаданных в области культурного наследия [Электронный ресурс] // Президентская библиотека имени Б. Н. Ельцина. – Режим доступа: https://www.prilib.ru/metadata_standards. – Дата доступа: 03.06.2022.

30. Стегаева, М. В. Каталогизация в цифровую эпоху: практика Президентской библиотеки им. Б. Н. Ельцина [Электронный ресурс] / М. В. Стегаева // Научные и технические библиотеки. – 2020. – № 2. – С. 24–38. – Режим доступа: <https://doi.org/10.33186/1027-3689-2020-2-24-38>. – Дата доступа: 03.06.2022.

31. СТБ 7.38–2017. Электронные издания. Основные виды и выходные сведения = Электронныя выданні. Асноўныя віды і выхадныя звесткі. – Введ. 2017–10–01.– Минск, 2017. – 16 с.

32. СТБ 7.53–2019. Издания. Международный стандартный книжный номер. Использование и издательское оформление = Выданні. Міжнародны стандартны кніжны нумар. Выкарыстанне і выдавецкае афармленне. – Взамен ГОСТ 7.53–2008; введ. 2019–11–01. – Минск, 2019. – 8 с.

Нарматыўныя прававыя акты

33. Кодэкс Рэспублікі Беларусь аб культуры [Электронны рэсурс] : 20 ліпеня 2016 г., № 413-З : прыняты Палатай прадстаўнікоў 24 чэрвеня 2016 г. : адобр. Саветам Рэсп. 30 чэрвеня 2016 г. – Режим доступа: <https://pravo.by/document/?guid=12551&p0=Hk1600413>. – Дата доступа: 03.06.2022.

34. Об авторском праве и смежных правах [Электронный ресурс] : Закон Респ. Беларусь, 17 мая 2011 г., № 262-З. – Режим доступа: <https://etalonline.by/document/?regnum=h11100262>. – Дата доступа: 03.06.2022.

35. Об информации, информатизации и защите информации [Электронный ресурс] : Закон Респ. Беларусь, 10 нояб. 2008 г., № 455-З. – Режим доступа: <https://pravo.by/document/?guid=3871&p0=h10800455>. – Дата доступа: 03.06.2022.

36. Государственная программа «Культура Беларуси» на 2021–2025 гг. [Электронный ресурс]. – Режим доступа: https://pravo.by/upload/docs/op/C22100053_1612558800.pdf. – Дата доступа: 03.06.2022.