

METADATA ARE STRUCTURED
INFORMATION ABOUT
ANY KIND OF RESOURCE,
WHICH IS USED TO IDENTIFY,
DESCRIBE, MANAGE OR GIVE
ACCESS TO THAT RESOURCE.



www.athenaeurope.org • info@athenaeurope.org

DIGITISATION STANDARDS LANDSCAPE FOR EUROPEAN MUSEUMS, ARCHIVES, LIBRARIES

eContentplus



ATHENA
Access to cultural heritage
networks across Europe



Europeana
digital gateway
to Europe's cultural
heritage

Оцифровка:

ландшафт стандартов

для европейских музеев,

архивов, библиотек



MINISTERO
PER I BENI E
LE ATTIVITÀ
CULTURALI

Оцифровка: ландшафт стандартов для европейских музеев, архивов, библиотек

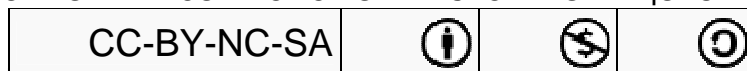
Публикация Рабочей группы 3 «Исследование стандартов и подготовка рекомендаций» проекта ATHENA.

Текст подготовлен Gordon McKenna, Collections Trust (UK); Chris De Loof, Royal Museums of Art and History (Belgium).

Тест этой брошюры является частью отчета проекта ATHENA и основан на анализе информационных ресурсов, которые партнеры проекта ATHENA предоставляют в Европейскую цифровую библиотеку Europeana, пользуясь сервисами, разработанными в проекте ATHENA.

Генеральный координатор проекта ATHENA
Rossella Caffo

Текст доступен на условиях лицензии Creative Commons Attribution Non-commercial Share Alike License¹, для некоммерческого использования, при условии указания авторства; распространение, копирование и модификация произведения разрешается, если модифицированные произведения также будут распространяться на условиях выполнения всех положений этой же лицензии.



Электронное издание на русском языке подготовлено Центром ПИК при финансовой поддержке Министерства культуры Российской Федерации.

Перевод: Н. Браккер
Редактор перевода: Л. Куйбышев
©2009 ATHENA EC Project
©2010 Центр ПИК

¹

http://ru.wikipedia.org/wiki/%D0%9B%D0%B8%D1%86%D0%B5%D0%BD%D0%B7%D0%B8%D0%B8_Creative_Commons (на русск. яз., прим. ред.)

Содержание

Предисловие к русскоязычному изданию	3
Предисловие.....	5
1. Основные положения	6
1.1. Стандарт	6
1.1.1. <i>Типы стандартов</i>	6
1.1.2. <i>Открытые стандарты</i>	7
1.2. Оцифровка	8
1.3. Совместимость	9
1.4. Метаданные	9
1.4.1. <i>Трудности дефиниции</i>	10
1.4.2. <i>Характеристики метаданных</i>	10
1.4.3. <i>«Более адекватная» дефиниция?</i>	13
2. Ландшафт стандартов.....	14
2.1. Описание стандартов.....	14
2.2. Схемы информации (метаданных)	15
2.2.1. <i>Музейные стандарты</i>	16
2.2.2. <i>Архивные стандарты</i>	18
2.2.3. <i>Библиотечные стандарты</i>	20
2.2.4. <i>Стандарты описания недвижимого наследия</i>	24
2.2.5. <i>Культурное наследие</i>	25
2.2.6. <i>Поиск ресурсов</i>	27
2.2.7. <i>Кодирование документов</i>	27
2.3. Мультимедиа форматы	28
2.3.1. <i>Текст</i>	28
2.3.2. <i>Изображение</i>	32
2.3.3. <i>Audio</i>	35
2.3.4. <i>Видео</i>	38
2.3.5. <i>Виртуальная реальность</i>	42
2.3.6. <i>Векторная графика</i>	44
2.4. Другие технические стандарты	45
2.4.1. <i>Поиск и вывод информации</i>	45
2.4.2. <i>Передача информации</i>	47
2.4.3. <i>Кодирование символов</i>	48
3. Выводы.....	50
3.1. Стандарты метаданных в сфере культуры	50
3.2. Технические стандарты.....	50
4. Указатель стандартов	51

Предисловие к русскоязычному изданию

Европейский проект ATHENA², инициированный партнёрами проектов серии MINERVA³, реализуется в рамках программы eContentplus в 2008 – 2011 гг.

Основная цель проекта – расширение контента европейской цифровой библиотеки EUROPEANA⁴, крупнейшего в Европе каталога информационных ресурсов по культуре. ATHENA собирает и объединяет (агрегирует) информационные ресурсы европейских музеев, библиотек, архивов и передает их в EUROPEANA для публикации. К началу 2011 года EUROPEANA уже предоставила открытый доступ к 14 миллионам информационных ресурсов по культуре. 1 757 029 информационных ресурсов поступило в EUROPEANA через проект ATHENA, который стал крупнейшим агрегатором информационных ресурсов для EUROPEANA.

EUROPEANA – это централизованный каталог культурного наследия, в нем хранятся и предоставляются пользователю так называемые «суррогаты» информационных объектов. Суррогаты содержат метаданные основных цифровых информационных ресурсов, необходимые для поиска, на том естественном языке, на котором они были предоставлены поставщиком информации. Метаданные сопровождаются небольшим изображением (иконкой) произведения искусства, обложки книги или документа, или небольшим аудио-видеофрагментом, а также ссылкой на основной ресурс на сайте источника информации – полное описание и большое изображение картины или фотографии, полный текст книги, аудио файл или оцифрованный фильм. Основные «полнотекстовые» ресурсы остаются на сайтах источников информации – музеев, библиотек, архивов; пользователь посмотрит суррогат и решит, хочет ли он обратиться по ссылке к сайту-держателю основного ресурса и получить доступ к полному тексту книги, фильму, аудиозаписи или большому изображению и описанию картины, рисунка, фотографии, документа.

Таким образом, в рамках проекта ATHENA требовалось разработать механизм сбора метаданных (а также изображений-иконок и/или аудио-видеофрагментов) из многочисленных и разнообразных учреждений культуры – поставщиков контента; привести собранную информацию к единообразию; трансформировать метаданные в формат централизованного каталога EUROPEANA и передать полученный материал на портал EUROPEANA для публикации и поиска.

Поэтому на первом этапе проекта ATHENA необходимо было понять, какими стандартами, форматами и протоколами пользуются учреждения культуры – потенциальные поставщики цифрового контента. Разработчиков ATHENA интересовали стандарты метаданных, форматы представления мультимедиа информации, протоколы передачи информации, которые используют учреждения культуры – поставщики контента. Большое значение имели также стандарты кодировки символов, так как ATHENA должна была агрегировать метаданные на 20 языках, в том числе, на языках, не использующих латинский алфавит.

В проекте ATHENA участвует более 120 учреждений культуры (в основном, музеи, но также и архивы, и библиотеки) из 21 страны (это страны – члены ЕС,

² www.athenaeurope.org

³ www.minervaplus.ru

⁴ www.europeana.eu

Россия и Израиль). Эти учреждения культуры приняли участие в исследовании, которое и легло в основу брошюры «Оцифровка: ландшафт стандартов для европейских музеев, архивов, библиотек».

Так как российские библиотеки не участвовали⁵ в проекте ATHENA, в «Ландшафт стандартов» не вошел российский коммуникативный формат предоставления библиографических данных RUSMARC⁶, который является национальной адаптацией формата UNIMARC⁷. UNIMARC также не упоминается в этой брошюре, так как не используется библиотеками европейских стран, принявшими участие в исследовании.

На втором этапе реализации проекта ATHENA для сбора информационных ресурсов и совмещения схем метаданных был разработан стандарт метаданных LIDO⁸, который не описан в «Ландшафте стандартов» просто потому, что он еще не был принят к моменту издания брошюры. LIDO был обнародован и одобрен на ежегодном заседании международного комитета по музейной документации CIDOC, которое состоялось на конференции международного комитета музеев ICOM в ноябре 2010 года в Шанхае.

LIDO (Lightweight Information Describing Objects) – это совместная инициатива создателей форматов CDWA Lite и museumdat, в которой сделана попытка объединить эти форматы и обеспечить совместимость с CIDOC-CRM и SPECTRUM.

LIDO хорошо зарекомендовал себя в проекте ATHENA и позволил агрегировать не только музейные, но и библиотечные, и архивные информационные ресурсы.

Разделы с описанием стандартов RUSMARC и LIDO добавлены к переводу оригинального текста редактором русскоязычного издания.

Брошюра «Оцифровка: ландшафт стандартов для европейских музеев, архивов, библиотек» издана рабочей группой проекта ATHENA в 2009 году. Электронное издание на русском языке подготовлено Центром по проблемам информатизации сферы культуры (Центром ПИК) в 2010 году при финансовой поддержке Министерства культуры Российской Федерации и адресовано специалистам российских музеев архивов, библиотек. Везде, где это было возможно, мы дополнили перевод ссылками на русскоязычные ресурсы, в которых разъясняются понятия и термины и описываются стандарты, форматы и протоколы, упомянутые в брошюре.

Мы надеемся, что эта брошюра поможет специалистам учреждений культуры, использующим информационные и коммуникационные технологии в своей повседневной работе, ориентироваться в многообразии понятий, стандартов, форматов и протоколов, без которых современная жизнь уже кажется невозможной.

Н.В. Браккер, главный специалист
Л.А. Куйбышев, генеральный директор
Центр по проблемам информатизации сферы культуры

⁵ Российская государственная библиотека и Российская национальная библиотека передают свои информационные ресурсы в EUROPEANA через проект «Европейская библиотека», см.

<http://www.rsl.ru/ru/s6/s373/s3734405/s37344054406/>

⁶ <http://www.rba.ru/rusmarc/>

⁷ <http://archive.ifla.org/VI/3/p1996-1/sec-uni.htm> (на англ. яз.)

⁸ <http://www.lido-schema.org> (на англ. яз.)

Предисловие

Цель проекта ATHENA – содействие музеям и другим учреждениям культуры в подготовке данных и передаче их для публикации в Europeana. Однако, известно, что музеи используют самые разнообразные стандарты. Поэтому для ATHENA было очень важно собрать сведения о тех стандартах, которые используют партнеры проекта. Рабочая группа 3 проекта ATHENA провела исследование стандартов, используемых участниками проекта, результаты которого были представлены в виде отчета⁹. При обсуждении стандартов мы пришли к выводу, что профессиональное сообщество проявляет большой интерес к этой проблеме, и решили опубликовать обзор стандартов в виде брошюры. Мы надеемся, что представленная в этой публикации информация будет тем инструментом, который поможет навести мосты между учреждениями культурного наследия и увеличит совместимость информационных ресурсов.

⁹ http://www.nytimes.com/2011/03/15/business/media/15libraries.html?_r=1 (на англ. яз., прим. ред.)

1. Основные положения

1.1. Стандарт

Вот какое определение стандарта дает Британский институт стандартов (The British Standards Institution - BSI), старейшая организация по стандартизации (1901): «Вот самое простое объяснение: стандарт – это согласованный, повторяемый способ что-то делать. Это опубликованный документ, содержащий техническую спецификацию или другие точные критерии, которые должны **систематически использоваться в качестве правила, руководства или определения**. Стандарты помогают **облегчить жизнь** и увеличить надежность и эффективность изделий и услуг, которыми мы пользуемся. При разработке стандартов **объединяются опыт и знания всех заинтересованных сторон**, таких как производители, продавцы, покупатели, пользователи и контролирующие органы относительно конкретного материала, продукта, процесса или услуги». К преимуществам стандартов, перечисленным выше, можно добавить **обеспечение совместимости**.

1.1.1. Типы стандартов

Существует несколько способов классификации стандартов.

С юридической точки зрения различаются следующие типы стандартов:

- *Фактические (De facto)*

Это стандарты, которые не утверждены стандартизирующим органом с помощью формальной процедуры, но широко используются и признаются в той области деятельности, к которой они относятся. Это обычная ситуация в индустрии информационных технологий, где, хорошо это или плохо, доминирование фирмы Microsoft привело к тому, что некоторые ее продукты фактически стали стандартами (например, *Word* для *Windows*). Такие стандарты, возможно, не самые лучшие для решения конкретной задачи, но часто наиболее успешные с экономической точки зрения.

- *Юридические (De jure)*

Стандарты, формально принятые стандартизирующим органом (например, ISO). Они выработаны на основе соглашения группы заинтересованных сторон, при этом ни одна сторона не доминирует. Однако для их разработки и утверждения требуется довольно много времени, поэтому иногда они не успевают за технологическими новшествами.

Стандарты можно классифицировать с точки зрения того сообщества, для которого они разрабатываются и в котором используются:

- *Служебные (In-house)*

Стандарты, разработанные и используемые в отдельной организации для определенных целей. Например, это географические названия данной местности. Такие названия расширяют национальную номенклатуру географических названий, которая содержит географические названия более крупных территориальных образований.

- *Секторальные (Community)*

Стандарты, разработанные несколькими организациями, работающими в одном секторе экономики, для использования в этом секторе. Например, стандарт музейной документации Соединенного Королевства *SPECTRUM* был разработан опытными специалистами музейного сектора.

- *Национальные (National)*

Стандарты, разработанные для использования в одной стране и признанные на национальном уровне. Терминология, признанная в пределах одной страны, является примером таких стандартов.

- *Международные (International)*

Это стандарты, признанные и используемые во всем мире, почти всегда подтвержденные международным стандартизирующим органом, например ISO 8601 – международный стандарт для даты и времени.

Некоторые стандарты, в начале относящиеся к одному какому-нибудь типу, в результате дальнейшей работы и процесса утверждения переходят в другую категорию. Например, концептуальная образцовая модель *CIDOC (Conceptual Reference Model - CRM)* первоначально была разработана рабочей группой по стандартам документации международного комитета CIDOC как секторальный стандарт, а сейчас утверждена международным комитетом по стандартизации (стандарт ISO 21127:2006).

Другой тип стандартов, известный как *открытые стандарты*, будет подробно рассмотрен в следующем разделе.

1.1.2. Открытые стандарты¹⁰

Кеннет Кречмер приводит десять признаков открытых стандартов¹¹:

1. Открытое обсуждение – все заинтересованные стороны могут участвовать в процессе разработки стандартов.

2. Консенсус – все предложения обсуждаются на равных, никакое мнение не доминирует, вырабатывается консенсус.

3. Процесс согласования – при согласовании используются процессы голосования и апелляции.

4. Открытые права на интеллектуальную собственность – держатели прав на интеллектуальную собственность, относящуюся к стандарту, предоставляют информацию о правах.

5. Единый стандарт для всего мир – один и тот же стандарт во всем мире для одной и той же функции (продукта).

6. Открытые изменения – все изменения представляются и согласовываются в соответствии с пятью вышеуказанными принципами.

7. Открытая документация – проекты документов и документация по завершеному стандарту должны быть легко доступными для внедрения и использования.

¹⁰ Об Открытых стандартах на русском языке см.

http://ru.wikipedia.org/wiki/%D0%9E%D1%82%D0%BA%D1%80%D1%8B%D1%82%D1%8B%D0%B9_%D1%81%D1%82%D0%B0%D0%BD%D0%B4%D0%B0%D1%80%D1%82 (прим. ред.)

¹¹ Kenneth Krechmer, *Open Standards Requirements*, “The International Journal of IT Standards and Standardization Research”, 4 (2006), No. 1. См: <http://www.csrstds.com/openstds.pdf>

8. Открытый интерфейс – поддержка внедрения; интерфейс доступен, внедрение стандарта не управляется извне; при внедрении поддерживается перенос (использование) на другие программно-аппаратные платформы.

9. Открытый доступ – наличие единообразных механизмов для тестирования внедрения и оценки стандарта пользователями.

10. Постоянная поддержка – стандарты поддерживаются до тех пор, пока пользователи проявляют к нему интерес, а не до тех пор, пока он интересен разработчику.

Большинство этих требований относится к разработке стандарта. Этот процесс должен быть прозрачным и демократичным. До конца не ясно, каждый ли стандарт, который считается открытым, отвечает всем вышперечисленным требованиям. Однако оформленные национальные и международные стандарты, вероятнее всего, хотя бы до некоторой степени отвечают большинству этих требований.

Для потенциального пользователя стандарта **открытые права на интеллектуальную собственность, открытый доступ и постоянная поддержка** - наиболее важные характеристики. Поэтому при описании стандартов мы каждый раз указываем, является ли стандарт «открытым» относительно этих трех критериев.

1.2. Оцифровка

Оцифровка – это процесс трансформации оригинального (аналогового) материала в цифровую форму. В зависимости от поставленной цели, различаются три типа оцифровки:

- **Репродукция** – оцифровка с целью воспроизведения исходного материала в цифровой форме с наибольшей точностью. Эта категория включает в себя оцифровку изображений, звука и видео;

- **Поиск** – оцифровка с целью поиска и извлечения оригинального материала. Эта категория включает в себя сканирование и индексацию документов, например, контрактов, писем и пр. Целью оцифровки является не точное воспроизведение, но увеличение возможностей поиска и использования материала.

- **Миграция** – оцифровка с целью захвата информации из аналоговой (бумажной) системы каталогизации и использования ее в системе автоматизированного управления коллекциями.

В данной работе мы, в основном, рассмотрим стандарты, относящиеся к оцифровке с целью репродукции. Однако мы также обсудим некоторые стандарты, связанные с двумя другими типами оцифровки.

1.3. Совместимость

Совместимость (Interoperability¹²) может быть определена следующим образом:

«Возможность систем, процедур и учреждений культуры функционировать таким образом, чтобы максимально обеспечить возможность обмена информацией и повторного использования информации как внутри учреждения, так и вне него»¹³.

Это определение было сформулировано таким образом, чтобы максимально увеличить возможности использования коллективного знания учреждения/организации. В данной работе мы делаем попытку максимально увеличить возможности европейских учреждений культуры обмениваться информационными ресурсами с помощью технологий информационного общества (и получать пользу от этого процесса).

Пол Миллер выделяет 6 типов совместимости¹⁴:

- **Техническая совместимость** – обеспечивается путем использования общих технических стандартов (например, форматов файлов, схем метаданных и пр.).
- **Семантическая совместимость** – обеспечивается путем использования общих терминологических словарей (например, тезаурусов).
- **Политическая/гуманитарная совместимость** – обеспечивается путем преодоления барьеров, связанных с разным культурным опытом, методами деятельности пользователей и провайдеров информации.
- **Межсекторальная совместимость** – обеспечивается в результате совместной деятельности разных профессиональных сообществ (например, музеев, архивов и библиотек).
- **Правовая совместимость** – обеспечивается путем соблюдения правовых ограничений, накладываемых на провайдеров информации (например, *свобода информации* и *законодательство по защите данных*).
- **Международная совместимость** – обеспечивается путем признания и преодоления барьеров, вызванных культурными и языковыми различиями.

В этой работе мы, в основном, касаемся вопросов технической и семантической совместимости. Другие типы совместимости могут стать темой других проектов.

1.4. Метаданные

Термин «метаданные» стал так часто использоваться, что имеет смысл остановиться на нем более подробно¹⁵.

¹² В русском языке в качестве термина часто используется калька с английского: «интероперабельность», определяемая как «способность к информационных систем к взаимодействию». См.

<http://wiki.services.openoffice.org/wiki/RU/kb/90000033> (на русск. яз., прим. ред.)

¹³ Helen Ashby – Gordon McKenna – Matthew Stiff. *SPECTRUM. Knowledge*. London: MDA, 2001, p. 63.

¹⁴ Paul Miller, *Interoperability: What is it and Why should I want it?*, “Ariadne”, 2000, No. 4.

См: <http://www.ariadne.ac.uk/issue24/interoperability/>

1.4.1. Трудности дефиниции

Термин «метаданные» построен по той же модели, что и термины «метафизика» и «мета-язык», в которых приставка «мета» означает:

«относится к концепциям и идеям дисциплины (об академической дисциплине)»¹⁶.

Элемент «данные» указывает на тенденцию ограничить использование термина цифровыми «объектами». Например, часто цитируется следующее определение: «Данные о данных»¹⁷.

В этом определении подразумевается, что метаданные относятся исключительно к текстам и мультимедиа в Интернете или в компьютере.

Термин «метаданные» в последние годы стал новым модным словом специалистов по информации, и оно в определенных случаях вызывает путаницу и/или неловкость. Например, Тони Гилл пишет: «... термин «метаданные» все чаще используется в таких контекстах, в которых термин «данные» был вполне достаточен всего несколько лет назад (например, описание людей, объектов или явлений)»¹⁸.

В этом более широком контексте метаданные могут быть данными, широко известными в секторе культурного наследия: данные по управлению коллекциями, каталожные записи и этикетки на выставках. Фактически, любые «данные» могут рассматриваться как «метаданные».

Итак, зачем вообще использовать термин «метаданные»? Возможно, обзор некоторых ключевых аспектов метаданных поможет найти ответ на этот вопрос.

1.4.2. Характеристики метаданных

Ключевой вопрос, связанный с метаданными, это вопрос о ресурсе. Ресурс - это та сущность, которая описывается метаданными. Опасно ограничивать ресурс текстами и мультимедийными «объектами», доступными на компьютере или в Интернете. Ресурс – это любая сущность, которую мы хотим описать и предоставить доступ к ней каким-либо способом.

Ресурсом могут быть¹⁹:

- Тексты (электронные или на бумаге)
- Физические объекты

¹⁵ О метаданных на русском языке см.

<http://ru.wikipedia.org/wiki/%D0%9C%D0%B5%D1%82%D0%B0%D0%B4%D0%B0%D0%BD%D0%BD%D1%8B%D0%B5>; http://www.elbib.ru/index.phtml?page=elbib/rus/methodology/md_rev ; <http://www.metadata.ru/content/view/46/63/> (прим. ред.)

¹⁶ Collins Concise Dictionary, Third Edition, 1995, p. 835.

¹⁷ См: <http://www.webopedia.com/TERM/m/metadata.html>

¹⁸ Gill, Tony. 'Metadata and the Web', in Baca, M. (ed.) Introduction to Metadata: Pathways to Digital Information. 3rd rev. Getty Information Institute. 2008.

См. http://www.getty.edu/research/conducting_research/standards/intrometadata/metadata.html

¹⁹ Метаданные также могут быть «ресурсом» в указанном выше смысле, т.е. метаданные также могут быть описаны с помощью других метаданных. См.

<http://traditio.ru/wiki/%D0%9C%D0%B5%D1%82%D0%B0%D0%B4%D0%B0%D0%BD%D0%BD%D1%8B%D0%B5> (прим. ред.)

- Мультимедиа (изображение, звук, видео и пр.)
- Программное обеспечение
- Люди
- Организации
- Территории
- События
- Идеи
- Собрания всего вышеперечисленного

Некоторые ресурсы являются **суррогатами** других ресурсов. Суррогат – это представление ресурса в какой-то другой форме. Например, цифровое изображение или фотография произведения искусства, или представление некоторой территории в виде виртуальной реальности, или факсимиле некоторого объекта.

При создании метаданных важно различать ресурс и его суррогат. *Метаданные суррогата не должны описывать исходный ресурс.* Например, оригинал художественного произведения может быть вне сферы действия авторского права, тогда как авторские права на фотографию этого произведения охраняются.

Ресурсы могут быть *связаны* друг с другом разнообразными отношениями, не только отношением между оригиналом и суррогатом. Например, создание, обнародование (в частности, публикация) и использование.

Известны различные классификации метаданных. Например²⁰:

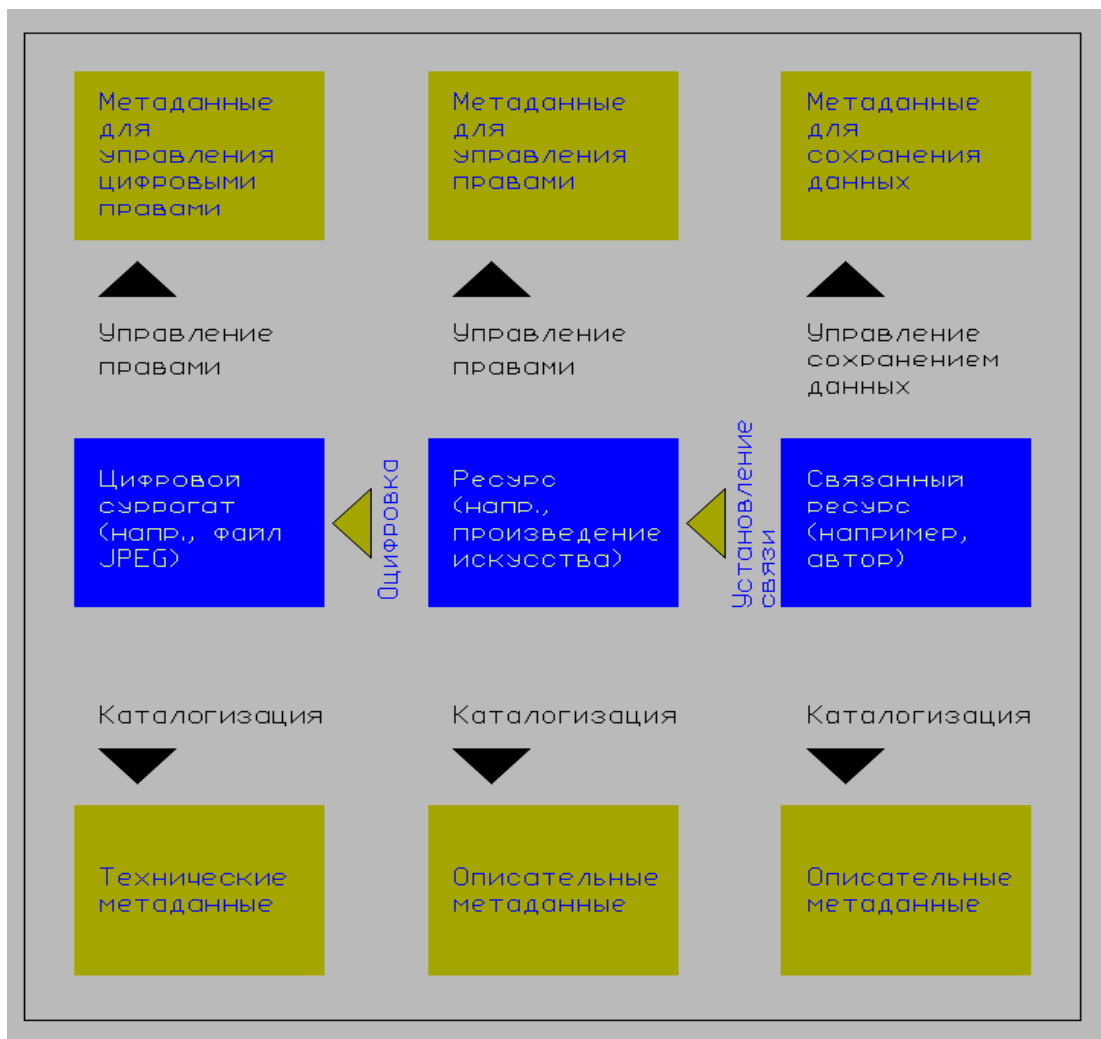
- **Административные** – управление и администрирование ресурсов (например, приобретение, права, место хранения).
- **Описательные** – описание или идентификация ресурсов (например, каталожные записи).
- **Метаданные для сохранения**²¹ – управление процессами сохранения (состояние, миграция данных).
- **Технические** – как функционирует система или метаданные (например, форматы, пароли, водяные знаки).
- **Метаданные, относящиеся к использованию** – уровень и тип использования ресурсов (пользователь и информация, относящаяся к нему).

Отношения между различными ресурсами и ассоциированными с ними метаданными можно представить в виде диаграммы:

²⁰ Anne J. Gilliland-Swetland, Setting the stage, in: Baca, M. (ed.), *Introduction to Metadata: Pathways to Digital Information*, 3rd rev. Getty Information Institute. 2008.

См. http://www.getty.edu/research/conducting_research/standards/intrometadata/setting.html

²¹ <http://www.minervaplus.ru/publish/Gureev.doc> (на русск. яз., прим. ред.)



Другой тип метаданных относится к **поиску ресурсов**. Такие метаданные помогают потенциальному пользователю ресурса найти информацию, которая дает ему возможность решить, хочет ли он получить доступ к этому ресурсу. Этот процесс аналогичен вводу какого-то слова или фразы в поисковую машину, просмотру найденных в Интернете результатов и переходу по ссылке. Цель метаданных – получить более точные и релевантные запросу пользователя результаты поиска.

Наиболее известные метаданные для поиска ресурсов – это *Дублинское ядро*²². Метаданные Дублинского ядра включают информацию о ресурсе, его *идентификатор*, который используется для *доступа* к ресурсу. Это прямой доступ, если идентификатор – это адрес в Интернете (URL), или непрямо́й доступ, если это, например, ISBN²³.

Наконец, существует тенденция группировать метаданные, создавая **схемы** метаданных. Схема – это описание набора отдельных метаданных (элементов), а также описание того, как эти элементы структурированы. При открытой разработке схем метаданных они тестируются и принимаются для внедрения организациями и правительствами. Такие схемы метаданных никак нельзя игнорировать.

1.4.3. «Более адекватная» дефиниция?

«Метаданные» - значительно более важное понятие, чем предполагает простое определение, приведённое выше. Мы можем предложить исчерпывающее определение:

«Метаданные – это структурированная информация о ресурсе любого типа, которая используется для идентификации ресурса, для описания ресурса, для управления ресурсом и для предоставления доступа к нему».

²² О Дублинском ядре на русском языке см.

http://ru.wikipedia.org/wiki/%D0%94%D1%83%D0%B1%D0%BB%D0%B8%D0%BD%D1%81%D0%BA%D0%BE%D0%B5_%D1%8F%D0%B4%D1%80%D0%BE

²³ ISBN – это уникальный международный порядковый номер книги (International Serial Book Number). ISBN состоит из 13 цифр и однозначно определяет издание (прим. ред.)

2. Ландшафт стандартов

2.1. Описание стандартов

Полностью разделяя идею использования стандартов, мы описываем каждый стандарт с помощью адаптированного для этой цели стандарта метаданных Дублинского ядра. В наших описаниях используется 9 из 15 элементов Дублинского ядра.

Вот эти элементы:

Название

Название, под которым данный стандарт известен. В большинстве случаев приводится полное и сокращенное название.

Создатель

Название организации или фамилия и имя лица, создавшего данный стандарт.

Издатель

Название организации, сделавшей данный стандарт публично доступным.

Дата

Дата, когда стандарт был *впервые* опубликован.

Идентификатор

Номер или другой идентификатор, под которым стандарт опубликован, или Интернет-адрес (URL), по которому можно найти описание стандарта.

Права

Правовые ограничения, например, патенты, относящиеся к данному стандарту.

Описание

Текстовое описание, разъясняющее сущность стандарта и его использование.

Содержание²⁴

Ключевые слова, описывающие сущность стандарта.

Связи

Другие стандарты, с которым связан данный стандарт, и ассоциированные веб-сайты.

Описание стандартов ориентировано на обычного читателя. Технические детали можно найти в Интернете. Цель данной брошюры – дать читателям возможность получить всю информацию о различных стандартах в одном месте.

²⁴ англ. Subject

2.2. Схемы информации (метаданных)

Большинство стандартов, перечисленных ниже, определяют описательные метаданные. Они направлены на описание информации о ресурсах: книгах, архивных документах, музейных предметах, людях, территориях и организациях.

SPECTRUM является не только описательным, но и процедурным стандартом. Этот тип стандартов определяет виды деятельности и требования к информации, необходимой для управления определенным сектором сферы культуры (административные метаданные).

Стандарты не обязательно содержат разъяснения того, как их технически применять. Часто их можно рассматривать как примеры положительного опыта в определенной области.

Описание стандартов структурировано в соответствии с тем, к какой области деятельности в сфере культурного наследия они относятся:

- музейные стандарты;
- архивные стандарты;
- библиотечные стандарты;
- стандарты описания недвижимого наследия;
- стандарты описания культурного наследия,

а также в соответствии с целевой направленностью:

- поиск ресурсов;
- кодирование документов.

EAD и TEI являются также стандартами разметки для кодирования документов. Первоначально они основывались на стандарте кодирования текста SGML, но в последнее время они были обновлены для использования XML (оба эти стандарта также описаны ниже).

2.2.1. Музейные стандарты

CDWA

Название: CDWA • Categories for the of Works of Art (Категории для произведений искусства)²⁵

Создатель: Art Information Task Force

Издатель: Getty Research Institute

Дата публикации: 1990

Идентификатор:

<http://www.getty.edu/research/institute/standards/cdwa/index.html>

Права: Getty Research Institute.

Описание: описывает информационные ресурсы баз данных по изобразительному искусству, предлагает концептуальный подход к описанию и доступу к произведениям искусства и изображениям. Предлагает терминологию, которая помогает совместить информацию из разных систем и сделать ее более доступной. Может служить основой для взаимодействия существующих информационных систем и для разработки новых систем.

Содержание: музейная документация

Связи:

http://www.getty.edu/research/conducting_research/standards/cdwa/cdwalite.html
(CDWA Lite)

MUSEUMDAT

Название: museumdat²⁶

Создатель: Fachgruppe Dokumentation im Deutschen Museumsbund Institut für Museumsforschung SMB-ПК
Zuse-Institut Berlin

Издатель: Fachgruppe Dokumentation im Deutschen Museumsbund / Institut für Museumsforschung SMB-ПК / Zuse-Institut Berlin

Дата публикации: 2006-2007

Идентификатор: <http://museum.zib.de/museumdat/museumdat-v1.0.xsd>

Права: Копирайт FG Dokumentation im Deutschen Museumsbund / Institut für Museumsforschung SMB-ПК / Zuse-Institut Berlin

Описание: формат для агрегирования (сбора) данных (XML схема), оптимизированный для поиска и публикации на музейных порталах, предназначен для автоматического сбора ключевых данных.

Содержание: описание объектов культуры

Связи: <http://museum.zib.de/museumdat/museumdat-v1.0-en.pdf>

(документация);

<http://www.museumdat.org/index.php?ln=en>

²⁵ <http://www.metadata.ru/content/view/46/63/> (на русск. яз., прим. ред.)

²⁶ О стандартах CDWA, CDWA lite и Museumdat см. <http://www.ifla.org/files/hq/papers/ifla75/107-coburn-ru.pdf> (прим. ред.)

ОБЪЕКТ ID

Название: Object ID²⁷

Создатель: Thornes, Robin (и др.)

Издатель: J. Paul Getty Trust

Дата публикации: 1999

Идентификатор: <http://archives.icom.museum/object-id/>

Права: Копирайт The J. Paul Getty Trust. Все права защищены.

Описание: стандарт для описания объектов культуры. Разработан в результате сотрудничества музейного сообщества, правоохранительных органов, таможи, коммерческих структур, страховых агентств и оценщиков произведений искусства и древностей.

Содержание: описание объектов культуры

Связи: <http://archives.icom.museum/object-id/checklist.html> (инструкция)

SPECTRUM

Название: SPECTRUM • SPECTRUM: The UK Museum Documentation Standard, 3rd Edition (Version 3.1) (Стандарт музейной документации Соединенного Королевства, версия 3.1)²⁸

Создатель: McKenna, Gordon (ред.), Patsatzi, Efthymia (ред.)

Издатель: Collections Trust [MDA]

Дата публикации: 2007

Идентификатор: 1 900642 14 X (ISBN)

Права: Копирайт Collections Trust. Все права защищены. [Открытый стандарт]

Описание: стандарт для музейной документации и управления коллекциями. Основан на 21 процедуре, типовой для музея, и понятии «единица информации» - данных, необходимых для реализации каждой процедуры. Существует версия на голландском языке для Фландрии и Нидерландов.

Содержание: музейная документация

Связи: <http://www.collectionstrust.org.uk/spectrum> (версия для Соединенного Королевства)

<http://www.collectionstrust.org.uk/spectrum-nl> (версия для Нидерландов)

<http://www.collectionstrust.org.uk/spectrum-nl-be> (версия для Фландрии)

²⁷ <http://www.metadata.ru/content/view/48/60/> (на русск. яз., прим. ред.)

²⁸ О предыдущих редакциях SPECTRUM на русском языке, процедурах и единицах информации см. <http://www.cpic.ru/Publicat/26-00.htm> (прим. ред.)

2.2.2. Архивные стандарты

EAD

Название: EAD • Encoded Archival Description²⁹ (Кодированное архивное описание)

Создатель: Encoded Archival Description Working Group of the Society of American Archivists Network Development and MARC Standards Office of the Library of Congress

(Рабочая группа общества американских архивистов и группа стандартов MARC Библиотеки Конгресса США)

Издатель: Society of American Archivists

Дата публикации: 2002

Идентификатор: <http://www.loc.gov/ead/ead.xsd> (W3C схема)

Права: Копирайт Society of American Archivists. [Открытый стандарт]

Описание: схема для кодирования архивных описей. Используется также для описания коллекций

Содержание: описание архивных коллекций и кодирование документов

Связи: <http://www.loc.gov/ead> (вебсайт EAD)

ISAD(G)

Название: ISAD(G): General International Standard Archival Description, Second Edition³⁰

(общий международный стандарт архивного описания, второе издание)

Создатель: Committee on Descriptive Standards (ICA/CDS) (организация, принявшая стандарт)

Издатель: International Congress on Archives (ICA) (Международный Конгресс архивов)

Дата публикации: 2000

Идентификатор: 0-9696035-5-X (ISBN)

[http://www.icacds.org.uk/eng/ISAD\(G\).pdf](http://www.icacds.org.uk/eng/ISAD(G).pdf)

Права: Копирайт International Congress on Archives

Описание: общие правила архивного описания, которые можно применять независимо от формы или носителя архивного материала. Эти правила содержат 26 элементов, которые можно комбинировать для описания единицы архивного хранения.

Содержание: архивное описание

Связи: <http://www.ica.org> (вебсайт)

²⁹ <http://www.metadata.ru/content/view/41/60/> (на русск. яз., прим. ред.)

³⁰ <http://www.metadata.ru/content/view/42/60/> (на русск. яз., прим. ред.)

ISAAR (CPF)

Название: ISAAR (CPF) • ISAAR (CPF): International standard archival authority record for corporate bodies, persons and families, Second Edition³¹

(Международный стандарт авторитетных записей об организациях, лицах и семьях, второе издание)

Создатель: ICA Committee on Descriptive Standard (разработчик)

Издатель: International Congress on Archives (ICA)

Дата публикации: 2004

Идентификатор: 2-9521932-2-3 (ISBN)

[http://www.icacds.org.uk/eng/ISAAR\(CPF\)2ed.pdf](http://www.icacds.org.uk/eng/ISAAR(CPF)2ed.pdf)

Права: Копирайт International Congress on Archives

Описание: общие правила создания архивных авторитетных записей для описания организаций, лиц и семей, которые могут быть значениями поля «Создатель» при описании архивных документов.

Содержание: архивное описание

Связи: <http://www.ica.org> (вебсайт ica)

³¹ http://www.archives.gov.ua/Problems/eurasica-2006-1_2.php (на русск. яз., прим. ред.)

2.2.3. Библиотечные стандарты

FRBR

Название: FRBR • Functional Requirements for Bibliographic Records³²

(функциональные требования к библиографическим записям)

Создатель: IFLA Study Group on the Functional Requirements for Bibliographic Records International

(исследовательская группа IFLA по функциональным характеристикам библиографических записей)

Издатель: International Federation of Library Associations and Institutions (IFLA) (Международная Федерация библиотечных ассоциаций и учреждений)

Дата публикации: 1998

Идентификатор: ISBN 3-598-11382-X

Права: Копирайт International Federation of Library Associations and Institutions

Описание: ориентированная на пользователя концептуальная модель, которая используется для поиска и доступа к библиотечным каталогам онлайн и библиографическим базам данных.

Содержание: библиографическое описание

Связи: <http://www.ifla.org/VII/s13/frbr/frbr.pdf> (окончательный отчет)

MAВ2

Название: MAВ2 • Maschinelles Austauschformat fur Bibliotheken

(автоматизированный формат для библиотечного обмена)

Создатель: Deutsche Nationalbibliothek (Национальная библиотека Германии)

Издатель: Deutsche Nationalbibliothek

Дата публикации: 2001

Идентификатор: <http://www.d-nb.de/standardisierung/txt/titelmaв.txt> [библиографические данные (MAВ-TITEL)]

Права: Deutsche Nationalbibliothek

Описание: немецкий формат для машиночитаемого обмена любыми данными, создаваемыми в библиотеках (библиография, авторитетные файлы и местные данные).

Содержание: библиографическое описание

Связи: <http://www.d-nb.de/eng/standardisierung/formate/maв.htm>

(веб страница MAВ Национальной библиотеки Германии на английском яз.)

³² http://www.gpntb.ru/win/ntb/ntb99/11/f11_02.html (на русск. яз., прим. ред.)

MARC

Название: MARC •MACHINE-Readable Cataloguing MARC 21 Concise Format for Bibliographic Data³³

(машиночитаемый формат каталогизации MARC 21 для библиографических данных)

Создатель: Network Development and MARC Standards Office of the Library of Congress

(Офис сетевого развития и стандартов MARC Библиотеки Конгресса США)

Издатель: Library of Congress

Дата публикации: 2002 (обновление)

Идентификатор: <http://www.loc.gov/marc/bibliographic/ecbdhome.html>

Права: [Открытый стандарт]

Описание: стандарт для представления и передачи библиографической информации в машиночитаемой форме

Содержание: библиографические описания

Связи: <http://www.loc.gov/marc/authority/ecadhhome.html> (авторитетные файлы)

<http://www.loc.gov/marc/holdings/echdhome.html> (фондовые данные)

<http://www.loc.gov/marc/classification/eccdhome.html> (классификационные данные)

<http://www.loc.gov/marc/community/eccihome.html> (данные о сообществе)

ISO 2709

Название: Information and documentation • Format for Information Exchange³⁴

(Информация и документация. Формат обмена данными)

Создатель: International Organization for Standardization (ISO)

(Международная организация по стандартизации)

Издатель: International Organization for Standardization (ISO)

Дата публикации: 1996

Идентификатор: ISO 2709:1996

http://www.iso.org/iso/catalogue_detail.htm?csnumber=7675

Права: [Открытый стандарт]

Описание: определяет требования к обменному формату библиографических описаний. Описывает структуру коммуникаций между системами обработки данных.

Содержание: библиографическое описание

Связи: MARC

³³ <http://marc21.rsl.ru/index.php?f=227> и <http://www.metadata.ru/content/view/45/61/> (на русск. яз., прим. ред.)

³⁴ <http://www.gpntb.ru/win/inter-events/crimea2005/disk/186.pdf> (на русск. яз., прим. ред.)

METS

Название: METS • Metadata Encoding and Transmission Standard³⁵

(стандарт кодирования и передачи метаданных)

Создатель: McDonough, Jerome (et al)

Издатель: Digital Library Federation (Федерация цифровых библиотек)

Дата публикации: 2007 (Version 1.7)

Идентификатор: <http://www.loc.gov/standards/mets/mets.xsd>

Права: Копирайт Digital Library Federation

Описание: XML схема – стандарт для кодирования описательных, административных и структурных метаданных об объектах, хранящихся в цифровой библиотеке.

Содержание: цифровая библиотека

Связи: <http://www.loc.gov/standards/mets/mets-schemadocs.html> (вебсайт METS)

MODS

Название: MODS • Metadata Object Description Schema³⁶

(схема метаданных для описания объекта)

Создатель: Library of Congress' Network Development MARC Standards Office (Офис сетевого развития и стандартов MARC Библиотеки Конгресса США)

Издатель: Library of Congress

Дата публикации: 2008 (версия 3.3)

Идентификатор: <http://www.loc.gov/standards/mods/v3/mods-3-3.xsd>

Права: [?]

Описание: XML схема для набора элементов библиографической записи, которая может использоваться для разных целей, в особенности для библиотечных систем

Содержание: библиографическое описание

Связи: <http://www.loc.gov/standards/mods/> (вебсайт)

³⁵ <http://www.metadata.ru/content/view/21/33/> (на русск. яз., прим. ред.)

³⁶ <http://www.gpntb.ru/libcom5/disk/doc/34.pdf> (на русск. яз., прим. ред.)

RUSMARC³⁷

Название: RUSMARC • Российский коммуникативный формат
Российская версия UNIMARC

Создатель: Российская библиотечная ассоциация

Издатель: Российская библиотечная ассоциация

Дата публикации: 1998

Идентификатор: <http://www.rba.ru/rusmarc/>

Права: [Открытый стандарт]

Описание: утвержден Приказом Министра Культуры РФ № 45 от 27.01.98 в качестве обязательного формата при обмене библиографическими записями среди библиотек сети Министерства Культуры. Это формат представления библиографических данных в машиночитаемой форме. Предназначен для обмена библиографическими данными.

Содержание: библиографическое описание, авторитетные/нормативные данные, классификационные данные

Связи: UNIMARC³⁸. Российский коммуникативный формат официально включен Постоянным комитетом по формату UNIMARC в список национальных адаптаций формата UNIMARC

³⁷ Раздел добавлен к переводу оригинального текста редактором русскоязычного издания

³⁸ <http://archive.ifla.org/VI/3/p1996-1/sec-uni.htm> (на англ. яз., прим. ред.)

2.2.4. Стандарты описания недвижимого наследия

MIDAS

Название: MIDAS •MIDAS Heritage (Monument Inventory Data Standard)

(Стандарт для инвентаризации монументов³⁹)

Создатель: English Heritage Forum on Information Standards in Heritage (FISH)
(форум по информационным стандартам в области наследия)

Издатель: English Heritage

Дата публикации: 2008

Идентификатор: <http://www.english-heritage.org.uk/publications/midas-heritage/>

Права: Копирайт Forum on Information Standards in Heritage (FISH)

Описание: Стандарт Соединенного Королевства для информации об историческом контексте. Он устанавливает, какая информация должна быть зарегистрирована для того, чтобы совместно использовать исторические сведения и для долговременного сохранения записей. Стандарт предназначен для описания отдельных элементов, формирующих исторический контекст (здания, места археологических раскопок, затонувшие корабли, исторические территории и артефакты); действия, предпринимаемые для понимания, защиты и управления изменениями этих ценностей.

Содержание: документация (историческое наследие)

Связи: <http://www.fish-forum.info> (вебсайт FISH⁴⁰),
<http://www.heritage-standards.org.uk> (инструментарий для обеспечения совместимости FISH)

³⁹ <http://www.metadata.ru/content/view/47/60/> (на русск. яз., прим. ред.)

⁴⁰ Форум по информационным стандартам для недвижимого наследия (Forum on Information Standards in Heritage) (прим. ред.)

2.2.5. Культурное наследие

Эти стандарты не связаны ни с каким конкретным сектором сферы культуры, но могут использоваться в любой области деятельности, связанной с культурой.

CIDOC-CRM

Название: CIDOC-CRM • CIDOC Conceptual Reference Model⁴¹

(концептуальная образцовая модель CIDOC)

Создатель: CIDOC Documentation Standards Working Group

(рабочая группа по стандартам документации CIDOC)

Издатель: International Organization for Standardization (ISO)

Дата публикации: 2006

Идентификатор: ISO 21127:2006.

Права: [Открытый стандарт]

Описание: концептуальная объектно-ориентированная модель, которая обеспечивает расширяемую онтологию для концепций и информации по культурному наследию и музейной документации.

Содержание: документация (музеи)

Связи: <http://www.iso.org> (вебсайт ISO)

<http://cidoc.ics.forth.gr> (вебсайт CRM)

VRA

Название: VRA⁴² • VRA Core (Version 4.0)

(Ядро VRA, версия 4.0)

Создатель: Visual Resources Association's Data Standards Committee

(Комитет стандартов данных Ассоциации визуальных ресурсов)

Издатель: Visual Resources Association

Дата публикации: 2007

Идентификатор: <http://www.loc.gov/standards/vracore/>

Права: Копирайт Visual Resources Association [?]

Описание: набор элементов метаданных (единицы информации, такие как название, местонахождение, дата и пр.), а также первоначальный план иерархической структуры этих элементов. Набор элементов дает возможность организации категорий описания произведений визуальной культуры, а также изображений.

Содержание: визуальная культура

Связи: <http://www.vraweb.org/projects/vracore4/index.html> (вебсайт)

<http://www.vraweb.org/projects/vracore4/vra-4.0.xsd> (XML схема)

⁴¹ http://www.worldlingo.com/ma/enwiki/ru/CIDOC_Conceptual_Reference_Model и <http://conf.cpic.ru/upload/eva2006/reports/%C8%E2%E0%ED%EE%E2%C2.pdf> (на русск. яз., прим. ред.)

⁴² <http://www.metadata.ru/content/view/49/60/> (на русск. яз., прим. ред.)

LIDO⁴³

Название: LIDO • Lightweight Information Describing Objects

(облегченный стандарт информации для описания объектов)

Создатель: проект Европейской Комиссии ATHENA⁴⁴

Издатель: международный комитет музейной документации CIDOC⁴⁵

Дата публикации: 2010

Идентификатор: <http://www.lido-schema.org>

Права: [Открытый стандарт]

Описание: описание и агрегирование ресурсов музеев, библиотек, архивов;
совмещение схем метаданных в сфере культуры

Содержание: музеи, библиотеки, архивы

Связи: CDWA Lite, museumdat, CIDOC-CRM, SPECTRUM

⁴³ Раздел добавлен к переводу оригинального текста редактором русскоязычного издания

⁴⁴ <http://www.minervaplus.ru/athena/athena.htm> (на русск. яз., прим. ред.)

⁴⁵ <http://cidoc.icom.museum/> (на англ. яз.). LIDO был обнародован и одобрен на ежегодном заседании CIDOC, которое состоялось на конференции международного комитета музеев ICOM в ноябре 2010 года в Шанхае (прим. ред.)

2.2.6. Поиск ресурсов

DUBLIN CORE

Название: Dublin Core • The Dublin Core Element Set Version 1.1⁴⁶

(Набор элементов Дублинского ядра, версия 1.1)

Создатель: Dublin Core Metadata Initiative

(Инициатива метаданных Дублинского ядра)

Издатель: Dublin Core Metadata Initiative

Дата публикации: 1999

Идентификатор: <http://dublincore.org/documents/1999/07/02/dces/>

Права: [Открытый стандарт]

Описание: Дублинское ядро – это простой набор метаданных, ориентированный на поиск электронных ресурсов. Элементы метаданных могут быть разделены на группы: контент – охват материала; описание, тип, связи; источник, содержание, название; интеллектуальные права – поставщик информации, создатель, издатель, права; детализация – дата, формат, идентификатор, язык. На использование этого стандарта дали мандат правительства нескольких европейских стран (например, Соединенного Королевства) и мира (например, Австралия).

Содержание: поиск ресурсов

2.2.7. Кодирование документов

TEI

Название: TEI • TEI P5: Guidelines for Electronic Text Encoding and Interchange

(руководство по кодированию и обмену электронными текстами)

Создатель: Burnard, Lou (ред.), Bauman, Syd (ред.)

Издатель: Text Encoding Initiative Consortium

(Консорциум по кодированию Текста)

Дата публикации: 2007

Идентификатор: <http://www.tei-c.org/Guidelines/>

Права: Копирайт TEI Consortium 2007. Универсальная общедоступная лицензия⁴⁷. Разрешается и поощряется копирование и распространение. [Открытый стандарт]

Описание: определяет набор тэгов (маркеров), которые вставляются в электронную форму документа (текст) для того, чтобы определить структуру и другие характеристики документа. Эти тэги используются при обработке текста компьютером. Они используются для текстов на любом естественном языке, любого времени создания и любого жанра. Эта версия совместима с XML.

Содержание: кодирование документов

Связи: XML, SGML, <http://www.tei-c.org> (вебсайт Text Encoding Initiative Consortium)

⁴⁶ <http://www.rba.ru/rusmarc/soft/dc.html> (на русск. яз., прим. ред.)

⁴⁷ http://ru.wikipedia.org/wiki/GNU_General_Public_License (на русск. яз., прим. ред.)

2.3. Мультимедиа форматы

2.3.1. Текст

Стандарты кодирования текста определяют, как тексты хранятся в компьютерных системах. Это промышленные стандарты (*de facto*), такие как Microsoft Word, а также стандарты, предложенные World Wide Web Consortium (W3C)⁴⁸, такие как XML.

DJVU

Название: DjVu⁴⁹

Создатель: AT&T Labs

Издатель: LizardTech Incorporated

Дата публикации: 1996 и далее

Идентификатор: <http://djvu.org/docs/DjVu3Spec.djvu> [версия 3]

Права: [Открытый стандарт]

Описание: разработан в основном для хранения сканированных изображений, в первую очередь, содержащих текст и рисунки.

Содержание: визуализация документов; структура документов

Связи: <http://djvu.org> [вебсайт DjVu]

HTML

Название: HTML • HyperText Markup Language⁵⁰

(язык разметки гипертекста)

Создатель: Berners-Lee, T

Издатель: World Wide Web Consortium (W3C)

Дата публикации: 1989 и далее

Идентификатор: <http://www.w3.org/TR/html4/> [версия 4.01]

Права: [Открытый стандарт]

Описание: язык разметки гипертекста был разработан Тимом Бернерсом Ли для сетевого взаимодействия документов. Этот язык базируется на SGML (является SGML приложением). Основное назначение HTML – это, скорее, визуализация (отображение) документов, а не представление структуры документа. Стандарт HTML поддерживается консорциумом World Wide Web. Стандарт каскадных таблиц стилей⁵¹ является дополнением к HTML, удобным для верстки и дизайна веб страниц. HTML версия 4 использует Unicode для стандартного кодирования символов.

Содержание: визуализация документов

Связи: SGML, XML, XHTML, Unicode

⁴⁸ [http://ru.wikipedia.org/wiki/World Wide Web Consortium](http://ru.wikipedia.org/wiki/World_Wide_Web_Consortium) (на русск. яз., прим. ред.)

⁴⁹ <http://djvu-reader.ru/> (на русск. яз., прим. ред.)

⁵⁰ <http://ru.wikipedia.org/wiki/HTML> (на русск. яз., прим. ред.)

⁵¹

http://ru.wikipedia.org/wiki/%D0%9A%D0%B0%D1%81%D0%BA%D0%B0%D0%B4%D0%BD%D1%8B%D0%B5_%D1%82%D0%B0%D0%B1%D0%BB%D0%B8%D1%86%D1%8B_%D1%81%D1%82%D0%B8%D0%BB%D0%B5%D0%B9 (на русск. яз.)

PDF

Название: PDF • **Portable Document Format**⁵²

(независимый от платформы формат документов)

Создатель: Adobe Systems Incorporated

Издатель: Adobe Systems Incorporated

Дата публикации: 1999

Идентификатор: [не известен]

Права: Копирайт Adobe Systems Incorporated. Все права защищены⁵³.

Описание: Стандарт для документов *de facto*. Разработан компанией Adobe как независимый от программной платформы способ отображения структуры и разметки документа. Он основан на стандарте визуализации документа Postscript⁵⁴. Формат pdf широко используется для публикации документов в сети и на CD-ROM.

Содержание: визуализация документа, структура документа

Связи: Postscript

<http://www.adobe.com/devnet/pdf.html> (сайт PDF)

RTF

Название: RTF • **Rich Text Format**⁵⁵ (формат обогащенного текста)

Создатель: Microsoft Corporation

Издатель: Microsoft Corporation

Дата публикации: 1999

Идентификатор: <http://msdn.microsoft.com/en-us/library/aa140277.aspx>
[версия 6.1]

Права: Копирайт Microsoft Corporation.

Описание: формат RTF предназначен для текста и графики и может использоваться различными устройствами вывода и операционными системами. Для форматирования документов на экране и при печати RTF использует наборы символов American National Standards Institute (ANSI), PC-8, Macintosh или IBM PC. С помощью формата RTF документы, созданные в различных операционных системах и с помощью разного программного обеспечения могут быть использованы в других операционных системах и приложениях.

Содержание: визуализация документа, структура документа

Связи: PDF

⁵² <http://ru.wikipedia.org/wiki/PDF> (на русск. яз., прим. ред.)

⁵³ PDF с 1 июля 2008 года является открытым стандартом ISO 32000 (прим. ред.)

⁵⁴ <http://ru.wikipedia.org/wiki/PostScript> (на русск. яз., прим. ред.)

⁵⁵ <http://ru.wikipedia.org/wiki/Rtf> (на русск. яз., прим. ред.)

SGML

Название: SGML • Standard Generalized Markup Language (стандартный обобщенный язык разметки⁵⁶)

Создатель: Goldfarb, Charles F

Издатель: International Organization for Standardization (ISO)

Дата публикации: 1986

Идентификатор: ISO 8879:1986

Права: [Открытый стандарт]

Описание: SGML – международный стандарт (ISO 8879) языка для структурированных данных и визуализации документов, лежащий в основе HTML и XML. Сам по себе SGML не очень широко распространен в области культуры. В начальный период внедрения SGML был реализован демонстратор CIMI CHIO⁵⁷. Позже в проектах в основном использовался XML, который является упрощением SGML.

Содержание: структура документа, кодирование документа

Связи: XML, HTML, <http://www.iso.org> (веб сайт ISO)

WORD

Название: MicrosoftWord Document Format⁵⁸ (формат для документов)

Создатель: Microsoft Corporation

Издатель: Microsoft Corporation

Дата публикации: 1983 и далее

Идентификатор: [недоступен]

Права: собственность Microsoft

Описание: формат для документов *de facto*. Формат .doc был разработан фирмой Microsoft для программного обеспечения Word и несколько раз изменялся. Формат широко используется для обмена документами. Спецификация формата не обнародована.

Содержание: визуализация документа, структура документа

⁵⁶ <http://ru.wikipedia.org/wiki/SGML> (на русск. яз.; прим. ред.)

⁵⁷ Для того, чтобы продемонстрировать использование SGML и Z39.50 Консорциум по обмену музейной информацией (CIMI, см. <http://xml.coverpages.org/cimi.html>) в 1994 году реализовал проект «Информация по культурному наследию онлайн» (CHIO, см. <http://www.cni.org/tfms/1994b.fall/perkins.html>) (прим. ред.)

⁵⁸ <http://office.microsoft.com/ru-ru/word/> (на русск. яз.; прим. ред.)

XML

Название: XML • eXtensible Markup Language (расширяемый язык разметки⁵⁹)

Создатель: World Wide Web Consortium (W3C)

Издатель: World Wide Web Consortium (W3C)

Дата публикации: 1996 и далее

Идентификатор: <http://www.w3.org/TR/xhtml1/>[revised 1 August 2002]

Права: [Открытый стандарт]

Описание: XML начинался как 'упрощенный SGML' для преодоления некоторых сложностей исходного формата. Вскоре он превзошел своего предшественника по популярности и теперь является наиболее распространенным языком для структурирования электронных документов. XML структурирует документ с помощью 'тэгов' (меток). Тэги могут быть любыми, но также могут определяться с помощью Определения типа документа (Document Type Definition - DTD⁶⁰) или с помощью XML-схемы. Для кодирования символов XML использует Unicode, поэтому он удобен в использовании в многоязычных и международных приложениях. Некоторые производные от XML языки также используются в качестве стандартов, например, XSLT⁶¹ (eXtensible Style Language and Transformation) и синтаксис для запросов XPath⁶².

Содержание: структура документа, кодирование документа

Связи: SGML, HTML, Unicode

⁵⁹ <http://ru.wikipedia.org/wiki/XML> (на русск. яз., прим. ред.)

⁶⁰ Document Type Definition (определение типа документа) – раздел текста, написанного на языке XML и других языках разметки, подробнее см. <http://ru.wikipedia.org/wiki/DTD> (на русск. яз., прим. ред.)

⁶¹ <http://ru.wikipedia.org/wiki/XSLT> (на русск. яз., прим. ред.)

⁶² <http://ru.wikipedia.org/wiki/XPath> (на русск. яз., прим. ред.)

2.3.2. Изображение

Для того, чтобы представить статическое изображение в цифровой форме, оно кодируется. В этом разделе мы рассмотрим 'растровые изображения', а не 'векторные изображения'. Растровые изображения обычно 'фотографические', т.е. изображение представлено определенным количеством горизонтальных и вертикальных элементов (пикселей). Каждый элемент имеет свое цветовое представление. Растровая графика⁶³ производится сканерами и цифровыми камерами. Если растровое изображение увеличить, то становятся видны отдельные пиксели.

BMP

Название: BMP • BitMap⁶⁴

Создатель: Microsoft Corporation

Издатель: Microsoft Corporation

Дата публикации: 1987

Идентификатор: [недоступен]

Права: Копирайт Microsoft Corporation

Описание: формат изображений bmp – это стандарт Microsoft Windows. Он кодирует черно-белые, 16-цветные, 256-цветные изображения, а также изображения Truecolor (24 бита). Изображения в палитре 16 цветов и 256 цветов можно сжать путем кодирования длин серий⁶⁵.

Содержание: формат изображения, растровая графика

DJVU

Название: DjVu⁶⁶

Создатель: AT&T Labs

Издатель: Lizardtech

Дата публикации: 1996 и далее

Идентификатор: <http://djvu.org/docs/DjVu3Spec.djvu>

Права: [Открытый стандарт]

Описание: разработан для хранения преимущественно сканированных изображений, особенно содержащих текст и рисунки. Содержит продвинутую технологию сжатия.

Содержание: визуализация документа, структура документа

Связи: PDF, <http://djvu.org> [вебсайт DjVu]

⁶³

http://ru.wikipedia.org/wiki/%D0%A0%D0%B0%D1%81%D1%82%D1%80%D0%BE%D0%B2%D0%B0%D1%8F_%D0%B3%D1%80%D0%B0%D1%84%D0%B8%D0%BA%D0%B0 (на русск. яз., прим. ред.)

⁶⁴ <http://ru.wikipedia.org/wiki/BMP> (на русск. яз., прим. ред.)

⁶⁵

http://ru.wikipedia.org/wiki/%D0%9A%D0%BE%D0%B4%D0%B8%D1%80%D0%BE%D0%B2%D0%B0%D0%BD%D0%B8%D0%B5_%D0%B4%D0%BB%D0%B8%D0%BD_%D1%81%D0%B5%D1%80%D0%B8%D0%B9 (на русск. яз., прим. ред.)

⁶⁶ <http://djvu-reader.ru/> (на русск. яз., прим. ред.)

GIF

Название: GIF⁶⁷ • **Graphical Interchange Format** (формат обмена графикой)

Создатель: CompuServe Interactive Services Incorporated

Издатель: CompuServe Interactive Services Incorporated

Дата публикации: 1987

Идентификатор: [недоступен]

Права: патент на компрессию LZW у компании Unisys.

Описание: GIF создан компанией CompuServe для онлайн-службы компании, но его спецификация была обнародована. GIF поддерживает до 256 цветов и алгоритм сжатия растровых данных Lempel Ziv Welch (LZW), запатентованный компанией Unisys) для уменьшения размера файла.

Содержание: формат изображения, растровая графика

JPG

Название: JPG • JPEG⁶⁸

Создатель: Joint Photographic Expert Group

Независимая JPEG Group

Издатель: International Organization for Standardization (ISO)

Дата публикации: 1990

Идентификатор: ISO/IEC 10918-1:1994

Права: [Открытый стандарт]

Описание: JPEG – это алгоритм сжатия статического изображения, основанный на том, что человеческий глаз не может зафиксировать тонкие различия цвета и контраста. JPEG – это алгоритм сжатия с потерей информации: чем больше сжатие, тем больше потери. Изображение, которое было сжато с помощью JPEG, полностью восстановить нельзя. Формат файла изображения, сжатого с помощью JPEG, называется JFIF. Именно этот формат имеется в виду, когда говорят о «JPEG».

Содержание: формат изображения, растровая графика

Связи: <http://www.iso.org> (веб сайт ISO)

⁶⁷ <http://ru.wikipedia.org/wiki/GIF> (на русск. яз., прим. ред.)

⁶⁸ <http://ru.wikipedia.org/wiki/JPEG> (на русск. яз., прим. ред.)

PNG

Название: PNG • **Portable Network Graphics** (независимая от платформы сетевая графика)⁶⁹

Создатель: World Wide Web Consortium (W3C)

Издатель: International Organization for Standardization (ISO)

Дата публикации: 2003

Идентификатор: ISO/IEC 15948:2003 (E)

Права: [Открытый стандарт]

Описание: PNG – расширяемый, независимый от платформы формат для хранения растровых изображений с хорошим сжатием без потерь. PNG обеспечивает свободную (незапатентованную) замену формата GIF и может также во многих случаях заменить формат TIFF. Поддерживаются изображения с индексированным цветом, Truecolor и черно-белые, а также альфа-канал⁷⁰ для прозрачности. Глубина цвета – от 1 до 32 бит.

Содержание: формат изображения, растровая графика

Связи: <http://www.w3.org/TR/PNG/>, <http://www.iso.org> (веб сайт ISO)

TIFF

Название: TIFF • **Tagged Image File Format** (формат файла для размеченного изображения)⁷¹

Создатель: Aldus Corp

Издатель: Adobe Systems Inc.

Дата публикации: 1992

Идентификатор:

<http://partners.adobe.com/public/developer/en/tiff/TIFF6.pdf> [версия 6.0]

Права: [Открытый стандарт]

Описание: TIFF – формат файла растрового статического изображения общего использования, совместимый с большинством сканеров и программного обеспечения для обработки изображений. Он не зависит от оборудования и используется в большинстве программных сред. Этот промышленный стандарт данных используется большинством производителей сканеров и программного обеспечения.

Содержание: формат изображения, растровая графика

⁶⁹ <http://ru.wikipedia.org/wiki/PNG> (на русск. яз., прим. ред.)

⁷⁰ Часть информации каждого пикселя, предназначенная для обеспечения прозрачности (прим. ред.)

⁷¹ <http://ru.wikipedia.org/wiki/TIFF> (на русск. яз., прим. ред.)

2.3.3. Audio

Оцифровка звука достигается путем дискретизации по времени. При преобразовании аналогового сигнала в цифровой отсчёт значений звукового сигнала берется через равные временные интервалы. При квантовании уровня каждый отсчёт заменяется числом, представляющим значение аналогового сигнала в момент дискретизации. Полученные значения уровня сигнала затем сохраняются в несжатом или сжатом виде.

Открытый стандарт для сжатого звука – MP3. Принадлежащие коммерческим компаниям стандарты дают лучшее качество звука или меньший размер файла (т.е. скорость передачи аудио информации).

AIFF

Название: **AIFF • Audio Interchange File Format**⁷² (формат файла обмена аудио информацией)

Создатель: Apple Computer Incorporated, Electronic Arts

Издатель: Apple Computer Incorporated

Дата публикации: 1988-1989

Идентификатор:

<http://www-mmsp.ece.mcgill.ca/Documents/AudioFormats/AIFF/Docs/AIFF-1.3.pdf> [Version 1.3]

Права: Apple Computer Incorporated [?]

Описание: аудио формат без сжатия, широко используемый на компьютерах Apple Macintosh. Формат, не дающий потерь качества, широко используется программным обеспечением для профессиональных аудио и видео систем.

Содержание: звуковой формат

⁷² <http://ru.wikipedia.org/wiki/AIFF> (на русск. яз., прим. ред.)

MP3

Название: MP3⁷³ • MPEG уровень 3

Coding of Moving Pictures and Associated Audio for Digital Storage Media

(кодирование движущихся изображений и связанного с ними звука для цифрового хранения)

Создатель: ITU-T (International Telecommunication Union Telecommunication Standardization Sector)

Экспертная группа Moving Pictures (ISO/IEC JTC1/SC29 WG11)

Издатель: International Organization for Standardization (ISO)

Дата публикации: 1993 и далее

Идентификатор: ISO/IEC 11172:1993, часть 3: Audio

ISO/IEC 13818, часть 3: Audio

ISO/IEC 14496, часть 3: Audio

(Дополнение 1: расширение аудио)

Права: [открытый стандарт]

Описание: аудио формат со сжатием, распространенный в Интернете. Являясь частью стандарта MPEG, этот стандарт может работать с аудиозаписями большого объема и уменьшает размер файла при незначительных потерях качества звука.

Содержание: звуковой формат

Связи: MPEG-1; MPEG-2; MPEG-3, <http://www.iso.org> (веб сайт ISO)

WAV

Название: WAV⁷⁴ • RIFF⁷⁵

Resource Interchange Format (формат обмена ресурсами)

Создатель: IBM

Microsoft Corporation

Издатель: Microsoft Corporation

Дата публикации: 1991

Идентификатор: <http://www.sonicspot.com/guide/wavefiles.html>

Права: [?]

Описание: содержит фрагментированную аудио информацию. Звуковая информация хранится в контейнере при помощи формата RIFF. Файл RIFF хранит данные фрагментами и содержит метаданные. Файл WAV может содержать звуковые клипы с различными шагами фрагментации, количеством каналов и пр.

Содержание: звуковой формат

Связи: AVI

⁷³ <http://ru.wikipedia.org/wiki/MP3> (на русск. яз., прим. ред.)

⁷⁴ <http://ru.wikipedia.org/wiki/WAV> (на русск. яз., прим. ред.)

⁷⁵ <http://ru.wikipedia.org/wiki/RIFF> (на русск. яз., прим. ред.)

WMA

Название: WMA •Windows Media Audio⁷⁶

Создатель: Microsoft Corporation

Издатель: Microsoft Corporation

Дата публикации: [?]

Идентификатор: [недоступен]

Права: Копирайт Microsoft Corporation

Описание: Принадлежащий фирме Microsoft конкурент формата MP3. Предназначен для распространения аудио информации через Интернет, особенно потокового аудио, используется программным обеспечением компании Microsoft. Включает управление правами, создает файлы небольшого размера и поэтому имеет преимущества в скорости передачи перед другими форматами, включая MP3. Кроме того, считается, что WMA дает лучшее качество звука.

Содержание: звуковой формат

Связи: WMV

REALAUDIO

Название: RealAudio

Создатель: RealNetworks Incorporated

Издатель: RealNetworks Incorporated

Дата публикации: [?]

Идентификатор: [недоступен]

Права: Копирайт Real Networks

Описание: формат для оптимизации доставки аудио через Интернет.

Содержание: звуковой формат

Связи: RealVideo, RealMedia

AU

Название: AU

Создатель: Sun Microsystems Incorporated

Издатель: Sun Microsystems Incorporated

Дата публикации: [?]

Идентификатор: [недоступен]

Права: Копирайт Sun Microsystems Incorporated

Описание: звуковой формат для систем Unix. Это стандартный аудио формат Java.

Содержание: звуковой формат

⁷⁶ http://ru.wikipedia.org/wiki/Windows_Media_Audio (на русск. яз., прим. ред.)

2.3.4. Видео

Оцифровка видео отличается от оцифровки статических изображений наличием шкалы времени. В принципе, процесс аналогичен получению растровых изображений, но для очень быстро меняющейся последовательности изображений. Количество изображений в секунду называется частотой кадров. Качество оцифрованного видео определяется тремя факторами: разрешением, глубиной цвета и частотой кадров. При оцифровке видео получаются данные большого объема, из-за этого очень важно сжатие. Сжатие основывается на том, что из всей последовательности изменяется лишь небольшое число кадров. С точки зрения стандартов кодирования, ситуация аналогична кодированию аудио, учитывая, что стандарты MPEG открытые. Кроме того, на рынке предлагается некоторое количество стандартов, принадлежащих частным компаниям.

AVI

Название: AVI⁷⁷ • **Audio Video Interleave** (чередование аудио и видео⁷⁸)

Создатель: Microsoft Corporation (для Intel)

Издатель: Microsoft Corporation

Дата публикации: [?]

Идентификатор: [недоступен]

Права: Копирайт Microsoft Corporation

Описание: самый старый видео формат для PC. Размер изображения при выводе на экран зависит от компьютера. Сжатие и развертывание функционально является частью программного обеспечения Microsoft *Video for Windows*, этот формат поддерживается большинством конфигураций программного и технического обеспечения. Размеры файлов очень большие, поэтому этот формат неудобен для доставки через Интернет.

Содержание: видео формат

FLV (Flash Video Format)

Название: FLV • **Flash Video Format**⁷⁹

Создатель: Macromedia (в настоящее время Adobe)

Издатель: Adobe Systems Incorporated

Дата публикации: 2002 и далее

Идентификатор:

http://www.adobe.com/devnet/devices/articles/flashlite_video_formats.html

Права: Копирайт Adobe Systems Incorporated

Описание: используется для доставки видео через Интернет. Просмотр файлов осуществляется с помощью специального плеера или программного обеспечения, встроенного в веб браузер. Становится *de facto* стандартом для видео, встроенного в веб страницы.

Содержание: видео формат

⁷⁷ <http://ru.wikipedia.org/wiki/AVI> (на русск. яз., прим. ред.)

⁷⁸ http://ru.wikipedia.org/wiki/Audio_Video_Interleave (на русск. яз., прим. ред.)

⁷⁹ http://ru.wikipedia.org/wiki/Flash_Video (на русск. яз., прим. ред.)

MOV (Quicktime)

Название: QuickTime⁸⁰

Создатель: Apple Computer Incorporated

Издатель: Apple Computer Incorporated

Дата публикации: 1991 и далее

Идентификатор: [недоступен]

Права: Копирайт Apple Computer Incorporated

Описание: Формат и система для видео и виртуальной реальности, принадлежащий компании Apple. Встроен в операционную систему Mac, может быть просмотрен на PC через специальный плеер, который можно скачать бесплатно. Частично технологии этого формата были использованы при разработке MPEG-4.

Содержание: видео формат

Связи: Quicktime VR, MPEG-4, <http://www.apple.com> (веб сайт Apple)

MP4

Название: MPEG-4⁸¹ • **Very-low bitrate audio-visual coding** (кодирование аудиовизуальной информации с очень низким битрейтом⁸²)

Создатель: Moving Pictures Expert Group (ISO/IEC JTC1/SC29 WG11)

Издатель: International Organization for Standardization (ISO)

Дата публикации: 1999 (версия 1), 2001 (версия 2)

Идентификатор: ISO/IEC 14496 части 1 - 10)

Права: [открытый стандарт]

Описание: Версия формата MPEG-2 с высоким сжатием. Версия 2 включает защиту данных и права на интеллектуальную собственность.

Содержание: видео формат

Связи: MPEG-2, MP3, QuickTime (на стадии разработки), <http://www.iso.org> (веб сайт ISO)

MPG

Название: MPEG-1⁸³ • **Coding of Moving Pictures and Associated Audio for Digital Storage Media** (кодирование движущегося изображения и связанного с ним звука для цифрового хранения)

Создатель: Moving Pictures Expert Group (ISO/IEC JTC1/SC29 WG11)

Издатель: International Organization for Standardization (ISO)

Дата публикации: 1993

Идентификатор: ISO/IEC 11172:1993 (Parts 1 to 5)

Права: [открытый стандарт]

Описание: создан как аналог формата видеозаписи в цифровой форме для записи на CD-ROM. Стандарт качественного телевизионного изображения с уровнем сжатия от 50 до 1.

Содержание: видео формат

Связи: MP3, <http://www.iso.org> (веб сайт ISO)

⁸⁰ <http://ru.wikipedia.org/wiki/QuickTime> (на русск. яз., прим. ред.)

⁸¹ <http://ru.wikipedia.org/wiki/MPEG-4> (на русск. яз., прим. ред.)

⁸² Характеристика цифрового аудио и видео: скорость прохождения битов информации (прим. ред.)

⁸³ <http://ru.wikipedia.org/wiki/MPEG-1> (на русск. яз., прим. ред.)

MPEG-2

Название: MPEG-2⁸⁴ • Coding of Moving Pictures and Associated Audio for Digital Storage Media

(кодирование движущегося изображения и связанного с ним звука для цифрового хранения)

Создатель: ITU-T (International Telecommunication Union Telecommunication Standardization Sector)

Moving Pictures Expert Group (ISO/IEC JTC1/SC29 WG11)

Издатель: International Organization for Standardization (ISO)

Дата публикации: 2000

Идентификатор: ISO/IEC 13818:2000 (части 1 - 11)

Права: [открытый стандарт]

Описание: усовершенствованный MPEG-1. Метод кодирования предназначен для записи высококачественного видео и аудио на DVD. Высококачественное телевизионное изображение с уровнем сжатия от 150 до 1.

Содержание: видео формат

Связи: MPEG-1, MP3, <http://www.iso.org> (веб сайт ISO)

RM (Real Media)

Название: RM • Real Media⁸⁵

Создатель: RealNetworks Incorporated

Издатель: RealNetworks Incorporated

Дата публикации: [?]

Идентификатор: [недоступен]

Права: Копирайт Real Networks

Описание: формат предназначен для доставки аудио информации через Интернет

Содержание: мультимедиа

Связи: RealAudio, RealVideo

SWF (Flash Movie)

Название: SWF • Small Web Format Flash Movie⁸⁶

Создатель: Macromedia (в настоящее время Adobe)

Издатель: Adobe Systems Incorporated

Дата публикации: 1996 и далее

Идентификатор:

http://www.adobe.com/devnet/swf/pdf/swf_file_format_spec_v10.pdf⁸⁷ [версия 10]

Права: Копирайт Adobe Systems Incorporated

Описание: первоначально предназначался для доставки анимационной векторной графики, сейчас используется для интерактивного аудио и видео. Может просматриваться через специальный плеер или через плеер, встроенный в веб браузер.

Содержание: анимационная векторная графика, интерактивность

⁸⁴ <http://ru.wikipedia.org/wiki/MPEG-2> (на русск. яз., прим. ред.)

⁸⁵ <http://ru.wikipedia.org/wiki/RealMedia> (на русск. яз., прим. ред.)

⁸⁶ http://ru.wikipedia.org/wiki/Adobe_Flash (на русск. яз., прим. ред.)

⁸⁷ недоступен (прим. ред.)

WMV (Windows Media Video)

Название: WMV • Windows Media Video⁸⁸

Создатель: Microsoft Corporation

Издатель: Microsoft Corporation

Дата публикации: [?]

Идентификатор: [недоступен]

Права: Копирайт Microsoft Corporation

Описание: принадлежащий Microsoft формат - конкурент MPEG-4. Предназначен для доставки видео, особенно потокового видео через Интернет, используется в продуктах Microsoft.

Содержание: видео формат

Связи: WMA

ASF

Название: ASF • Advanced Streaming Format⁸⁹ (продвинутый стриминговый формат)

Создатель: Microsoft Corporation, RealNetworks Incorporated

Издатель: Microsoft Corporation

Дата публикации: 2004

Идентификатор:

<http://www.microsoft.com/windows/windowsmedia/forpros/format/asfspec.aspx>
[версия 01.20.03]

Права: Копирайт Microsoft Corporation. Все права защищены

Описание: формат предназначен для сжатия и доставки потокового видео/аудио через Интернет.

Содержание: видео формат

⁸⁸ http://ru.wikipedia.org/wiki/Windows_Media_Video (на русск. яз., прим. ред.)

⁸⁹ http://ru.wikipedia.org/wiki/Advanced_Systems_Format (на русск. яз., прим. ред.)

2.3.5. Виртуальная реальность

VRML97

Название: VRML97 • Virtual Reality Modelling Language⁹⁰ (язык моделирования виртуальной реальности)

Создатель: VRML Consortium Incorporated

Объединенный технический комитет ISO/IEC JTC 1, Информационные технологии, подкомитет 24, компьютерная графика и обработка изображений

Издатель: International Organization for Standardization (ISO)

Дата публикации: 1997

Идентификатор: ISO/IEC 14772-1:1997

Права: [открытый стандарт]

Описание: формат для описания объектов и миров 3D. Разработан для доступа через Интернет, интерактивности, а также для других сценариев (например, автономных компьютеров). Может представлять статические и динамические 3D проекты и позволяет делать ссылки на другие мультимедиа объекты, такие как текст, статические изображения, аудио и видео.

Содержание: виртуальная реальность

Связи: X3D, <http://www.web3d.org> (веб сайт консорциума Web3D), <http://www.iso.org> (веб сайт ISO)

X3D / EXTENSIBLE 3D

Название: X3D⁹¹ • Extensible 3D (расширяемый 3D)

Создатель: Web3D Consortium

Издатель: International Organization for Standardization (ISO) [в процессе разработки]

Дата публикации: 2007

Идентификатор: ISO/IEC FDIS 19775-1.2:2008

Права: [открытый стандарт]

Описание: формат основан на XML и выполняет те же функции, что и VRML97. В настоящее время находится в разработке как стандарт ISO.

Содержание: виртуальная реальность

Связи: VRML97, XML, <http://www.web3d.org> (веб сайт консорциума Web3D), <http://www.iso.org> (вебсайт ISO)

⁹⁰ <http://ru.wikipedia.org/wiki/VRML> (на русск. яз., прим. ред.)

⁹¹ <http://ru.wikipedia.org/wiki/X3D> (на русск. яз., прим. ред.)

QUICKTIME VR

Название: QuickTime VR⁹²

Создатель: Apple Computer Incorporated

Издатель: Apple Computer Incorporated

Дата публикации: 1995 и далее

Идентификатор: [недоступен]

Права: Копирайт Apple Computer Incorporated

Описание: формат для виртуальной реальности, принадлежащий компании Apple, и система для создания контента. Встроен в операционную систему Mac, можно просматривать на PC через плеер, который предоставляется для скачивания бесплатно. Является часть системы QuickTime, 'склеивает' отдельные фотографии в панораму, производит 3D представления объектов и позволяет связывать их вместе, создавая 'декорацию'.

Содержание: виртуальная реальность, панорамы

Связи: QuickTime, <http://www.apple.com> (веб сайт Apple)

⁹² <http://ru.wikipedia.org/wiki/QuickTime> (на русск. яз., прим. ред.)

2.3.6. Векторная графика⁹³

EPS

Название: EPS • Encapsulated Postscript⁹⁴ (инкапсулированный Postscript)

Создатель: Adobe Systems Inc.

Издатель: Adobe Systems Inc.

Дата публикации: 1992

Идентификатор:

http://partners.adobe.com/public/developer/en/ps/5002.EPSF_Спец.pdf [Версия 3.0]

Права: [открытый стандарт]

Описание: хотя, строго говоря, EPS – это не формат для изображений, он часто используется для представления рисунков. EPS основан на языке PostScript - это язык верстки страниц, часто используемый в принтерах. EPS дает возможность 'инкапсулировать' в другие документы страницы, закодированные с помощью Postscript.

Содержание: язык верстки страниц, векторная графика

SVG

Название: SVG • Scalable Vector Graphics⁹⁵ (масштабируемая векторная графика)

Создатель: World Wide Web Consortium (W3C)

Издатель: World Wide Web Consortium (W3C)

Дата публикации: 2003

Идентификатор: <http://www.w3.org/TR/SVG/index.html> [версия 1.1]

Права: [открытый стандарт]

Описание: SVG – это язык, описывающий двухмерную графику с помощью XML. SVG предназначен для трех типов графических объектов: формы векторной графики (например, схемы, состоящие из прямых и кривых линий), изображения и текст. Графические объекты могут быть сгруппированы, трансформированы и встроены в подготовленные ранее объекты. Хотя первоначально стандарт SVG предназначался для документов в векторной графике, он может также использоваться для растровых изображений.

Содержание: векторная графика

Связи: XML

⁹³

http://ru.wikipedia.org/wiki/%D0%92%D0%B5%D0%BA%D1%82%D0%BE%D1%80%D0%BD%D0%B0%D1%8F_%D0%B3%D1%80%D0%B0%D1%84%D0%B8%D0%BA%D0%B0 (на русск. яз., прим. ред.)

⁹⁴ http://ru.wikipedia.org/wiki/Encapsulated_PostScript (на русск. яз., прим. ред.)

⁹⁵ http://ru.wikipedia.org/wiki/Scalable_Vector_Graphics (на русск. яз., прим. ред.)

2.4. Другие технические стандарты

Технические стандарты относятся к реализации различных функций:

- **Поиск и вывод информации** – средства формулировки запроса к информационной системе, получение и обработка результатов поиска.
- **Передача информации** – доставка цифровых данных. Передача информации может быть неспецифической (низкого уровня, например, TCP/IP) или более связанная с конкретной программной средой (например, http). Сегодня TCP/IP – это основной протокол низкого уровня. Более специализированные протоколы конструируются на базе TCP/IP.
- **Кодировка символов** – кодировка символов определяет, как хранятся тексты в компьютерной системе.

2.4.1. Поиск и вывод информации

OAI-PMN

Название: OAI-PMN⁹⁶ • **Open Archives Initiative Protocol for Metadata Harvesting (версия 2.0) (протокол для сбора метаданных)**

Создатель: Sompel, Herbert van de, Lagoze, Carl

Издатель: Open Archives Initiative

Дата публикации: 2002

Идентификатор: <http://www.openarchives.org/OAI/openarchivesprotocol.html>

Права: [открытый стандарт]

Описание: протокол инициативы «Открытые архивы» предназначен для сбора информации, хранящейся в базах данных или хранилищах, которые невозможно собрать с помощью ‘стандартных’ протоколов http/html/. Протокол основан на запросах http, результаты даются в форме XML. Поддерживается разный синтаксис записей, но при этом обязательно использование метаданных Дублинского ядра.

Содержание: протокол сбора данных

Связи: XML, Dublin Core

⁹⁶ <http://socionet.ru/docs/infohubs.htm> (на русск. яз., прим. ред.)

SQL

Название: SQL • Structured Query Language⁹⁷ (язык структурированных запросов)

Создатель: International Organization for Standardization (ISO)

Издатель: International Organization for Standardization (ISO)

Дата публикации: 1987

Идентификатор: ISO 9075

Права: [открытый стандарт]

Описание: SQL – язык запросов для реляционных баз данных. Он был разработан компанией IBM в 1970-е годы и прошел процесс стандартизации в 1980-е годы, а второе издание SQL (SQL2) было принято в 1992 году. Синтаксис основан на операциях с двумерными таблицами, которые лежат в основе реляционных баз данных.

Содержание: язык запросов, реляционные базы данных⁹⁸

Связи: ODBC, <http://www.iso.org> (веб сайт ISO)

Z39.50

Название: Z39.50⁹⁹ • Information and documentation — Information retrieval (Z39.50) — Application service definition and protocol specification (информация и документация – информационный поиск – приложение и спецификация протокола)

Создатель: ANSI/NISO

Издатель: International Organization for Standardization (ISO)

Дата публикации: 1988

Идентификатор: ISO 23950:1998

Права: [открытый стандарт]

Описание: Z39.50 – протокол поиска и сбора информации в распределенной неоднородной среде с архитектурой «клиент - сервер». Запросы формулируются с помощью Reversed Polish Notation (RPN)¹⁰⁰, результаты поиска могут быть в разных форматах, включая формат MARC (Machine Readable Catalogue)¹⁰¹. Поисковый компонент занимается обработкой и выполнением запроса, компонент сбора информации занимается обработкой результатов поиска.

Содержание: протокол поиска и сбора информации

Связи: MARC, <http://www.iso.org> (веб сайт ISO)

⁹⁷ http://ru.wikipedia.org/wiki/Structured_Query_Language (на русск. яз., прим. ред.)

⁹⁸ http://ru.wikipedia.org/wiki/%D0%A0%D0%B5%D0%BB%D1%8F%D1%86%D0%B8%D0%BE%D0%BD%D0%BD%D1%8B%D0%B5_%D0%B1%D0%B0%D0%B7%D1%8B_%D0%B4%D0%B0%D0%BD%D0%BD%D1%8B%D1%85 (на русск. яз., прим. ред.)

⁹⁹ <http://z3950.ruslan.ru/> (на русск. яз., прим. ред.)

¹⁰⁰ Обратная польская запись была предложена в 1920 году Яном Лукасевичем как способ записи математического выражения без использования круглых и квадратных скобок. (на русск. яз., прим. ред.)

¹⁰¹ <http://ru.wikipedia.org/wiki/MARC> (на русск. яз., прим. ред.)

2.4.2. Передача информации

ТСР/IP

Название: **ТСР/IP**¹⁰² • **Transmission Control Protocol/Internet Protocol**
(протокол управления передачей данных/Интернет протокол)

Создатель: Internet Engineering Task Force

Издатель: World Wide Web Consortium (W3C)

Дата публикации: 1981

Идентификатор: RFC793, RFC1122

Права: [открытый стандарт]

Описание: протокол управления транспортом (TCP) и интернет-протокол (IP) работают вместе и обеспечивают основу движения информации через Интернет. Протокол IP работает на сетевом уровне. На этом уровне не дается никаких гарантий корректной доставки данных. TCP работает 'поверх' IP и обеспечивает гарантированную доставку корректных данных. TCP, в отличие от IP является протоколом, ориентированном на коннективность.

Содержание: Интернет, передача данных

Связи: HTTP, FTP

HTTP

Название: **HTTP** • **HyperText Transfer Protocol**¹⁰³ (протокол передачи гипертекста)

Создатель: Berners-Lee, Tim

Издатель: World Wide Web Consortium (W3C)

Дата публикации: 1991

Идентификатор: <http://www.w3.org/Protocols/>

Права: [открытый стандарт]

Описание: протокол передачи гипертекста работает 'поверх' TCP/IP и предназначен для передачи запрошенных гипертекстовых документов по каналам Интернет. В рамках запроса http передается набор параметров, который можно использовать для протоколов высокого уровня, таких как OAI.

Содержание: передача гипертекста

Связи: TCP/IP, OAI

FTP

Название: **FTP** • **File Transfer Protocol**¹⁰⁴ (протокол передачи файлов)

Создатель: Neigus, Nancy J.

Издатель: Internet Engineering Task Force

Дата публикации: 1973

Идентификатор: <http://www.ietf.org/rfc/rfc542.txt>

Права: [открытый стандарт]

Описание: протокол FTP – это независимый от компьютерной / операционной системы протокол для передачи файлов. Он связан с началом Интернета и может считаться одним из элементов фундамента сети.

Содержание: передача файлов

Связи: TCP/IP

¹⁰² <http://ru.wikipedia.org/wiki/TCP/IP> (на русск. яз., прим. ред.)

¹⁰³ http://ru.wikipedia.org/wiki/HyperText_Transfer_Protocol (на русск. яз., прим. ред.)

¹⁰⁴ http://ru.wikipedia.org/wiki/File_Transfer_Protocol (на русск. яз., прим. ред.)

2.4.3. Кодирование символов

ASCII • American Standard Code for Information Interchange

Название: ASCII¹⁰⁵ • American Standard Code for Information Interchange
(американский стандартный код для обмена информацией)

Создатель: ANSI (American National Standards Institute)

Издатель: ANSI

Дата публикации: 1967

Идентификатор: [?]

Права: [открытый стандарт]

Описание: 7-битовый код для символов (букв, цифр, специальных знаков) в компьютерных системах. ASCII был создан в 1963 году, но окончательная версия была выпущена в 1967 году. Набор символов ASCII содержит 128 различных знаков, из которых 95 можно напечатать, а 33 являются управляющими знаками.

Содержание: кодирование символов

Связи: ISO 8859-1, Unicode

ISO 8859-1 / ISO LATIN 1

Название: ISO 8859-1 • ISO Latin 1¹⁰⁶

Создатель: International Organization for Standardization (ISO)

Издатель: International Organization for Standardization (ISO)

Дата публикации: 1998

Идентификатор: ISO-8859-1, ECMA-94

Права: [открытый стандарт]

Описание: 95 знаков ASCII недостаточно для представления всех европейских языков. Для решения этой задачи European Computer Manufacturers Association (ECMA-8) и фирма IBM создали 8-битовую версию. IBM предложила разные *кодовые страницы* для разных языков. 8-битовый стандарт ECMA позднее был принят международной организацией по стандартизации (ISO) под регистрационным номером ISO-8859-1.

Содержание: кодирование символов

Связи: ASCII, Unicode, <http://www.iso.org> (веб сайт ISO)

¹⁰⁵ <http://ru.wikipedia.org/wiki/ASCII> (на русск. яз., прим. ред.)

¹⁰⁶ <http://ru.wikipedia.org/wiki/ISO-8859-1> (на русск. яз., прим. ред.)

UNICODE

Название: Unicode¹⁰⁷

Создатель: Unicode Incorporated

Издатель: International Organization for Standardization (ISO)

Дата публикации: 1992

Идентификатор: ISO/IEC 10646 ECMA-94

Права: [открытый стандарт]

Описание: 16-битовое представление символов, таких как буквы и цифры, в компьютерных системах. Стандарт Unicode был создан для расширения 8-битового набора знаков и перехода к универсальной кодировке знаков для западных языков, а также для письменностей, не использующих латинский алфавит, таких как китайская, арабская, кириллическая и иврит. Версия 3.0 стандарта Unicode содержит 49,194 знаков. Unicode использует различные схемы кодирования, включая широко используемую переменную длину кода UTF-8. Unicode – набор символов для XML по умолчанию.

Содержание: кодирование символов, нелатинские письменности

Связи: ASCII, XML, <http://www.iso.org> (веб сайт ISO)

¹⁰⁷ <http://ru.wikipedia.org/wiki/Unicode> (на русск. яз., прим. ред.)

3. Выводы

3.1. Стандарты метаданных в сфере культуры

Существует ограниченное число ключевых стандартов, которые широко используются в Европе и в мире. Они используются в наиболее успешных проектах, но, как показало исследование, для достижения совместимости надо пройти еще очень долгий путь. В будущих проектах также необходимо принимать во внимание национальные стандарты некоторых стран.

Дублинское ядро (DC) интенсивно используется как 'минимальный стандарт' метаданных для поиска ресурсов, в том числе и в Европейской цифровой библиотеке Europeana. Однако необходимо установить соответствие DC и секторальных стандартов метаданных и на основе этого организовать обмен данными. В этом отношении возможное решение состоит в использовании XML и в особенности связанного с XML языка XSLT¹⁰⁸ (Extensible Stylesheet Language Transformations – трансформации расширяемого языка стилей).

XSLT используется для трансформации одного типа документа XML в другой тип документа XML (или в форму, которую читает человек). Т.е. необходимо разработать механизм преобразования метаданных из специфических для определенного сектора культуры стандартов в метаданные стандарта Дублинского ядра. Однако, вряд ли возможна обратная трансформация: из стандарта Дублинского ядра в секторальный стандарт, за исключением ограниченного числа элементов метаданных.

3.2. Технические стандарты

Существует много технических стандартов для разнообразных мультимедиа ресурсов, в том числе, много открытых стандартов.

Хороший современный справочник стандартов, которые должно использовать учреждение культуры, разработан и опубликован проектом Minerva: «Техническое руководство для программ создания цифровых информационных ресурсов по культуре»¹⁰⁹.

Этот справочник имеет много преимуществ:

Многоязычие – для учреждения культуры большое преимущество - пользоваться справочником на родном языке. Приветствуются переводы на другие языки.

Справочник предназначен для специалистов по культуре – чисто техническое руководство трудно для понимания гуманитариям, составляющим большинство сотрудников учреждений культуры.

Обновление – очень важно, чтобы техническое руководство постоянно обновлялось, учитывая скорость изменения технологий.

¹⁰⁸ <http://ru.wikipedia.org/wiki/XSLT> (на русск. яз., прим. ред.)

¹⁰⁹ Technical Guidelines for Digital Cultural Content Creation Programmes
<http://www.minervaeurope.org/interoperability/technicalguidelines.htm> [с ссылками на разные версии]
(на английском, немецком, голландском, французском, итальянском, греческом языках)

4. Указатель стандартов

AIFF	стр. 35
ASCII	стр. 48
ASF	стр. 41
AU	стр. 37
AVI	стр. 38
BMP	стр. 32
CDWA	стр. 16
CIDOC-CRM	стр. 25
DJVU (текст)	стр. 28
DJVU (изображение)	стр. 32
DUBLIN CORE	стр. 27
EAD	стр. 18
EPS	стр. 44
FLV (Flash Video Format)	стр. 38
FRBR	стр. 20
FTP	стр. 47
GIF	стр. 33
HTML	стр. 28
HTTP	стр. 47
ISAAR (CPF)	стр. 19
ISAD (G)	стр. 18
ISO 2709	стр. 21
ISO 8859-1 / ISO LATIN 1	стр. 48
JPG	стр. 33
LIDO	стр. 26
MAB2	стр. 20
MARC	стр. 21
METS	стр. 22
MIDAS	стр. 24
MODS	стр. 22
MOV (Quicktime)	стр. 39
MP3	стр. 36
MP4	стр. 39
MPEG1	стр. 39
MPEG2	стр. 40
MUSEUMDAT	стр. 16
OAI-PMH	стр. 45
OBJECT ID	стр. 17
PDF	стр. 29
PNG	стр. 34
QUICKTIME VR	стр. 43
REALAUDIO	стр. 37

RM (Real Media)	стр. 40
RTF	стр. 29
RUSMARC	стр. 23
SGML	стр. 30
SPECTRUM	стр. 17
SQL	стр. 46
SVG	стр. 44
SWF (Flash Movie)	стр. 40
TCP/IP	стр. 47
TEI	стр. 27
TIFF	стр. 34
UNICODE	стр. 49
VRA	стр. 25
VRML97	стр. 42
WAV	стр. 36
WMA	стр. 37
WMV (Windows Media Video)	стр. 41
WORD	стр. 30
X3D / EXTENSIBLE 3D	стр. 42
XML	стр. 31
Z39.50	стр. 46