Министерство образования и науки Российской Федерации

Государственное образовательное учреждение высшего профессионального образования

Горно-Алтайский государственный университет

Экономический факультет

Кафедра прикладной информатики в экономике

Реферат

По курсу «Мировые информационные ресурсы»

«Стандарты ODA и SGML»

Выполнил:

Разгоняев Александр Сергеевич

3 курс, 837 группа

Проверил:

Кыров Виктор Николаевич

Реферат защищён

«\_\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_\_2010 года

Оценка \_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(подпись руководителя)

г. Горно-Алтайск 2010

Содержание

Введение 3

Глава 1. Object Data Base Markup Language 4

1.1 Основные положения 4

1.2 Технология ODA 6

1.3 Практическое применение 7

Глава 2. Стандарт SGML 9

2.1 Что такое SGML 9

2.2 Терминология 11

2.3 Особенности SGML 12

Заключение 16

Список литературы 17

**Введение**

Информация является одним из ценнейших ресурсов человеческой деятельности, представляет собой сведения о лицах, предметах, фактах, событиях, явлениях и процессах независимо от формы их представления.

Своевременное получение и обработка полной, достоверной информации определяют эффективность управления предприятием, учреждением или организацией. Большая часть информации, используемой в управленческой деятельности, фиксируется в документах. Документированная информация (документ) зафиксированная на материальном носителе путем документирования информация с реквизитами, позволяющими определить такую информацию или ее материальный носитель

Современный уровень развития компьютерной техники обеспечивает существование и повсеместное распространение информационных технологий, позволяющих преобразовывать документную информацию в электронный вид. От повышения оперативности и четкости в работе с документами во многом зависит эффективность работы организации в целом. В последнее время на смену бумажным документам стремительно приходит электронный документооборот и вводится понятие электронного документа как документа, «в котором информация представлена в электронно-цифровой

форме».

В документационном обеспечении управления и производства при подготовке многих видов документов в настоящее время используются различные системы, в основном это – текстовые процессоры, которые позволяют быстро и качественно подготавливать как электронные, так и бумажные документы.Параллельно используются единые стандарты применяемые к электронным документам которыми и являются SGML и ODA. О которых и будет написано в данной работе.

**Глава 1. Object Data Base Markup Language**

**1.1 Основные положения**

В настоящее время общепринятым и совместно используемым единым стандартом хранения и передачи структурированных данных является стандарт ODA (Object Data Base Markup Language), т.е. такой формат данных, который не привязан ни к платформе, ни к конкретному программному пакету.

ODA – язык описания структурированных данных объектного типа, который разработан на базе языка XML, рекомендованного в качестве стандарта. Документы формата ODA являются объектными хранилищами данных любого типа, структуры и содержания. Структура описания данных ODA позволяет представить в данном формате любые данные из существующих информационных систем, хранилищ и баз данных. Благодаря универсальному описанию информационных структур хранение, передача, анализ, интерпретация данных может осуществляться инструментами, создаваемыми любыми независимыми разработчиками без привязки к информационному контексту, но при этом все они будут полностью совместимы между собой.

Формат ODA равнозначно воспринимается всеми операционными системами, в связи с чем, отсутствует платформенная привязанность и сложности с межплатформенными отношениями и поэтому является основой для единого формата хранения и передачи структурированных данных

ODA (Object Data Base Markup Language) – язык описания структурированных данных объектного типа, который разработан на базе языка XML (Extensible Markup Language), рекомендованного консорциумом W3C в качестве стандарта, с целью наращивания функциональных возможностей и повышения уровня взаимодействия в Сети 6 октября 2000 года.

Технология ODA изначально задумывалась как некий инструмент, способный быстро и легко создать информационные модули для каждого процесса, существующего внутри предприятия, а затем объединить (интегрировать) все модули в единую систему. Поэтому сфера применения этого продукта очень широка.

ODA не повторяет традиционные, стандартные решения, а позволяет использовать мощный и современный принцип объектной информации.

Вот основные тезисы, положенные в основу технологии ODA.

Любую информацию можно представить в виде объектов, имеющих собственное название и некий набор свойств, описывающих их.

Применяя различные наборы и типы свойств можно описать абсолютно любые информационные единицы и объекты.

Однородные по свойствам объекты можно объединить в информационные разделы, в том числе с возможностью наследования и преемственности объектов и их свойств.

В качестве объектов информационной системы может выступать абсолютно любая информация, уже накопленная в электронном виде (файлы, документы и т.д.). Достаточно только присвоить им имя и код и указать системе их физическое местонахождение и логическую связь с другими разделами и объектами системы.

В формировании информационной структуры должны принимать участие не программисты, а непосредственно те, кто будет ее использовать. Только в этом случае система не будет отторгнута пользователями и приобретет прикладное значение, наиболее соответствующее требованиям специалистов и задач, стоящих перед ними.

Технология настройки системы должна быть простой, доступной и понятной для любого пользователя, независимо от его знаний в области создания и программирования информационных систем и баз данных.

Система должна быть мобильной. Это значит, что она должна позволять вносить любые изменения и добавления в структуру в любой момент времени без потерь информации и функциональности.

Система должна иметь возможность подключаться к различным внешним источникам информации и использовать их информацию так же легко, как свою.

Документы формата ODA являются объектными хранилищами данных любого типа, структуры и содержания.

Структура описания данных ODA позволяет представить в данном формате любые данные из существующих информационных систем, хранилищ и баз данных.

Благодаря универсальному описанию информационных структур хранение, передача, анализ, интерпретация данных может осуществляться инструментами, создаваемыми любыми независимыми разработчиками без привязки к информационному контексту, но при этом все они будут на 100% совместимы между собой.

Формат ODA равнозначно воспринимается всеми операционными системами, в связи с чем, отсутствует платформенная привязанность и сложности с межплатформенными отношениями.

Для полноценной работы с данными ODA необходим ODA сервер.

Для доступа к ODA данным достаточно любого Web-броузера, поддерживающего HTML, XML и JavaScript.

**1.2 Технология ODA**

Технология описания данных с помощью стандарта ODA позволяет создавать информационные структуры любого уровня сложности, сохраняя при этом простоту чтения и интерпретации.

В основе стандарта лежат 5 принципов построения данных:

1. Описание строения базы в отдельном файле.

Параметры и структура каждой базы данных описывается в файле заголовка этой базы в формате xml. Этот файл содержит параметры самой базы, описание всех полей и табличных частей, используемых событий (действий).

2. Создание из каждой записи базы отдельного файла.

В отличие от существующих объектных баз данных, в котором все записи базы составляют один файл, по стандарту ODA каждая запись базы данных - это отдельный файл в формате xml. Все файлы-объекты хранятся в папке OBJECTS. Структура каждого файла состоит из заполненных полей, определенных заголовком данной базы.

3. Возможность использования необязательных полей.

Кроме полей, определенных в заголовке данной базы, в файле объекта возможны так называемые «необязательные» поля. Это могут быть поля другой базы данных; поля, доступные для чтения другой программой; поля, доступные для чтения другим пользователем и т.п.

4. Использование глобального уникального идентификатора (GUID) для каждой базы данных.

В момент создания базы ей присваивается глобальный уникальный идентификатор (GUID).

5. Использование связей между базами через GUID.

Использование GUID’а позволяет работать с различными базами не только на отдельном компьютере или в локальной сети, но и в сети Интернет.

**1.3 Практическое применение**

Основное назначение технологии ODA - создание комплексных информационных систем для предприятий различных видов деятельности и масштабов.

С помощью ODA могут быть автоматизированы практически все информационные процессы и операции предприятия: бухгалтерский учет, документооборот, кадровая деятельность, юридическое обеспечение, складской учет, управление производством, управление сбытом, управление маркетингом, финансовый анализ, бюджетирование, управленческое планирование и т.д. и т.п.

Другими словами, с помощью технологии ODA можно создать комплексную систему управления предприятием.

Технология ODA может иметь различные сферы применения. создание информационных хранилищ данных распределенных в сети. создание информационных систем для предприятий и организаций, использующих как собственные хранилища, так и размещенные в сети. организация взаимодействия между любыми информационными системами и приложениями, разработка универсальных аналитических и прочих инструментов, способных обрабатывать данные ODA без учета контекста хранимой информации, разработка информационных сайтов и порталов с сервисом не только по доступу к данным, но и по их добавлению и изменению, любые сферы деятельности, где необходима эффективная работа со структурированными данными.

Существует масса общепринятых стандартов и форматов для различных типов информации (DOC, HTML, TXT, XLS, JPEG, BMP, MP3, AVI и т.д.). Но, как ни странно, до сих пор не существовало самого главного стандарта – единого стандарта структурированных данных. Имеется в виду такой формат данных, который не привязан ни к платформе, ни к конкретному программному пакету, а являлся бы просто общепринятым и совместно используемым стандартом хранения и передачи структурированных данных.

При определенных действиях по внедрению и продвижению, ODA имеет все шансы стать основой для единого формата хранения и передачи структурированных данных.

**Глава 2. Стандарт SGML**

**2.1 Что такое SGML**

Стандарт ISO 8879 Information Processing - Text and Office System - Standard Generalized Markup Language (SGML) определяет обобщенный стандартный язык разметки текста, способ описания структуры документа, а также формат вставляемых в документ описательных меток. С точки зрения стандарта SGML, документ рассматривается как совокупность:

1. содержания (информации, содержащейся в документе в текстовой, графической и мультимедийной форме);
2. данных о структуре документа (взаимосвязи глав, разделов, параграфов, ссылок, прав доступа к элементам документа);
3. данных о стиле оформления документа (используемых шрифтах, интервалах, размерах полей, способе нумерации и т.д.).

Структура документа задается при помощи таблицы «Определения типа документа» (DTD) (в терминах стандарта – Document Type Definition), описывающего его структуру. В DTD указывают соответствие символов и их кодов, максимальные длины используемых идентификаторов, способ представления ограничителей для тегов, другие возможные соглашения, синтаксис DTD, а также тип и версию документа. DTD задает взаимосвязь глав, заголовков глав, разделов и других фрагментов текста, образующих документ. Стандарт SGML устанавливает такие множества символов и правил для представления информации, которые позволяют различным системам правильно распознавать и идентифицировать эту информацию, следовательно, SGML можно назвать метаязыком для семейства конкретных языков разметки. В частности, подмножествами SGML можно считать языки разметки XML (Extensible Markup Language - расширяемый язык описания документа) и HTML. При этом XML более удобен, чем SGML: легче воспринимается, приспособлен для использования в современных браузерах, сохраняет возможности SGML.

Для конкретных приложений создаются свои варианты (словари) XML. Известны варианты для математики, химии, медицины. Применение стандарта SGML для создания структурированных документов дает значительные преимущества. Фактически документ преобразуется в базу данных, допускающую манипуляции с элементами документа в соответствии с заданным DTD. В свою очередь, DTD разрабатывается в соответствии с назначением документа.

Стандарт ISO/IEC 10179 Document Style Semantics and Specification Language (DSSSL) определяет язык для описания правил и формата отображения SGML-документов при выводе на экран, печать или иное устройство отображения.

Сегодня все чаще можно встретить аббревиатуру SGML. Многие виды документации поставляются в формате SGML. Популярный язык разметки Web-страниц HTML является приложением SGML.

Стандарт SGML (ISO 8879). SGML - Standard Generalized Markup Language, стандартный обобщенный язык разметки. Разработан компанией IBM в 80-е гг. для составления электронной документации на системы вооружения Министерства обороны США.

Стандарт предназначен для унификации представления документации в автоматизированных системах. Первоначально ориентирован на текстовые данные, но охватывает и гипермедиа.

Появляется все больше прикладных программ для работы с SGML: редакторы, средства форматирования, системы документооборота, использующие SGML в качестве формата хранения документов... SGML-продукты выпускают Adobe, Corel, SoftQuad, Microsoft. В SGML готовятся книги издательства O'Reilly, техническая документация IBM, Sun и OSF. Основывается на SGML информационная технология многих ведущих машиностроительных и авиакосмических фирм всего мира.

Появившаяся в начале восьмидесятых SGML-технология переживает сейчас период расцвета. Что это такое? Что она дает своим пользователям? Как с ней работать? На эти и другие вопросы я пытаюсь ответить этой статьей.

**2.2 Терминология**

SGML (Generalized Markup Language) разработан в недрах вездесущей корпорации IBM. Его наследник SGML (Standard Generalized Markup Language) принят в 1986 году в качестве международного стандарта[1] для определения независимых от устройств ввода/вывода, независимых от вычислительной среды методов представления текстов в электронной форме. Более точно, SGML -- это метаязык, то есть средство формального описания языка, в данном случае, языка разметки.

Исторически слово разметка использовалось для описаний аннотаций или других обозначений внутри текста, предназначенных для указаний составителю или "верстальщику" того, как именно конкретное место должно быть напечатано или сверстано. Примеры включают подчеркивание волнистой чертой, обозначающее курсив, специальные значки для пропуска фраз или их печати конкретным шрифтом, и так далее. Когда форматирование и печать текстов стали автоматизированными, этот термин стал охватывать все виды специальных кодов разметки, вставляемых в электронные тексты для управления форматированием, печатью или другой обработкой.

Обобщая, разметку, или кодировку, определяют как любое средство сделать явным интерпретацию текста. На банальном уровне все напечатанные тексты кодированы в этом смысле: знаки препинания, использование заглавных букв, расположение букв по странице, даже интервалы между словами можно считать в какой-то степени разметкой, функция которой -- помочь человеку, читающему текст, определить, где кончается одно слово и начинается другое, или как идентифицировать особенности структуры, такие как заголовки, или простые синтаксические единица вроде подчиненных предложений. Кодирование текста для компьютерной обработки, в принципе, как расшифровка манускрипта с пергамента, -- процесс делания явным неявного или подразумеваемого, процесс указания пользователю того, как должно интерпретироваться содержимое текста.

Под языком разметки понимают набор соглашений о разметке, применяемых для кодирования текстов. Язык разметки должен специфицировать, какая разметка допустима, какая разметка обязательна, как отличить разметку от текста и что разметка значит. SGML предоставляет решения для первых трех задач, отдельная документация обычно необходима для последней.

**2.3 Особенности SGML**

Три характеристики SGML отличают его от прочих языков разметки.

В первую очередь это описательная разметка. Система с описательной разметкой использует коды разметки, которые просто предоставляют названия для категоризации частей документа. Коды разметки, такие как <para> или \end{list}, просто идентифицируют порцию документа и утверждают, что "она является параграфом", или что "это -- конец последним начатого списка" и т.п. С другой стороны, система с процедурной разметкой определяет, какая обработка должна выполняться в конкретной точке документа: "в этом месте вызвать процедуру PARA с параметрами 1, b и x", или "передвинуть левую границу на 2мм левее, правую границу -- на 2мм правее, пропустить одну строку и встать на новую левую границу" и т.п. В SGML инструкции, необходимые для обработки документа с какой-либо конкретной целью (например, для форматирования), четко отделяются от описательной разметки, которая встречается внутри документа. Обычно они собраны вне документа в отдельных процедурах или программах.

С описательной, а не процедурной, разметкой один и тот же документ может быть обработан разнообразными программами, каждая из которых может применять различные инструкции обработки к тем его частям, которые она считает важными. Например, программа анализа содержимого может полностью игнорировать сноски, тогда как программа форматирования может извлекать и собирать их для печать в конце каждой части. Различные виды инструкций обработки могут ассоциироваться с одной и той же частью файла. Например, одна программа может извлекать из документа фамилии людей и географические названия для создания индекса или базы данных, тогда как другая, обрабатывающая тот же самый текст, может печатать фамилии и названия отличающимся шрифтом.

SGML вводит понятие типа документа, и, соответственно, определения типа документа (document type definition, DTD). Документы считаются типизированными, так же, как и другие обрабатываемые компьютерами объекты. Тип документа формально определяется его составными частями и их структурой. Определение, например, отчета может быть таким, что он состоит из заголовка и, возможно, автора, за которыми следует аннотация и последовательность одного или более абзацев. Любой документ в отсутствие заголовка, в соответствии с этим формальным определением, не будет формально являться отчетом, так же как не будет им являться и последовательность абзацев, за которой следует аннотация, невзирая на то, насколько похож на отчет такой документ с точки зрения читателя-человека.

Поскольку документы относятся к известным типам, можно использовать специальную программу, называемую анализатором (parser), для того, чтобы обработать документ, утверждающий, что он относится к конкретному типу, и проверить, действительно ли все элементы, требуемые для данного типа документов, присутствуют и находятся в правильной последовательности. Что еще более важно, разные документы одного типа могут обрабатываться унифицированным образом. Можно писать более интеллектуальные программы, использующие знания, заключенные в информационной структуре документа.

Основная цель проектирования SGML была в создании гарантий того, что документ, закодированный согласно его положениям, будет переносимым с одной аппаратной и программной среды в другую без потери информации. Два его свойства, описанных выше, отвечают этому требованию на абстрактном уровне; третье свойство -- на уровне строчек байтов (символов), которые составляют документ. SGML предоставляет обобщенный механизм строковой подстановки, то есть, простой машинно-независимый способ указания, что конкретная строка символов в документе в момент обработки документа должна заменяться на некоторую другую строку. Одно очевидное применение этому механизму - обеспечение единой терминологии; другое, и более значительное, противодействие известной неспособности разных компьютерных систем понимать наборы символов друг друга, например, одной системе представлять все графические символы, необходимые приложению, путем описательного отображения непередаваемых символов. Строки, определяемые этим механизмом, называются сущностями (entities).

SGML даёт возможность производить документы в любом требуемом виде. Удобно оказалось и написание "вордовых" бумажек в привычной среде XEmacs без использования Windows. Все использованные программы- бесплатны и распространяются в исходных текстах, что позволило установить их на всех используемых мной системах. Причем оказалось, что начать работать с SGML вовсе не сложно, каждый, кто хоть раз редактировал WWW страничку, не испугается вида конструкций SGML.

Потом уже, по мере постепенного изучения SGML, я осознал, что эта техника дает громадное количество возможностей по сравнению с типичной "настольной типографией":

Четко разделенные процессы ввода информации и ее форматирования позволяют автору сосредоточиться на изложении мыслей, не отвлекаясь на двигание текста по экрану и подбор стилей.

Легко выдерживать различные документы в едином стиле, используя единую терминологию. Если стиль или термины нужно изменить, это делается разом во всех документах, не трогая их содержимое.

Повторное использование для SGML означает наличие возможности использовать детали старых проектов в новых, с минимальными изменениями. Часть документа, оформленная в виде SGML-элемента, может переноситься в другие документы, легко повторяться в разных местах текста.

Долговечность информации. Из-за того, что SGML -- простой и стандартный формат хранения данных, отсутствует необходимость переформатировать их ввиду устаревания аппаратной или программной платформы. Информация просто доступна навсегда. Она несет с собой все необходимое для создания документа.

Лучшее управление данными. С SGML можно определять информационные элементы и манипуляции с ними с произвольной степенью детальности. Размеченные элементы могут иметь атрибуты, определяющие характеристики и свойства элементов. Эта информация не предназначена для печати, но может помочь в управлении элементами данных. Например, атрибут ID (идентификатор) может уникальным образом идентифицировать один абзац, или целый раздел, примечание, иллюстрацию, задание, -- любой элемент, как в этом примере: <para id=431>Информация</para>

Так как идентификаторы являются машинно-читаемыми, они могут связывать между собой информацию и использоваться для разнообразного управления ей. Например: контролировать безопасность доступа к информации, позволяя только определенным людям просматривать или изменять ее. Автоматизировать перемещение информации - например, обновление данных в одном месте может инициировать обновление той же информации в других приложениях.

Следующая особенность SGML разделимость. Возможность работы со структурированными компонентами документа позволяет строить целый документ из составных частей, разбросанных по организации. Это позволяет пользователям делиться информацией без ее дублирования.

Мобильность В информационных сетях, объединяющих разнообразные компьютеры, операционные системы и приложения, мобильность становится ключом ко всеобщей доступности информации. Поскольку SGML не зависит от аппаратуры и приложений, можно легко обмениваться документами между различными системами.

Гибкость в приложении SGML позволяет использовать информацию далеко за рамками "настольной типографии". Например: страницы WWW; информационные базы данных; диагностические / экспертные системы; электронная почта; гипертекстовая документация; публикации на CD-ROM; интерактивные электронные мануалы.

**Заключение**

Стандарты ODA и SGML по своей сути не являются сравнимыми, каждый из них выполняет свои определённые функции. В частности как уже говорилось технология описания данных с помощью стандарта ODA позволяет создавать информационные структуры любого уровня сложности, сохраняя при этом простоту чтения и интерпретации. Основное назначение технологии ODA - создание комплексных информационных систем для предприятий различных видов деятельности и масштабов.

С помощью ODA могут быть автоматизированы практически все информационные процессы и операции предприятия: бухгалтерский учет, документооборот, кадровая деятельность, юридическое обеспечение, складской учет, управление производством, управление сбытом, управление маркетингом, финансовый анализ, бюджетирование, управленческое планирование и т.д. и т.п. теперь что касается SGML то здесь, на мой взгляд, стандарт предназначен для унификации представления документации в автоматизированных системах и первоначально ориентирован на текстовые данные, но охватывает и гипермедиа. В совокупности же они являются довольно слаженными инструментами обработки, обслуживания, стандартизации сфер применения документооборота. SGML и ODA нельзя рассматривать только как стандарты это и языки и технологии и инструменты, целые программные обеспечения создание для выполнения определённого класса задач, которые были описаны выше.

**Список литературы**

Сеймон. Ш. Об SGML и HTML. [Электронный ресурс]/ http://pyramidin.narod.ru/html401/sgmltut.html

Черняк Л.А. [Электронный ресурс]http://www.i2r.ru/static/479/

1. http://pyramidin.narod.ru/html401/sgmltut.html - Об SGML и HTML
2. http://www.artbis.ru/index.php?mact=CGBlog,cntnt01,detail,0&cntnt01articleid=33&cntnt01returnid=72 - всё об ODA
3. http://www.oda.su/index.php?page=tehnologiya-oda – технология ODA
4. http://ru.wikipedia.org/wiki/SGML -SGML. Материал из Википедии.
5. http://en.wikipedia.org/wiki/Standard\_Generalized\_Markup\_Language - Standard Generalized Markup Language
6. http://mcc.hydromet.ru/docs.rus/html4/sgml/intro.html - Справочная информация SGML для HTML
7. http://webcache.googleusercontent.com/search?q=cache:M2OtLUgk8ZoJ:www.intuit.ru/department/internet/html/3/+SGML&cd=10&hl=ru&ct=clnk&gl=ru&lr=lang\_ru&client=opera –Введение в SGML. Intuit.ru
8. http://www.freebsd.org/ru/docproj/sgml.html -
9. http://www.oasis-open.org/cover/sgml-xml.html - Страница SGML/XML. Содержит бесчисленное множество ссылок на информацию о SGML.
10. http://xml.coverpages.org/sgml.html - SGML and XML as (Meta-) Markup Languages