

УДК 528:004.057.2

ИСПОЛЬЗОВАНИЕ ТЕХНОЛОГИЙ ESRI ДЛЯ СОЗДАНИЯ ИНФРАСТРУКТУРЫ ПОДДЕРЖКИ МЕТАДААННЫХ

Николаев В.М, Салтовец А.А., Сорока В.А.

В статье рассматриваются общие вопросы реализации системы поддержки метаданных пространственных данных. Предложены пути реализации системы поддержки метаданных на базе программного обеспечения фирмы ESRI. Рассмотрены вопросы соответствия ПО ESRI текущему уровню стандартов ISO в области метаданных.

Ключевые слова: Метаданные, стандарты, профили, пространственные данные, XML

ВВЕДЕНИЕ

В работе «Методические вопросы реализации метаданных на основе профиля стандарта ISO 19115» [1] приведена обобщённая процессная модель методики создания и внедрения системы метаданных. Предметом рассмотрения данной статьи являются процессы разработки приложений для работы с метаданными, а также процессы организации хранилища и публикации метаданных.

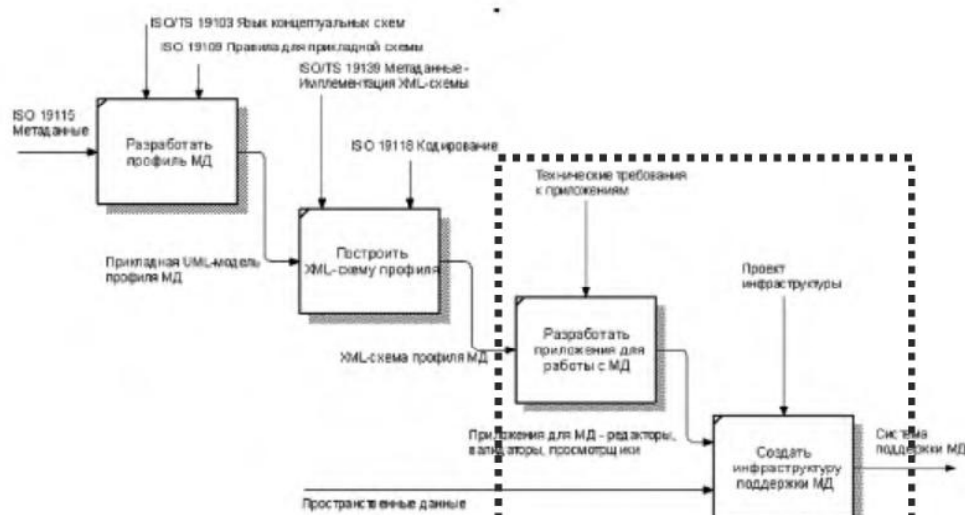


Рис. 1 Общая функциональная модель процессов создания системы поддержки метаданных

Рассмотрим более детально элементы «Разработать приложения для работы с МД» и «Создать инфраструктуру поддержки МД» приведенной на рис.1 диаграммы.

Результатом выполнения предыдущих процессов является XML-схема, соответствующая разрабатываемому профилю метаданных. XML-схема (XML-scheme, файл XSD) профиля метаданных используется для разработки программных средств поддержки метаданных – редакторов, валидаторов и т.п. При этом XML-схема профиля является источником для получения дерева описания метаданных, заполнения обязательных и дополнительных полей метаданных, проверки корректности получаемых МД в формате XML.

Разрабатываемые программные средства работают с метаданными в формате XML, создавая их, проверяя правильность их составления, обеспечивая просмотр или обращаясь с запросами к хранилищу метаданных. Для отображения содержания XML в удобочитаемом виде используется, обычно, технология XSLT (Extensible Style Language Transformation) [2].

Завершающим этапом создания системы поддержки метаданных (если не включать в рассмотрение этап промышленной эксплуатации) является разворачивание инфраструктуры этой поддержки, состоящей из хранилищ метаданных, средств их публикации, опоискования и связи с самими пространственными данными.

Процесс разворачивания инфраструктуры поддержки профиля метаданных включает, в общем виде, обеспечение доступности эталонной схемы, включение в технологический процесс подготовки ПО для генерации корректных описаний пространственных данных, создания хранилища метаданных, публикацию метаданных. Создание хранилищ и центров публикации метаданных подразумевает, в общем случае, создание системы взаимодействующих узлов системы. В данной статье ограничимся рассмотрением задачи создания единичного узла поддержки метаданных, так как задача создания системы взаимодействующих узлов основывается скорее на определении и реализации правил взаимодействия субъектов системы, чем на внесении существенных дополнительных требований к хранению и публикации метаданных.

РЕДАКТОР МЕТАДААННЫХ

Как было сказано выше, задача создания программного средства, способного создавать и редактировать XML-описание данных на основе профиля метаданных в формате XML-схемы, является первоочередной.

Рассмотрим пример реализации «независимого» приложения для работы с метаданными. Под «независимым» будем понимать приложение, не требующее специализированного программного окружения, кроме, быть может, расширенной среды исполнения операционной системы.

При проектировании редактора метаданных первостепенное внимание было уделено автоматизации обработки всех элементов входных данных и описаний на основе XML. Точнее, за основу был принят тезис, что вся необходимая для формирования экземпляра метаданных информация содержится непосредственно в XML-схеме профиля метаданных, а необходимая для адаптации пользовательского интерфейса информация может содержаться в дополнительных XML-описаниях.

Собственно XML-схема профиля метаданных является исчерпывающим источником информации о составе метаданных, а следовательно и о возможных формах представления и отображения генерируемых XML. Таким образом разработка программного обеспечения может быть выполнена любым в достаточной мере квалифицированным разработчиком с применением средств разработки, поддерживающим работу с XML. Более того, ориентация на использование XML-схемы в качестве основы реализации открывает путь к созданию универсального ПО, ориентированного на генерацию корректных, по отношению к используемой XML-схеме, описаний в формате XML.

Созданный редактор метаданных инвариантен по отношению к используемой в нём XML-схеме, шаблону экранного представления и шаблону локализации имён используемых ключевых слов.

В целом разработанное приложение позволяет создавать и редактировать XML-описания на основе заданной XML-схемы, сохранять XML-описания в файлы XML, отображать содержимое текущего XML-описания с применением подключаемой схемы трансляции XSLT, проводить проверку XML-описания на соответствие заданной XML-схеме.

Приложение реализовано в пакете MS VisualStudio 2005 на языке C#. Для работы с XML используется штатный парсер Microsoft msxml.dll (MSXML 4.0).

Общий вид интерфейса пользователя приложения приведен на рис.2.

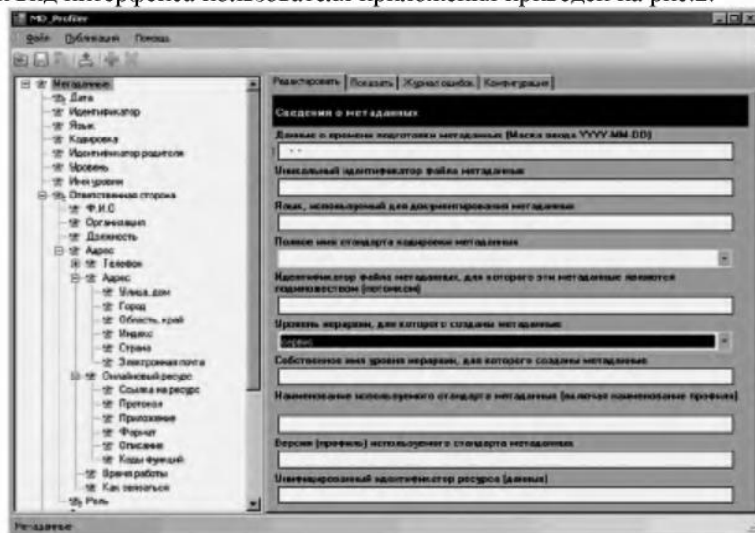


Рис.2 Общий вид редактора метаданных

Левую сторону рабочего поля приложения занимает дерево элементов используемой XML-схемы. Процесс построения дерева элементов экземпляра метаданных состоит из интерпретации заданной в файле конфигурации XML-схемы (далее - схемы) и трансляции имён отображаемых сущностей и атрибутов схемы в локализованные имена, в соответствии со специально подготовленным для

выбранной схемы файлом отображения имён. В случае отсутствия файла отображения имён или отсутствия в нём нужных пар трансляции, имена сущностей и атрибутов отображаются латиницей, в исходном для схемы виде. Обязательные для заполнения элементы схемы выделены в дереве специальным значком.

Правая часть рабочего поля приложения представляет собой набор «закладок» переключающих приложение в режимы редактирования, просмотра текущего экземпляра XML, просмотра журнала ошибок и просмотра значений из файла конфигурации. Ключевой особенностью приложения является отсутствие заготовленных форм на закладке редактирования. Все элементы редактирования генерируются «на лету», на основании информации из XML-схемы. При этом для редактирования всегда доступны все терминальные (имеющие простой тип) атрибуты выбранной в дереве схемы сущности метаданных.

Процесс проверки XML-описания на соответствие заданной XML-схеме запускается автоматически при переходе пользователя на закладку «Журнал ошибок». И в этом процессе, опять-таки, используется текущая XML-схема.

Для обеспечения успешной проверки корректности метаданных при размещении их в сервисе метаданных ArcGIS ArcIMS применяется специальный XML-файл, обеспечивающий добавление к генерируемому экземпляру метаданных тегов, проверяемых ArcIMS при публикации.

На рис.3 приведено изображение страницы конфигурации приложения, на котором видно, что все элементы входных данных, участвующие в обработке экземпляра XML-описания представляют собой файлы форматов XSD, XML и XSL.

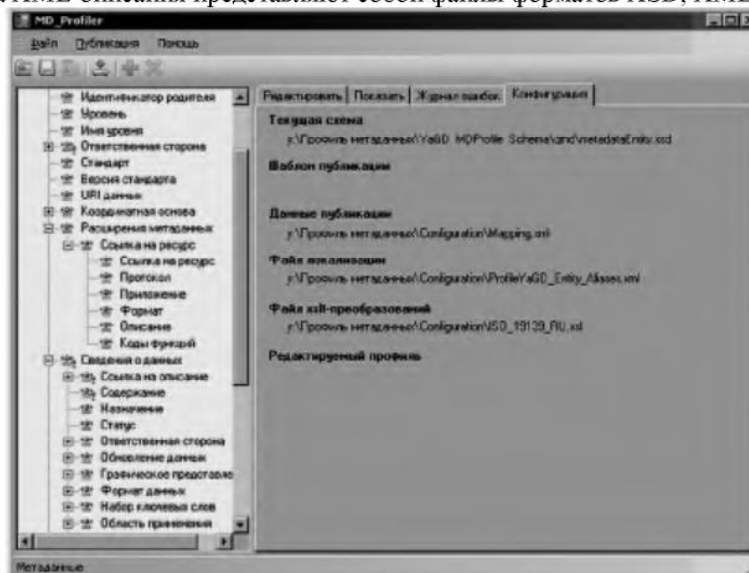


Рис.3 Параметры конфигурации редактора метаданных

В целом, разработанный редактор позволяет получать экземпляры метаданных, в виде XML-файлов, в соответствии с задаваемой XML-схемой профиля

метаданных. В качестве XML-схемы для тестирования была выбрана XML-схема из стандарта ООО ЯМБУРГГАЗДОБЫЧА «Профиль метаданных пространственных данных», разработка которой проводилась в соответствии с методикой, рекомендованной для создания профилей метаданных в серии стандартов ISO 19100.

СИСТЕМА ПОДДЕРЖКИ ИНФРАСТРУКТУРЫ МЕТАДАННЫХ

Итак, обеспечив систему инструментом подготовки метаданных, можно двигаться дальше.

Как было сказано выше, необходимыми элементами поддержки метаданных являются хранилища метаданных и системы публикации метаданных.

Учитывая количество накопленных пространственных данных (ПД), число их создателей и владельцев, а также постоянно возрастающую потребность в поиске нужных ПД, можно с уверенностью утверждать, что для создания хранилищ метаданных необходимо использовать промышленные СУБД.

Подготовка и реализация информационной структуры хранилищ могут осуществляться несколькими путями. С одной стороны, профиль метаданных, с учётом методики его создания, представляет собой, по сути, информационную структуру, которую достаточно просто отобразить в структуру базы данных. С другой стороны, корректные XML-описания, полученные на базе XML-схемы профиля, представляют собой чётко структурированные и организованные наборы данных, достаточно просто обрабатываемые современными программными средствами. Сказанное определяет два возможных подхода к организации хранилища метаданных. Первый подход реализуется имплементацией информационной структуры профиля метаданных в структуру той или иной базы данных, второй – использованием для хранения и поиска метаданных непосредственно XML-описаний.

Первый подход характеризуется возможностью организации исключительно гибкого поиска, но требует дополнительных затрат на создание модели базы. Кроме того, создаваемое при таком подходе хранилище трудно адаптировать для метаданных, созданных на основе разных профилей. Преимуществом второго подхода является универсальность работы в условиях применения различных стандартов. Разумеется возможен и третий, смешанный, подход, сочетающий, в разных пропорциях, оба из ранее указанных. Представляется, что именно он и является наиболее продуктивным.

В рамках статьи, под публикацией метаданных, будем понимать обеспечение доступности информации хранилищ метаданных для пользователей. Доступность понимается в широком смысле слова и подразумевает не только возможность физического подсоединения к хранилищу, но и предоставление инструментов поиска, просмотра и получения данных.

Исходя из потребности в поиске ПД у различных, «несвязанных» пользователей, ведущих разнородную деятельность, можно с уверенностью сказать,

что доступ к хранилищу должен быть удалённым и организован на основе WEB-технологий.

При создании системы пространственных данных неизбежно встаёт вопрос о выборе необходимых решений и технологий, обеспечивающих реализацию всех перечисленных задач. Так как информационная составляющая и основные элементы системы метаданных достаточно ясны и формализованы, можно начать проектирование соответствующего решения с нуля. Но можно попробовать использовать готовое решение, адаптировав его под текущие нужды. Именно второй путь и рассматривается далее в данной статье. Далее - об использовании технологий ESRI для поддержки инфраструктуры метаданных, создаваемых на основе профиля стандарта ISO 19115.

СИСТЕМА ПОДДЕРЖКИ ИНФРАСТРУКТУРЫ МЕТАДААННЫХ И ТЕХНОЛОГИИ ESRI

Система поддержки метаданных в линейке программных продуктов фирмы ESRI является одним из самых сложных и разветвлённых элементов, в работе которого принимают участие практически все уровни клиентских и серверных приложений. Это определяет высокую сложность управления и настройки системы, а также достаточно высокую стоимость её внедрения.

Для подготовки, хранения и публикации метаданных ESRI предусматривает использование следующего программного обеспечения и материалов:

- профиль метаданных ESRI [3], доступный по URL <http://www.esri.com/metadata>, а также сопутствующие профилю материалы, в виде описаний, файлов определения типов, таблиц стилей и трансляции;
- ArcCatalog и набор инструментов для обработки метаданных – создание, редактирование, просмотр и поиск метаданных;
- ArcSDE и промышленная СУБД (MS SQL Server, Oracle, DB2) с поддержкой полнотекстового индексирования – хранилище метаданных;
- ArcIMS, WEB-сервер (MS IIS, Apache), сервлетная машина (ServletExe, Tomcat) – создание сервиса метаданных и публикация метаданных;
- MetadataExplorer – удалённый клиент для доступа к сервису метаданных (начиная с версии 9.2 выведен из состава системы метаданных и поставляется в виде примера приложения).

Несколько слов о **профиле метаданных ESRI** [3]. Этот документ датирован мартом 2003 года. О стандарте «International Standard ISO 19115 Geographic information – Metadata» в документе сказано как о окончательной черновой версии. Документ в явном виде указывает на то, что он является профилем «Стандарта на содержание метаданных цифровых геопространственных данных» (Content Standard for Digital Geospatial Metadata, CSDGM) FGDC (FGDC-STD-001-1998). Следует отметить, что профиль ESRI радикально отличается от понятия «профиль», используемого в данной статье по форме представления. В профиле ESRI присутствует ряд тегов, по структуре и информационному смыслу соответствующих стандарту ISO 19115, правда, для их наименования выбраны

короткие имена. В версии 9.2 программных продуктов ESRI упоминается стандарт ISO 19139, который на самом деле является технической спецификацией кодирования XML-схемы, и даже подготовлен ряд таблиц трансляции тегов.

В целом, профиль метаданных ESRI является руководством для создания совместимых метаданных и программных продуктов их поддержки, но абсолютно не учитывает сегодняшнее состояние дел ряда стандартов ISO, связанных с метаданными.

Далее рассмотрим применимость перечисленного выше ПО с точки зрения поддержки метаданных, созданных на основе профиля стандарта ISO 19115.

ArcCatalog является основным инструментом управления пространственными данными. В нём присутствуют средства создания и редактирования метаданных на основе профиля ESRI с возможностью выбора режима заполнения метаданных на базе CSDGM или ISO. В ПО предусмотрено использование различных редакторов метаданных, расширение их функциональности и даже подключение собственных редакторов. Просмотр содержимого метаданных возможен с помощью подключаемых стилей (файлов XSLT). При подключении к сервису метаданных ArcIMS, ArcCatalog может работать в режиме клиента этого сервиса и служит основным средством для размещения подготовленных метаданных в сервисе метаданных. ArcCatalog способен выполнять поиск пространственных данных с использованием метаданных, включая поиск в подключенных сервисах метаданных. Правда, поиск с использованием сервиса ограничивается поиском по тегам, явно прописанным в теле программы, что фактически не позволяет выполнять поиск метаданным, созданным на основе профиля стандарта ISO 19115.

Для адаптации ArcCatalog к решению задач, рассматриваемых в статье, учитывая присутствие в системе программирования ESRI специальных объектов, предназначенных для создания редакторов метаданных, возможно создание специализированного расширения ArcCatalog. Оно может сочетать функциональность рассмотренного выше редактора метаданных в формате XML-схемы с информацией о пространственных данных, доступной при помощи программных объектов ArcObjects. В завершение краткого обзора ArcCatalog хотелось бы обратить внимание создателей ПО ESRI на явно недостаточное использование мощнейшего организующего элемента в виде XML-схем.

Сервер приложений **ArcSDE** предназначен для организации хранения пространственных данных под управлением промышленных СУБД. Роль ArcSDE в системе метаданных ограничивается размещением в геобазе довольно простых информационных структур поддержки сервисов метаданных и взаимодействием с выбранной СУБД для выполнения запросов от сервисов метаданных. Непосредственно к хранилищу метаданных предъявляется требование поддержки полнотекстового индексирования. В целом, система реализует подход к созданию хранилища, при котором метаданные хранятся и опоиcковываются как полные XML-описания. Для задания полнотекстовых индексов служит специальный конфигурационный файл, который содержит имена и краткие описания тегов поиска, что даёт возможность полностью адаптировать поиск метаданных к XML-схеме профиля метаданных.

Публикацию метаданных в системе ESRI обеспечивает сервер ArcIMS. Один сервер ArcIMS может обеспечить публикацию нескольких сервисов метаданных, использующих собственные информационные структуры в геобазе под управлением ArcSDE. Сервис метаданных IMS может публиковать любые XML-описания при отключении в файле конфигурации проверки корректности по некоторым предопределённым ESRI тегам.

Особенностью сервиса метаданных ArcIMS является взаимодействие с клиентами исключительно при помощи фирменного, тщательно документированного, подмножества XML – ArcXML, что даёт возможность реализации собственных модификаций клиентов работы с метаданными и отлично согласуется с общим подходом к созданию системы метаданных. Пример запроса на поиск метаданных приведен на рис.4. Кроме того, сервер ArcIMS имеет внутреннюю систему разграничения прав доступа, которая позволяет регулировать уровень доступности метаданных в зависимости от отношений ролевой функции подключаемых клиентов.

```
<?XML VERSION="1.0" ENCODING="UTF-8"?>
<ARXML VERSION="1.1">
<REQUEST>
<GET_METADATA>
<SEARCH_METADATA
  FOLDERMASK="4" GNDXTENT="SEARCH" MAXRESULTS="10" SORT="LOCAL_AREA"
  SORT2="CONTENTTYPE" STARTRESULT="0">
<ENVELOPE MAXX="60.0" MAXY="42.0" MINX="-30.0" MINY="-17.0" SPATIALOPERATOR="OVERLAPS"/>
<SEARCH_METADATA GNDXTENT="NONE" OPERATOR="OR">
  <TAGTEXT TAG="METADATA/IDINFO/KEYWORDS/THEME/THEMKEY" WORD="IMAGERYBASEMAPSEARTHCOVER"/>
  <TAGVALUE ATTRIBUTE="VALUE" EQUALTO="010" TAG="METADATA/DATAIDINFO/TPCAT/TOPICCATCD"/>
</SEARCH_METADATA>
<FULLTEXT WORD="AFRICA"/>
<DOCUMENTINFO CONTENT="OFFLINE DATA"/>
</SEARCH_METADATA>
</GET_METADATA>
</REQUEST>
</ARXML>
```

Рис.4 Пример запроса на поиск метаданных в формате ArcXML

В целом, несмотря на некоторые трудности в администрировании, ArcIMS является готовым решением для публикации метаданных.

В заключение обзора программных средств ESRI несколько слов о «фирменных» удалённых клиентах. До версии 9.1 включительно, в состав ArcIMS входил «тонкий» клиент Metadata Explorer, позволявший очень быстро обеспечить доступ к публикуемым сервисам метаданных удалённым клиентам при помощи Internet-браузера. В версии 9.2 аналогичный по функциональности клиент присутствует в виде примера WEB-приложения. Недостатками обеих реализаций является использование тегов поиска (см. рис.4), весьма ограниченный набор которых встроен в тело скриптов WEB-приложений. Для адаптации «тонких» клиентов системы метаданных ESRI к работе с метаданными, созданными на основе профиля стандарта ISO 19115 нужно провести работу по локализации WEB-приложений и настройки их на поддержку имён из пространства используемой XML-схемы.

ВЫВОДЫ

Рассмотрение системы поддержки метаданных ESRI позволяет сделать вывод, что эта система является готовым промышленным решением и отвечает, в целом, потребностям создания инфраструктуры поддержки метаданных, создаваемых на основе профиля стандарта ISO 19115, однако для полноценной поддержки профилей стандарта ISO 19115 требуется адаптация части её ключевых элементов.

К сказанному следует добавить, что в работах [4, 5] явно обозначена заинтересованность ESRI в поддержке международных инициатив в области интеграции систем поддержки метаданных пространственных данных, что с учётом перспективности стандартов ISO должно привести к качественному пересмотру подхода ESRI к использованию профилей метаданных.

Список литературы

1. Салтовец А.А. Николаев В.М. Методические вопросы реализации метаданных на основе профиля стандарта ISO 19115: - Ученые записки Таврического национального университета им. В.И.Вернадского. Серия «География». – 2008.
2. XSL Transformations (XSLT) Version 2.0: W3C Recommendation - <http://www.w3.org/TR/2007/REC-xslt20-20070123/>
3. ESRI Profile of the Content Standard for Digital Geospatial Metadata - An ESRI® Technical Paper, March 2003
4. ESRI® Geospatial Portal Technology - An ESRI White Paper, September 2007
5. ESRI® Technology and INSPIRE - An ESRI White Paper, June 2007

Салтовец О. О., Николаев В. М., Сорока В. О. Використання технологій ESRI для створення інфраструктури підтримки мета даних // Вчені записки Таврійського національного університету ім. В. І. Вернадського. – 2008. – Серія «Географія». – Т. 21 (60). – № 1. – С. 97-105

У статті розглядаються загальні питання реалізації системи підтримки метаданих просторових даних. Запропоновані засоби реалізації системи підтримки метаданих на базі програмного забезпечення фірми ESRI. Представлені питання відповідності програмного забезпечення ESRI поточному рівню стандартів ISO в галузі метаданих.

Ключові слова: Метадані, стандарти, профілі, просторові дані, XML

Saltovets A., Nikolaev V., Soroka V. Application of ESRI technologies for the metadata supporting infrastructure creation // Uchenye zapiski Tavricheskogo Natsionalnogo Universiteta im. V.I. Vernadskogo. – 2008. – Series «Geography». – V. 21 (60). – № 1. – P. 97-105

The article shows general aspects of realization of the systems supporting the metadata of spatial data. The approaches of the metadata supporting system realization based on the ESRI software are suggested.

Key words: Metadata, Standards, Profiles, Spatial data, XML

Поступила в редакцию 24.04.2008 г.