

Информационная система регионального центра распространения геоданных

М.А. Свидрицкая, А.В. Козлов, К.В. Игушев

*Югорский НИИ информационных технологий
628011 г. Ханты-Мансийск, ул. Мира, 151
E-mail: sma@uriit.ru*

На сегодняшний день при ежедневно возрастающих объемах поступающей, а, следовательно, и хранящейся в архивах геоинформации различных типов ее пользователи испытывают трудности при подборе необходимых данных. По оценкам специалистов компании Core Software Technology, занимающихся разработкой коммерческого продукта imageNet, около 75% времени работы над геоинформационными проектами тратится исключительно на подбор необходимых данных из различных источников.

Даже в условиях одной крупной организации поиск геоинформации усложняется тем, что геоданные хранятся на разных носителях: жестких дисках, магнитных лентах, CD и DVD-дисках. Кроме того, организации часто не имеют обобщенного каталога архивных данных с возможностью удобного поиска и предварительного просмотра информации о данных без необходимости обращения к архивному носителю информации.

Если рассматривать эту проблему на региональном уровне, то она усугубляется территориальной удаленностью держателей информации, частой неосведомленностью заинтересованных организаций о уже имеющихся в наличии геоданных в организациях этого региона и трудностями поддержания данных в актуальном состоянии на текущий момент времени.

В целях рационального использования геоинформационных данных организациями Ханты-Мансийского автономного округа (ХМАО), Югорским научно-исследовательским институтом информационных технологий для «Югорского центра геопространственных данных» [<http://www.ucgd.ru>] с начала 2005 года ведется разработка системы сбора, обработки, хранения и распространения геоданных (ИС). Некоммерческая организация «Югорский центр геопространственных данных» (ЮЦГД) учрежден заинтересованными организациями (ХМАО). Основными предпосылками к созданию ЮЦГД можно выделить следующие:

1. Многократное дублирование геоданных, как на уровне региона, так и на уровне отдельно взятых организаций. Это в свою очередь приводит, с одной стороны, к сложности, а иногда и невозможности отслеживания различных версий одного и того же экземпляра данных. Как правило, в результате возникает множество разрозненных версий одного и того же ресурса данных, которые невозможно синхронизировать между собой. С другой стороны, многократное дублирование данных может привести к их утрате в результате того, что каждый из держателей может уничтожить свой экземпляр, не заинтересовавшись судьбой этого ресурса у других пользователей.
2. Высокая стоимость пространственных данных часто не позволяет отдельным организациям приобретать их для собственных нужд. Решением может быть совместное приобретение необходимой информации заинтересованными в ней организациями.
3. Возрастающий спрос на геоданные требует максимально быстрого реагирования обладателей востребованных ресурсов на запросы потребителей.
4. Вследствие повышения спроса на пространственную информацию встает вопрос организации быстрого и качественного поиска данных. Зачастую, заинтересованная в геоданных сторона не имеет возможности оперативно найти и определить степень пригодности данных для своих проектов, что приводит к увеличению времени работы над промышленными и научными проектами.

Югорский центр геопространственных данных позволит решить следующие задачи:

- создание библиотеки геоданных различных форматов;
- создание и поддержание в актуальном состоянии каталога геопространственных данных, хранящихся у разных держателей геоданных;
- организация взаимодействия держателей пространственных данных в Ханты-Мансийском автономном округе;
- совместная разработка нормативных документов, регулирующих работу участников проекта.

В рамках решения первых двух из перечисленных задач было принято решение о создании автоматизированной информационной системы по каталогизации и поиску геоданных.

Для того, чтобы проиллюстрировать примерные объемы геопространственных данных, хранящихся в организациях ХМАО и которые необходимо каталогизировать, приведем размеры архивов геоданных в Югорского НИИ информационных технологий (ЮНИИ ИТ). На конец 2005 года в ЮНИИ ИТ были накоплены данные в следующем объеме:

Сканерные изображения

- Метеор-3М/МСУ-Э (2002-2005гг) - 700Гб
- Ресурс-01-3/МСУ-Э, МСУ-СК (1997-2000гг) - 252,4Гб
- NOAA/AVHRR (2003-2005гг) - 380Гб
- TERRA/MODIS (2003-2005гг) - 1600Гб
- TERRA/ASTER - 15Гб
- Landsat/TM(MrSid) -12Гб
- Landsat/ETM (MrSid) - 12Гб
- Landsat/TM+ETM (GeoTiff) (1982-2003гг) - 450Гб

Космофотоснимки

- Ресурс Ф1М - 4Гб
- Ресурс Ф2М/МК4 (1993г) -108Гб

Радарные изображения

- Shuttle/SIR-C -16Гб
- Shuttle/SRTM -6,8Гб
- JERS-1/SAR -6,4Гб
- ENVISAT/ASAR - 30Гб
- ERS-2/SAR – 50 лент по 30Гб

Итого объем информации хранящейся в ЮНИИ ИТ превышает 5 Тб.

Следует отметить, что в этот список не включены различные векторные данные, данные сканированных изображений, геологической сейсморазведки и другие.

Очевидно, что, обладая таким огромным объемом данных, остро встает вопрос об их каталогизации и организации поиска по каталогу. Для решения этой задачи сотрудниками ЮНИИ ИТ в рамках работ над ИС для ЮЦГД разработан электронный каталог.

К информационной системе были введены следующие функциональные требования, ИС должна:

- хранить метаданные о геопространственных данных (БМД – база метаданных), то есть вести базу данных метаописаний, следить за ее целостностью и непротиворечивостью;
- позволить конечному пользователю системы производить поиск метаданных по различным критериям;
- производить операции по созданию, актуализации и удалению данных из БМД, а также для случаев, где это возможно, автоматизировать этот процесс, вплоть до исключения из него оператора;
- обеспечивать защиту информации, хранимой в базе метаданных, как с точки зрения метаданных, так и самих геоданных;
- организовать систему заказов данных, найденных при поиске.

С технической точки зрения система должна быть:

- масштабируемой (качество функционирования ИС не должно ухудшаться с увеличением

количества обрабатываемых данных);

- кроссплатформенной (система должна функционировать на любой аппаратно-программной платформе);
- функциональная наращиваемость (система должна позволять вводить в себя дополнительные функциональные компоненты без модификации общей архитектуры системы).
- возможность интеграции с подобными системами (возможность интеграции системы каталогизации и поиска с системами других производителей).

На основании выделенных технических и функциональных требований к системе было принято решение вести разработку системы с использованием технологии Java 2 Enterprise Edition (J2EE) [<http://java.sun.com/javaee/index.jsp>, <http://java.sun.com/javase/index.jsp>] и использовать для хранения базы метаданных СУБД Oracle. Это, с одной стороны, позволило сделать систему максимально масштабируемой и наращиваемой, с другой стороны, обеспечило кроссплатформенность и широкие возможности интеграции с другими системами, используя J2EE Connector Architecture.

В результате проведенного анализа требований и выбранных технологий их реализации было выделено четыре основных компоненты системы:

- База метаданных
- Сервер приложений
- Автоматизированное рабочее место (АРМ)
- Web-интерфейс

Для построения базы данных системы была выбрана СУБД Oracle 10g.

СУБД Oracle, с одной стороны, обладает высокими показателями производительности и надёжности, с другой – имеет встроенные средства работы с геоданными – Oracle Spatial. Использование этого модуля позволило переложить часть функций информационной системы, связанных с пространственными запросами, на СУБД.

В качестве основы модели данных БД был выбран стандарт ISO 19115, так как он наиболее полно описывает структуру метаданных, затрагивая различные аспекты связанной с геоданными информации. Этот стандарт описывает следующие параметры данных:

- пространственно временные характеристики данных;
- вопросы хранения и распространения данных;
- информация о качестве и сертификации данных;
- информация об идентификации данных;
- информация об изменении данных;
- информация о секретности данных и метаданных;
- персональная информация о лицах, тем или иным способом связанных с данными.

В качестве метода доступа к данным был выбран объектный подход, то есть реляционная модель данных СУБД была приведена к объектному виду с использованием Enterprise Java Beans. Это позволило, с одной стороны, значительно ускорить процесс разработки системы, с другой стороны, обеспечить максимальную гибкость и наращиваемость системы. Данная объектная модель данных была размещена на сервере приложений под управлением Sun Java System Application Server, который соответствует спецификации J2EE 1.4.

Кроме объектной модели данных на сервере приложений была размещена вся бизнес-логика системы, что в значительной степени снижает нагрузку на клиентское программное обеспечение и позволяет производить проверку данных на соответствие стандарту на стороне сервера и вместе с тем обеспечивает защиту информации системы.

Другая компонента разрабатываемой информационной системы – автоматизированное рабочее место. На данную компоненту возлагается решение следующих задач:

- разграничить доступ пользователей к данным информационной системы с целью минимизировать возможные потери информации в результате работы с ними не квалифицированных лиц, а также

обеспечить соответствующий данным уровень секретности, вместе с этим обеспечить возможность журналирования всех операций, производимых с системой;

- администрировать БМД, производить комплексный анализ целостности и актуальности имеющихся данных;
- администрировать базу пользователей ИС с целью выявления потенциально узких мест, делегирование пользователям тех или иных прав на доступ к данным метаописаний;
- производить поиск данных по заданным критериям;
- импортировать данные из других информационных систем и производить мониторинг подсистем автоматического импорта данных, где это возможно.

Система WEB-интерфейса должна обеспечивать следующую функциональную нагрузку:

- осуществлять интерактивный и атрибутивный поиск геопространственных данных;
- осуществлять доступ к данным ИС в зависимости от прав пользователей;
- подавать заявку на предоставление данных.

Подсистема поиска ИС логически разделяется на две части: во-первых, это непосредственно подсистема поиска геоданных (атрибутивного и интерактивного поиска) и подсистема отображения найденных данных. Подсистема атрибутивного поиска позволяет искать данные, основываясь на атрибутивных фильтрах. Атрибутивный фильтр включает в себя поля:

- тип данных;
- временной интервал данных;
- пространственные координаты объекта;
- дополнительные атрибутивные данные;
- ключевые слова, с которыми данные связаны.

Атрибутивный поиск требует максимальной осведомленности пользователя системы о данных, которые он ищет. В случае, если пользователь не имеет точных данных о том, что он ищет, система предлагает пользователю воспользоваться интерактивным поиском данных с использованием картографического интерфейса. Фактически при интерактивном поиске пользователь обозначает область интересов примитивными атрибутивными фильтрами и географической областью на карте, на основе которых система осуществляет поиск. Очевидно, что в результате интерактивного поиска вместе с полезной информацией пользователю будет найдена значительная часть ненужной информации, в результате потребуется визуально отсортировать найденные данные, однако ввиду простоты составления запроса на поиск системе, интерактивный поиск является более удобным.

Другая функция WEB-интерфейса – это предоставление пользователю только тех данных, которые ему разрешены ограничениями, находящимися в метаданных. Для этого, перед тем как начать работу с системой, пользователю предоставляется возможность идентифицировать себя, пройдя авторизацию, после чего все действия пользователя в системе протоколируются для обеспечения мониторинга за системой и последующего анализа этих данных для оптимизации поиска.

Одной из важных подсистем системы поиска является возможность осуществления заказа выбранных данных. Для этого пользователю предоставляется возможность выбрать необходимые ему объекты и отправить заявку администратору системы, которая в дальнейшем обрабатывается и направляется реальному владельцу данных.

На сегодняшний день компоненты описанной информационной системы интенсивно разрабатываются и внедряются Югорском НИИ информационных технологий. Работу внедренных компонент информационной системы можно посмотреть на ресурсе: <http://www.ucgd.ru>.