

УДК 528.85

М. А. Попов, С. Ю. Марков, С. А. Станкевич, А. Л. Бодня

Науковий центр аерокосмічних досліджень Землі Інституту геологічних наук НАН України, Київ

Система каталогизации и распределенного доступа к данным дистанционного зондирования Земли: концепция, архитектура, реализация

Надійшла до редакції 09.02.06

Викладено концептуальний підхід до створення системи каталогування та розподіленого доступу до даних дистанційного зондування, який реалізовано в Науковому центрі аерокосмічних досліджень Землі Інституту геологічних наук НАН України у вигляді пілотного проекту. Підхід враховує світовий досвід будівництва інфраструктур геопросторових даних. Розроблено профіль стандарту на метадані для аеро- і космічних зображень, який засновано на міжнародних стандартах та реальних вимогах споживачів геопросторової інформації. Описано архітектуру системи та розглянуто особливості її програмно-апаратної реалізації.

ВСТУПЛЕНИЕ

Высокое качество видовой (иконической) информации, получаемой сенсорами, которые устанавливаются на борту воздушных и космических летательных аппаратов, способствует все более интенсивному ее использованию при решении самых разнообразных задач из областей науки, промышленности, сельского хозяйства и др. Поэтому круг потребителей аэро- и космической информации и в мире, и в Украине неуклонно расширяется.

В любой организации, которая постоянно занимается решением разнообразных тематических задач, связанных с анализом данных дистанционного зондирования Земли (ДЗЗ), мониторинга территорий, картографирования местности с использованием аэро- и космических снимков и т. д., по мере решения подобных задач постепенно накапливается собственный фонд видовой (иконической) информации, эффективное использование которого невозможно без его систематизации.

С систематизацией фонда изображений связано обеспечение таких важных для практической работы условий, как:

- простой доступ для легитимных пользователей с одновременным запретом для всех остальных пользователей;
- оперативность доступа к метаданным и изображениям;
- возможность пополнения и обновления фонда как через внутренние, так и через внешние (включая международные) информационные каналы;
- возможность взаимного информационного обмена с другими организациями через современные коммуникационные каналы;
- информационная устойчивость (минимальная вероятность потери или искажения данных).

Удовлетворение перечисленных условий возможно в рамках функционирования системы каталогизации и распределенного доступа к данным ДЗЗ. Такая система создана в Научном центре аэрокосмических исследований Земли Института геологических наук НАН Украины

(ЦАКИЗ) как пилотный проект. Ниже излагается концепция ее построения, а также описываются архитектура системы и особенности программно-технической реализации.

ИНФРАСТРУКТУРА ГЕОПРОСТРАНСТВЕННЫХ ДАННЫХ КАК КЛЮЧЕВОЙ ЭЛЕМЕНТ ЭФФЕКТИВНОЙ СИСТЕМЫ ИСПОЛЬЗОВАНИЯ ДАННЫХ ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ ЗЕМЛИ

Составными частями решения проблемы создания в организации системы каталогизации и распределенного доступа к данным ДЗЗ, отвечающей современным требованиям, являются:

- формирование базы метаданных об изображениях, которыми располагает организация;
- организация хранения данных ДЗЗ;
- обеспечение взаимодействия данной организации с другими организациями в вопросах обмена сведениями об информационных ресурсах и их получения.

Как известно, метаданные — это документированная информация, которая улучшает техническое и смысловое понимание данных и связанных с ними процессов [<http://www.tdan.com/>].

В условиях все увеличивающегося количества способов хранения данных, многообразия интерфейсов пользователей, гетерогенности физических устройств обработки данных необходимость выработки стандартных способов доступа до данных становится все более важной. Решение проблемы гомогенного доступа до гетерогенных источников данных лежит в стандартизации метаданных, однако в настоящее время в мире не выработан единый универсальный подход к формированию метаданных. Вместе с тем рядом международных организаций инициированы работы, направленные на создание документов, устанавливающих стандарты на пространственные метаданные.

Наиболее известные из подобного рода работ следующие.

1. FGDC Content Standard for Digital Geospatial Metadata — стандарт на пространственные метаданные, разработанный Геологической службой США (U.S. Geological Survey) в рамках программы построения Национальной инфраструктуры геопространственных данных (НИГД)

[<http://www.fgdc.gov/metadata/constan.html>]. Данный стандарт является одним из первых документов подобного рода, поэтому он уже довольно длительное время используется в различных странах мира при создании подобных документов.

2. CEN/TC 287 Env 12657 — европейский эквивалент стандарта на метаданные FGDC, разработанный Европейским комитетом по стандартизации (CEN) [<http://www.cenorm.be>]. Данный стандарт широко используется в европейских странах при создании профилей национальных стандартов на пространственные метаданные.

3. Стандарт ISO 19115 Geographic information — Metadata Международной организации по стандартизации [<http://www.isotc211.org/>]. Разработан ее техническим комитетом ISO/TC 211. Поскольку этот стандарт появился позже стандартов FGDC и CEN, его создатели учитывали требования упомянутых стандартов, и существенных различий между этими документами нет. Стандарты FGDC и ISO ближе друг к другу и более конкретны по содержанию по сравнению со стандартом CEN/TC [<http://www.ec-gis.org/Workshops/7ec-gis/papers/html/ruzicka/ruzicka.htm>].

В целом тенденции развития стандартов на геопространственные метаданные ведут к постепенному приближению содержания всех национальных и отраслевых стандартов к ISO 19115. Этот стандарт официально принят также консорциумом OpenGIS (www.opengeospatial.org).

Учитывая, что различным предметным областям характерна своя специфика работы с данными и метаданными, возникает необходимость создания так называемых «профилей» стандартов, т. е. некоторых расширений базовых стандартов, которые позволяют учесть особенности данной предметной области. В частности, в самом стандарте ISO 19115 описывается методика создания подобного «профиля». В исходном стандарте выделяются некоторые ключевые метаданные (рис. 1), которые полностью должны войти в профиль. Кроме того, из остальных метаданных выделяется та часть, которая должна войти в профиль, а также формируется дополнительный набор метаданных, отсутствующих в базовом стандарте, но необходимый для данной предметной области.

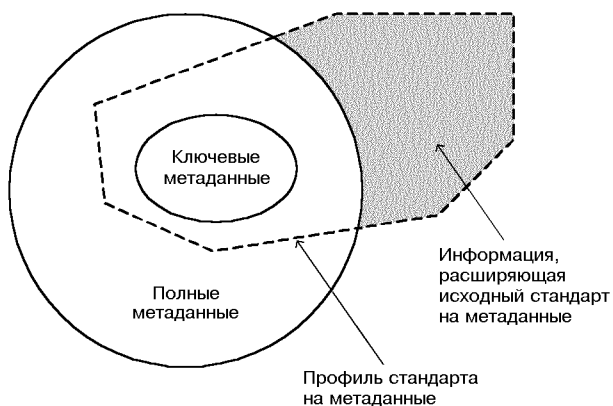


Рис. 1. Механизм формирования профиля стандарта

Таким образом, профиль должен включать:

- ключевые метаданные;
- обязательные элементы метаданных в обязательных информационных группах;
- условные элементы метаданных во всех обязательных информационных группах, если выполнено соответствующее условие;
- обязательные элементы метаданных во всех условных информационных группах, если выполнено соответствующее условие;
- условные элементы метаданных во всех условных информационных группах, если выполнено соответствующее условие.

Перечисленные требования должны безусловно учитываться при разработке профиля метаданных для любой организации, заинтересованной в эффективном использовании данных ДЗЗ. При этом необходимо учитывать, что вышеназванные стандарты на пространственные метаданные не совсем подходят для описания метаданных для данных ДЗЗ, поскольку в них отсутствуют некоторые информационные группы, важные для данной предметной области. В первую очередь это:

- идентификационные данные о миссиях, платформах и сенсорах;
- информация об орбите носителя;
- данные об уровне обработки изображения;
- полная информация о спектральном диапазоне (поддиапазонах) съемки;
- информация об условиях, в которых получено изображение;

- координатные данные, необходимые для привязки изображения;
- свойства сенсора, необходимые для привязки изображения.

По этой причине различные организации несколько лет назад начали работу по созданию специализированных стандартов на метаданные для данных ДЗЗ, которые являются расширениями базовых стандартов на пространственные метаданные. Наиболее важными являются:

- FGDC Content Standard for Digital Geospatial Metadata: Extensions for Remote Sensing Metadata — расширение стандарта на пространственные метаданные FGDC для данных ДЗЗ [http://www.fgdc.gov/standards/status/csdgm_rs_ex.html]. В него добавлены названные выше информационные группы для изображений, а также внесены некоторые другие изменения. В настоящее время идет доработка этого стандарта в направлении его гармонизации с требованиями соответствующего стандарта ISO;
- ISO 19115-2 Geographic Information — Metadata — Part 2: Extensions for imagery and gridded data — расширение стандарта ISO 19115 для метаданных пространственной информации, представленной в форме растровых изображений [http://www.isotc211.org/Outreach/Overview/Factsheet_19115-2.pdf]. Работа над этим документом пока не закончена. Планируется, что он будет содержать следующие дополнительные информационные группы, по сравнению с ISO 19115:
 - данные об уровне обработки данных;
 - информация об условиях, в которых получено изображение;
 - размеры и пространственное положение пикселей;
 - информация о спектральном диапазоне съемки;
 - идентификационные данные о миссиях, платформах и сенсорах.

Указанные данные вводятся как подклассы уже существующих классов стандарта ISO 19115, чтобы не нарушать его общую структуру классов.

При разработке стандартов на метаданные для данных ДЗЗ использовался также стандарт ISO 19130: Geographic Information — Sensor and data

model for imagery and girded data (Модели сенсоров и данных и для изображений и другой растровой информации) [http://www.isprs.org/istanbul2004/comm2/comm2.html]. Материалы данного стандарта применялись для определения информационных групп, специфичных для данных ДЗЗ.

Вопросы организации хранения геопространственных данных и обмена ими были в начале 1990-х гг. предметом всестороннего и обстоятельного изучения в США при создании НИГД. В директиве (№ 12906 от 13 апреля 1994 г.), подписанной президентом Б. Клинтонем, Национальная инфраструктура геопространственных данных определяется следующим образом: «НИГД представляет собой совокупность технологических решений, политики, стандартов и человеческих ресурсов, необходимых для сбора, обработки, распространения и эффективного использования геопространственных данных».

Накопленный опыт создания и использования НИГД в США и других странах подсказывает, что успешное решение вопросов организации хранения геопространственных данных и обмена ими связано с:

- разработкой системы стандартов на пространственные данные и на технологии их обработки, а также на метаданные;
- созданием базовых наборов пространственных данных, включающих цифровую картографическую информацию;
- отработкой технологической инфраструктуры работы с пространственными данными, которая обеспечит пользователям возможность нахождения необходимых им базовых наборов пространственных данных и получения доступа к ним.

Технологическая инфраструктура работы с пространственной информацией реализуется путем создания так называемых клиринговых центров управления геопространственными данными, которые представляют собой «совокупность программных и организационных средств, предназначенных для упрощения поиска, оценки пригодности для решения некоторой задачи, а также получения цифровых геопространственных данных» [1]. Данному определению соответствует концептуальная архитектура, показанная на рис. 2. Основными компонентами клирингового центра являются локальный сер-

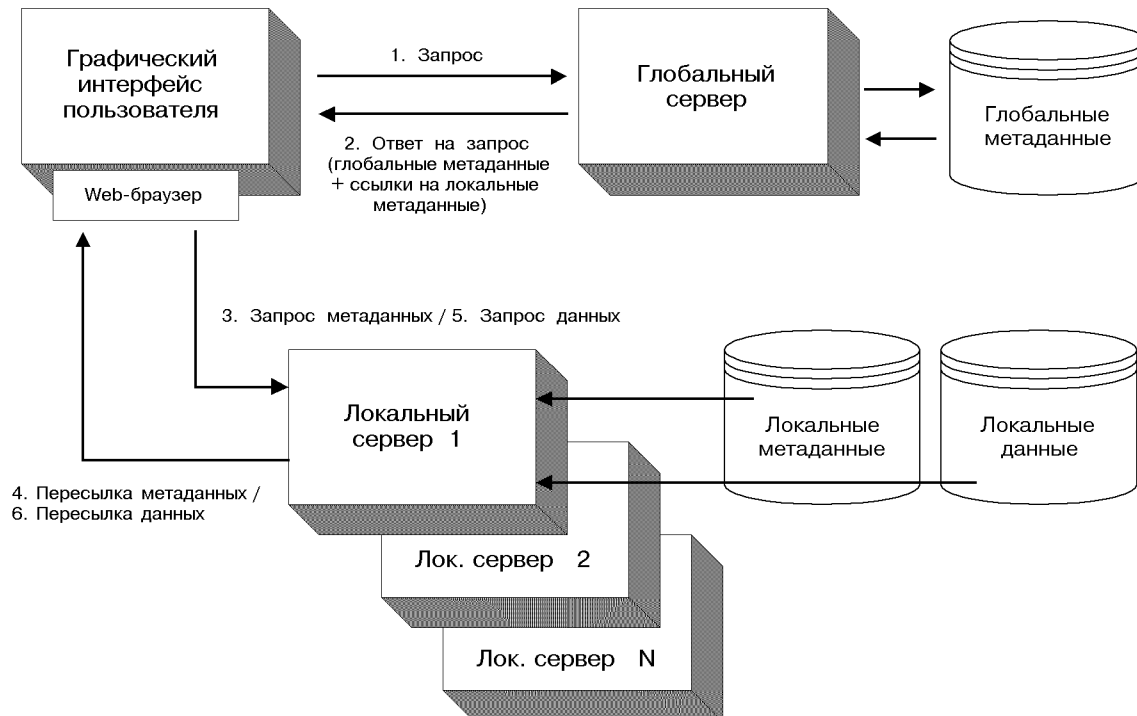


Рис. 2. Концептуальная архитектура клирингового центра

вер, в котором хранится информация о геопространственных ресурсах организации (локальные геопространственные данные и метаданные), а также глобальный сервер, который обеспечивает доступ к глобальным метаданным. Локальные метаданные — это данные о геопространственных данных, находящихся в организации, а глобальные метаданные — это их генерализованный (обобщенный) вариант, доступный для ознакомления внешним пользователям.

Пользователь, работающий с клиринговым центром, через графический интерфейс может обратиться с запросом к глобальному серверу (1), который даст ответ на этот запрос в виде глобальных метаданных либо ссылки на локальные метаданные (2). Пользователь также может обратиться к какому-либо локальному серверу с запросом на локальные метаданные либо данные (3 либо 5). В ответ на этот запрос ему пересылаются запрошенные метаданные либо данные (4 либо 6). Необходимость в глобальных метаданных и соответствующих архитектурных элементах возникает тогда, когда организация осуществляет информационное взаимодействие с другими организациями. В противном случае можно ограничиться локальной частью данной архитектуры, причем ее конкретный вид зависит от задач организации и, соответственно, от требуемой для этого функциональности. Как видно из концептуальной архитектуры клирингового центра, основным предназначением локального и глобального серверов является реализация каталожных функций.

Эталонная архитектура, позволяющая реализовать каталожный сервис, представлена на рис. 3.

Эта архитектура описана в соответствующей спецификации консорциума OpenGIS [<http://www.opengeospatial.org/specs/>] и, таким образом, может претендовать на статус стандарта «де-факто», например в области геоинформационных систем (ГИС).

Многие государственные и коммерческие организации как за рубежом, так и в Украине, занимающиеся обработкой и распространением данных ДЗЗ, уже реализовали в том или другом виде клиринговые центры управлением данными ДЗЗ.

Так, по данным [<http://www.isprs.org/istanbul2004/comm2/papers/202.pdf>] в Корей-



Рис. 3. Эталонная архитектура каталожного сервиса

ском институте исследования проблем электроники и телекоммуникаций создан и функционирует на национальном уровне клиринговый центр управления космическими снимками, архитектура которого включает три основных сервера:

- 1) регистрация данных ДЗЗ;
- 2) управление данными;
- 3) передача данных ДЗЗ.

Используемый профиль стандарта на метаданные разработан на основе стандартов ISO 19115 и ISO 19115-2 (часть 2).

Французской компанией SPOT Image клиринговый центр реализован как часть web-портала компании [<http://www.spot.com/>]. Каталог поставляемых изображений доступен в режиме online. Структура метаданных, которые использует SPOT Image для описания изображений, не полностью совпадает с требованиями стандарта ISO 19115-2. Есть некоторые отличия в типах данных некоторых полей, в синтаксисе и семантике информации.

Одним из немногих украинских серверов управления данными ДЗЗ является web-сервер Государственного научно-производственного цент-

ра аэрокосмической информации, дистанционного зондирования Земли и мониторинга окружающей среды «Природа» [http://www.pryroda.gov.ua/]. В разделе «Услуги» имеется возможность просмотреть каталог снимков и выполнить заказ выбранных изображений.

Поиск изображения может быть выполнен по критериям:

- тип спутника;
- диапазон дат;
- координаты углов экстенда изображения;
- выделение области на карте.

После ввода значений критериев выполняется поиск, в результате которого пользователь получает таблицу со списком найденных изображений, для которых выводится следующая информация:

- номер снимка;
- тип спутника;
- тип датчика;
- тип изображения;
- погодные условия (облачность);
- регион;
- дата.

Интерфейс пользователя производит хорошее впечатление: с одной стороны он достаточно

прост, с другой — позволяет выбрать необходимое изображение. Накопленные технологические, аппаратные, протокольные, интерфейсные и другие решения, которые нашли свое применение в известных клиринговых центрах, могут быть использованы в определенной мере при создании новых систем, работающих с данными ДЗЗ.

С учетом вышеизложенного при построении системы каталогизации и распределенного доступа к данным ДЗЗ необходимо было решить следующие основные задачи:

- разработать и обосновать профиль стандарта на метаданные данных ДЗЗ, хранящихся в ЦАКИЗ;
- разработать и обосновать архитектуру и концепцию развития системы;
- разработать предложения по программно-аппаратной реализации системы;
- построить пилотный образец системы.

Далее кратко описывается порядок решения названных задач.

Описание профиля стандарта на метаданные ДЗЗ. При создании профиля были использованы имеющиеся публикации в этой предметной области, опубликованные стандарты

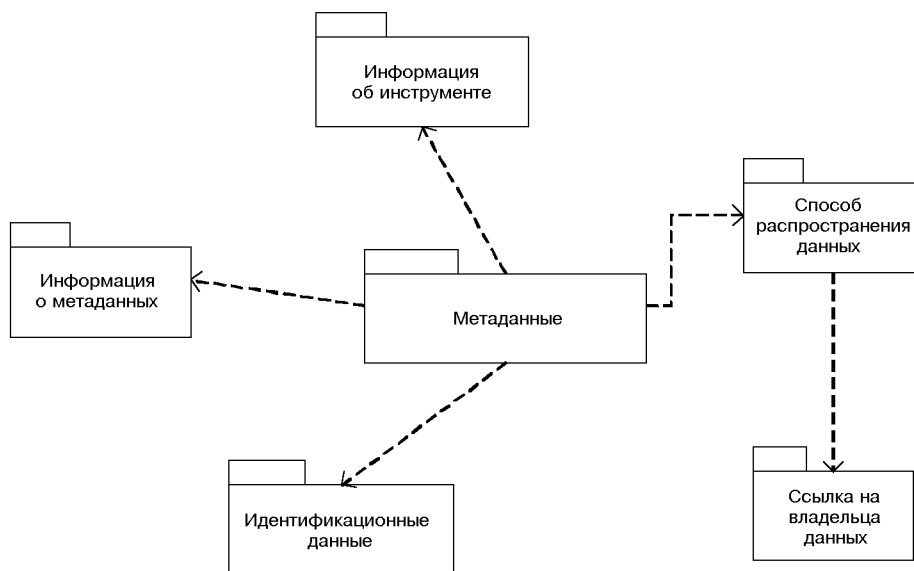


Рис. 4. UML-диаграмма профиля стандарта ЦАКИЗ

на метаданные для данных ДЗЗ, описанные выше подходы к созданию профилей стандартов. Был проведен анализ текущих и будущих потребностей ЦАКИЗ в отношении метаданных об изображениях, которые хранятся в организации. Сначала был произведен выбор международного стандарта на метаданные, который наиболее целесообразно использовать в качестве основы профиля стандарта ЦАКИЗ. Разумеется, наилучшим вариантом для этого мог быть стандарт ISO 19115-2 [http://www.openspatial.org], однако на момент начала выполнения описываемого проекта он еще находился в стадии разработки, и его окончательный вариант не был опубликован. Поэтому было решено в качестве исходной основы использовать аналогичный стандарт Геологической службы США.

При создании профиля стандарта были сохранены все наименования полей и типы данных, используемые в исходном стандарте. Поскольку на первом этапе построения системы от него не требуется сложной функциональности, содержание профиля стандарта было максимально упрощено. Содержание основных информационных групп разработанного профиля стандарта изображено на рис. 4 в виде диаграммы, созданной с использованием универсального языка моделирования UML (Universal Modelling Language). Из большого количества информационных групп, предлагаемых базовым стандартом, было выбрано четыре основных, которые содержат только самые необходимые данные об изображениях. Информация о метаданных организована по иерархическому принципу.

Для облегчения работы пользователя, а также для обеспечения лучшей поддержки целостности базы метаданных для некоторых полей вместо типа данных «текст» использован перечисляемый тип с конечным числом возможных значений соответствующего атрибута (либо доменное ограничение «любой текст» заменено на перечисляемую совокупность допустимых значений атрибута).

Чтобы реализовать такую возможность, для полей «Название миссии», «Датчик», «Уровень обработки» созданы словари, которые включают весь перечень допустимых значений соответствующих атрибутов. В качестве примера в таблице приведено описание поля «Уровни обработки».

Уровни обработки снимков

Название уровня обработки	Пояснение
Уровень 0	Декодированный и сформированный снимок с присоединенными к нему метаданными
Уровень 1	Откорректированный снимок, привязанный по данным орбиты носителя
Уровень 2	Геореференцированный снимок
Уровень 3	Дешифрированный снимок с легендой
Уровень 4	Тематическая карта, созданная с использованием снимка

Архитектура системы. Архитектура системы каталогизации и распределенного доступа к данным ДЗЗ, реализованной в ЦАКИЗ, показана на рис. 5.

При разработке архитектуры системы учитывались следующие обстоятельства:

- аппаратно-программная архитектура компьютерных ресурсов ЦАКИЗ;
- перспективы развития этих ресурсов на ближайшее будущее;
- текущие и перспективные потребности ЦАКИЗ в отношении управления накопленными в организации аэро- и космическими снимками (изображениями) и их метаданными.

Конечные пользователи, используя web-браузер, связываются с некоторым ресурсом (Внешний источник данных) в интернете, предоставляющим космические изображения (или аэроснимки), и получают эти изображения. Далее снимок помещается в локальное хранилище данных компьютера пользователя («Локальные данные»). Кроме того, пользователь должен передать метаданные полученного изображения администратору метаданных. Формирование этих метаданных выполняется в стандартном формате, который определяется описанным выше профилем стандарта на метаданные. Администратор метаданных проверяет полученное от пользователя сообщение на предмет отсутствия формальных ошибок и помещает новую запись о метаданных в базу метаданных, расположенную на его компьютере.

Таким образом, только администратор имеет право вносить любые изменения в базу метадан-

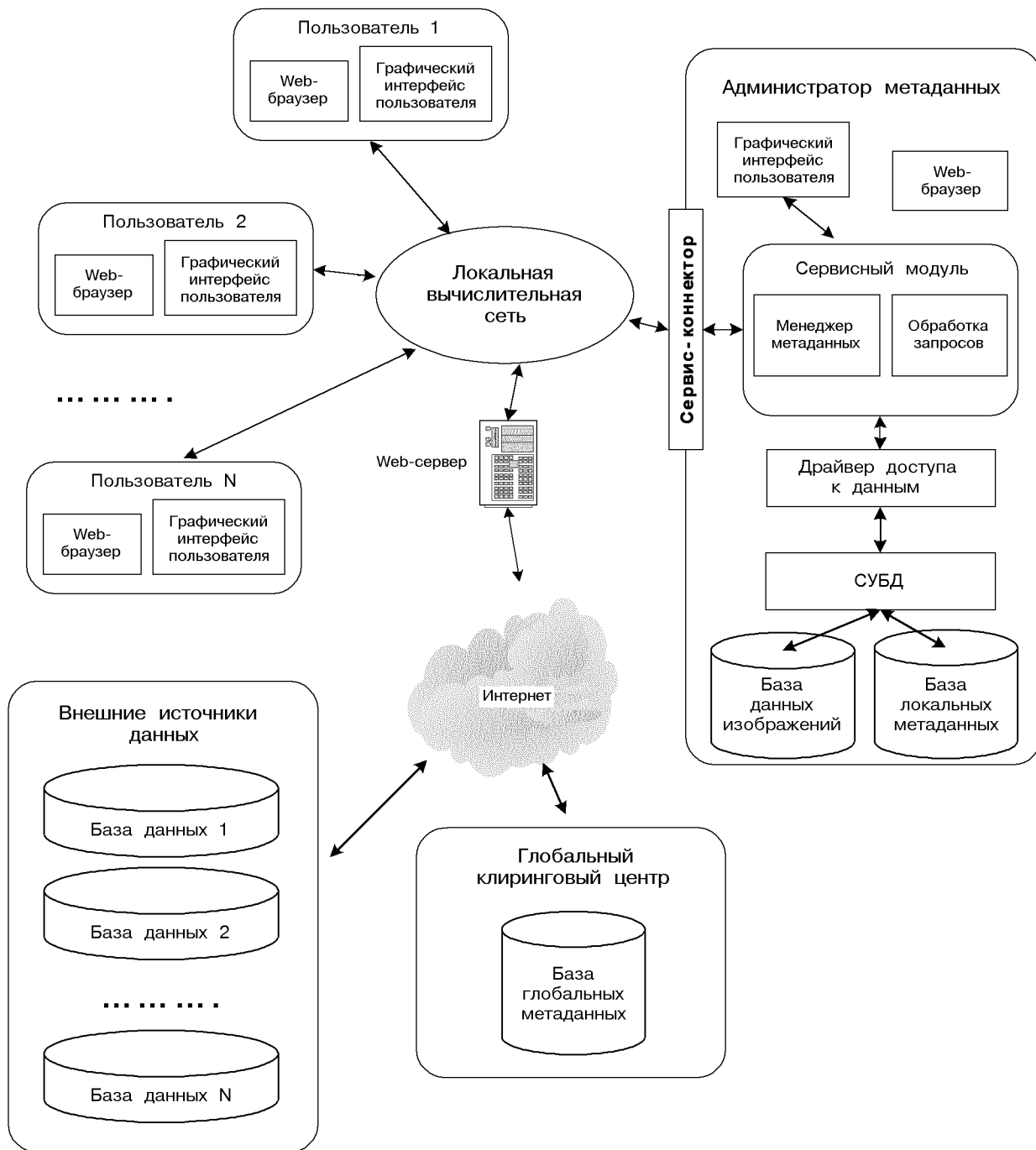


Рис. 5. Архитектура системы

ных (используется подмодуль «Менеджер метаданных» сервисного модуля). Если пользователь замечает ошибку в метаданных хранящегося у него изображения, исправления производятся также через администратора.

Если пользователь хочет сделать запрос к базе метаданных, чтобы узнать о наличии в базе ЦАКИЗ интересующих его изображений, он, используя графический интерфейс (GUI — Graphical User Interface), формирует запрос. В качестве параметров запроса могут выступать:

- дата съемки;
- географическое место съемки;
- географические координаты экстенда, покрывающего район съемки;
- носитель;
- сенсор (датчик).

Запрос через локальную сеть пересылается на компьютер администратора метаданных и далее в подмодуль «Обработка запросов» сервисного модуля. В соответствии с параметрами запроса из базы метаданных извлекается информация, которая пересылается назад, на компьютер пользователя. В числе прочих данных там содержится ссылка на местоположение самого изображения.

В будущем предполагается несколько усовершенствовать архитектуру созданной системы. Основные отличия новой информационной инфраструктуры по сравнению с существующей будут состоять в следующем.

- Хранение всех изображений на одном сервере (сервере администратора базы метаданных). На том же сервере будет храниться, как и в настоящее время, локальная база метаданных этих изображений. К этим ресурсам обеспечивается централизованный доступ пользователей организации, а управление будет выполняться, как и сейчас, администратором базы метаданных.
- Создание сокращенной копии базы метаданных (глобальной базы метаданных), которая помещается на некоторый внешний (или глобальный) сервер. Предназначение этого сервера — предоставление информации о наличных в ЦАКИЗ изображениях для внешних пользователей.
- Более совершенный профиль стандарта ЦАКИЗ с точки зрения более полного описания изображений (расширения перечня

метаданных).

- Обеспечение записи информации о метаданных в XML-совместимый формат для стандартизации процедур обмена метаданными с другими организациями;
- Улучшение графического интерфейса пользователя, реализация в ней картографического компонента.

Предполагается, что периодически функциональность системы будет совершенствоваться и далее, и будут создаваться новые версии соответствующего программного обеспечения.

Программно-аппаратная реализация системы. Пилотная версия системы реализуется на аппаратных ресурсах, наличных в ЦАКИЗ. Для упрощения используется одноранговая сетевая инфраструктура на платформе Windows, в которой один из компьютеров играет роль администратора метаданных.

Для программной реализации сервисов работы с метаданными используются бесплатно лицензируемые программные продукты:

Apache — для реализации серверного компонента приложения;

MySQL — для реализации СУБД.

Это обеспечивает наименьшие расходы на создание соответствующих приложений, и в то же время снимает проблему лицензирования средств разработки.

ЗАКЛЮЧЕНИЕ

Описанная система каталогизации и распределенного доступа к данным дистанционного зондирования Земли может быть развернута в любой организации, использующей в своей работе аэрокосмическую информацию.

Если же рассматривать проблему повышения эффективности использования распределенных геопространственных ресурсов более широко, то очевидной становится необходимость применения технологий создания геоинформационных порталов, на которых будут храниться ссылки на все эти ресурсы и краткое их описание. Очевидно, пришло время подумать о создании такого портала в масштабах Украины для данных ДЗЗ. При построении такого портала целесообразно использовать рекомендации консорциума OpenGIS, а также опыт, накопленный

геоинформационным сообществом. Несомненно, важнейшим элементом портала будет система метаданных о данных ДЗЗ, поэтому вопросы, обсуждаемые здесь, представляются нам довольно актуальными.

В настоящее время пилотный проект системы проходит опытную эксплуатацию. После этого можно будет ознакомиться с версией системы через сайт ЦАКИЗ www.casre.kiev.ua, и даже получить ее демонстрационную версию.

1. Geospatial data infrastructure: concepts, cases and good practice / Ed. by R. Groot, J. McLaughlin. — New York: Oxford Univ. Press, 2000.—286 p.

THE SYSTEM OF CATALOGUING AND DISTRIBUTED ACCESS TO REMOTE SENSING DATA: CONCEPTION, ARCHITECTURE, REALIZATION

*M. A. Popov, S. Yu. Markov,
S. A. Stankevich, A. L. Bodnia*

Our conceptual approach to system of cataloguing and distributed access to remote sensing data creation is considered. At present the system is realized as pilot project at the Scientific Centre for Aerospace Researches of the Earth of IGS of NAS of Ukraine. The approach takes into account the world experience of geospatial data infrastructures designing. The profile of standard for airborne and spaceborne imagery metadata is developed based on the international standards and real requirements of geospatial information consumers. The architecture of the system is described and some features of its hardware and software realization are considered.