На правах рукописи

Дубов Сергей Сергеевич

РАЗРАБОТКА АРХИТЕКТУРЫ И АЛГОРИТМОВ РАБОТЫ ФЕДЕРАЛЬНОГО ХРАНИЛИЩА ПЛАНЕТАРНЫХ ДАННЫХ

Специальность 25.00.35 – Геоинформатика

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук

Москва – 2011

Работа выполнена в ФБГОУ ВПО «Московский государственный университет геодезии и картографии (МИИГАиК)»

Научный руководитель: доктор физико-математических наук, профессор
Шингарева Кира Борисовна

Официальные оппоненты: доктор технических наук, профессор
Чугреев Игорь Григорьевич

Кандидат технических наук
Тихонов Александр Дмитриевич

Ведущая организация: **Государственный астрономический институт им. П.К. Штернберга МГУ**

Защита состоится «23» декабря 2011 г. в 12 часов на заседании диссертационного совета Д 212.143.03 при Московском государственном университете геодезии и картографии (МИИГАиК) по адресу**:** 105064, Москва, Гороховский пер., 4, зал заседаний ученого совета.

С диссертацией можно ознакомиться в библиотеке Московского государственного университета геодезии и картографии (МИИГАиК).

Автореферат разослан «22» ноября 2011 г.

Ученый секретарь
диссертационного совета Климков Ю.М.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы.

В связи с ростом интенсивности и масштабов космической деятельности, в том числе исследований планет и внеземных территорий, возросли разнообразие и объем получаемых планетарных данных. Это породило проблемы, связанные с хранением значительных объемов территориально распределенной информации, а также ее доступностью для удаленных пользователей. Суть этих проблем сводится, во-первых, к предметизации планетарных данных по космическим телам и процессам, направлениям исследовательской деятельности, способам изучения свойств космических тел, во-вторых, к систематизации, классификации, каталогизации и индексации информационных объектов, содержащих планетарные данные, в- третьих к созданию рациональной электронной информационной инфраструктуры хранения и доступа к ним. Анализ инфраструктуры хранения планетарных данных Российской Федерации показал, что на сегодняшний день она дезинтегрирована. Коллективное использование и синтез планетарных на сегодняшний день существенно затруднены для данных для анализа, обработки и построения новых, более адекватных моделей небесных тел. Таким образом, задачи разработки архитектуры, алгоритма и работы федерального хранилища планетарных данных являются важными и актуальными.

Объект исследования и предмет исследования. Объектом исследования являются информационные процессы хранения и доступа к планетарным данным. Предметом исследования являются архитектуры, алгоритмы и сценарии работы федерального хранилища планетарных данных.

Целью исследования является разработка архитектуры, алгоритма и сценария работы федерального хранилища планетарных данных.

Для достижения данной цели необходимо решить следующие ***задачи***:

- Провести анализ процессов хранения и доступа к планетарным данным;
- Сформулировать требования к федеральному хранилищу планетарных данных;
- Разработать архитектуру федерального хранилища планетарных данных;
- Разработать сценарии работы федерального хранилища планетарных данных;

Научная новизна диссертации определяется спецификой предметной области исследования и заключается:

Для архитектуры федерального хранилища планетарных данных в применении метода анализа иерархий для формирования рациональных вариантов организации хранилищ проектируемого федерального хранилища планетарных данных, а также в обосновании механизма прогнозирования значений глобальных приоритетов при сравнении альтернативных вариантов.

Для алгоритма работы федерального хранилища планетарных данных, в реализации механизмов поиска, добавления и обновления информации в нем.

На защиту выносятся следующие **основные результаты**:

1. Архитектура федерального хранилища планетарных данных;
2. Алгоритм работы федерального хранилища планетарных данных.

Практическое значение заключается в том, разработанные требования, архитектурное решение и алгоритмы работы федерального хранилища планетарных данных позволят обеспечить систематизированное хранение, а также доступность для удаленных пользователей.

Методы исследования. Решение поставленных задач выполнено с использованием методов научно-теоретического анализа, объектно-ориентированного проектирования и системного анализа, аналитических сетей и анализа иерархий.

Структура и объем работы. Диссертация состоит из введения, трех глав, заключения, списка литературы, включающего 47 наименования источников и 2 приложений. Общий объем диссертации: 90 страницы текста и 10 страниц приложения, в том числе 5 таблиц и 12 рисунков в тексте.

Апробация работы и публикации. По теме диссертации опубликована 1 статья в журнале, рекомендованном ВАК России, промежуточные и окончательные результаты работы представлены в 4 сборниках международных конференций.

Апробация архитектуры федерального хранилища планетарных данных проведена на третьем этапе научно исследовательской работы «Разработка обобщенной открытой

программной платформы для обработки и хранения пространственных данных» № УДК 01201068386 выполнявшейся по заказу Министерства образования и науки Российской Федерации.

Апробация алгоритмов работы федерального хранилища планетарных данных проведена на третьем этапе научно исследовательской работы «Разработка модели геоданных и способов её применения при построении единой инфраструктуры пространственных данных Российской Федерации»№ 01201057638 выполнявшейся по заказу Министерства образования и науки Российской Федерации. Обе работы завершены. Положительные экспертные заключения по этим этапам НИР позволяют сделать выводы о достоверности полученных результатов.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во **введении** раскрыты актуальность темы, цель и задачи, научная новизна и практическая ценность научных результатов.

В первой главе проведен анализ процессов хранения планетарных данных. Исследованы основные характеристики и выявлены особенности их хранения. Рассмотрены нормативные документы, определяющие требования к структуре и содержанию планетарных данных и метаданных к ним. Исследованы существующие системы хранения планетарных данных в России и за рубежом, сформулированы требования к федеральному хранилищу планетарных данных.

Под **планетарными данными** будем понимать данные, отражающие характеристики небесных тел (за исключением Земли) и деталей рельефа внеземных территорий, измеренные с использованием научной аппаратуры и их производные.

Рисунок 1. Планетарные данные.

Анализ процессов хранения показал, что сегодня планетарные данные хранятся в различных организациях, в каждой из которых они имеют собственную, удобную для данной организации, классификацию, структуру и формат хранения, что затрудняет возможность их совместного использования для построения более точных моделей небесных тел и внеземных территорий. В процессе анализа основных характеристик были выявлены уникальные свойства планетарных данных: узкий круг потребителей, высокая научная ценность, большой объем единицы данных.

Анализ нормативных документов, определяющих требования к структуре и содержанию планетарных данных и метаданных к ним, показал, что данный вопрос проработан в США. Разработаны стандарты, четко регламентирующие формат и структуру данных и атрибутивной информации (метаданных).

В состав метаданных входят: Общие сведения, информация о использованном стандарте метаданных, орбитальные данные, информация о инструменте с помощью которого были получены данные, информация о космической программе, информация о пространственно временных характеристиках, информация о качестве, информация о правилах распространения и информация об авторе.

Исследованы существующие системы хранения планетарных данных в России (см. таблицу 1) и за рубежом (см. таблицу 2)

Таблица 1. Наиболее известные Российские информационные ресурсы по планетной тематике.

http://www.iki.rssi.ru/pe.html	Информационный ресурс, содержащий научную информацию по планетологии, астрофизике космическим миссиям, и приборам для исследования космического пространства.	ИКИ РАН, Россия
http://www.sai.msu.ru/basa/index.html	Информационный ресурс, содержащий научную информацию по астрономии, астрофизике, небесной механике и т.д. различные базы данных, каталоги звёзд и другую научную информацию по исследованию небесных тел.	ГАИШ МГУ, Россия
http://intranet.geokhi.ru	Информационный ресурс, содержащий научную информацию по геохимии планет, сравнительной планетологии, космохимии и т.д.	ГЕОХИ РАН, Россия
www.astronet.ru	Информационный ресурс, содержащий обзорную информацию по планетологии.	ГАИШ МГУ Россия
http://planetmap.ru/	Информационный ресурс, содержащий обзорную информацию по планетологии.	МИИГАиК, Россия
http://galspace.spb.ru	Информационный ресурс, содержащий обзорную информацию по планетологии, советские снимки Луны	Лобанов А.В. Россия
http://phobos.cosmos.ru	Информационный ресурс, содержащий информацию по космической программе Фобос-грунт	ИКИ, НПО им. Лавочкина, Россия
http://venera-d.cosmos.ru/	Информационный ресурс, содержащий информацию по космической программе Венера Д	ИКИ, НПО им. Лавочкина, Россия

Таблица 2.Наиболее известные зарубежные системы хранения планетарных данных.

Название системы	описание	Разработчик
Planetary Data System (PDS)	Распределенная система, содержащая наиболее полную коллекцию планетарных данных и информационных ресурсов по планетной тематике.	NASA, США
Map-a-Planet	Геоинформационная система отображения планетарных данных. Входит в состав PDS	USGS, США
United State Geological Survey (USGS)	В нем собраны Изображения планет и спутников, сделанные в результате различных миссий по изучению космоса	USGS, США
http://www.rssd.esa.int	Planetary Science Archive, европейский аналог PDS, использует стандарт PDS для описания структуры и формата планетарных данных	ESA, Евросоюз

США накапливают, структурируют и предоставляют свободный доступ к результатам всех своих миссий и научных исследований в области освоения космоса.

Евросоюз имеет хранилище планетарных данных, к которому обеспечен свободный доступ через Интернет.

Россия имеет большое количество справочных интернет ресурсов по планетологии, астрофизике, геохимии и другим направлениям космических исследований. В некоторых из них имеются планетарные данные в удобной для организации-владельца структуре и формате. Центрального хранилища, аккумулирующего знания по исследованию небесных тел, в России не существует. Российские исследователи в качестве источника планетарных данных зачастую используют информационные ресурсы других государств.

Первой стадией разработки федерального хранилища планетарных данных является формирование требований к нему. Требования приведены на рисунке 2.

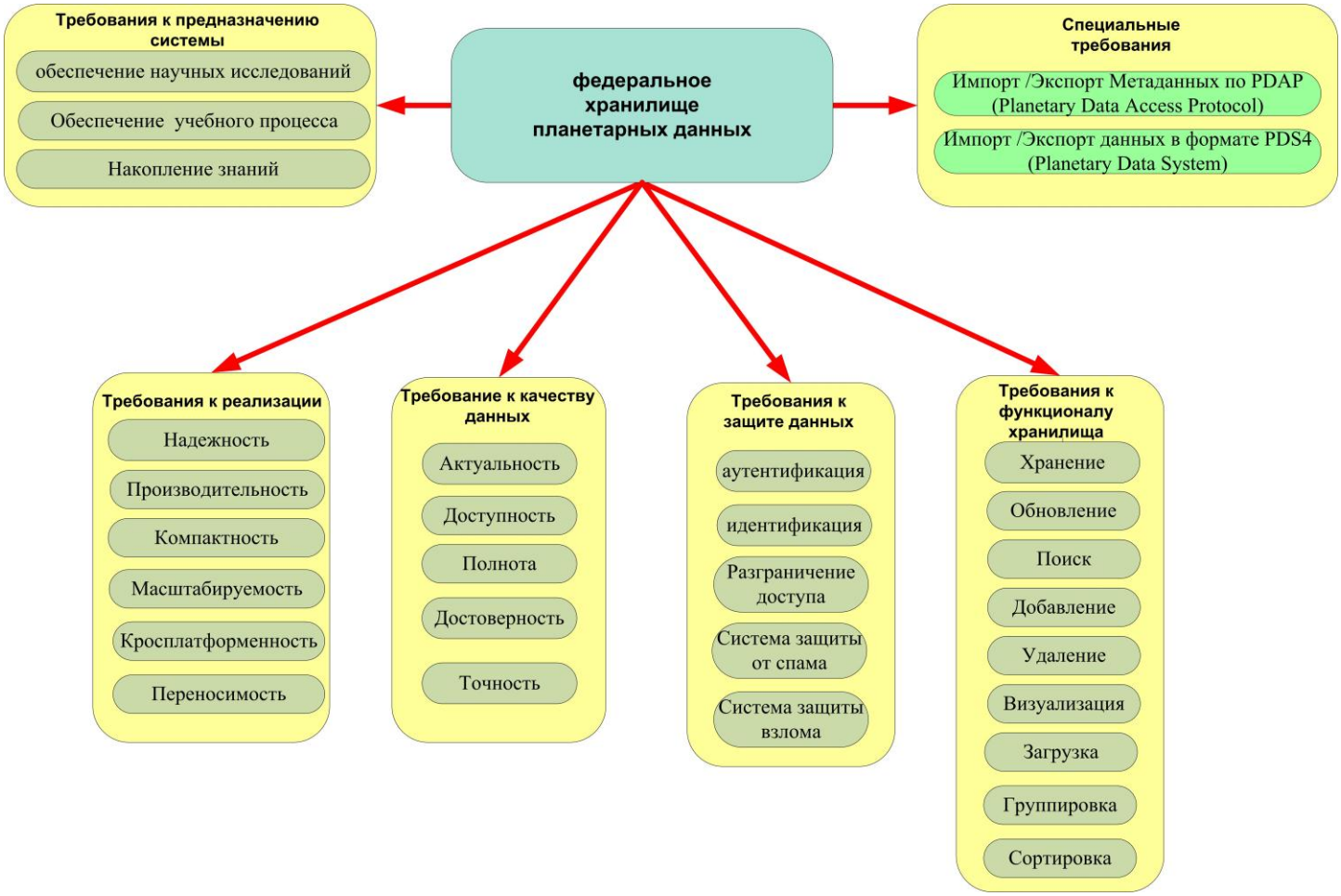


Рисунок 2. Требования к федеральному хранилищу планетарных данных

Вторая глава посвящена разработке архитектуры федерального хранилища планетарных данных.

Для обоснованного выбора архитектурного решения системы управления метаданными федерального хранилища планетарных данных проанализируем организацию среды управления метаданными, которая, в общем случае, включает в себя источники и репозиторий метаданных, а также средства их интеграции, управления, доставки и публикации. Упрощенная схема среды управления метаданными приведена на рисунке 3.

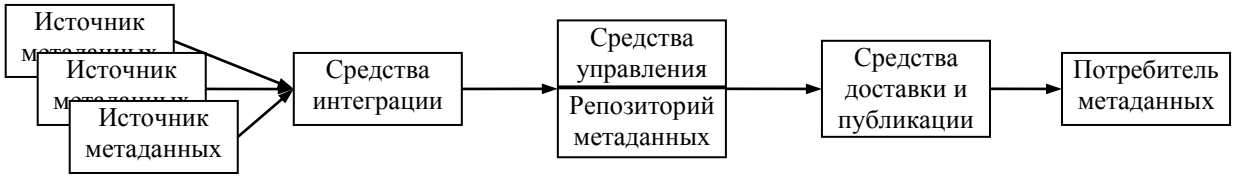


Рисунок.3 – Упрощенная схема среды управления метаданными федерального хранилища планетарных данных.

Источники метаданных – это все организации, которые включены в систему управления метаданными федерального хранилища планетарных данных.

Средства интеграции предназначены для извлечения метаданных из источников, их интеграции и размещения в репозитории метаданных;

Средства управления метаданными обеспечивают определение и соблюдение прав доступа к содержимому репозитория.

Средства доставки и публикации должны переслать метаданные (в соответствии с правами доступа, графиком доставки и требованиями к составу) в точки их потребления субъектами федерального хранилища.

Репозиторий содержит метаданные, организованные по единой и согласованной модели, и может, в свою очередь, иметь собственную архитектуру (централизованную, децентрализованную или распределенную), каждая из которых имеет свои особенности. Проанализируем эти архитектуры и их особенности для построения репозитория метаданных.

Централизованная архитектура предполагает наличие глобального репозитория, который построен на основе единой модели метаданных и обслуживает все хранилища организаций источников. Локальные репозитории отсутствуют. Имеется единственная, единообразная и согласованная модель метаданных. С позиции федерального хранилища это архитектурное решение потребует доступа всех организаций (источников планетарных данных) к единому центральному репозиторию метаданных. Это весьма проблематично с позиции организации телекоммуникаций, а работы по созданию собственно репозитория натолкнутся на согласование между организациями единой модели метаданных.

В распределенной архитектуре глобальный репозиторий содержит метаданные всех хранилищ организаций источников планетарных данных. Локальные репозитории, содержащие подмножество метаданных одного хранилища организации, обслуживают локальные планетарные данные этой организации. Модель метаданных является

единообразной и согласованной. Все метаданные локальных репозиториев проходят обработку и согласование в глобальном репозитории. Достоинства локальных репозиториев компенсируются требованиями к их синхронизации с центральным репозиторием метаданных. Распределенная архитектура является более предпочтительной для организации репозитория метаданных ввиду территориальной распределенности самих хранилищ организаций - источников планетарных данных.

Децентрализованная архитектура предполагает, что центральный репозиторий содержит только ссылки на метаданные, которые ведутся независимо в локальных репозиториях. Отсутствие затрат на согласование терминов и понятий значительно сокращает стоимость разработки. Применимость этой архитектуры для организации репозитория метаданных ограничена множеством и разнообразием моделей, возможно несовместимых друг с другом. Это имеет место, когда интегрируются информационные системы организаций источников планетарных данных внутри пересекающихся направлений космической деятельности.

Таким образом, наиболее приемлемым архитектурным решением репозитория метаданных в нашем случае является распределенная архитектура.

С учетом вышеизложенного, специфика организации репозитория для федерального хранилища диктует необходимость выбора архитектурного решения системы управления метаданным по типу «двойная звезда», как наиболее адекватно реализующего необходимые связи между интегрируемыми информационными системами организаций - источников планетарных данных. При таком решении, с одной стороны, сохраняется необходимая территориальная распределенность, а с другой – обеспечивается требуемая интегрированность метаданных. Абстрактная модель такого архитектурного решения представлена на рис. 4. Для выбранного архитектурного решения системы управления метаданными, источниками метаданных являются метаданные хранилищ организаций - источников планетарных данных. Хранение самих планетарных данных осуществляется в тематических БД, входящих в состав хранилищ,

находящихся в ведении соответствующих организаций.

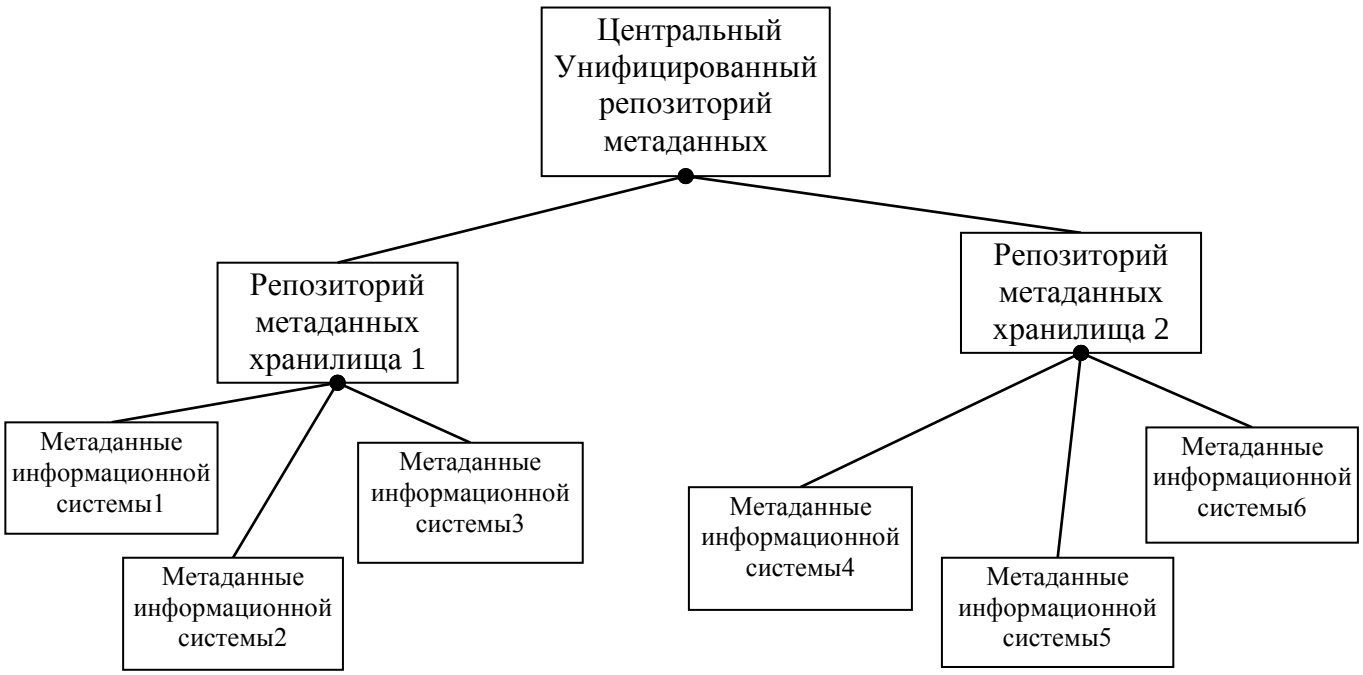


Рисунок 4. – Абстрактная модель архитектурного решения системы управления метаданными федерального хранилища планетарных данных

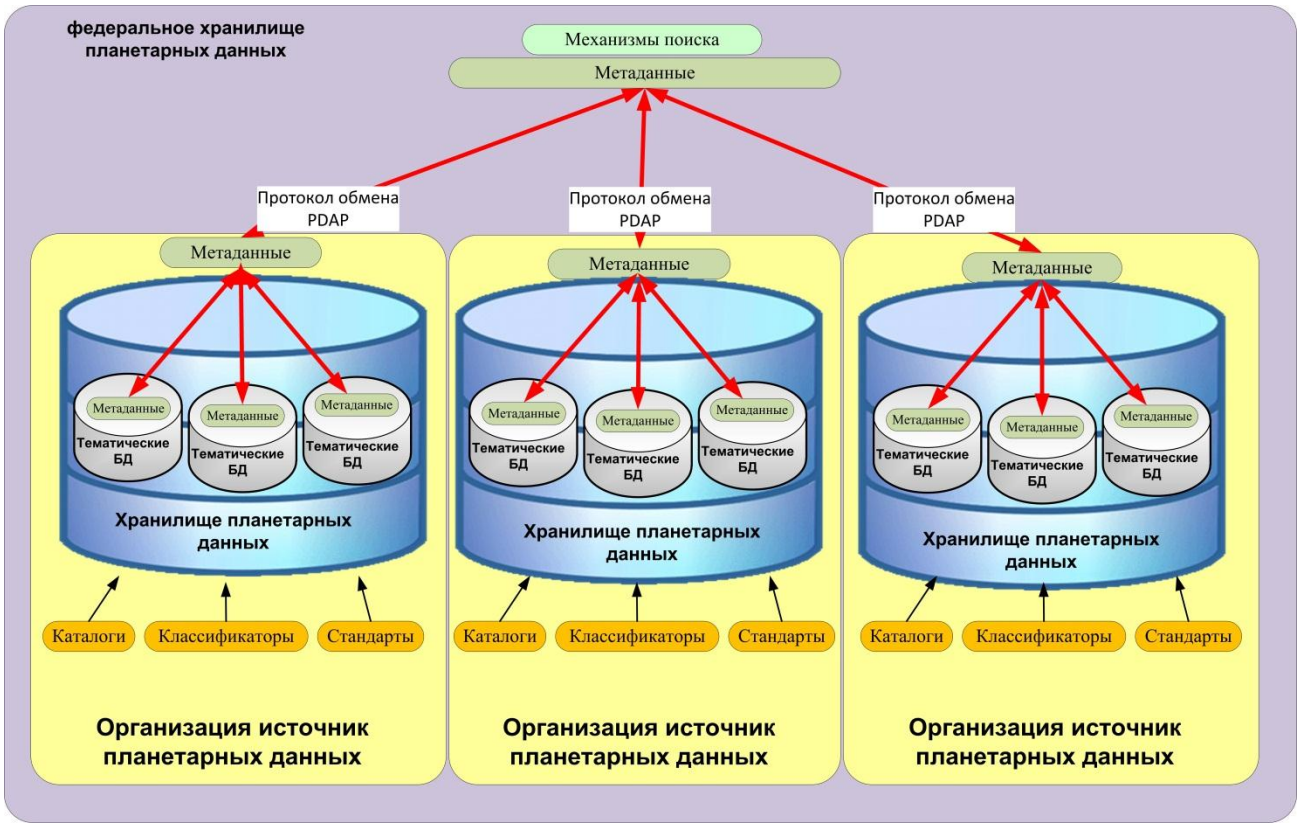


Рисунок 5. Архитектура федерального хранилища планетарных данных

Такая архитектура федерального хранилища позволяет;

- производить поиск планетарных данных из одной точки;
- избежать дублирования планетарных данных;
- снизить затраты, связанные с хранением планетарных данных.

Рассмотренные требования и архитектура федерального хранилища планетарных данных описывают его облик и позволяют перейти к рассмотрению алгоритмов его работы.

В **третьей** главе рассмотрены алгоритмы работы федерального хранилища планетарных данных.

Алгоритмы работы федерального хранилища планетарных данных, как самостоятельной информационной системы представлен на рисунке 6.

Рассмотрим алгоритм работы федерального хранилища планетарных данных на примере базовых операций в хранилища.

Поиск информации будет осуществляться посредством центрального унифицированного репозитория метаданных для данных, хранящихся во всех информационных системах организаций - источников планетарных данных.

Эффективность поиска планетарных данных зависит от многих факторов. Их можно разделить на системные и внесистемные.

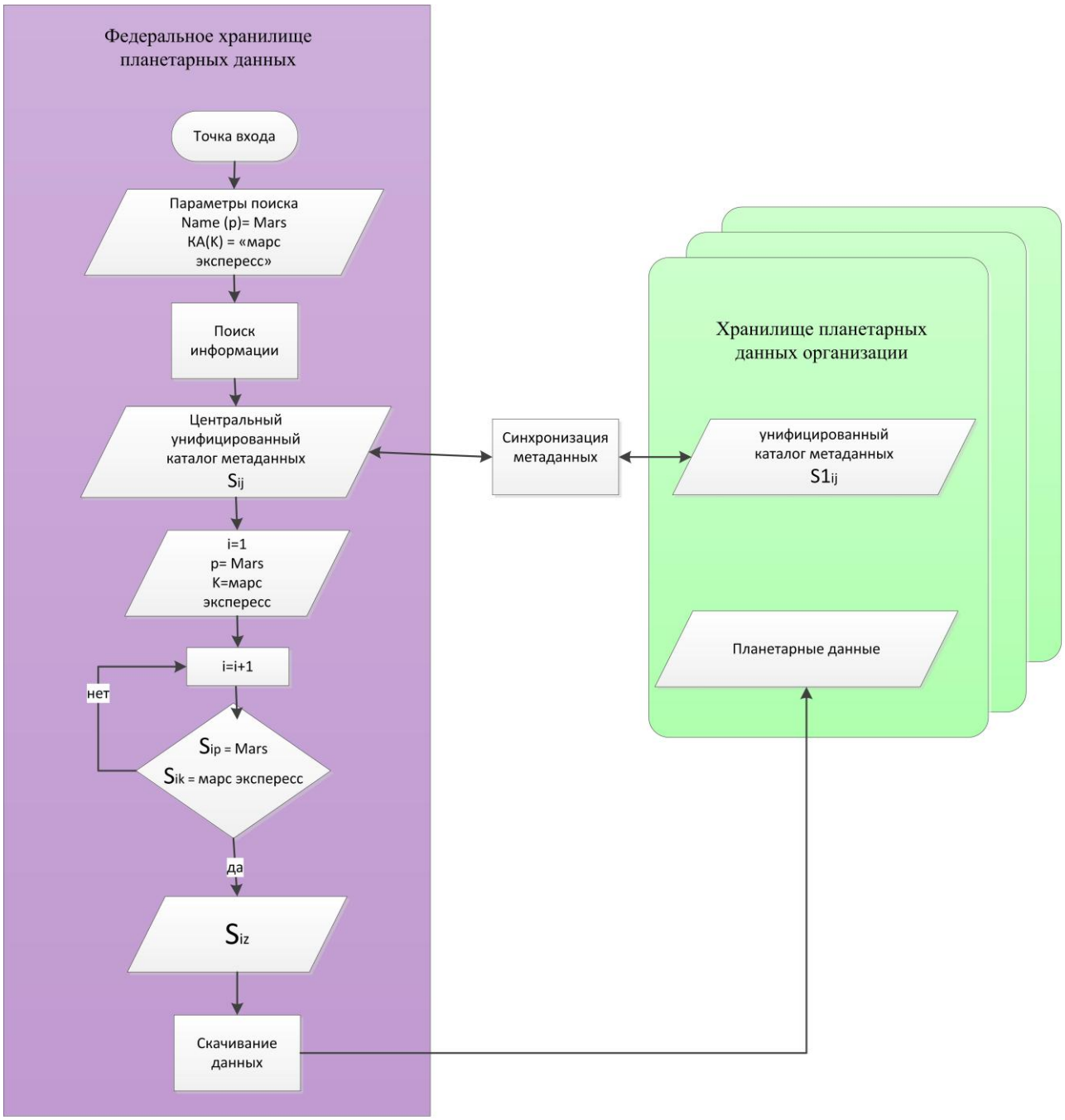
К внесистемным факторам относятся достоверность, актуальность и полнота данных, на которые мы повлиять с точки зрения системного подхода не можем.

К системным факторам относятся систематизация, классификация и унификация, производительность аппаратно-программной части.

Для повышения эффективности поиска при разработке федерального хранилища планетарных данных использовались методы систематизации, классификации и унификации метаданных составляющих центральный репозиторий метаданных.

Поиск информации в федеральном хранилище планетарных данных сводится к выбору критериев, заданию в них требуемых значений и активации механизма поиска. Поиск планетарных данных по выбранным критериям осуществляется через центральный каталог метаданных, центральный каталог метаданных содержит информацию из каталогов метаданных организаций и каталогов метаданных других систем хранения планетарных данных. Результат поиска

выводится в виде ключевых характеристик планетарных данных и ссылки на файл для скачивания.



P= Столбец Name (Название небесного тела)
K= Столбец КА (название космического аппарата)
Sji Множество значений унифицированного каталога
Siz Ссылка на загрузку файла, где Sip = Mars, Sik=марс экспересс

Рисунок 6. Алгоритм работы федерального хранилища планетарных данных, как самостоятельной структурной единицы.

Механизмы добавления, обновления и удаления информации из федерального хранилища планетарных данных отражены в алгоритме его работы (рисунок 6) и

заключаются в наличии отдельной точки входа в хранилище организации - источника этих данных. Только в этом хранилище оператор будет иметь права на редактирование.

В 2006 году NASA и ESA создали Международный альянс планетарных данных (International Planetary Data Alliance или IPDA), целью которого является интеграция всех хранилищ в единую централизованную систему, использующую единые стандарты, определяющие структуру и формат этих данных. IPDA ведется активная работа по созданию единого формата обмена метаданными между различными информационными системами хранения планетарных данных. Сегодня в состав альянса входят США, Германия, Франция, Индия, Италия, Китай, Япония, Россия (ИКИ РАН).

Интеграция хранилищ планетарных данных в рамках международного альянса планетарных данных реализована через полносвязную архитектуру, в которой все хранилища связаны друг с другом напрямую. Центральное хранилище в архитектурном решении данного типа отсутствует, что позволяет обеспечить достаточно большую надежность системы.

Алгоритм работы федерального хранилища в рамках международного альянса планетарных данных представлен на рисунке 7.

Методы интеграции и обмена данными с внешними системами хранения планетарных данных продемонстрированы на рисунке 8. Они позволяют быстро и с высокой периодичностью обновлять информацию в базах данных, не загружая сетевой трафик.

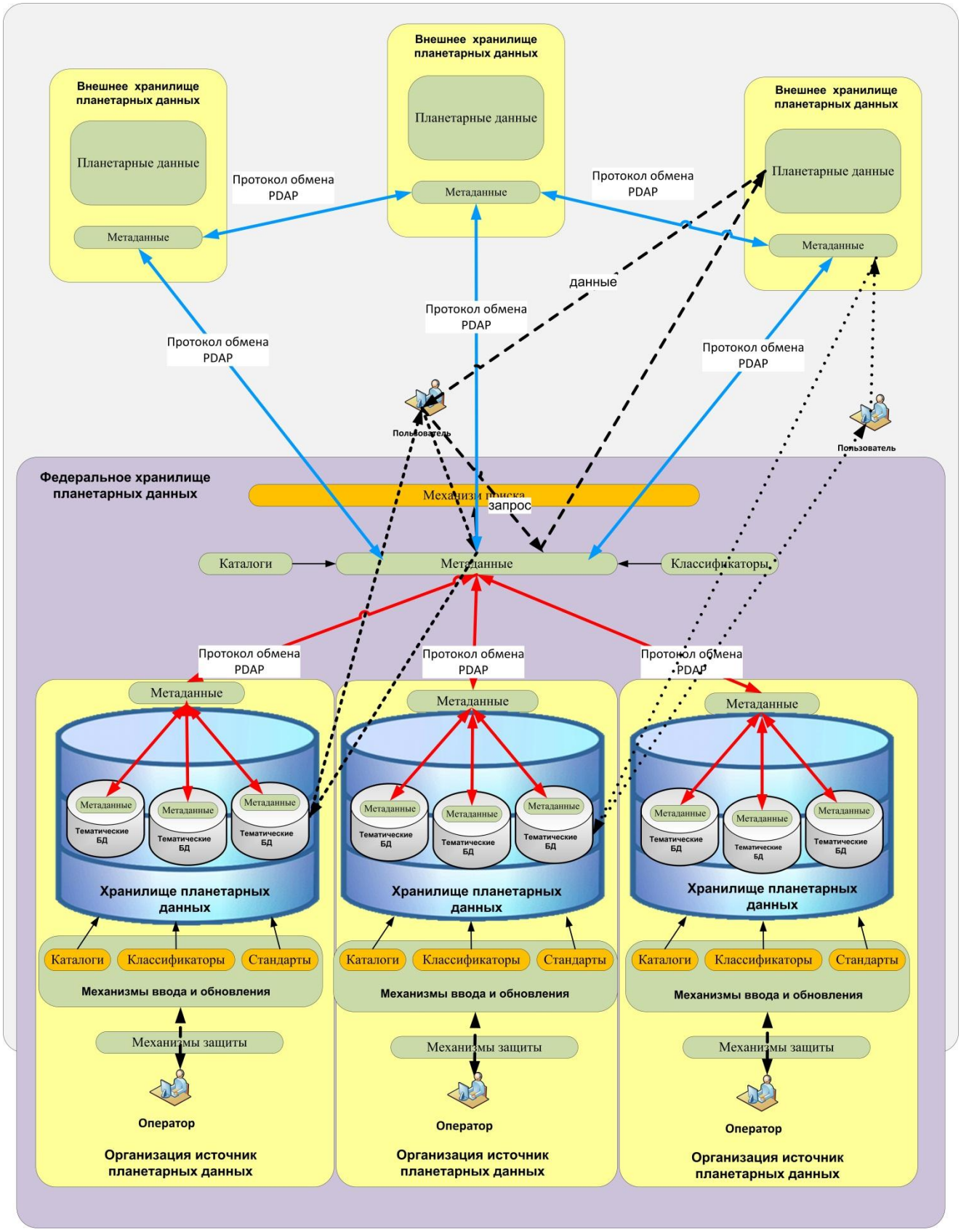


Рисунок 7 Сценарий работы федерального хранилища планетарных данных

Заключение

Анализ информационных систем показал, что в зарубежных странах существуют федеральные системы хранения планетарных данных, как правило, они прикреплены к федеральному космическому агентству. В настоящее время рядом стран ведется активная работа по объединению планетарных баз данных. На сегодняшний день Россия имеет большое количество справочных ресурсов по планетологии. В некоторых из них встречаются единичные планетарные данные. Федеральной системы хранения планетарных данных на государственном уровне не существует. Российские исследователи в качестве источников планетарных данных часто используют информационные ресурсы других государств. Доступ к планетарным данным, полученным в ходе реализации Российских космических программ, сильно затруднен.

В процессе исследования были решены следующие задачи:

- Проведен анализ процессов хранения и доступа к планетарным данным;
- Сформулированы требования к федеральному хранилищу планетарных данных.
- Разработана архитектура работы федерального хранилища планетарных данных;
- Разработан алгоритм работы федерального хранилища планетарных данных.
- Разработан сценарий работы федерального хранилища планетарных данных.

Решение вышеперечисленных задач говорит о том, что цель исследования была достигнута.

Достоверность исследования определяется апробацией научных результатов в рецензируемом журнале, на международных конференциях, в научно-технических отчетах имеющих государственную регистрацию и получивших положительную экспертную оценку.

В ближайшем будущем планируется разработка макета картографического сегмента федерального хранилища планетарных данных на базе комплексной лаборатории планетной картографии Московского государственного университета геодезии и картографии

Разработка федерального хранилища планетарных данных позволит увеличить скорость исследований, осуществить систематизацию и накопление знаний по планетной тематике, обеспечит доступность планетарных данных для научных исследований, что позволит перейти к построению более точных и разноплановых моделей.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ РАБОТЫ

1. Архитектура федерального хранилища планетарных данных;
2. Алгоритм работы федерального хранилища планетарных данных.

РАБОТЫ, ОПУБЛИКОВАННЫЕ ПО ТЕМЕ ДИССЕРТАЦИИ:

1. BILINGUAL MULTIMEDIA ELECTRONIC VERSION ON "ATLAS OF TERRESTRIAL PLANETS AND THEIR MOONS" // S.S. Dubov, I.Y. Rozhnev, S.M. Leonenko, 2007.
2. Development of a Database Model on a Cartographic Studyness of the Moon // S.S. Dubov - 2nd International Conference on Cartography & GIS, January 2008, Bulgaria.
3. "Выбор поверхности относимости и системы координат для картографо-геодезического обеспечения посадки на Фобос КА "Фобос-Грунт" // Дубов С.С., Зельков К., Конопихин А.А., Краснопевцева Б.В., Лобанов А.А., Рожнев И.Ю., Шингарева К.Б. - Международная научно-техническая конференция, посвященная 230-летию МИИГАиК, май 2009, Москва.
4. Geodetic and Cartographic Aspects of «Phobos-Grunt» Projects on the Base of GIS-technologies // S.S.Dubov, K.M.Zel'kov, A.A.Konopikhin, B.V.Krasnopevtseva, A.A.Lobanov, Ivan Rozhnev, K.B.Shingareva - 3rd International Conference on Cartography & GIS, June 2010, Bulgaria.
5. Майоров А.А., Соловьёв И.В., Цветков В.Я., Дубов С.С, и др. «Отчёт о научно-исследовательской работе по теме ««Разработка модели геоданных и способов её применения при построении единой инфраструктуры пространственных данных Российской Федерации»» (промежуточный отчёт (3-й этап)), МИИГАиК, 2010.
6. Майоров А.А., Соловьёв И.В., Кудж С.А. , Дубов С.С, и др. «Отчёт о научно-исследовательской работе по теме Разработка обобщенной открытой программной платформы для обработки и хранения пространственных данных» (промежуточный отчёт (3-й этап)), МИИГАиК, 2011.
7. «Современные задачи организации хранения планетарных данных Российской Федерации». Известия ВУЗов. Геодезия и аэрофотосъемка, №6, 2011.