

Инфраструктура пространственных данных для научных исследований

О.М. Атаева, А.В. Кошкарёв, А.А. Медведев, В.А. Серебряков, К.Б. Теймуразов

Аннотация

Обсуждаются планы создания Академической инфраструктуры пространственных данных как элемента единого научного пространства РАН. Приводятся результаты ее реализации, в том числе геопорталы будущей распределенной сети узлов «ГеоМета» ВЦ РАН (<http://www.geometa.ru>) и ИГ РАН (<http://asdi.igras.ru>), и проект по особо охраняемым природным территориям России на основе интегрированной прикладной схемы в соответствии со спецификацией Директивы INSPIRE Европейского союза.

Ключевые слова: пространственные данные, инфраструктура пространственных данных, геопортал, особо охраняемые природные территории, метаданные.

1. Введение

В середине 1990-х гг. в США, Канаде, Австралии и некоторых европейских странах были разработаны долгосрочные программы создания национальных инфраструктур пространственных данных (ИПД) как информационно-телекоммуникационных систем, объединяющих национальные ресурсы пространственных данных (геоинформационные ресурсы) и обеспечивающих доступ к ним. Это означало начало нового этапа развития геоинформатики как науки, технологии и производства. Современное состояние геоинформатики – это переход от эпохи ГИС к эпохе ИПД. На новом витке развития информационно-телекоммуникационных технологий речь идет уже о сетевых ГИС, обеспечивающих равный, простой и свободный доступ к геоинформационным ресурсам: данным, сервисам, приложениям, опыту, знаниям.

К настоящему времени в учреждениях РАН накоплен большой опыт использования геоинформационных технологий, реализованы многочисленные геоинформационные проекты, созданы базы и банки пространственных данных. Академические ресурсы пространственных данных составляют значительную часть национальных информационных ресурсов. Основным производителем пространственных данных являются учреждения геологического, геофизического, геохимического, географического и экологического (природоохранного) профиля. В то же время данные рассредоточены, их использование ограничено зачастую рамками того проекта, где они созданы, затруднены или невозможны поиск существующих данных и доступ к ним, не налажен обмен ими. Причина этого – отсутствие эффективной системы управления пространственными данными. Ее создание позволило бы интегрировать данные и знания о территории, строить и использовать модели природных и социально-экономических явлений и процессов, их взаимодействия в системе «общество – природная среда», использовать методы пространственного анализа, обеспечивать территориальное планирование и управление.

С этих позиций может быть сформулирована стратегия создания Академической ИПД (АИПД) [8, 9], общая цель которой – объединение распределенных ресурсов пространственных данных и результатов научных исследований, информации и знаний о Земле, ее природе, населении, хозяйстве и социокультурных аспектах территориальной организации общества, обеспечение свободного доступа к ним пользователей в среде Интернет и их многократного использования. С более общих позиций ее создание нацелено на повышение эффективности научной деятельности, рост инновационного потенциала науки, возможность внедрения ее результатов в практику и образовательные программы. Предметная область АИПД – науки о Земле, ее технологические основы – информационно-телекоммуникационные технологии, в частности, геоинформационные технологии, основная целевая аудитория – научно-образовательное сообщество.

Рассмотрим, как общие принципы и подходы к созданию ИПД и ее компонентов, заложенные в стратегии создания АИПД, могут быть адаптированы научно-образовательным сообществом для решения научных и прикладных задач, иллюстрируя ее примерами работ, выполняемых Вычислительным центром им. А.А. Дородницына РАН при участии Института географии РАН и других научных учреждений на платформе Единого научного информационного пространства РАН.

Стандартизация как обязательное условие обеспечения организационной, технологической и семантической интероперабельности данных и сервисов (геосервисов) в распределенной среде, что предполагает использование стандартов ИСО серии 19100 «Географическая информация/геоматика», спецификаций консорциума OGC Inc., подобно стандартам ИСО нацеленных на унификацию операций с пространственными данными, других международных и отечественных де-факто и де-юре стандартов в сфере геоинформатики и информационно-телекоммуникационных технологий в целом.

Система управления пространственными метаданными (геометаданными), которая позволяет их готовить, редактировать (в том числе обновлять), автономно или непосредственно на геопортале. Опыт построения такой системы иллюстрируется примером разработки академического профиля на содержание метаданных «ГеоМета», наследующего пакеты и элементы стандартов ИСО, включая ISO 19115:2003 Geographic information – Metadata [12], ISO 19119:2005 Geographic information – Services [13], и национальному стандарту ГОСТ Р 52573–2006 «Географическая информация. Метаданные» [5] – российскому профилю ISO 19115:2003.

Геопорталы как единые точки входа в распределенную среду пространственных данных, геосервисов и приложений. Геопортал поддерживает две основные функции: 1) поиск и отыскание пространственных данных по их метаданным в сетевой среде; 2) визуализацию пространственных данных, в том числе в форме карт [6]. К другим типовым функциям геопорталов относят сервисы загрузки данных их трансформации, вызов удаленных сервисов и механизмы электронной торговли.

Приводятся результаты реализации АИПД, в том числе геопорталы будущей распределенной сети узлов «ГеоМета» на примере ВЦ РАН (<http://www.geometa.ru>) и ИГ РАН (<http://asdi.igras.ru>), и проект по особо охраняемым природным территориям России.

2. Геопортал «ГеоМета»

2.1. Основные возможности системы

Геопортал «ГеоМета» – это стандартизированная и децентрализованная среда управления пространственной информацией, разработанная для доступа к пространственным данным, картографическим продуктам и связанным с ними метаданным из различных источников, облегчающая обмен пространственной информацией между организациями и ее совместное использование посредством интернета. Главная цель геопортала – увеличить доступность разнообразных междисциплинарных данных различного масштаба вместе с сопутствующей информацией, организованных и документированных стандартным и непротиворечивым способом, улучшить кооперацию и координацию усилий при сборе данных.

Геопортал «ГеоМета» представляет собой платформу для создания распределенной среды интеграции неоднородных источников геоинформационных данных и предоставления к этой среде единой точки входа (веб-портала), которая позволит ученым в сфере наук о Земле легко находить специализированные данные и приложения, производить вычислительные эксперименты, визуализировать результаты деятельности (рис. 1).

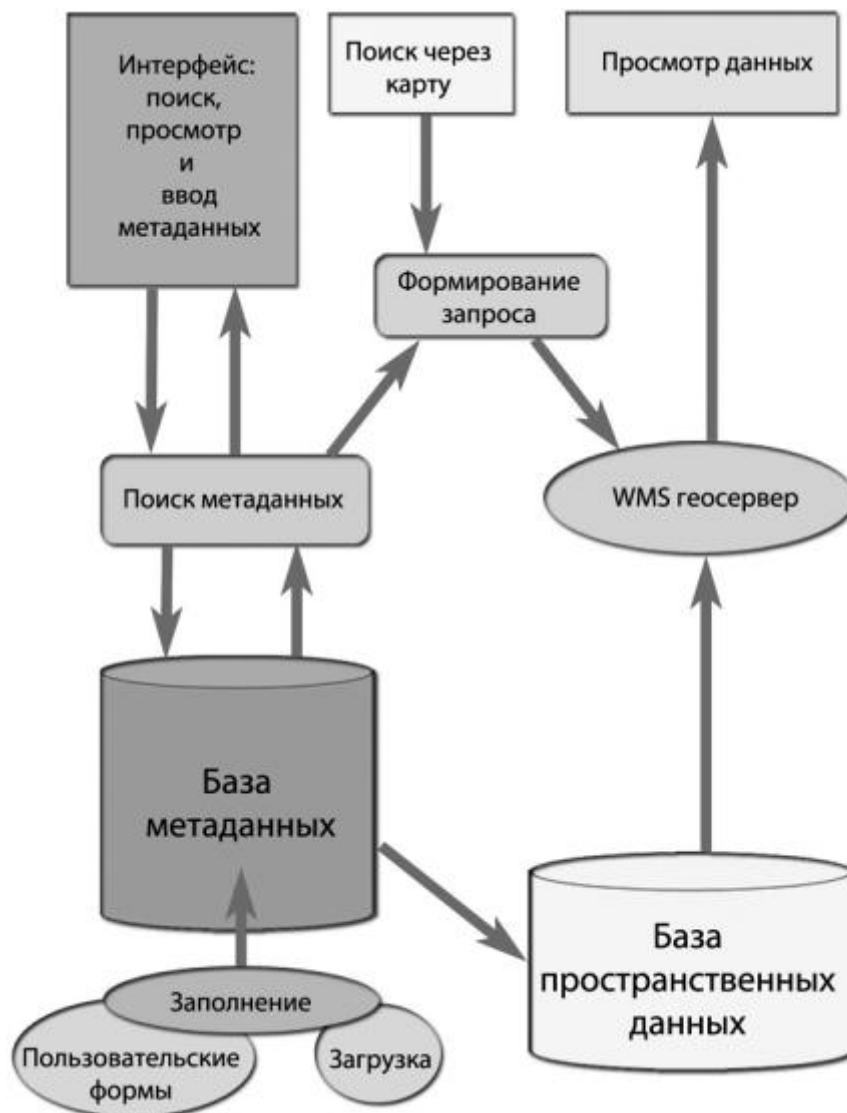


Рис. 1. Общая архитектура системы

Благодаря тому, что портал «ГеоМета» построен на базе ИС «НИ РАН» [2], являющейся базовым инфраструктурным компонентом ЕНИП [3], он может интегрироваться в ЕНИП с предоставлением расширения схемы пространственными данными и метаданными.

Функциональность системы включает:

- каталогизацию, сбор, поиск пространственных метаданных;
- размещение данных в собственном хранилище и предоставление к ним доступа;
- предоставление доступа к распределенным пространственным данным по стандартизованным протоколам;
- визуализацию карт, редактирование их элементов.

Интерфейс системы представлен веб-порталом, поэтому основным методом доступа пользователя к информации является обычный доступ к веб-страницам портала через любой распространенный браузер. Ядро системы предоставляет следующие возможности: управление статическим содержанием; хранение объектов системы (представленных RDF-тройками) в реляционных СУБД; индексирование и полнотекстовый поиск; обеспечение безопасности системы.

2.2. Пространственные данные и метаданные

Задача интеграции распределенных неоднородных источников пространственных данных и сервисов решается путем выделения базовой схемы метаданных, реализованной в виде OWL-онтологии [4] на основе стандартов ISO 19115:2003 и ISO 19139:2007 [14] и полностью включающей российский стандарт ГОСТ Р 52573-2006 (а также учитывающей профили стандартов и схемы метаданных, лежащие в

основе наиболее известных национальных ИПД), для обеспечения семантической интероперабельности систем, участвующих в распределенном взаимодействии.

3. Проект по особо охраняемым природным территориям

В соответствии с положениями Директивы INSPIRE [10], а также международной организации по стандартизации ИСО, в основу идеи ИПД положено информационное моделирование предметной области. Согласно стандарту ИСО 19101:2002 «Географическая информация – эталонная модель» [11] важным компонентом представления информационных и вычислительных уровней компьютерных систем является концептуальное моделирование, представляющее собой процесс создания абстрактных описаний части окружающего мира и относящихся к ней онтологических понятий. Иными словами, необходимо разработать концептуальную схему, которая определит, какие базовые характеристики, а также взаимосвязи между ними необходимо иметь особо охраняемым природным объектам для наиболее полного пространственного моделирования в системе ИПД.

С учетом Федерального закона «Об особо охраняемых природных территориях» построена схема данных, содержащая минимальный набор обязательных объектов и их характеристик (в виде классов, их атрибутов и связей), создаваемый для особо охраняемого природного объекта.

Анализ спецификации данных INSPIRE заключается в выборе наиболее подходящей прикладной схемы из трех предлагаемых.

Следующий этап концептуального моделирования – это интеграция полученных схем данных в единую прикладную схему. При этом базовой схемой считается прикладная схема INSPIRE как наиболее детально проработанная. Интеграция заключается в отображении объектов и их атрибутов схемы данных шейп-файлов и схемы пространственным данным обязательных элементов на наиболее релевантные объекты и атрибуты схемы данных INSPIRE. При отсутствии аналогичных объектов в схеме данных INSPIRE производится ее расширение путем введения дополнительных объектов и атрибутов. Общий процесс моделирования иллюстрируется рис. 2.

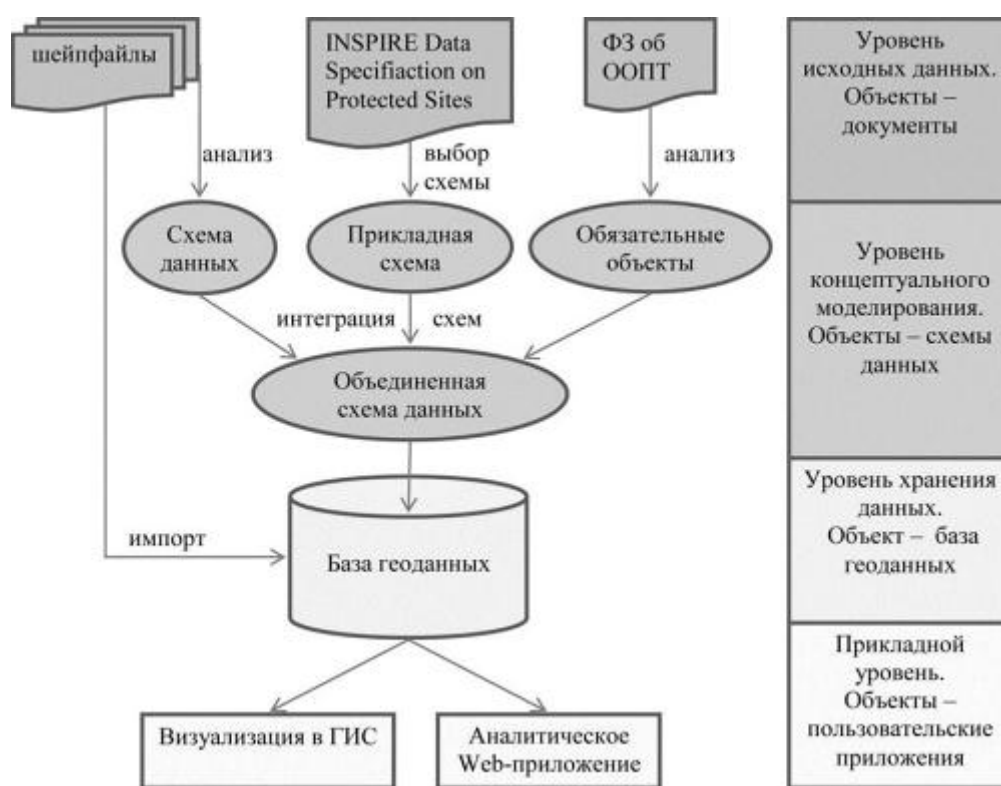


Рис. 2. Уровни моделирования предметной области «ООПТ»

При выполнении проекта была разработана прикладная схема базовых пространственных данных для особо охраняемых природных объектов. На основе базы геоданных было разработано сетевое аналитическое приложение для осуществления запросов и визуализации результатов запросов [7]. Аналитическое приложение на основе открытых серверных технологий, что позволяет легко интегрировать в различные веб-сайты.

На основе имеющихся наборов пространственных данных был создан портал «ИПД ООПТ России» (<http://www.ecoarea.ru>). На портале представлены метаданные пространственной информации по заказникам, национальным паркам, заповедникам России, которые были собраны в Институте географии РАН и на Географическом факультете Московского государственного университета им. М.В. Ломоносова. В портал интегрирован Штутгартский картографический сервер (Suas Mapserver) для визуализации пространственных данных, что позволило совмещать наборы данных портала с другими открытыми WMS-серверами.

Литература

- [1] Атаева О.М., Кузнецов К.А., Серебряков В.А., Филиппов В.И.. Портал интеграции пространственных данных «ГеоМета». Препринт ВЦ РАН, 2010 г. 106 стр.
- [2] Бездушный А.А., Бездушный А.Н., Нестеренко А.К., Серебряков В.А., Сысоев Т.М., Теймуразов К.Б., Филиппов В.И. Информационная Web-система «Научный институт» на платформе ЕНИП. М.: ВЦ РАН, 2007. 254 с.
- [3] Бездушный А.А., Бездушный А.Н., Серебряков В.А., Филиппов В.И. Интеграция метаданных Единого Научного Информационного Пространства РАН. М.: ВЦ РАН, 2006. 238 с.

- [4] Бездушный А.Н., Вершинин А.В., Дьяконов И.А., Динь Ле Дат, Серебряков В.А. Пространственные метаданные в системе «ГеоМета»// Пространственные данные. 2008. № 2. С. 16 – 25, 68; №3. С. 26 – 29.
- [5] ГОСТ Р 52573–2006 «Географическая информация. Метаданные». [Электронный ресурс]. — Режим доступа: <http://protect.gost.ru/document.aspx?control=7&id=128907>.
- [6] Кошкарёв А.В. Геопортал как инструмент управления пространственными данными и геосервисами // Пространственные данные. 2008. № 2. С. 6 – 14. (электронная версия: [Электронный ресурс]. — Режим доступа: <http://www.gisa.ru/45968.html>).
- [7] Кошкарёв А.В., Медведев А.А., Атаева О.М., Серебряков В.А., Теймуразов К.Б. Интеграция пространственных данных в распределенной академической инфраструктуре // Розвиток тематичної складової інфраструктури геопросторових даних в Україні: Зб. наук. праць. К., 2011. С. 84-90.
- [8] Кошкарёв А.В., Медведев А.А., Серебряков В.А. Интеграция пространственных данных в распределенной академической инфраструктуре // Труды Математического центра имени Н.И. Лобачевского: Лекционные материалы Девятой молодежной научной школы-конференции «Лобачевские чтения – 2010», Казань, 1 - 6 октября 2010 г.; Казан. матем. об-во. – Казань: Казань. матем. об-во, 2010. – Т. 42. С. 134-154.
- [9] Кошкарёв А.В., Ряховский В.М., Серебряков В.А. Инфраструктура распределенной среды хранения, поиска и преобразования пространственных данных. [Электронный ресурс]. — Режим доступа: http://seminar2010.fegi.ru/tezis/cat_view/7-.
- [10] Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE). [Электронный ресурс]. — Режим доступа: <http://eur-lex.europa.eu/JOHtml.do?uri=OJ:L:2007:108:SOM:EN:HTML>.
- [11] ISO 19101:2002 Geographic information – Reference model. [Электронный ресурс]. — Режим доступа: http://www.iso.org/iso/iso_catalogue/catalogue_ics/catalogue_detail_ics.htm?ics1=35&ics2=240&ics3=70&csnumber=26002.
- [12] ISO 19115:2003 «Geographic information – Metadata». [Электронный ресурс]. — Режим доступа: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=26020&commid=54904.
- [13] ISO 19119:2005: «Geographic information – Services». [Электронный ресурс]. — Режим доступа: http://www.iso.org/iso/catalogue_detail.htm?csnumber=39890.
- [14] ISO/TS 19139:2007. Geographic information – Metadata – XML schema implementation. [Электронный ресурс]. — Режим доступа: http://www.iso.org/iso/catalogue_detail.htm?csnumber=32557.

Об авторах

Атаева Ольга Муратовна - Вычислительный центр РАН e-mail: oli@ultimeta.ru

Кошкарёв Александр Владимирович - Институт географии РАН e-mail: akoshkarev@yandex.ru

Медведев Андрей Александрович - Институт географии РАН e-mail: a.a.medvedeff@gmail.com

Серебряков Владимир Алексеевич - Вычислительный центр РАН e-mail: serebr@ccas.ru

Теймуразов Кирилл Борисович - Вычислительный центр РАН e-mail: kbt@ccas.ru
