

Инфраструктура научных информационных ресурсов для непрерывного доступа к спутниковым данным исследования Земли из космоса*

Кудашев Е.Б.

*Данная статья публикуется на английском языке под названием "Science Data Infrastructure for Access to Earth Observation Satellite Data" и размещена целиком в разделе "English" (<http://www.elbib.ru/index.phtml?page=elbib/eng/journal/2012/part1>) .

Аннотация

В статье рассматриваются проблемы формирования e-Science Infrastructure для организации непрерывного доступа к спутниковым данным и обмена огромными потоками данных в связи с актуальной задачей интеграции информационных ресурсов. В России спутниковые данные рассредоточены в региональных спутниковых центрах данных. Их использование зачастую ограничено рамками конкретного проекта, в котором они были созданы. Чрезвычайно затруднены поиск ресурсов, доступ к данным и обмен данными. Многие ведомства и спутниковые центры пользуются собственными форматами представления данных, протоколами обмена, созданными достаточно давно и потому не полностью гармонизированными с современными международными правилами и стандартами. Одной из основных проблем, связанных с организацией хранения больших объемов данных Дистанционного Зондирования Земли (ДЗЗ) и непрерывного доступа к спутниковым ресурсам, является потребность интеграции различных поставщиков данных в единую информационную систему, позволяющую повысить эффективность научных исследований в области ДЗЗ. Предложены и обоснованы принципы и архитектура объединенной электронной e-Инфраструктуры непрерывного доступа к спутниковым данным. Описание геопространственных данных определено международными стандартами ISO 19115:2003 Geographic information – Metadata, ISO 19115-2:2005 Geographic information – Metadata. – Part 2: Extensions for imagery and gridded data. Существуют прототип FGDC-STD-001-1998 названных стандартов США и российский профиль стандарта ISO 19115 в виде документа ГОСТ Р 52573-2006 “Географическая информация. Метаданные.” Метаданные на текстовые материалы в сетях формируются с использованием стандарта Dublin Core. Единая политика администрирования e-Инфраструктуры обеспечивает согласованные между участниками правила доступа и использования ресурсов, унифицированные подходы к обеспечению информационной безопасности. Гарантией доступности и безопасности данных является проведение единой политики управления данными. Требование свободного информационного обмена реализуется путем задействования стандартизированных сетевых протоколов обмена данными и языков описания данных и построения запросов (прежде всего, SQL и XML). Интероперабельность обеспечивается глобальной унификацией технических регламентов, протоколов передачи и форматов данных, а также метаданных и поддерживается целой группой специально разрабатываемых стандартов (ISO/IEC 11179 Information technology — Specification and standardization of data elements; ISO 19113:2002 Geographic information — Quality principles; ISO 19114 Geographic information — Quality evaluation procedures и др.). Разработана архитектура e-Инфраструктуры непрерывного доступа к спутниковым данным, состоящая из компонент: 1) ГИС-сервер управляет геоинформационными ресурсами (карты, изображения, цифровые модели местности, текстовальная информация и т. д.) и транслирует их веб-приложениям в виде отдельных геосервисов; 2) Веб-сервер предназначен для хранения и управления геопорталом. Структурно и функционально Веб- и ГИС-сервера могут быть развернуты на отдельных компьютерах; 3) Геопортал - платформа для создания распределенной среды интеграции геоинформационных данных. Геопортал как единая точка входа в распределенную среду данных представляет из себя совокупность отдельных приложений и сервисов. Используется для публикации, администрирования и поиска стандартизированных геоинформационных ресурсов. Геопортал расположен на веб-сервере и является клиентским приложением ГИС-сервера. Пользователи, взаимодействуя с ним, вызывают сервисы, которые являются посредниками между веб-сервером и ГИС-сервером и обращаются к ГИС-серверу через его компонент SOM за необходимым инструментарием (геосервисами); 4) Сервисы каталогов - предназначены для доступа, поиска, накопления, обработки метаданных геопространственных ресурсов; сервисы обеспечивает общий механизм для классификации, регистрации, публикации, поиска, хранения и доступа к метаинформации об имеющихся в сети геоинформационных ресурсах. Сервисы каталога предоставляют метаданные о зарегистрированных геоинформационных ресурсах, поддерживают формирование запросов на информацию по типам ресурсов. Они позволяют публиковать, создавать и редактировать наборы метаданных для геопространственных данных, геопространственных сервисов и геоссылок; 5) Сервис данных - обеспечивает доступ к содержанию ресурсных информационных материалов, которые находятся в базах и банках геоданных. Каждой единице ресурсов, с которыми работает этот сервис, присваивается уникальное имя. Для повышения скорости поиска осуществляется индексация ресурсов; 6) База моделей тематических задач ДЗЗ - предназначена для хранения моделей тематических задач ДЗЗ и позволяет с помощью интерфейса программы-планировщика выбрать наиболее адекватную (релевантную) модель в соответствии с поставленной задачей; 7) Программное обеспечение как сервис (Software as a Service (SaaS)). SaaS подразумевает предоставление приложений для конечного пользователя в виде сервиса «по требованию» вместо его установки на конкретном рабочем месте или на собственном сервере. В рамках “облачных” вычислений существуют и другие направления: платформа как сервис и инфраструктура как сервис. 8) Инфраструктура как сервис (Infrastructure as a Service (IaaS)) - охватывает аппаратные средства и технологии для компьютерных вычислений и хранения данных, операционные системы и другую инфраструктуру, которые предоставляются не как локальные ресурсы, а опосредованно - через обращение к сервисам, размещенным на стороне провайдера. 9) Клиентские ГИС-приложения - устанавливаются на конечных устройствах локальных и удаленных пользователей и подключаются по протоколу HTTP к геопорталу по сетям LAN или WLAN для поиска, публикации и сохранения стандартизированных геоинформационных данных.

Ключевые слова: инфраструктура научных данных, электронная наука, данные наблюдения Земли из космоса, научная электронная инфраструктура, открытая инфраструктура данных, управление данными.

Работа выполнена при финансовой поддержке РФФИ (грант № 11-07-00006).

Об авторе

Кудашев Ефим Борисович - член-корреспондент Российской Инженерной Академии, д.т.н., ведущий научный сотрудник Института космических исследований Российской академии наук, профессор Механико-Математического Факультета МГУ им. М.В. Ломоносова,

