

Е. Б. Кудашев

Институт космических исследований РАН, Москва, Россия

**Электронная библиотека спутниковых данных:
доступ к коллекциям экологического
мониторинга**

Розглядаються проблеми створення єдиної системи автоматичної обробки інформації та розподіленого пошуку по всіх архівах космічної інформації. Описано структуру Міжнародної інформаційної системи INFEO, розробленої ESA, яка об'єднує національні архіви супутникових даних, надає користувачам спеціалізовані інтерфейси доступу до даних, забезпечує довготривале зберігання даних і дозволяє провадити пошук супутникових знімків і колекцій даних.

Формирование информационной инфраструктуры и интеграция электронных архивов спутниковых данных сейчас особенно актуальны, так как значительно вырос объем спутниковых данных дистанционного зондирования Земли. Возникает задача их систематизации, структуризации, создания средств каталогизации и поиска.

Единая система доступа ко всем существующим архивам позволяет эффективно использовать информацию, накопленную человечеством со времен развития космических технологий.

**ЭЛЕКТРОННАЯ БИБЛИОТЕКА КОСМИЧЕСКОЙ
ИНФОРМАЦИИ**

Тенденция развития электронных библиотек в задачах исследования Земли из космоса определяется тем, что электронные библиотеки представляют широкий класс распределенных информационных систем. Электронные библиотеки (ЭБ) — качественный скачок в области информационных технологий, связанный с развитием интернет-технологий и международных стандартов на программное обеспечение этих технологий.

Наиболее эффективными методами решения проблем доступа к данным, таким как разработка технологии централизованно-распределенного хранения данных; разработка схем метаданных с уче-

том международных и отечественных стандартов содержания и обмена данных; анализ корпоративных стандартов и согласование используемых стандартов хранения и представления информации и метаданных, являются информационные технологии, получившие название ЭБ.

Будем рассматривать ЭБ как область исследований и разработок, направленных на развитие методов обработки, хранения, поиска и анализа цифровых данных различной природы. Итак, во-первых, ЭБ — распределенные однородные коллекции цифровых объектов. Во-вторых, ЭБ — распределенное интегрированное пространство электронных ресурсов.

Методы и технологии ЭБ особенно важны для развития информационной поддержки космического мониторинга объектов природной среды и глобальных изменений климата из-за тенденции к интеграции национальных информационных ресурсов, создания единой системы автоматической обработки информации и распределенного поиска по всем национальным архивам космической информации.

Проект «Электронная библиотека спутниковых данных изменения окружающей среды и климата» развивается в Институте космических исследований РАН (ИКИ) [1, 3] как инфраструктура системы международного обмена космической информацией на основе технологий INFEO (Information about Earth Observation). Электронная библиотека

ИКИ предоставляет пользователям специализированные интерфейсы доступа к данным, обеспечивает долговременное хранение данных и позволяет производить поиск спутниковых снимков и коллекций данных.

Информационно-поисковая система построена на стандартах ISO для работы со спутниковыми данными и форматах электронного обмена массивами данных международных информационных систем распределенного поиска данных в области исследования Земли из космоса INFEO и EOSDIS (The Earth Observing System Data and Information System) для обмена спутниковыми данными. Система обеспечивает поиск по заданным атрибутам и поддерживает форматы электронного обмена массивами данных, а также унифицированный сетевой доступ к данным космического экологического мониторинга.

ИНДИВИДУАЛЬНЫЕ АРХИВЫ ДАННЫХ И ПРОБЛЕМА ИНТЕГРАЦИИ

Формирование информационной инфраструктуры и интеграция электронных архивов спутниковых данных особенно актуальны, так как в региональных Центрах спутникового мониторинга созданы и развиваются богатейшие электронные ресурсы. Наиболее яркими примерами таких информационных ресурсов являются спутниковые данные по лесным пожарам и динамике восстановления растительного покрова на территории России [1, 3, 4].

Созданные электронные коллекции включают данные по обширной территории России, начиная от Подмоскovie и центральных регионов страны, включая Западную и Восточную Сибирь, Дальний Восток и Приморье. Оперативный спутниковый мониторинг Тихого Океана привел к созданию электронной коллекции данных по развитию тайфунов и по диагностике параметров тайфунов; карты поверхностных температур моря и тепловых структур морской поверхности. Важное направление развития информационных ресурсов космического мониторинга связано с изучением состояния окружающей среды урбанизированных территорий, построением тепловых карт и обнаружением температурных аномалий мегаполисов. Примеры созданных в рамках проекта электронных коллекций спутниковых данных представлены в работах [1, 3, 4]; с информационными ресурсами можно также познакомиться на Web-сайте ИКИ РАН «Электронная библиотека космической информации» [<http://iris.iki.rssi.ru>].

С течением времени объем космической информации значительно возрастает, в связи с чем воз-

никает естественная потребность в систематизации спутниковых данных, их структуризации, создании средств каталогизации и поиска. Первым шагом на пути решения этой проблемы является выделение из всех свойств хранимых объектов только тех, которые необходимы в рамках решаемой задачи. Конечное число выделенных атрибутов позволяет произвести классификацию хранимых данных.

Бурное развитие компьютерных и интернет-технологий создает возможности объединения информационных ресурсов в единую информационную среду и организации эффективного хранения данных [2]. Для более продуктивного использования ресурсов архива данных создаются Web-сервера баз данных, что позволяет получать необходимую информацию о спутниковых данных через сеть интернет, обеспечивая наибольший охват информации. Для обеспечения потребителей адекватными и высококачественными данными, полученными в результате долговременных дистанционных исследований, разрабатываются наземные комплексы приема и обработки космической информации.

Структура системы ввода, архивации и распределения данных и функциональная схема инфраструктуры наземного комплекса обработки, хранения и распространения спутниковых данных, разработанные в Институте радиотехники и электроники РАН, были представлены нами ранее и анализировались в докладе на Первой Всероссийской конференции электронной библиотеки [1, С. 231, 235].

Существующие архивы спутниковых данных содержат, как правило, информацию по одному научному проекту. Огромное количество и разрозненность серверов баз данных (БД) осложняют поиск полной информации по интересующей тематике. Развитие серверов баз данных основано на технологиях управления метаданными информационных систем. Более низким уровнем представления информационных ресурсов в электронной библиотеке космической информации (см. рис. 1) является уровень, предназначенный для создания и поддержки справочника (Directory Level) спутниковых информационных ресурсов: IDN (International Directory Network), FGDC (Federal Geographic Data Committee) ISITE software, NGDF (UK). На этом уровне пользователь может лишь узнать, в каком архиве данных хранится необходимая ему информация, и вынужден, перейдя по ссылке в этот архив, уже в нем повторить процедуру поиска.

Более высоким является уровень каталога; в информационной системе электронной библиотеки космической информации этот уровень представления спутниковых данных обеспечивают INFEO и IMS EOSDIS (Information Manager System). Пре-

имущество такого подхода очевидно: конечная информация по запросу пользователя ищется уже непосредственно по всем источникам; отпадает необходимость в повторных запросах по каждому источнику.

МЕЖДУНАРОДНАЯ СИСТЕМА INFEO

В Международных программах исследования Земли из космоса перспективной задачей является создание единой системы автоматической обработки информации и распределенного поиска по всем национальным архивам космической информации. На пути реализации такой системы возникает множество проблем, связанных с разнородностью применяемого программного обеспечения для реализации хранилищ и организации их внутренней структуры, разнообразия интерфейсов, моделей данных.

Международная информационная система INFEO (Information of Earth Observation), разработанная Европейским космическим агентством (ESA) и ориентированная на работу со спутниковыми данными, объединяет архивы данных, расположенные в различных частях земного шара. Система INFEO, являясь активным многоцелевым архивом космических данных, предоставляет пользователям специализированные интерфейсы доступа к данным, обеспечивает долговременное хранение данных и позволяет производить поиск спутниковых снимков и коллекций данных.

Все пользователи могут получить доступ к ресурсам INFEO через единые узлы доступа (MWND). INFEO разработана для того, чтобы поддерживать различные пользовательские интерфейсы, т. е. WWW, Z39.50, CIP и другие специфичные интерфейсы.

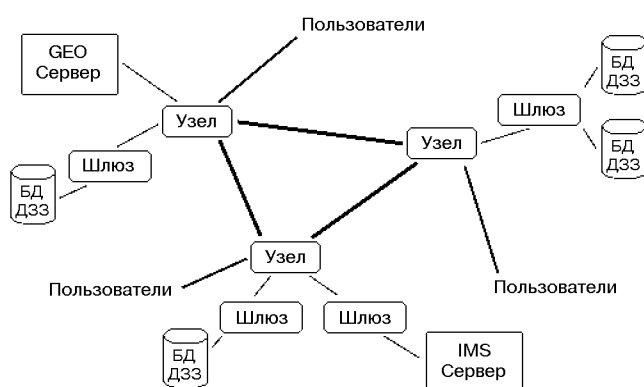


Рис. 1. Общая схема работы системы INFEO

Поставщики данных, как показано на рис. 1, имеют доступ к узлу MWND через ftp или Web для того, чтобы предоставить описание и дополнительную информацию о тех данных, которые у них имеются.

Рассмотрим INFEO Gateway — Шлюз поставщика данных. Существует большое количество возможных поставщиков для INFEO; некоторые из них могут уже иметь данные исследования земли из космоса в базе данных on-line. Поставщики, у которых лишь небольшое количество элементов, или те, кто не желает делать их доступными on-line, могут через INFEO сообщить клиентам о своей информации и предоставить им доступ через простой URL.

В системе INFEO для поставщиков с большими наборами данных предоставляется множество шлюзов. Это делается для того, чтобы минимизировать усилие, требуемое для того, чтобы соединиться с INFEO, и обеспечения механизма перевода между методом доступа к данным и INFEO стандартами.

Шлюз в информационной системе INFEO обеспечивает преобразование между протоколами INFEO, а именно Z39.50, CIP, WAIS и GILS, в формат, не зависящий от базы данных (то есть ODBC). Реализация такого механизма в INFEO возможна благодаря конфигурационным файлам.

Единые узлы доступа (MWND) обеспечивают доступ к ресурсам INFEO для поставщиков и клиентов. Узел информационной системы MWND решает следующие задачи.

- Описание коллекций данных — сюда также входит постоянный контроль за актуальностью информации.
- Библиографический поиск — поиск в библиотеках узла доступа и удаленных z39.50-совместимых узлах доступа.
- Поиск данных — поиск по локальным и удаленным CIP-коллекциям.
- Обслуживание пользователей — регистрация и обслуживание пользователей.
- Обслуживание словаря данных — управление словарем данных INFEO.
- Помощь — поддержка помощи об использовании INFEO.
- Мониторинг — ведение логов и общий контроль системы.

Наиболее важной частью системы INFEO являются шлюзы (см. рис. 2), состоящие из двух основных компонентов: брокера (Retrieval Manager) и переводчика (Translator). Брокер выполняет следующие функции:

- обработка запросов полученных от узла доступа (MWND);

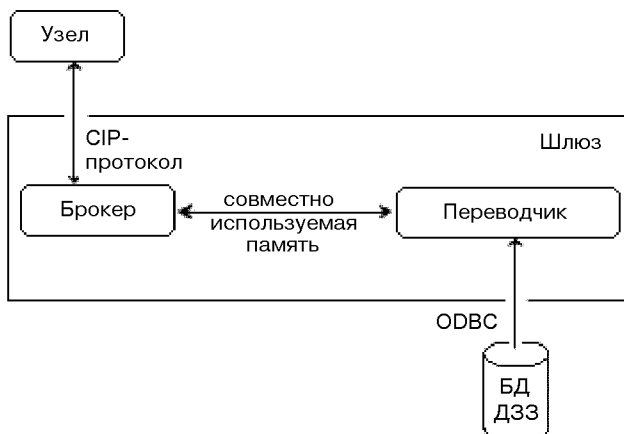


Рис. 2. Схема шлюза

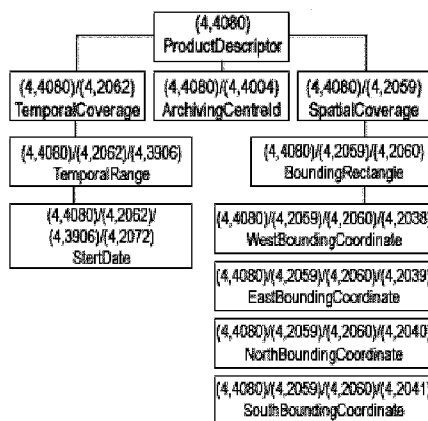


Рис. 3. Иерархическая структура атрибутов данных в протоколе СIP

- передача обработанного запроса переводчику;
- интерпретация ответа, полученного от переводчика;
- формирование и передача ответа узлу доступа.

Полученный от брокера запрос переводчик преобразует по правилам, описанным в конфигурационных файлах, в SQL запрос. Сформированный запрос передается через ODBC-драйвер серверу баз данных. Такая схема действий позволяет работать практически с любым сервером баз данных и любой

схемой данных. Полученный от сервера ответ переводчик передает брокеру.

СIP-ПРОТОКОЛ

Catalogue Interoperability Protocol основан на протоколе Z39.50 и является внутренним языком для работы системы INFEO. Все атрибуты данных в протоколе СIP находятся в иерархической взаимосвязи, показанной на рис. 3.

Например, параметру «Самая северная координата (NorthBoundingCoordinate)» соответствует номер (4.4080)/(4.2059)/(4.2060)/(4.2040). Глядя на этот номер, можно согласно схеме рис. 3 точно определить, свойства каких атрибутов унаследовал рассматриваемый параметр.

В нашем примере это «Прямоугольное ограничение (BoundingRectangle)» «Пространственное покрытие (SpatialCoverage)» «Атрибут данных (ProductDescriptor)».

1. Арманд Н. А., Кравцов Ю. А., Кудашев Е. Б. и др. On-line Электронная библиотека космической информации по экологическим и природо-ресурсным фундаментальным программам // Электронные библиотеки: Перспективные методы и технологии, Электронные коллекции: Тр. Первой Всерос. науч. конф. — С.-Петербург: Изд-во С.-ПБ Гос. ун-та, 1999.—С. 239—235.
2. Бездушный А. Н., Жижченко А. Б., Кулагин М. В., Серебряков В. А. Интегрированная система информационных ресурсов РАН и технология разработки электронных библиотек // Программирование.—2000.—26, № 4.—С. 177—185.
3. Мясников В. П., Арманд Н. А., Кравцов Ю. А. и др. Информационные технологии и информационные ресурсы космического экологического мониторинга // Вестник РФФИ.—2000.—2.—С. 30—37.
4. Kravtsov Yu. A., Kudashev E. B., Golomolzin V. V., Sharakhmanyan M. A. The satellite techniques in early warning systems for large towns and megacities // Early Warning Systems for Natural Disaster Reduction / Ed. J. Zschau and A. Kuppers. — Berlin, Heidelberg: Springer-Verlag, 2003.—Chapter 8.2.—P. 741—744.

AN ELECTRONIC LIBRARY OF SPACE DATA — THE ACCESS TO COLLECTIONS OF ECOLOGICAL MONITORING

E. B. Kudashev

The structure of the international information system INFEO is described.