

ИНФОРМАЦИОННЫЙ ОБМЕН В РАСПРЕДЕЛЕННОЙ ИНФОРМАЦИОННОЙ СИСТЕМЕ "МОЛЕКУЛЯРНАЯ СПЕКТРОСКОПИЯ"

Е.М. Козодоева

Институт оптики атмосферы имени В.Е. Зуева СО РАН

e-mail: klen@iao.ru

Аннотация

Рассматриваются разработанные подсистемы обмена метаданными и новостями между узлами распределённой информационной вычислительной системы "молекулярная спектроскопия".

Описываемая информационная система «Молекулярная спектроскопия» ориентирована на представление информационных ресурсов и вычислительных приложений предметной области, в которой данные измерений и вычислений, их обработка и классификация, являются основными задачами. Значения этих данных связаны с большим числом молекул, представляющих интерес для приложений, используемых в астрономии, радиационной физике, атмосферной оптике и других науках. Достоверность спектральных данных, поставляемых профессиональными спектроскопистами для таких приложений, имеет огромное значение.

Одновременно с развитием вычислительной техники в Институте оптики атмосферы СО РАН ведутся работы по созданию информационных систем для молекулярной спектроскопии. С появлением персональных компьютеров работы Головки В.Ф. и др. [1] привели к созданию информационных систем работающих на клиентском месте.

Развитие Интернет-технологий позволило сделать следующий шаг в развитии ИС по молекулярной спектроскопии [2-4]. Доступный в Интернет информационный ресурс (<http://spectra.iao.ru>) опирается на известные банки спектроскопических данных Hitran и Geisa. Большая часть решаемых в информационно-вычислительной системе SPECTRA задач связана с поиском параметров спектральной линии в базе данных.

Следующим этапом эволюции стало создание распределенной информационно-вычислительной системы (РИВС) "Молекулярная спектроскопия" (<http://saga.atmos.iao.ru>). Пользователь этой системы может внести собственные данные для проведения расчетов по алгоритмам, поддерживаемых ИВС, или их сравнения со спектроскопической информацией имеющейся в базе данных ИВС.

Одним из вариантов формирования пользовательского набора спектроскопических данных в информационной системе является их загрузка из опубликованных журнальных статей. Появляется необходимость ведения не только баз данных наборов пользовательских данных, но и создания библиографической базы данных. Формируемый набор данных обязательно привязывается к библиографической ссылке, это дает возможность последующего отслеживания происхождения данных, что является очень важным моментом для описания результатов расчетных задач, в которых использованы данные конкретного набора данных.

Загруженные данные доступны только самому пользователю и не видны остальным пользователям. Для того чтобы эта информация стала доступна широкой общественности

пользователю необходимо пройти процедуру публикации, состоящую в проверке данных экспертами, и присвоения им статуса «публичные». В нашей системе предполагается, что реплицированы с других узлов системы могут быть только «публичные» данные по решению администратора узла.

В разных городах на распределенных узлах информационной системы загрузка данных выполняется независимо, и возникает ситуация дублирования информации. Поскольку процесс загрузки является очень затратным по временным и человеческим ресурсам, возникает необходимость информирования пользователей распределенной системы о том, какие данные и из каких публикаций уже имеются в системе. Для этого в РИВС «Молекулярная спектроскопия» разработан механизм формирования метаданных описывающих имеющиеся на локальном узле источники данных [5]. Метаданные для опубликованных источников информации должны быть одинаковыми на всех узлах распределенной системы.

На каждом из узлов РИВС "Молекулярная спектроскопия" имеется система публикации новостей. Эти новости создаются автоматически при появлении публичного ресурса, или составляются вручную, например, в случае не программируемых ситуаций. Подробно система публикации новостей описана в работе [6].

Таким образом, поскольку работа по сбору, обработке и классификации данных измерений и вычислений, получаемых в разных группах специалистов, становится осмысленной только при их коллективной работе, необходимо средство обмена между узлами распределенной системы такой информацией как: опубликованные наборы данных, библиографические ссылки, метаданные "публичных" наборов данных и тематические новости.

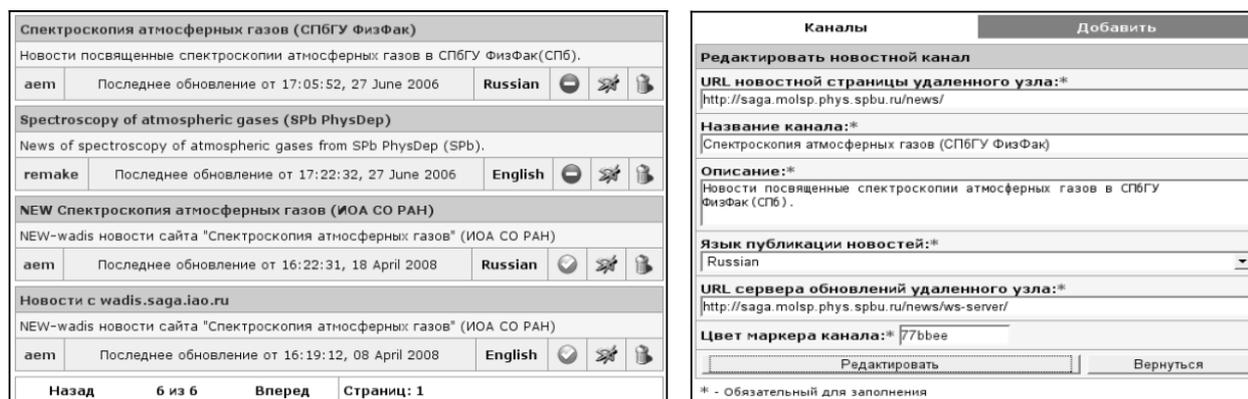
На данный момент разработаны и реализованы подсистемы обмена метаданными и новостями между узлами РИВС. Настройка этих подсистем идентична. Для каждого удаленного узла системы администратору локального узла необходимо создать информационный канал, задав такие параметры как: URL корневой страницы информационной системы на удаленном узле, название канала, описание, URL сервера обновлений удаленного узла (URL веб-сервиса), цвет маркера канала и периодичность запуска веб-сервиса.

Наша информационная система разработана с поддержкой интерфейсов на двух языках: русском и английском и в системе публикации новостей предусмотрена возможность создания новостей на двух языках, что явилось причиной задания для новостного канала такого параметра как язык публикации новостей. (см. рис.1)

Поскольку каждая группа специалистов работает с ограниченным набором задач спектроскопии [7], и собирает информацию об одном или нескольких веществах, то необходимо администратору узла дать возможность указывать область интересов его группы предметных специалистов, задавая с помощью представленного ниже интерфейса (см. рис. 2) список задач спектроскопии и список веществ, для которых необходимо получать метаданные.

Приведенные выше интерфейсы позволяют производить включение или отключение канала, если необходимо временно прервать поступление новых новостей или метаданных, получение метаданных вне графика и удаление или помещение в архив информационного канала. Последнее зависит от того существуют ли в базе данных метаданные полученные с

выбранного канала. Если да, то канал приобретает статус архивного для сохранения исторической информации о происхождении метаданных с ним связанных.



а)Список новостных каналов

б)Редактирование параметров канала.

Рис.1 Интерфейсы для просмотра, создания и редактирования каналов.

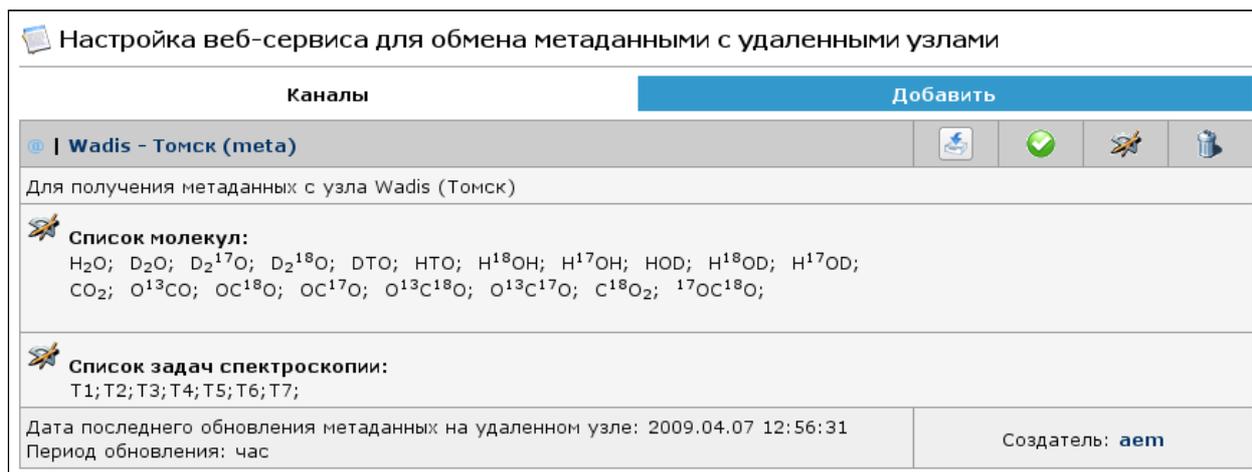


Рис. 2 Интерфейс настройки параметров веб-сервиса.

В отличие от метаданных новости не используются для построения базы знаний, и поэтому, по решению администратора узла, новостной канал может быть полностью удален вместе со связанными с ним новостями, либо только очищен от новостей, то есть новости связанные с этим каналом будут удалены, а настройки самого канала останутся.

Подсистемы обмена новостями и метаданными основаны на использовании технологии веб-сервисов [8, 9]. Веб-сервисы для обмена новостями и метаданными состоят из клиентской и серверной частей. Для запуска клиента веб-сервиса необходимо обратиться к веб-странице, на которой он расположен. На каждом узле распределенной системы настраивается системный процесс ("cron" в операционной системе Linux), который, по заданному ему расписанию, запускает на выполнение программу, скачивающую веб-страницу и добавляющую полученное содержимое страницы в конец файла-журнала. Если клиент отработал без ошибок, то возвращается пустая строка, в противном случае возвращается сообщение о возникших ошибках исполнения.

Клиентское приложение обращается к БД и получает список информационных каналов, с которых надо импортировать свежую информацию будь то новости или метаданные. Новизна и для метаданных и для новостей определяется полной датой их создания. Для каждого канала из сформированного списка выполняется вызов метода веб-сервиса. Параметрами вызываемого метода для сервиса обмена новостями являются язык, на котором опубликована новость, и время, начиная с которого возвращать новости. Параметрами метода веб-сервиса обмена метаданными являются: время, начиная с которого возвращать метаданные, список веществ и список задач спектроскопии.

Серверная часть веб-сервиса, получив запрос, извлекает из базы данных новости или метаданные соответственно с фактическим временем создания/изменения большими, чем входной временной параметр. Извлекаются только те новости и метаданные, которые были созданы на данном узле. Результат выборки возвращается клиентской части веб-сервиса.

Клиентское приложение, получив затребованную информацию, заносит ее в БД, устанавливая значение идентификатора информационного канала и значение времени последнего обновления канала меняет на максимальное значение времени среди записей полученной информации. Таким образом, в поле "время последнего обновления канала" хранится локальное время удаленного узла, что позволяет учесть разницу между часовыми поясами, в которых расположены узлы РИВС.

Практическое использование такой РИВС для коллективной работы с разделяемыми данными позволит решить следующие две задачи. С одной стороны, объединить имеющиеся в России данные с международными информационными ресурсами по спектроскопии, а с другой - стать местом сбора результатов научных проектов в этой предметной области. Нижний Новгород, Москва, Санкт Петербург и Томск будут являться основными узлами этой системы.

ЛИТЕРАТУРА:

1. V.F. Golovko, etc, Information system AIR-SENTRY for modeling atmospheric IR-spectra and radiation transmission in the atmosphere, ADBIS'95 Proc. The 2-nd Int. Workshop, v.2, Moscow, 1995, p.12-14.
2. Yu.L. Babikov , etc, WEB information system: atmospheric spectroscopy, Proc. SPIE "7-th Int. Symp. on Atmos. and Ocean Optics", v. 4341, 2000, p. 604-615.
3. Ю.Л. Бабикив и др., Интернет-коллекции по молекулярной спектроскопии, Сборник трудов 3 Всероссийской конференции по электронным библиотекам, Петрозаводск, 2001, с.183-187.
4. Ю.Л. Бабикив и др., Сайт SPECTRA - информационный ресурс по молекулярной спектроскопии, VII Международная конференция по электронным публикациям "EL-Pub2002".
5. <http://www.ict.nsc.ru/ws/elpub2002/4384>
6. А.В. Козодоев, А.И. Привезенцев, А.З. Фазлиев Аннотирование информационных ресурсов в распределенной информационной системе "Молекулярная спектроскопия", Труды Седьмой Всероссийской научной конференции RCDL-2005 «Электронные библиотеки: перспективные методы и технологии, электронные коллекции», 2005г., стр. 80-86
7. А.В. Козодоев, Е.М. Козодоева, А.И. Привезенцев, А.З. Фазлиев Система ввода и обмена данными в распределенной информационной системе "Молекулярная спектроскопия", Труды 13-й Всероссийской конференции "Информационные и математические технологии в науке и управлении", Иркутск, 2008г.

8. А.И. Привезенцев , А.З. Фазлиев Прикладная онтология задач для систематизации информационных ресурсов молекулярной спектроскопии, Труды 9^{ой} Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» - RCDL'2007, Переславль Залесский, ч.1, 2007, с.201-210.
9. А.Феррара, М.Мак-Дональд Программирование web-сервисов для .NET. Библиотека программиста. - Киев:ВНУ; СПб.: Питер, 2003, 430с.
10. Э.Ньюкомер Веб-сервисы. СПб.: Питер, 2003, 256с.