

WEB-СИСТЕМА УПРАВЛЕНИЯ ЗНАНИЯМИ ОБ ОКРУЖАЮЩЕЙ СРЕДЕ*

Е. П. ГОРДОВ

*Институт мониторинга климатических
и экологических систем СО РАН, Томск, Россия*

e-mail: gordov@scert.ru

С. П. КОВАЛЕВ, Ю. И. МОЛОРОДОВ, А. М. ФЕДОТОВ

Институт вычислительных технологий СО РАН, Новосибирск, Россия

e-mail: kovalyov@nsc.ru, yumo@ict.nsc.ru, fedotov@sbras.ru

Needs of the modern society require principally novel types of the informational systems — informational and informational-computational systems including the distributed ones. Important task is to establish facilities for access and effective use of these resources. Nowadays, it is particularly actual to use these resources for interdisciplinary projects in environmental sciences.

Введение

Обеспечение использования научных и информационных ресурсов мирового научного сообщества, распространение собственных достижений в виде порталов, атласов и информационных систем, создание и организация доступа к ним являются важнейшими задачами информационной поддержки исследований окружающей среды. Актуальной задачей является также создание средств доступа и технологий распределенного использования вычислительных и информационных ресурсов.

Развитие математических и вычислительных моделей для исследования окружающей среды показало явную недостаточность существующих данных о ее состоянии, что стимулировало интенсивное накопление временных рядов пространственно-распределенных данных инструментальных наблюдений. Эти обстоятельства предопределили особую роль создания информационных технологий в этой области — Web-систем, используемых для сбора, анализа и управления знаниями об окружающей среде. Для этого необходимо провести теоретические исследования и разработать принципы создания информационных систем на основе архитектуры распределенной сети баз данных. Нужно разработать сервисные программы, метаданные и структуру данных для эффективной обработки, хранения, поиска, представления и визуального наблюдения данных об окружающей среде.

*Работа выполнена при частичной поддержке Президентской программы (грант № НШ 2314.2003.1) и интеграционного проекта СО РАН № 169.

© Институт вычислительных технологий Сибирского отделения Российской академии наук, 2005.

На климатические процессы в окружающей среде существенно влияет в том числе и крупномасштабный перенос атмосферных загрязнений, формируя ухудшение экологических условий обширных регионов. Для получения достоверных данных о пространственно-временных изменениях окружающей среды необходимо провести комплексные мультидисциплинарные исследования, дополненные проведением мониторинга окружающей среды, разработкой аналитических моделей, проведением математического моделирования и лабораторно-стендовыми исследованиями. Это позволит оценить влияние атмосферных аэрозолей на качество воздуха, уровни загрязнения растительности, почвы и воды, скорости миграции различных веществ и элементов в объектах окружающей среды и выяснить степень воздействия атмосферных аэрозолей различной природы на атмосферные процессы, здоровье людей и животный мир.

На сегодняшний день существует ряд информационных систем, предназначенных для решения различных задач из области наук об окружающей среде и экологии: HYSPLIT (<http://www.arl.noaa.gov/ready/hysplit4.html>) — обширная база данных наблюдений, позволяющая определить перенос воздушных масс за определенный период времени; GISS (http://www.giss.nasa.gov/data/rsp_air/) комбинирует и сравнивает результаты набора данных космических наблюдений с глобальными моделями атмосферы, земной поверхности и океанических процессов; PANGAEA (www.pangea.de) — библиотека данных и публикаций в области геофизики; мультимедийный информационный ресурс, открывающий интерактивный доступ к собранной информации по оценке влияния атмосферных примесей на окружающую среду (<http://www.ess.co.at/AIR-EIA>). Аналогичным примером такой системы можно назвать Российский научный портал для атмосферных наук ATMOS (<http://atmos.iao.ru>), который представляет собой интегрированный набор множества распределенных, но координируемых предметных сайтов, содержащих типовую мультимедийную информацию с исследовательскими базами данных, моделями и аналитическим инструментарием для прямого использования и визуализации [1, 2].

В настоящее время в сети Интернет можно найти информационные ресурсы, содержащие обширный экспериментальный материал о физических характеристиках окружающей среды. С их помощью можно провести верификацию моделей и, используя современные глобальные климатические модели, провести численные эксперименты. Например:

— ECMWF reanalysis ERA-15 (1979–1993 годы), ERA-40 (1957–2001 годы)

<http://www.ecmwf.int/research/era>;

— NCEP/NCAR (National Center of Environment Protection/National Center of Atmospheric Research, USA), 1958–1997 годы

<http://wesley.wwb.noaa.gov/reanalysis.htm>;

— Precipitation from 1979 to present time

<http://www.cpc.ncep.noaa.gov/products/globalprecip>;

— Archive NDP048, containing multiyear data of routine observations on 225 meteorological stations of the former USSR

<http://cdiac.esd.ornl.gov/ftp/ndp048>.

1. Способы получения данных об окружающей среде

С 1991 года специалисты институтов СО РАН (Институт химической кинетики и горения, Институт неорганической химии, Лимнологический институт, Институт водных и экологических проблем), Кемеровского, Красноярского и Томского научных центров проводили регулярные измерения временных характеристик атмосферных аэрозолей (АА). Монито-

ринг проводился на территории Западной и Восточной Сибири, Алтайского и Красноярского краев.

Специалисты ИХКиГ СО РАН производили отбор проб и наблюдения на 10 станциях: Белое море — корабль, п-ов Картеш; о. Мудьюг — Архангельская область; Завьялово, Карасук, Новосибирск, Ключи — Новосибирская область; Красноселькуп, Самбург, Тарко-Сале — Тюменская область.

В рамках совместных исследований с Сибирским отделением РАН ИКИ РАН (г. Москва) проводит сбор данных со станций приема космической информации на основе низкоорбитальных спутников серии NOAA, Метеор с многоканальной аппаратурой низкого, среднего и высокого разрешения, представляющих собой спутниковые снимки различных участков Земли. При этом используются возможности бортовых радиометров систем наблюдения Earth Observing System (EOS) на MODIS/Terra, MODIS/Aqua, AMSR-E/Aqua платформах, осуществляющих измерения атмосферных и поверхностных характеристик, а также и изменения свойств подстилающей поверхности.

1.1. Измерения характеристик аэрозоля

К физическим свойствам атмосферного аэрозоля, рассматриваемым в данной работе, относятся счетная и массовая концентрация аэрозольных частиц, спектр размеров частиц и полное аэрозольное светорассеяние, измеряемое нефелометрией. Для измерений полного спектра размеров атмосферного аэрозоля обычно используются несколько приборов, чтобы перекрыть весь диапазон размеров аэрозольных частиц. Для оригинальных измерений, результаты которых представлены в данной работе, использовался следующий набор инструментов. Для диапазона размеров менее 0.1 мкм по радиусу использовались приборы фирмы TSI (TSI GmbH, Ziegelstrasse 1, 5100 Aachen, Germany): счетчик ядер конденсации (СЯК), модель 3020, сетчатая диффузионная батарея (СДБ), модель 3040. Для отбора частиц с последующим определением счетной концентрации и спектра размеров в диапазоне радиусов 0.1–100 мкм использовались одно- и двухступенчатые импакторы и импактор открытого типа. Импакторы различных конструкций и фильтры использовались для определения массовой концентрации и отбора проб для последующего анализа на элементный и химический состав. Нефелометрические измерения по определению аэрозольного светорассеяния проводились с помощью промышленного нефелометра ФАН-А. Для измерения концентрации субмикронных и околомикронных или так называемых оптических частиц использовались оптические счетчики аэрозоля. Для калибровки использовался промышленный счетчик ПКЗВ-6.

По информации, полученной с космических аппаратов, находились характеристики отражения света от земной поверхности — альbedo, равные отношению отраженного потока к падающему. Данные со спутников позволяют выявить влияние АА на способность поверхности отражать падающий на нее поток электромагнитного излучения или частиц.

Проводимый комплекс измерений позволил определить следующие физические свойства атмосферных аэрозолей: массовую концентрацию, счетную концентрацию, химический состав, полное аэрозольное светорассеяние, концентрация ядер конденсации ($r \leq 0.1$ мкм), форму частиц, спектр размеров частиц ($0.1 \leq r \leq 0.4$ мкм), спектр размеров частиц ($0.4 \leq r \leq 10.0$ мкм), коэффициент преломления вещества частиц.

Этот огромный объем экспериментальных данных неизмеримо расширил и во многом изменил существующие представления о техногенном влиянии индустриальных центров на окружающую среду региона.

На основе выполненных наблюдений в соответствующих районах юга и севера Западной Сибири установлена возможность идентификации источников АА по измерению пространственно-временной изменчивости их химического состава в различных регионах Западной Сибири. В частности, отношение содержания органического и неорганического углерода примерно постоянно независимо от пунктов наблюдения: Карасук — фоновая территория, Ключи — пригородная зона г. Новосибирска, Новосибирск — промышленный центр. Поэтому этот компонент может служить трассером для определения источника. Появление массовой концентрации этого трассера и образование траектории воздушных масс указывают на возможные источники регионального и глобального масштаба.

2. Структура атласа “Атмосферные аэрозоли Сибири”

До недавнего времени вся эта эмпирическая информация хранилась в бумажном виде: в виде записей в рабочих журналах, на лентах самописцев и фотошленках, лабораторных отчетах и статьях. К сожалению, весь этот ценный эмпирический материал был доступен только его владельцу и, как правило, терялся с уходом хозяина.

Повышение эффективности использования огромного объема эмпирической информации может достигаться за счет разработки, развития и использования информационных технологий, на основе разработки как новых специализированных моделей, ориентированных на решение относительно узких классов задач, так и путем разработки технологий организации процесса хранения и обработки данных с использованием существующих методов и моделей.

В этих условиях крайне необходимо выработать и последовательно реализовать единую информационную политику, обеспечивающую развитие и эффективное использование современных информационных технологий и информационных ресурсов. Важной задачей в области повышения эффективности использования региональных информационных ресурсов является формирование единой информационной среды, которая должна включать два основных компонента. Первый обеспечит методологическую базу хранения, обработку и анализ. Это модели систем, процессов и задач предметной области. Второй представляет собой технологическую базу моделирования в виде распределенной интегрированной информационно-вычислительной среды, позволяющей реализовать эти модели и разрешить работу с ними достаточно широкому кругу пользователей [3, 4].

Для решения поставленной задачи необходимо, во-первых, разработать подходы и технологии, обеспечивающие виртуальную интеграцию описаний разнородных информационных ресурсов, расположенных на серверах различных организаций, в единую базу данных на основе открытых международных стандартов, во-вторых, создать систему базовых понятий для описания электронных коллекций атмосферных аэрозолей, в-третьих, построить информационные модели описаний микрофизических и химических характеристик на основе результатов измерения характеристик атмосферных аэрозолей различной природы. Для этого необходимо провести анализ технологий использования, хранения и обработки данных, построить модели связей биогеохимических циклов с факторами физической среды, обосновать информационную безопасность взаимодействия пользователей: конфиденциальность, целостность и доступность. Современные Интернет-технологии позволяют по-новому организовывать хранилища данных и доступ к ним [2].

Анализ существующего состояния сбора, хранения и обработки результатов измерений характеристик атмосферных аэрозолей, а также разработки и внедрения информаци-

онных технологий и средств вычислительной техники в природоохранную деятельность показывает, что различными организациями СО РАН созданы и создаются собственные информационные компьютерные системы. Они разрабатываются, внедряются и функционируют, как правило, для себя, разобщенно, без координации и согласования с кем-либо структуры, содержания способов и форматов хранения информации и доступа к ней. Это приводит к дублированию работ, избыточности первичной информации, многократному ее вводу, приводящему к нарушению ее целостности и низкой достоверности, значительному удорожанию разработок и эксплуатации систем. Ведомственная разобщенность затрудняет обмен информацией и доступ к ней, что осложняет информационное обеспечение заинтересованных лиц.

Специалисты Института вычислительных технологий и Института химической кинетики и горения СО РАН разработали информационную модель и структуру метаданных, используя обобщенный подход для формирования и заполнения входных данных, включая их унификацию и связи для создания новой информационно-вычислительной системы, — атлас “Атмосферные аэрозоли Сибири” (<http://web.ict.nsc.ru/aerosol/>). В нем аккумулируется весь накопленный до настоящего времени эмпирический материал, обеспечивая решения различных задач в области сбора, обработки и публикации сведений об атмосферных аэрозолях. Важной особенностью созданной информационной системы является коллективная работа с архивами данных.

Атлас решает различные задачи в области сбора, обработки и публикации сведений об атмосферных аэрозолях, такие как:

- добавление, хранение, редактирование исходных данных об атмосферных аэрозолях, включая процедуры авторизации пользователей и ограничение доступа к различным информационным структурам;
- реализация функций математической обработки данных о химическом составе аэрозолей: статистические характеристики, корреляционный и факторный анализ временных рядов и другие функции;
- разработка Web-интерфейса для отображения исходных данных и результатов обработки, включая табличное и графическое представление;
- в Web-интерфейсе и на уровне данных реализация поддержки русского и английского языков.

3. Данные об аэрозолях

Атмосферный аэрозоль является сложной структурой, состоящей из различных химических элементов и их соединений. Сбор исходного материала АА осуществляется на стационарных точках наблюдений (станциях) с периодичностью несколько раз в год. Полученные пробы разделяют на элементные, ионные и углеродные фракции, определяют их количественное и качественное содержание.

Элементный состав аэрозоля включает в себя отдельные химические элементы, такие как, например, Са, Вг, К и т.д. (в настоящее время имеющаяся аппаратура способна выделить более 60 элементов). В ионном составе выделяют катионы (NH_4^+ , K^+ , H^+ , Na^+) и анионы (HCO_3^- , Cl^- , F^- , NO_3^-), так как только эти химические структуры определяемы на современном оборудовании. Для углеродного состава определяют содержание органического и неорганического углерода на фильтре, которым берут пробы. Система вычисляет и другие значимые углеродные характеристики: суммарную относительную концентрацию

и соотношение органического и неорганического углеродов.

Наиболее важными для дальнейшей обработки и моделирования являются полная массовая концентрация покомпонентного состава и концентрация грубодисперсной фракции, а также субмикронная массовая концентрация и полная счетная концентрация аэрозолей, особенно отличающиеся нерегулярностью измерений. Физико-химический анализ проб осуществляется в научно-исследовательских институтах (Институте химической кинетики и горения, Институте ядерной физики и Институте неорганической химии СО РАН), после чего все полученные данные экспортируются в хранилище атласа.

3.1. Содержание таблиц

База данных атласа “Атмосферные аэрозоли Сибири” содержит таблицы с результатами анализа проб приземных аэрозолей, собранных на территории Западной и Восточной Сибири, на границе с Казахстаном и в Белом море.

В таблицы внесены как исходный эмпирический материал, так и результаты простейшей математической обработки:

- 1) время, место и длительность отбора конкретного элемента;
- 2) среднее геометрическое значение и отклонение концентрации элемента;
- 3) среднее арифметическое значение массы элемента;
- 4) минимальное, максимальное и их отношение значения массы;
- 5) коэффициент Рана (environment factor), например, Ca/Fe;
- 6) коэффициент обогащения (enrichment factor).

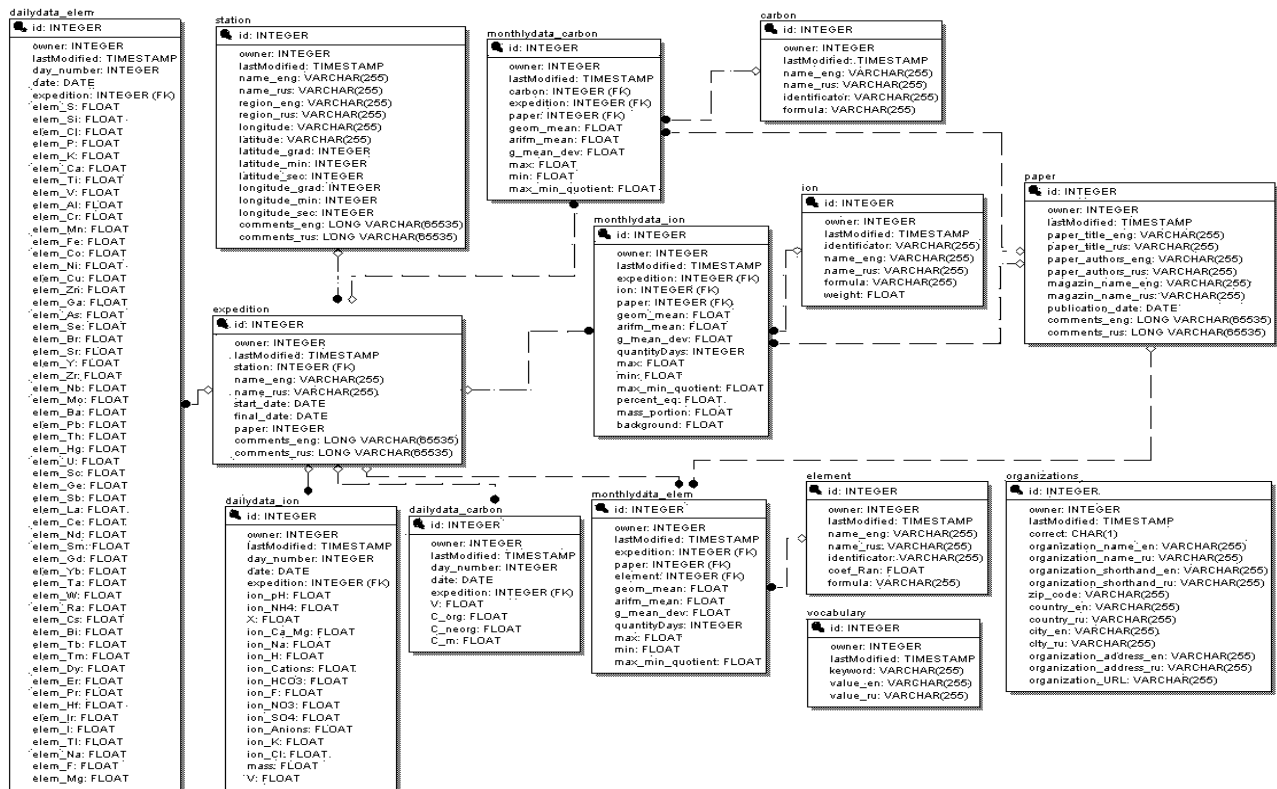


Схема данных в виде E – R-диаграммы.

Для построения атласа использовался программный продукт SMART [5]. Структура атласа представляет собой набор документов, входящих в коллекции. Каждая из коллекций изначально представлена в виде RDF-описания. На выходе по запросу клиента формируется документ в HTML-формате, при этом стилевое оформление документов физически отделено от их содержательной части. При этом хранимая информация представляет собой набор классов, являющихся объединением объектов. Например, класс экспедиций, станций наблюдения, статей. Классы, в свою очередь, взаимосвязаны. На E — R-диаграмме (Entity — Relationship) представлены схема данных информационной системы “Аэрозоли Сибири” и состав таблиц, используемые для хранения эмпирических данных (см. рисунок).

4. Обработка данных

Информационная система предоставляет не только средства хранения данных, визуализацию их зависимостей от времени в табличном и графическом виде, но и возможность обработки данных специализированными алгоритмами для решения различных задач исследования окружающей среды. При этом были выделены несколько классов задач, каждый из которых порождает собственные модели хранения и обработки данных.

4.1. Статистическая обработка данных

Для интерпретации полученных данных используются различные статистические методы: корреляционный, факторный, дискриминантный, кластерный и фурье-анализы временных рядов. К настоящему времени в информационно-вычислительной системе реализованы следующие функции:

- подсчет наиболее значимых статистических характеристик покомпонентного химического состава аэрозолей: коэффициентов обогащения, среднего арифметического и геометрического, геометрического отклонения;
- корреляционный анализ, показывающий взаимосвязь химических элементов и их влияние друг на друга;
- факторный анализ временных рядов, основанный на корреляционном анализе и позволяющий выделить ряд факторов, которые обуславливают присутствие тех или иных химических соединений в составе аэрозоля.

4.2. Теория Ми

Мониторинг состояния и поведения аэрозоля осуществляется, как правило, оптическими методами, поэтому необходима информация об оптических свойствах и характеристиках аэрозоля в зависимости от углов наблюдения. Сечения поглощения и рассеяния, относительное светорассеяние аэрозольными частицами сферической формы в зависимости от размера и функции, описывающие зависимость коэффициентов рассеяния от угла, могут быть рассчитаны на основе теории Ми [6]. Эта теория широко используется, а в ряде случаев и для несферических частиц, для интерпретации экспериментальных данных об оптических свойствах атмосферных аэрозолей: рассеяния, поглощения и ослабления солнечного излучения аэрозольными частицами в атмосфере, что влияет на видимость, радиационный обмен и пр.

В качестве масштаба за единицу бралось рэлеевское рассеяние света чистым воздухом без аэрозольных частиц. Расчеты проводились с использованием программы ВНМІЕ [6]. Результаты этих расчетов использовались для определения эквивалентного оптического диаметра и спектра размеров при измерении оптических характеристик: светорассеяния для нефелометрии, ослабления, обратного рассеяния при лидарном зондировании.

Заключение

Функционирующая система обеспечивает проведение комплексного мониторинга газоаэрозольных эмиссий в районах Западной и Восточной Сибири и работу с данными комплексного мониторинга различных территориально разобобщенных групп исследователей, доступ и проведение совместного анализа данных наземных измерений массовой концентрации и химического состава аэрозольных частиц и газообразных примесей, спектра размеров аэрозолей в диапазоне от 10^{-3} до 10^2 мкм, морфологии частиц и их пространственно-временных изменений в районах наблюдений Новосибирской, Томской, Кемеровской областей, Алтайского и Красноярского краев, в районе оз. Байкал, Пуровском и Красноселькупском районах Ямало-Ненецкого автономного округа и спутниковых измерений (данных, поступающих со спутников NOAA, Meteor-ЗМ, Ресурс, Terra, Aqua, Spot).

Разрабатываемая Web-структура позволяет не только формировать систему знаний об окружающей среде на основе обширного эмпирического материала, но и разрабатывать средства управления этими знаниями.

Список литературы

- [1] Гордов Е.П., DE RUDDER A., ФАЗЛИЕВ А.З., FEDRA K. Веб-портал Атмос как основа для выполнения интегрированных исследований по окружающей среде Сибири // Вычисл. технологии. 2004. Т. 9, ч. 2. С. 3–13.
- [2] ФАЗЛИЕВ А.З. Информационные ресурсы и Интернет-технологии для наук об окружающей среде // Вычисл. технологии. 2004. Т. 9, ч. 1. С. 11–20.
- [3] КУЦЕНОГИЙ К.П., КУЦЕНОГИЙ П.К., МОЛОРОДОВ Ю.И., ФЕДОТОВ А.М. Разработка структуры метаданных по атмосферным аэрозолям на основе информационной модели // Тр. Междунар. конф. “Математические методы в геофизике”. Ч. II. Новосибирск, 2003. С. 438–444.
- [4] МОЛОРОДОВ Ю.И., КУЦЕНОГИЙ П.К., СУДАРИКОВА И.А. Содержательное наполнение атласа “Атмосферные аэрозоли Сибири” // География и природные ресурсы. Спецвыпуск. Новосибирск: Изд-во СО РАН, филиал ГЕО, 2004. С. 86–90.
- [5] ШРАЙБМАН В.Б., ГУСЬКОВ А.Е. Разработка информационных систем на основе RDF-технологии // Тр. ХLI Междунар. науч. студенческой конф. “Студент и научно-технический прогресс”. Новосибирск: НГУ, 2003. Ч. 1. С. 153.
- [6] БОРЕН К., ХАФМЕН Д. Поглощение и рассеяние света малыми частицами. М.: Мир, 1986. 660 С.

Поступила в редакцию 2 июня 2005 г.