

# ИНФОРМАЦИОННО-СПРАВОЧНАЯ СИСТЕМА СО РАН\*

А. М. ФЕДОТОВ, В. Б. БАРАХНИН, А. Е. ГУСЬКОВ, О. Л. ЖИЖИМОВ,  
О. А. КЛИМЕНКО, Ю. В. ЛЕОНОВА, Е. В. РЫЧКОВА,  
Ж. С. ТУНДУКПАЕВ, И. В. ШАБАЛЬНИКОВ

*Институт вычислительных технологий СО РАН, Новосибирск, Россия*  
e-mail: klimenko@ict.nsc.ru

This article is devoted to the description of the main principles for construction of the Information System of the Siberian Branch of the Russian Academy of Sciences. Dynamic system of publications [1] and the collection model for construction of the information systems [2] set the basis of the presented Information System.

## Введение

Создание корпоративных информационно-справочных систем поддержки научных исследований, формирования и реализации управленческих решений для крупного территориально распределенного научного сообщества, каким является Сибирское отделение РАН (СО РАН), это один из наиболее действенных способов интеграции научных коллективов. Основной задачей таких систем является интеграция разнородной информации, так или иначе относящейся к деятельности СО РАН. При создании системы решаются задачи обеспечения доступа (в том числе и удаленного) пользователей к разнородным информационным ресурсам, защиты авторских прав на электронные документы, систематизации большого объема разнородных типов документов. Вследствие территориальной распределенности Сибирского отделения РАН для реализации эффективного функционирования таких систем наиболее остро стоит проблема получения и использования оперативной информации.

Интеграция разнородных информационных ресурсов в единую среду и организация доступа к ним — одни из важнейших направлений развития современных информационных технологий. При этом наблюдается переход к исключительно распределенной схеме создания, поддержания, хранения ресурсов<sup>1</sup> и в то же время стремление к виртуальному единству посредством предоставления свободного доступа к любым ресурсам сети через ограниченное число “точек доступа”.

---

\*Работа выполнена при частичной финансовой поддержке Российского фонда фундаментальных исследований (гранты № 06-07-89060, № 06-07-89038), президентской программы “Ведущие научные школы РФ” (грант № НШ-9886.2006.9) и интеграционных проектов СО РАН.

<sup>1</sup>Эффективная эксплуатация информационных ресурсов возможна только в случае, если они постоянно поддерживаются авторами, т. е. на основе технологий использования распределенных информационно-вычислительных ресурсов, которые получили название GRID-технологий.

© Институт вычислительных технологий Сибирского отделения Российской академии наук, 2006.

Доступ к информационно-справочной системе СО РАН реализован через информационный сервер (портал) СО РАН [3]. В ее основу положена коллекционная модель информационной системы, позволяющая достаточно прозрачно унифицировать работу с документами [1, 2]. Организационно-технологическая часть системы построена на принципе использования электронных библиотек. Электронные библиотеки — это набор коллекций и правила организации и каталогизации (построения метаописаний) информационных ресурсов. А основные задачи, решаемые электронными библиотеками, — “интеграция информационных ресурсов (включая поддержку унифицированного доступа к ним) и эффективная навигация в них” [4].

**Информационное наполнение системы.** Информационная система представляет собой набор связанных между собой коллекций, имеющих общие правила каталогизации. Под коллекцией мы понимаем множество документов одинаковой тематической направленности с выделенной фиксированной структурой.

**База данных организаций и сотрудников.** Основой системы является набор коллекций под общим названием “База данных организаций и сотрудников СО РАН” (БД “СО РАН”) [5, 6], куда входят:

- коллекция “Персоны”;
- коллекция “Организации”;
- коллекция “Неструктурированные документы”, содержащая дополнительную информацию как по персонам, так и по организациям;
- “каталог” (публикаций, проектов и т. д.).

Стартовая страница этой базы данных содержит ссылки на четыре раздела:

- список организаций;
- поиск сотрудников СО РАН;
- поиск организаций СО РАН;
- авторизованный вход.

Первые три раздела предназначены для доступа к публичной информации и относятся к пользовательской части базы данных. Административная часть обеспечивает навигацию, модификацию и просмотр информации, а также управление правами пользователей базы.

**Представление информации.** Информация, которая содержится в информационной системе, представлена пользователям на портале СО РАН. Его основные разделы:

- общая характеристика СО РАН (история, выдающиеся ученые, награды);
- события и мероприятия (общие собрания, конференции, семинары);
- проекты и программы;
- документы;
- новости;
- инфраструктура (издательства, библиотеки, газета “Наука в Сибири”, Дом ученых).

Каждый из разделов содержит подразделы (например, “Награды”, “Документы”, “Конференции”, “Диссертационные советы”), в которых помещены ссылки на коллекции.

Информация о наградах и присуждении наград сотрудникам состоит из трех коллекций, которые связаны между собой внешними связями. Это:

- коллекция со сведениями о наградах;
- информация о присуждении наград в таблице объявлений СО РАН;
- классификатор наград.

Для хранения информации о присуждении использована коллекция объявлений СО РАН. Это связано с тем, что любое присуждение награды есть некоторое событие, ко-

торое связано с документом в форме постановления. Поэтому было решено хранить эту информацию в виде объявлений с соответствующим заголовком и связывать с информацией о самой награде и персоне, ее получившей.

В разделе **“События и мероприятия”** содержится информация о прошедших, текущих и предстоящих событиях в научном сообществе. Событиями могут быть важные административные решения и документы, значительные достижения отдельных коллективов, юбилеи и встречи с выдающимися деятелями, объявления о конкурсах, грантах и премиях. Особую ценность и популярность имеют сведения о таких мероприятиях, как конгрессы, симпозиумы, конференции, совещания, семинары и школы. Востребованы как материалы уже прошедших конференций, так и сведения о предстоящих мероприятиях.

В настоящее время можно выделить следующие электронные источники информации о событиях.

1. Регулярный источник официальных сведений — общие собрания РАН и СО РАН. Их результатом являются отчеты, где представлены основные итоги организационной деятельности, постановления, приоритетные направления научных исследований. Важность этой информации обусловила необходимость создания отдельного ресурса для ее публикации.

2. Газета **“Наука в Сибири”** — официальное печатное издание СО РАН, где с 1961 года освещаются практически все стороны жизни научного сообщества Сибири. С 1998 года газета издается и в электронном виде по адресу: <http://www.sbras.ru/НВС>. Данный ресурс целесообразно использовать в качестве общедоступного архива наиболее значимых событий.

3. Электронная доска объявлений (<http://www.sbras.ru/win/anons>) — это ресурс, где в зависимости от прав доступа посетитель может поместить свое объявление, отнеся его к соответствующему разделу рубрикатора.

4. Информационная система **“Конференции”** [7] (<http://www.sbras.ru/ws>) разработана как вспомогательный инструмент для поддержки проведения конференций и совещаний и включает в себя средства публикации соответствующей информации в Интернете. Кроме материалов конференций данная система имеет сервис для размещения анонсов конференций, материалы которых публикуются другими средствами и системами.

В соответствии с проведенным анализом сформирована структура раздела **“События и мероприятия СО РАН”**, интегрирующая информацию из рассмотренных источников. В разделе **“Проекты и программы”** представлены тематические коллекции, связанные с научной деятельностью СО РАН, такие как **“Виртуальный музей науки и техники СО РАН”**, **“Биоразнообразие животного и растительного мира Сибири”** [8], **“Химия в СО РАН”** [9], **“Web-ресурсы математического содержания”** [10, 11].

В качестве примера рассмотрим **“Web-ресурсы математического содержания”**. Документы в коллекции упорядочиваются по их типу (персона, общество, институт, отдел, лаборатория, группа, факультет, кафедра, научная школа, конференция, семинар, издательство, журнал, книга, статья, проект, пакет программ, библиотека, коллекция, база данных, форум). Кроме того, применяется упорядочивание в соответствии с классификатором математических сущностей, используемым Американским и Европейским математическими обществами. Дополнительно предусмотрено установление внутренних связей между документами, что позволяет:

— при отображении информации об организации давать информацию о персонах, в ней работающих, и о публикациях этих персон;

— при отображении информации о персоне давать гиперссылку на сайт соответствующей организации и список ресурсов — публикаций данной персоны и т. д.

Все это относится к ресурсам СО РАН (в том числе с использованием информации из базы данных “СО РАН”), что позволяет существенно облегчить поиск нужной информации, а также дает пользователю системы более полное представление о состоянии математических исследований.

В разделе “**Инфраструктура**” представлена информация о сети диссертационных советов СО РАН, открытых в институтах по определенной специальности, а также описания паспортов специальностей научных работников.

Раздел “**Документы**” состоит из законодательных актов, указов, постановлений, рекомендаций и т. п., принимаемых органами государственной власти Российской Федерации, президиумами РАН, СО РАН, а также органами государственной власти субъектов федерации и президиумами научных центров СО РАН. Документ включается в коллекцию, если его содержание непосредственно затрагивает деятельность СО РАН или представляет интерес для широкого круга его сотрудников. До недавнего времени элементы коллекции “Документы” упорядочивались по дате принятия документа (в порядке убывания). Однако рост числа документов, размещенных на портале, привел к разработке новой, более детальной классификации. Прежде всего документы классифицируются по их регионально-ведомственной принадлежности:

- федеральный уровень;
- Российская академия наук;
- Сибирское отделение РАН;
- регионы и соответствующие научные центры СО РАН.

Внутри каждого раздела классификация ведется с учетом вида источника права:

- законы и кодексы (с выделением в отдельный подраздел Конституции и Уставов организаций);
- указы, постановления и распоряжения;
- международные договоры.

Наконец, внутри каждого подраздела документы упорядочиваются по дате принятия (в порядке убывания). Описанная классификация заметно облегчает поиск пользователем нужного документа.

## 1. Систематизация ресурсов системы

Основным объектом (ресурсом) хранения данных в информационной системе является документ, снабженный метаданными<sup>2</sup>. Понятие метаданных уже давно и успешно используют в таких контекстах, как информационные хранилища и системы аналитической обработки данных, электронный документооборот и управление потоками работ, управление знаниями. Метаданные позволяют описывать классы объектов, представляющие типы

---

<sup>2</sup>Под понятием “документ” в контексте данной статьи мы понимаем следующее. В информационном пространстве все информационные ресурсы: события, факты и любые другие сущности реального или виртуального мира — существуют только в форме некоторых информационных ресурсов (объектов). В этом смысле информационный ресурс, снабженный метаописанием, превращается в документ, который является основным “объектом”, с которым оперирует любая информационная система. Таким образом, документ — это структурированное описание реальной сущности (объекта, факта или понятия, включая алгоритмы и программы), которые составляют информационное наполнение системы.

информационных документов, структурные элементы документов, атрибуты и отношения между различными документами (ресурсами). Электронный документ — это наиболее близкое представление обычного документа, т. е. такого элемента данных, с которым удобнее всего работать человеку. Множество документов, принадлежащих одному классу, объединяются в коллекцию документов.

При разработке системы все коллекции были разбиты на следующие группы:

- словари (этот тип делится на два подтипа — словарь описания **элементов** и словари-классификаторы);
- информационные коллекции;
- каталоги.

В качестве базовой части словаря структурных элементов для всех коллекций в информационной системе выбран набор элементов в соответствии с международным стандартом Дублинского ядра (Dublin Core) [12].

Для описания дополнительных атрибутов документов конкретных коллекций этот словарь расширяется в зависимости от нужд этой коллекции. Например, для коллекций персон и организаций использовались элементы из рекомендаций vCard и профилей X.500. Для тематических коллекций допускается использование дополнительных элементов, необходимых для описания класса документов, ее составляющих. К обязательным требованиям относится регистрация этих элементов в словаре информационной системы.

Рассмотрим, как указанные положения реализуются в одной из коллекций, входящих в информационную систему СО РАН, — в сайте газеты “Наука в Сибири”.

**Словарь элементов** здесь расширен элементами краткого библиографического описания статей газеты. Для описания связей между документами и другими коллекциями используется дополнительный элемент link. Это служебный элемент информационной системы.

**Словари-классификаторы** — специальные коллекции с жестко зафиксированной структурой в соответствии с международными стандартами. Они включают в себя словарь номеров и словарь рубрик.

**Информационные коллекции** формируются из коллекции документов, отображаемых в газете, не имеющих четко фиксированной структуры. В документы этих коллекций можно добавлять новые объекты, можно удалять существующие. Были выделены следующие информационные коллекции: статьи и иллюстрации.

**Каталог** — специальная коллекция, представляющая метаописания всех документов всех информационных коллекций. Это вспомогательная информационная структура со следующими функциями: навигация (организация эффективного поиска документа по всем коллекциям) и установление связей (установка связей между документами, позволяющая вставлять в один документ информацию из другого).

Газета в целом имеет иерархическую структуру: на верхнем уровне имеется список всех номеров газеты, каждый номер состоит из статей. В случае необходимости возможно связать рубрику (выбрав ее из словаря рубрик). Кроме того, наличие элементов типа link позволяет осуществлять взаимосвязь статей газеты с другими коллекциями информационной системы СО РАН.

Связи между документами формируются либо по назначению (для формирования структуры газеты или для организации поиска), либо по типам коллекций (связь информационной коллекции со словарем, связь словаря со словарем).

Одним из важнейших словарей-классификаторов в информационной системе является “Рубрикатор информационных ресурсов”. Он предназначен для однозначной системати-

зации документов по рубрикам (разделам) информационной системы. Рубрикатор имеет фиксированные и дополнительные классификаторы, связанные с историей и научными достижениями СО РАН и т. п.

## 2. Технологии построения системы

В качестве базовой технологии работы портала используется концепция обеспечения доступа к информационным ресурсам в соответствии с трехуровневой моделью на основе открытых международных стандартов. Основными технологиями построения портала выступают:

— **WWW** — для организации пользовательских интерфейсов доступа к информации, организации шлюзов (Z39.50, LDAP) и информационных порталов, основанных на технологиях XML, RDF и web-сервисах;

— **Z39.50** — для организации унифицированного доступа к базам данных и реализации механизмов сквозного поиска в распределенной информационной системе;

— **LDAP** — для хранения простой информации в реплицируемых каталогах, предоставления доступа к ней по стандартному протоколу, для аутентификации и авторизации пользователей.

Коллекции хранятся на сайтах тех организаций, где они создаются (в Новосибирске, Томске, Иркутске и т. д.). На центральном сервере информационной системы ведется интегрированный каталог коллекций. Для его поддержки разработана распределенная система ведения каталогов (метаданных) информационных ресурсов СО РАН.

## Заключение

В заключение уместно отметить постоянно развивающуюся тенденцию к интеграции уже существующих информационных ресурсов. Одно из необходимых условий для этого — каталогизация информационных ресурсов и документов на основе существующих открытых стандартов.

## Список литературы

- [1] Шокин Ю.И., Федотов А.М., Леонова Ю.В. Принцип динамического формирования документов в информационных системах на примере интегрированной распределенной информационной системы (ИРИС) СО РАН // Тр. Четвертой Всерос. науч. конф. “Электронные библиотеки: перспективные методы и технологии, электронные коллекции”, Дубна, 15–17 октября 2002 г. Т. 2. С. 159–169.
- [2] Гуськов А.Е. О модели информационных цифровых систем // Вычисл. технологии. 2005. Т. 10. Спецвыпуск. С. 58–70.
- [3] Информационная система СО РАН. <http://www.sbras.ru>
- [4] Шокин Ю.И., Федотов А.М., Гуськов А.Е. и др. Электронные библиотеки — путь интеграции информационных ресурсов Сибирского отделения РАН // Вест. КазНУ. Спецвыпуск. Алматы, Каз. нац. ун-т им. аль-Фараби. 2005. № 2. С. 115–127.

- [5] ИНФОРМАЦИОННАЯ СИСТЕМА “База данных и организаций сотрудников СО РАН”.  
<http://www.sbras.ru/sbras/db>
- [6] ШОКИН Ю.И., ФЕДОТОВ А.М., КЛИМЕНКО О.А., ЛЕОНОВА Ю.В. Содержательное наполнение справочно-информационной системы научного сообщества // Вычисл. технологии. (Совместный выпуск). Вест. КазНУ им. аль-Фараби. Сер. Математика, механика, информатика. 2004. Ч. 4. Т. 42, № 3. С. 346–350.
- [7] ФЕДОТОВ А.М., ГУСЬКОВ А.Е., МОЛОРОДОВ Ю.И. Информационная система поддержки проведения конференций СО РАН // Матер. выездного заседания координ. науч. совета СО РАН по целевой программе “Информационно-телекоммуникационные ресурсы СО РАН”, Иркутск, 29–30 июля, 2003 г. С. 91.
- [8] ЭЛЕКТРОННЫЙ АТЛАС “Биоразнообразие животного и растительного мира Сибири”.  
<http://www.sbras.ru/win/elbib/bio/>
- [9] ИНФОРМАЦИОННАЯ СИСТЕМА “Химия в СО РАН”. <http://www.catalysis.nsk.su/chem/>
- [10] ИНФОРМАЦИОННАЯ СИСТЕМА “Web-ресурсы математического содержания”.  
[http://www.sbras.ru/win/elbib/data/show\\_page.dhtml?2+184](http://www.sbras.ru/win/elbib/data/show_page.dhtml?2+184)
- [11] БАРАХНИН В.Б., ГУСЬКОВ А.Е., КЛИМЕНКО О.А. и др. Информационная система “Web-ресурсы математического содержания” // Матер. конф. молодых ученых, посвященной М.А. Лаврентьеву. Новосибирск, 17–19 ноября, 2004. Ч. I. С. 23–27.
- [12] DUBLIN Core Metadata Initiative (DCMI). <http://www.dublincore.org>

*Поступила в редакцию 31 октября 2006 г.*