

# Информационная система Сибирского отделения РАН\*

Шокин Ю.И.  
Сибирское отделение РАН  
Новосибирск  
shokin@ict.nsc.ru

Федотов А.М.  
Институт вычислительных технологий СО РАН  
Новосибирск  
fedotov@ict.nsk.su

## Аннотация

В докладе дан обзор основных работ, проводимых в Сибирском отделении РАН, в области сохранения и поддержки информационных ресурсов отделения и создания “Электронной библиотеки Сибирского отделения РАН”.

## 1 ИНФОРМАЦИОННЫЕ РЕСУРСЫ

В Сибирском отделении РАН накоплена и постоянно пополняется уникальная научная информация. Но, к сожалению, пока отсутствует единая технология сбора информации, разобщены как места ее хранения, так и формы представления и, что самое главное, практически отсутствует информация об информации. Громаднейший оригинальный материал находится в виде бумажных архивов и в большинстве случаев не представляет организованную информационную среду, которая является основой для современных научных исследований, что является существенным препятствием на пути развития интеграционных проектов и других форм научного взаимодействия. Все это не позволяет широко использовать информационные ресурсы для научных исследований.

В силу не зависящих от нас причин, постоянно уменьшаются внешние контакты сотрудников Отделения, особенно молодых, число командировок и поездок на конференции. Сократилось и количество приходящих в библиотеки научных журналов, что приводит к сильному уменьшению количества информации об информации. Существенной проблемой, связанной с информационными ресурсами Отделения, является факт “гибели” информации или недоступности информации, вызванный постоянным

\*Работы поддержаны грантами РФФИ № 00-07-90335 и СО РАН № 88

©Вторая Всероссийская научная конференция  
ЭЛЕКТРОННЫЕ БИБЛИОТЕКИ:  
ПЕРСПЕКТИВНЫЕ МЕТОДЫ И ТЕХНОЛОГИИ,  
ЭЛЕКТРОННЫЕ КОЛЛЕКЦИИ  
26-28 сентября 2000г., Протвино

оттоком кадров исследователей и технического персонала, ответственного за ее хранение. Зачастую жизненный путь информации заканчивается с “уходом” ее создателя. В настоящий момент перед нами стоит задача сохранения всего того громадного информационного материала, который был накоплен в Отделении на протяжении многих лет.

Для решения проблемы информационной обеспеченности необходимо создание и поддержка собственной универсальной информационной системы Отделения, включая создание и поддержку “Электронной библиотеки Сибирского отделения РАН”([22]), содержащей полнофункциональную систему об интеллектуальном потенциале Отделения (информационную систему об институтах, сотрудниках [2], об основных достижениях [18] и др. аспектах, связанных с работой Отделения) и систему электронной поддержки сбора и накопления информации (системы электронных коллекций, баз данных и т.п.).

Электронные публикации научных коллекций представляют собой новую форму хранения и обмена информацией. Для нее характерны прежде всего динамичность (возможность обновления) и глобальный доступ (через компьютерные сети). На сегодняшний день электронные публикации не преобладают в общем объеме опубликованных информационных ресурсов, но их доля стремительно растет. При этом возрастает и качество таких публикаций. Изначально электронная публикация появилась как электронный вариант бумажной публикации. В последние годы она приобретает все более самостоятельное значение. Более того, часть научных проектов в процессе их выполнения переориентируется с бумажной формы на электронную.

В настоящий момент в Сибирском отделении сформирована целевая программа создания и развития информационных ресурсов под общим названием “Электронная библиотека Сибирского отделения РАН” ([33], см. рис. 1), в рамках которой разработан и осуществляется целый ряд проектов поддержки работ по созданию информационных ресурсов с использованием Internet/Intranet технологий [27] и организации системы доступа к информации через WWW сервис Internet (часть которых поддерживается РФФИ или другими фондами). Основные направления программы связаны с формированием собственных электронных ресурсов по основным отраслям наук (науки о земле, химия, биология, археология и др.), созда-

нию и поддержке электронных коллекций и электронных публикаций, организации удобных систем доступа к библиотечным и библиографическим базам данных ГПНТБ СО РАН и базам данных институтов Отделения, организации зеркал наиболее значимых мировых информационных ресурсов, необходимых для работы сотрудников Отделения.

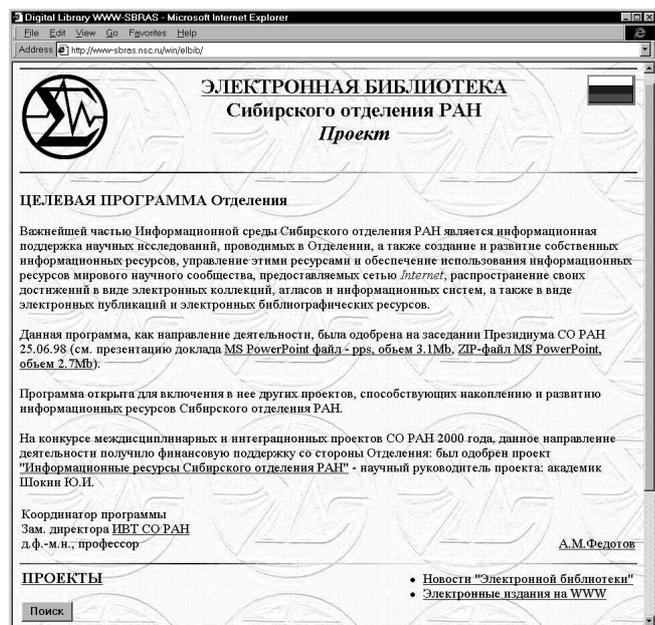


Рис. 1: Титульная страница программы “Электронная библиотека СО РАН”.

Информационная система Сибирского отделения РАН должна обеспечить:

- Единую информационную среду Отделения, основанную на современных сетевых средствах и перспективных информационных технологиях.
- Информационную поддержку исследований по фундаментальным и прикладным направлениям, проводимым в институтах Отделения, а также межинститутских междисциплинарных научных исследований.
- Поддержку профессионально-ориентированных систем подготовки и обмена научных документов с элементами удаленной совместной работы.
- Поддержку профессионально-ориентированных систем доступа и интерфейсов с банками данных и автоматизированными библиотеками.
- Поддержку перспективных систем телекооперации исследователей на базе современных технологий.
- Коллективное использование приобретаемой электронной литературы, каталогов, баз данных и библиографических изданий.

- Поддержку электронных версий научных журналов, издаваемых институтами Отделения. Издание собственных электронных журналов, книг, препринтов и дайджестов по различным направлениям научных исследований.
- Поддержку принятия и реализации организационных и управленческих решений в Отделении. Создание среды, обеспечивающей электронный документооборот Отделения.

Важной проблемой выступает организация разнородной информации в удобном для конечного пользователя виде, что требует новых исследований и разработок интерфейсов для корректного отражения предметной области. Другой задачей обслуживания электронных коллекций является стандартизация данных, разработка технологических решений и юридических аспектов использования информации, включая вопросы интеллектуальной собственности. Отличительной чертой электронной библиотеки является возможность параллельного использования различных поисковых механизмов и средств доступа к банкам электронных данных. Так в качестве ответа на запрос к электронной библиотеке пользователю может быть представлен не единственный электронный документ (или его фрагмент) и совсем не обязательно в текстовой форме, необходимы такие информационные системы, которые бы обеспечивали эффективный комплексный поиск и анализ информации в коллекциях разнородных объектов.

В одном из основных пунктов постановления Президиума СО РАН о “Сети Интернет Новосибирского Научного Центра” [26] (N 81 от 22.03.99) сказано: “Считать первоочередной задачей создание и активное расширение объемов собственных информационных ресурсов сети, используя, в том числе, конкурс интеграционных проектов”.

Организационно-технологическое обеспечение процесса создания полнофункциональной информационной системы включает в себя большой спектр работ, связанных с организацией системы доступа пользователей к информационно-вычислительным ресурсам и к базам данных, сохранение, поддержку и создание информационных ресурсов Отделения и, что самое главное, *воспитание нового пользователя, способного работать в современном информационном мире.*

## 2 ПРОЕКТЫ

Создание информационной системы Отделения основано на реализации целого ряда проектов<sup>1</sup> связанных одной общей информационной средой, которая базируется на “Информационном сервере Отделения” [10] (см. рис. 2).

Дадим краткое описание некоторых из них:

- **Создание электронной библиотеки ГПНТБ СО РАН ([6]).**

Проект направлен на решение проблем информационного сопровождения фундаментальных научных

<sup>1</sup>Описания основных проектов даны на сайте “Электронной библиотеки Отделения” [33] и приведены в списке литературы [6, 7, 9, 13, 16, 17, 21, 23, 24, 29, 31].

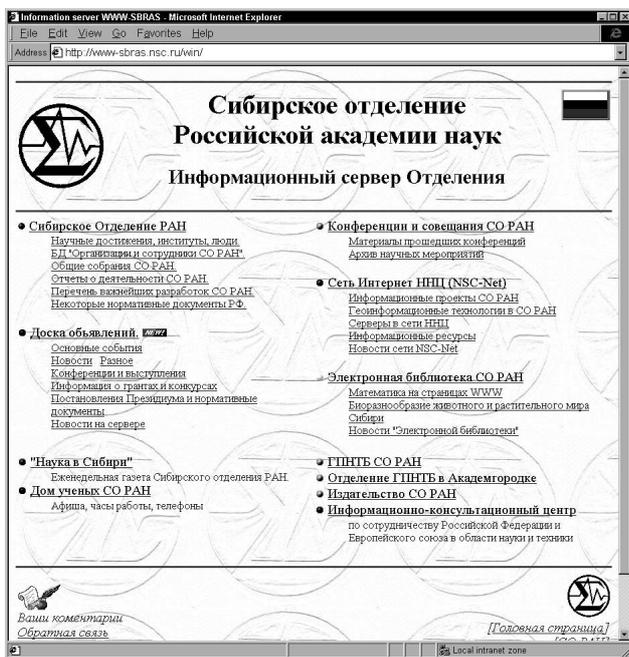


Рис. 2: Информационный сервер отделения.

исследований, проводимых в Сибирском регионе и предусматривает создание и внедрение многофункциональной автоматизированной системы распределенных баз данных для поддержки научных исследований в Отделении, обеспечивающей интегральное использование разнородных информационных ресурсов, работающей в сетевом режиме и поддерживающей более 50 тематических баз данных с глубокой ретроспективой от 5 до 10 лет. В качестве источников информации для создаваемой электронной библиотеки предполагается использовать в сетевом режиме накопленные в ГПНТБ СО РАН базы данных по физике, химии, механике, экологии, медицине и др. (общий объем более 26 млн. документов).

В результате электронные каталоги, базы данных научно-технической информации регионального характера, полные тексты изданий ГПНТБ СО РАН, сборники научных трудов и монографии ученых библиотеки, серии аналитических обзоров мировой литературы, коллекции фонда редких книг и рукописей предоставлены мировому научному сообществу через Интернет [3].

• **Полнотекстовые базы данных научных публикаций: Электронные журналы СО РАН ([31]).**

Данный проект ориентирован на создание полнофункциональной системы одновременной подготовки электронных и печатных версий математических журналов Отделения и создание электронных версий некоторых журналов, издаваемых издательством СО РАН [34].

В качестве первого этапа математической части проекта выбраны журналы “Вычислительные техно-

логии” [32] и “Сибирский математический журнал”. Планируемая работа состоит их двух частей:

1. Создание актуализируемой полнотекстовой базы данных с доступом через Интернет с расширенными элементами поиска авторов и статей, включая контекстный поиск по содержанию статей и аннотаций, а также поиск математических формул и выражений, записанных в Т<sub>Р</sub>X’овской нотации.
2. Создание системы поддержки полного документооборота редакции (автоматическое получение статьи, ее регистрация, отправка рецензенту, редподготовка, а также справочная система для авторов о прохождении статей) и системы автоматического формирования электронной версии журнала одновременно с подготовкой печатной.

Созданная система позволяет реализовать расширенный поиск конкретных фрагментов опубликованных статей.

Электронные версии журналов Издательства СО РАН пока представляют собой содержания журналов с ссылками на полные тексты в PDF формате.

• **Электронный атлас “Биоразнообразие животного и растительного мира Сибири” ([13]).**

Электронный атлас призван обеспечить систематизацию и организацию широкого доступа к разнообразной информации по биоразнообразию растительного мира Сибири на основе современных информационных технологий. Предполагается разработать общие подходы по систематизации и сохранению разнородной ботанической и экологической информации, интегрировать эту информацию в Геоинформационную систему, сделать доступными через сеть Internet массивы данных материалов и коллекций.

Основные источники информации - данные, собранные исследователями в течение нескольких десятилетий в биологических институтах Отделения, научные журналы, монографии, учебники, материалы, хранящиеся в зоологических музеях и гербариях, а также результаты экспериментальных исследований и полевые журналы экспедиций.

В рамках атласа создаются базы данных со средствами поиска по важнейшим группам растений и растительных сообществ. В частности, базы данных “Исчезающие и нуждающиеся в охране растительные сообщества Сибири” [1] (см. рис. 3), “Редкие виды растений Сибири”, “Охраняемые природные территории”, “Определители таксономической принадлежности”, “Лекарственные растения”, “Природно-очаговые болезни”, “Генофонд сельскохозяйственных растений и животных” и т.д.

Основой атласа являются уникальные коллекции и массивы данных по биоразнообразию растительного мира Сибири. Так например, в гербариях ЦСБС СО

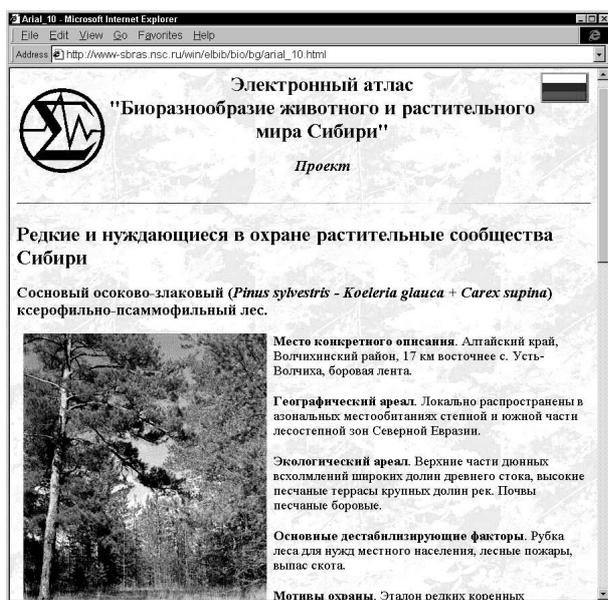


Рис. 3: “Зеленая книга Сибири” — пример документа.

РАН хранится около 310 тыс. образцов высших сосудистых растений, собранных на территории Сибири, в числе которых и уникальные типовые образцы. Имеются коллекции мхов, грибов, лишайников, низших растений. В фитоценоотеке хранятся материалы более чем 10 тыс. описаний разнообразных растительных сообществ. Собрание древесных растений ЦСБС СО РАН насчитывает около 800 видов, гибридов и форм, что в 2.5 раза превышает количество видов дендрофлоры Сибири.

- **Интегральная электронная библиотека по пространственным структурам и функциям ДНК, РНК и белков ([9]).**

Проект направлен на изучение принципов структурно-функциональной организации биологических макромолекул (ДНК, РНК, белков), молекулярных механизмов формирования их пространственных структур и, что не менее важно, на исследование взаимосвязей между пространственными структурами этих макромолекул и их специфической биологической активностью. Решение именно этих задач является исключительно важным и определяющим для широкого круга генно-инженерных, биотехнологических и медико-биологических исследований.

- **Объединенная информационная система по наукам о Земле СО РАН ([7]).**

Проект направлен на создание в Новосибирском научном центре СО РАН интегрированной информационной системы по наукам о Земле коллективно пользования внутренних и внешних информационных ресурсов и современных сетевых технологий, интегрированных с Internet. В рамках этой системы предполагается организация доступа к распределен-

ным базам данных формата CDS/ISIS, предоставленные электронных копий журналов “Геология и геофизика”, “Криосфера Земли”, “География и природные ресурсы”.

- **Объединенная информационная система по химическим наукам СО РАН ([17]).**

Создание объединенного центра по химическим наукам СО РАН предполагает развитие унифицированной технологии ведения электронной библиотеки по химическим наукам, включающей в себя создание системы электронных библиографических каталогов, позволяющих осуществлять эффективный поиск первоисточников информации; приобретение и создание собственных фактографических баз данных по свойствам химических соединений, встраивание их в общую информационную систему; создание аннотированных электронных каталогов химических журналов, издаваемых в СО РАН; создание полнотекстовых электронных версий химических журналов СО РАН [12].

- **Электронные ресурсы сибирской и мировой археологии и этнографии ([23]).**

Электронная библиотека предназначена для накопления, систематизации, сохранения и организации широкого доступа к информации о древней истории, культуре и искусству Северной Азии, в том числе:

- Создание и поддержка электронных версий уникальных коллекций СО РАН в области археологии и этнографии: “Виртуальные музей и энциклопедии Сибири” [5].
- Информационная система об интеллектуальном потенциале археологии, этнографии и антропологии Сибири. Информационно-справочная система по археологам, этнологам, антропологам мира.
- Создание информационных систем и пространственных баз данных по археологии Сибири и сопредельных территорий. Разработка системной классификации археологического знания. Создание информационных систем по фактографическим и статистическим базам данных палеолитических памятников Северной Азии и базы данных по абсолютному датированию археологических памятников Сибири, Дальнего Востока и Китая. Создание банка данных археологических памятников.
- Создание полнотекстовых баз данных археологических и этнографических публикаций (журналов, бюллетеней “Компьютеры и археология”, “Siberian Archaeological Herald”, монографий, сборников научных трудов, дайджестов, оглавлений и резюме ведущих археологических и этнографических журналов), а также полнотекстовых баз данных полевых отчетов.

- **Обеспечение унифицированного доступа к разнородным коллекциям и информационным ресурсам на основе технологии CORBA ([16]).**

Проект направлен на организацию доступа к информационным ресурсам Интернет, которые, как правило, являются разнородными и несовместимыми, тогда как пользователю часто требуется однородный доступ к информации вне зависимости от ее источника. При этом существенными являются две составляющих доступа: запрос на поиск документа (или ресурса) и форма получаемого документа. В основу проекта положена разработка технологии, позволяющая объединить информационные источники, родственные по предоставляемой информации в единую пользовательскую среду доступа к ним.

### 3 ИНФОРМАЦИОННЫЙ СЕРВЕР ОТДЕЛЕНИЯ

Информационный сервер Отделения [10] является интегрирующим звеном для системы поддержки информационных ресурсов Отделения.



Рис. 4: БД — основные разработки СО РАН.

В настоящий момент сервер содержит разнообразную информацию о деятельности Сибирского Отделения, включая: научно-исследовательские и конструкторско-технологические институты (База данных СО РАН [2]); публичная доска объявлений [19]; основные результаты исследований и годовые отчеты СО РАН; интеллектуальный (научный) потенциал Сибири; конференции СО РАН; электронная библиотека СО РАН; поддержка исследований в области математики [29]; важнейшие разработки Институтов СО РАН, предлагаемые для широкого использования [18] (см. рис. 4); Сеть Internet Новосибирского научного центра; информационные проекты СО РАН [26]; Геоинформационные технологии в СО РАН (информационный бюллетень) [8]; Дом Ученых СО РАН; газета “Наука в Сибири”; справочные материалы по информационным ресурсам; нормативные документы РФ, касающиеся науки, информатизации и средств свя-

зи; схемы автомобильных дорог Новосибирской области; фотоальбом Сибирского отделения РАН; история Сибирского отделения РАН и т.д.

Из наиболее важных информационных систем отметим:

- Информационная система о научном потенциале Сибири.
- База данных по научным организациям и сотрудникам отделения Отделения [2].
- Информационная система по важнейшим разработкам СО РАН [18].
- Справочная система по информационным ресурсам Internet.
- Информационная система по правовой базе науки и информатизации.
- Информационная система по конференциям СО РАН.
- Информационная система “Математика на WWW” [29]. Система включает доступ к зеркалу Электронной библиотеки Европейского математического общества (EMIS), в которой реализован доступ к математической базе данных.
- Справочная система о СО РАН (общие собрания и отчеты о научной деятельности).
- Электронная версия газеты “Наука в Сибири”.

### 4 БИБЛИОТЕЧНО-ИНФОРМАЦИОННОЕ ОБСЛУЖИВАНИЕ

Среди других немаловажных ресурсов Отделения следует отметить разработку системы библиотечно-информационного обслуживания сотрудников СО РАН, включающую

- Проведение поиска вторичных источников информации по электронным каталогам и библиографическим базам данных ГПНТБ СО РАН и институтов СО РАН;
- Предоставление первоисточников информации (книг, журналов и др. документов) средствами межбиблиотечного обмена, а также в виде твердых и электронных копий;
- Предоставление доступа к полнотекстовым электронным базам данных ведущих мировых издательств (Springer, Elsevier и др.);
- Создание собственных электронных ресурсов в рамках проекта “Электронная библиотека СО РАН”.

Среди таких систем в Отделении наиболее развитыми являются:

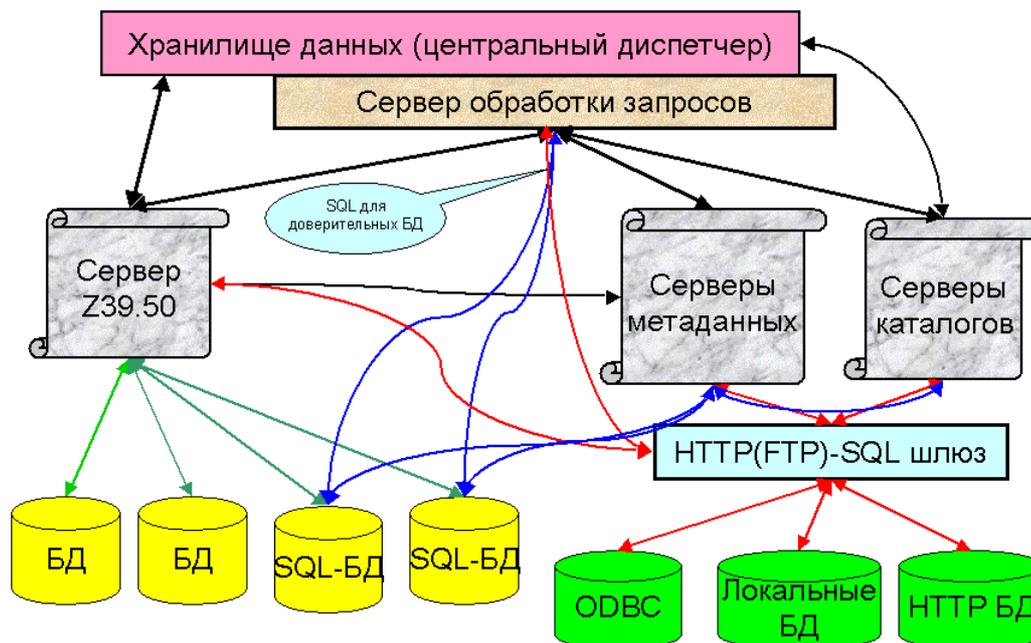


Рис. 5: Общая схема, централизованно-распределенной системы хранения информации.

- информационная система по различным областям знаний на базе ГПНТБ СО РАН: Базы данных и электронные каталоги (из действующих отметим электронный каталог ГПНТБ СО РАН) [3];
- объединенная информационная система по химическим наукам Сибирского Отделения РАН на базе ИК СО РАН: Базы данных по Химии и Биологии;
- информационная система по наукам о Земле Сибирского Отделения РАН на базе Информационно-библиотечного центра и сектора Вычислительных технологий ОИГГиМ СО РАН: (Базы данных по Геологии, Геофизике и др) [4].

## 5 НАПРАВЛЕНИЯ РАБОТ

Для решения задач преодоления информационной недостаточности и организации системы хранения и поддержки информационных ресурсов в Отделении разработан проект *“Информационные ресурсы Сибирского отделения РАН”* [25], направленный на организацию работ по созданию *“Информационной системы Отделения”*. Основными задачами проекта являются:

- создание средств хранения, доступа и обработки больших массивов распределенной информации;
- создание системы поддержки интеллектуального интерфейса;

- создание системы формирования и эксплуатации корпоративных информационных ресурсов и информационного обеспечения;
- создание системы защиты и авторизованного доступа к информации;
- лицензионная поддержка;
- обеспечение нормативно-правовой базы создания и использования информационных ресурсов;
- создание и поддержка интегрированной системы предоставления фактографической информации.

В настоящий момент при создании электронной библиотеки уже решается целый ряд технологических задач хранения и накопления информации, с одной стороны, и задачи организации удобного доступа к информации, с другой стороны. Принципиально здесь можно выделить следующие направления работ:

- создание информационных систем, обеспечивающих эффективный комплексный поиск и анализ информации в коллекциях гетерогенных объектов;
- представление разнородной информации в удобном виде для конечного пользователя;
- разработка удобных человеко-машинных интерфейсов, направленных в том числе на поиск и актуализацию данных;

- согласование используемых стандартов хранения и представления информации и мета информации и технологических решений;
- решение вопросов защиты информации и авторизации доступа к информации;
- решение юридических вопросов использования электронных информационных ресурсов, включая вопросы авторского права и интеллектуальной собственности.

Из первоочередных задач, которые решаются в настоящий момент отметим следующие:

1. Инвентаризация наиболее существенных информационных ресурсов. Это касается в первую очередь информационных ресурсов, а так же ресурсов, связанных с интеграционными проектами СО РАН.
2. Адаптация существующих разработок институтов СО РАН в области построения распределенных систем для Информационной Системы СО РАН.
3. Организация централизованно-распределенной системы хранения информации, ориентированной на распределенные базы данных и хранилища информации.
4. Разработка корпоративных стандартов хранения, поиска и представления информации на основе существующих международных и отечественных стандартов. Разработка технологии хранения и представления через сеть разнородной информации (включая пространственные БД, карты, схемы и т.д.).
5. Разработка и адаптация технологий коллективной работы исследователей, в том числе в рамках международных и других интеграционных проектов.

При реализации концептуальной схемы проекта мы столкнулись с необходимостью поддержки функционирования старых СУБД, работающих под различными платформами. Поэтому в основу организации централизованно-распределенной системы хранения информации положена концепция информационных хранилищ с учетом поддержки уже функционирующих технологий, которая представлена на схеме, изображенной на рис. 5.

Следующий этап работы, связан с созданием полнофункциональной информационной системы (рассчитанную в первую очередь на Президиум СО РАН), которая должна состоять из различных баз данных, характеризующих деятельность Отделения:

- БД институты и научные направления, основные достижения и результаты, отчеты, планы, проекты и публикации — частично эти базы уже функционируют на информационном сервере СО РАН [2] (см. рис. 6).
- Интегрированная распределенная БД кадры (сотрудники, проекты результаты, публикации), объединяющая БД институтов и Управление кадров.

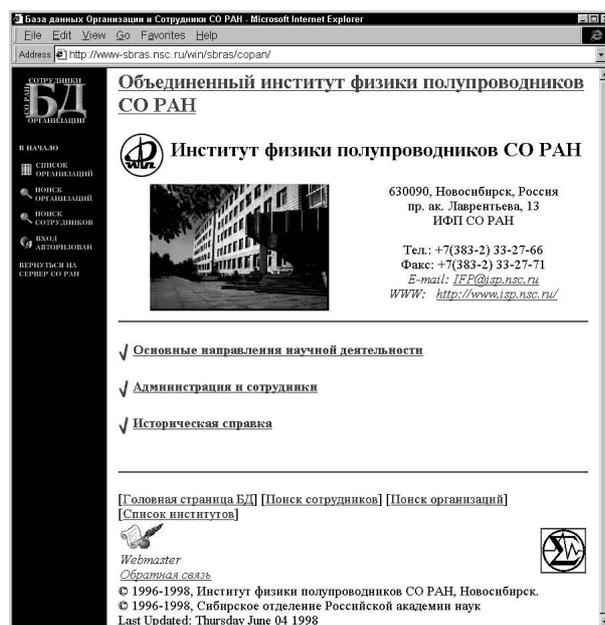


Рис. 6: База данных по организациям СО РАН.

- БД постановления и распоряжения СО РАН, официальная рассылка (реализуется на информационном сервере СО РАН в разделе “Доска объявлений” [19]) и документооборот.
- Поиск и электронный заказ информационных ресурсов (включая копии статей) — создание зеркал наиболее значимых мировых ресурсов (например, “Математика на страницах WWW” [29], а также электронные каталоги ГПНТБ СО РАН [3], ОИГГМ СО РАН [4] и ИК СО РАН) и электронный заказ ресурсов, который должен обеспечить гарантированную доставку документов через сеть Интернет и снизить нагрузку на каналы за счет исключения повторных запросов).
- Распределенная электронная библиотека с единой системой поиска и доступа к информации через сеть Интернет.
- Информационно-справочные системы (конференции и гранты, объявления, новости и т.д. — реализуется на информационном сервере СО РАН [10], SICC [11] и на серверах институтов).
- Системы представления пользователю информации о данных.

## 6 ТЕХНОЛОГИИ

В основу создания Электронной библиотеки СО РАН положен принцип Internet/Intranet технологий [27], использование которых при создании информационных ресурсов и построении информационных систем различного назначения в ближайшее время станет доминирующим в ми-

ровом информационном пространстве. Это связано с тем, что указанные технологии:

- Позволяют организовать с достаточной простотой для пользователя системы поиска нужной информации.
- Предъявляют минимальные требования как с технической стороны, так и со стороны программного обеспечения к рабочему месту клиента (клиент работает со стандартным программным обеспечением).
- Поддерживают распределенные системы хранения информации и множественные методы ее хранения.
- Поддерживают работу с практически неограниченным объемом разноплановых данных (текст, графика, изображение, звук, видео, векторные карты и др.).
- Предоставляют технологически простой способ администрирования информационных систем с одного рабочего места.
- Поддерживают удаленные методы редактирования и пополнения информации.

Internet технология позволяет удачно сочетать возможности гипертекстового оформления информации с использованием возможностей современных систем управления базами данных, причем со стороны клиента полностью унифицируются запросы на поиск и представление информации, а также получение аналитических справок и данных из информационных систем. Что самое главное, Internet технология позволяет оперативно управлять и актуализировать информацию, хранящуюся в базах данных через просмотрщик (браузер) WWW страниц и создавать динамические системы формирования коллекций.

Используемые технологии позволили реализовать два основных принципа, декларируемых при создании больших информационных систем: независимость функционирования системы от платформы (на стороне клиента) и независимость от используемой СУБД.

## 7 ДИНАМИЧЕСКАЯ МОДЕЛЬ ДОКУМЕНТА

В основе создания электронных коллекций в наших работах положена концепция “динамической системы” формирования документов. Согласно данной концепции каждый тип документов, содержащих информацию о конкретных фактах, представляется в виде набора объектов со своими характеристиками и атрибутами (аналогично тому, как это принято в объектных языках программирования). Разработанная технология позволяет оперативно управлять и актуализировать информацию, хранящуюся в разнородных и распределенных по сети базах данных, организовать гибкий поиск, что самое главное, создать достаточно удобный интерфейс для ее наполнения. При описании документа, составляющего коллекцию, выделяются два основных класса объектов, характеризующих документ:

1. **Специальные объекты**, например, объект-заголовок, описывающий внешний вид документа, при выдаче его пользователю или навигационный объект, описывающий гипертекстовые связи между отдельными документами (или его частями).
2. **Информационные объекты**, содержащие фактическую информацию (текст, графика, измерения и т.п.).

Очевидно, что первый тип объектов связан с WWW технологией представления информации.

Содержательная часть информационных объектов, как правило, хранится в базе данных, однако возможно их хранение и в файловой системе. Объекты подразделяются на простые объекты, которые не имеют связей с другими коллекциями документов или с другими документами или объектами и сложные объекты, которые такие связи имеют. Использование связей позволяет создавать коллекции с иерархической структурой (например, **виды**  $\Leftrightarrow$  **популяции**  $\Leftrightarrow$  **особи**).

Для занесения в электронную коллекцию фактографической информации, последняя подлежит паспортизации — формальному описанию структуры объектов, составляющих документы (см., например, унифицированный паспорт описания сообществ в информационной системе “Зеленая книга Сибири” [1], содержащей информацию о редких и нуждающихся в охране растительных сообществах Сибири).

По данной технологии в настоящий момент создаются и уже функционируют коллекции электронного атласа “Биоразнообразие животного и растительного мира Сибири” [13], коллекции полнотекстовых документов (см. электронную версию журнала “Вычислительные технологии” [31], публичная доска объявлений СО РАН [19], информационная система “Геоинформационные технологии в СО РАН” [8], и др.

**Электронная коллекция** в нашем понимании — это набор документов, которые содержат фактографическую информацию, имеющую одинаковое формальное описание структуры — “паспорт”. Работа с любой электронной коллекцией разделяется на две части: работа с описаниями коллекции и документов (работа с метainформацией) и работа с содержательной частью коллекции (работа с фактографической информацией). Словесная характеристика основных терминов следующая:

**Коллекция:** набор однотипных документов — характеризуется своим описанием и стилем коллекции, а также описанием структуры, входящих в нее документов.

**Документ:** — характеризуется своим описанием и стилем документа, а также описанием характеристик (атрибутов, свойств и функций) объектов его составляющих.

**Объект:** — определяется заданием типа объекта и описанием его атрибутов, свойств и функций.

Метainформация и интерфейсные модули хранятся на специализированном сервере метаданных, а фактографическая информация может храниться на любом сервере

в сети Интернет под управлением СУБД, обрабатывающей SQL запросы через TCP/IP порт. В настоящий момент система может работать с СУБД Oracle, Informix, PostgreSQL, MySQL, SyBase, MS SQL, dBase, InterBase, AdabasD, DB2 и MS SQL server. Тип используемой СУБД может повлиять только на некоторые специфические характеристики объектов, например, разные СУБД имеют разные ограничения на размер неструктурированных текстовых полей, максимальный размер которых колеблется от 255 символов до 4 Mb и более.

В системе описания документов могут использоваться простые (элементарные) объекты и сложные объекты: составленные из других объектов или являющиеся ссылками на объект или документ.

Система имеет трехуровневую структуру:  
 объект (метаинформация) → документ (метаинформация)  
 → коллекция (метаинформация).

Функционирование электронной коллекции основывается на информационной структуре данной коллекции документов — метаинформации, поэтому каждый информационный объект должен в своем метаописании нести “полную характеристику” как его “использовать”.

Характеристики внешней структуры и стиля документа описываются в стандарте SGML (XML), что позволяет представлять выходную информацию в любом из используемых в настоящий момент форматах электронных публикаций (HTML, L<sup>A</sup>T<sub>E</sub>X, RTF).

## 8 ЗАКЛЮЧЕНИЕ

В заключении отметим, что в конце апреля 1999 года в Новосибирском Академгородке состоялось “**Четвертое рабочее совещание по электронным публикациям**”, которое проводилось с использованием динамической системы формирования документов. На этом совещании представлено большинство результатов работ Отделения, посвященных технологии создания электронных коллекций и библиотек (см. более подробную информацию в материалах совещания и в докладах [14]).

Следующее (5-е) совещание состоялось в июне 2000 года. На этом совещании основное внимание было уделено разработке программы создания корпоративной информационной системы СО РАН (см. материалы и тексты докладов совещания [15]).

## Список литературы

- [1] База данных “Зеленая книга Сибири”.–  
<http://www-sbras.nsc.ru/win/elbib/bio/green/>.
- [2] База данных СО РАН.–  
<http://www-sbras.nsc.ru/win/sbras/copan/>.
- [3] Базы данных и электронные каталоги ГПНТБ СО РАН.–  
<http://info.spsl.nsc.ru/>.
- [4] Библиографический WEB сервер ОИГГиМ СО РАН.–  
<http://geolibr.uiggm.nsc.ru/LWS/>.
- [5] Виртуальный музей “Древняя История, Культура и Искусство Северной Азии”.–  
[http://sati.archaeology.nsc.ru/virtual\\_e.htm](http://sati.archaeology.nsc.ru/virtual_e.htm).
- [6] Елепов Б.С. *Создание электронной библиотеки ГПНТБ СО РАН.*–  
<http://www-sbras.nsc.ru/libr/win/Rffi98.htm>.
- [7] Жижимов О.Л. *Объединенная информационная система по наукам о Земле СО РАН.*–  
<http://www-sbras.nsc.ru/win/elbib/geo/>.
- [8] *Информационная система “Геоинформационные технологии в СО РАН.”*–  
<http://www-sbras.nsc.ru/win/gis/>.
- [9] Колчанов Н.А. *Интегральная электронная библиотека по пространственным структурам и функциям ДНК, РНК и белков.*–  
<http://wwwmgs.bionet.nsc.ru/mgs/>.
- [10] *Информационный сервер Сибирского отделения РАН.*–  
<http://www-sbras.nsc.ru/win/>.
- [11] *Информационный сервер Сибирского информационно-консультативного центра по сотрудничеству Российской Федерации и Европейского союза в области науки и техники.*–  
<http://www-sbras.nsc.ru/sicc/>.
- [12] *Информационный сервер “Химия в СО РАН.”*–  
<http://www.catalysis.nsk.su/chem>.
- [13] Коропачинский И.Ю., Шокин Ю.И., Шумный В.К., Ермаков Н.Б., Колчанов Н.А., Федотов А.М. *Электронный атлас “Биоразнообразие животного и растительного мира Сибири.”*–  
<http://www-sbras.nsc.ru/win/elbib/bio/>.
- [14] *Материалы VI рабочего совещания по электронным публикациям El-PUB’99.*–  
<http://www-sbras.nsc.ru/ws/elpub99/>.
- [15] *Материалы V рабочего совещания по электронным публикациям El-PUB’2000.*–  
<http://www-sbras.nsc.ru/ws/el-pub-2000/>.
- [16] Марчук А.Г., Осипов А.Е. *Обеспечение унифицированного доступа к разнородным коллекциям и информационным ресурсам на основе технологии CORBA. // Тезисы докладов семинара-совещания “Электронные библиотеки: перспективные методы и технологии, электронные коллекции”, М.:1998.*
- [17] Пармон В.Н. *Объединенная информационная система по химическим наукам СО РАН.*–  
<http://www-sbras.nsc.ru/win/ellib/parmon.html>.
- [18] *Перечень важнейших разработок СО РАН, предлагаемых для широкого использования.*–  
<http://www-sbras.nsc.ru/sbras/adv/>.
- [19] *Публичная доска объявлений СО РАН.*–  
<http://www-sbras.nsc.ru/win/anons/>.
- [20] Федотов А.М., Артемов И.А., Ермаков Н.Б., Красников А.А., Потемкин О.Н., Рябко Б.Я., Федотов А.А., Хорев А.Г. *Электронный атлас “Биоразнообразие растительного мира Сибири.” // Вычислительные технологии, т. 3, 5, 1998.*

- [21] Федотов А.М., Рябко Б.Я. *Информационная безопасность полнотекстовых баз данных в среде Интернет.*–  
[<http://www-sbras.nsc.ru/win/elbib/security.html>].
- [22] Федотов А.М., Шокин Ю.И. *Электронная библиотека Сибирского отделения РАН.* // Информационное общество, №2, 2000.
- [23] Холюшкин Ю.П. *Электронные ресурсы сибирской и мировой археологии и этнографии.*–  
[<http://www-sbras.nsc.ru/win/elbib/arch/>].
- [24] Целищев В.В. *Создание нормативно-правовой базы использования информационных ресурсов Отделения.*–  
[<http://www-sbras.nsc.ru/win/elbib/law.html>].
- [25] Шокин Ю.И. *“Информационные ресурсы Сибирского отделения РАН”.*– Интеграционный проект СО РАН №88 (2000)  
[<http://www-sbras.nsc.ru/win/elbib/InfRes.html>].
- [26] Шокин Ю.И., Белов С.Д., Бредихин С.В., Мушер С.Л., Никульцев В.С., Федотов А.М. *“Сеть Интернет Новосибирского научного центра”.*–  
[<http://www-sbras.nsc.ru/win/nsc-net/nsc.html>].
- [27] Шокин Ю.И., Федотов А.М. *Информационные технологии Internet* // Вычислительные технологии, т. 2, с 3, 1997.
- [28] Шокин Ю.И., Федотов А.М. *Распределенные информационные системы* // Вычислительные технологии, т. 3, с 5, 1998.
- [29] Шокин Ю.И., Федотов А.М. *Информационная система поддержки исследований в области математики.*–  
[[http://www-sbras.nsc.ru/mathpub/math\\_www.html](http://www-sbras.nsc.ru/mathpub/math_www.html)].
- [30] Шокин Ю.И., Федотов А.М. *Библиотека, работающая круглосуточно* // ЭКО, №6, 2000.
- [31] Шокин Ю.И., Федотов А.М., Богомяков П.А. *Электронные журналы по математике (на примере электронной версии журнала “Вычислительные технологии”.*–  
[<http://www-sbras.nsc.ru/win/elbib/journals/>].
- [32] Шокин Ю.И., Федотов А.М., Богомяков П.А. *Электронные журналы по математике* // Доклад на конференции “Электронные издательские системы для науки и образования” (EPS4SE-98).  
[<http://www-sbras.nsc.ru/ws/elpub98/577/>].
- [33] *Электронная библиотека Сибирского отделения РАН (проект).*–  
[<http://www-sbras.nsc.ru/win/elbib/>].
- [34] *Электронные версии журналов издательства СО РАН.*–  
[<http://www-psb.ad-sbras.nsc.ru/elversw.htm>]