

Московским государственный университет им. М.В. Ломоносова
Южный Федеральный Университет
Институт вычислительной математики РАН

Труды Всероссийской научной конференции

**Научный сервис в сети ИНТЕРНЕТ:
многоядерный компьютерный мир.
15 лет РФФИ**



ИНФОРМАЦИОННАЯ СИСТЕМА ПО БИОЛОГИЧЕСКОМУ РАЗНООБРАЗИЮ ЖИВОТНЫХ РОССИИ В ИНТЕРНЕТЕ

И.С. Смирнов, А.Л. Лобанов, А.Ф. Алимов, О.Н. Пугачев, В.А. Кривохатский, А.Г. Кирейчук, Е.П. Воронина

В настоящее время в мире предпринимаются активные шаги по интеграции данных о видовом биологическом разнообразии в глобальной сети Интернет [1].

Зоологический институт РАН (ЗИН) разрабатывает различные проекты, направленные на поддержку исследований по биологическому разнообразию России, сопредельных территорий и полярных регионов планеты [2, 3, 4, 5, 7].

Уже пять лет на веб-портале ЗИН РАН размещена информационно-поисковая система (ИПС) <Биоразнообразие России> - <БИОДИВ> (BIODIV), начальное развитие которой поддерживалось с 2002 по 2004 год Федеральной целевой научно-технической программой <Исследования и разработки по приоритетным направлениям развития науки и техники> (государственный контракт № 43.073.11.2510) [6]. За год страницы ИПС <БИОДИВ> посещает около 1 миллиона пользователей (<http://www.zin.ru/BioDiv/index.html>).

Любая информация по биоразнообразию в любой ее составляющей части одинаково важна и должна быть одинаково доступна при обращении к ней потенциального пользователя. Самый простой и логичный способ организации информации - это таксономический классификатор, отражающий систематическое положение любого из включенных таксонов и состоящий из частных классификаторов, связанных между собой. Все классификации - авторские; у разных специалистов они отличаются между собой. Основной классификатор проекта <БИОДИВ> включает 5 таксономических царств, объединяющих все живые организмы: бактерии, грибы, растения, протисты и животные. Раздел <Бактерии> связан непосредственно с базой данных (БД) Всероссийской Коллекции Культур Микроорганизмов (ВКМ).

Для 4 других царств живых организмов специально для проекта "БИОДИВ" были разработаны и выставлены в сети Интернет следующие макроклассификации:

1. грибы - 8 макротаксонов и ссылки с них на сайты Ботанического института РАН (БИН) и ВКМ;
2. растения - 15 макротаксонов и ссылки с них на сайты БИН РАН;
3. протисты - 29 типов и собственные разделы сайта для каждого из них на сайтах БИН РАН и ЗИН РАН (оригинальная классификация С.А.Карпова (С.-Петербургский государственный университет) и его коллег);
4. животные - В.В.Малаховым (Московский государственный университет) совместно со специалистами ЗИН РАН разработана новая оригинальная классификация животных - 227 таксонов, включающая 35 типов и 150 классов. Для каждого типа создан собственный иллюстрированный раздел (некоторые разделы доведены полностью или частично до видов).

Для развития только одной ветви классификатора, посвященной царству животных (фауна), в ЗИН РАН задействовано более 50 баз данных, выполненных в формате ZOOCOD. Доведенные до семейств и далее до видов, модельные таксоны (наиболее полно даны классификации пауков, насекомых, рыб) имеют разное назначение, объемы информации, географические границы (Мир, СНГ, Россия) и иллюстративное сопровождение. В дальнейшем предусматривается разработка системы фильтров, позволяющих манипулировать с базами данных таким образом, чтобы отсортировать демонстрируемые таксоны по определенным заданиям или наборам признаков. Пользователь, таким образом, сможет отсортировать, например, из классификации мировой фауны список таксонов фауны России или другого региона.

Для некоторых таксонов от основных классификаторов BIODIV сделаны ссылки на классификаторы других ИПС. Так, классификации некоторых отрядов насекомых, сопровождающиеся коллекционными данными, оригинально располагаются на страницах проекта ZInsecta, но открываются и со страниц ИПС "Биоразнообразие России".

В ходе работы над проектом "BIODIV" было создано 132 базы данных, содержащие сведения о 45 тысячах таксонов разного систематического уровня.

Эта информация доступна на 423 веб-страницах портала ЗИН (HTML- и ASP-страницы).

Страницы проиллюстрированы 1700 фотографиями и рисунками. Специальный фотоальбом включает галерею художественных снимков животных и растений разных авторов.

Общий объем сайта BIODIV (ноябрь 2006 г.) - 77 мегабайт.

В 2006 году новый проект "Разработка Всероссийской Информационной Системы по Биоразнообразию животных России в Интернете (ИСИБР или ZooDiv)" получил поддержку РФФИ по конкурсу "Ориентированные фундаментальные исследования" на развитие зоологической составляющей ИПС, в том числе на базы данных, развиваемые и поддерживаемые в Зоологическом институте РАН (№ 06-04-08020). Разработка проекта предполагала: создание и постоянное обновление в сети Интернет ИПС по биоразнообразию животных России с общей таксономической базой данных свыше 20 тысяч видов, не имеющей аналогов в Интернете, поскольку представление российской фауны на современном уровне под силу только специалистам, в основном ЗИН РАН; создание электронной библиотеки по таксономии и систематике, как одной из составляющих ИСиБР, которая должна значительно обогатить информационную систему;

разработку гипертекстового таксономического справочника, базирующегося на уже имеющихся в системе данных, за счет большого числа перекрестных ссылок облегчающих поиск необходимой информации, а также иллюстрирование таксонов всех уровней.

Гипертекстовый справочник с глобальным поиском по всем таксономическим базам данных ЗИН РАН частично реализован на сайте проекта ZOOINT (http://www.zin.ru/projects/zooint_r/all_db.htm). ИПС ZOOINT - "ветеран" среди информационных проектов ЗИН РАН. Проект "ЗООИИИТ" почти ровесник РФФИ (поддержка была получена в 1993 году) и дал начало практически всем представленным на вебпортале базам данных. Все эти БД построены на одинаковых принципах, с использованием стандарта ZOOCOD, который совершенствовался в рамках проекта ZOOINT. Эти базы данных содержат сведения о десятках тысяч таксонов и включают в настоящее время ~ 25 тысяч записей и рассредоточены по отдельным страницам сайта ЗИН РАН, что затрудняет поиск необходимой информации. Проект ZooDiv, получающий поддержку РФФИ, призван облегчить эту задачу за счет осуществления быстрого поиска на основе единой таксономической базы данных. Первые результаты работы над проектом представлены на сайте программы (<http://www.zin.ru/ZooDiv/index.html>).

Некоторые группы животных, внесенные в общую базу данных представлены списками, составленными для мировой фауны.

Начато создание библиографической базы данных для крупных таксонов животных с дальнейшей трансформацией ее в электронную библиотеку по биоразнообразию.

В настоящий момент фрагмент базы данных, выставленный в Интернет, насчитывает более 8 тысяч записей (http://www.zin.ru/ZooDiv/zdbib_r.htm).

Переход на платформу SQL Server, проводимый в настоящее время, позволит расширить возможности реализуемого проекта: приложения становятся более масштабируемыми и надежными, облегчается процедура обмена данными для разных платформ, а также труд разработчиков, получающих гибкую и современную среду разработки.

В ходе работы над проектом все больше внимания уделяется метаданным и способам описания создаваемого продукта.

Информационно-поисковая система "Информационная система по биоразнообразию животных России" может и должна способствовать решению следующих фундаментальных проблем и задач зоологии: 1) созданию единой классификации животных фауны России с постоянно обновляемыми данными по новейшим ревизиям; 2) выявлению фаунистических особенностей отдельных территорий и акваторий; 3) выяснению происхождения и эволюции фаун отдельных регионов; 4) проблемы распределения по горизонтали и вертикали отдельных представителей фауны и фаунистических комплексов; 5) изучению биоценотических отношений таксонов (животное-животное или животное-растение; паразитарные, симбиотические, трофические и энергетические); 6) исследованию изменений, происходящих в фауне регионов, на основе сравнения современных сборов и старых коллекций, хранящихся в зоологических музеях страны; 7) созданию динамичного экологического мониторинга; 8) созданию ИПС по музейным коллекциям для активно развивающихся отраслей молекулярной биологии - молекулярной генетики и филогенетики.

Необходимость создания единой системы компьютерной классификации животных мировой и региональных фаун с постоянно обновляемыми данными, отражающей современный уровень биологического знания, сейчас уже не оспаривается никем. В рамках бумажной технологии создание подобной информационной системы было практически невозможно. В компьютерной системе можно использовать и обновлять даже несколько классификаций, соответствующих пониманию структуры фауны разными учеными. Тем самым появляется возможность проверять совместимость тех или иных представлений с данными географического распространения, образом жизни и т.д., т.е. верифицировать выводы систематиков. Появление технологии мультимедиа должно еще больше приблизить кропотливые таксономические исследования к конечному потребителю, поскольку создание веб-сайтов и компакт-дисков сокращает время подготовки публикаций и позволяет интерактивно использовать имеющуюся информацию или в сети Интернет, или записанную на лазерный диск. Включение иллюстраций, карт распространения, а в дальнейшем звуковых и видео записей обогатит создаваемую систему и расширит ее возможности и привлекательность.

Сложность задач, на которые наталкиваются специалисты, заставляет многих оставить попытки разработки всеобъемлющей системы учета названий животного мира. В Зоологическом институте, благодаря в основном поддержке РФФИ, ведется работа по созданию зоологических БД и интеграция разрозненной информации в единое информационное пространство.

На первых этапах идеология развития интегрированной информационной системы привела к необходимости организации локальной сети. В дальнейшем были предприняты попытки создания информационного ресурса уже в сети Интернет. К сожалению, медленное внедрение компьютерных технологий в работу зоологов-систематиков (и технически, и психологически), отсутствие нормативных актов, в частности, закона об интеллектуальной собственности, тормозит создание функциональной и общедоступной версии ИСиБР. Реально работающая общеинститутская локальная сеть, активное объединение имеющихся баз данных в Институте и представление интегрированной системы сбора и обработки таксономических и коллекционных данных в сетях Интранет и Интернет, позволяют рассчитывать на создание функциональной информационно-поисковой системы по биологическому разнообразию животных России на портале Зоологического института. На Западе начинают создаваться подобные системы, но в случае европейских государств (Fauna Europaea и др.) - таксономический охват ограничивается Европой, а для глобальных проектов

(Fishbase, GBIF и др.) или характерна специализация на конкретной группе животных, или степень детализации фауны еще не достигла определенного уровня для России и ее морей и вряд ли когда-либо достигнет, так как на Западе катастрофически не хватает специалистов-систематиков.

Существующие компиляции таксономических списков, созданные неспециалистами, скорее вносят путаницу, чем облегчают поиск нужной информации по таксонам.

Новизна данного проекта состоит в соединении усилий всех специалистов-систематиков и новых компьютерных технологий для создания дееспособной и надежной системы описания биоразнообразия животных России с упором на таксономический аспект.

В дальнейшем предполагается обеспечить весь имеющийся массив данных адекватной поисковой системой, завершить работу по наполнению главных ветвей классификатора животного мира, проиллюстрировать основные таксоны и перевести титульные страницы на английский язык.

Поддержка: гранты РФФИ N 05-07-90179-в, 05-07-90354-в и 06-04- 08020, проект "Антарктика" (ФЦП "Мировой океан") и программа "Биоразнообразие".

ЛИТЕРАТУРА:

1. Смирнов И.С., А.Л. Лобанов, А.Ф. Алимов, А.А. Голиков, А.Г. Кирейчук. Международные Интернет-проекты по созданию биологических электронных коллекций // Научный сервис в сети ИНТЕРНЕТ: Труды Всероссийской научной конференции (18-23 сентября 2006 г., г. Новороссийск). - М.: Изд-во МГУ, 2006. с. 216-218.
2. Смирнов И.С., А.Л. Лобанов, А.Ф. Алимов, А.Г. Кирейчук, А.Е. Вахитов. Электронные зоологические коллекции и технологии электронных библиотек // Электронные библиотеки: перспективные методы и технологии, электронные коллекции: Сборник тезисов постерных докладов Восьмой Всероссийской научной конференции (RCDL'2006). Суздаль, 17-19 октября 2006 г. - Ярославль: Ярославский государственный университет им. П.Г. Демидова, 2006: 60-61.
3. Смирнов И.С., Лобанов А.Л., Алимов А.Ф., О.Н. Пугачев, А.Г. Кирейчук, В.А. Кривохатский, Граничин О.Н., Вахитов А.Т. Зоологические электронные публикации: коллекции и идентификационные системы // Интернет и современное общество. Труды IX Всероссийской объединенной конференции, (14-16 ноября 2006 г., г. СПбГУ), Санкт-Петербург, 2006: 115-118.
4. Смирнов И.С., Лобанов А.Л., Пугачев О.Н., Гельтман Д.В. К истории совещаний по биологическим (зоологическим и ботаническим) базам данных и информационно-поисковым системам в СССР и России // "Информационные системы и web-порталы по разнообразию видов и экосистем". Материалы международн. симпозиума. Борок, 28 ноября - 1 декабря 2006 г., Москва, Тов-во науч. изд. КМК, 2006. С. 18-20.
5. Лобанов А.Л., Кирейчук А.Г., Смирнов И.С., Дианов М.Б. Суперсайту веб-портала Зоологического института "Жуки и колеоптерологи" - 7 лет // "Информационные системы и web- порталы по разнообразию видов и экосистем". Материалы международн. симпозиума. Борок, 28 ноября - 1 декабря 2006 г., Москва, Тов-во науч. изд. КМК, 2006. С. 53-55.
6. Пугачев О.Н., Алимов А.Ф., Лобанов А.Л., Кривохатский В.А., Смирнов И.С. Первые итоги разработки информационной системы по биоразнообразию России (BIODIV-ZOODIV) // "Информационные системы и web-порталы по разнообразию видов и экосистем". Материалы международн. симпозиума. Борок, 28 ноября - 1 декабря 2006 г., Москва, Тов-во науч. изд. КМК, 2006. С. 170- 173.
7. Смирнов И.С., Лобанов А.Л., Пугачев О.Н., Алимов А.Ф., Воронина Е.П. Электронные коллекции в зоологии и электронные библиотеки // Электронные библиотеки, 9 (4). 2006. http://www.elbib.ru/index.phtml?page=elbib/rus/journal/2006/p_art4/SLPAV