

Московский государственный университет им. М.В.Ломоносова
Ростовский государственный университет
Институт вычислительной математики РАН

Труды Всероссийской научной конференции

Научный сервис в сети ИНТЕРНЕТ

г. Новороссийск, 22-27 сентября 2003 г.

Издательство Московского университета
2003

ИНФОРМАЦИОННАЯ СИСТЕМА ПО БИОЛОГИЧЕСКОМУ РАЗНООБРАЗИЮ РОССИИ

И.С. Смирнов, А.Л. Лобанов, А.Ф. Алимов, О.Н. Пугачев, В.А. Кривохатский

Число только известных видов живых существ, обитающих на нашей планете превышает 1434700. Для России эта цифра составляет по разным оценкам от 94000 до 97000 видов. Число видов животных фауны России, образцы которых хранятся в Зоологическом институте РАН (ЗИН) более 77600 видов (более 80% от общего числа видов фауны России). Крупнейшие в России и одни из крупнейших в мире фондовые коллекции ЗИНа насчитывают (только определенных) более 30 млн. единиц хранения. В связи с началом работ по созданию в ЗИН АН СССР банков данных по систематике, фаунистике, морфологии и экологии различных групп животных; а также в связи с необходимостью выполнения Институтом функций координатора научно-исследовательских работ по систематике животных СССР 13 июня 1988 года появился приказ о создании межлабораторной группы по учету состава и изученности фауны СССР. Руководителем этой группы был О.А.Скарлато. Этот приказ можно считать началом официальной истории информатизации зоологических исследований в Зоологическом институте.

Неофициальная история компьютеризации исследований в ЗИН РАН началась в 1973 г. с создания системы ввода орнитологических данных по кольцеванию птиц на перфокарты и магнитные ленты для дальнейшей обработки на больших машинах (БЭСМ-6, а позднее ЕС-1430).

В 1978 г. в лаборатории пресноводной гидробиологии начались работы по использованию удаленной ЭВМ для моделирования продукционных процессов в пресноводных водоемах. В 1985 г. на терминале БЭСМ-6 по инициативе А.Ф.Алимова начались работы по созданию зоологических баз данных (БД) и машинным определительным ключам. Результатом деятельности межлабораторной группы по учету состава и изученности фауны явилось создание БД по различным таксонам фауны СССР и России с данными по численности каждого из них (Алимов, Старобогатов, Кержнер, Лобанов, Степаньянц, 1996). Для СССР была получена цифра в 129432 вида, для фауны России - 93980.

В 1989 г. началось широкомасштабное внедрение машинных методов с помощью персональных компьютеров. Это ускорило разработку прикладных программ для зоологов и создало предпосылки для накопления зоологических данных на машинных носителях.

В 1995 г. было положено начало развитию локальной сети Института в рамках проекта ЗООИНТ (ЗООлогические ИНТегрированные системы), поддержанного Российским Фондом Фундаментальных Исследований. В 1997 г. произошло подключение к сети Интернет и началось использование рассеянной там информации через академическую сеть РОКСОН,

которая просуществовала до весны 1998 г. Это позволило начать разработку собственного сайта Института. До этого существовала версия на английском языке, выставленная в 1995 г. на сайте американских коллег в Иллинойсе (США).

В 1999 г. был разработан, с привлечением профессионалов, проект общеинститутской локальной сети и с этого года началось ее интенсивное развитие и широкомасштабное подключение к ней пользователей. В настоящее время к локальной сети и Интернету подключено более 100 компьютеров, включая 4 сервера. В 1999 г. дирекции ЗИНа удалось добиться получения более стабильного выхода в Интернет через компанию Петерлинк и в настоящее время пользователи Института имеют доступ в Интернет по выделенному каналу FrameRelay с пропускной способностью 256 кбит/сек и возобновленный доступ в тестовом режиме через оптоволоконный канал академической сети РОКСОН - 20 Мбит/сек. Эти благоприятные условия привели к созданию десятков тематических сайтов и трансформировали сайт Института в большой веб-портал.

Особую роль в зоологии играют базы данных, ибо вся зоологическая информация опирается на списки таксонов и коллекционных экземпляров, хранить которые логичнее всего в БД (другие (некомпьютерные) варианты создания таких списков сейчас уже можно считать анахронизмом).

На основе развития таксономических БД в Институте расцвел целый букет успешных проектов: Банк данных по кольцеванию птиц (более 4 млн. записей с разнообразными возможностями анализа); Зоологическая ИНТегрированная информационно-поисковая система ЗООИНТ; Информационно-аналитическая система по блохам Земного шара и их паразито-хозяйным связям; Банк данных о паразитах пресноводных рыб, много лет разрабатываемый в лаборатории паразитических червей; Распределенный банк данных по морским животным Арктики и Антарктики "Океан"; Уникальная компьютерная диагностическая система РІСКЕУ, которая позволяет накапливать зоологическую информацию в специализированных БД; Информационно-поисковая система по пресноводным рыбам России, на основе которой уже функционирует сайт и планируется создание электронной монографии; Интерактивный банк данных по коловраткам; Виртуальные коллекции насекомых в сети Интернет (проект ZInsecta) Почти все эти разработки представлены на портале Института - www.zin.ru

Использование уникальных оригинальных данных и мультимедийных технологий на веб-портале ЗИН сделало его одним из самых популярных научных ресурсов Рунета, посещаемым сотни тысяч раз за год. К маю 2003 г. портал насчитывал 38 тематических сайтов, более 7500 веб-страниц и более 11500 изображений с суммарным объемом более 560 Мб. В 2002 г. на портал было осуществлено более 650000 визитов и обработано более 7000000 запросов. Самым популярным на портале ЗИНа был сайт о жуках и колеоптерологах, который считается самым динамичным и привлекающим внимание самых разнообразных групп населения: от школьников и любителей природы до маститых ученых. Над его развитием при поддержке РФФИ трудится большой виртуальный коллектив в России и за рубежом.

В результате осуществленной компьютеризации и информатизации зоологических исследований, в 2002 г. Зоологический институт РАН смог принять участие в конкурсе объявленном Министерством промышленности, науки и технологий, по Блоку 2 "Поисково-прикладные исследования и разработки" в разделе "Технологии живых систем", подразделе "Биология". Выиграв конкурс, 31 января 2002 г. ЗИН заключил государственный контракт на выполнение научно-исследовательских и опытно-конструкторских работ по теме "Информационная система по биоразнообразию". Соисполнителями стали Институт проблем экологии и эволюции РАН (ИПЭЭ), Ботанический институт РАН (БИН) и Институт цитологии и генетики СО РАН (ИЦГ). В 2003 г. к этому проекту подключился Институт биохимии и физиологии микроорганизмов РАН (ИБФМ).

Основная цель проекта "Информационная система по биоразнообразию" (ИСБР) - создание комплекса программных средств и БД для работы с классификацией животного и растительного мира России, которые станут основой формирования информационно-поисковой системы (ИПС) по биоразнообразию России, поддерживающей неоднородные

коллекции распределенных информационных ресурсов ЗИНа, БИНа, ИПЭЭ, ИЦиГа и ИБФМ, содержащих сведения систематического, коллекционного и экологического характера.

Основные задачи ИСБР - разработка стандартов, форматов и методологии создания распределенной БД по биологическому разнообразию в рамках российского сегмента глобальной сети Интернет, и создание пилотной информационной сети по видовому разнообразию России, объединяющей БД, поддерживаемые институтами-соисполнителями, по видовому составу всех таксонов (микроорганизмов, растений и животных) и метаданные о коллекциях, хранящиеся в этих учреждениях.

Исходными данными для проведения работы служат результаты исследований структурных подразделений ЗИН, БИН, ИПЭЭ, ИБФМ и ИЦиГ в области биологии, в том числе уже созданный и частично опубликованный ими научный продукт в области систематики животных и растений.

Объектами создаваемых БД являются названия бактерий, простейших, грибов, растений и животных, обитающих на территории России и в сопредельных регионах. Особое место в этих исследованиях занимает создание БД по фауне Антарктики. На первом этапе работы над проектом коллективами перечисленных выше институтов РАН сделаны следующие шаги в осуществлении проекта:

Разработаны стандарты построения системы и создания и ведения классификаторов (специализированных БД) названий животных и растений для зоологических и ботанических банков данных (ZOOCOD и PLANTCOD).

Написан пакет оригинальных программ ZooDiv (BioDiv), предназначенный для создания БД по таксономии, номенклатуре и географическому распространению животных и растений, а также для иерархического заполнения таксономических БД. Удобные программы - хорошее подспорье в сложном процессе создания биологических баз данных.

Открыт в Интернете сайт ИСБР с "систематическим древом" высших таксонов всех царств животного и растительного мира с демонстрацией возможности выхода из отдельных его ветвей на следующие систематические уровни вплоть до конкретных биологических видов и даже на уникальные экземпляры коллекций музеев страны (<http://www.zin.ru/BioDiv/index.html>).

Работа поддержана грантами РФФИ NN 02-07-90217, 02-07-90105 и 01-07-90284, проектом "Информационная система по биоразнообразию", программами "Биоразнообразие" и "Антарктика".