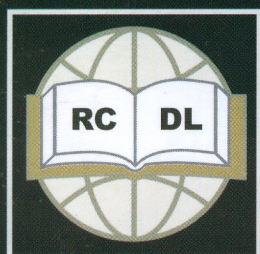


**ЭЛЕКТРОННЫЕ БИБЛИОТЕКИ:
ПЕРСПЕКТИВНЫЕ МЕТОДЫ И ТЕХНОЛОГИИ,
ЭЛЕКТРОННЫЕ КОЛЛЕКЦИИ**

Труды XI Всероссийской научной конференции

Петрозаводск, Россия



2009

Petrozavodsk, Russia

Proceedings of the XIth All-Russian Research Conference

**DIGITAL LIBRARIES:
ADVANCED METHODS AND TECHNOLOGIES,
DIGITAL COLLECTIONS**

81

Российская академия наук
Российский фонд фундаментальных исследований
Карельский научный центр РАН
Институт прикладных математических исследований
Петрозаводский государственный университет
Институт проблем информатики РАН
Московская секция ACM SIGMOD

Электронные библиотеки: перспективные методы и технологии, электронные коллекции

XI Всероссийская научная конференция RCDL'2009

Петрозаводск, 17–21 сентября 2009 г.

Труды конференции

Russian Academy of Sciences
Russian Foundation for Basic Research
Karelian Research Center of the Russian Academy of Sciences
Institute of Applied Mathematical Research
Petrozavodsk State University
Institute of Informatics Problems of the Russian Academy of Sciences
Moscow ACM SIGMOD Chapter

Digital Libraries: Advanced Methods and Technologies, Digital Collections

XI All-Russian Research Conference RCDL'2009

Petrozavodsk, September 17–21, 2009

Proceedings of the Conference

Петрозаводск • 2009

УДК 002:004.9 (063)

ББК 78.3

Э 45

Электронные библиотеки: перспективные методы и технологии, электронные
Э 45 коллекции: Труды XI Всероссийской научной конференции RCDL'2009
(Петрозаводск, Россия, 17–21 сентября 2009 г.). – Петрозаводск: КарНЦ РАН, 2009. –
487 [1] с.: ил.

ISBN 978-5-9274-0374-5

Электронные библиотеки – область исследований и разработок, направленных на развитие теории и практики обработки, распространения, хранения, анализа и поиска цифровых данных различной природы. Основная цель серии конференций RCDL (<http://rcdl.ru/>) заключается в формировании сообщества специалистов России, ведущих исследования и разработки в области электронных библиотек и близких областях. Всероссийская научная конференция 2009 г. (RCDL'2009) является одиннадцатой конференцией по данной тематике (1999 г. – Санкт-Петербург, 2000 г. – Протвино, 2001 г. – Петрозаводск, 2002 г. – Дубна, 2003 г. – Санкт-Петербург, 2004 г. – Пушкино, 2005 г. – Ярославль, 2006 г. – Суздаль, 2007 г. – Переславль-Залесский, 2008 г. – Дубна).

Настоящий сборник включает тексты докладов, коротких сообщений и стендовых докладов, отобранных Программным комитетом RCDL'2009 в результате проведенного рецензирования.

Конференция организована при поддержке Российского фонда фундаментальных исследований и Российской академии наук.

Digital Libraries: Advanced Methods and Technologies, Digital Collections:
Proceedings of the XI All-Russian Research Conference RCDL'2009 (Petrozavodsk,
September 17–21, 2009). – Petrozavodsk: KRC RAS, 2009. – 487 [1] p.: il.

ISBN 978-5-9274-0374-5

Digital Libraries is a field of research and development aiming to promote the theory and practice of processing, dissemination, storage, search and analysis of various digital data. The purpose of the series of All-Russian Research Conferences on Digital Libraries (RCDL, <http://rcdl.ru/>) is to stimulate consolidation of the Russian digital libraries community and encourage research in this field. All-Russian Research Conference RCDL'2009 is the Eleventh Conference on this subject (1999 – St. Petersburg, 2000 – Protvino, 2001 – Petrozavodsk, 2002 – Dubna, 2003 – St. Petersburg, 2004 – Puschino, 2005 – Yaroslavl, 2006 – Suzdal, 2007 – Pereslavl-Zalessky, 2008 – Dubna). The RCDL'2009 Proceedings include the texts of reports, short papers and posters selected by Program Committee RCDL'2009.

The conference was organized with the support of the Russian Foundation for Basic Research and the Russian Academy of Sciences.

Виртуальные коллекции животных и интерактивные определители биологических объектов*

© А.Г.Кирейчук, А.Л.Лобанов, И.С.Смирнов, А.Т.Вахитов, Е.П.Воронина, О.Н.Пугачев

Учреждение Российской Академии наук Зоологический институт РАН, Санкт-Петербург
smiris@zin.ru

Аннотация

Разработка виртуальных коллекций, определительных ключей и систем экологического мониторинга получает за последние годы более интенсивное развитие, связанное с совершенствованием компьютерных технологий и возрастающей ролью оценки и сохранения биоразнообразия под действием антропогенной нагрузки. В связи с распространением операционной системы Windows и Интернета, становится актуальным перевод программ ввода и работы с ключами в эту среду, а также представление данных систем в Интернете для более оперативного использования в целях управления ресурсами и в образовании.

1 Введение

В своей книге «Электронные библиотеки» Вильям Армс дает: «Информационное определение ЭБ: управляемая коллекция информации в совокупности с соответствующими сервисами, причем информация хранится в цифровых форматах и доступна по сети» [2, стр. 10]. Большинство компьютерных разработок в Зоологическом институте РАН тесно связано с идеологией электронных библиотек или коллекций [34].

Важным аспектом зоологических исследований является использование реальных коллекций, наличие возможности перепроверить то или иное описание вида, выявить новый признак, рассмотреть детали строения под микроскопом или даже просто удостовериться в наличии именно данного вида в данной точке планеты, что особенно становится актуальным в свете интенсивного распространения чужеродных и, особенно, вредоносных видов (классические примеры: колорадский жук и двустворчатый моллюск-дрейссена). Можно еще упомянуть недавние катастрофические последствия заносов в США из Китая жука-усача *Anoplophora*

glabripennis и жука-златки *Agrilus planipennis*. Несмотря на усилия множества служб и специалистов, эти два вида принесли многомиллионный ущерб и продолжают расселяться. Виртуальные коллекции позволяют ускорить привлечение настоящих коллекций для анализа материала в данных ситуациях, а в ряде случаев и заменить их.

Во всех научно-исследовательских зоологических музеях, где хранятся собрания образцов животных, как ныне живущих, так и вымерших, которые собраны и поддерживаются учеными для более полного изучения и документирования биологического разнообразия, ведутся традиционные каталоги и картотеки и во многих начинают приступать к созданию или электронных каталогов, или специализированных баз данных, или даже информационно-поисковых систем [29, 31, 32]. Появляются компьютерные каталоги, которые выставляются в глобальной информационной сети. Перевод списков экспонатов, а затем и музейных каталогов, в цифровую форму и создание коллекционных баз данных служит первым шагом на пути создания виртуальных коллекций. Вторым шагом в создании электронных или цифровых коллекций является накопление изображений экспонатов и создание электронных фотогалерей и фотоальбомов для зоологических образцов [19, 25]. Соединение цифровых коллекций с информацией о музее, его истории, кураторах, специалистах (которые определяли материал), постепенно приводит к созданию виртуального музея [8, 21, 26-29].

Интернет-технологии ускорили процессы представления разнообразной музейной информации.

2 Электронные публикации

Среди компьютерных технологий определенное место занимают электронные публикации на компакт-дисках CD-ROM и в виде файлов на серверах сети Интернет [1, 22]. Первая серьезная электронная публикация на веб-портале ЗИН - интерактивный каталог коловраток пресных вод Северо-запада России [9]. Исследованиями авторов каталога охвачены около 100 озер, более 70 рек, свыше 10 водохранилищ и другие водоемы.

Труды 11^й Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» - RCDL'2009, Петрозаводск, Россия, 2009.

В ЗИНе разработана концепция построения компьютерных зоологических руководств типа «Фауна России» и «Определители по фауне России» [3, 12-14, 17, 18, 24]. Эта концепция реализована в пакете программ, получившем название DIALOBIS (DIALOGue Biological Identification Systems). В полном объеме эта идеология и оригинальный пакет программ, созданный сотрудниками ЗИНа, впервые на практике использованы немецкой фирмой «dialobis edition» для подготовки серии биологических изданий на лазерных дисках [16 и др.].

Функциональная основа DIALOBIS - многоаспектное представление об исходном наборе таксонов, который можно исследовать и редуцировать разными способами. Для этого используются специализированные прикладные программы, одновременно являющиеся инструментами исследования какого-то аспекта и фильтрами. Эти программы получают на входе набор таксонов (либо в виде копии исходного полного набора, либо как результат одной из предыдущих редуциций, сам текущий набор или то его подмножество, которое проходит через все фильтры), а на выходе могут редуцировать его в соответствии с желанием пользователя. Такие действия отдельных программ не сразу влияют на состав текущего набора, а накапливаются в виде совокупности фильтров, информация о которых постоянно выводится на экран главной управляющей программой пакета. Такой пакет программ и управляемая им информация получили название - гипербаза данных. Гипербаза дает возможность пользователю с помощью отдельных шагов многоаспектного поиска активно конструировать требуемый для детального изучения массив данных. Широкое применение такой технологии позволяет создавать очень эффективные зоологические электронные руководства. К сожалению, на практике бумажные технологии пока преобладают в сознании над электронными. С использованием базы данных по рыбам, разработанной Е.П.Ворониной, было подготовлено и опубликовано несколько каталогов фондовой коллекции Зоологического института [4, 5, 6, 7].

Создание биологических (зоологических, ботанических) информационно-поисковых систем - способствует быстрой и качественной экологической экспертизе, как в интересах рационального природопользования, так и при обеспечении экологической безопасности, например, при освоении различных месторождений полезных ископаемых. Прошедшее в 2004 году международное совещание по информационным системам, применяемым в изучении морского биоразнообразия, показало значительную востребованность ИПС в биологии и, в частности, в морской экологии [30, 33, 35, 36].

Видовой состав организмов той или иной территории или акватории необходимо знать не только из чисто академического интереса. Очень

остро последнее время стоит проблема видов-вселенцев, которые способны серьезно нарушать экологическую обстановку и вести к колоссальным экономическим потерям. Эта проблема может быть решена, только тогда, когда хорошо известны местные фауна и флора, т.е. состав населения водоемов, и тогда, когда на протяжении многих лет можно проследить динамику появления чужеродного вида и захват им новых участков и мест обитания. Без современных информационных технологий механизм подобного мониторинга создать невозможно, учитывая современные средства передвижения и пассивного переноса нежелательных чужеродных видов.

2.1 Электронные справочные пособия и определители. Краткая предистория

Важным моментом развития коллекционных баз данных и ИПС является создание на их основе полноценных справочных пособий и определителей. Своевременная идентификация тех же чужеродных вселенцев и принятие соответствующих мер, предотвращающих их нежелательное распространение, зачастую может сэкономить колоссальные финансовые и человеческие ресурсы.

Первые опыты применения ЭВМ для диагностики таксономической принадлежности биологических объектов были осуществлены в середине 60-х годов, когда ЭВМ еще были для биологов экзотической техникой. Интересно, что на заре компьютерной биологической идентификации отставания отечественных исследований от зарубежных как по качеству, так и по количеству диагностических систем практически не было. Позднее за рубежом, благодаря развитию компьютерной индустрии, в этой области произошел всплеск активности американских, английских и австралийских биологов, который привел к созданию сразу нескольких машинных систем, и таким образом у зарубежных коллег появился количественный перевес. Пиком этой активности можно считать выход в 1975 году сборника "Biological Identification With Computers".

К началу 80-х годов сложилось представление о специфических функциях компьютерных диагностических систем. Их полный набор включал:

1. Предварительную обработку диагностической информации о таксонах.
2. Накопление, хранение и анализ информации на машинных носителях.
3. Диалоговый диагноз с оптимизацией со стороны ЭВМ.
4. Автоматическое составление оптимизированных текстовых определителей.
5. Анализ параметров готовых определителей.

Первая в СССР действующая полная машинная диагностическая система "Диагностика-1" (т.е. выполняющая все перечисленные выше основные специфические функции компьютерных диагностических программ) была создана

А.Л.Лобановым в 1974 году на ЭВМ "Наири-2". Затем эта система расширялась, модернизировалась в соответствии с растущими возможностями доступных биологам ЭВМ, - сначала ЕС "Ряд-1", затем БЭСМ-6, СМ-4, СМ-1420 и, наконец, IBM PC. Начиная с пятой версии системы ("Диагностика-5") программы разрабатывались только для IBM PC, они использовали базы данных формата DBF и состояли из модулей, написанных на языке Фортран-88 и на внутреннем языке СУБД FoxPro. С 1992 г. к работе над диагностическими компьютерными системами подключился М.Б.Дианов. Усовершенствованная версия системы "Диагностика-5" получила новое название "BiKey5", а входящая в нее специализированная диалоговая программа была названа "PicKey". Последний вариант комплексной системы "BiKey7" (1996-1998 гг., языки программирования FoxPro и Fortran) был создан совместно с М.Б. Диановым [21].

Разработка последней версии системы (Bikey8b/PicKey8b) закончена в 2005 г. Все блоки системы программировались на языке C++ для Windows и использовали базы данных формата DBF. Программа PicKey8b является одной из лучших в мире в своем классе (интерактивные диагностические программы без специальной ориентации на использование в сети Интернет). С помощью перечисленных программ были созданы определители, включающие до 400 таксонов самых разнообразных организмов (жуки - А.Л.Лобанов, деревья - Б. и В.Шиловы, медузы - С.Д.Степаньянц, офиуры - И.С.Смирнов, циклопы - В.Р.Алексеев, нематоды - А.Ю.Рысс и др.). О программе PicKey можно узнать на специальном веб-сайте: <http://www.zin.ru/projects/pickey>. Этой программе и созданным с ее помощью определителям посвящены десятки публикаций. Ссылки на них имеются на сайте PicKey и веб-странице: http://www.zin.ru/Animalia/Coleoptera/rus/all_ref.htm.

Эти определители охватывали лишь небольшое число таксонов и работали на отдельных персональных компьютерах. Для более совершенного механизма определения необходимо было создать алгоритм, позволяющий производить диагностику значительно большего числа предполагаемых объектов идентификации.

2.2 Интернет-определители

С появлением и широким распространением Интернета появилась идея написания программы, которая бы давала возможность дистанционного определения [23].

В ходе подготовки к реализации задуманного, при поддержке РФФИ (грант N 02-07-90105), в течение 2002-2004 гг. был создан электронный Атлас жуков России и сопредельных стран, который размещен на сайте Зоологического института РАН: <http://www.zin.ru/Animalia/Coleoptera/index.htm>. В 2005 г. была получена поддержка РФФИ специально на разработку программного обеспечения к многоходовым политомическим

определителям с использованием сети Интернет. Работа над этим проектом стала возможной благодаря участию специалистов в области сетевых компьютерных систем на основе баз данных - О.Н.Граничина и А.Т.Вахитова. Так родился проект WebKey-X:

<http://www.zin.ru/projects/webkey-x/index.html>, в котором участвовали специалисты по разным группам насекомых и офиурам (грант РФФИ N 05-07-90179a, руководитель А.Г. Кирейчук). Связь определителей с атласами, а также необходимыми ссылками на литературные источники по каждому таксону позволяет достичь высокой эффективности на всех этапах определения и проверки (уточнения) его надежности. В настоящее время совершенствуются методы диагностики по медузообразным (С.Д. Степаньянц), иглокожим (А.В. Смирнов, И.С. Смирнов), насекомым (А.Г. Кирейчук, А.Л. Лобанов, А.И. Халаим) (грант РФФИ N 09-04-00789a).

Проведенный в 2005 г. обзор существующих диагностических систем, показал наличие довольно развитых программных продуктов, обеспечивающих определение различных групп животных и растений [23]. На основе анализа компьютерных систем создана таблица, идея и часть содержания которой заимствованы у M. Dallwitz'a: <http://delta-intkey.com/www/comparison.htm>. Таблица сильно модифицирована: в нее добавлены новые характеристики (они предложены А.Лобановым и Д.Дмитриевым и выделены в раздел "New features") и ряд отсутствовавших в ней программ (в том числе наши разработки - PICKEY и WebKey-X, а также программа Д.Дмитриева - 3I): <http://www.zin.ru/projects/webkey-x/index.html>.

Таблица активно дополняется оценками для новых характеристик и для добавленных программ. В таблицу включены следующие программы: [WebKey-X](#), [PICKEY](#), [3I \(by D.Dmitriev\)](#), [Flora Search](#), [Intkey](#), [LucID](#), [MEKA](#), [PollyClave](#), [IdentifyIt](#), [Visual Key](#), [NaviKey](#), [XID](#), [DAP](#), [DAWI](#).

На основе полученных данных шла разработка структуры типового интерактивного определителя (компьютерного ключа) и создание пилотных вариантов различных определителей. Принципиальные свойства интерактивного определителя или компьютерного ключа могут быть условно разделены на 2 группы - структурные и динамические. Структурные свойства - особенности структуры базы данных ключа. Динамические свойства - специфические особенности шага идентификации, представляющего собой итеративно повторяющийся цикл диагноза (под шагом, более точно, понимается набор процедур, обычно включающий выбор признака, ввод в компьютер информации о состоянии признака и получение ответа компьютера с текущим набором таксонов, обладающих данным состоянием признака.). Главные структурные свойства ключа: число входов и число состояний в одном признаке. Важнейшее свойство - число входов для начала каждого шага диагноза.

Таблица.

ОПРЕДЕЛИТЕЛИ ХАРАКТЕРИСТИКИ	MAX_VALUE	WEBKEY_X	PICKEY	DMTR_3I	FLORA_SRCH	INTKEY	LUCID	MEKA	POLLYCLAVE	IDENTIFYIT	VISUAL_KEY	NAVIKEY	XID	DAP	DAWI
	НОВЫЕ ХАРАКТЕРИСТИКИ														
Использование традиционных баз данных	16	16	16	16	0	8	0	0	0	0	0	0	0	0	0
Указание на недопустимые состояния вместо удаления	4	0	4	4	0	0	0	0	0	0	0	0	0	0	0
Демонстрация числа таксонов с определенным состоянием признака	2	0	0	2	0	0	0	0	0	0	0	0	0	0	0
Возможность различных путей определения на основе одной матрицы данных	2	0	2	2	0	0	0	0	0	0	0	0	0	0	0
Многоязычные ключи (весь текст на нескольких языках)	8	0	8	8	0	0	0	0	0	0	0	0	0	0	0
Хранение, использование и поиск синонимических названий	2	0	2	2	0	0	0	0	0	0	0	0	0	0	0
Использование на компьютерах MAC и IBM	4	0	0	4	0	0	0	0	0	0	0	0	0	0	0
Поиск в Интернете дополнительной информации по каждому таксону	4	0	0	4	0	0	0	0	0	0	0	0	0	0	0
Хранение и поиск библиографических источников по каждому из таксонов	2	0	0	2	0	0	0	0	0	0	0	0	0	0	0
Хранение и поиск информации по кормовым растениям	2	0	0	2	0	0	0	0	0	0	0	0	0	0	0
Хранение и поиск данных по распространению с использованием карт	4	0	4	4	0	0	0	0	0	0	0	0	0	0	0
ПРЕИМУЩЕСТВА ПЕРЕД ТРАДИЦИОННЫМИ КЛЮЧАМИ															
Использование признаков с нефиксированными значениями	8	0	0	8	0	8	8	8	8	8	8	8	8	0	0
Изменение и удаление признаков	4	4	4	4	0	4	4	4	4	4	0	4	4	0	0
Допустимость ошибок определяющего	8	0	8	8	0	8	8	0	0	0	0	0	0	0	0
Указание программы на ошибки пользователя	4	0	0	4	0	4	0	0	0	0	0	0	0	0	0
Оценка неопределенности	8	0	0	8	0	8	8	4	8	0	8	0	4	0	0
Использование непрерывных числовых признаков	8	0	0	8	0	8	8	0	8	0	0	8	0	0	0
ПОМОЩЬ В ВЫБОРЕ ПРИЗНАКОВ															
Отображения иерархии признаков	8	0	0	0	0	0	0	0	0	0	8	0	0	0	0
Указание оптимальных признаков	8	8	8	8	0	8	4	0	4	8	8	0	8	0	0
Указание признаков для выделения таксона	4	4	4	0	0	4	0	0	0	0	0	0	2	0	0
Оптимальные пути определения	2	0	0	0	0	0	2	0	0	0	0	0	0	0	0
Удаление избыточных признаков	2	0	0	2	0	2	2	2	0	2	2	2	2	0	0
Удаление избыточных состояний признаков															
Оценка надежности признаков	4	0	0	4	0	4	0	0	4	0	0	0	4	0	0
Оценка надежности состояний признаков	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Составление списка признаков	4	4	4	4	0	4	4	0	2	0	4	0	0	0	0
ЗАПИСЬ И ОПРЕДЕЛЕНИЕ ЦЕННОСТИ ПРИЗНАКОВ															
Сохранение невыясненных признаков	16	0	0	16	0	16	16	16	16	0	16	16	8	0	0
Использование взаимосвязи признаков	4	0	0	0	0	4	2	0	0	0	0	0	0	0	0
Автоматический контроль признаков	4	0	0	0	0	4	0	0	0	0	0	0	0	0	0
Использование интервалов для числовых признаков	4	0	0	0	0	4	4	0	4	0	0	4	0	0	0
Возможность текстового заполнения признаков	2	2	2	0	0	2	0	0	1	0	0	1	0	0	0
Специальные оценки ключей	4	0	0	0	0	0	2	0	0	0	0	0	0	0	0
Вероятностное определение	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Указание неприемлемых комбинаций состояний признаков	2	0	0	0	0	2	0	0	2	0	1	1	0	0	0
Расширенные диапазоны числовых признаков	2	0	0	0	0	1	0	0	1	0	0	0	0	0	0
Допустимость неизвестных состояний признаков	2	0	0	0	0	0	2	2	0	0	0	0	0	0	0
Точные признаки	2	0	0	0	0	2	0	0	0	0	0	0	0	0	0
Определение ценности признака	2	0	0	0	0	2	2	0	0	0	0	0	0	0	0

ИСПОЛЬЗОВАНИЕ ГРУПП ПРИЗНАКОВ															
Поименованные наборы признаков	4	0	4	4	0	4	4	4	0	0	0	0	4	0	0
Общие наборы признаков	4	0	4	4	0	4	4	4	0	0	0	0	0	0	0
Конкретные наборы признаков	2	0	2	0	0	2	0	0	0	0	0	0	1	0	0
Поименованные наборы (группы) таксонов	4	0	4	0	0	4	0	0	0	0	4	0	0	0	0
Общие группы таксонов	4	0	4	0	0	4	0	0	0	0	0	0	0	0	0
Конкретные группы таксонов	2	0	2	0	0	2	0	0	0	0	0	0	0	0	0
ИНТЕРПРЕТАЦИЯ ПРИЗНАКОВ															
Комментарии к признакам	4	0	4	4	0	4	4	0	0	2	0	0	4	0	0
Глоссарии (словари признаков)	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Иллюстрации признаков	4	4	4	4	0	4	2	0	4	2	4	0	4	0	0
Выбор состояния признака по иллюстрации	4	4	4	4	0	4	2	0	0	2	4	0	0	0	0
ИЗОБРАЖЕНИЯ И ЗВУКИ															
Иллюстрации таксонов	4	4	4	4	0	4	4	4	2	4	4	4	2	0	0
Иллюстрации таксонов по группам	4	0	4	4	0	4	4	0	0	4	0	0	0	0	0
Произвольный показ иллюстраций	4	0	0	0	0	4	0	0	0	0	2	0	0	0	0
Текст с иллюстрациями	2	0	0	2	0	2	2	0	0	0	2	0	2	0	0
Звуки	2	0	2	2	0	2	2	0	0	2	2	0	0	0	0
Видеоизображения	2	0	2	2	0	2	2	0	0	2	2	0	0	0	0
Определение без рисунков	2	0	0	2	0	2	0	0	0	0	2	0	0	0	0
СОПРЯЖЕННЫЕ КЛЮЧИ															
Интегральные иерархические ключи	2	0	0	2	0	2	0	0	0	0	0	0	0	0	0
Самостоятельные иерархические ключи	2	0	0	2	0	2	2	0	0	0	0	0	0	0	0
ПОИСК ИНФОРМАЦИИ															
Поиск по названиям таксонов	2	0	0	2	0	2	1	0	1	0	0	0	0	0	0
Составление описаний таксонов	2	0	2	2	0	2	2	0	2	2	2	2	2	0	0
Различия между таксонами	4	0	0	0	0	4	2	0	0	2	0	0	0	0	0
Сходство таксонов	2	0	0	0	0	2	1	0	0	1	0	0	0	0	0
Составление дифференциальных описаний	4	0	0	0	0	4	0	0	0	0	0	0	0	0	0
Поиск таксонов по сочетаниям признаков	2	0	0	0	0	2	1	0	0	0	0	0	0	0	0
Контроль правильности значений состояний признаков	2	0	0	0	0	2	1	1	0	0	0	0	1	0	0
Распределение признаков по ценности	2	0	0	0	0	2	0	0	0	0	0	0	1	0	0
Наиболее сходные таксоны	2	0	0	0	0	0	2	0	0	0	0	0	0	0	0
Текстовые файлы, прикрепленные к таксонам	2	0	0	2	0	2	2	0	0	2	0	0	0	0	0
СОВМЕСТНОЕ ИСПОЛЬЗОВАНИЕ ДАННЫХ															
Импорт данных в формате DELTA	2	0	0	0	0	2	1	0	2	0	0	2	0	0	0
Экспорт данных в формате DELTA	2	0	0	0	0	2	1	0	2	0	0	2	0	0	0
Вывод данных	2	0	2	0	0	2	0	0	0	0	0	0	0	0	0
Связь с составлением описаний таксонов	2	0	0	2	0	2	0	0	2	0	0	2	0	0	0
Связь с генерацией одноходовых ключей	2	0	2	0	0	2	0	0	2	0	0	2	0	0	0
Связь с классификацией	2	0	0	0	0	2	0	0	2	0	0	2	0	0	0
УДОБСТВА ДЛЯ ПОЛЬЗОВАТЕЛЯ															
Интерактивная справка	2	0	0	1	0	2	1	0	1	1	0	1	2	0	0
Использование командных файлов или макросов	2	0	0	0	0	2	1	1	0	0	0	0	0	0	0
Использование инструментальной панели по выбору	4	0	0	0	0	4	0	0	0	0	0	0	0	0	0
Доступ к тексту программы	2	0	0	0	0	2	2	0	0	0	0	0	0	0	0
Папки с документацией	2	2	2	0	0	2	0	0	0	0	0	0	0	0	0
Неограниченный объем данных	4	4	4	4	0	4	4	4	4	2	4	4	4	0	0
Неограниченная длина полей	2	2	2	2	0	2	1	0	2	0	2	2	2	0	0
Отсутствие специальных требований к памяти компьютера	2	0	0	2	0	2	2	2	2	2	2	2	2	0	0
Скорость ответа программы на действия пользователя	4	0	0	4	0	4	0	4	0	4	4	0	4	0	0
Возможность работы в Интернете	8	8	4	8	0	4	0	0	4	0	4	2	0	0	0
Отсутствие необходимости инсталляции программ	2	2	0	2	0	2	0	2	0	0	2	0	2	0	0

Существуют одновходовые (пользователю предлагается один признак) и многовходовые (предлагается несколько признаков для выбора наиболее надежного и удобного) ключи. Число состояний признака принципиально менее важное свойство. Существуют дихотомические (каждый признак имеет только 2 состояния) и политомические (признак включает 3 и более состояний) ключи. Практически важна способность оперировать изображениями признаков и состояний признаков; существуют ключи, управляемые образами (изображения признаков и их состояний - экранные кнопки управления диагнозом) и ключи, управляемые словесными формулировками (используются альтернативные текстовые формулировки). Изображения дают возможность мгновенно уяснить признак и верно выбрать его состояние; текстовая формулировка требует времени для понимания, сравнения, выбора соответствующего признака и состояния. Повышает возможности определительной программы способность оперировать количественными признаками; возможность фильтровать набор таксонов по диапазону значений количественного признака может существенно сократить путь диагноза [20].

В дальнейшем была предпринята попытка представить себе идеальный компьютерный определитель или «определитель-идеал», к которому нужно стремиться при разработке подобных систем. На основе сравнения всех свойств компьютерных определителей, созданных к настоящему времени как за рубежом, так и в СНГ, можно смоделировать оптимальный идентификационный ключ. Это интерактивный, многовходовой, политомический, управляемый изображениями ключ со следующими динамическими свойствами: ранжированием признаков на каждом шаге в зависимости от их диагностической ценности; возможностью видеть названия и изображения таксонов текущего набора, а также значения вероятности их идентичности определяемому объекту; возможностью возврата на один или несколько шагов диагноза для коррекции ошибок определяющего; возможностью выбрать несколько признаков в каждом шаге и отметить "невозможные" состояния признаков; активным оперированием как качественными, так и количественными признаками с возможностью использования для последних математических функций (диапазон минимум-максимум, среднее, формулы дискриминантного анализа). Позже прибавилась возможность интерактивного представления определителей или ключей в глобальной сети Интернет. По этим направлениям и ведутся поисковые работы, так как ни одна из существующих систем пока не обладает полным набором описанных свойств [10, 11, 15, 20, 23].

Литература

- [1] Алимов А.Ф., Смирнов И.С., Рысс А.Ю., Дианов М.Б., Лобанов А.Л., Голиков А.А. Современные биологические электронные публикации: коллекции, идентификационные системы и базы данных // Информационные и телекоммуникационные ресурсы в зоологии и ботанике. Труды 2-го международного симпозиума. 2001. С. 13-19.
- [2] Армс В. Электронные библиотеки. (Перевод С.А.Арнаутова). ПИК ВИНТИ, 2001. - 275 с.
- [3] Дианов М.Б., Лобанов А.Л. PISKEY - Программа для определения организмов с интерактивным использованием изображений // В: С.Степаньянц, А.Лобанов, М.Дианов, ред., Базы данных и компьютерная графика в зоологических исследованиях // Труды Зоологического института РАН, 1997. Т. 269. С. 35-39.
- [4] Каталог фондовой коллекции Зоологического института РАН. Класс Костистые рыбы (Osteichthyes). Отряд Камбалообразные (Pleuronectiformes). Сост. Воронина Е.П., Волкова Г.А. Исследования фауны морей. Т. 55(63). - СПб., 2003. 198 с.
- [5] Каталог фондовой коллекции Зоологического института РАН. Класс Костистые рыбы (Osteichthyes). Отряд Скорпенообразные (Scorpaeniformes). Подотряд Cottoidei. Часть I. Сиделева В.Г., Неелов А.В., Воронина Е.М., Волкова Г.А. Исследования фауны морей. Т. 57(65). - СПб., 2006. С. 1-223.
- [6] Каталог фондовой коллекции Зоологического института РАН. Класс Костистые рыбы (Osteichthyes). Отряд Скорпенообразные (Scorpaeniformes). Подотряд Cottoidei. Часть II. Сиделева В.Т., Неелов А.В., Воронина Е.П., Волкова Г.А. Исследования фауны морей. Т. 57(65). - СПб., 2006. С. 225-349.
- [7] Каталог фондовой коллекции Зоологического института РАН. Класс Костистые рыбы (Osteichthyes). Отряд скорпенообразные (Scorpaeniformes). Подотряды Scorpaenoidei, Congiopodoidei, Platycephaloidei, Anoplomatoidei, Hexagrammoidei, Scorpaenoidei. Воронина Е.П., Волкова Г.А. Исследования фауны морей. Т. 58(66). - СПб., 2007. 189 с.
- [8] Кривохатский В.А., Лобанов А.Л., Медведев Г.С., Белокобыльский С.А., Дианов М.Б., Смирнов И.С., Халиков Р.Г. Информационная система по энтомологическим коллекциям в Интернете // Труды Русского энтомологического общества, Т. 74, СПб., 2003: С. 59-70.
- [9] Кутикова Л.А., Николаева И.П. Каталог видов коловраток (Rotifera) пресных вод Севера-Запада России // [Электрон. ресурс]. СПб, ЗИН РАН, 2002. (Рус.).

- <http://www.zin.ru/books/rotcatalog/default.asp> [22 ноября 2006]
- [10] Лобанов А.Л. Проблемы создания единой системы диагностической информации в биологии // Единая система и информационно-поисковых языков. Тезисы докладов Всесоюзной научной конференции. Юрмала, 6-8 сентября 1977 г. 1977. С. 84-87.
- [11] Лобанов А.Л. Принципы построения определителей насекомых с использованием электронных вычислительных машин. - Автореферат диссертации на соискание ученой степени канд. биол. наук. Л.: ЗИН АН СССР, 1983. С. 1-19.
- [12] Лобанов А.Л. Диалоговые компьютерные биологические диагностические системы ВІКЕУ5 и ВІКЕУ6 // В: Степаньянц, Лобанов, Дианов, ред., Базы данных и компьютерная графика в зоологических исследованиях // Труды Зоологического института РАН, Т. 269. 1997а. С. 61-65.
- [13] Лобанов А.Л. Компьютерные определители в биологии: результаты 30-летней эволюции // Компьютерные базы данных в ботанических исследованиях. Сборник научных трудов. 1997б. С. 51-55.
- [14] Лобанов А.Л. Компьютерные определители животных и растений: современное состояние и перспективы // В: Рысс, Смирнов, ред., Информационно-поисковые системы в зоологии и ботанике. Тезисы международного симпозиума, май 1999 // Труды Зоологического института РАН. Т. 278. 1999. С. 79-80.
- [15] Лобанов А.Л., Дианов М.Б. Компьютерная диагностическая система ВІКЕУ и возможности ее использования в защите растений // Защита растений в условиях реформирования агропромышленного комплекса: экономика, эффективность, экологичность. Всеросс. съезд по защите растений. Тезисы докл. 1995. С. 548-549.
- [16] Лобанов А.Л., Дианов М.Б. Мир жуков ("Wir bestimmen Käefer") - CD-ROM и краткое руководство. 1996. - Berlin: dialobis edition.
- [17] Лобанов А.Л., Дианов М.Б. CD-ROM: новый инструмент изучения биологического разнообразия // Компьютерные базы данных в ботанических исследованиях. Сборник научных трудов. 1997. С. 55-57.
- [18] Лобанов А.Л., Дианов М.Б. Комплекс программ для создания компьютерных зоологических монографий на компакт-дисках // Отчетная научная сессия по итогам работ 1997 г. Тезисы докладов. 1998. С. 27-28.
- [19] Лобанов А.Л., Дианов М.Б. Средства мультимедиа в электронных зоологических и ботанических публикациях // Информационно-поисковые системы в зоологии и ботанике (Тезисы международного симпозиума, май 1999). Труды Зоологического института РАН. 1999. Vol. 278. P. 100.
- [20] Лобанов А.Л., Рысс А.Ю. Компьютерные идентификационные системы в зоологии и ботанике: современное состояние и перспективы // Информационно-поисковые системы в зоологии и ботанике (Тезисы международного симпозиума, май 1999). Труды Зоологического института РАН. 1999. Vol. 278. P. 17-29.
- [21] Лобанов А.Л., Смирнов И.С. Место и роль информационных технологий в исследованиях Зоологического института РАН // Фундаментальные зоологические исследования: Теория и методы. (По материалам Международной конференции «Юбилейные чтения, посвященные 170-летию Зоологического института РАН», 23-25 октября 2002 г.), М.-СПб.: Товарищество научных изданий КМК. 2004: 283-318 (резюме на англ. яз.).
- [22] Лобанов А.Л., Дианов М.Б., Рысс А.Ю. Современные типы биологических электронных публикаций: CD-ROM и Internet // Информационно-поисковые системы в зоологии и ботанике (Тезисы международного симпозиума, май 1999). Труды Зоологического института РАН. 1999а. 278. С. 39-44.
- [23] Лобанов А.Л., Кирейчук А.Г., Смирнов И.С., Дианов М.Б., Граничин О.Н. Интернет и интерактивные определители биологических объектов // Научный сервис в сети ИНТЕРНЕТ: Труды Всероссийской научной конференции (19-24 сентября 2005 г., г. Новороссийск). - М.: Изд-во МГУ, 2005. с. 132-134.
- [24] Смирнов И.С., Лобанов А.Л. Компьютерный определитель по офиурам как база данных для хранения таксономической информации // Бюллетень Московского общества испытателей природы (МОИП). Отд. геологии. Т. 72, Вып. 1. 1999. С. 87-88.
- [25] Смирнов И.С., Рысс А.Ю. Биологические коллекции и базы данных // Рысс А.Ю., Смирнов И.С. (ред.). Информационно-поисковые системы в зоологии и ботанике // Труды Зоологического института РАН, Т. 278, Санкт-Петербург. 1999. С. 30-38.
- [26] Смирнов И.С., Лобанов А.Л., Дианов М.Б. Зоологические виртуальные музеи // Научный сервис в сети Интернет. Тезисы докладов Всероссийской научной конференции (20-25 сентября 1999 г., г. Новороссийск), Изд-во Моск. ун-та, 1999а. С. 185-187.
- [27] Смирнов И.С., Лобанов А.Л., Голиков А.А., Дианов М.Б. Электронные зоологические коллекции // Электронные библиотеки: перспективные методы и технологии, электронные коллекции. Труды Первой Всероссийской научной конференции (19-22 октября 1999 г., г. С.-Петербург), Изд-во Санкт-Петербургского ун-та, 1999б. С. 236-240.
- [28] Смирнов И.С., Лобанов А.Л., Алимов А.Ф., Голиков А.А. От электронных коллекций к

виртуальным коллективам зоологов в сети Интернет // Интернет и современное общество. Тезисы докладов II Всероссийской научно-методической конференции, (29 ноября-3 декабря 1999 г., г. Санкт-Петербург), Санкт-Петербург, 1999в. С 61-62.

[29] Смирнов И.С., Лобанов А.Л., Дианов М.Б., Голиков А.А., Алимов А.Ф. Зоологические виртуальные музеи: настоящее и будущее // Научный сервис в сети Интернет: Труды Всероссийской научной конференции (24-29 сентября 2001 г., г. Новороссийск). – М.: Изд-во МГУ, 2001. С. 22-24.

[30] Смирнов И.С., Лобанов А.Л., Дианов М.Б., Голиков А.А., Алимов А.Ф., Неслов А.В., Гаврило М.В. Создание информационно-поисковой системы по экологии бентоса и птиц Антарктики (ECOANT) на основе электронной коллекции беспозвоночных, рыб и птиц. «Электронные библиотеки: перспективные методы и технологии, электронные коллекции». Сборник докладов Третьей Всероссийской конференции RCDL'2001. Петрозаводск, 11-13 сентября 2001 г. – Карельский научный центр РАН, 2001. С. 197-198.

[31] Смирнов И.С., Лобанов А.Л., Алимов А.Ф., Кривохатский В.А. Электронные коллекции Зоологического института РАН. Электронные библиотеки: перспективные методы и технологии, электронные коллекции: Труды Пятой Всероссийской научной конференции RCDL'2003, (Санкт-Петербург, 29-31 октября 2003 г.): – Санкт-Петербург: НИИ Химии СПбГУ, 2003: 275-278.

[32] Смирнов И.С., Лобанов А.Л., Алимов А.Ф., Пугачев О.Н., Кривохатский В.А.. Информационная система по биологическому разнообразию России // Научный сервис в сети ИНТЕРНЕТ: Труды Всероссийской научной конференции (22-27 сентября 2003 г., г. Новороссийск). – М.: Изд-во МГУ, 2003. С. 12-14.

[33] Смирнов И.С., Воронина Е.П., Лобанов А.Л., Голиков А.А., Неслов А.В. Создание информационно-поисковых систем по коллекциям морских животных (рыб и беспозвоночных) в Зоологическом институте РАН // Электронные библиотеки: перспективные методы и технологии, электронные коллекции: Труды Шестой Всероссийской научной конференции RCDL'2004, (Пушино, 29 сентября - 1 октября 2004 г.): – Москва, типография ООО «Мультипринт», 2004: 30-33.

[34] 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/part4/SLPAV>

[35] Smirnov I.S., Voronina E.P., Lobanov A.L., Neyelov A.V. The information system of the marine animals collection (fish and invertebrates) in the Zoological Institute Russian Academy of Sciences // Ocean Biodiversity Informatics. International Conference on Marine Biodiversity Data Management. Hamburg, Germany: 29/11-1/12/2004, 2004. p. 27.

[36] Ocean Biodiversity Informatics International Conference on Marine Biodiversity Data Management Hamburg, Germany: 29 November to 1 December 2004. (Engl.). <http://www.vliz.be/obi/> [22 November 2006]

Digital animal collection and interactive keys of biological objects

A.G.Kireitchuk, A.L.Lobanov, I.S.Smirnov,
A.T.Vakhitov, E.P.Voronina, O.N.Pugachev

Elaboration of digital collections, identification keys and ways of ecological monitoring requires more intensive usage of more complex computer technologies. Increasing necessity of estimation and preservation of a biodiversity under anthropogenic action for last years makes this problem particularly important. Wide expansion of the Windows operational system and Internet makes also actual a translation of programs of input and work with keys on this environment, and also data presentation of systems on the Internet for more efficient use of available resources of management and information. Identification keys allow to structure diverse information and to fulfil a fast-access retrieval of it in different sources.

* Работа по теме осуществляется частично при поддержке грантов РФФИ N 09-04-00789a и 07-04-00514a, проекта N11 «Исследование Антарктики. Проведение комплексного изучения антарктической биоты», Федеральной Программы «Мировой Океан» и программы «Биоразнообразие».