

**В.А. Кривохатский, А.Л. Лобанов, Г.С. Медведев,
С.А. Белокобыльский, М.Б. Дианов, И.С. Смирнов, Р.Г. Халиков**

ИНФОРМАЦИОННАЯ СИСТЕМА ПО ЭНТОМОЛОГИЧЕСКИМ КОЛЛЕКЦИЯМ В ИНТЕРНЕТЕ

Начало XXI века совпало с наступлением «эры всеобщей интеграции в деле создания, хранения и использования баз данных (БД) в зоологии» (Лобанов и др., 2002). Эта интеграция связана с унификацией способов представления информации и систем управления БД. Базы данных играют в зоологии особую роль, поскольку вся зоологическая информация опирается на списки таксонов и списки коллекционных экземпляров, хранить которые логичнее всего в БД (другие, некомпьютерные, варианты создания таких списков сейчас уже можно считать научным расточительством, не имеющим будущего). Сегодня разработаны правила для создания БД, потенциально совместимых друг с другом в рамках одной реляционной информационной системы (ИС). Созданные специально для этого стандарты в системе ZOOCOD (Лобанов, 1986; Лобанов, Смирнов, 1997) были положены в основу большинства БД, разрабатываемых российскими зоологами (Алимов и др., 1993; Смирнов и др.,



Рис. 1. Титульная страница информационной системы ZI в Интернете.

2002). На портале Зоологического института РАН (ЗИН) <<http://www.zin.ru>>, размещается информация по самым разным отраслям зоологии, посещаемость его превышает 650 тыс. визитов в год (2002). Сейчас на этом портале действуют 10 интерактивных ИС, включая ИС ZInsecta.

Проект ZInsecta (ZI), открытый в 2001 г. при поддержке РФФИ (грант № 01-07-90284, руководитель В.А. Кривохатский) на сервере ЗИН <<http://www.zin.ru/projects/ZInsecta>> (рис. 1), задумывался как тестовая информационная система, предназначенная для поисков и проверки методов представления в Интернете научных данных из ряда уже имеющихся БД. База данных, отражающая классификацию насекомых (Insecta) до уровня отряда и выполненная целиком в идеологии ZOOCOD (разработчик – А.Л. Лобанов), легла в основу первичного классификатора, представленного в виде таксономического дерева (рис. 2), в котором разворачиваются (+) и сворачиваются (-) отдельные фрагменты классификации по принципу работы папок (Folders – компонент Tree View, WINDOWS) (рис. 2).

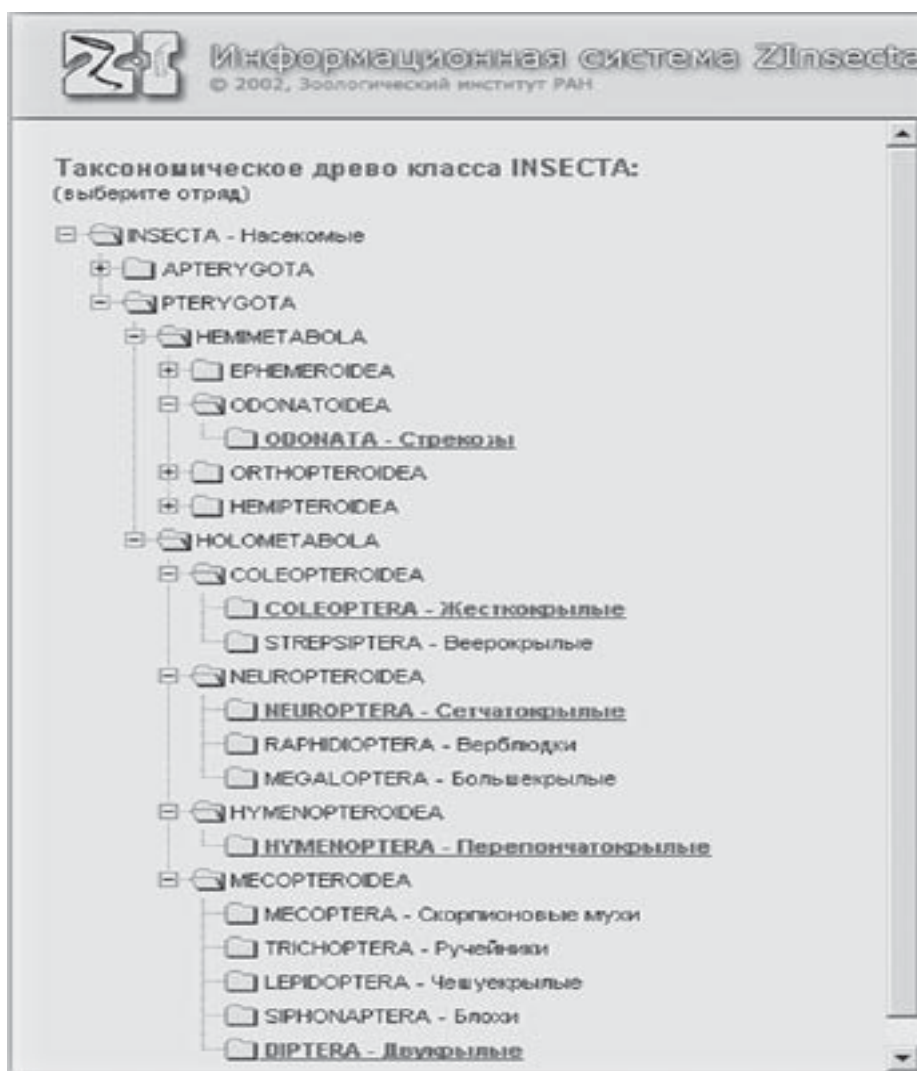


Рис. 2. Основной интерактивный классификатор ZI в Интернете, генерируемый из базы данных.

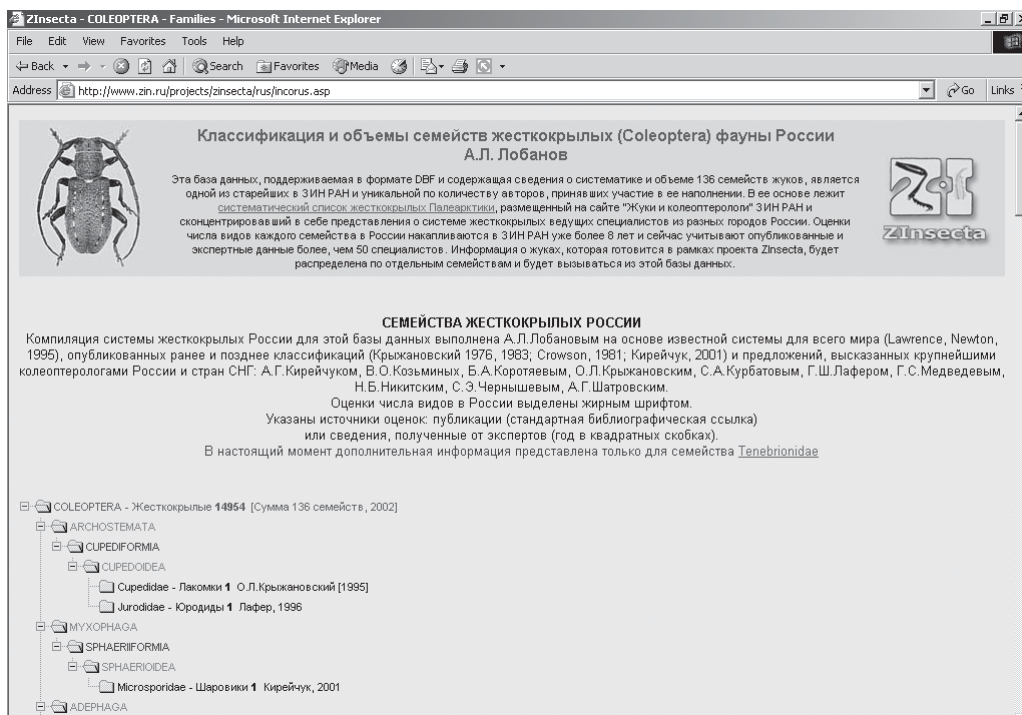


Рис. 3. Классификатор семейств отряда Coleoptera, генерируемый из базы данных, с авторскими оценками числа видов в фауне России.

К основному классификатору привязываются БД следующих порядков посредством скрытого буквенного кода. Часть из них – это классификаторы следующих порядков, например авторский (А.Л. Лобанов) каталог семейств жуков (Coleoptera) с множественными оценками разных авторов числа видов для каждого из семейств в фауне России (рис. 3) и всего мира. Другая часть – это самостоятельные БД специального назначения. Среди них оригинальный систематический каталог мировой фауны муравьиных львов (Neuroptera, Myrmeleontidae) В.А. Кривохатского (рис. 4) и сопряженная с ним через буквенные коды коллекционная БД (рис. 5, 6), включающая географические этикетки изученных автором экземпляров муравьиных львов, хранящихся в коллекциях мира.

Основной и уникальной единицей информации в ZI был принят определенный коллекционный экземпляр или серия экземпляров с одной и той же этикеткой и местом хранения (Белокобыльский и др., 2002). Остальные единицы информации, включая научные названия и ранги таксонов, являются вспомогательными. Цель ИС – объединить основные единицы информации в единой классификации, построенной на вспомогательных единицах. Эта цель достигается в рамках определенных условий и решается в комплексе научных задач, поставленных перед всем многообразием БД, включаемым в ИС, и перед способами реализации связей между ними (Белокобыльский и др., 2002):

Обязательными условиями являются: использование бинарной номенклатуры, реализуемое на основе МКЗН; авторство и дата каждой БД, включенной в ИС;

Описание базы данных



Сетчатокрылые (Neuroptera) мировой фауны

Базы данных этого раздела, поддерживаемые в формате DBF, содержат сведения о систематике, фаунистике и важнейших коллекционных экземплярах сетчатокрылых - отряда насекомых, представленного в коллекции ЗИН РАН более, чем 20 тысячами экземпляров.

Система отряда до уровня семейств и подсемейств (по Oswald, Penny, 1991, с изменениями) представлена ниже. Просмотр дальнейшей информации возможен пока только для семейства Myrmeleontidae.

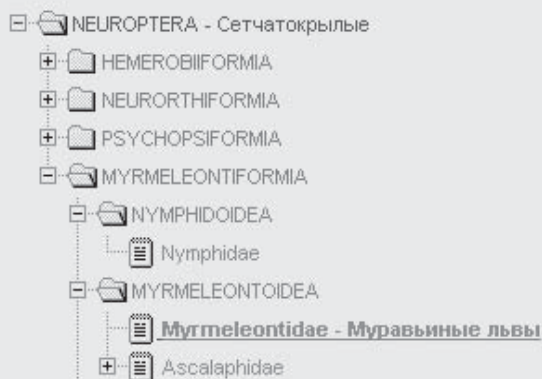


Рис. 4. Классификатор таксонов группы семейства отряда Neuroptera мировой фауны.

иерархическая связь, ранговость, авторство и фиксированный объем таксонов во всех включаемых БД. Условия по выбору языка, формата БД, объема и перечня реализуемой информации в ИС из каждой БД являются предметом соглашения.

Задачи ИС несут гносеологическую нагрузку и отражают попытку совместить тенденции стандартизации, формализации и объективизации ИС с субъективностью и авторской интуицией, реализованными в каждой БД. Традиционно (Белокобыльский и др., 2002) нами ставятся следующие задачи:

- совмещение объективного и субъективного;
- совмещение таксономии и филогении;
- поливариантное отображение ступеней иерархии;
- использование способов унификации и формализации, не нарушающих субъективную уникальность БД;
- динамичность и прерогатива развития ИС над БД.

Вполне научные условия и полностью утилитарные задачи могут находиться в противоречии между собой. В основном эти противоречия находят свои решения. Приведем для примера те случаи, когда возникающие противоречия нам удавалось разрешить.

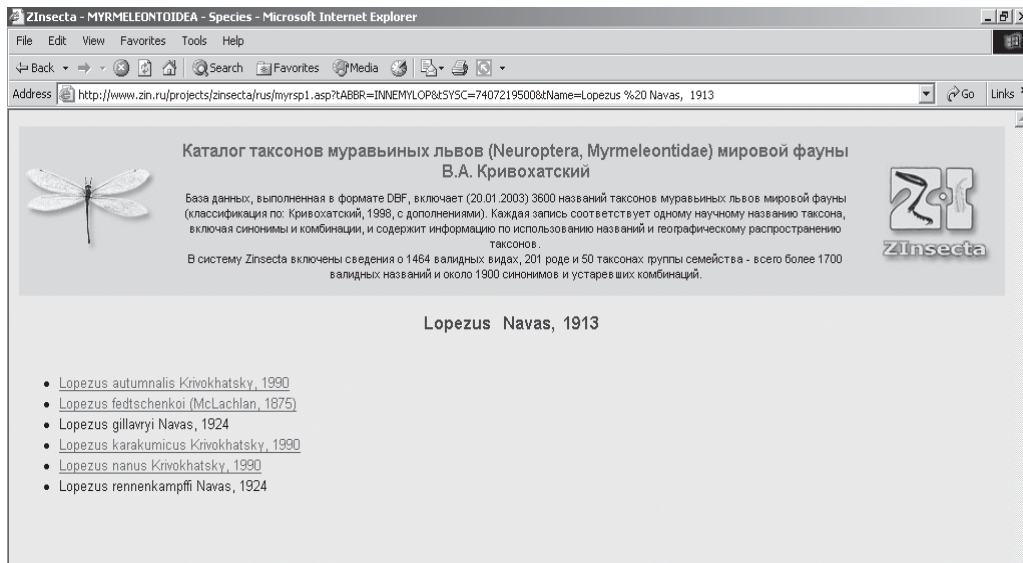


Рис. 5. Классификатор последнего, видового, уровня, генерируемый из базы данных по систематике муравьиных львов (Neuroptera, Myrmeleontidae) мировой фауны.

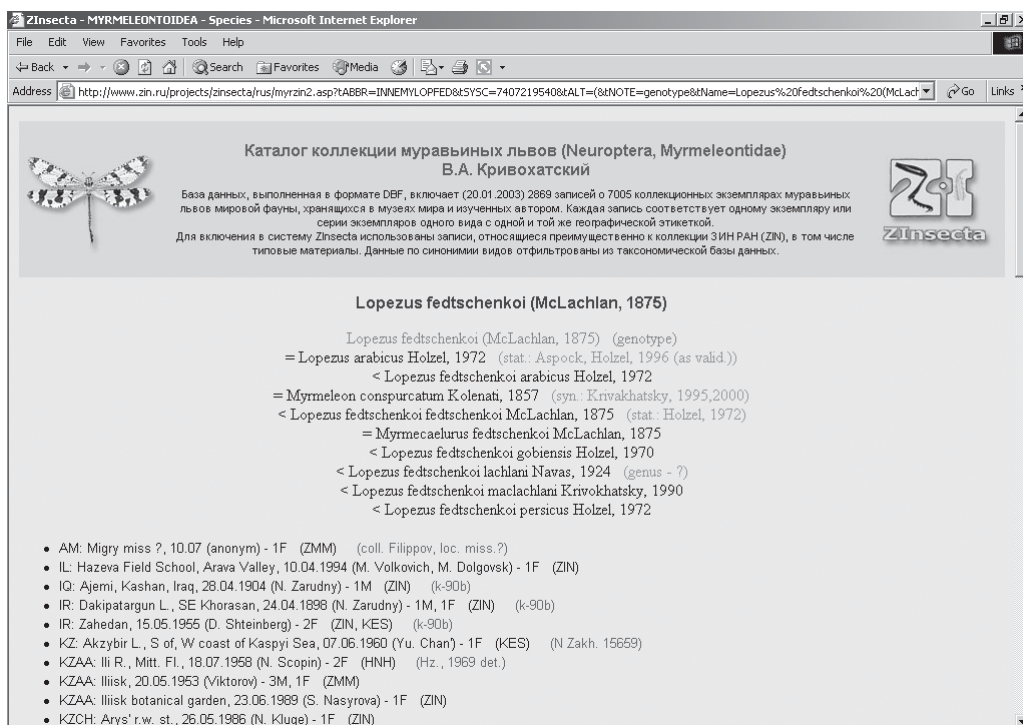


Рис. 6. Каталог коллекционных экземпляров муравьиных львов (Neuroptera, Myrmeleontidae) мировой фауны, совмещенный с таксономической базой данных по муравьиным лвам мировой фауны, включающей синонимы и употреблявшиеся комбинации.

Серьезные трудности возникли при попытке включить в демонстрационную версию классификатора данные по синонимии таксонов из систематических БД. Выбранный нами способ отображения иерархии таксонов (Tree View) позволял нам включать все синонимы таксона только в перечень подчиненных таксонов, что приводило к путанице. Отсутствие решения этой проблемы на странице ASP с раскрывающимися папками натолкнуло нас на попытку ее решения на страницах уже следующего уровня, где сразу после валидного названия таксона приводились все его синонимы, если таковые были включены в БД (рис. 6).

Другая сложность возникла при попытке совместить на одном классификаторе базы различного назначения. С.А. Белокобыльским, кроме БД классификатора одного из подсемейств наездников-браконид, был подготовлен каталог типов видов всего семейства браконид, описанных Н.Р. Кокуевым и хранящихся в ЗИН РАН. К классификатору только одного подсемейства мы, по вполне понятным причинам, могли подсоединить данные этикеток только тех таксонов Кокуева, которые сегодня относятся к этому подсемейству. Решение проблемы на этот раз лежало как раз на иерархическом классификаторе, где удалось развести функционально различные БД от одной ветки таксономического дерева, соответствующей семейству Braconidae (рис. 7).

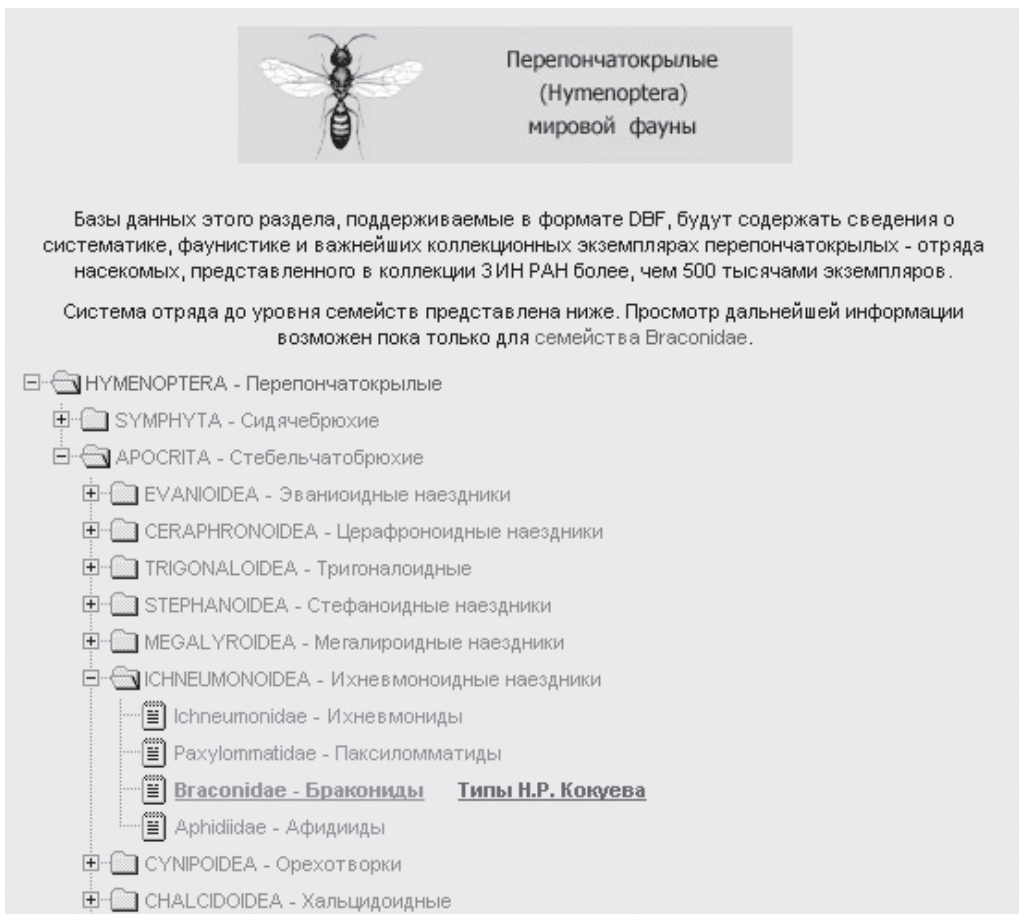


Рис. 7. Пример разведения разных БД от одной ветки таксономического дерева.



Рис. 8. Представление классификатора в частично раскрытом виде.

Некоторые из противоречий нами до сих пор не решены. Так, при программном обеспечении изображения раскрывающегося древа таксонов – классификатора, нами принималось, что как на уровне БД, так и на страницах ASP и HTML, количество рангов таксонов (на экране оно соответствует числу уровней включаемых папок WINDOWS) должно быть одинаковым хотя бы в пределах одной БД. В результате мы пошли на искусственное включение промежуточных таксонов (триб) в пределах основного классификатора муравьиных львов (рис. 8, 9). Здесь подсемейства Myrmecaelurinae (рис. 8) и Pseudimarinae (рис. 9) включают 3 и 2 трибы соответственно. Для подсемейств, имеющих в своем составе лишь по одной трибе (читай: не имеющих в своем составе триб), мы вынуждены были включать эти трибы в классификатор (рис. 9: Dimarinae – Dimarini; Stilbopteryginae – Stilbopterygini; Echthromyrmicinae – Echthromyrmicini). В противном случае, как это мы

сделали для Palparinae (рис. 9), уровень ветвления последующего таксона (рода) нарушался на один такт. Заметим при этом, что новые названия нами не создавались – перечисленные названия триб, а также название трибы Palparini, использовались ранее в научной литературе, они вошли в БД, в которой рассматриваются как синонимы к соответствующим названиям подсемейств.

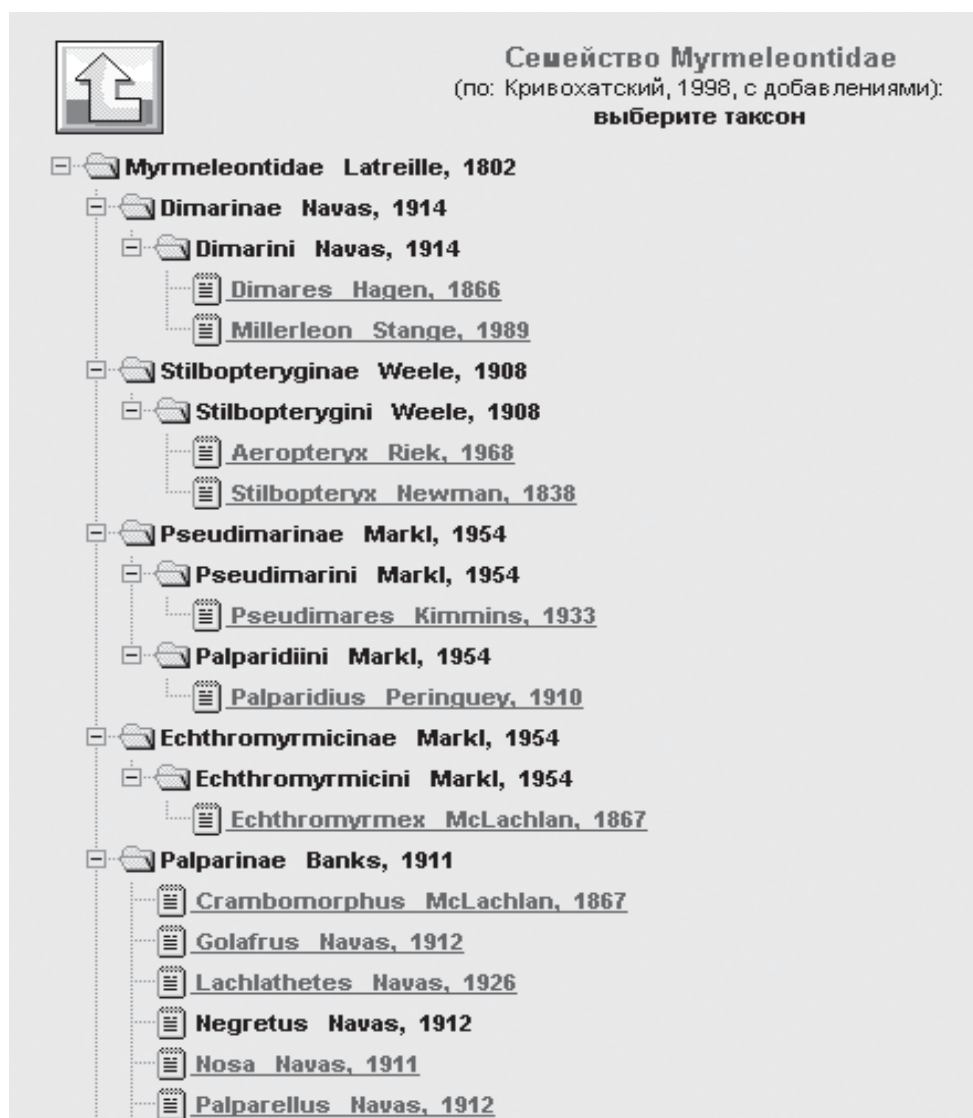


Рис. 9. Представление классификатора в полностью раскрытом виде.

В указанном случае Интернет не служит для первого опубликования каких либо данных, требующих фиксированного авторства и даты опубликования. Однако в ИС выставляются нами и заверненные работы, заслуживающие авторских публикаций на традиционных бумажных носителях. Так, в БД «Жуки-чернотелки фауны России» (рис. 10) Г.С. Медведев по сути не включает новых, ранее не известных данных, но сам список видов России представлен здесь впервые и ранее

опубликован не был. Точно также каталог стрекоз коллекции ЗИН (В.А. Криво-хатский, Н.В. Мяэкиви, Е.А. Голикова) (рис. 11), каталог двукрылых фауны Карелии (А.В. Полевой) (рис. 12) и многие другие самостоятельные разделы, вошедшие в ZI, никогда не публиковались целиком, но, возможно, послужат в будущем для публикаций на бумажных носителях.

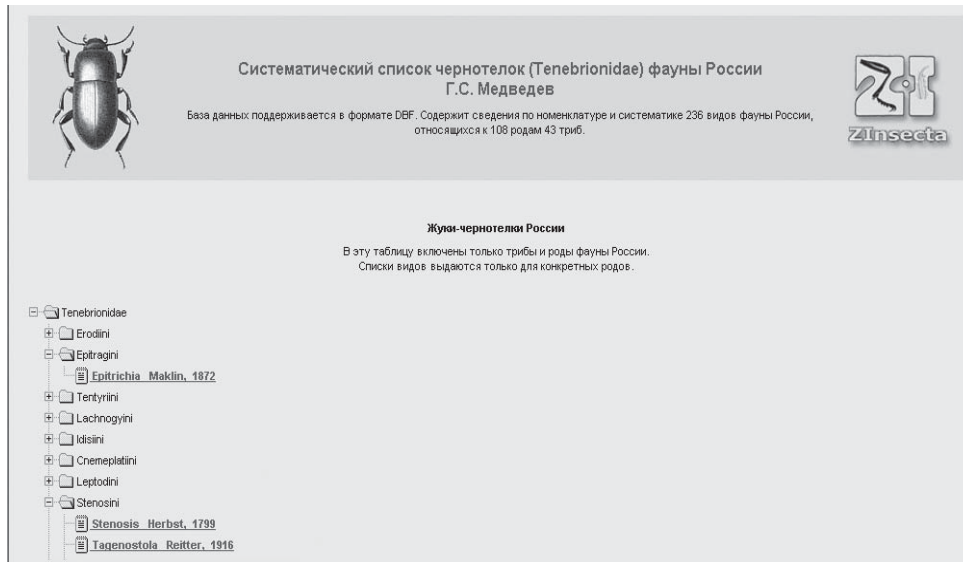


Рис. 10. Жуки-чернотелки (Coleoptera, Tenebrionidae) фауны России.

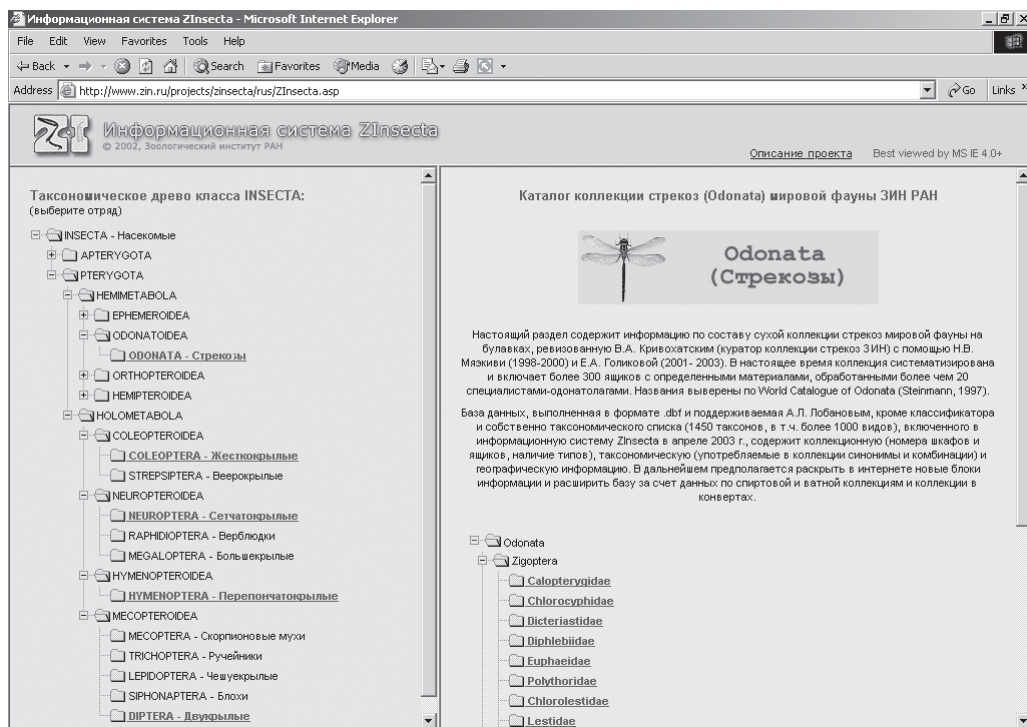


Рис. 11. Состав коллекции стрекоз (Odonata) Зоологического института РАН.

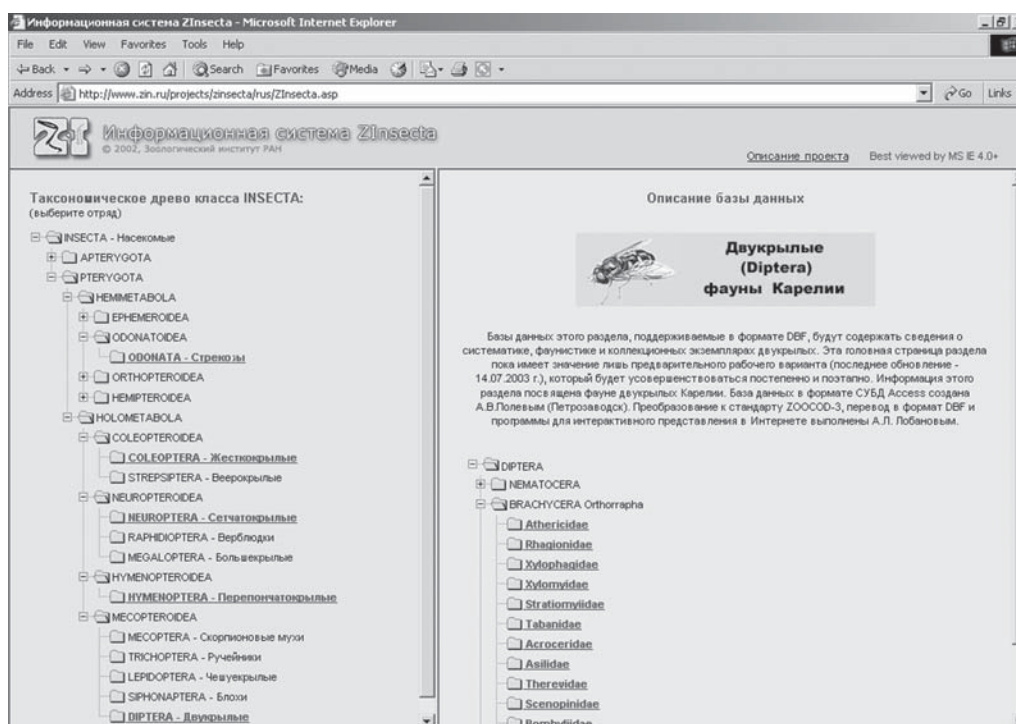


Рис. 12. Двукрылые (Diptera) фауны Карелии.

Создание энтомологических баз данных началось в ЗИН в 1985 году на удаленном терминале ЭВМ БЭСМ-6. Первыми были БД по систематике паразитических перепончатокрылых семейства Encyrtidae (Лобанов, Тряпицын 1985), географическому распространению жуков-жужелиц (Лобанов, Михайлов, 1987) и некоторые другие БД, сведения о которых появились в печати значительно позже. С тех пор в ЗИН были разработаны разными авторами десятки баз данных, затрагивающих почти все крупные отряды насекомых. Однако в этих разрозненных хранилищах информации в основном не был отражен коллекционный аспект, и они не были приспособлены для интерактивного представления в сети Интернет. ZI была призвана исправить эти недостатки.

Несомненно, что успешному распространению в ЗИН технологии баз данных способствовала оригинальная разработка концепции представления сложных таксономических иерархий в плоских, моделирующих обычную таблицу реляционных баз данных. Эта концепция получила название "ZOOCOD" и стала практически стандартом для многих БД в институте, она эволюционировала вместе с развитием компьютеров и СУБД и успешно выдержала экзамен при переходе ПК с DOS на Windows (Лобанов, Зайцев, 1991; Лобанов, Смирнов, 1997). Этот стандарт был разработан в конце 80-х годов в ЗИН РАН А.Л. Лобановым для преобразования иерархических классификаций в плоскую реляционную таблицу таким образом, чтобы все сведения, которые содержатся в систематическом списке, были отражены в соответствующих полях классификатора (Лобанов, 1986). Усовершенствование и многократные испытания стандарта ZOOCOD привели к

его широкому распространению (ЗИН, БИН, ИПЭЭ, Московский государственный педагогический университет, Нижегородский государственный университет и др.) и появлению последней версии - ZOOCOD3. Таблицы, создаваемые в системе ZOOCOD, не ограничиваются использованием только текстовой и цифровой информации. Они с успехом могут содержать и ссылки на иллюстративные материалы, что уже испытывается в ZI в реальном Интернете (рис. 13).

Кроме того нами предполагается дальнейшее развитие географической составляющей баз данных, начатое более 10 лет назад в целях генерации на компьютерные карты местообитаний насекомых из БД по географическим координатам, высчитываемым из сопряженных БД. Эта работа проделывалась в целях зоогеографического районирования небольших территорий, например Монголии (Медведев, Лобанов, 1990; Медведев, 1997; Krivokhatsky et al., 1997) последняя из этих работ уже имела демонстрационную компьютерную версию. Сегодня на страницах других сайтов авторами настоящего сообщения опробован новый способ генерирования карт распространения видов в Интернете по сведениям, включенным в авторские БД. Этот метод предполагается использовать и на страницах ZI.

ZInsecta – это информационная система по энтомологии, теоретически не имеющая никаких ограничений по представлению научных данных в Интернете, генерируемых из БД. Ее развитие теоретически бесконечно в отличие от практической конечности человеческих возможностей и исчерпаемости имеющихся средств.

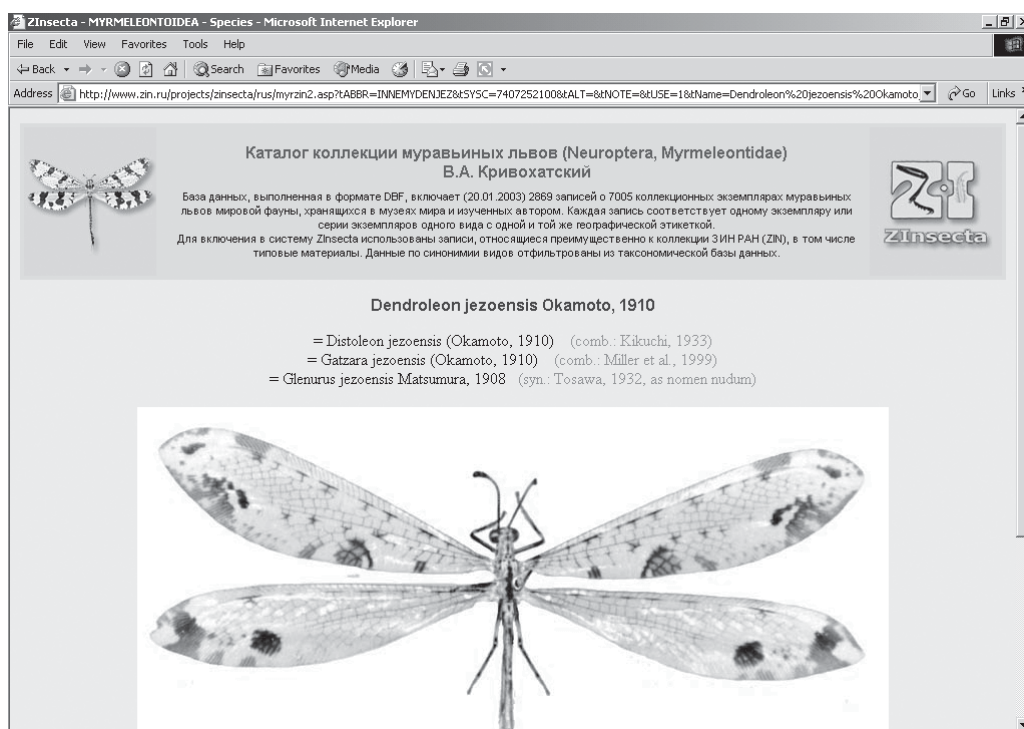


Рис. 13. Пример страницы каталога муравьиных львов (Neuroptera, Myrmeleontidae) мировой фауны с изображением коллекционного экземпляра, генерируемым по ссылке из БД.

Литература

- Алимов А.Ф., Лобанов А.Л., Пугачев О.Н.** Сравнительный анализ реляционного и сетевого подходов к созданию банков данных по систематике, экологии и географическому распространению животных // Журнал общей биологии. 1993. Т. 54, вып. 1. С. 96-103.
- Белокобыльский С.А., Дианов М.Б., Кривохатский В.А., Лобанов А.Л., Медведев Г.С.** Проект ZInsecta - информационная система по энтомологическим коллекциям в Интернете // XII Съезд Русского энтомологического общества. Санкт-Петербург, 19-24 августа 2002 г. Тезисы докладов. СПб., 2002. С. 35.
- Лобанов А.Л.** Линейно-иерархическая структура баз данных о таксонах животных // Принципы и методы экоинформатики. 1986. С. 293-295.
- Лобанов А.Л., Зайцев М.В.** Создание компьютерных баз данных по систематике млекопитающих на основе классификатора названий животных "ZOOCOD" // Труды Зоологического института РАН. 1991. Т. 243. С. 180-198.
- Лобанов А.Л., Михайлов В.А.** Компьютерный банк данных о жуках-жужелицах юго-востока Средней Азии на основе линейно-иерархической структуры баз данных. Душанбе, 1987. С. 1-21.
- Лобанов А.Л., Смирнов И.С.** Принципы построения и использования классификаторов животных в стандарте ZOOCOD // Степаньянц, Лобанов, Дианов, ред., Базы данных и компьютерная графика в зоологических исследованиях (Труды Зоологического института РАН. Т. 269). 1997. С. 66-75.
- Лобанов А.Л., Тряпичин В.А.** Автоматизированная информационно-поисковая система для родов энциртид (Hymenoptera, Encyrtidae) мировой фауны на базе ЭВМ БЭСМ-6 // Энтомологическое обозрение. 1985. Т. 64, вып. 3 С. 649-659.
- Лобанов А.Л., И.С.Смирнов, М.Б.Дианов, А.Ф.Алимов, А.Г.Кирейчук, В.А.Кривохатский.** Российские зоологические базы данных в Интернете // Научный сервис в сети ИНТЕРНЕТ. Труды Всероссийской научной конференции (23-28 сентября 2002 г., г. Новороссийск). 2002. С. 51-53.
- Медведев Г.С.** Опыт эксплуатации компьютерного банка данных по систематике и географическому распространению жуков-чернотелок (Coleoptera, Tenebrionidae) Монголии // Степаньянц, Лобанов, Дианов, ред., Базы данных и компьютерная графика в зоологических исследованиях (Труды Зоологического института РАН.Т. 269). 1997. С. 79-85.
- Медведев Г.С., А.Л.Лобанов.** 1990. Фаунистический список чернотелок (Coleoptera, Tenebrionidae) Монгольской Народной Республики с указанием координат местонахождения // Насекомые Монголии. 1990. Вып. 11. С. 100-200.
- Смирнов И.С., Лобанов А.Л., А.Ф.Алимов, Медведев С.Г., Голиков А.А.** Итоги развития проекта ЗООИНТ и его дальнейшие перспективы // Электронные библиотеки: перспективные методы и технологии, электронные коллекции. Труды Четвертой Всероссийской научной конференции RCDL'2002 (Дубна, 15-17 октября 2002 г.). Дубна: ОИЯИ. 2002. Т. 2. С. 308-315.
- Krivokhatsky V.A., Emeljanov A.F. & Lobanov A. L.** The distribution of antlions in Mongolia (Insecta: Neuroptera: Myrmeleontidae) // In: Canard M., Aspöck H. & Mansell M.W. (Eds). Pure and Applied Research in Neuropterology. Proceedings of the Fifth International Symposium on Neuropterology. Cairo, Egypt, 1994. Toulouse, France. 1996. P. 147-159.