



## **ИНТРОДУКЦИЯ И АККЛИМАТИЗАЦИЯ**

УДК 631.529 : 58.006 : 581.543

### **ИНФОРМАЦИОННО-ПОИСКОВАЯ СИСТЕМА В БОТАНИЧЕСКИХ САДАХ СССР**

*П. И. Латин, З. Е. Кузьмин, Г. Н. Зайцев, С. В. Сорокин*

В современных условиях при непрерывно возрастающем потоке информации огромное значение имеет своевременное и эффективное использование знания, накопленного человечеством. Для хранения и поиска информации стали все более широко применять автоматизированные информационно-поисковые системы (ИПС), основанные на использовании электронно-вычислительных машин и дистанционных средствах связи. В настоящее время ИПС используются практически во всех сферах народного хозяйства и областях науки. Ставится задача объединить эти ИПС в автоматизированные сети национального или международного масштаба [1]. Важно, чтобы подобные системы нашли широкое применение также и в ботанических исследованиях, многие из которых нуждаются в координации и обобщении по региональному признаку.

Первые биологические ИПС были разработаны еще в 1954—1958 гг. на основе электронной обработки данных (ЭОД), и в 1966—1968 гг. они широко внедрились в научные исследования. К 1970 г. было предложено более 40 проектов ИПС для нужд ботаники, зоологии, палеонтологии и прикладной биологии [2]. В ботанических исследованиях ЭВМ наиболее эффективно применяются при создании автомати-

зированной банка данных гербарных этикеток с последующим его использованием для флористических и фитогеографических исследований [3]. Перспективно их использование в поиске новых лекарственных растений [4].

Для хранения и обработки данных о коллекциях живых растений ЭВМ применяются лишь в Соединенных Штатах Америки. В 1966 г. на Международном конгрессе по садоводству в Мэриленде (США) обсуждались вопросы обработки данных наблюдений, проводимых в ботанических садах и арборетумах, с помощью новейших статистических методов и применением ЭВМ. После конгресса в работу включилась Американская ассоциация ботанических садов и арборетумов, несколько институтов выделили для этого специальные средства. Были разработаны специальная программа и план этих работ на десятилетний период времени. Уже через пять лет были обработаны данные по более чем 50 тыс. таксонам из 27 коллекций. Большинство из них включено в главную картотеку данных Научного центра сведений о культурных растениях Американского садоводческого общества в Маунт-Верноне (Вирджиния). Программой предусмотрена выдача сведений об отдельных коллекциях в виде любого числа отпечатков. Главная картотека данных позволяет получать любую записанную информацию по любому таксону. Фонд данных можно поддерживать, систематически дополняя и изменяя его [5].

На Международной конференции по применению электронной вычислительной техники при обработке ботанических коллекций, состоявшейся в Королевском ботаническом саду в Кью (Англия) в октябре 1973 г., при обсуждении проектов применения ЭВМ в гербарном деле и флористических исследованиях рассматривались также пути использования компьютеров при ЭОД живых растений в ботанических садах [6].

Важность и необходимость проведения таких работ в СССР вполне очевидна [7, 8]. В ботанических садах Советского Союза накоплен огромный многолетний фактический материал по интродукции многих тысяч растений. Однако он не может быть в полной мере использован для разработки теоретических и практических вопросов из-за трудностей учета и научной обработки такого большого материала.

Организация информационно-поисковой системы по коллекционным фондам растений всех ботанических садов и дендрариев страны создает возможность решить эту проблему. Создание такой системы позволит наладить унифицированный учет и своевременную обработку экспериментальных данных, организовать должным образом службы информации, даст возможность готовить общие списки интродуцированных растений в СССР с указанием места их произрастания, списки растений по каждому ботаническому саду, а также по таксономическим показателям. При помощи ИПС станут возможны обобщение и классификация хранимых материалов, оперативная их обработка и соответствующий анализ, разработка рекомендаций по улучшению состава коллекций, а также по практическому использованию интродуцированных растений.

Во исполнение решений Совета ботанических садов СССР от 1972 и 1979 гг. Главный ботанический сад АН СССР (ГБС) совместно с Вычислительным центром АН СССР (ВЦ) создает информационно-поисковую систему по коллекционным фондам 115 ботанических садов и дендрариев Советского Союза на базе ЭВМ БЭСМ-6, находящейся в ВЦ, и математического обеспечения, применяемого там. Оперативная связь будет поддерживаться по телефону через терминал, устанавливаемый в ГБС. Терминал состоит из дисплея типа «Видеотон-340» в комплексе с алфавитно-цифropечатающим устройством и преобразователем импульсов.

## **ЕДИНАЯ СИСТЕМА УЧЕТА КОЛЛЕКЦИОННЫХ РАСТЕНИЙ**

Для учета и передачи через терминал данных о большом числе растений и последующей их обработки очень важно выбрать определенную систему унифицированных признаков, записываемых в электронную память машины. Такая единая для всех ботанических садов система разработана. Она предусматривает регистрацию растений природной флоры до образца, а культурных растений до сорта.

Определено 39 показателей (реквизитов) единой системы учета. Показатели включают таксономические (названия семейства, рода, вида, культивара, формы и гибрида, автора), географические (ареалы, места сборов образцов), морфологические признаки, фенологические данные, характерные биологические особенности (тип посадочного материала, листопадность для древесных, продолжительность жизни, периодичность плодоношения, отношение к увлажнению, зимостойкость, способность к возобновлению и т. д.). Показатели выбраны таким образом, чтобы по возможности полно и объективно характеризовать интродуцируемые растения любой жизненной формы (деревья, кустарники, лианы, многолетники и др.).

Запись показателей производится на специальную анкету, которая обеспечивает в случае необходимости перенос данных на три перфокарты [9].

## **МАТЕМАТИЧЕСКОЕ ОБЕСПЕЧЕНИЕ**

## ИНФОРМАЦИОННО-ПОИСКОВОЙ СИСТЕМЫ

ИПС базируется на ЭВМ БЭСМ-6 с операционной системой ДИСПАК [10, II], магнитных дисках емкостью 7, 25 и 29 мегабайт, на базе данных МАРС [12]. ИПС, рассчитанная на 100 тыс. анкет, ориентировочно займет 35 мегабайт памяти. Диалог на естественном языке и управление ИПС предполагается осуществить через систему ДИЛОС [13]. Программная часть ИПС написана С. В. Сорокиным на языке ПАСКАЛЬ [14, 15]. Отладка программ производится в системе ПУЛЬТ [16].

Общий объем программ в виде текстов 0,2 мегабайта. Сейчас заполнено около половины всего объема программ. ИПС состоит из пяти основных программ:

1. «Ввод анкет». Программа вводит поисковые образы документов (анкеты), их исправляет, исключает.
2. «Поиск документов». Программа осуществляет поиск документов по запросу (поисковое предприятие).
3. «Преобразование поисковых образов». Программа предназначена для ускорения поиска.
4. «География». Программа модифицирует граф географических районов.
5. «Диалоговый запуск предыдущих программ».

Прежде чем запустить программу «ввод анкет», содержимое анкет переносится на машинные носители информации. Это примерно 4/5 всей ручной работы по информационному заполнению ИПС. Анкеты можно заносить непосредственно через клавиатуру терминала на магнитные диски. При этом скорость занесения информации в среднем примерно в 2—3 раза меньше, чем скорость перепечатки на машинке из-за того, что пользователи, запуская свои программы, занимают различные ресурсы ЭВМ и заставляют другие программы ждать освобождения этих ресурсов.

Другой способ занесения информации — перенос (пробивка) информации на перфокарты. С перфокарт информация через вводное устройство ЭВМ переписывается на магнитный диск. Но при этом возникает дополнительный этап ручной работы по вводу перфокарт, хотя скорость перфорации приближается к скорости перепечатки на пишущей машинке. Через терминал удобнее вводить информацию, когда объем работ не больше недели, иначе лучше пользоваться перфокартами.

После ввода информации с анкет на магнитный диск запускается программа «ввод анкет», которая просматривает анкеты и проверяет их правильность (синтаксический контроль). Там, где должны быть числа, не должно быть букв, числа должны быть не слишком большие (например, там, где месяц, число должно быть не больше 12), названия ареалов совпадать с теми, которые записаны в таблицах ИПС и т. п. Обнаруженные ошибки через терминал исправляются, и программа «ввод анкет» запускается повторно. Выдаются данные о вновь обнаруженных ошибках, последние снова исправляются оператором ИПС, и так до тех пор, пока не будут исправлены все ошибки. Когда ошибки исправлены, программа «ввод анкет» переписывает содержимое анкет в предварительную базу данных. Анкеты удобно объединить в файлы и записывать в базу данных по одному файлу. Практика показала, что в файл удобно объединять от 25 до 50 анкет.

После ввода анкет в предварительную базу данных запускается программа «поиск документов». Данными к этой программе являются запросы. Запрос, как и анкета, может состоять из 39 показателей (реквизитов), в число которых входят номер показателя и его содержимое. По содержимому показателя программа определяет, какие анкеты (поисковые образы документов) найдены, а какие — нет. В простейшем случае, когда содержимое показателя запроса и поискового образа документа совпадает, анкета считается найденной.

Содержимым показателя может быть некоторое множество объектов. Например, диапазон дат, список ареалов, диапазон оценок по шкале зимостойкости и т. п. Для облегчения заполнения показателей анкет и поиска вводится несколько критериев найденности поискового образа.

1. Тожество содержимого показателя в поисковом образе и в запросе.
2. Непустое пересечение этих множеств.
3. Вхождение запроса в поисковый образ.
4. Вхождение поискового образа в запрос.

Поэтому в запросе указывается и критерий найденности. Кроме того, имеется возможность через логические операции: конъюнкция, дизъюнкция и отрицание — объединять различные показатели запроса и получать более сложную формулу для определения, какие поисковые образцы документа считать найденными.

Далее в программе «поиск документов» указывается, какие сведения о найденных документах нас интересуют. Можно указать количество найденных документов, распечатку определенных показателей, запись показателей в некоторый файл для последующей обработки их какими-либо программами вне ИПС.

Счетчик программы «поиск документов» определяет, по каким показателям шел поиск и сколько раз использовался данный показатель. Когда подобная статистика будет достаточно большой, оператор ИПС

запустит программу «преобразование поисковых образов». Эта программа упорядочит поисковые образы по тем показателям, которые встречаются часто. Тогда программа «поиск документов» сможет находить поисковые образы не перебирая все подряд.

В показателе 7 (ареал) используются названия географических районов. В методических указаниях [9] предусмотрен 131 район. Эти районы объединены в граф, представленный в виде дерева, в основании которого находится земной шар: ветви дерева — районы, каждый район делится на подрайоны, которые в свою очередь могут подразделяться дальше. Список географических районов в методических указаниях ориентировочный, поэтому система открыта для пополнения новыми районами. Тут граф в виде дерева уже недостаточен, и поэтому в систему встроены направленные графы. Это значит, что один подрайон может входить не в один, а в несколько районов. Для изменения графа географических районов имеется программа «география», которая запускается оператором ИПС.

Таким образом, оператор ИПС запускает по очереди определенные программы со своими данными. Для запуска таких программ необходима диалоговая программа, общение с которой ведется на некотором алгоритмическом диалоговом языке. Возможно, наряду с алгоритмическим языком будет частично использоваться и естественный язык (русский). Целесообразность введения естественного языка будет определена опытной эксплуатацией ИПС.

#### ЛИТЕРАТУРА

1. Соколов А. В. Информационно-поисковые системы. М.: Радио и связь, 1981. 152 с.
2. Мальшев Л. И. Электронная обработка данных в гербарном деле и флористике.— Ботан. журн., 1977, т. 62, № 5, с. 713—732.
3. Лякавичюс А. А., Красаускас А.-И. И. Применение электронно-вычислительной техники в гербарии Института ботаники АН Литовской ССР.— Ботан. журн., 1977, т. 62, № 4, с. 523—527.
4. Шретер А. И., Терехин А. Т. Некоторые итоги и перспективы использования ЭВМ при поиске новых лекарственных растений.— Раст. ресурсы, 1980, т. 16, вып. 4, с. 481—493.
5. Говард Р. А., Браун Р. А. Вопросы регистрации и обработки данных по интродукции растений.— Бюл. Гл. ботан. сада, 1976, вып. 100, с. 29—34.
6. Утехин В. Д. ЭВМ в ботанических коллекциях.— Новые книги за рубежом, 1977 (В), № 1, с. 31—33.
7. Лапин П. И. Об организации централизованного учета коллекционных фондов ботанических садов СССР.— В кн.: Методика фенологических наблюдений в ботанических садах СССР. М.: ГБС АН СССР, 1972, с. 4—17. Рукопись деп. в ВИНТИ, № 5494—73 Деп.
8. Цицин Н. В. Деятельность и задачи ботанических садов в свете решений XXV съезда КПСС.— Бюл. Гл. ботан. сада, 1976, вып. 102, с. 3—9.
9. Кузьмин З. Е., Зайцев Г. Н., Сорокин С. В. Методические указания по учету коллекционных растений ботанических садов СССР с помощью ЭВМ. М.: ГБС АН СССР: ВЦ АН СССР, 1979. 50 с.
10. Бокова И. Д., Зельдинова С. А., Зуев В. И.; Корякин В. К., Кошкина Л. В., Озорнин Ю. В., Тюрин В. Ф., Шулепов Н. И. Операционная система ДИСПАК для БЭСМ-6 (пользователю). М.: ИПМ АН СССР, 1973. 80 с.
11. Балакирев Н. Е., Зельдинова С. А., Копытов М. А., Рогов Ю. П., Тюрин В. Ф., Шебанов В. Ф. Руководство пользователей по работе с операционной системой ДИСПАК. М.: ВЦ АН СССР, 1982. 128 с.
12. Филиппов В. И. Руководство по СУБД Компас. М.: ВЦ АН СССР, 1981. 48 с.
13. Абрамов В. Г., Брябрин В. М., Пховелишвили М. Г., Сенин Г. В., Элигулашвили А. А. ДИЛОС—диалоговая система для взаимодействия с ЭВМ на естественном языке. М.: ВЦ АН СССР, 1979. 80 с.
14. Вирт Н. Язык программирования ПАСКАЛЬ (пересмотренное обобщение) / Пер. с англ. Покровского С. Б. Алгоритмы и организация решения экономических задач. М.: Статистика, 1977, вып. 9, с. 52—86.
15. Пирин С. И. Язык ПАСКАЛЬ-МОНИТОР и его использование. М.: ВЦ АН СССР, 1978. 56 с.
16. Брябрин В. А., Сметанин В. Л., Филиппов В. И., Юфа В. М. Система ПУЛЬТ-78. Руководство к использованию. М.: ВЦ АН СССР, 1978. 100 с.