Introducing BioBot, the Biological Search Engine for the Internet

Today, the Internet can be likened to an information bazaar. The good news is, the content has exploded in recent years. The bad news is, the content has exploded in recent years—along with the number of information sources on biological resources. At times, the bazaar is almost bizarre. Those who need to learn more about a particular species or, perhaps, broader environmental issues can be overwhelmed by the sheer volume of data and information available.

To deal with this dilemma, the National Biological Information Infrastructure (NBII) product development team now offers a new and truly unique tool for the biological community called “BioBot” (the NBII is a broad, collaborative program to provide increased access to data and information on the nation’s biological resources).

NBII customers can use BioBot to:

- Search NBII information as well as other biological information available on the Internet through sources such as SnapBiology, AltaVista, and Yahoo. These vendors and others are currently indexing and harvesting information from a number of biological sites. BioBot builds upon this information, allowing users to easily access the biological information they want.
- Define (“profile”) specific information of interest — which BioBot will then deliver to the user’s desktop daily, weekly, biweekly, or monthly, as the user prefers (search results are presented back to the user in a ranked, concise format); and
- Combine BioBot with other biological information systems

The Global Biodiversity Information Facility: An International Network of Interoperable Biodiversity Databases

Biodiversity is distributed all over the Earth, with concentrations primarily in developing countries. In contrast, scientific knowledge about biodiversity is concentrated in major centers in developed countries. To be useful in the management and use of biodiversity, information about biodiversity should be available when and where it is needed, by decision makers and scientists alike.

With these observations in mind, in 1996 the Megascience Forum of the Organization for Economic Cooperation and Development (OECD) convened a Working Group on Biological Informatics with representatives from more than 20 countries, including Australia, Belgium, Canada, Denmark, Finland, France, Germany, Israel, Italy, Japan, Korea, Mexico, Netherlands, Norway, Poland, Portugal, Russia, Sweden, United Kingdom, United States, and the European Union.

In its report to the OECD in January 1999, one of the Working Group’s major recommendations was that an international mechanism is needed to make biodiversity data and information accessible worldwide, and it proposed the formation of the Global Biodiversity Information Facility or GBIF (the report of the OECD Working Group on Biological Informatics can be found at <www.oecd.org/dsti/sti/s_tms/prod/index.htm>). GBIF will be a distributed system of interlinked databases (continued on page 6)
United States Hosts First IABIN Council Meeting

The U.S. Geological Survey hosted the first official meeting of the Council of the Inter-American Biodiversity Information Network (IABIN), in Miami, FL, on December 13-14, 1999. The IABIN Council, which for this meeting included the official Focal Point from each participating country, adopted rules of procedure for IABIN, officially establishing a governance structure for the network. The current Council agreed to extend Council membership in the near future to selected regional biodiversity initiatives, international organizations, and non-governmental organizations in the biodiversity information community. A process for this Council expansion is being developed.

The IABIN Council elected Gladys Cotter, United States, as Council Chair and Elaine Fisher, Jamaica, as Vice-Chair. The Council appointed an Executive Committee to coordinate activities between meetings. For 2000, the Executive Committee includes IABIN Focal Points from Argentina, Brazil, Colombia, Dominican Republic, Guatemala, and Mexico; a representative from the non-governmental/inter-governmental organizations and initiatives (Marcos Silva, of the North American Biodiversity Information Network, is that representative for the next year); and a representative from the Organization of American States (OAS). The Council Chair and Vice-Chair are also members of the Executive Committee.

Two Working Groups were established to address the network activities planned for calendar year 2000. One will address developing technical projects, establishing cooperative linkages with other regional and global initiatives, and evaluating software initiatives, and evaluating software tools for possible network-wide use. The other Working Group will address a financial sustainability strategy for IABIN as a whole and will help prepare funding proposals.

The IABIN Council, which for this meeting included the official Focal Point from each participating country, adopted rules of procedure for IABIN, officially establishing a governance structure for the network.

OAS Committee Endorses IABIN

Another major milestone was achieved in October 1999 with the endorsement of IABIN by the Inter-American Committee on Sustainable Development (CIDS) of the OAS. CIDS passed a resolution which recognized the importance of IABIN. The resolution called for IABIN to develop a long-term program reflecting the priorities of participants and focusing on technological tools needed for information exchange. It also requested that the OAS Unit for Sustainable Development and Environment enhance fund-raising efforts to support IABIN. The request for fund-raising support is particularly pertinent because the proposal to OAS for FY 2000 support for IABIN, submitted by Brazil on behalf of IABIN, was not selected for funding.

IABIN is an international initiative to promote greater coordination among Western Hemisphere countries in the collection, sharing, and use of biodiversity information. Information on IABIN can be found at <www.iabin.org> (international site) and at <www.nbii.gov/iabin> (U.S. site).
Partners in the Spotlight

“Partners in the Spotlight” highlights the activities and contributions of a wide range of NBII partners. We are pleased to welcome Jeff Waldon and the Fish and Wildlife Information Exchange as this issue’s Partner in the Spotlight. If you’re interested in producing a similar article about your organization, please contact Ron Sepic, Access Editor, at ron_sepic@usgs.gov.

Conserving Information at the Fish and Wildlife Information Exchange

Conservationists are awash in data, but making information out of that data is a continual challenge. The Fish and Wildlife Information Exchange (FWIE, a division of the Conservation Management Institute, College of Natural Resources at Virginia Tech, <fwie.fw.vt.edu>) specializes in the research and development of techniques and information technologies that will assist in the conservation of wildlife and plants worldwide. We help agencies and institutions develop and modernize databases, take advantage of emerging technologies like Web databases, address management needs like planning and metadata development, and in general help conservation organizations—both private and public—get a maximum return on their investment in data collection.

The FWIE works at many different levels within organizations from the individual data set, to statewide information systems, to multi-agency networks that address large geographic areas. As examples, we are assisting Shenandoah National Park to develop a unified inventory and monitoring database to bring together the many different data sets maintained by the Park. At the state level, we are helping the New Mexico Game and Fish Department and several other states to develop statewide faunal species information systems, available on the Web, and used for environmental assessment, planning, extension, and education. We also have several projects underway to bring together multiple states into a federated data network on the Web to improve access to regional data. The most advanced example is the Multi-State Aquatic Resource Information System, an effort by several state fisheries agencies in the Lake States region to provide fish catch/unit effort data over the Web. This effort is one component of the Aquatic Resource Coalition initiative spearheaded by the U.S. Geological Survey (USGS).

We have been an active participant in the NBII since 1995. The collaboration includes a survey of state-level biological databases, development of Web-based systems that can query across multiple state fish and wildlife databases, and cooperation on providing NBII metadata training to state fish and wildlife agency personnel.

By working with other divisions within the Conservation Management Institute, the FWIE is able to bring a broad set of capabilities to new problems. We have a strong geographic and remote sensing capability and programmers that can integrate observation data sets with spatial data sets. We have field crews who are involved in species surveys, vegetation work, and a variety of other activities in the field. We have field biologists, wildlife biologists, botanists, an entomologist, a geologist, computer programmers, and staff with international experience (Madagascar, Botswana, Egypt, Belize, and others) to bring a multi-disciplinary approach to problem solving. We work closely with major federal programs like the USGS/National Park Service Vegetation Mapping Program, the Gap Analysis Program, the NBII, and others to make sure products and procedures comply with policies and best management practices.

We also provide training by including students in real-world projects and holding continuing education sessions on a variety of topics in cooperation with the Organization of Fish and Wildlife Information Managers (OFWIM) <www.fwie.fw.vt.edu/ofwim/>. The OFWIM is the professional organization for fisheries and wildlife data managers in the United States. The FWIE works closely with this organization to sponsor national meetings (the latest was in Lake Tahoe, where 121 analysts and data managers converged) and continuing education.
NBII Hosts “Designing a Report on the Nation’s Ecosystems”

On October 27, “Designing a Report on the Nation’s Ecosystems” made its debut on the Web at <http://www.us-ecosystems.org>. The product of a public-private effort, the report is hosted on a virtual Web server at the USGS Center for Biological Informatics and is part of the growing number of partnerships between the NBII and the biology and environmental communities. The report is also featured on the non-governmental organizations page of the NBII at <http://www.nbii.gov/programs/nongov.html>.

The H. John Heinz III Center for Science, Economics, and the Environment created the study as a prototype, intended to elicit comments, criticism, and suggestions on the project’s initial approach and preliminary findings. “This is the first step toward truly reliable and regular reports on the condition of our natural resources,” said William Merrell, President of the Heinz Center.

Under the leadership of William Clark, professor at Harvard’s Kennedy School of Government, the study brought together a wide range of experts from industry, conservation groups, academia, and government agencies. More than 100 individuals joined various committees and working groups, contributed information, or reviewed the work in progress. In following a core set of indicators across ecosystems and over time, the report mirrors the approach used in the collection and reporting of economic data. “Periodic reports on leading economic indicators, such as employment, production, and trade are fundamental tools for making investments and determining economic policy,” Clark said. “Developing analogous indicators for the environment is long overdue and fundamental to the quest for sustainable environmental policies.”

Although environmental information is produced by many public agencies and private interests, it is rarely synthesized or brought together to improve overall understanding of ecosystems. Indeed, different interests often report only what they find most useful. “I get opinions every day,” said Rep. Saxton, Chairman of the House Subcommittee on Fisheries Conservation, Wildlife, and Oceans, “but I need timely, accurate information in a meaningful context, and that is why I support this project and applaud its report.”

“If the point of the exercise is to develop a reporting system that is perceived as fair and accurate, then it is essential that every stakeholder comes out of the process with a sense of ownership,” said Tom Jorling.

Introducing BioBot, the Biological Search Engine for the Internet

(continued from page 1)

MyNBII Filter. BioBot then scours the Web for “invasive species” and delivers this information to a personal Web page, electronically notifying the user at the specified time interval that the information is available. Further customization of these search and information delivery features is also possible.

BioBot also offers online assistance. In fact, it has been designed to help both novice and expert searchers find information of interest. Sample search terms and online/context sensitive help are both available.

“MyNBII Filter. BioBot can save the biotic Web surfer valuable time,” says Mike Frame, Director of the NBII’s BioBot team. “It lets NBII customers receive on-the-fly, up-to-date biological information from their Web pages.”

“We urge NBII users to look for BioBot on the main page of the NBII <www.nbii.gov>, then try it out,” says Frame. “Whether they’re in search of serious scholastics, commonsense help, or even info-tainment with a biological edge, I think they’ll be pleased with the results. They may even find that their interest in searching for biological information starts to grow as fast as the Web’s biological content!”

BioBot has two parts. The first, shown above, is the main search screen, which offers users a wide variety of options for finding the information they want. For instance, users can indicate which of the Internet’s leading biological information sources they want to employ (such as BioLinks, GoNetworkBiology, or YahooBiology). Search types include a simple search, Boolean search, and phrase searches.

MyNBII Filter Home Page

Options: Add A Query | Delete All Queries | Help | Sign Off

<table>
<thead>
<tr>
<th>Query String</th>
<th>Edit Options</th>
<th>View Results</th>
<th>Delete Query</th>
<th>Rerun Now</th>
</tr>
</thead>
<tbody>
<tr>
<td>invasive species</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The second principle part of BioBot, MyNBII Filter, gives users the opportunity to define their interest(s) and how often information of interest will be delivered to them (daily, weekly, biweekly, or monthly).
When the Texas/Mexico Borderlands Information Center (BIC) received a grant from the USGS Biological Resources Division for setting up a metadata clearinghouse node in 1997, Miguel Pavon, Administrator of the Information Center, and his supervisor, Charles Palmer, Coordinator, Border Affairs, attended a two-day NBII metadata training workshop to prepare themselves for this endeavor.

The BIC is a clearinghouse and referral center for information about both sides of the Texas/Mexico border. Metadata expertise is critical to the BIC’s mission: to help researchers, planners, and others locate data and information related to the natural resources and demographics of the border region; and to promote data sharing among entities with borderlands responsibilities.

Since attending the NBII metadata training workshop, Miguel has been able to teach researchers and planners from both sides of the border how to document their data sets with metadata in compliance with the Federal Geographic Data Committee (FGDC) Content Standard for Digital Geospatial Metadata. Using material from the NBII metadata training workshop, Miguel trained 45 workshop participants on the basics of metadata in 1998.

Charles adds that the metadata training has aided the BIC’s efforts to document its historical in-house data. For example, the BIC was able to develop metadata to provide detailed information about the Mexican Mapping Agency’s topographical maps and data about geographical land use and soils.

Miguel says that before attending training, he found metadata concepts difficult to understand and “not intuitive.” He credits his “excellent instructors” (Jennifer Gaines and Sharon Shin, NBII metadata training workshop coordinators) with making metadata comprehensible by introducing the concepts first and then providing hands-on exercises so that the participants could “learn at their own pace” and practice what they learned.

The metadata training familiarized Miguel with metadata standards, easy-to-use metadata templates, and the NBII metadata creation software, MetaMaker. By implementing what he learned at the training, Miguel says that “now our data can be discovered by a user.”

He also says that descriptive metadata is essential for “advertising your data set.” “If you don’t describe it, you can’t sell it,” he says. Miguel compares metadata to washing dishes: “If you wash dirty dishes immediately, it’s easy. But if you wait a couple of weeks, you have a mess.” In other words, if you document data sets with identifying metadata right away, the data can be found and used productively.

BIC hosts two metadata nodes: the Texas Natural Resources Information System (TNRIS) node and the Transboundary Metadata Clearinghouse Node; between them, the metadata collection totals close to 4,000 entries. Miguel and Charles say that the next goal of the BIC metadata project is to translate the metadata training materials into Spanish to train their Mexican colleagues.
and interoperable modules (databases, software and networking tools, search engines, analytical algorithms, and so forth) that will enable users to navigate and put to use the world’s vast quantities of biodiversity information. GBIF will be distributed, while encouraging cooperation and coherence; global in scale, though implemented nationally and regionally; and open to participation and benefit by all countries.

GBIF will be organized into a series of work programs. As currently envisioned, these will include:

- Access/interoperability/search capabilities—developing the software and hardware infrastructure to enable, for instance, interoperability of specimen-level and species-level databases, linkages among with molecular databases, linkages between biological and non-biological information (for example, climate, geography), algorithms to search multiple databases simultaneously, and standards for data and software.
- Electronic Catalog of Names of Known Organisms—encouraging and speeding up of processes already in place, such as Species 2000, to produce a reasonably complete catalog of the scientific names (including synonyms) of all described organisms.
- Digitization of natural history collections data, and provision of access to those databases—GBIF will work with its member governments to provide the funds for the digitization of specimen data, thereby aiding in the repatriation of data from the developed to the developing world and providing an excellent curatorial tool for participating museums.
- “Species Bank”—developing a means of accessing information of all sorts about individual species, including species descriptions, information about geographic distribution, and so forth. Much of GBIF’s efforts in this area will build on existing initiatives such as the Tree of Life Project.
- Digital Biodiversity Literature Resources—developing a digital library of biodiversity information compiled from information available in print libraries, including type descriptions.
- Training—helping to develop proposed curricula for biodiversity informatics training programs, as well as providing assistance in how to use GBIF and acting as an interface between the computer science and biodiversity communities.
- Outreach—working with other organizations to help provide the hardware, software, and Internet connectivity to allow full participation in GBIF by the developing world.

Funding for these projects will largely be provided through existing national and regional mechanisms. In addition, a small GBIF secretariat will provide a clearinghouse for information about past, present, and proposed projects; promote interoperability among databases by, for example, holding standards-setting workshops and developing protocols for linking databases; provide coordination for national funding programs as well as

(continued on page 7)
CHM Prepares for SBSTTA-5

The pilot phase of the Clearing-House Mechanism (CHM) was a major topic on the agenda of the fifth meeting of the Subsidiary Body on Scientific, Technical, and Technological Advice (SBSTTA) of the Convention on Biological Diversity (CBD), held January 31–February 4, 2000, in Montreal, Canada. A Note prepared by the CHM Secretariat was presented to the SBSTTA delegates for consideration. The Note, entitled “Pilot Phase of the Clearing-House Mechanism: Note by the Executive Secretary,” outlines the achievements to date in the implementation of the CHM and invites the SBSTTA to recommend to the CBD Conference of Parties (which next meets in Nairobi, Kenya, in May 2000) a number of actions to further the objectives of the CHM. The Note also covers the highlights of a consultation process among the Informal Advisory Committee (IAC) to the CHM, the national focal points, and the CHM partners which produced an independent review of the CHM, a strategic plan, and a longer-term program of work. The full text of the Note is available at <http://www.biodiv.org/sbstta5/docs.html>. Click on report UNEP/CBD/SBSTTA/5/3.

The IAC met in Copenhagen, Denmark, in September 1999, to offer recommendations to be included in the Note; the IAC also met just before the SBSTTA meeting to refine those recommendations and prepare to offer others from the floor when the Note was considered. The U.S. Geological Survey represents the Inter-American Biodiversity Information Network on the IAC and participated in the Copenhagen meeting, hosted by the European Environment Agency, as well as the pre-SBSTTA IAC meeting in Montreal.

Recommendations included in the Note, plus those offered from the floor by the IAC and by SBSTTA delegates were incorporated into the draft recommendations to be submitted by the SBSTTA Chair to the Conference of Parties. The final draft recommendations were issued as paper UNEP/CBD/SBSTTA/5/L.3. Results from the SBSTTA meeting should be available shortly through the CBD Home Page <http://www.biodiv.org>.

The CHM is an international initiative of the Convention on Biological Diversity (1992). The CHM is designed to facilitate technical and scientific cooperation among countries and to provide global access to and exchange of information on biological diversity. Additional information is available at <www.biodiv.org/chm/>.

The Global Biodiversity Information Facility: An International Network of Interoperable Biodiversity Databases (continued from page 6)

searching for additional funding from international and regional bodies; and facilitate practical applications that demonstrate the utility of biodiversity data for addressing critical social concerns.

GBIF will work in close cooperation with established programs and organizations that compile, maintain, and use biological information resources such as Diversitas, Species 2000, the Integrated Taxonomic Information System, the Inter-American Biodiversity Information Network, Man and the Biosphere, and the Clearing-House Mechanism of the Convention on Biological Diversity. Data sets will remain under the control and management of those entities that compiled them. GBIF will be defined and supervised by a Governing Board consisting of representatives of the countries that join GBIF. The Board will be open to full participation by any countries that are willing to make the required financial contribution.

In order to provide easy access to GBIF data sets, it is recommended that each member country develop a national node (or nodes) where those data sets can be registered. Within the United States, the most likely national node is the National Biological Information Infrastructure (NBII). The proposed development of NBII-2, as outlined in the report by the President’s Committee of Advisors on Science and Technology titled Teaming with Life: Investing in Science to Understand and Use America’s Living Capital is highly congruent with GBIF. GBIF is also a component of the U.S. interagency initiative on Integrated Science for Ecosystem Challenges, through which funding for GBIF activities will be sought by federal agencies.

The science ministers from OECD countries endorsed GBIF in June 1999 and called on interested countries to develop it. Consequently, an Interim Steering Committee, composed of countries that are likely to be founding members of GBIF, is finalizing GBIF’s operating principles and intends, by the middle of 2000, to issue an open invitation to the countries of the world to join GBIF. The target date for establishing GBIF is January 1, 2001.

Please note: Portions of this article first appeared in the ASC Newsletter, June/August 1999, vol. 27 (3-4), pp. 6-7.
NBII Metadata Training Schedule

A metadata training workshop is usually a full two-day course in which participants engage in in-depth discussions, and receive hands-on training, on metadata activities related to the Federal Geographic Data Committee’s metadata standard and the NBII biological metadata profile, MetaMaker, and the NBII Clearinghouse.

Training Schedule

<table>
<thead>
<tr>
<th>2000</th>
</tr>
</thead>
<tbody>
<tr>
<td>National Park Service, San Antonio, TX</td>
</tr>
</tbody>
</table>

New Metadata Training URL!

The URL for NBII metadata training has changed. Now, you can get the latest information about the location and dates of metadata training classes and related matters at <http://www.nbii.gov/metadata/training/index.html>. Or, you can contact Sharon Shin, NBII Metadata Training Coordinator, USGS Center for Biological Informatics, P.O. Box 25406 MS 302, Denver, CO 80225; phone 303/202-4230; fax 303/202-4219; e-mail: sshin@nbii.gov.

See You At ESA in Snowbird

Are you going to the Ecological Society of America Annual Conference in Snowbird, Utah, which is being held from August 6-10. If so, please stop by the NBII exhibit and say hello. We’ll be offering a variety of NBII handouts, demonstrations, and special prize drawings.