

## BRIEFING NOTE OF THE GLOBAL BIODIVERSITY INFORMATION FACILITY SCIENCE SYMPOSIUM: 5 OCTOBER 2011

The Global Biodiversity Information Facility (GBIF) Science Symposium 2011, was held in Buenos Aires, Argentina on 5 October 2011, under the title “GBIF at 10: Reaping Benefits for Science and Society.” Once a year, the GBIF sponsors a Science Symposium in association with its Governing Board meeting and associated committees. The 2011 Symposium focused on the experience gained during GBIF’s first decade of work and the future of biodiversity informatics. Topics addressed during the symposium included insights from eco-informatics studies on climate and earth’s biodiversity, and presentations on lessons learned and experiences in building and managing biodiversity information networks in different countries across the globe.

GBIF was created in 2001 with the objective of making the world’s biodiversity information freely and universally available for science, society and a sustainable future. GBIF unites networks in 57 countries and 47 organizations throughout the world and has been successful in building an open and global infrastructure for biodiversity informatics, allowing for data exchange and networking, and helping participants build their national informatics capacities and capabilities. GBIF is governed by: a Governing Board, where countries that make financial contributions to GBIF have a vote; an Executive Committee, composed of the different Committee Chairs; several Advisory Committees to the Board, for example, on budget, nodes, rules and science; as well as a Secretariat.

GBIF held its 18<sup>th</sup> Governing Board Meeting on 4-6 October 2011. The Governing Board endorsed the appointment of Donald Hobern as GBIF’s new Executive Secretary (taking office in January, 2012), and approved the GBIF Work Programme 2012-13, and a revised budget for 2012. Joanne Daly was re-elected as Chair of the GBIF Executive Committee.

More than 100 participants, from academia, museum collections, governments and biodiversity networks took part in the Science Symposium in parallel to the Governing Board meeting, which was also accompanied by GBIF committee meetings, a three-day training session and a two-day meeting of GBIF node managers and staff.



Participants during the Symposium

### OPENING REMARKS

Leonard Krishtalka, Chair of the GBIF Science Committee, opened the Science Symposium, “GBIF at 10: Reaping Benefits for Science and Society,” recalling GBIF’s mission and first decade of success in creating networks among a diversity of scientists, technologists, government officials and science administrators, to advance science and society for a sustainable future.

Marta Rovira, President of Argentina’s National Council of Science and Technological Research (CONICET) recalled Charles Darwin was one of the pioneer naturalists who recorded Argentina’s biological diversity, and said the country is now striving to improve data availability, both in qualitative and quantitative terms, in light of technological advances. She highlighted how GBIF helped Argentina put in place a national network of biodiversity information and build an open and collective effort towards biodiversity information conservation.

### GBIF EBBE NIELSEN PRIZE WINNER 2011

Jens-Christian Svenning, Department of Biological Sciences, Aarhus University, Denmark, was awarded the 2011 Ebbe Nielsen Prize for his work on climate and biodiversity using eco-informatics. He presented his latest work on species’ reactions to climate change, which looks at scientific scenarios for climate change over the next century, and, based on past climate disturbances, estimates how species are likely to react to temperature increases, finding patterns that are non-linear and even include species moving in opposite directions.

He explained science is facing a new paradigm driven by the rapid acceleration of data availability and computer power, allowing the analysis of phenomena at a global scale. For



Jens-Christian Svenning, Department of Biological Sciences, Aarhus University, Denmark

example, he explained that studies on biodiversity's sensitivity to climate change in the tropics looked at the effects of massive changes in temperature on biodiversity in the past, for example during the glacial ages. He said that, in this respect the findings show some species respond very strongly to drying, but it is not clear how much they react to temperature *per se*, noting biological interactions and precipitation may be more important than temperature in driving species movement. He cited the example of a study on African Palms that combined all spatial data available on palm occurrences and highlighted the relevance of water-related factors to palm occurrence, as well as barriers to species dispersal that may prevent a species from expanding even in the presence of favorable climatic conditions.

He highlighted another question addressed in his research as being the degree to which endemism is linked to past climate change, and said his finding was that areas with high endemism coincided with those that did not suffer rapid climate change in the past. He noted that this finding raises concern because areas with high endemism, such as the Western Amazon, are among those identified by scientific models as likely to suffer rapid climate change during the next decades.

Responding to questions from the audience, Svenning said his findings did not show a homogeneous pattern of movement by communities of species, but rather individual reactions by each species. He said some species' reactions to climate change may be related to the response of related species, but this was not a unified trend, and noted further studies on species interactions are needed.

The audience also discussed the inevitability of massive extinctions in light of climate change that is already occurring, and whether the label 'invasive species' made sense in the context of climate change. They concluded that 'invasive species' should be targeted only when they generate problems (such as vectors for disease), as climate change will likely bring a widespread movement of species beyond their present ranges.

## **GBIF AND TEN YEARS OF BIODIVERSITY INFORMATICS IN ARGENTINA**

Edgardo Romero, Director of the Bernardino Rivadavia Natural Science Museum of Argentina (MACN), Argentina, and head of delegation for Argentina to the GBIF Governing Board, presented on Argentina's experience in digitizing and linking data available in museum collections, noting GBIF's significant contribution to the success of this endeavor. He noted the value of GBIF's training and technological tools, as well as the positive contribution of university students to the digitization efforts.

Martin Ramirez, General Curator of MACN, presented on several projects for biodiversity information digitization in Argentina, including work at Instituto Lillo to design algorithms to detect areas of high endemism, as well as research on insects and sharks using eco-informatics.

Responding to questions from the audience on how to design accurate models when a large amount of biodiversity information is still unavailable, Ramirez explained that the algorithms developed are able to account for data gaps.

## **DEVELOPING A BIODIVERSITY INFORMATION FACILITY FOR ECOLOGICAL SUSTAINABILITY AND ECONOMIC GROWTH: THE INDIAN INITIATIVE**

Vishwas Chavan, GBIF Secretariat, on behalf of the Indian delegation, presented on India's experience in developing a national biodiversity information facility. Noting much of existing biodiversity data is still locked up and not freely available in electronic format, he highlighted the importance of generating systems that: encompass linguistic variations; compensate data collection efforts; and are easy to use. Highlighting that India's Biological Diversity Act of 2002 mandates every village and local government to develop, update and inventory its own biodiversity, he said the challenge ahead is how to ensure all this information is digitized and available for free and open access.

Chavan also emphasized the importance of ensuring datasets are relevant to decision-making needs, to ensure long term funding for biodiversity collections, and highlighted the need to change a past culture of closeness in this field.

Responding to questions, Chavan said the Indian GBIF node is targeting a broad-based constituency, working to incorporate data from other spheres, like the remote sensing community and data custodians at all levels.

## **THE SOUTH AFRICAN BIODIVERSITY ADVISOR: IN SUPPORT OF THE USER**

Selwyn Willoughby, Director of Biodiversity Information Management at the South African National Biodiversity Institute (SANBI) analyzed how to make information available to users in a way that can be useful for modern needs. He said the

information presented by South Africa's Biodiversity Advisor is retrieved from a variety of sources, and SANBI manages and disseminates the information, highlighting they have more than 40 million free-access records that, together with geo-referencing and mapping capabilities, are able to present information valuable to decision makers, for example on threatened ecosystems, land uses, etc. He also highlighted the value of GBIF's technical support and seed grants to provide impetus to the South African network, which now receives around 30,000 queries per month and is used by authorities and the private sector for land planning decisions and assessments.

In response to a question from the audience Willoughby highlighted that the private sector is actively contributing to the system as publishing data is perceived as a means for validating their findings and assessments.

### ***THE ATLAS OF LIVING AUSTRALIA - A NATIONAL BIODIVERSITY INFORMATION FACILITY***

Donald Hobern, forthcoming GBIF Executive Secretary, presented on the "The Atlas of Living Australia" and highlighted the size of this platform, which includes data from different sources, such as museum collections, photographic records and observational data. He said a user-needs analysis led them to ascertain that main trends in information demand relate to species distribution and spatial information, used for land- and species-management decisions. Other issues that were identified as important were resolving scientific nomenclature discrepancies and dealing with the vast number of unnamed species in Australia. He also highlighted the value of amateur observation, as well as the need to be aware that some communities fear providing localization data may lead to valuable species being disrupted or collected.

He commented on Australia's experience in automating data digitization, for example through scanning of herbaria. He also referred to the many levels of data made available on the same platform, which in the near future will allow the retrieval of taxonomic records of a species, together with photographs and even literature references, allowing users to use biodiversity information for a vast array of uses.

Responding to a question on the impact of the digitization effort, he said, while impact is obviously difficult to measure, there are clear savings in terms of the efficiency of information retrieval and processing.

### ***GBIF SPAIN***

Benjamin Sanchez, Spanish Ministry of Science, presented on the history of the Spanish GBIF node, noting both political commitment and efforts to develop a wide network of centers encompassing Spain, Portugal and several countries in Latin

America. He noted GBIF experience will contribute to the European LifeWatch initiative and highlighted GBIF's reputation for neutrality, credibility and openness.

Arturo Ariño, Universidad de Navarra, noted their database reaches seven million registries with a network of 152 databases in Spain. He highlighted efforts to audit the quality of the information, as well as positive experiences in providing relevant data for decision making by political and management authorities, for example, on the effect of nitrogen pollution on ecosystems.

Responding to a question from the audience, Ariño explained that their system allows them to integrate data from different sources for modeling purposes.

### ***SEP-CEPDEC: LESSONS LEARNED FROM A CONTRIBUTION TO GBIF OUTREACH IN AFRICA, THE INDIAN OCEAN AND SOUTH-EAST ASIA***

Eric Chenin, Sud Expert Plantes (SEP), France, presented on the experience of SEP working with GBIF's Capacity Enhancement Programme for Developing Countries (CEPDEC) through which they engage in activities on training, workshops, research and support for 22 countries in Africa and Asia.

Pierre Radji, Director and Curator of the Botanical Garden and Herbarium of Togo, shared his experience in creating a network of biodiversity information nodes in Africa, noting challenges in digitizing and collecting specimens. He highlighted, however, that data provided through the network is already being used for research on malaria and other vector-borne diseases.

On lessons learned, Eric Chenin, was satisfied with the results of efforts to train node managers, hoping to further enhance the availability of information and engage smaller institutions in joining the platform. He said future work would include private sector and industry.

Questions and comments from the audience highlighted the relevance of biodiversity information to complying with REDD+ requirements, and the possibility of retrieving important amounts of data on African biodiversity from European collections.

### ***AN OVERVIEW OF ANTARCTIC BIODIVERSITY NETWORKS***

Bruno Danis, Royal Belgian Institute of Natural Sciences, presented an overview of two Antarctic biodiversity networks: the Marine Biodiversity Information Network for the Scientific Committee on Antarctic Research (SCAR MarBIN); and the Antarctic Biodiversity Information Facility (ANTA BIF). He explained SCAR MarBin was the result of the census on Antarctic marine life, which engaged over 400 scientists, and that the idea behind these nodes is to offer free and open access

to the data and technology on Antarctic biodiversity and expose the data and metadata in the most number of contexts, to enhance their usefulness. He noted they have geospatial data for 1.3 million records and have received more than eight million hits by users retrieving such data. He also highlighted advances in providing single pages that compile all data available on the network on a particular species, including distribution maps, microscopic details and photographic images, while noting a large information gap regarding the biodiversity of the deep sea.

He also provided an overview of future activities, which include developing predictive maps with more than 90 environmental parameters. Stating the objective to achieve 100% “open source” and to improve data retrieval efficiency, he explained cloud computing will significantly improve efficiency by storing datasets in a single place and including simple badging tools to ease the retrieval process. He said the biggest challenge nowadays is not technological but cultural, saying the Antarctic system provides an example of collaboration to provide information that is “open, discoverable, linked, useful, interoperable and safe.”

Responding to a question from the audience he said until now the databases have not retrieved information ‘from,’ but only provided information to, GBIF, and also clarified they do not have paleontological data at present but have the capacity to hold it.

#### **BUILDING A DATA SHARING COMMUNITY: VERTNET**

David Bloom, VertNet, United States, recalled the efforts to create a national biological information infrastructure linking universities, museums and hundreds of biological collections in his country. He explained how individual networks came together to share information and applications, thus enhancing the value of collections and conservation of resources.

At present, he said one of VertNet’s main challenges is how to manage a growing interest by other collections to join the network, while maintaining data quality and ensuring speed of retrieval and technological integration are not impaired. In this respect, he highlighted the building of a new architecture to create a new cloud-based platform that will allow a significant growth in the network’s capacity to handle data from different sources, providing both data storage and processing services, while reducing costs of maintenance of the system. He explained Vertnet mobilizes more than 85 million specimen records, with around 2.5 million used each week.

Responding to a question from the audience on user feedback, Bloom said feedback from scientists is usually very positive in terms of data accessibility, while usual queries relate to fitness for use and validity of the data.

#### **CLOSING PANEL**

A final panel discussion on the expectations for the next decade was led by Mark Graham, Canadian Museum of Nature, who highlighted radical technological improvements, such as the human genome project, generate a new wave of challenges to preserving and managing vast quantities of information and providing for links between science and policy making. Among his predictions for the next decade, he challenged the audience to debate whether in ten years all major museum collections will be accounted for digitally and shared through GBIF. David Bloom, agreed that it is likely that the majority of collections will be digitized, and added that in order to validate and verify the massive amounts of data, citizen science will grow in relevance as there are not enough scientists and PhD’s for the enormous task ahead. Edgardo Romero hoped not just major collections, but also small collections in towns throughout Latin America, will be able to access and provide data to GBIF. Arturo Ariño added that automated classification, for example through software interpreting imagery, would be increasingly relied on to incorporate new data. Participants and panelists also referred to photo-documentation through hand-held devices as a growing source of information, alongside traditional collection or observation.

Participants in the audience, however, cautioned that given resource constraints not all information is worth digitizing, and there will be an unavoidable priority setting when deciding what to digitize within collections. They debated how to prioritize digitization of data, with some considering that priority datasets are those of interest as indicators of climate change, food security or ecosystem resilience, and Svenning cautioning on the risks of *a priori* defining what species may be of interest as eco-informatic models benefit from the widest variety of datasets.

Participants also discussed the role of society in digitizing data, addressing language barriers, and ensuring information is tailored to provide answers to key policy questions, and Svenning emphasized that citizen science will be the only way to get the largest possible amount of data in electronic format

On how to manage vast quantities of data, participants discussed current storage capacity limits, and the need to learn from other sciences that manage large amounts of data, like astronomy, noting this is a field where scientists are used to interacting with super computers, as well as to receive input from citizen science.

The discussion concluded with a question from Leonard Krishtalka on how to deal with the two facets of modern biology: that governed by Darwinian principles, and “quantum biology.” Participants then gathered for a cocktail and gave their best wishes to the current GBIF Executive Secretary, Nicholas King, who will leave in December 2011 after serving in his post for four and a half years.