

From the World Directory of Collections of Cultures of Microorganisms towards Mash-up of Biodiversity and Sequence Databases

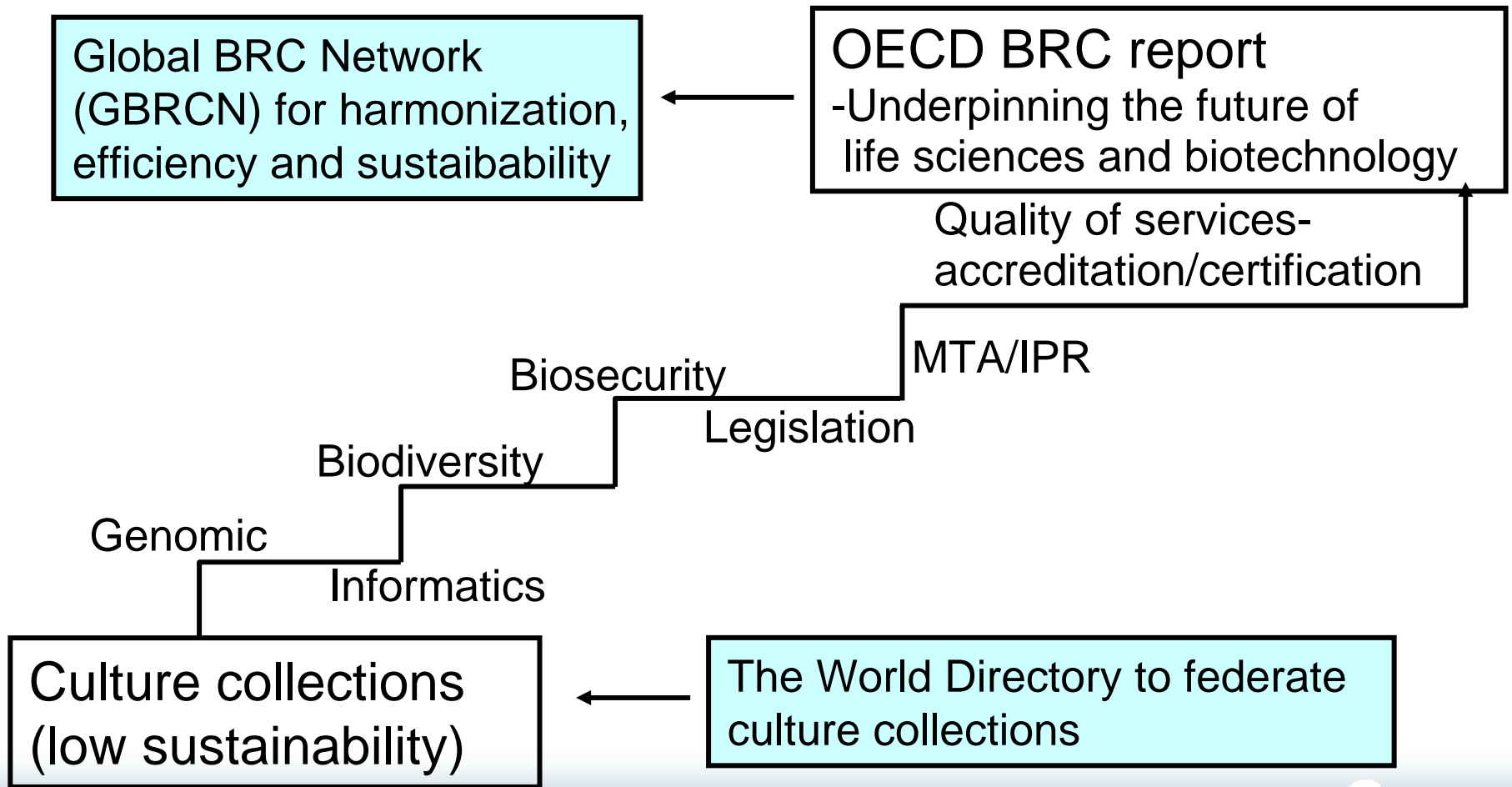
Director, WFCC-MIRCEN World Data Centre for Microorganisms
Vice-chair, Governing Board of Global Biodiversity Information Facility (GBIF)
Center for Information Biology and DNA Data Bank of Japan (DDBJ)

National Institute of Genetics

Hideaki Sugawara






Evolution of Culture Collections to BRC



Prototypes of information network for GBRCN.

- ◆ Global networks, e.g.
 - WFCC-MIRCEN* World Data Centre for Microorganisms (WDCM)
 - Global Biodiversity Information Facility (GBIF)
- ◆ Regional networks, e.g. EBRCN
- ◆ National networks, e.g. CCCM, JSCC, TNCC, UKNCC

* WFCC: World Federation for Culture Collections
MIRCEN: Microbial Resources Centers Network

HOME About WFCC What's New Statutes and Bylaws Strategic Plan (PDF) Newsletter Publications Workshops Conferences	
	<ul style="list-style-type: none">• Annual Conference of the Association for General and Applied Microbiology (2006/3/19-22) [PDF] [Word] <i>uploaded!</i>• EBRCN transport paper WP5 Revision 15Jun05.doc <i>uploaded!</i>• The Biology of Vibrios: Biodiversity, Genomics, Disease/Epidemiology, Ecology, and Applications, Ghent, Belgium (2005/11/07-08) <i>uploaded!</i>• Australian Society for Microbiology 2005 (2005/09/25-29)• VI International Conference on Environmental Pollution (2005/09/20-25)• II International Conference on Microbial Diversity: current situation, conservation strategy and biotechnological potentialities (2005/09/20-25)• Congress of PERMI and ACLAB-3 (2005/08/25-27)• What's New
About WFCC <ul style="list-style-type: none">• Please join us (WFCC)• WFCC Executive Board• WFCC Task Forces and Activities• WFCC Statutes• WFCC Bylaws• Brochures	<ul style="list-style-type: none">• Messages/Proposals from WFCC Executive Board• WFCC Newsletter• Publications on culture collections and microbial resource centers• WFCC Workshops and Conferences• Documents on bioterrorism and biosecurity• Calendar of Events relevant to microbes and more
 <p>We are here to help collections in difficulty (Please visit What's New!)</p> <ul style="list-style-type: none">• WFCC membership application•  Contact us!	<ul style="list-style-type: none">• Home Pages of Culture Collections in the world• Statistics on Culture Collections in the world• Federations, Societies and Networks of Microbial Resource Centers• WFCC-MIRCEN World Data Centre for Microorganisms

501 culture collections in 66 countries

175 culture collections publish catalogues

154 culture collections have Web pages



WFCC-MIRCEN World Data Centre for Microorganisms (WDCM) provides a comprehensive directory of culture collections, databases on microbes and cell lines, and the gateway to biodiversity, molecular biology and genome projects.

Registration of culture collections and their holdings

WDCM assigns culture collections WDCM numbers, cf. accession number of sequence data assigned by INSDC

- About WFCC
- What's New from WFCC
- Search Engines and Portal Sites
- Nomenclature and Phylogeny
- Databases of Resource Banks
- e-Learning System
- Sequence/ Phylogenetic Analysis
- Genome Projects
- Biodiversity
- Biosafety
- Publishers
- Others
- Inter-Union Bioinformatics Group Report
- OECD BRC Report (PDF)
- Global Biodiversity Information Facility (GBIF)
- Download e-Workbench

SEARCH ENGINE

- Simple Search
- Search by Fields
- AHMI
- Lineage search *NEW!!*

Home Pages of Culture Collections in the World

REGISTRATION/ UPDATE

- CCINFO
- STRAIN

WDCM is sponsored by/in cooperation with:

- | | |
|-------------|----------|
| ■ WFCC | ■ NEDO |
| ■ MIRCEN | ■ CODATA |
| ■ CIB | ■ UNESCO |
| ■ JST/BRnet | ■ UNEP |

Please search menu(s) in this site by entering keyword in the window below



Global search by

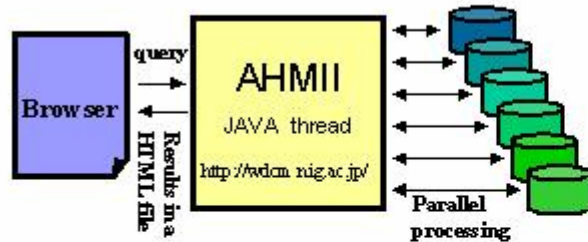


Home Page in Japanese

Search engines

WDCN wrapped several databases for one-stop query for a search engine named AHMII

A parallel access to multiple data sources in the Internet



The following 3 databases, which would be useful for microbiologist, are available now!!

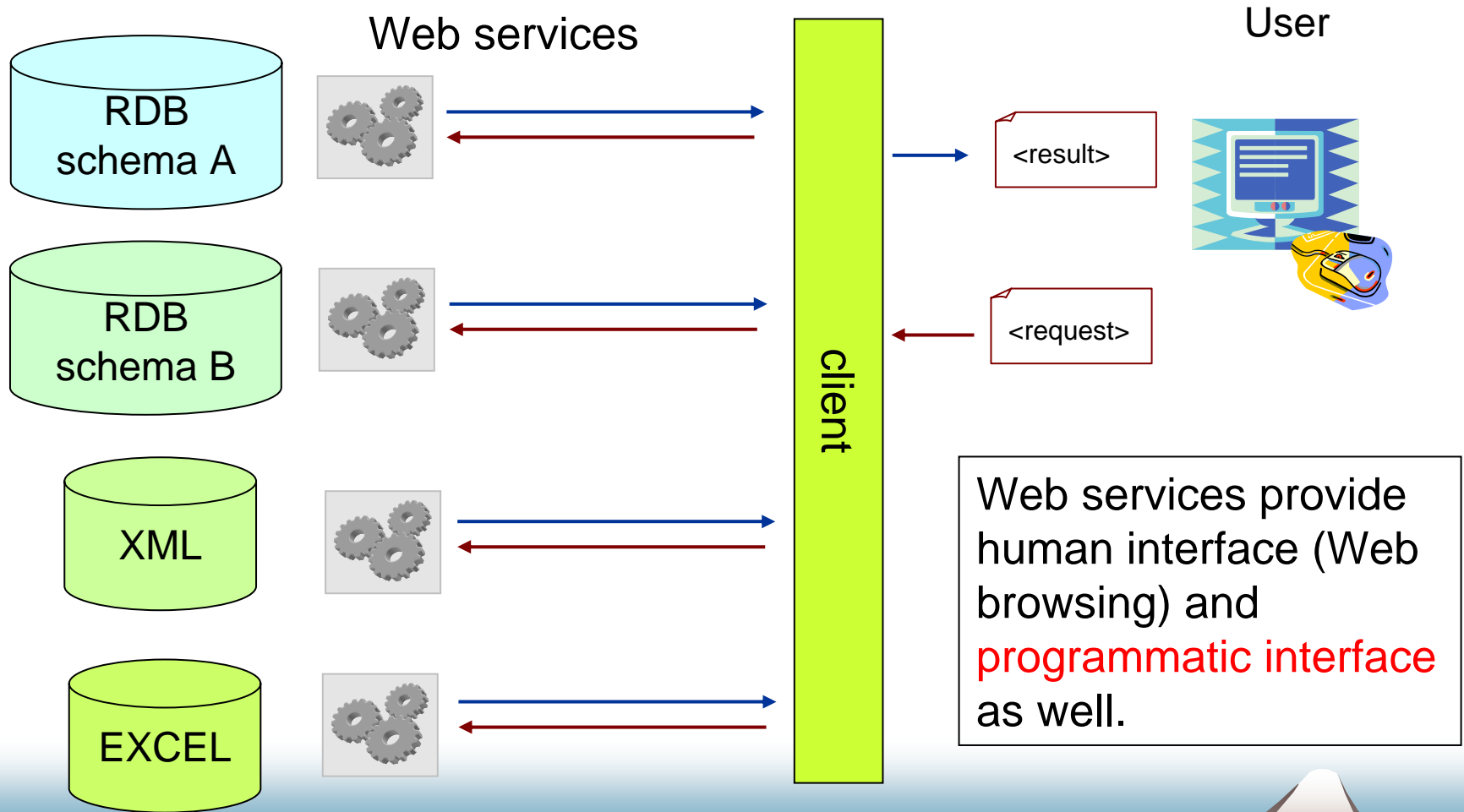
For Bacteria

For Fungi & Yeasts

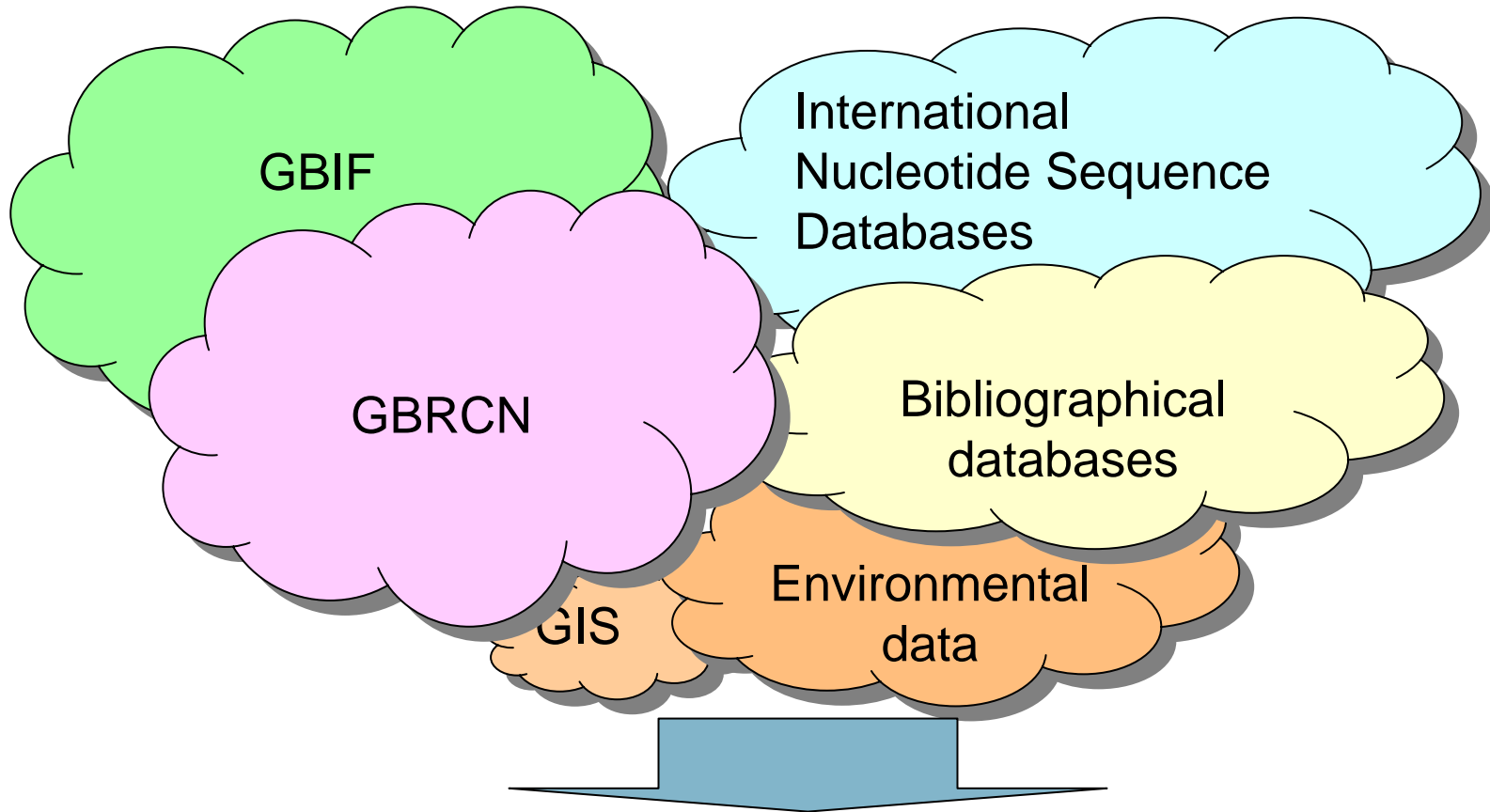
For Cell Lines

- ◆ AHMII call CGIs of the member Web sites.
 - ◆ AHMII requires neither change nor expansion of the member Web sites.
 - ◆ AHMII counts on query functions of the member Web sites.
 - ◆ AHMII does not display results in a standardized format.
 - ◆ WDCM has to maintain the program and the list of the member Web sites.
- More sophisticated system than AHMII is now feasible

Web services provide more sophisticated way to aggregate biological data resources

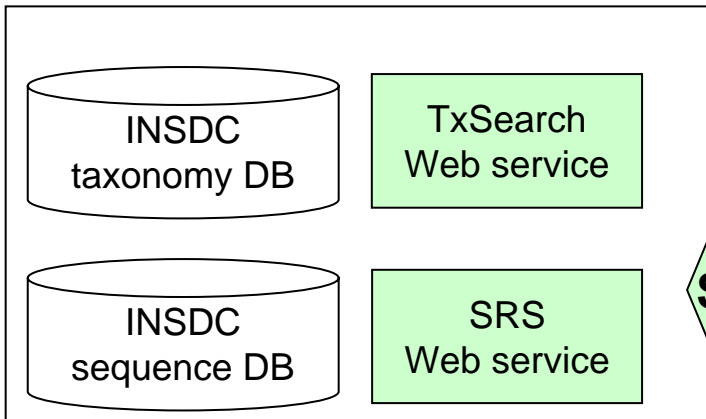


Web services expands beyond BRCs

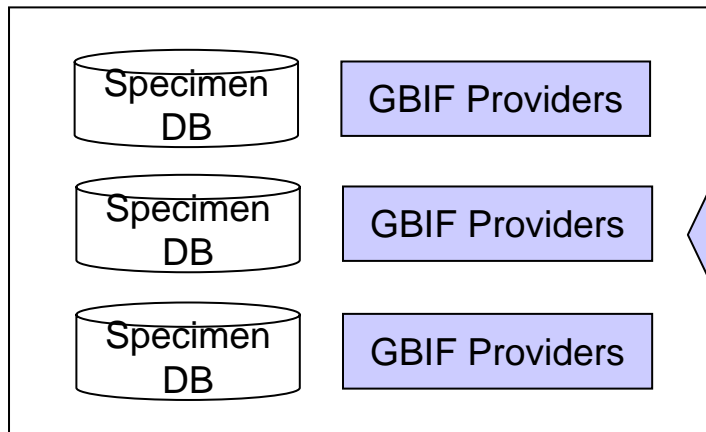


Aggregation of taxonomy, phenotypes, sequence data, environmental condition, location, references and even socio-economical data will be the base of the problem solving environment for R&D and decision makers.

Examples of mash-up: INSDC and GBIF



SOAP



DiGIR

Dynamic linkage of the sequence database and GBIF data providers

Dynamic Linkage of the Taxonomy File used by the International Nucleotide Sequence Databases (DDBJ/EMBL/GenBank) and the data provided by Global Biodiversity Information Facility (GBIF)

[Japanese Page](#)

Scientific name:

Option of the output

without accession number of sequence data of INSD (default)

with accession number of sequence data of INSD:
The system will not display the list of accession numbers, if it takes more than 2 minutes to return the result. It happens if the number of hits is comparatively large.

- Sample of query: Please specify the latin name only
[Homo sapiens](#) - human
[Pieris rapae](#) - cabbage butterfly
[Apis mellifera](#) - honeybee
- Wild card usable in the query: *

[GBIF Data Use Agreement](#)

Search

Scheme of the dynamic linkage

Web services by SOAP servers

TxSearch() Web service

Local Index

MySQL

DiGIR protocol (XML)

More fine linkage of specimen including microorganisms cultures and sequence data by “structured vouchers”

<institution-code>	- abbreviation of the archiving institution
<collection-code>	- collection within the institution (possibly null)
<specimen-id	- specimen identifier within the collection

/specimen_voucher=“<institution-code>|<collection-code>|<specimen-id>”

- museums
- herbaria
- culture collections
- stock centers
- germplasm repositories (seed banks)
- frozen tissue banks
- zoos/aquaria/botanical gardens
- DNA banks
- personal collections
- e-voucher archives

Example of mash-up: GBIF and mapping tool

Please select your interested one :

- All
- Aster ageratoides
- Aster brevipes
- Aster hispidus
- Aster microcephalus



Example of mash-up: GBIF, data analysis tool and mapping tool

Biodiversity Analysis Tool - Define Analysis

Data Source

Please choose one of the data sources from the pulldown menu first.
Then you can set the parameters in the following pulldown menu.

Grid Size (e.g. 1 Degree)

Class (e.g. Magnoliopsida)

Order (e.g. Asterales)

Family (e.g. Asterac

Analysis Type Richness
 Endemism
 Taxonomic Diversity

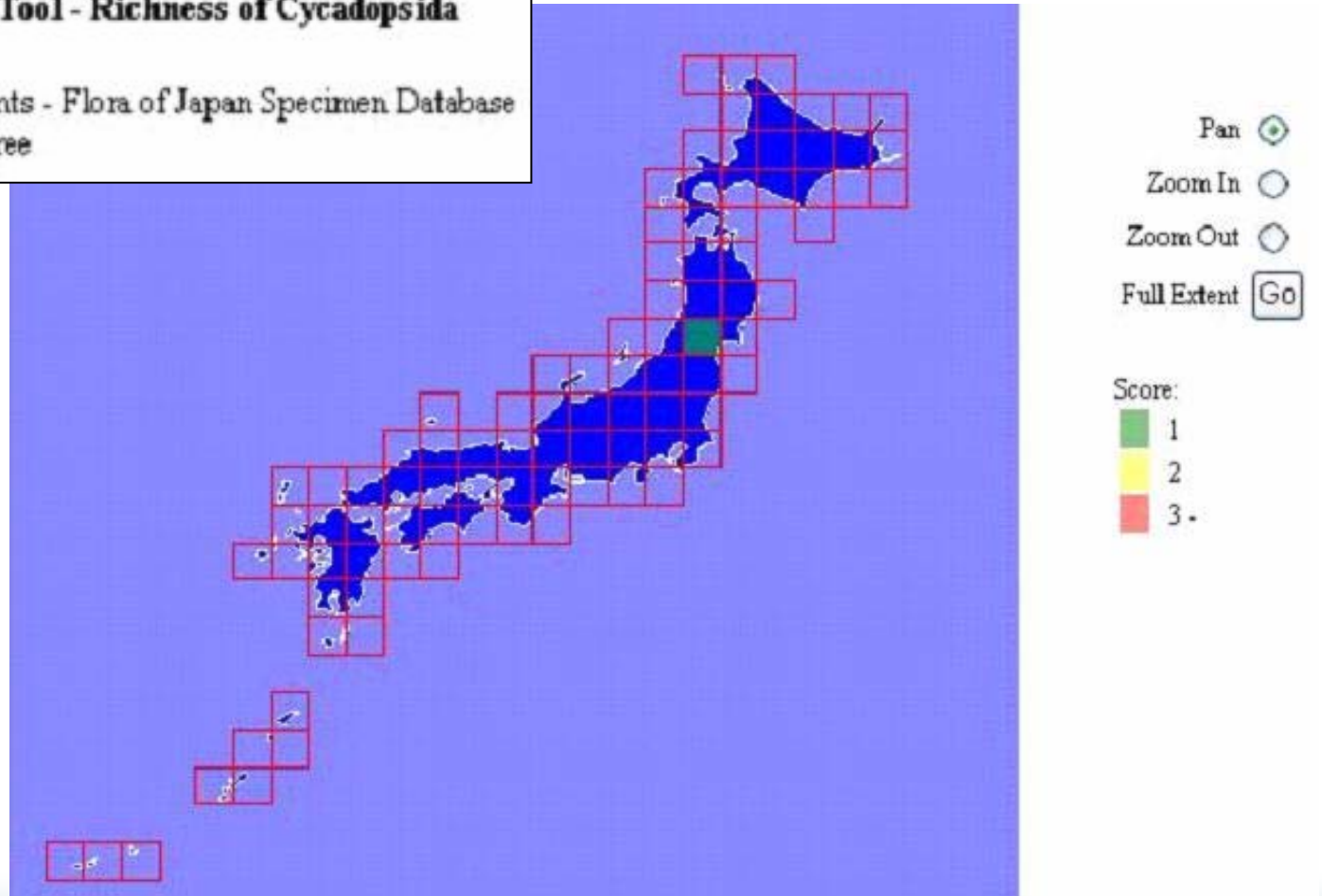
Output MapServer
 Google Earth

BAT has been implemented in NIG based on collaborative work with Australia.

With MapServer

Biodiversity Analysis Tool - Richness of Cycadopsida

- Data Source: Plants - Flora of Japan Specimen Database
- Grid Size: 1 Degree



Display richness of species in Google earth

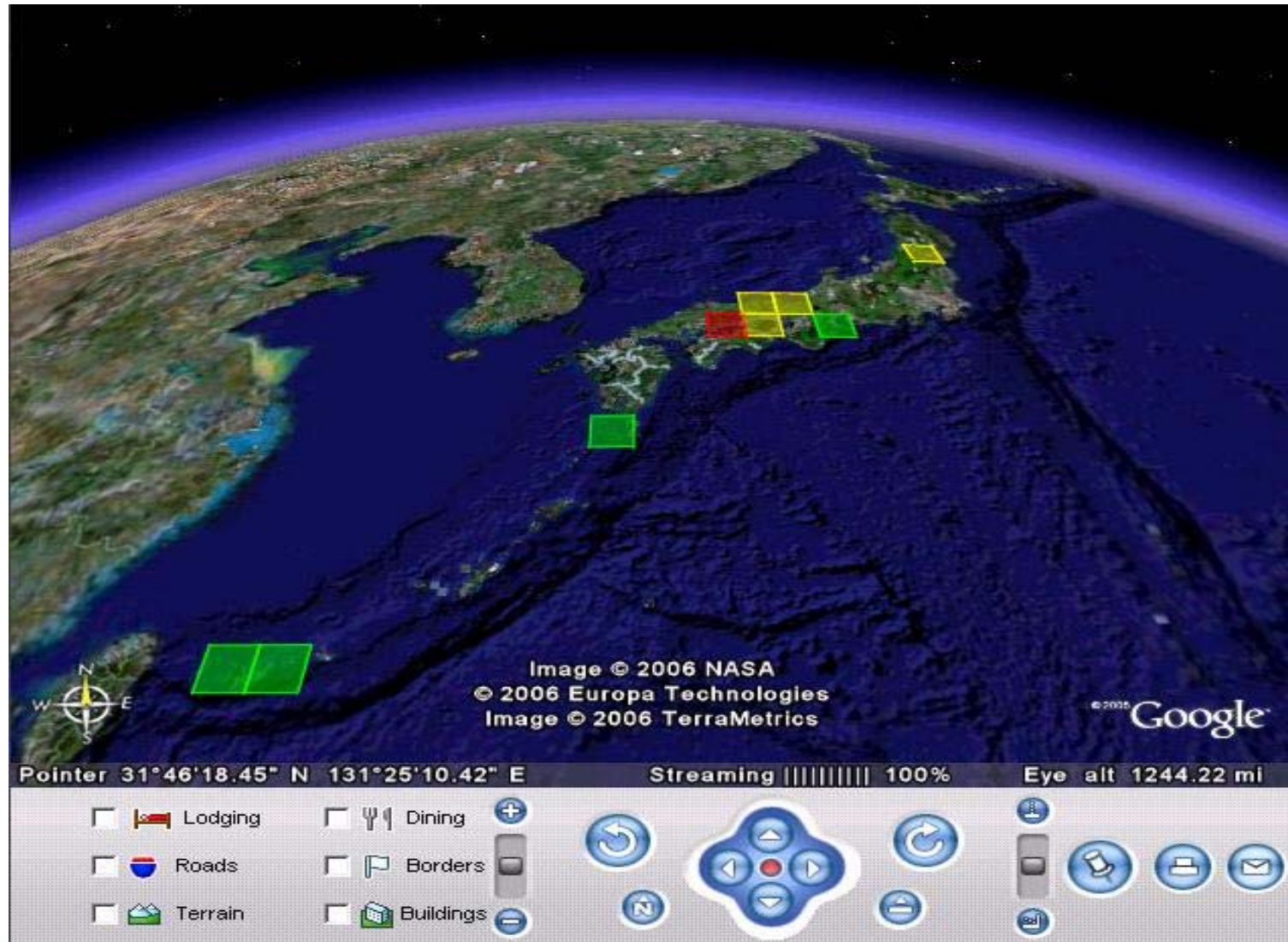
Specimen data were from a plant data file in Japan GBIF node.



Display richness of species in Google earth

Specimen data were from a plant data file in Japan GBIF node.

Grids and color codes are displayed based on outputs from BAT system



Summary

- ◆ The next generation WDCM will depend on a distributed system based on standard data schema and data transfer protocol.
- ◆ The question is who designs and maintains the standards.
- ◆ The next generation WDCM will provide a stable database that supports LSID (Life Science IDentifier) for the linkage of cultures and their data on genotypes and phenotypes.
- ◆ The question is who funds these activities for a long term.