

Biodiversity Data Management Pilot Project University of Otago Library

October 2008



We arrived at this Project as a result of informal interviews with Otago researchers

We don't have a National strategy or an Institutional policy on data management

Acknowledge:

New Zealand Digital Strategy: 2005

New Zealand Digital Content Strategy: 2007 (few references to data)

Images by University of Otago researchers: Keith Probert (Marine Science), Janice Lord (Botany), Phil Bishop (Zoology)

Otago Biodiversity Data Project

1. Who
2. What
3. Why
4. Where
5. How
6. When

University of Otago Library

1. **Who** – the people
2. **What** – the challenge
3. **Why** – the discipline
4. **Where** – the landscape
5. **How** – **the Project**
6. **When** – the goals

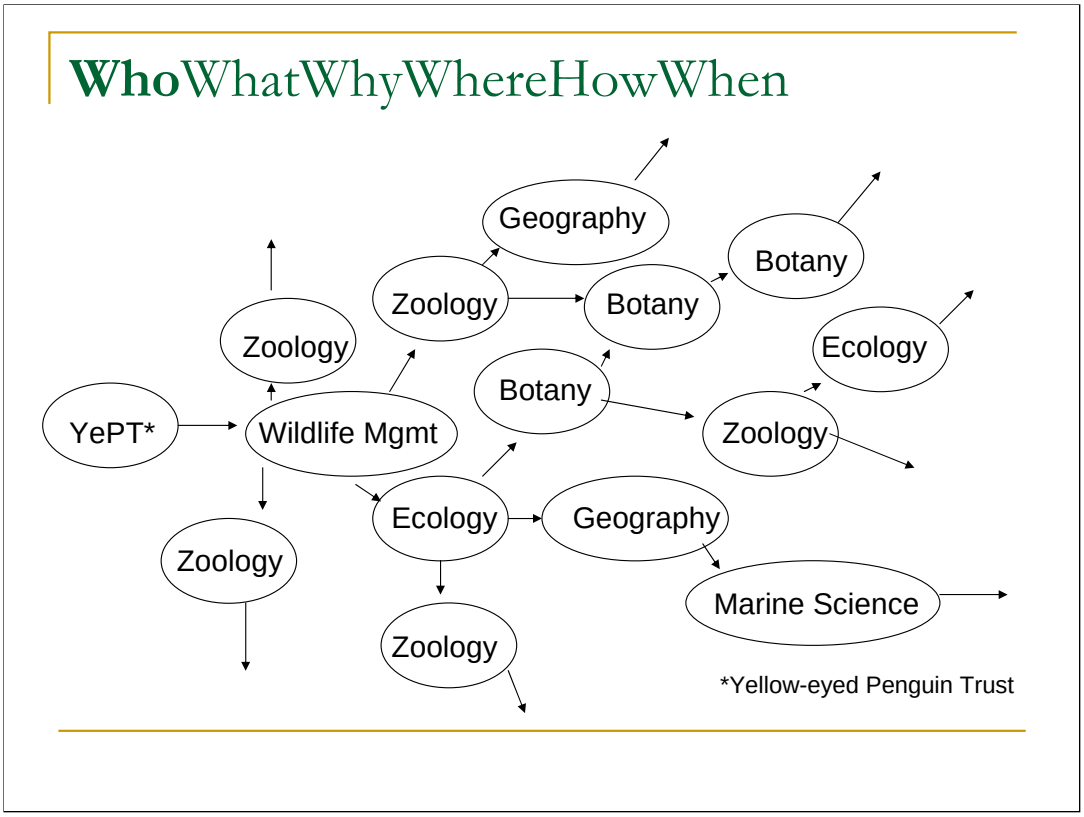
Image by Russell Poulter, Biochemistry Dept, Uni of Otago

Project genesis

If the Library could do one thing that would better support your research what would that be?

Open ended interviews with Life Scientists: June-August 2007

Image by Russell Poulter, Biochemistry Dept, Uni of Otago



Conversations with Otago Scientists – arrows representing connections and further conversations (off the slide)

Who**What**WhyWhereHowWhen

Documents to be scanned and made **available:** Zoology Wildlife series; Otago theses and dissertations; Postgraduate reports, dissertations; Fish and Game Society documents

Database to be **created and managed:** Otago marine resources – linking: card index, photos, specimens, research; Yellow-eyed Penguin resources; Orokonui resources; Lepidoptera specimens

Databases to be managed and **updated:** Frog disease database; National amphibian database

Databases to be **made available:** National frog identification image database; Dunedin land mapping project; Topo-cadastral maps; Otago mollusc database; Dune database – access

Species level approach to: Otago Giant and Grand Skink repository; Takahe, Saddleback

Enhanced access, linking: Biodiversity portal to existing OU biodiversity databases; Otago academic bibliographies; Specialist collaboration – academics (content) and library (infrastructure)

Collaboration with other local database providers: Repository for DOC publications, access base maps; DCC District Plan survey data – access

Summary of 'wish-list' of researchers (unpublished print – raw data)

Who**What**WhyWhereHowWhen

Informal survey of Otago life scientists (2007) researchers have stated:

The University of Otago is a significant producer of biodiversity data, but that this data can be difficult to access and to share beyond immediate colleagues.

Image: Keith Probert, Marine Sciences, Uni of Otago

WhoWhat**Why**WhereHowWhen

New Zealand is a '**Biodiversity Hotspot**'

Endemic (native species, unique to NZ)

Bats (only native mammals)	100%
Frogs (world's most endangered list)	100%
Lizards (largest vertebrate group)	96%
Insects & marine molluscs	90%
Vascular Plants	80%
Birds	25%

Note:The UK has 1 endemic bird species & a few plants

NZ statistics from DOC website:
http://www.biodiversity.govt.nz/picture/biodiversity/index.html

Image by Ian Jamieson, Zoology Dept, Uni of Otago

Images supplied by University of Otago researchers: L to R, top to bottom:

1. Torrentfish by Bruno David (courtesy of Gerry Closs)
2. Mollymawk by Keith Probert
3. Saddleback by Ian Jamieson
4. Stalk-eyed Mud Crab by Keith Probert
5. Takahe by Ian Jamieson
6. Serolis by Keith Probert – deep sea isopod from Otago Canyon
7. Snares Island Penguins by Ursula Ellenberg
8. Giant Kokopu by Eric Hansen (courtesy of Gerry Closs) - the largest galaxiid in the world (up to 40 cm)
9. Brown Tree Frog by Phil Bishop – introduced from Tasmania to Greymouth only once (1875), now common in South Island
10. Yellow-eyed Penguins by Ursula Ellenberg

Tree of Life: information about the diversity of organisms on Earth, their evolutionary history (phylogeny), and characteristics. www.tolweb.org/tree/

EDIT (European Distributed Institute of Taxonomy): to help to reduce the fragmentation in European taxonomic research and expertise and to coordinate the European contribution to the global taxonomic effort

www.e-taxonomy.eu/

Wikispecies: free species directory http://species.wikimedia.org/wiki/Main_Page

GBIF (Global Biodiversity Information Facility): http://www.gbif.org/

Amphibiaweb: retrieve information relating to amphibian biology and conservation www.amphibiaweb.org/about/index.html

BHL: Ten major natural history museum libraries, botanical libraries, and research institutions…digitize the published literature of biodiversity held in their respective collections. www.biodiversitylibrary.org/Default.aspx

GLORIA (Global Observation Research Initiative in Alpine Environments): to establish and maintain a world-wide long-term observation network in alpine environments. http://www.gloria.ac.at/

uBio (Universal Biological Indexer and Organizer): to create and utilize a comprehensive and collaborative catalog of known names of all living (and once-living) organisms. http://www.ubio.org/

Atrium – biodiversity information system: (by Uni of Texas) provides researchers with tools to collaborate worldwide in real-time using web technology and purpose-built tools to streamline the research process.

www.atrium-biodiversity.org/

EOL (Encyclopedia of Life): ecosystem of websites that makes all key information about all life on Earth accessible to anyone, anywhere in the world. http://www.eol.org/

OBIS (Ocean Biographic Information System): to make marine biogeographic data, from all over the world, freely available over the Web. http://www.iobis.org/

Atlas of Living Australia: (planned) the Atlas will support research and decision making in national resource management and conservation planning, biosecurity, biodiscovery, and health and education. NCRIS, CSIRO

Don Hobern - Director, *Atlas of Living Australia* (ALA), CSIRO, Canberra The ALA is not yet available online, but contributions to this initiative are coming from CSIRO, state museums and herbaria. Funding, via NCRIS, has targeted infrastructural development, rather than databases or content. When established, the ALA aims to be a 'Yellow Pages' for species, providing information and directing to further resources.

ARKive: images of life on earth (still, video) http://www.arkive.org/

Convention on Biological Diversity (CBD): an international treaty to sustain the rich diversity of life on earth. 1991. NZ is a signatory. www.cbd.int/

Biodiversity data are being managed; databases are being built – but what about university research data? What about interoperability, standards? What about access – OA?

CRIs (9):

AgResearch; Crop & Food Research; ESR (Institute of Environmental Science and Research Ltd); HortResearch; Industrial Research Limited; Institute of Geological & Nuclear Sciences Ltd; Landcare Research; NIWA (National Institute of Water & Atmospheric Research); Scion (Forest Research).

Databases: NZ Fungi Db, Ant Distribution Db, National Vegetation Survey Db (Landcare), Freshwater Biodata Information System (NIWA)

Government departments:

DOC, LINZ: Land Information NZ

Databases: Land Environments of New Zealand. Hector's dolphin incident db, Bioweb & herpetofauna db's

Museums and NZ organisations:

Databases: Te Papa online collections, NZ Frog

WhoWhatWhy**Where**HowWhen

New Zealand universities?

Eprint:*BUGZ* – Harvester:*KRIS* – Data:*NZSSN*

BUGZ: Bibligraphy of New Zealand Terrestrial invertebrates: CU, Landcare Research, TFBIS, DOC

Full-text electronic archive of the *Bibliography of New Zealand Terrestrial Invertebrates 1775 - 1993*

BUGZ is the first New Zealand biodiversity database to allow dynamic matching of its entire full-text database against the taxonomic namebank of uBio – the universal Biological indexer and organiser. Namebank is a reconciled list of over 8,000,000 taxonomic names (including homonyms, synonyms and common names) and creates a virtual link to an ever-increasing number of international biodiversity databases (e.g. GBIF, NCBI, ITIS, Species 2000) that may contain additional biodiversity information useful to the user.

http://entdocs.landcareresearch.co.nz/WebForms/SearchForm.aspx

KRIS: Kiwi Research Information Service

Gateway to the open-access research documents produced at universities, polytechnics, and other research institutions throughout New Zealand.
http://nzresearch.org.nz/

NZSSN: New Zealand Social Sciences Network

Set up in August 2007 (TEC funded). Started with 15 survey data sets, from the area of politics (all NZ uni's are participants EXCEPT Otago)

http://www.nzssds.org.nz/links

WhoWhatWhyWhere**How**When

Otago Biodiversity Data Management Project:

- Aims to establish a **framework** for managing and sharing Otago biodiversity research data
- Aims to **connect** with national and international biodiversity strategies and information sharing initiatives
- Is driven by **interests** of University of Otago researchers
- Supports the **University of Otago** 'Emerging Research Theme': *Ecology, Conservation and Biodiversity in NZ*

University of Otago **Emerging Research Theme:** *Ecology, Conservation and Biodiversity in New Zealand*

http://www.otago.ac.nz/research/themes/theme_ecology.html

Takahe by Ian Jamieson, Zoology Dept, Uni of Otago

WhoWhatWhyWhere**How**When

Project Priorities:

1. **Communication** ongoing within/beyond the University
   share Project findings
2. **Survey** questionnaire (quantitative/qualitative)
   interviews (qualitative)
3. **Capabilities** technical, legal/cultural, organisational
   'Pilot within a Pilot'*

*Frog Image Database: *"New Zealand's Archey's frog is the World's most evolutionarily distinct and globally endangered amphibian (out of more than 6,240 species)* " Bishop 2008

Pilot

Test case for some of the general and subject related issues e.g.

General: IP/copyright, ownership

Subject: geospatial & taxonomic metadata requirements, standards (new to the Library)

Archey's frog Image by Phil Bishop, Zoology, Uni of Otago

WhoWhatWhyWhere**How**When

Survey:1

Otago Biodiversity Data Questionnaire is

- **to raise awareness** / inform about data mgmt issues
- **to quantify interest** / gather preliminary data
- **based on** three recent questionnaires (UK, USA, Aust)
- **presented** both online/print
- **contains** both multi-choice and free text options

Project StORe survey: cross-discipline report, Graham Pryor, University of Edinburgh, Aug 2006 www.era.lib.ed.ac.uk/handle/1842/1419

Data Needs Assessment Survey, Melissa Cragin, University of Illinois, 2008

University of Queensland (QUT, Melbourne Uni) survey, available via the APSR website, 2007

http://www.apsr.edu.au/orca/data_management_at_uq_blank_survey.pdf

Buttercup image by Janice Lord, Botany Dept, Uni of Otago

1. Please identify your own role
2. Please indicate your area of research:
3. How would formal management of biodiversity research data be useful to you?
4. What kinds of **non-digital data** do you generate or collect for your research (or have you generated or collected in the past)?
5. What kinds of **digital data** do you generate or collect for your research?
6. In what formats are these digital information sources held?
7. Are the research data you generate sometimes a combination of different **data formats**?
8. How **large** (in total) are your digital research data? Please estimate:
9. How long do you think your research data will have value?
10. What **data storage** and back up system(s) do you currently have in place?
11. Who currently manages your data?
12. Do you currently have a formal Research Data Management Plan in place and, if not, please indicate the reasons.
13. To which **repositories** – and how often - do you submit your research data?
14. Can you please indicate what types of metadata you consider important to assign to your data.
15. At what stage are **metadata** assigned to your research data?
16. Who assigns metadata to your research data?
17. Why might you wish to **access** the research data generated by other researchers?
18. How would you normally access the research data of other researchers?
19. How do you make your research data available to others?
20. What formal **restrictions** do you apply to your research data to control access?
21. What factors would **encourage you to share** your research data?
22. What factors would **discourage you from sharing** your research data?
23. If support were available, would you be interested in training or **advice** on any of the following?
24. Would you like to receive ongoing email updates about the current Otago Biodiversity Project and the findings of this research data questionnaire?
25. Please feel free to add any other comments regarding data management, long term data storage and access, digitisation, training, etc

25 multi-choice questions in University of Otago Biodiversity Data Questionnaire

WhoWhatWhyWhere**How**When

Data questionnaire - key findings:

- There is considerable **interest in formal research data management** among Otago researchers (70% of respondents wish to stay in touch with this Project and further developments)
- Data management is **not widely understood** by all researchers (indicated by the high number of 'don't know' responses and also through free text comments)
- The majority of researchers **manage their own data** and apply their own metadata

Buttercup image by Janice Lord, Botany Dept, Uni of Otago

WhoWhatWhyWhere**How**When

key findings (cont.)

- The majority of Life Science researchers have relatively **small amounts of data** (<1TB) and this data typically has a **long life** (>10 years)
- Most researchers indicate interest in **sharing data** with others but typically do so **informally** only (email, personal contact and direct F2F data sharing remain popular)
- Practical considerations (time/support) are **barriers to sharing data** (along with IP/ownership)

Buttercup image by Janice Lord, Botany Dept, Uni of Otago

WhoWhatWhyWhere**How**When

Otago Biodiversity Data Questionnaire

www.library.otago.ac.nz/services/projects.html

Buttercup image by Janice Lord, Botany Dept, Uni of Otago

WhoWhatWhyWhere**How**When

Survey: 2
Otago Biodiversity Data Survey will

- be **based on interviews/communications** with >70 researchers/advisors (F2F, phone, email)
- add a **qualitative** dimension ('real voices')
- **further enhance understanding** of researcher interests & concerns, regarding data management
- be available **Nov 2008** on Project website

Project website: http://www.library.otago.ac.nz/services/projects.html

Buttercup image by Janice Lord, Botany Dept, Uni of Otago

WhoWhatWhyWhere**How**When

Data Survey – emerging themes

1. **Otago Biodiversity Project** (support, recommendations)
2. **Metadata** (descriptive)
3. **Sharing data** (accessing others/sharing own)
4. **Role of the Library in data management**
5. **Research process & the University of Otago**
6. **Data management** (understanding, issues)
7. **Collaboration & funding** (models, issues, potentials)

Buttercup image by Janice Lord, Botany Dept, Uni of Otago

WhoWhatWhyWhere**How**When

Capabilities
'Pilot within a pilot'

- Working with national frog expert, Dr Phil Bishop's digital image collection
- Examining the technical, legal/cultural and organisational issues of managing a test case research data collection
- Reporting on findings, capabilities and requirements

50 digital images, various file formats (JPG, TIFF, Media Player…).
Experiences of existing 'Digital Collections'
Obligations to Treaty of Waitangi (in addition to legal considerations)

Archey's frog image by Phil Bishop, Zoology Dept, Uni of Otago

WhoWhatWhyWhereHow**When**

Where to from here…?

- Project (have we done enough?)
- Support (is there any?)
- Funding (who pays now – in future?)
- Organisation (where do we want to go?)
- Capabilities (can we do it?)
- Connections (have we made any?)
- …

BLRG: Biodiversity Library Reference Group

Education & Information Services Manager, Access & Development Services Manager, Digital Initiatives Librarian, Biodiversity Project Co-ordinator, Bibliographic Access & Metadata Librarian, Science Faculty Librarian

BDMAG: Biodiversity Data Management Advisory Group

Lawyer (Environmental Science) – Chair

Project Co-ordinator (Librarian) – Convener

Zoology, Botany, Marine Science, Geography, Geology, Chemistry, Information Science, Maori Research Office, HEDC, ITS, Phd & student rep/Ecosanctuary Community Liaison Officer

Dialogue:

Divisional: Pro-Vice-Chancellor, Sciences

Research: Deputy-Vice-Chancellor, Research

Academic: Deputy-Vice Chancellor, Academic

Presentations/workshops/conference:

NZ MoRST/GBIF Workshop: Wellington, June 2008

Otago Ecology/Botany Seminar: April 2008

Otago Zoology Board of Studies: August 2007

Digital Curation Centre Conference: Washington DC, Dec 2007

Otago Biodiversity Data Project

This is NOT about:

- National strategy
- Big science (large data size)
- One size fits all

This IS about:

- **Institutional** – Otago University research community focus
- **Regional** - Otago researcher community/connections
- **Small science** or ‘Long tail of science’ (small, dispersed data)
- **National** collaboration ‘by stealth’
- **International** collaboration by standards

Thank you

For further information contact:
Gillian Elliot
Biodiversity Project Co-ordinator

University of Otago Library
gillian.elliot@otago.ac.nz

Image by Russell Poulter, Biochemistry Dept, Uni of Otago