



Le réseautage de bases de données comme outil facilitant la recherche : deux initiatives du Service Canadien des Forêts

*Isabelle Aubin, Ph.D.
Service Canadien des Forêts, RNCAN
Sault Ste. Marie, ON
isabelle.aubin@RNCAN.gc.ca*

Montréal – Décembre 2011



Natural Resources
Canada

Ressources naturelles
Canada

Canada

Outline

- 1. Networking ecological Databases, ... why?**
- 2. Challenges of data networking in ecology**
- 3. Two CFS initiatives : TOPIC & Forest Change**

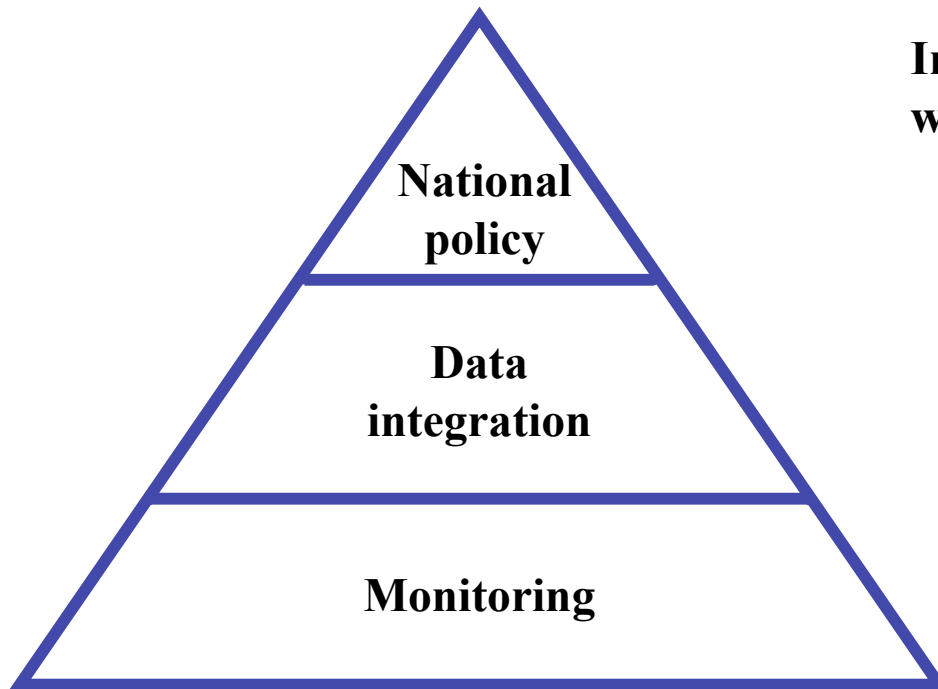


Networking ecological Databases, ... why?

- “Ecology increasingly relies on data synthesis and integration” (Kueffer et al. 2011).
- Today, forest ecosystems are undergoing changes at the global scale. This requires new approaches and tools which in turn require vast amounts of data from various sources, environments and disciplines (IPCC 2007).
- Database networking :
 - enhances the availability of data
 - facilitates the search for data
 - facilitates the sharing of data
 - promotes collaboration
 - May ensure data perennity



Networking ecological Databases, ... why?



In recent years, increased effort within governmental agencies:

At Canadian Forest Service :

Data integration used as a link between on the ground regional research and national policy

via data integration: up-scaling of monitoring outputs to relevant information for the drafting of national policies,.



Challenges of data networking in ecology

The complexity challenge

- Ecological data usually displays great variation in :
 - the type of data (Globe and Stevens 2008)
 - spatial and temporal scales (Davidson et al. 1995)
 - nomenclature (Levy 2000, Stein 2003, Globe and Stevens 2008)
 - protocols of measurement (Globe and Stevens 2008)
 - ...



Challenges of **data networking** in ecology

The financial challenge

- Funding for data management and standardization is scarce.



Challenges of data networking in ecology

The operational challenge

- Data is scattered and in great amounts.
- Databases are built in different formats and languages, and are supported by different technologies (Davidson et al. 1995, Globe and Stevens 2008).
- The structure and content of databases continually evolve with the development of experimental techniques and the discovery of new generalizations and laws (Davidson et al. 1995, Stein 2003).
- The generalization of terms and concepts for integration may cause a loss of precision in data (Davidson et al. 1995).
- ...



Challenges of data networking in ecology

The human challenge

- Intellectual property issues
- Cultural issues
- Motivation & cooperation issues



Challenges of data networking in ecology

The human challenge

Motivation & cooperation issues

- *“ As there is no reward system in place that recognizes the contributions made by sharing data, it is difficult to convince researchers to contribute simply because there will be a benefit to their discipline ” (Jones et al. 2006).*

*“ No-one will win a Nobel Prize for defining a workable format standard ”
(Globe and Stevens 2008).*



Challenges of data networking in ecology

The human challenge

- In order to ensure their success and perennity, initiatives of data integration must strike a **balance between the needs, rights and obligations of data users, managers and providers.**

Users : “ *All data should be public, free and accessible.* ”

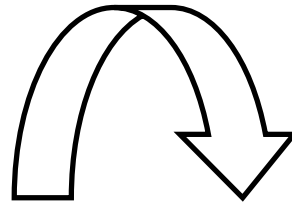
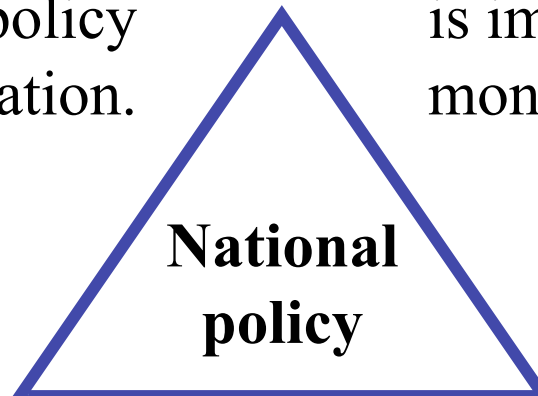
Providers : “ *I spent years collecting these data. I don't like the attitude of data users. They build their work on my data and publish in Science without any acknowledgment of my work.* ” **OR** “ *I don't have time to reformat my data to reach the database requirements!* ”

Data managers : “ *I have to struggle with all these challenges to help others to publish. While I do that, my own work stagnates!* ”



Monitoring can hardly influence national policy without data integration.

However, data integration is impossible without monitoring!

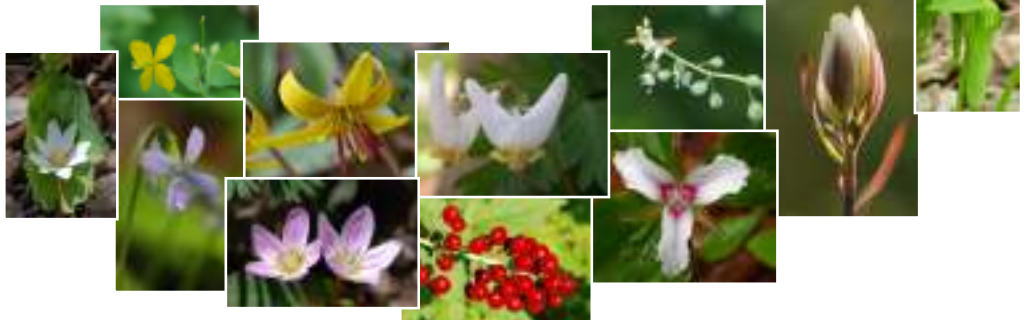


Toward a new research culture:

- Project designed with re-use of data in mind
- Reward system:
 - Citation of the original work
 - Co-authorship
 - Incentive for data sharing from funding agencies



Two initiatives in development of the Canadian Forest Service



Forest Change



Natural Resources
Canada

Ressources naturelles
Canada

Canada 

The TOPIC Network

Facilitating research collaboration
on key ecological issues through
the **functional traits approach**



Natural Resources
Canada

Ressources naturelles
Canada

Canada 

The functional traits approach has been underexploited in North America

Because of the :

- **Gaps in knowledge**
- **Fragmentation of available information**
- **Lack of an ecoinformatics structure**



Objectives of the TOPIC Network

- Integrate trait data into a regional database
- Facilitate the exchange of trait data
- Increase the collaboration between researchers
- Promote the use of the functional traits approach in Canada



From TOPIQ to TOPIC – a brief History

TOPIQ : Traits of Plants in Quebec (2004-2008)

- Created at the University of Montreal
- Reference : Aubin et al. 2007. *Ecoscience* 14 (2): 259-271.

TOPIC : Traits of Plants in Canada (2009-present)

- Managed by the Canadian Forest Service (CFS) – Great Lakes Forestry Centre
- More than 700 species and 50 traits are documented
- Data has been used in over 30 research projects

TOPIC network: Spring 2011

Networking or Integration?

Should we thrive towards a **network of regional databases or a single global database?**

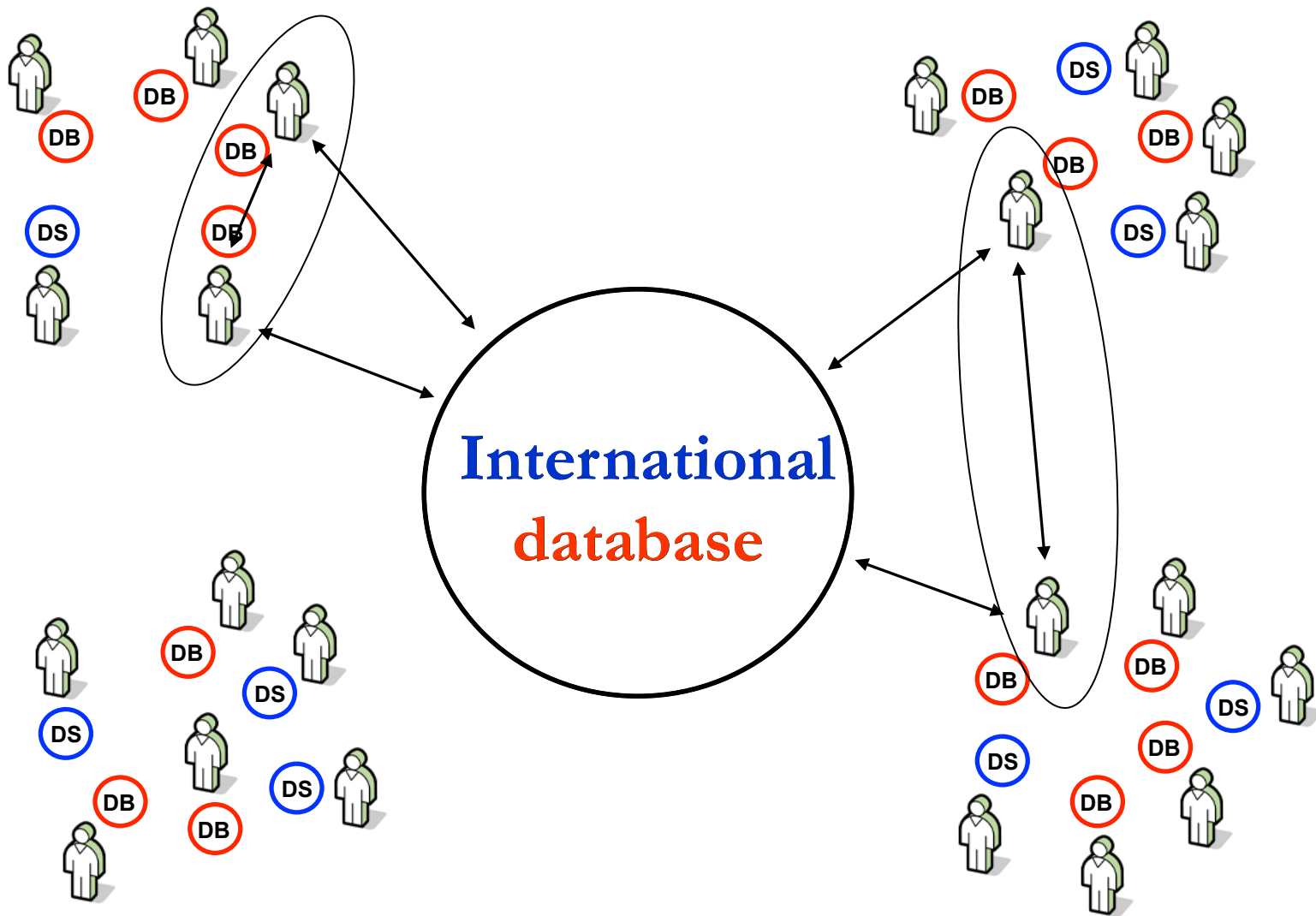
“ Although it is tempting to treat the integration of biological databases as a technological problem, in fact the main impediment to achieving this goal is not technological but sociological ” (Stein 2003).

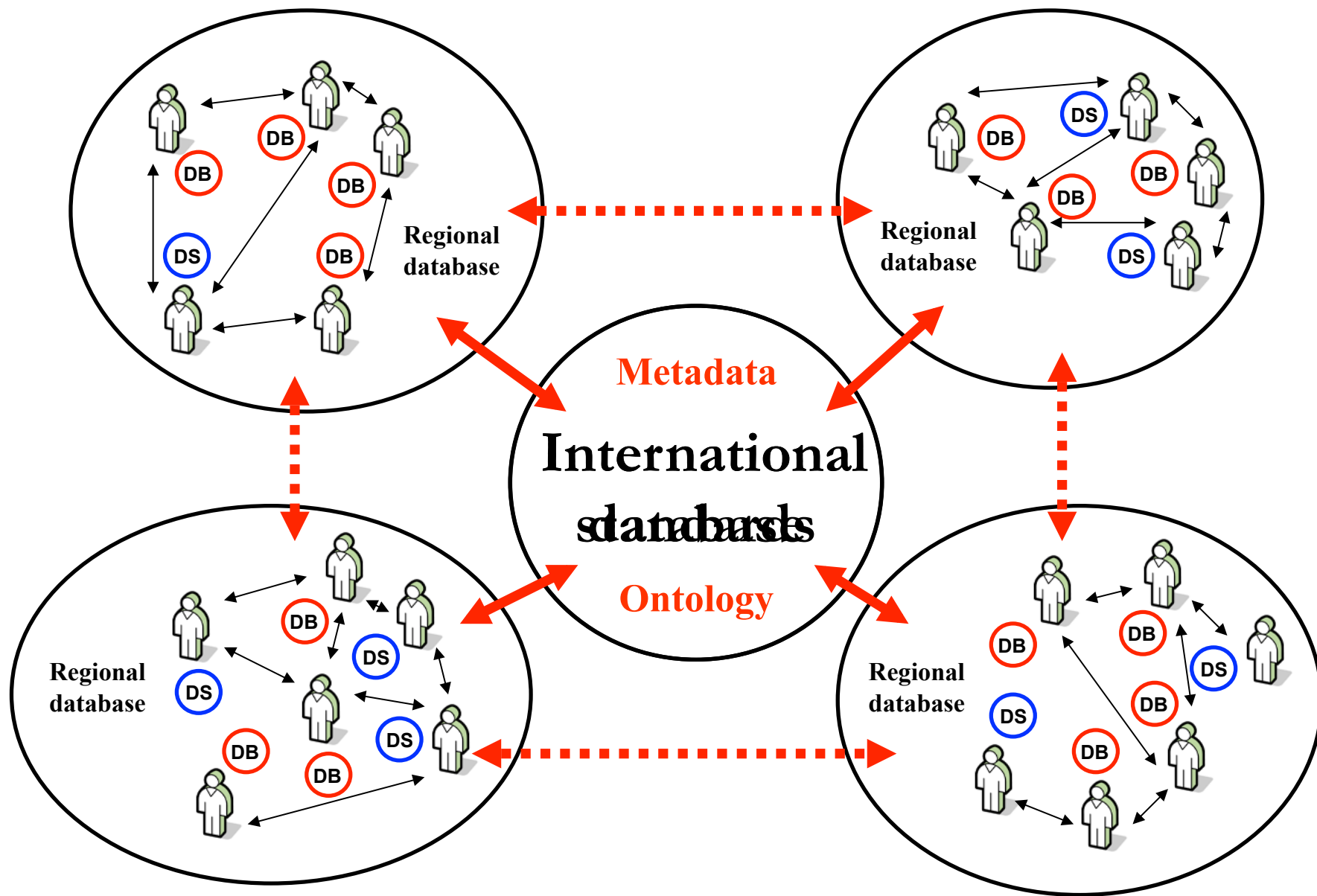


Natural Resources
Canada

Ressources naturelles
Canada

Canada





Members of the TOPIC Network

Orientation committee

- Aubin, Isabelle, CFS
- Bell, Wayne, OFRI
- Gachet, Sophie, IMEP (France)
- Hébert, François, MRNF
- Messier, Christian, UQAM
- D'Astous, Amélie
- Bachand, Marianne
- Boiffin, Juliette

Coordination committee

- Representatives from:
- **Universities**, UQAM
 - **Task force**
 - **Provincial**
 - **Industry**
 - **Editorial board**
 - **Provincial**
 - **Industry**

- Roulin, Monique, ULaval
- Ricard, Jean-Pierre, Usherbrooke
- Simpley, Bill, Usherbrooke
- Sirin, Industry, UQAR
- Thiffault, Nelson, MNRF
- Vellend, Mark, USherbrooke
- Dan McKenney, CFS
- Louis DeGrandpre, CFS
- Kevin Lawrence, CFS

Tasks:

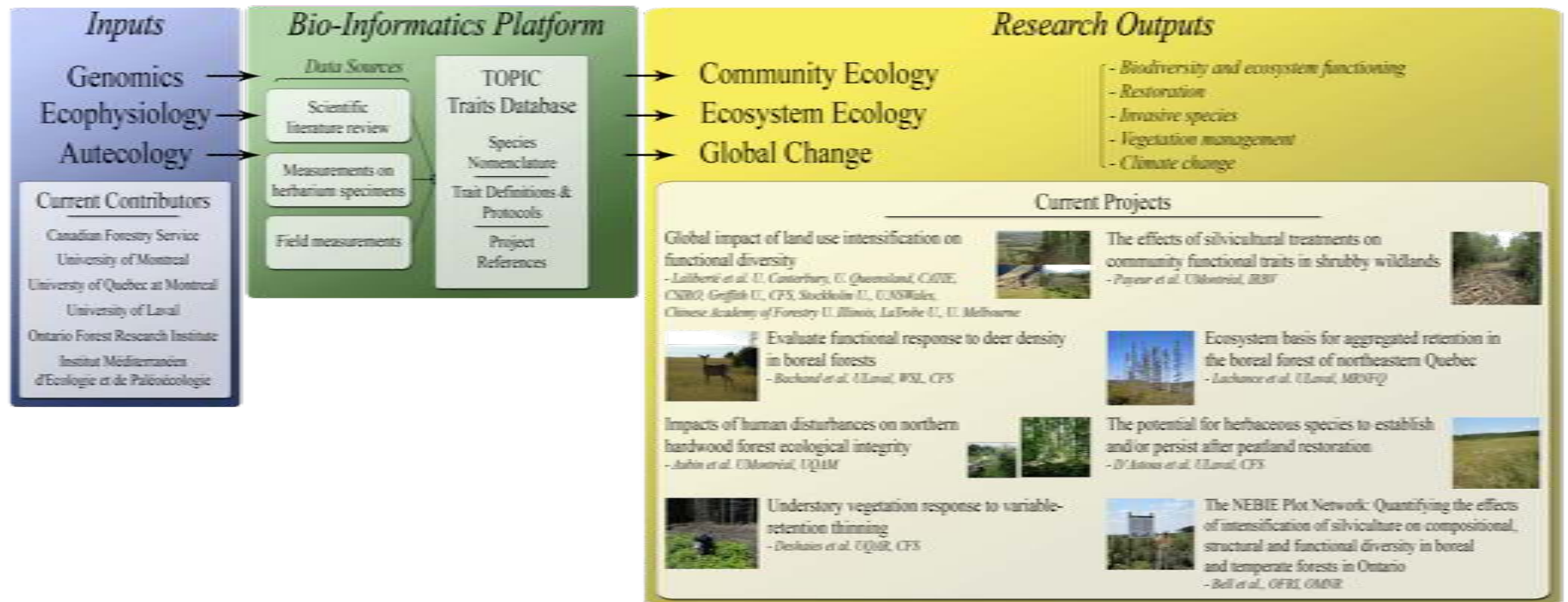
- Data acquisition
- Data validation
- Data distribution

Tasks:

- for the network
- Match for funding
- Promote the network



Structure and functioning of the TOPIC database



Structure and functioning of the TOPIC database

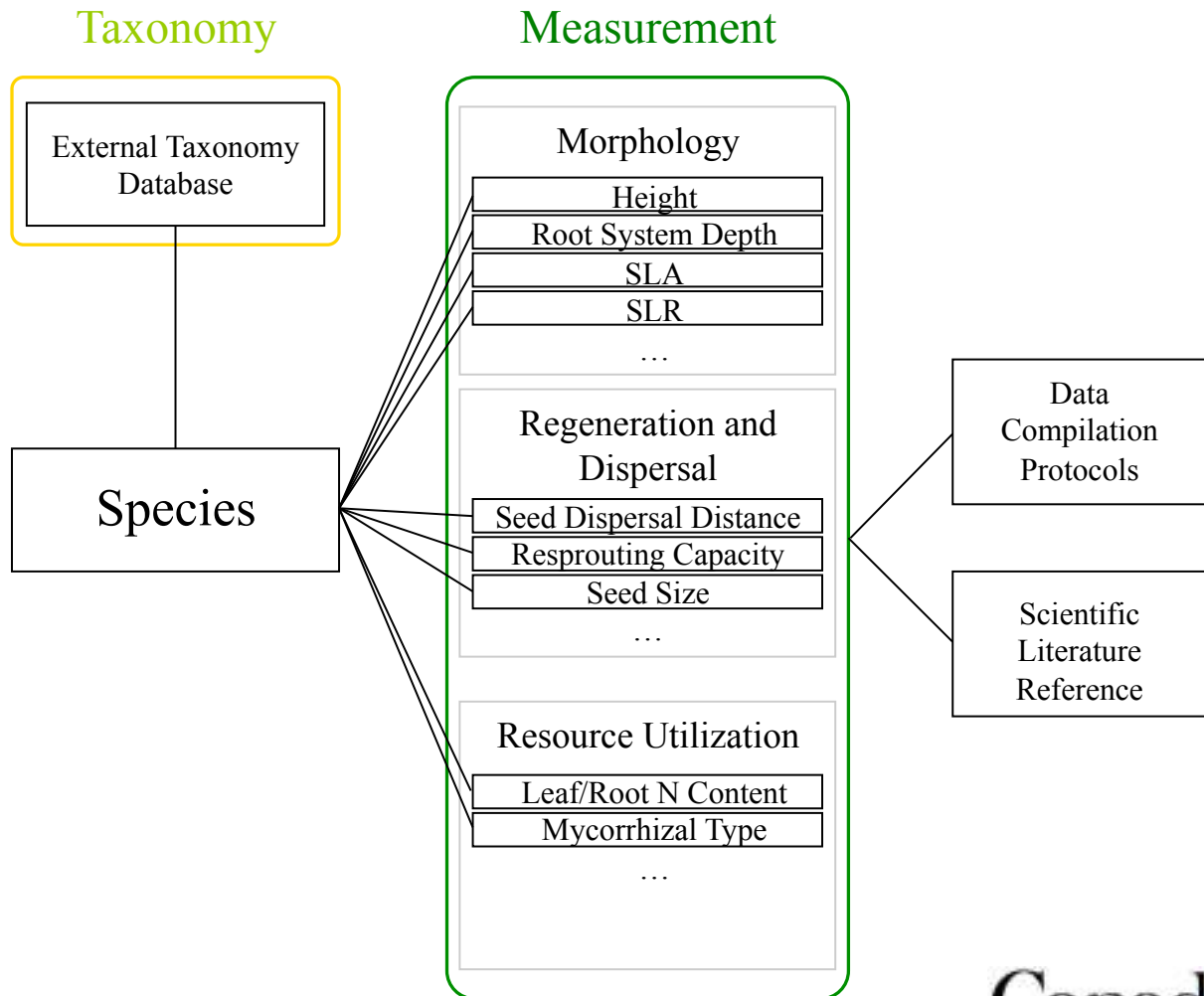
Two modules :

Module	Framework	Level of integration	Object	Advantage
Literature review	Relational Database	Fully integrated and standardized	Species	Ready-to-use data
Georeferenced field measurements	Metadata-driven Database	Low integration	Observation entity x time	Flexible data

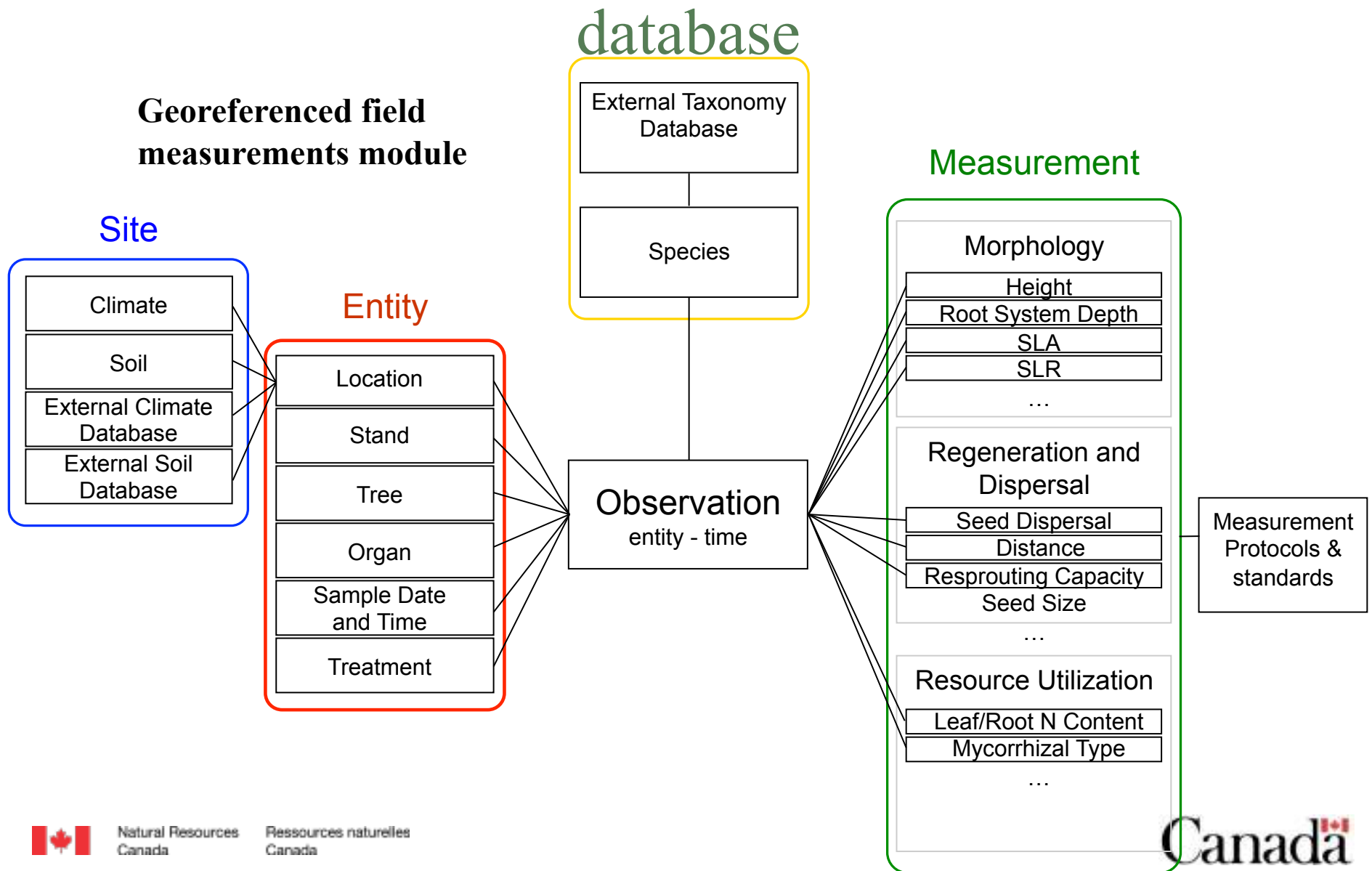


Structure and functioning of the TOPIC database

Literature review module



Structure and functioning of the TOPIC database



The TOPIC Website

Natural Resources Canada
www.nr.gc.ca

French | Home | Contact Us | Help | Search | canada.gc.ca

Traits of Plants in Canada (TOPIC)

TOPIC

TOPIC is a network of research scientists aiming to aggregate, promote and facilitate research in plant ecology and community ecology. The core of the network is a geoinformatics platform that contains data on the plant functional traits of the vascular flora of Canada. Data is provided by the members of the network and is available to anyone interested in contributing to the network.

The TOPIC geoinformatics platform currently contains data classified in four main categories, namely: morphology and structure of the adult plant, ecology for regeneration and dispersal, resource utilization and distribution. More than 700 vascular plant species are documented. The platform is organized around two main data modules: scientific literature review and geo-referenced field measurements. The data is categorized using international standards.

The TOPIC network was developed at the University of Montreal and is now hosted by the Canadian Forest Service of Natural Resources Canada. Members of the network include representatives from universities, provincial and national agencies and the industry.

Decomposed plant functional traits (November 28th 2011)

Morphological type
Type of data: qualitative
Format documented: KML

Maximum life span
Type of data: quantitative
Format documented: KML

Maximum height
Type of data: quantitative
Format documented: KML

Life cycle
Type of data: qualitative
Format documented: KML

Foliage type
Type of data: qualitative
Format documented: KML

Foliage persistence
Type of data: qualitative
Format documented: KML

Data request

Data will be released only to members of the TOPIC network. In order to become a member of the network, a request must circulate in the network by providing plant functional traits data which will be added to the TOPIC database. Please consult the 'Data contribution' tab of the present web site for details on data contribution. The amount of plant functional traits data to be provided by a requester will depend on the amount and type of data requested. Once a request has been made, the database manager will contact the requester to discuss the request.

Here is the detailed procedure for data request:

- The requester enters in the database manager its email and agrees by the internet to deliver data from the TOPIC database.
- The database manager sends a 'Data request' package to the requester via email.
- The requester provides a list under standardized format of the species and plant functional traits (traits) which he/she wishes to obtain from the:
- depending on the number of plant functional traits and species requested, the database manager (to clarify the amount of data to be provided by the requester) or data to TOPIC or member of the TOPIC network; the database manager then contacts the requester to discuss the request (on email and provide a form to the requester on the acquisition of data from the TOPIC database);
- the requester fills and signs the form and mails it to the database manager;
- within a month of receiving the filled and signed 'agreement', the database manager provides the data to the requester (on email under standardized format (Excel format));
- within two months of receiving the data from the TOPIC database, the requester provides the plant functional traits data to the requester in a format to be provided to the database manager (on email under standardized format);
- As soon as the requester provides plant functional traits data for the TOPIC database, the data provided from the TOPIC database becomes available to the requester;

The database is not entitled to use the data from the TOPIC database for publication or presentation and neither provide plant functional traits data to the database manager.

Data contribution

The development of the TOPIC network requires the contribution of the scientific community. Scientists are asked to contribute to the network by providing plant functional traits data which will be added to the TOPIC database. Contributors of at least a minimum amount of plant functional traits data and species will be invited to become members of the network.

In this context, the TOPIC network can be used as a working platform to further enhance collaboration between researchers, public sector agencies and use of the data contributed to its members. The network will allow data exchange and thus facilitate studies on critical issues related to plant ecology, community ecology and other related fields such as taxonomy, floristics, climate change and forest productivity. Furthermore, the network will allow inter-agency and inter-national plant functional traits comparisons and lead to interesting among the scientific community.

The TOPIC network is also of high importance in the formation of graduate students that have the opportunity to apply the geoinformatics approach to their research projects by using data from the TOPIC database. This also has the opportunity to build projects for their future careers.

The TOPIC database is organized around two main data modules which are related to the data and format of the data:

- SCIENTIFIC literature review
- georeferenced field measurements

The type and extent of the plant functional traits data initially contributed by the members define the right of access of the members to data from the TOPIC database.

Contributions of data from a scientific literature review

- Public access to records
- Open access for members that have contributed to the network

Contributions of data from georeferenced field measurements

- Public access to records



Present development of TOPIC

- **Automated procedure & make the portail fully operational**
 - e.g. data entry and diffusion via the web
- **Data acquisition and consolidation**
 - coordination of data acquisition campaigns
- **Development of bridges to external databases**
 - e.g. TRY, VASCAN-Canadensys
- **Funding and staffing strategy**
 - Development within CFS, help to user within member's organisation
 - Post-doc position
- ...



FOREST CHANGE

A portal collating information and tools on
**the risks and opportunities posed by climate
change to Canada's forests**



Natural Resources
Canada

Ressources naturelles
Canada

Canada 

Forest Change – Context and objectives

Context:

- Managing and mitigating the effects of climate change on Canada's forests require informed decision-making, which creates **an increasing demand for relevant knowledge, information, products and services.**

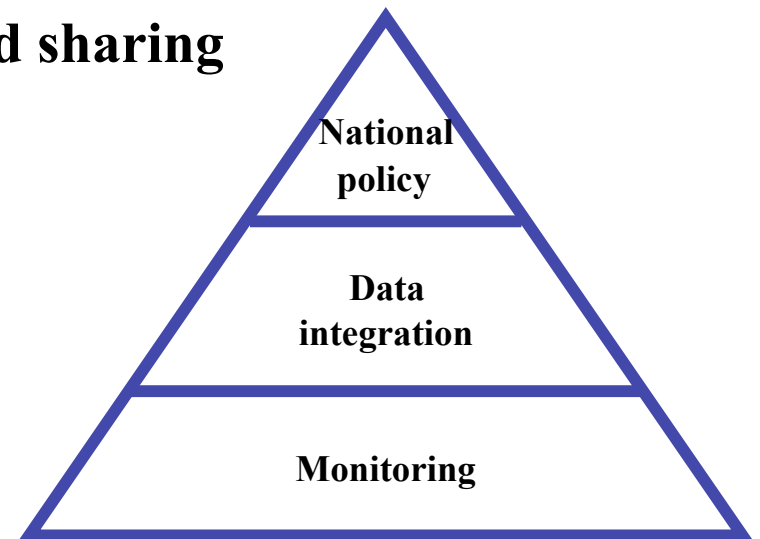
Objectives:

- **KNOWLEDGE TRANSFER** – Tailoring information to meet the needs of users and enabling the incorporation of climate change considerations into forest management
- **INTEGRATED ASSESSMENT** – Integrating data in order to enable the integrated assessment of multiple interacting impacts of climate change on Canada's forests
- **CORPORATE PLANNING FRAMEWORK** – Identifying potential synergies in research activities, knowledge gaps and research priorities, as well as areas for collaboration in order to maximize the efficiency and relevance of research



Present development of Forest Change

- **Design of the Forest Change Portal** to host a diversity of information from general information (for public, education, etc.) as well as more detailed scientific data/knowledge.
- **Technological infrastructure** to enable the integration of climate change impacts knowledge from multiple disciplines.
- **Policies and guidelines of data use and sharing**



References

- Davidson, S.B., Overton, C. and Buneman, P. 1995. Challenges in Integrating Biological Data Sources. *Journal of Computational Biology* (4): 557–72.
- Globe, C. and Stevens, R. 2008. State of the nation in data integration for bioinformatics. *Journal of Biomedical Informatics* 41: 687–693.
- IPCC 2007. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.), Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
- Jones, M.B., Schildhauer, M.P., Reichman, O.J. and Bowers S. 2006. The New Bioinformatics: Integrating Ecological Data from the Gene to the Biosphere. *Annu. Rev. Ecol. Evol. Syst.* 37: 519–44.
- Levy, A.Y. 2000. LOGIC-BASED TECHNIQUES IN DATA INTEGRATION. In: Minker, J. (ed.), *Logic-based artificial intelligence*, Kluwer Academic Publishers Norwell, USA.
- Madin, J.S., Bowers, S., Schildhauer, M.P. and Jones, M.B. 2007. Advancing ecological research with ontologies. *Trends in Ecology and Evolution* 23: 159–168.
- Stein, L.D. 2003. INTEGRATING BIOLOGICAL DATABASES. *NATURE REVIEWS* (4): 337–345.
- Aubin, I., Gachet, S., Messier, C. and A. Bouchard. 2007. How resilient are northern hardwood forests to human disturbance? *Ecoscience* 14 (2): 259-271.



Questions?

