Managing a Fine Balance between New Geoscience Data Acquisition and Exploitation of Existing Legacy Data

John Broome

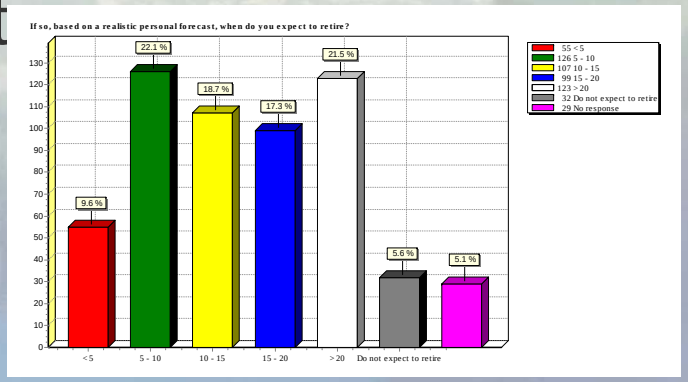
IAMG, 2005

The Legacy Data Challenge

To preserve existing investments in geoscience information and manage that information in a manner that meets current and future requirements a reasonable cost.

Environment:

- Geological Survey = Information Agency
- Surveys typically small but information rich
- Increasing cost of new data acquisition
- Need to preserve and make accessible
- Future : Real time dynamic online delivery.
- Key role of IM in new environment
- 30% Retirements in 10 years



Legacy Data Issues

- Many diverse types of data; physical, analog and digital maps, etc …
- Variable standardization & quality.
- Multiple interpretations & scientific conflict.
- Missing information/metadata.
- Understandable only to select experts.
- Demographics: retirements in next 5 years.
- IM is not valued in the geoscientific culture.
- Adequate funding not available.

The Balancing Act

To move forward on resolution of the legacy data problem recognizing that costs of this activity must be balanced with other business activities.

Geological Survey of Canada Approach

1. Divide task into 2 phases:
    a) Information Life Cycle Management (preservation)
    b) Prioritized upgrading
2. Fund through temporarily 50% fieldwork reduction.
3. Establish corporate IM strategy and policies.
4. Governance: establish corporate IM Branch.
5. Establish federated Information Architecture.
6. Standards
7. Tools

1) Data Life Cycle Management

- Not all legacy data are useful.
- Assessment of value and utility is a first priority
- Actions range from disposal, to static management, to upgrading.
- Different degrees of upgrading are appropriate depending on the anticipated value and usage.

2) Prioritized Upgrading

- Digitizing : Raster and Vector depending on requirements.
  - Raster for archival purposes & print on demand.
  - Vector to support current and future projects (GIS)
- Conversion to standard legend, data model, and science language.

3) Reduced New Data Acquisition

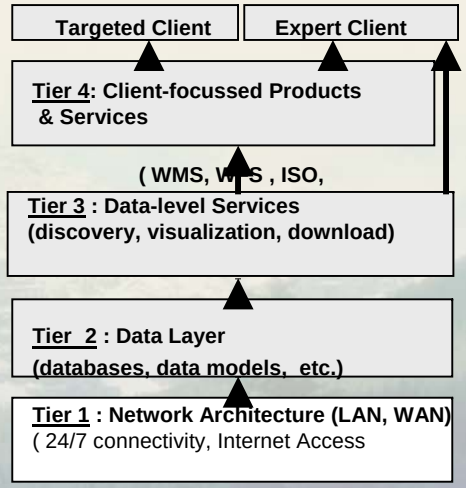
- 3-year diversion of funds to improved IM and addressing the legacy data problem.
- Reassigned field geoscientists to quality assessment and regional compilation.
- Compilation product is digital database not map.
- Major project to complete development of Geoscience Data Repository.

4) Corporate IM Strategy & Policy

- Essential to define the principles and guidelines.
- “Service” rather than “product” orientation
- Accessible, documented, policies, standards, and plans.
- Defined responsibilities/governance
- Corporate Information Infrastructure
- “Federated” approach NOT “Centralized”

5) Information Architecture

- A common corporate approach to IM developed from business requirements analysis.
- Previously, local centers of IM excellence with diverse approaches (“silos”)
- “Horizontally-integrated” approach based on reusable components based on commonalities.
- Analyze enterprise requirements; develop logic model, THEN implement.



6) Standards

- Essential to upgrading of legacy data.
- Conformity to documented standards facilitates use by future application.
- Adoption of national and international standards.
- Participation in international data model and science language (NADM, IGC,CGMW, etc.

7) Tools

- Provide easy to use software tools
- Facilitate population of “standard” databases.
- Insulate user from the complexity of the model
- Make it easier to work the “standard way”.

The Future:

- New data collected and managed in an ad-hoc fashion creates future legacy data problems.
- Information must be managed for the long term
- Legacy data evaluation is Step 1 for future mapping.
- Future : Remote Predictive Mapping approach supported by legacy data and samples.

Summary

- Resolving legacy data problems requires commitment at all levels in the organization.
- All legacy data are not worth preserving and upgrading, Assessment is “Step 1”
- Corporate strategies, policies and systems play key roles in addressing the problem.
- Ad-hoc project expenditures on data upgrading or acquisition are costs. Proper upgrading and IM is an investment in the future.