Standards and the International Polar Year

Siri Jodha Singh Khalsa
National Snow and Ice Data Center
IEEE GRS-S liaison to ISO TC211

A Paul Cooper
British Antarctic Survey
SCAR liaison to ISO TC211
Summary

• The IPY is a major international, interdisciplinary scientific event.
• The size and scope of the IPY provide unprecedented opportunities and challenges.
• Standards will play an important role in achieving the goals of the IPY.
What is IPY?

• An intense, internationally coordinated campaign of research in polar science
• Covers a wide range of scientific disciplines, including the social sciences
• Emphasis on public education and training of the next generation of engineers, scientists, and leaders.
Background to IPY

• First IPY in 1882/83, 2nd in 1932/33
• The 3rd IPY became the IGY in 1957/58
  – Fostered a new era in Antarctic collaboration
  – Resulted in The Antarctic Treaty
  – Left a network of World Data Centres
• Fifty years later, IPY 2007/08
  – Simultaneous with eGY, IHY and International Year of Planet Earth
World Data Centres

• 52 WDCs located in Europe, South Africa, Russia, Japan, India, China, South America, Australia, Canada, and the United States
• Discipline based
  – E.G. Antarctic Glaciological Data Center
• Provide access to data and publications
• Often hosted within larger archives or bodies
  – National Snow and Ice Data Center
  – Scott Polar Research Institute
• Active in international standards coordination
Size and Scope of IPY

- 60 nations
  - most have representation on TC211
- Estimated 50,000 participants
  - includes technicians, crew, etc., in addition to scientists
- Vast in scientific scope: genomic sequences to the geoelectric circuit
- Very wide range in knowledge of and involvement with standards
Goals of the IPY

• Have an interdisciplinary emphasis, with active inclusion of the social sciences
• Link researchers across different fields to address questions and issues lying beyond the scope of individual disciplines
• Strengthen international coordination of research and enhance international collaboration and cooperation
Goals of the IPY - continued

• Leave a legacy of observing sites, research facilities, data, and systems to support ongoing polar research and monitoring.

• Engage the awareness, interest and understanding of schoolchildren, the general public and decision-makers worldwide.

• Effective data management important to all of these goals.
Data Management and the IPY

• To achieve its aims, IPY must build a strong, cross disciplinary international data sharing community.
• Data products must be accessible
• Data products will need to be at many levels
  – “Raw data” products for domain experts
  – Consolidated products for researchers in other fields
  – Summary information for policy makers
IPY Data Management Workshop

- Held at the British Antarctic Survey in Cambridge, England 3-4 March 2006
- Hosted by NSIDC and the IPY Programme Office
- Recommendations from this workshop form the basis of this presentation
- http://nsidc.org/events/ipydis/
The IPY DIS

• The primary goal of the workshop was to develop an implementation plan for the IPY Data and Information Service
• Includes a Data Coordination Office
  – assistance with standards compliance
  – development of union catalog
• Education - “Data Stories”
• Innovative data discovery and access
Models for Data Discovery

- Central catalogue service
  - Single point of failure
  - Would not serve community-specific needs
- Multiple catalogs interconnected through XML-based metadata harvesting protocols
  - Information shared through “union” catalogue
  - Individual catalogs harvest metadata they require
  - Can develop discovery and access interfaces most appropriate to each community
  - Capture user interaction patterns to enhance future use
Enabling Interoperability

• The IPYDIS must encourage interoperability at all levels.
• At its most basic, “interoperability” is the ability for different software and hardware to share and use data — definition belies the complexity of the problem, especially when sharing data across cultures and scientific disciplines.
The Standards

• Suitable standards exist for many of the components needed for an IPY system
  – Existing standards may need extending
  – Some new standards may be required

• IPY is in a position to influence standards development
  – Via national standards bodies
    • e.g. BSI, ANSI, DIN
  – Via Liaison bodies
    • e.g. SCAR, WMO, IEEE, UNEP
Why Geospatial Standards?

• Most environmental data are geospatial.
• Many other data have a geographic component.
• Geography is a common linking theme between disciplines.
• Geospatial Standards (ISO, OGC) are mature and have many implementations.
Recommendations

• Require projects to provide ISO 19115 compatible metadata using XML-based transport formats where possible.

• Assist in developing and encouraging community-specific profiles of ISO 19115.

• In cases where ISO 19115 is inappropriate or inadequate (e.g. artifacts, multimedia) the IPYDIS should encourage the use of international library and archival standards.
IPY DIS Services

- Data Discovery
- Mapping services
- Data delivery services
- Service chaining
  - Image processing
  - Reprojection, reformatting
What We Have

- **WMS, WFS and WCS** well specified by OGC
- **WMS** delivers maps via simple URL requests
  - Suitable for data exploration
  - OGC SLD and ISO 19117 (Portrayal)
- **WFS, WCS** deliver data via XML requests
  - Suitable for detailed analysis
- **Use needs to be refined in light of other standards**
  - ISO 19110 and ISO 19126 (Feature Catalogue and Data Dictionary)
  - ISO 19113 and ISO 19114 (Data quality principles and evaluation)
Need Steps

• Data Discovery services
  – Tailored to specific disciplines
  – Requires adequate metadata (and ISO 19139)
• Semantic interoperability
• Need services built on standards such as
  – ISO 19110 (Feature Catalogues)
  – ISO 19126 (Data Dictionary)
  – ISO 19115 and ISO 19139 (Metadata)
  – ISO 19119 (Services)
  – NWIP on Cross-Domain Vocabularies
Semantics

• Metadata and data standards, while essential, only partially address the issues of interoperability in the cross-cultural, interdisciplinary IPY domain.
• Technical terminology is often discipline specific
• Semantic interoperability desirable when a large range of disciplines are attempting to share data and information.
Statement of Problem

• Each science domain or community develops its own terminology to describe concepts, resources (objects, data) and relationships
• Data discovery and data sharing depend critically on being able to attach unambiguous meaning to the terms used to describe domain knowledge
• Different domains have used different means of codifying domain knowledge
Knowledge Organization Systems

- **Controlled vocabularies**
  - Glossaries, Dictionaries, Thesauri
  - Gazetteers
- **Classification Schemes**
  - Taxonomies
  - Feature catalogs
- **Ontologies**
Ontologies

• Expressed in a formal conceptual language (UML, ERD, RDF, OWL,...)
• Uses symbols, text, rules, axioms to express:
  – classes (conceptualizations of objects)
  – instances of classes
  – properties of classes
  – relationships between classes
Process

• There is no one correct way to model a domain
  – there are always viable alternatives.
• Ontology development is necessarily an iterative process.
Approaches - Top Down

- Rarely based on consensus of community
- Our intent is to draw upon existing domain knowledge representations in each IPY discipline
  - Investigate tools for bringing these knowledge bases into a common system
Approaches - Bottom Up

- System for assigning subject metadata
  - “tagging” interface, “folksonomy”
  - High level terms from defined domain specification
  - Leave discovery to web services such as Google
  - Analyze relationships of tags and users to extract a “light” ontology
  - Allow community review and editing
Recommendations

• The IPYDIS should encourage the development of formal semantic approaches to interoperability (e.g., ontologies) in areas where needed.

• This will require a phased, community-based approach that could include formal use cases and informal approaches to soliciting community input (e.g. wikis, social bookmarking, etc.)
Final Words

• Many nations around the world will be cooperating
• Many organizations with different infrastructures
• IPY systems must be open
• Specify interoperability at the interfaces
• Must not depend on proprietary software or other intellectual property.
Summary

- The IPY is a major international, interdisciplinary scientific event.
- The size and scope of the IPY provide unprecedented opportunities and challenges.
- Standards will play an important role in achieving the goals of the IPY.
Resources

- http://nsidc.org/events/ipydis/
- http://www.ipy.org/
- http://www.ipy.gov/
- http://www.ngdc.noaa.gov/wdc/
- http://www.antarctica.ac.uk/