Data content standards in Africa

Antony K Cooper (CSIR)
Sives Govender (EIS-Africa)
Kate Lance (USGS/EROS)
Larry L Tieszen (USGS/EROS)
Data content standards in Africa

- Background to the project
- The nature of data content standards
- Data dictionaries and feature catalogues
- Standards assessed
- Data content components
- Recommendations
- Conclusions
Background to the project

• US Geological Survey EROS Data Center (USGS/EROS) initiated a project with EIS-Africa:
  – “Guidelines for data content standards for Africa”

• Funded by:
  – US AID
  – CSIR
Nature of data content

- Documentation specifying the information in a data set:
  - Metadata
  - Reference models
  - Data dictionaries, feature catalogues and classification
  - Feature instances
  - Data organization

- Formal description of a model, eg: using UML
  - Hopefully embedded in the data content standards
Nature of data content

- Metadata
  - Data about data, including data quality
- Reference models
  - Scope of standardization activity and the context
- Data dictionaries, feature catalogues and classification
  - Feature types (classes), attribute types, attribute domains, feature relationships
- Feature instances
  - Unique, definitive versions of features
- Data organization
  - eg: XML, GML
Why data content standards?

• Data content standards tend to be more accessible
  – Easier to understand
  – Used directly by many end users

• Immediately applicable to Africa

• More “susceptible” to culture and language
  – Hence, more important to have local standards
Data content standards

- Documentation specifying the information in a data set:
  - Metadata (including data quality)  
  - Reference models
  - Data dictionaries, feature catalogues and classification
    - Feature types, attribute types, attribute domain, feature relationships
  - Feature instances (unique, definitive versions of features)
  - Data organization (eg: XML, GML)

- Formal description of a model, for example using UML
  - Hopefully embedded in the data content standards

Many projects
Data dictionary or feature catalogue

- Both contain the types of geographical features
  - Classes or feature types
- Both contain feature attributes
  - Types and domains
- Conceptual relationships between feature types
  - Eg: an instance of the feature type ‘bridge’ can carry an instance of the feature type ‘road’ over an instance of the feature type ‘river’
Do you have a data dictionary or feature catalogue?
Data dictionary vs feature catalogue

• Need proper definitions to differentiate between feature types, and not merely use the label (name)
• Data dictionary
  – An unstructured collection of feature types
• Feature catalogue
  – A structured collection of feature types
    • Eg: as a hierarchical classification
    • Hence, easier to use
• Typically, a feature catalogue is constructed from a data dictionary
  – Perhaps as a profile (subset) of the data dictionary
Data content standards [4]

- 160+ standards were assessed
  - ISO/TC 211, OGC, FGDC, South Africa, Zimbabwe, etc

<table>
<thead>
<tr>
<th>Name</th>
<th>Source</th>
<th>ID</th>
<th>Date Publish</th>
<th>Status</th>
<th>Data Content?</th>
<th>Metadata?</th>
<th>Relevant?</th>
<th>Publicly Accessible?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feature Instance Identification Standard</td>
<td>Stan SA</td>
<td>SABS 1876</td>
<td>draft</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Cost</td>
<td></td>
</tr>
<tr>
<td>Land Cover Classification Scheme for Remote Sensing Applications in South Africa</td>
<td>Stan SA</td>
<td>SABS 1877</td>
<td>2003</td>
<td>published</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Cost</td>
</tr>
<tr>
<td>South African Spatial Metadata Standard</td>
<td>Stan SA</td>
<td>SABS 1878</td>
<td>draft</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Cost</td>
<td></td>
</tr>
<tr>
<td>South African Geospatial Data Dictionary (SAGDaD) and Its Application</td>
<td>Stan SA</td>
<td>SANS 1880</td>
<td>draft</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Cost</td>
<td></td>
</tr>
</tbody>
</table>
Data content standard components

- Data content components from selected standards were compared

<table>
<thead>
<tr>
<th>SANS 1880</th>
<th>FGDC Cadastral</th>
<th>FGDC Hydrographic</th>
<th>Zimbabwe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isoline</td>
<td></td>
<td>Depth contour</td>
<td>Height contour</td>
</tr>
<tr>
<td>Administrative Area</td>
<td></td>
<td>Administration Area</td>
<td>Adm Area</td>
</tr>
<tr>
<td>River</td>
<td></td>
<td>River</td>
<td>River</td>
</tr>
<tr>
<td>Cadastral Property</td>
<td>Parcel</td>
<td></td>
<td>Property Parcel</td>
</tr>
</tbody>
</table>
Project outline

• Project plan
  – Existing data content standards and current practices
  – User requirements
  – Evaluation of existing standards
  – Data content components for each theme
  – Consult with specialists
  – Draft document on guidelines and best practices
  – Circulate guidelines for feedback
  – Analyse comments and revise guidelines
  – List server for the project
    • http://www.gsdci.org/

• Unfortunately, we had a very limited response
Recommendations on which standards to use [1]

- Any feature catalogues used should conform to ISO 19110:2005, *Geographic information – Methodology for feature cataloguing*
- ISO 19110 has some limitations
  - Mechanisms for cultural and linguistic adaptability (CLA)
  - Particularly useful in multi-lingual environments
    - Applies to most, if not all, countries in Africa
  - Aliases allow feature types with labels (names) in multiple languages
Recommendations on which standards to use [2]

• Currently, ISO 19110 limitations unlikely to affect most users

• Few feature catalogues will use ISO 19110 optional constructs
  – Eg: feature operations and feature associations

• ISO/TC 211 will continue to maintain and enhance ISO 19110
Recommendations on which standards to use [3]

- Not possible to recommend one, definitive data dictionary or feature catalogue
  - To be used for all digital geographical information
  - By all users across Africa
  - Under all circumstances
- Use a widely used feature catalogue that meets most of one’s needs
  - Adding more detailed feature types for in-house use, if necessary
  - Does one want a specialist or a general purpose one?
Conclusions [1]

- We have completed the project
- EIS-Africa has published a CD-ROM of available data content standards and other resources
- EIS-Africa will make these available on their Web site: http://www.eis-africa.org/
- Be realistic in our ambitions …
Conclusions [2]

• What do you think should be the next step?
• Regular reporting at recognised events
  – Africa GIS
    • Including a workshop on standards
  – AARSE
  – CODI
• UN ECA should ID and fund pilots
  – Produce a draft standard and pilot it!
Thank you!

Contact details:

Antony Cooper
CSIR, PO Box 395, Pretoria, South Africa
Email: acooper@csir.co.za
Telephone: +27 12 841 4121
Facsimile: +27 12 841 3037