International Journal of Spatial Data Infrastructures Research, 2012, Vol.7, 508-535

Okavango Collections: Sharing Environmental Information Resources of the Cubango-Okavango River Basin*

Richard T. Cooper[1], Monica Morrison[2]

[1]Southeast Asia START Regional Centre, Chulalongkorn University, Bangkok, Thailand, rcooper@start.or.th

[2]OKACOM Secretariat / Secretariado da OKACOM, Permanent Okavango River Basin Water Commission / Comissão Permanente das Águas da Bacia Hídrográfica do Rio Okavango, Maun, Botswana, monica@okacom.org

Abstract

This paper provides an account of the development of the web-based *Okavango Collections (OC)* metadata catalogue. In light of climate change, growing population, and developmental pressures on the transboundary Cubango-Okavango river basin, there is an urgent need to ensure timely access to reliable environmental data and information for environmental decision-making. Commissioned by the Permanent Okavango River Basin Water Commission (OKACOM), *OC* provides a searchable directory of collections of data, data sets, image archives, books, reports and other information available both in the region and in other parts of the world. It is based on GeoNetwork Opensource, a standards-based and open source web catalogue, implementing international standards for metadata content and system interoperability. Its main features include a geospatial catalogue application, providing metadata editing and advanced search and discovery capabilities, and an integrated map viewer. Key development challenges included the creation of a regional geographic names thesaurus as a controlled vocabulary for place keywords, an email notification system to facilitate ongoing maintenance of metadata records, improved contacts management, and enhanced search and discovery functionality. In addition to providing an important environmental information service for the Cubango-

DOI: 10.2902/1725-0463.2012.07.art23

Okavango region, and a key tool to facilitate negotiation of river basin data sharing among the riparian states, *OC* potentially offers a model system for implementation elsewhere.

Keywords: transboundary river basin, Cubango-Okavango river basin, information systems, metadata, spatial data infrastructure.

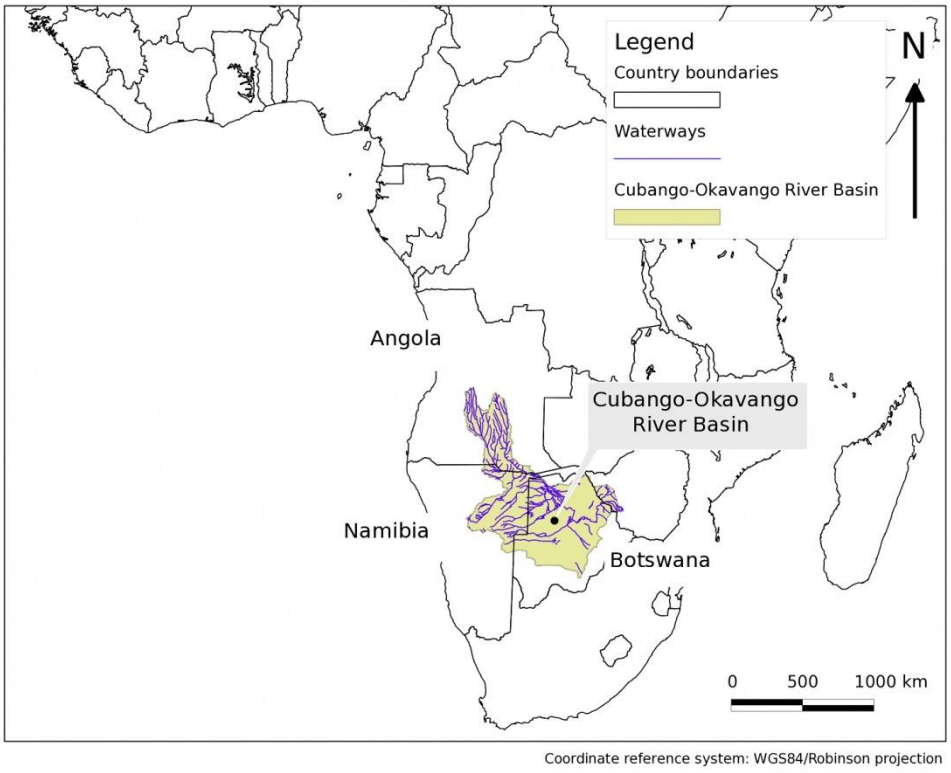
1. INTRODUCTION

The Cubango-Okavango river basin (Figure 1), a major transboundary system of Angola, Botswana, and Namibia, is a region of high national, regional and global environmental significance (OKACOM, 2011a). The river is one of the least developed in Africa, one of the longest in southern Africa (extending some 1000km), and the inland Okavango Delta is one of the largest Ramsar sites in the world (Kgathi et al, 2006, OKACOM, 2011a); the Okavango Delta System was designated as a *Wetland of International Importance* in 1996 (Ramsar, 2012). While the basin is described as currently being in 'near-pristine status' (OKACOM, 2011a), it is threatened by population growth, land use change, poverty and climate change (OKACOM, 2011a). In both the medium and long-term, climate change poses a serious risk to people and their environment in southern Africa (Biggs et al, 2004), and according to Andersson et al (2006), the hydrological impact of climate change on the Okavango River basin could be greater than development; the latter authors however acknowledge the uncertainty of their approach based on scenario modelling. The governments of the three member states, through the Permanent Okavango River Basin Water Commission (OKACOM) are looking for ways to improve livelihoods of people living in the basin and to increase joint economic benefits to the countries without degrading the river system.

Effective joint planning and management of the Cubango-Okavango river basin requires timely access to up to date and reliable information. Collections of pertinent data and information are held at various government agencies and organisations throughout the region, with metadata records of these scattered collections either lacking or inadequate. Since there was not a single access point enabling search and discovery of information resources about the river basin, necessary to facilitate negotiation of data sharing arrangements among the riparian states, an online geospatial metadata catalogue and integrated web map viewer - *Okavango Collections (OC)* - was commissioned by OKACOM. This paper provides some background to the development of the portal in relation to ongoing efforts to facilitate information access for environmental decision-making,

and describes its key features and functions including challenges to its development and sustainability.

Figure 1: Cubango-Okavango River Basin



Source: The waterways and catchment boundary shown in Figure 1 are derived from GIS files available at http://www.fao.org/geonetwork/srv/en/metadata.show?id=12 and http://iwlearn.net/iw-projects/842/maps_graphics/okavango-river-basin

2. FACILITATING ACCESS TO ENVIRONMENTAL INFORMATION

This section places the development of *OC* within the context of international and regional approaches to data and information access, including development of spatial data infrastructure, implementation of international standards for system interoperability and the benefits of open source software.

2.1. International and Regional Context

The need for improving access to data and information for environmental decision-making has been articulated on a number of occasions at international

and regional meetings in Africa over the last decade (e.g. WSSD, 2002; UNECA, 2007; AWCS, 2011). It was highlighted in the 'Plan of Implementation of the World Summit on Sustainable Development' (WSSD, 2002), which identified the need to 'develop information systems that make the sharing of valuable data possible, including the active exchange of Earth observation data'. In a broader context, linking Information and Communication Technologies (ICT) to sustainable development, it was reported at the Thirteenth Meeting of the Intergovernmental Committee of Experts for Southern Africa (ICE), held in Lusaka (UNECA, 2007) that 'Information and Communication Technologies (ICT) were central to the creation of a global knowledge based economy and could play an important role in accelerating growth, promoting sustainable development and eradicating poverty in developing countries'.

More recently, in the 'White Paper on GEO Capacity Building and Water Resources in Africa' (AWCS, 2011), water information systems were described as 'fundamental for improving water governance and implementing Integrated Water Resource Management (IWRM) successfully', and that there 'must be improvements to the data sets that are available to decision makers and to the systems whereby these data are distributed and displayed'. However, access to information is only one step towards improved decision-making. The availability of information and the capacity to properly interpret and apply it will determine the success of river basin management (Roehrig, 2002). Key linkages identified between data and information exchange and democracy in transboundary water management settings reinforce the need to develop systems that facilitate this vital exchange (Gerlak et al, 2011).

2.1.1. Spatial Data Infrastructure

> *Countries should document their geo-information resources and publish this information widely. Adequate documentation of existing information and data is needed to ensure its continued use. Documentation, too, should be standardised. A system enabling a potential user to search and retrieve documentation on information resources should be implemented and maintained (EIS-AFRICA, 2002).*

The latter quotation is one recommendation from EIS-AFRICA (2002) for implementing an African spatial data infrastructure (SDI), where developing a mechanism to provide improved access to geospatial information can be considered as one component for building an SDI. While there is no single definition of an SDI, it can be considered as 'the mechanisms for efficient production, management, dissemination and use of geospatial information' or 'the framework of elements/factors that are needed by a community, to *make effective use of spatial or geographic data* (UNECA, 2012). The EIS-AFRICA (2002) policy paper identified various factors required for ensuring the use of geographic

information (jointly called the spatial data infrastructure), including data availability and access; provision of core data sets; availability of metadata; standards to ensure interoperability; policies and practices encouraging the sharing of data and information; and adequate human and technical resources. At the international level, the Global Spatial Data Infrastructure (GSDI) Association seeks to 'promote international cooperation and collaboration in support of local, national and international spatial data infrastructure developments that will allow nations to better address social, economic, and environmental issues of pressing importance' (GSDI, 2012).

The *OC* provides OKACOM's stakeholders with a service to facilitate access to environmental information resources, and can be considered as a key component contributing to a spatial data infrastructure for the region. For example, as a national spatial data infrastructure initiative, GeoNetwork Opensource software[1], the platform used for *OC*, was adopted for developing the Rwanda Metadata Portal[2] (Akinyemi and Kagoyire, 2010).

2.1.2. Environmental Information Sharing for the Cubango-Okavango River Basin

It is within this international and regional context calling for improved access to information resources that the *OC*[3] online metadata catalogue has been developed by OKACOM[4]. With regard to managing the Cubango-Okavango river basin, OKACOM provides technical guidance to the three government parties of the 1994 OKACOM Agreement (Angola, Botswana and Namibia), and its permanent Secretariat, established in 2008, directs information sharing and communication activities (OKACOM, 2011b).

Recognizing that the sources of data and information relevant to its work were diverse and geographically widespread, in 2005 OKACOM requested assistance from the USAID Okavango Integrated River Basin Management (IRBM Project) (USAID, 2009) to propose an alternative to a centralised repository of such material. A key objective of this work was to provide a base to facilitate negotiation of data sharing among OKACOM's three riparian states to support joint planning and management. The result was a metadatabase: a set of descriptions of data sources that identified their content, thematic and temporal coverage and their ownership and management. These descriptions were in the format of electronic documents and it was proposed that the content be migrated to a database, incorporating the international ISO 19115 standard.

---

[1] GeoNetwork Opensource: http://geonetwork-opensource.org
[2] Rwanda Metadata Portal: http://www.cgis.nur.ac.rw/geonetwork/srv/en/main.home
3 Okavango Collections (*OC*): linked from OKACOM's homepage at http://www.okacom.org
4 The Permanent Okavango River Basin Water Commission (OKACOM): http://www.okacom.org

In 2010, OKACOM's recently established Secretariat picked up this work and, in the context of developing its new interactive web site, created *OC*, the starting point for discovery of collections of knowledge resources related to the Cubango-Okavango river basin in Angola, Botswana and Namibia. OKACOM's Secretariat has identified and updated relevant sources of information to ensure access to reliable and appropriate data to support researchers, planners and managers working in the Basin. The resource comprises a searchable directory of collections of data, data sets, image archives, books, reports and other information available both in the region and in other parts of the world (OKACOM, 2011c).

2.2. Okavango Collections: a Standards-Based and Open Source Web Catalogue

> *Standardization is one of the essential building blocks of the Information Society. There should be particular emphasis on the development and adoption of international standards. ... International standards aim to create an environment where consumers can access services worldwide regardless of underlying technology (Geneva Declaration of Principles; World Summit on the Information Society; WSIS, 2003).*

The *OC*, which is based on GeoNetwork Opensource, is a standards-based and open source geospatial metadata catalogue. System interoperability, generally referred to as the 'capability to communicate, execute programs, or transfer data among various functional units in a manner that requires the user to have little or no knowledge of the unique characteristics of those units' (ISO/IEC, 1993 cited in Lassoued et al, 2011), is supported by the use of standards which facilitate the search and discovery of data and information pertinent to management of the Okavango region. Geonetwork is an established open source software system, with a decade of development and implementation worldwide (GeoNetwork, 2011). Initially developed as a prototype in 2001 by the UN Food and Agriculture Organisation, it has since received support from the UN World Food Programme, the UN Environment Programme, the UN Office for the Coordination of Humanitarian Affairs (UNOCHA), the Consultative Group on International Agricultural Research (CSI-CGIAR), and many other contributors (GeoNetwork, 2011). The present Okavango Collections is based on version 2.6.3 of GeoNetwork with additional OKACOM-designed customisations, as described later in this paper.

2.2.1. Metadata Standards

Metadata, commonly referred to as 'structured data about data' (e.g., DCMI, 2012), provide a descriptive record of a data set or other information resource. Metadata elements typically provide a record of a resource's characteristics,

including details such as its title, creator, date, content description, language, rights/constraints, and spatial extent (e.g., DCMI, 2005; FGDC, 2011). A metadata record can be considered comparable to a library card containing the description of a book, or might be thought of as the who, what, why, when, where, and how of the resource (FGDC, 2012a). Metadata are crucial for resource search and discovery, access, and interpretation of quality and applicability; inadequate metadata can make data unusable.

Many organisations managing metadata catalogues have adopted international or national metadata standards. Standards ensure that a structured set of terms or definitions are captured to describe an information resource. Commonly used standards include those that are used to describe general documents, such as the Dublin Core Metadata Initiative (DCMI)[5], and geographic resources, including ISO19115:2003/ 19139:2007 and the Content Standard for Digital Geospatial Metadata (CSDGM) of the U.S. Federal Geographic Data Committee. ISO19115:2003 and ISO19139:2007 were developed by the TC211 committee on geographic information/geomatics of the International Organization for Standardization (ISO, 2012), and the CSDGM is commonly referred to as the 'FGDC Metadata Standard' (FGDC, 2012b). The U.S. is moving towards a national profile of ISO 19115, which will enable it to join a global spatial data infrastructure, where geographic resources are described under a common standard (FGDC, 2010).

The metadata standards adopted by *OC* are ISO 19115:2003 (ISO/TC211, 2003) and ISO/TS 19139:2007 (ISO/TC211, 2007). The ISO 19115:2003 standard comprises a set of mandatory, optional and conditional elements to describe the information resource, and ISO/TS 19139:2007 defines its XML implementation. ISO 19115 specifies a core set of seven mandatory, and eleven optional and four conditional elements for improving interoperability. The XML or eXtensible Markup Language was selected as it is the data exchange format used on the Web (Lassoued et al, 2011). While ISO19115 is directed toward description of digital geographic data (datasets, dataset series, and individual geographic features), 'its principles can be extended to many other forms of geographic data such as maps, charts, and textual documents as well as non-geographic data' (ISO/TC211, 2003). In other words, it can be applied to all types of digital and hardcopy resources, whether geographic or not, but 'certain mandatory metadata elements may not apply to these other forms of data' (ISO/TC211, 2003). This potential broad applicability of the standard is leveraged by *OC* as a means to describe hardcopy documents, such as the national atlases of OKACOM member countries.

5 Dublin Core Metadata Initiative: http://www.dublincore.org

2.2.2. Additional Standards Supporting System Interoperability

In addition to the metadata standards described above, the *OC* application makes use of a variety of other standards to support system interoperability. The overall structure of GeoNetwork is mostly compliant with the Open Geospatial Consortium (OGC) Portal Reference Architecture, by incorporating a geospatial portal, catalogue services and map viewer (OGC, 2004; GeoNetwork, 2011); the OGC is a worldwide consortium comprising government agencies, universities, research organisations and the private sector, that work together to produce open geospatial standards (OGC, 2012).

A number of mechanisms are incorporated to support harvesting of metadata from remote servers and to facilitate searching of multiple metadata catalogues from a single GeoNetwork application (i.e., collecting/copying for local storage). This is especially useful where poor connectivity hampers searching across multiple servers (i.e., a distributed search) (GeoNetwork, 2011). Standards used for retrieving metadata include the OGC Catalog Services for the Web (CSW); z39.50 protocol that is maintained and registered by the Library of Congress U.S.A; Open Archive Initiative Protocol for Metadata Harvesting (OAI-PMH); Web-based Distributed Authoring and Versioning (WebDAV); and OGC Web Services (GeoNetwork, 2011; 2012).

2.2.3. Adoption of Open Source Software

The *OC* leverages the benefits of GeoNetwork Opensource and should offer a long-term sustainable solution in the provision of an online service for the search and discovery of information resources pertinent to the Cubango-Okavango river basin. GeoNetwork Opensource is an established, tested, and widely implemented open source software system which is supported by an active community of users and developers[6]. Moreover, the application is well documented for users and developers[7], and being open source in nature (licenced under the GNU General Public Licence), the software could be customised to meet the specific needs of OKACOM. A key benefit of open source software being that innovation can be driven by end user collaboration, rather than vendors, which is typically the case with proprietary software (AGIMO, 2011).

3. KEY FEATURES AND FUNCTIONS OF OKAVANGO COLLECTIONS

The *OC* draws on the archival approach to collection level description, where a collection of material is described by one record. It is an approach that can let

6 GeoNetwork community: http://geonetwork-opensource.org/community.html [last accessed 27 April 2012]

7 GeoNetwork documentation: http://geonetwork-opensource.org/docs.html [last accessed 27 April 2012]

potential users of information know that an organisation has a collection of useful material, even if that hasn't been described at the level of specific items. This is especially relevant in a developing country environment where collections of legacy materials ranging from natural history specimens to reports and sound recordings have not yet been considered for digitization and, in many cases, do not even have electronic catalogues. Furthermore, even when databases and other electronic resources have been developed, they are still not frequently made available on the open internet. A resource like *OC* can alert and guide the user to the people who manage such collections, providing enough of a description to allow a user to decide whether it is worth their time and effort to investigate further.

The *OC* is an online geospatial catalogue application, providing metadata editing and advanced search and discovery capabilities, and an integrated interactive map viewer for displaying uploaded information layers and geographic content from remote WMS servers (GeoNetwork, 2011). The following section gives an overview of the key features and functions of *OC*, highlighting additional customisations integrated into GeoNetwork by OKACOM. The GeoNetwork User Manual (GeoNetwork, 2011) provides a comprehensive guide on its standard features and functions, and may be accessed from the '*Help'* link in *OC's* top banner. A number of OKACOM's modifications, all of which are built into OC, plan to be integrated by the GeoNetwork community into a subsequent official release of GeoNetwork Opensource software (Table 1).

Table 1: Key OKACOM Customisations that Plan to be Integrated into a Future Official Release of GeoNetwork Opensource

1. Email notification tool - to alert users on the maintenance status of metadata records (see section 3.6)
2. Improved functionality to save metadata search functionality. Searches of metadata records conducted in OC now include search terms that can be output as a PDF file; search terms are not included in the default GeoNetwork application (see section 3.5).
3. Ability to search by organization name. An additional search field for organisations was inserted into the home page interface (see section 3.1, Figure 2).
4. Ability to readily switch between languages for a given metadata record (e.g., toggle between English and Portuguese, the two official languages of OKACOM) (see section 3.4).
5. Ability to associate a clickable hyperlink to a word/phrase in the metadata abstract field (see section 3.4).

6. System loading indicator (especially useful if accessing site from area with poor internet connectivity). A small rotating image and the phrase 'searching for metadata' becomes visible when conducting a metadata search from the home page.
7. Useability of the metadata thesaurus was refined, including: automatic insertion of translated keywords (i.e., a Portuguese translation of a manually input English term will be automatically inserted into a metadata record if existing in the thesaurus, and vice versa if initially inputting Portuguese keywords); and in the theme and disciplines thesauri (based on FAO's AGROVOC thesaurus) there is indication of broader, narrower and related terms (see section 3.7).
8. Incorporation of ISO Topic category dropdown list, replacing the list of user-defined categories accessible after clicking on 'Advanced' search/ 'Restrict to' on the home page.

3.1. Home Page

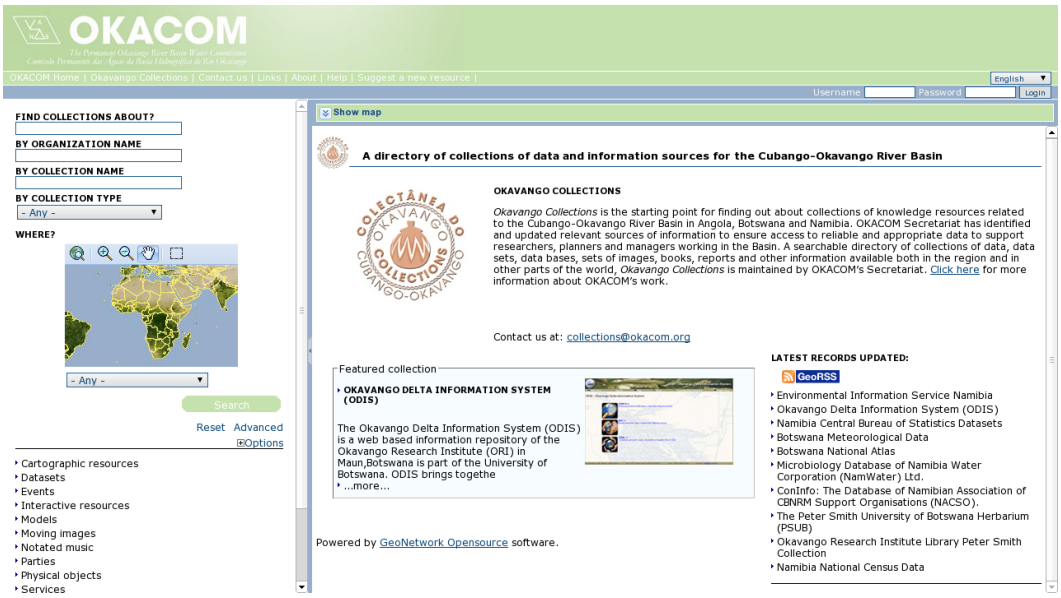
Key features of the home page (Figure 2) include: the top banner with various navigational links; selection of search options in the left column; introduction to *OC*; automatic changeable display of a featured collection; and GeoRSS feed listing the most recently updated metadata records. The top banner includes links to the home page of *okacom.org,* and *Contact us* and *Suggest a new resource* form where users can suggest additional data and information resources and/or providing suggestions and comments. The language of the site can be readily switched between English and Portuguese using the drop down menu above the login button located in the top right corner.

Additional search options were integrated into the site to enhance the ease of finding information resources and are accessible from the home page. Options include free text search of all metadata fields (labelled *Find collections about*), and searching by organisation name, collection name and collection type; these replace GeoNetwork's default *What* search field. Resources can also be searched geographically by country or region, or by clicking on the frame icon and selecting the area of interest interactively on the map; the modified list of predefined geographic areas in OC comprises all countries of Africa, plus OKACOM and South African Development Community (SADC) regions. Records can also be listed according to category (as specified in the 'Category Management' section accessible via the administration page) at the bottom of the left hand column.

Under the advanced search options (click on *Advanced*/*Restrict to*), the dropdown list of user-defined categories was replaced by an ISO Topic category

dropdown list. ISO Topic category codes are inserted into each metadata record and thus provide an additional means of resource discovery.

Figure 2: *OC* Home Page



3.2. Administration Page

The administration page allows users to access a range of functions dependent on the type of user logged in (e.g., whether an administrator or editor). Access to the administration page is via a link in the upper banner made visible after logging into the site. Key default administration functions include, among others: creating a new metadata record; user management; group management; category management (each record can be assigned to one to more user specified categories) ; importing metadata records; localization to enable multilingual possibilities for the portal (*OC* is bilingual English/Portuguese), and harvesting management (linking to other metadata catalogues). Two new OKACOM developed functionalities accessible through the administration page include: *Mail notifications* (to permit configuration of notifications of metadata changes) and *Contacts* (to provide management of metadata contacts). These are further described below.

3.3. Interactive Viewer

A map viewer is accessible by clicking on the *Show map* link located just beneath the banner. The map viewer allows geographic data to be viewed and overlain.

Images generated by both geographic data hosted on the GeoServer application integrated into *OC* and from remote servers can be displayed (via OGC Web Map Services). A PDF output of the map can be created by clicking on the printer icon.

The local *OC* GeoServer application is identified in the map viewer as *OKACOM GeoServer*, and three layers are currently being served by default. Additionally, four remote WMS servers have been added to enable users to display additional information layers. There is future potential to add more layers relevant to the Okavango region, by adding layers to the local *OC* GeoServer or from remote web services.

3.4. OKACOM's Metadata Template

All *OC* metadata records use a modified version of the ISO 19139 template provided in the default installation of GeoNetwork - *Template for Vector data in ISO19139 (multilingual).* This has been modified by inserting new elements into the template to permit inclusion of new *Collection type* (Figure 3) and MARC *Nature of contents* (Figure 4) categories (see below). New functionality also permits contacts to be readily entered from a pre-created list of contact records (see section 3.8).

Additionally, the default view of the template has been modified to make editing easier, a useful feature of GeoNetwork to simplify editing of metadata records. When a new record is created, the editor is presented with a record displaying important fields, some of which are populated with information and a few with duplicate fields to facilitate editing. For example, two fields for entering metadata in English and Portuguese are visible; the geographic area is set to OKACOM's region of interest; contact details of metadata authors are automatically included; and multiple copies of some elements (e.g., contact details, topic category code, online resources) are automatically displayed.

Functionality to permit switching the display of a given record from one language to another was implemented in OC. Previously, switching languages would require the metadata editor to exit the record, and they would then need to search for the record again in order to view in a different language. It is now possible to view in both languages by simply clicking on a button located at the top of a record.

The abstract field was modified to allow hyperlinks to be associated with text or a phrase, instead of only a link to the visible URL itself. There are two ways to insert hyperlinks into the abstract field using the OC template:

1. insert the actual hyperlink, e.g., http://www.okacom.org, which will display the URL as written (GeoNetwork's default method);

2. use the following syntax, including square brackets, to insert text which is linked to a URL (this will link the text 'the phrase to be linked')

[http://www.okacom.org|the phrase to be linked]

Figure 3: Collection Type Categories

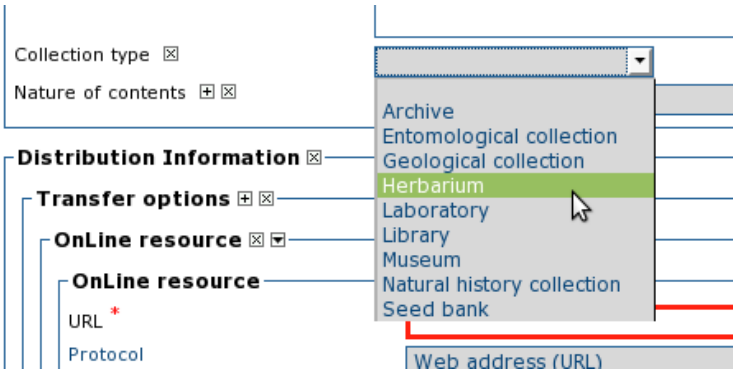
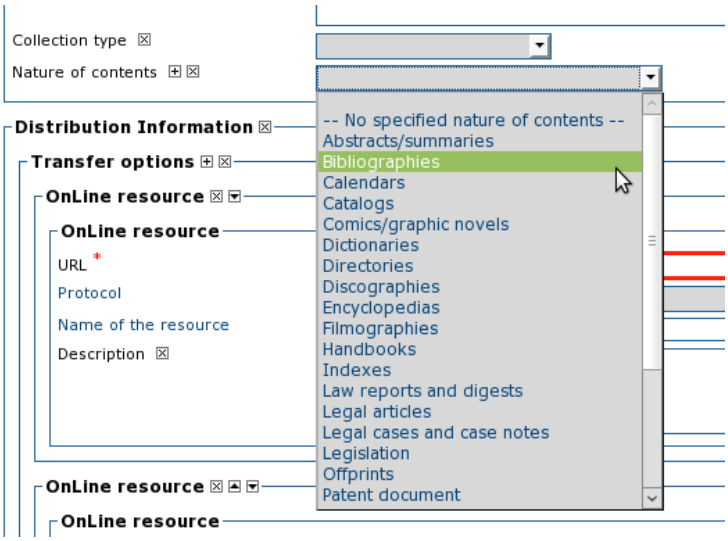


Figure 4: MARC Nature of Contents Categories



The metadata language list, from which languages used within the collections are specified, is derived from the Open Language Archives Community (OLAC)[8], a global partnership of organisations and individuals that are creating an online library of language resources. Sixty-four languages from Botswana, Namibia and Angola were inserted into the list, with language codes and spelling matching entries specified in the ISO639-3 standard (Codes for the representation of names of languages)[9]. The languages listed in the latter standard are based on two sources – Ethnologue (2012) and The Linguist List[10] (cited in SIL, 2012). The *OC* languages list will likely be further refined through consultation with OKACOM partners.

3.5. Search Results

Search results are displayed in the right panel of the page display (Figure 5). Each record listed provides an excerpt of the summary of the information resource and associated contact details of the person and organisation. This layout has been refined for OC, where the GeoNetwork default list of keywords associated with a displayed record is replaced by *Point of Contact* details. The selection of buttons displayed beneath the record, which are default features of GeoNetwork, will vary according to the privileges set for the record and the permissions of the user.

Clicking on the *More* button will reveal all of the details of the metadata record. Associated resources may be accessed via the *Download* button if available (e.g., data, documents). The *Create*, *Edit, Delete* and *Other actions* buttons are only visible to logged in users with sufficient permissions. Clicking on the latter reveals additional functionalities including *Privileges*, *Categories*, and *Create child* record to allow for description of sub-collections. Privileges set for a record determine whether a particular user group established within OC can read, download or edit the record, and whether the resource is displayed in the map viewer or featured on OC's home page in the *Featured collection* box. A record can be assigned to one of the categories listed in the left column beneath the small map, enabling users to quickly browse records so defined. For a given record being edited, the *Create* child functionality enables a new metadata *child* record to be created and linked to the latter *parent* record.

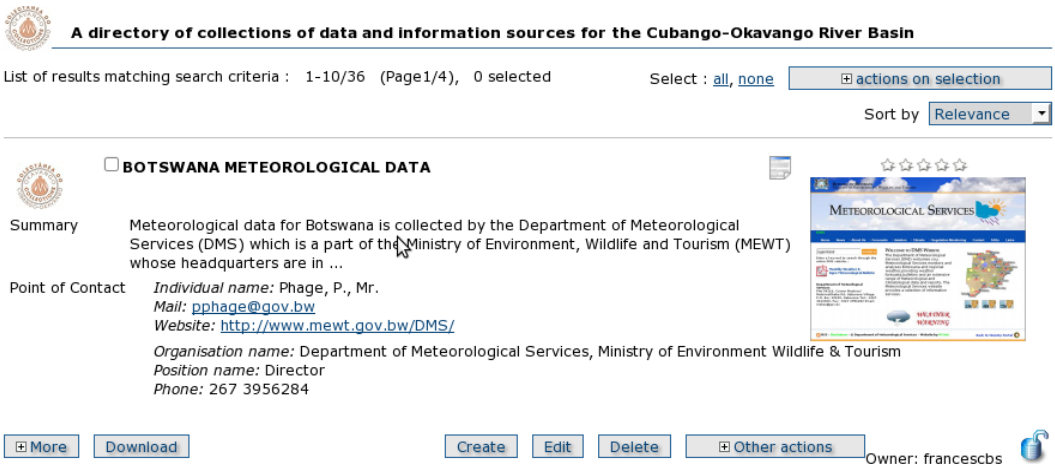
---

8 Open Language Archives Community: http://www.language-archives.org [last accessed 27 April 2012]

9 ISO 639-3 list of languages and codes: http://www.sil.org/iso639-3/iso-639-3_20100707.tab. A three-letter code is specified for known human languages. SIL International is the Registration Authority for ISO639-3: *Codes for the representation of names of languages - Part 3: Alpha-3 code for comprehensive coverage of languages*.

10 The Linguist List, International Linguists Community Online: http://linguistlist.org [last accessed 27 April 2012]

Figure 5: Search Results Layout



A variety of actions can be applied to the displayed list of results, via the *actions on selection* button located above the topmost record in the right corner of the page. Options are mostly self-explanatory with the *Export (ZIP)* a useful means of exporting (also backing up) existing metadata records, and *Export (TXT)* allows records to be viewed in a spreadsheet. The details of a search can be recorded by selecting the *Print to PDF* option. The functionality of *Print to PDF* is modified in OC to better record search strategies by including search terms as a list at the beginning of the output PDF file; for example, a simple search for ' water' would appear as follows with 14 records identified:

14 results selected from this search

Any of the words: water

All of the words:

The exact phrase:

Without the words:

Collection name contains:

Collection type:

Abstract contains:

Keywords:

Metadata change date:

Metadata temporal extent:

Topic category:

Fuzziness used: .8

An integral feature of GeoNetwork that will be later exploited by OC, is its capability to harvest (i.e., copy and store) metadata from remote catalogues and make available on the local server for searching. This is a powerful metadata sharing mechanism, with OC capable of harvesting metadata from a range of external sources including other GeoNetwork nodes and catalogues implementing a variety of protocols such as the OGC Catalog Services for the Web (CSW) (see section 2.2.2). Similarly, OC metadata can be harvested and made searchable in external catalogues. In terms of query performance, harvesting and storing external metadata on a single machine ensures that users experience a faster response than might be the case if search requests were made to two or more geographically distributed catalogues.

The OC potentially offers a valuable mechanism for sharing data and information from other environmental data initiatives in Africa and internationally. For instance, pertinent data sets and information products may be accessible from the Regional Centre for Mapping of Resources for Development (RCMRD, 2012), SERVIR-Africa (SERVIR, 2012), and the GEONETcast programme of the Group on Earth Observations (GEO, 2012a). Another opportunity for increased data availability is GEO's recent launch of the AfriGEOSS initiative that aims to implement GEO activities and coordinate with other Earth observation activities in Africa (GEO, 2012b).

3.6. Email Notification Tool

To ensure currency of information in *OC*, the OKACOM-designed email notification tool alerts administrators and content reviewers of the maintenance status of metadata records. Email notification settings (Figure 6) are accessed through the administration page.

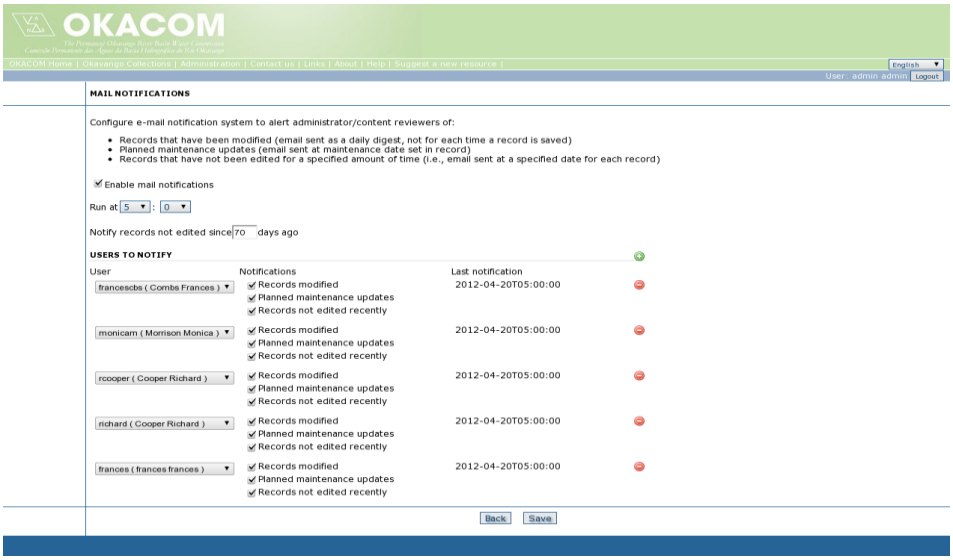
Email alerts are sent as a daily digest indicating which records:

(i) have been modified during the last 24 hours;

(ii) require updating (maintenance date set in metadata record);

(iii) have not been edited for a specified amount of time (number of days set in mail notifications tool).

Figure 6: Mail Notifications Configuration Page



3.7. Keywords and Thesaurus Development

Metadata keywords can be selected from one of five types of thesauri developed for *OC* (as per the ISO 19115:2003 standard): theme; place; discipline; temporal; and stratum. The *theme* thesaurus comprises the FAO's AGROVOC thesaurus[11]; this is a managed resource, comprehensive, and applicable to OKACOM's catalogued information resources. The *place* thesaurus is based on place names derived from the GeoNames geographical database[12], which will subsequently be refined through consultation with the naming authorities of Angola, Botswana and Namibia. The *discipline* thesaurus is based on terms derived from the AGROVOC thesaurus. Temporal and stratum terms are derived from vocabulary lists of the Oregon Geospatial Enterprise Office of the Oregon State Government[13] and the GIS unit of the Oregon Department of Forestry[14]; the stratum keyword list gives an idea of the vertical location represented by the dataset.

11 FAO's AGROVOC thesaurus: Downloaded in December 2010 from http://aims.fao.org/website/Download/sub.
12 GeoNames geographical database: http://www.geonames.org. Data sources used by GeoNames: http://www.geonames.org/data-sources.html
13 Oregon Geospatial Enterprise Office of the Oregon State Government: http://gis.oregon.gov/thesaurus
14 GIS unit of the Oregon Department of Forestry: http://www.odf.state.or.us/gis/docs/metadata_keywords.doc

3.8. Contacts Management

OKACOM customisation permits insertion of contact details into a metadata record from a list of contact records previously created via the *Contacts Management* page (accessible through the administration page). A user with administrator or editor role can add contact details via the *Contacts Management* page, and these records can be subsequently inserted into a metadata record; a popup box listing contacts is accessible from the latter. The benefits of this approach are: modifications made to the contacts management list will update all linked metadata records, and inserting contact details into a metadata record is faster and more accurate. If a person is not present in the contacts management list, then the metadata editor can enter new contact details.

4. CHALLENGES AND LESSONS LEARNED

The OC provides a unique service for facilitating information sharing about the Cubango-Okavango river basin. Prior to the implementation of OC, there was no single catalogue, offline or online, available to environmental practitioners and stakeholders. Key outcomes of the project include making environmental information both more discoverable and accessible as each resource includes a point of contact. Furthermore, by helping ascertain the extent of current information resources, OC helps new projects and initiatives avoid duplication of effort in creating new data sets and loss of scarce financial resources.

Integral to its development, OC also aims to provide a *sustainable solution* for sharing environmental metadata, where the system offers scope for addressing future technical, management and stakeholder needs. Key requirements for the metadata catalogue included: developing a user friendly system that offers flexibility for customisation of design and functionality; the capacity to integrate bilingual metadata; and incorporation of international metadata standards to ensure effective sharing of metadata and discovery of information resources. As such, the open source GeoNetwork application was identified as the best available option to address these needs both in the context of its functionality and useability, but also financially as OC builds on an existing system developed over many years thereby avoiding the need for extensive and costly programming effort. During OC's development however, a number of challenges were encountered which are further elucidated below.

4.1. Outdated Information

The data gathered in 2005 as part of the IRBM project (USAID, 2009) was the starting point for development of the web-based resource. Since the beginning, it was clear that updating of the data would be required: in six years managers would have changed, collections enlarged and digitized, and new collections would have been established. OKACOM Secretariat contracted a dedicated

researcher for six months to establish the scope of the updating work required, using telephone and email communications.

It was found that, while the organizations and collections described in 2005 still existed, many had merged and changed names. Contacting and eliciting information from government institutions – OKACOM's most important stakeholders – was a challenge. This was especially so in the context of Angola, where work on many collections had been suspended during years of civil war and new institutional structures for their support were not yet fully functional. Explaining the nature of the *OC* work was made easier because of data owners' knowledge of OKACOM's mandate and responsibility, and the requisite level of trust was established fairly quickly. Nevertheless, simply updating information about a collection previously identified required at least three working days of contact and follow-up, as draft descriptions were compiled, edited and sent to the owners for review.

4.2. Item Versus Collection

Parallel to updating of data about the collections was adaptation of the GeoNetwork templates for entry of data. This turned out to be more of an iterative process than had been anticipated. At the beginning of the development work, it had become clear that the work carried out in 2005 had not distinguished between item and collection level description. For example, single document-based datasets and individual maps had been treated at the same level as whole collections of reports and images. While this approach did capture the existence of useful resources, no matter what type, and was well suited to the Geographic Information Systems (GIS) nature of other GeoNetwork resources, it caused some delay in *OC*'s development while the development team clarified what was to be included in the work. Eventually it was agreed that *OC* would focus on collections made up of multiple resources, and any individual resources of value identified in the process would be catalogued as items through OKACOM's library catalogue and online document retrieval systems.

4.3. Adapting GeoNetwork's GIS Metadata Bias

Another challenge related to the structure of the GeoNetwork database was the shortage of metadata capacity for description of the nature of content of a variety of types of collections. Templates designed for specific types of spatially referenced data such as maps were not necessarily well suited for describing other types of materials. Adaptation of the templates eventually revealed the need for addition of fields and search capacity to the final GeoNetwork resource.

The *OC* metadata template was therefore modified to incorporate two new elements defining MARC *Nature of contents* categories[15] and *Collection type* categories. MARC refers to the MAchine-Readable Cataloging standard for recording bibliographic information and is used by libraries worldwide (LOC, 2012). Nature of contents categories, drawn from 30 MARC categories, help identify whether a collection contains certain types of material such as legislation, statistics and treaties. Collection type categories specified by the OC metadata template include: Archive; Entomological Collection; Geological collection; Herbarium; Laboratory; Library; Museum; Natural history collection; and Seed bank.

4.4. Identifying Existing Naming Authorities

Unrelated to GeoNetwork's database structure and built-in assumptions about the nature of content was the challenge presented by the lack of lists of standard names for description of regional resources. Name of places and of languages in particular were problematic as none of OKACOM's three riparian member states had established authorities for these. Eventually compromise solutions were found to enable *OC*'s development, but the political implications of this work will continue to be a challenge for some time to come. Overall, the standards and supporting institutions for systematic content description at a regional and local level are lagging far behind the capacity of technology to assist in managing this information.

4.5. Useability

OC is intended to be a tool for those seeking information to support work in the river basin and understanding of its nature and natural resources. These users range from academic researchers to government officials to business people. A concern is that many of these non-technical users, familiar with user-friendly and aesthetically designed mass market online services would find the detailed and heavily structured information in the resource difficult to manipulate. As a resource discovery tool, it was important that it be easy to use and its powerful data management tools be transparent to the user. Considerable time and effort was spent in making the user interface clean, suppressing display of information such as lists of keywords in the results display (replaced by point of contact information in *OC*) that was not central to a user's quest. Furthermore, training on creating and editing metadata records, user and system management were provided to OKACOM by the GeoNetwork consultant, together with detailed documentation.

15 MARC 21 Bibliographic - Full. URL: http://www.loc.gov/marc/bibliographic/bd008b.html [last accessed 26 April 2012]

Other potential factors affecting useability included low bandwidth and electricity shortage in Maun, the Secretariat's location. These were addressed by hosting *OC* with a commercial web-hosting company in South Africa[16], who also provide IT support for OKACOM's main website (www.okacom.org).

4.6. Software Customisation

A key benefit of GeoNetwork is that is open source software which allows programming source code to be modified and the system customised to meet user needs. Such customisation can be considered at two levels of complexity: the first addressing the design and layout of the system, and the second concerning system functionality. At the first level, a person with competent IT skills but lacking specialized programming skills would be capable of refining certain features of the system, such as changing colours, logos, text, predefined geographic areas selectable under the small map and categories in the home page's left column, and the list of data set languages accessible through the metadata template. At the second level of complexity, programming skills are needed to add new elements into the metadata template (e.g., as with the creation of the OKACOM template) and all of the additional functionalities incorporated into *OC* as listed in Table 1. Incorporating new functionality can be a time-consuming process, as experienced with *OC*, necessitating substantial testing and evaluation to ensure the system is free from any programming error.

4.7. *OC* Sustainability

Key to sustainability of a resource like *OC* is the relevance of its content to the ongoing work of the organization that maintains it. Management of information resources can only make sense in the context of an organization's activities. *OC* meets this criterion, by providing OKACOM with a container and tracking system for discovered resources that are relevant to its work and by functioning as a networking tool to stay actively engaged with the organization's large and varied stakeholder community. Awareness of *OC* should be further advanced through a proposed future linkage to the SADC-Water Information Sharing Hub (SWISH) System[17].

Another desired feature that could be developed as OC is used and grows is a visualisation tool that quickly displays the relative likelihood of discovering needed data in a collection.

Training of stakeholders in use of *OC*, and in contributing information about their collections, is an essential element in OKACOM's information management

[16] http://www.upfrontsystems.co.za
[17] http://www.sadcwaterhub.org

planning, but the assumption is that the amount and quality of dedicated support and promotion from the Secretariat will determine its success in meeting OKACOM's objectives.

There is, however, a cost to all this utility. Systems to manage codified knowledge resources such as large databases still require substantial human input. OKACOM's small secretariat may be challenged to dedicate the resources needed to maintain *OC* as a tool with current, high quality content. Such work does not lend itself to outsourcing since it is OKACOM's core business to be aware of the best information available about the river basin, and the constant flow of new information into the secretariat from partner projects is best captured when it reaches OKACOM.

These factors were kept in mind when *OC* was developed. Features for partial decentralization and automation of content maintenance were built in as customizations: a facility to assign data editing permissions to owners and managers of collections, an alerting system to remind *OC*'s managers that a particular resource had not been updated for some time and provision of an online form to allow users to identify new resources for inclusion in *OC*.

OKACOM learned through creating *OC* that there is a role for dedicated information management expertise to provide the human interface between owners and users of collections of data and information. With this expertise in place, currency and relevance will be ensured and the full power of the technology used.

5. RECOMMENDATIONS

Based on challenges encountered and lessons learned, the following key recommendations for future metadata catalogue development are presented as follows.

1. Sustainability. From the outset of development, long-term management of the catalogue requires careful consideration. Critically, metadata needs updating to ensure its reliability and value to stakeholders. Financial planning is needed to ensure continuity of service.

2. Flexibility of design/useability. In the initial stages of assessing potential metadata catalogue software options, include open source systems as part of the evaluation. Existing open source systems, such as GeoNetwork, can offer sophisticated and well-used systems that can be modified to meet specific requirements.

3. Capacity building. Staff will need training with regard to creating and editing metadata (i.e., ensuring details of core elements are captured), and more generally the metadata catalogue software including user management.

4. IT administrative support. If IT expertise is not present within an organisation, then consider outsourcing. This is the option adopted by OKACOM, where *OC* is hosted and administratively supported by a regionally-located company. Furthermore, external hosting may help address issues of poor internet connectivity and electricity supply.

5. Assess metadata capacity. For example, the ISO19139 metadata standard was modified by incorporation of extra fields to better meet OKACOM's information needs.

6. The rapidly changing global information environment requires adaptation and repurposing of standards from different disciplines. A tool designed to manage a specific type of information can be adapted and the planning stage is of key importance. Patient testing and re-testing of the appropriateness of metadata schema should be part of this process.

6. CONCLUSIONS

The *OC* represents the first online metadata catalogue focusing on the Cubango-Okavango river basin. Collections of data and information are now more accessible to OKACOM's stakeholders. The resource offers a single point of entry to search, discover, and retrieve (where access constraints permit) relevant information about the region. Details about the contents of various collections of data and information can be readily accessed and their potential value determined, and the names and contact details of people and agencies responsible for their management should allow for more effective use of existing information resources.

In addition to *OC* providing a key environmental information service for the Cubango-Okavango region, it potentially offers a model system for implementation elsewhere. From a technical perspective, *OC* builds additional features and functionality on top of the already well established GeoNetwork Opensource software. In particular the email notification tool will facilitate the updating of metadata records and their management. As with any web-based system, it is crucial post-development that the system is actively maintained, to sustain its services and also the confidence of users that the information is up to date and accurate. Maintaining system sustainability is arguably as great a challenge as initial development given the need for ongoing financial, technical, and management support.

ACKNOWLEDGEMENTS

The support of OKACOM Secretariat in facilitating this work is acknowledged. Special thanks go to Frances Murray-Hudson for her tireless work in helping compile the metadata records of *OC*. The completeness and quality of the records reflects Frances' commitment and effort, including extensive consultations with stakeholders in member countries to ensure the currency of records.

REFERENCES

AGIMO - Australian Government Information Management Office (2011). *A Guide to Open Source Software for Australian Government Agencies*, Version 2.0, at http://www.finance.gov.au/e-government/infrastructure/docs/AGuidetoOpenSourceSoftware.pdf [accessed 27 April 2012].

Andersson, L., Wilk, J., Todd, M.C., Hughes, D.A., Earle, A., Kniveton, D., Layberry, R. and H.H.G. Savenije (2006). Impact of climate change and development scenarios on flow patterns in the Okavango River, *Journal of Hydrology*, 331(1-2): 43–57.

Akinyemi, F. and C. Kagoyire (2010). The Rwanda Metadata Portal: A Web Catalogue Service, *International Journal of Spatial Data Infrastructures Research*, 5: 382-401.

AWCS (2011). *White Paper on GEO Capacity Building and Water Resources in Africa, The African Water Cycle Coordination Initiative (AfWCCI)*, 2nd GEOSS African Water Cycle Symposium, February 23-25 2011, Addis Ababa, Ethiopia, at http://www.earthobservations.org/documents/meetings/201102_2nd_awcs/20110223_25_2nd_awcs_cb_white_paper.pdf [accessed 20 April 2012].

Biggs, R., Bohensky, E., Desanker, P.V., Fabricius, C., Lynam, T., Misselhorn, A.A., Musvoto, C., Mutale, M., Reyers, B., Scholes, R.J. Shikongo, S. and A.S. van Jaarsveld (2004). “Nature Supporting People, The Southern African Millennium Ecosystem Assessment”, Council for Scientific and Industrial Research, Pretoria, South Africa.

DCMI - Dublin Core Metadata Initiative (2005). *Using Dublin Core - The Elements*, at http://dublincore.org/documents/usageguide/elements.shtml [accessed 5 August 2012].

DCMI - Dublin Core Metadata Initiative (2012). *DCMI Frequently Asked Questions (FAQ)*, at http://dublincore.org/resources/faq/#whatismetadata [accessed 30 July 2012].

EIS-AFRICA (2002). *Geo-information supports decision-making in Africa, An EIS-Africa position paper*, at http://geoinfo.uneca.org/sdiafrica/Chap_HTML/..\Reference\Ref0\SDI Declaration for WSSD_en.pdf [accessed 20 April 2012].

Ethnologue (2012). *Languages of the World*, at http://www.ethnologue.com [accessed 27 April 2012].

FGDC - Federal Geographic Data Committee (2010). *Federal Geographic Data Committee Educational Resource: Preparing for International Metadata. National Spatial Data Infrastructure*, Version 2, at http://www.fgdc.gov/metadata/documents/preparing-for-international-metadata-guidance.pdf [accessed 26 April 2012].

FGDC - Federal Geographic Data Committee (2011). Geospatial metadata. Federal Geographic Data Committee. July 2011, at http://www.fgdc.gov/library/factsheets/documents/GeospatialMetadata-July2011.pdf [accessed 5 August 2012].

FGDC - Federal Geographic Data Committee (2012a). *Geospatial metadata*, at http://www.fgdc.gov/metadata [accessed 21 August 2012].

FGDC - Federal Geographic Data Committee (2012b). *Geospatial Metadata Standards*, at http://www.fgdc.gov/metadata/geospatial-metadata-standards [accessed 26 April 2012].

GEO - Group on Earth Observations (2012a). *GEONETCast, Delivering operational services to users*, at http://www.earthobservations.org/geonetcast.shtml [accessed 7 August 2012].

GEO - Group on Earth Observations (2012b). *AfriGEOSS: An initiative to reinforce Earth observation in Africa*, at http://www.earthobservations.org/art_019_002.shtml [accessed 7 August 2012].

GeoNetwork (2011). GeoNetwork Opensource, *GeoNetwork User Manual*, Release 2.6.3, at http://geonetwork-opensource.org/manuals/2.6.3/users/GeoNetworkUserManual.pdf [accessed 27 April 2012].

GeoNetwork (2012). *Functionality provided by GeoNetwork*, at http://geonetwork-opensource.org/functionality.html#functionality-provided-by-geonetwork [accessed 27 April 2012].

Gerlak, A., Lautze, J. and M. Giordano (2011). Water resources data and information exchange in transboundary water treaties, *International Environmental Agreements: Politics, Law and Economics*, 11(2):179-199.

GSDI (2012). *Global Spatial Data Infrastructure Association*, at http://www.gsdi.org [accessed 22 April 2012].

ISO - International Organization for Standardization (2012). *International Standards for Business, Government and Society*, at http://www.iso.org/iso/home.html [accessed 28 April 2012].

ISO/IEC (1993). *ISO 2382-1: Information technology – Vocabulary – Part 1: Fundamental terms, International Organization for Standardization, 3.0 edition* at http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=7229 [accessed 26 April 2012].

ISO/TC211 (2003). *ISO 19115:2003, Geographic Information - Metadata, International Organization for Standardization*, at http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=26020 [accessed 26 April 2012].

ISO/TC211 (2007). *ISO/TS 19139:2007. Geographic information - Metadata - XML schema implementation, International Organization for Standardization*, at http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=32557 [accessed 26 April 2012]

Kgathi, D.L., Kniveton, D., Ringrose, S., Turton, A.R., Vanderpost, C.H.M., Lundqvist, J. and M. Seely (2006). The Okavango; a river supporting its people, environment and economic development. *Journal of Hydrology,* 331: 3-17.

Lassoued, Y., Pham, T.T., Bermudez, L., Stocks, K., O'Grady, E., Isenor, A. and P. Alexander (2011). "Coastal Atlas Interoperability", in Wright, D.J., Dwyer, N. and V. Cummins (Eds.) *Coastal Informatics: Web Atlas Design and Implementation*. USA: Information Science Reference.

LOC - Library of Congress (2012). *What is a MARC record, and why is it important?*, at http://www.loc.gov/marc/umb/um01to06.html [accessed 26 April 2012].

OGC - Open Geospatial Consortium (2004). *Geospatial Portal Reference Architecture, A Community Guide to Implementing Standards-Based Geospatial Portals, OGC 04-039, version 0.2*, at http://portal.opengeospatial.org/files/?artifact_id=6669 [accessed 26 April 2012]

OGC - Open Geospatial Consortium (2012). *OGC Press Releases*, at http://www.opengeospatial.org/pressroom/pressreleases [accessed 26 April 2012].

OKACOM - The Permanent Okavango River Basin Water Commission (2011a). *Cubango-Okavango River Basin Transboundary Diagnostic Analysis,* Maun, Botswana, at http://iwlearn.net/iw-projects/842/reports/okavango-river-basin-transboundary-diagnostic-analysis-tda-2011 [accessed 20 April 2012].

OKACOM - The Permanent Okavango River Basin Water Commission (2011b). *The Permanent Okavango River Basin Water Commission, (OKACOM)*, at http://www.okacom.org/okacom-commission [accessed 4 August 2012].

OKACOM - The Permanent Okavango River Basin Water Commission (2011c). *Okavango Collections. A directory of collections of data and information sources for the Cubango-Okavango River Basin*, at http://www.okacom.org [accessed 4 August 2012].

Ramsar (2012). *Ramsar Sites Information Service*. Botswana 1BW001. Okavango Delta System, at http://ramsar.wetlands.org/Database/Searchforsites/tabid/765/ Default.aspx [accessed 30 July 2012].

RCMRD - Regional Centre for Mapping of Resources for Development (2012) *Data dissemination*, at http://www.rcmrd.org/index.php?option=com_content&view=article&id=68 &Itemid=64 [accessed 7 August 2012].

Roehrig, J. (2002). Information Interoperability for River Basin Management, *Journal of Technology Resource Management and Development*, 2: 127-134.

SERVIR (2012). *SERVIR Africa*, at http://www.servir.net/africa/index.php?option=com_content&task=view&id=26&Itemid=46 [accessed 7 August 2012].

SIL - Summer Institute of Linguistics international (2012). *ISO 639-3. This is the home page for Part 3 of the ISO 639 family of standards, Codes for the representation of names of languages*, at http://www.sil.org/iso639-3/default.asp [accessed 27 April 2012].

UNECA - United Nations Economic Commission for Africa (2007). *Accelerating Africa's development to meet the MDGs: Challenges and the way forward for Southern Africa*. Report of the 13th Intergovernmental Committee of Experts Meeting for Southern Africa (ICE), March 14-16 2007, Lusaka, Zambia.

UNECA - United Nations Economic Commission for Africa (2012). *Spatial Data Infrastructure (SDI) Explained", in SDI-Africa: An Implementation Guide*, at http://geoinfo.uneca.org/sdiafrica/Chap_Pdf/02SDIEXplained.pdf [accessed 20 April 2012].

USAID (2009). *Okavango Integrated River Basin Management Project (IRBM). Final Report*, at http://www.okacom.org/okacoms-work/partners-and-projects/projects/partner-projects/irbm/IRBM%20Final%20Report%20-%20August%2031%202009.pdf [last accessed 30 July 2012]

WSIS - World Summit on the Information Society (2003). *Geneva Declaration of Principles*, WSIS-03/GENEVA/DOC/0004, at: http://www.itu.int/wsis/docs/geneva/ official/dop.html [accessed 22 April 2012].

WSSD - World Summit on Sustainable Development (2002). *Plan of Implementation of the World Summit on Sustainable Development*, Johannesburg, at http://www.un.org/esa/sustdev/documents/WSSD_POI_PD/English/WSSD_PlanImpl.pdf [accessed 20 April 2012].