Open Geospatial Consortium, Inc.

Date: 2010-08-18

Reference number of this document: OGC 10-086r1

Category: OGC Public Engineering Report

Editor(s): Andrew Turner

OGC® OWS-7 - Authoritative Data Source Directory Engineering Report

Warning

This document is not an OGC Standard. This document is an OGC Public Engineering Report created as a deliverable in an OGC Interoperability Initiative and is not an official position of the OGC membership. It is distributed for review and comment. It is subject to change without notice and may not be referred to as an OGC Standard. Further, any OGC Engineering Report should not be referenced as required or mandatory technology in procurements.

Document type:	OpenGIS® Engineering Report
Document subtype:	NA
Document stage:	Approved for public release
Document language:	English

Preface

This document presents the Authoritative Data Source Directory (ADSD) engineering suggestions and results of the OGC OWS-7 ADSD thread. This group focused on creating a workflow for geospatially referencing, finding, and federating data sources with associated authority and relevance.

Contents

Page

Figures

Page

Tables

Page

OGC® OWS-7 - Authoritative Data Source Directory Engineering Report

1 Introduction

1.1 Scope

The Authoritative Data Source Directory (ADSD) is a resource capable of organizing and discovering a wide variety of types of data such as web sites, books, pictures/images, et al. as well as available web services. The directory will have the ability to identify and query for data sources based on socio-cultural themes, geographic area (either coordinates or geographic name), temporal relevance, and data quality (*e.g.* precision, fitness for use).

The ADSD concept will be implemented as an OGC catalog service supporting all the interfaces of CSW 2.0.2. ADSD extends CSW, adding the following features:

- Enforces certain metadata requirements on data entries, such as inclusion of geography, time, data quality, authoritativeness, usefulness rating, provenance, and keywords.
- Provides an OpenSearch query interface with support for the Suggestions extension.
- Catalogs OWS Context (OWC) documents
- Provides access to taxonomic lists (such as socio-cultural themes) via the OpenSearch interface with support for the Suggestions extension.

1.2 Document contributor contact points

All questions regarding this document should be directed to the editor or the contributors:

Name	Organization
Andrew Turner	FortiusOne
Roger Brackin	Envitia
Raj Singh	Open Geospatial Consortium
Mark Lawrence	Compusult
Phillippe Duchesne	ERDAS
Roberto Lucchi	ESRI

1.3 Revision history

Date	Release	Editor	Primary clauses modified	Description
3/25/10	0.1	Andrew Turner		Initial draft
6/12/10	0.2	Andrew Turner		Updated Word Draft

6/25/10	0.3	Raj Singh, Roger Brackin	All	Update of Enterprise VP/Use Cases, Technology VP
6/26/10	0.4	Andrew Turner	All	Final Edits and Corrections
6/30/10	0.9	Raj Singh, Andrew Turner, Roger Brackin	Computation and Information VPs	Final contributions from participants

1.4 Future work

Improvements in this document are desirable to address open issues; to correct errors or enhance existing document content.

The goal of this testbed was to provide for standards in finding, storing, and sharing authoritative data. Through the work of the testbed these techniques were demonstrated and key problem areas identified. Future work includes mechanisms for sharing authoritative metrics across servers in order to provide for consistent scoring and ranking.

1.5 Forward

The purpose of this engineering report (ER) is to document the design concepts, equivalent standards, and implementation guidelines as pursued in the OWS-7 Testbed. It encompasses a variety of viewpoints, each of which explains the considerations and requirements based on the viewpoint.

The Enterprise Viewpoint provides the context of the use of an ADSD within operational processes and goals of an organization and mission. The Computational Viewpoint concerns the functional decomposition of the system into a set of distinct components that interact at interfaces. The Information Viewpoint documents the information models and encodings of the content and services that were outlined in the Computational Viewpoint. Lastly, the Engineering Viewpoint describes the specific components linked by a communications network and the implementable interactions of distinct computational objects. A summary of issues identified and further work outlines areas of future research and exploration.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. The Open Geospatial Consortium Inc. shall not be held responsible for identifying any or all such patent rights.

Recipients of this document are requested to submit, with their comments, notification of any relevant patent claims or other intellectual property rights of which they may be aware that might be infringed by any implementation of the standard set forth in this document, and to provide supporting documentation.

2 References

The following documents are referenced in this document. For dated references, subsequent amendments to, or revisions of, any of these publications do not apply. For undated references, the latest edition of the normative document referred to applies.

1. Department of Defense Discovery Metadata Specification (DDMS) (http://metadata.dod.mil/mdr/irs/DDMS/)

2. Dublin Core Metadata Element Set, Version 1.1 (http://dublincore.org/documents/dces/)

3. IETF RFC4287, *The Atom Syndication Format* (http://www.ietf.org/rfc/rfc4287.txt)

4. ICIQ-04, *Proceedings of the Ninth International Conference on Information Quality*, Frank J. Ponzio, Jr., 2004

5. ISO 15836:2009, Information and documentation--The Dublin Core metadata element set (http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=52142)

6. OGC 05-062, *OGC Web Services Context Interoperability Experiment Final Report*

7. OGC 07-124r2, *OpenGIS® OWS-5 KML Engineering Report*

8. OGC 07-006r1, *OpenGIS Catalogue Service Implementation Specification* (http://portal.opengeospatial.org/files/?artifact_id=20555)

9. OGC 09-084r1, *OpenSearch Geospatial Extensions Draft Implementation Standard* (http://portal.opengeospatial.org/files/?artifact_id=35983)

10. OGC 10-035, *OWS-7 Information Sharing ER* (http://portal.opengeospatial.org/files/?artifact_id=39520&version=1)

11. OGC Geospatial Fusion Services Testbed [working documents provided on request from techdesk@opengeospatial.org]

12. OGC Web Services Context Schema v0.3.1 (http://www.ogcnetwork.net/schemas/owc/0.3.1/)

13. OpenSearch 1.1 draft 4 (http://www.opensearch.org/Specifications/OpenSearch/1.1/Draft_4)
14. OpenSearch Extensions (http://www.opensearch.org/Specifications/OpenSearch/Extensions)

3 Terms and definitions

For the purposes of this report, the definitions specified in Clause 4 of the OWS Common Implementation Specification [OGC 06-121r3] and in OpenGIS® Abstract Specification shall apply. In addition, the following terms and definitions apply.

3.1 Authorititive Data Source Data Directory (ADSD)

A repository which stores links to data services and assigns an authoritive ranking to the catalogues entities based on a ranking scheme (e.g. human terrain, dataset reliability etc).

3.2 OpenSearch

OpenSearch, Ref[13], is a collection of simple formats for the sharing of search results. Specifically provides a templated interface to allow search services to document query parameters in a simple search request. It is most prominently used in web browsers to enable the quick search feature for querying popular search engines and websites.

3.3 Ranking

A quantitive measure of the 'authoritiveness' of a dataset based on a ranking scheme relevant to a given user or community. Includes a provinence (sponsor) for the ranking scheme (e.g. NGA).

3.4 Tagging

(from Wikipedia) A non-hierarchical keyword or term assigned to a piece of information (such as an Internet bookmark, digital image, or computer file). This kind of metadata helps describe an item and allows it to be found again by browsing or searching. Tags are generally chosen informally and personally by the item's creator or by its viewer, depending on the system. Tagging was popularized by websites associated with Web 2.0 and is an important feature of many Web 2.0 services.

4 Conventions

4.1 Abbreviated terms

API	Application Program Interface
ADSD	Authoritative Data Source Directory
CSW	Catalog Services for the Web
DDMS	Department of Defense Discovery Metadata Specification (see Ref [1])
FDF	OWS-7 Feature and Decision Fusion Thread
HT	Human Terrain
JSON	JavaScript Object Notation
KML	Keyhole Markup Language
OGC	Open Geospatial Consortium
OWC	OGC Web Context
OWS-7	OGC Web Services Testbed, Phase 7
WCS	Web Coverage Service
WFS	Web Feature Service
WMS	Web Mapping Service

5 Enterprise Viewpoint

The context for feature and decision fusion can be one of critical urgency. Lives, resources, diplomatic efforts, and civil order may depend on the outcomes of these fusion processes. Complicating the analytical processing is the need for coherent and consistent communications and data throughout a hierarchical chain of command. This is true in both defense and civil applications. The regional and national organizations involved could be emergency response agencies, centers for disease control, theater and national defense commands, and so on.

A key part of supporting feature and decision fusion are catalog, or registry services, providing sophisticated capabilities to discover, organize and access relevant data sources. One currently popular term for this kind of information service is an Authoritative Data Source. Ponzio defines the motivation for an ADS as follows:

"Throughout the Federal Government, including the Department of Defense (DoD), there is a pervasive need to have current, reliable, and trusted data from what are termed Authoritative Data Sources (ADSs). This requirement has grown increasingly important as the military transforms. The transformation requires individual organizations, each using an array of independently developed stovepipe systems, to function on the battlefield with other weapons systems, services, and friendly nations, sharing communications networks and data. To ensure the accuracy of the data being provided by ADSs, new methodologies and metrics must be initiated."

"Authoritative Data Source (ADS) Framework and ADS Maturity Model,"
Proceedings of the Ninth International Conference on Information Quality (ICIQ-04)
Frank J. Ponzio, Jr., 2004

Authoritative data sources, however, are rarely comprehensive enough to fulfill all the needs of analysts. They still need to look beyond the confines of "official" or enterprise data sets, reaching out into other structured data repositories and also leveraging the vast, unstructured open source content available on the Internet.

In the ADSD sub-thread of the OWS-7 Feature and Decision Fusion thread, participants analyzed relevant existing technologies from within the OGC standards baseline as well as the general information technology space. The goal was not only to provide services for geospatial data, but also to integrate geospatial data discovery methods with those from non-geospatial domains such as web sites, books, pictures/images, etc. Due to the emphasis on Humanitarian Assistance and Disaster Relief operations undertaken by military organizations, particular attention was paid to discovering, organizing, and categorizing information that falls into the discipline of Human Terrain Analysis (see Table 1 and Table 5).

Detailed search requires metadata. Even seemingly unstructured searches use metadata, they just use custom metadata schemes that aren't communicated to users. Web search engines, for example, closely guard the secret of what metadata is considered important because they

consider that a key facet of their market advantage. The change from traditional geospatial and information sharing is that web sites and "open source" data often don't include highly detailed metadata. Relying on detailed metadata, as is used in previous OGC efforts is not always possible. Therefore it was necessary to consider more generalized mechanisms for searching and federating data.

The ADSD is a resource capable of organizing and discovering a wide variety of types of structured and unstructured data and services. It has the ability to identify and query for data sources based on socio-cultural themes, geographic area (either coordinates or geographic name), temporal relevance, and data quality (fitness for use).

For this purpose, OpenSearch was considered as an attractive solution that is commonly used amongst web services and catalogs that include both detailed metadata records as well as publicly available, open-sourced data. OpenSearch provides mechanisms for defining loosely coupled services. In order to support the various facets the Geo, Time, Suggestions and Rankings extensions are included.

The ADSD concept was implemented as an OGC catalog supporting a custom OpenSearch interface, which is comprised of the basic OpenSearch API plus a number of extensions which provide the following functionality:

- Enforcing certain metadata requirements on data entries, such as inclusion of geography, time, data quality, authoritativeness, usefulness rating, provenance, and keywords.
- Providing "type-ahead" content completion hints to the user via the OpenSearch Suggestions extension.
- Cataloging of OGC Web Services Context (OWC) documents.
- Providing access to taxonomic lists (such as socio-cultural themes).

Information to be discovered falls into a number of categories spanning both geospatial and non-geospatial as well as from highly-structured to unstructured narratives or media as described below in Table 1.

	Geospatial	**Non-geospatial**
Highly Structured	vector, raster data with well defined format/content model	Well defined database schema/file specification
	e.g VMAP, AIXM	e.g tabular, Database Records
Loosely Structured	Loosely structured data, geospatial, with some format constraints.	Loosely structured, non-geospatial data with some format constraints
	OpenStreetMap	Semantic content, RDF+OWL

Unstructured	No formal geospatial structure but clearly a geospatial entity	No formal structure, general text, pdf or html. Data without an obvious geospatial reference abiguous or not
	VGI – SMS or Twitter reports (*e.g.* Ushahidi)	Web sites, PDFs, books and binary (video, images)

Table 1: Types of Information shared through an ADSD

The goal of the ADSD is to provide the ability to discover, share, publish, and consume this array of data in such a way as to include measures of quality and authority, both explicit and implicit. The ADSD is not a single service or software component. It is the synergistic combination of search interfaces, information models for describing content, and user-interface design.

5.1 Use Cases

The ADSD provides two key roles. Firstly it allows information discovery and then supports information publication. These two roles are outlined in the various use cases for ADSD (1-7) that are summarized below.

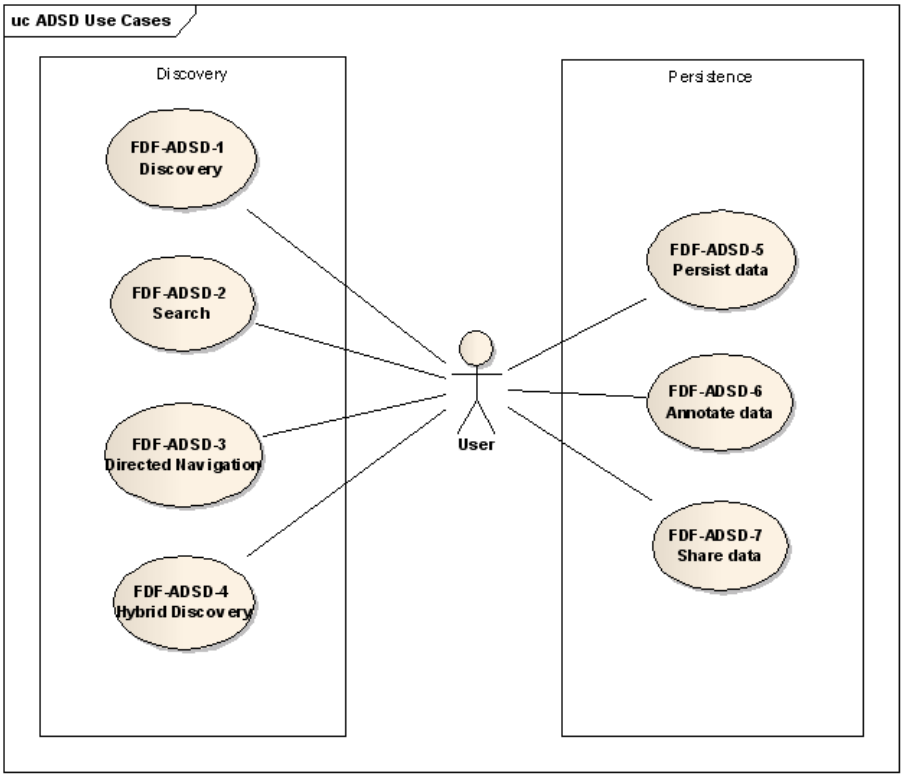


Figure 1: Overall Use Cases defined for ADSD

The detail of these use cases is given in the sections below.

5.1.1 Use Case: ADSD Discovery (FDF-ADSD-1)

Description: Describes how a user discovers an ADSD compliant catalog.

Goal: Obtain access to a "catalog" resource

Actors (Initiators): Human User, Email, Web search, etc.

Actors (Receivers): Human User

Pre-Conditions: None

Post-Conditions: Access to ADSD

System Components:

- ADSD service
- "Decision Fusion" client application

Alternate Courses of Action:

- User gets a URL to the ADSD offline (*e.g.* via email, etc.)
- ADSD access built into User's "Decision Fusion" client application

5.1.2 Use Case: ADSD Search (FDF-ADSD-2)

Description: Describes how a user browses metadata by searching on keywords, geography, and time (traditional OGC CSW search activity).

Goal: Find relevant data sources

Actors (Initiators): Human User Actors (Receivers): Human User

Pre-Conditions: FDF-ADSD-1

Post-Conditions: Client seeded with search results

System Components:

- ADSD service
- ADSD client (or integrated "Decision Fusion" client)

Basic Course of Action:

1. User enters search parameters in a GUI
   - keywords (tags)
   - area of interest (geographic bounding box)
   - temporal range (min date / max date)
2. GUI client submits the user's search request to ADSD

3. ADSD service discovers datasets from one or more registries (Controlled Artifacts) and more general URLs and these reference objects of various mime types.
4. ADSD service responds with results
5. The results include ordered listings of results based on measures of authority as well as capabilities for a client or user to discover related data sources.

5.1.3 Use Case: ADSD Directed Navigation (FDF-ADSD-3)

Description: Describes how a user browses metadata using pre-defined thematic relationships

Goal: Find relevant data sources

Actors (Initiators): Human User **Actors (Receivers):** Human User

Pre−Conditions: FDF-ADSD-1

Post-Conditions: Client seeded with search results

System Components:

- ADSD service
- ADSD client (or integrated "Decision Fusion" client)
- Taxonomy of data themes (provided via an ADSD service operation?)

Basic Course of Action:

- User requests theme list from ADSD service
- ADSD client sends request for themes
- ADSD service responds with themes
- ADSD client uses the response to construct a GUI that allows the user to do any of the following:
    - view all data sets for a particular theme
    - restrict the result set (drill-down) by selecting any of the following:
        - a sub-theme
        - an authoritativeness ranking
        - a specific publisher
        - a "community-sourced" usefulness ranking
    - change the result set by selecting a related theme

5.1.4 Use Case: ADSD Hybrid Discovery (FDF-ADSD-4)

Description: Describes how a user combines traditional search with directed navigation

Goal: Find relevant data sources

Actors (Initiators): Human User **Actors (Receivers):** Human User

Pre-Conditions: FDF-ADSD-1, FDF-ADSD-2, FDF-ADSD-3

Post-Conditions: Client seeded with search results

System Components:

- ADSD service
- ADSD client (or integrated "Decision Fusion" client)
- Taxonomy of data themes

Basic Course of Action:

- User enters search parameters in a GUI as in FDF-ADSD-2. However, as the user types each character, the ADSD client queries the ADSD service for matching themes and/or tags using OpenSearch Suggestions, and presents the user with a search completion list. This list changes with each character the user types.
- Alternatively, or in addition, the user may start with a directed navigation as in FDF-ADSD-3. In this case, the GUI should allow keyword/tag/geography/time search to be combined with the directed navigation at any point in the user's search activity.

5.1.5 Use Case: Persist Data (FDF-ADSD-5)

Description: User has one or more data sets in their workspace and now wants to save/persist this information.

Goal: The state of a geospatial-centric work session is described in an OWS Context document

Actors (Initiators): Human User **Actors (Receivers):** Human User

Pre-Conditions: TBD

Post-Conditions: TBD

System Components:

- ADSD client (Integrated Client)
- OWS Context document (OWC)

Basic Course of Action:

- Not included in OWS-7

5.1.6 Use Case: Annotate Data (FDF-ADSD-6)

Description: User has one or more data sets in their workspace and now wants to annotate this information with metadata such as authoritativeness, usefulness, thematic tags, lineage, quality, etc.

Goal: Rich metadata is associated with data service binding information

Actors (Initiators): / Actors (Receivers): Human User

Pre−Conditions: FDF-ADSD-5

Post−Conditions: Annotations are stored in the ADSD and available in subsequent rankings and queries.

System Components:

- ADSD client (Integrated Client)
- OWS Context document (OWC)

Basic Course of Action:

- Not included in OWS-7

5.1.7 Use Case: Share Data (FDF-ADSD-7)

Description: User has one or more data sets in their workspace and now wants to share this information.

Goal: Data resources can be published and become available in OpenSearch

Actors (Initiators)/(Receivers): Human User

Pre−Conditions: FDF-ADSD-6

Post−Conditions: TBD

System Components:

- ADSD client (Integrated Client)
- OWS Context document (OWC)
- ADSD service

Basic Course of Action:

- Not included in OWS-7

5.1.8 Use Case: OWC Map Annotation (FDF-OWC-3)

Description: Describes how a user can draw on a map and store that information's shape in geographic and paper space, and also attach textual information to the shape.

Goal: Store a map annotation

Actors (Initiators/Actors (Receivers): Human User

Pre/Post−Conditions: TBD

System Components:

- ADSD

Basic Course of Action:

- Not included in OWS-7

5.1.9 Use Case: Exploit Data (FDF-ADSD-8)

Success Scenario:

User takes search response and other data that didn't come through the search and produces a product related to this and persist it.

- The client evaluates the results are stored in an OGC Context ready for evaluation.
- Results are stored as a context document.
- The client reads the context document and evaluates the relevance of each one (probably without loading them all) in an evaluation process (e.g. user can look at the thumbnails).
- The client loads the data related to the context and exploits it.

New services can be published and become available in OpenSearch with geo/temporal/tagging/authority and sort by authority

The remainder of this report outlines the design and implementation of the components of an ADSD with close attention paid to the treatment of user-generated content. The information models used are described in the *Information Viewpoint* section of this document. The search/discovery interfaces are described in the *Computational Viewpoint* section. And their interactions are described in the *Engineering Viewpoint*.

6 Computational Viewpoint

6.1 Overview

The Computational Viewpoint is concerned with the functional decomposition of the system into a set of distinct components that interact at interfaces. For OGC, such components are generally realized as web services. With regards to the ADSD this section will focus on the overall use cases of a client discovering, searching, and parsing the results to a service that contains a list of authoritative documents or data sources. The processes of discovering a search interface, parsing the interface and performing a query, parsing the response and following links to data or service endpoints is discussed. In addition, the concept of publishing search results to a directory is presented.

For details on the encodings, models, and implementations of these computational viewpoints, see the corresponding information and engineering viewpoints.

The typical use case for this is shown in the figure below.

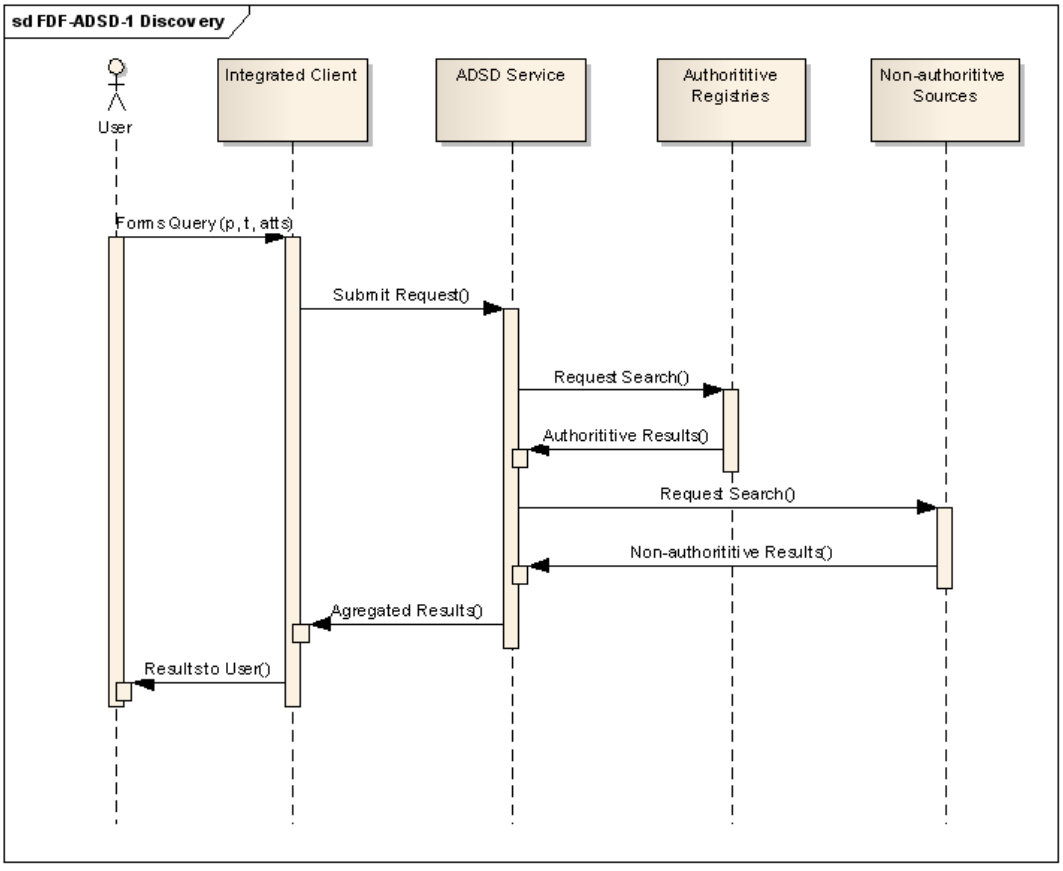


Figure 2: ADSD Interactions for Search

A User performs a query to a Client application or interface by specifying any of the following criteria: search term, categories, attribute ranges, geographic bounds, or authoritativeness.

Additional criteria could be supported by a client or subsequent services as provided but are beyond the scope of OWS-7.

The Client application formulates this combined query and requests processing from one or more ADSD servers. The query is formed as specified by the ADSD server following the proposed template interface described in the Information and Engineering viewpoints.

The Servers respond with a list of data sources that meet the specified Query criteria along with additional information on metrics of authoritativeness, quality, reliability, relevance and also additional parameters as supported by the server.

Additional Client query requests can be made to non-authoritative data source directories as might be available from publicly available web sites or search engines. These results are also formulated and returned with available result measurement criteria.

The Client application will aggregate these multiple results from the ADSD Servers and additional network services into an interface or set that is then viewable the User.

Details on the processes of internal search and sorting of Servers, or the aggregation of results in the Client are beyond the scope of OWS-7.

6.2 Interfaces

The current OGC Catalogue query interface is predicated on forming fairly formal queries using a specialist interface (reverse polish notation, and XML heavy syntax). The ADSD on the other hand needs a query interface that is both less heavily structured and is more accessible to other clients.

A more accessible and “open” search interface is considered an additional requirement rather than a replacement option. The interface chosen to experiment with this is the OpenSearch model that will be described in further detail in the Information and Engineering Viewpoints.

To meet the overall requirements, the ADSD has to have a permissive search interface, offering the ability to discover a range of content as opposed to very formalized content (i.e. it should not preclude delivery of both well defined, exploitable content and loose content only qualified by MIME-type). In addition it should support the publication of content, so that new content can be registered.

6.2.1 Query (OpenSearch ADSD)

Query of the ADSD is performed using OpenSearch. OpenSearch works by defining the service’s query API through a short XML document called an OpenSearch description document.

Search clients can use OpenSearch description documents to learn about the public interface of a search engine. These description documents contain parameterized URL templates that indicate how the search client should make search requests. Search engines can use the OpenSearch response elements to add search metadata to results in a variety of content formats.

Example of a simple OpenSearch description document:

```
<?xml version="1.0" encoding="UTF-8"?>
 <OpenSearchDescription xmlns="http://a9.com/-/spec/opensearch/1.1/">
   <ShortName>Web Search</ShortName>
   <Description>Use Example.com to search the Web.</Description>
   <Tags>example web</Tags>
   <Contact>admin@example.com</Contact>
   <Url type="application/rss+xml"
        template="http://example.com/?q={searchTerms}&amp;pw={startPage?}&amp;format=rss"/>
 </OpenSearchDescription>
```

6.2.2 Publish (CSW)

Publication and/or harvesting of resources was not deemed a requirement for the OWS-6 work on ADSDs. However, some participants used the CSW 2.0.2 publishing mechanism.

Publishing into a CSW 2.0.2 catalog can be done using two operations. These include the 'Transaction' operation and the 'Harvest' operation. The 'Transaction' operation is made up of three actions including 'Insert', 'Update', and 'Delete' which are used to create, modify and delete catalog records respectively.

The 'Insert' action acts as a container for the one of more records that are to be inserted into the catalog.

The 'Update' action contains one or more records with modified data that will be changed in the catalog.

The 'Delete' action identifies a set records that will be removed from the catalog.

The 'Harvest' operation involves the resolving a URI pointing to a metadata resource, parsing it, and finally creating or modifying metadata records in the catalog to register the resource.

6.3 Context Document Publishing

In addition to performing queries and aggregating and parsing results there is the need to publish context documents or search results from queries to additional ADSD servers. A context document contains a list of data services and sources with associated metadata. It should be possible to publish a Context Document to an ADSD as a means to register the data sources as well as provide additional metadata on quality based on annotations and usage of the source within the context document.

In OWS-7, Context Document publishing was done in an ad hoc, non-standardized way. Once the Context Document encoding format reaches OGC Standard status, the publishing of these documents in CSW catalogs should be addressed. However, CSW publication and/or harvesting is not a requirement for an ADSD. Since Context Documents are Atom resources, they can be easily discovered and indexed by mainstream search engines as long as they are made accessible on the Web.

6.4 Creating and Querying Annotations

As part of the Authoritative Data Source Directory, it is beneficial to provide mechanisms to allow client users to update and annotate data sources with quantitative and qualitative evaluations of authority and fitness for use.

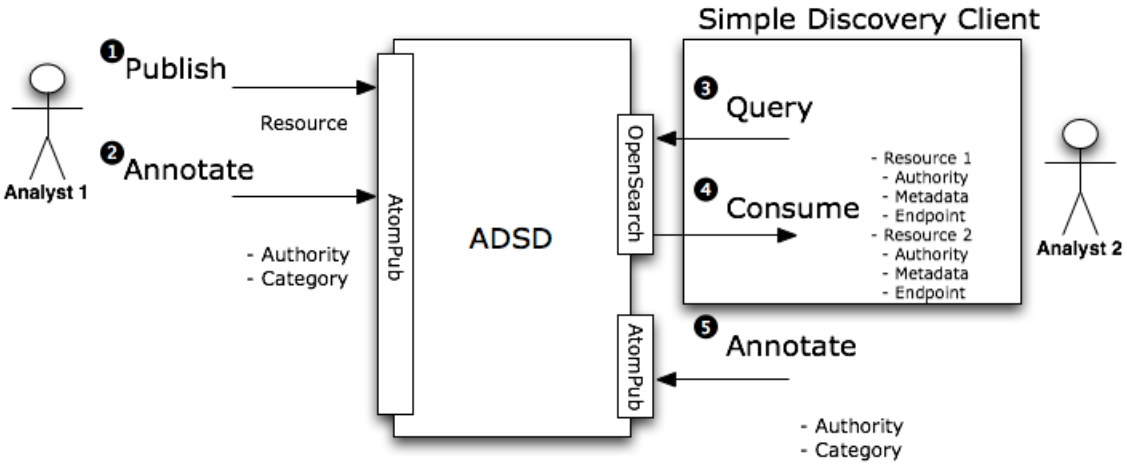


Figure 3: Creating and Querying Annotations of an ADSD

The content of annotations is outside the scope of OWS-7, but could consist of a free text note, or a type of rating such as accuracy, reliability, completeness, and relevance. These annotations would be stored in the ADSD and reference the specified data source. Subsequent calculations of authority and rank would incorporate these annotations based on internal calculations and provide ordered results based on these internal metrics and user requested filtering and sort orders.

In addition, a client consumer could optionally request the specific annotations. Based on specific ADSD restrictions of access to annotations and attributions, these annotations would be provided to clients as lists of elements.

In this way, organizations providing ADSD software can provide customers with a standard search interface—ADSD OpenSearch—while still providing competitive differentiation through the "smarts" of their internal ranking and rating systems.

7 Information Viewpoint

7.1 Overview

The Information Viewpoint considers the information models and encodings that will make up the content of the services and exchanges discussed in the Computational Viewpoint. With regards to the ADSD this includes the Query encoding in either OpenSearch for endpoint discovery and interface templating, or Catalog Services for the Web (CSW). In addition, this section describes the search response as either a general Atom response, Ref[3], or as a more specific OWS Context document.

As part of the OWS-7 FDF thread, extensions are suggested to Atom and OWS Context that provide for encoding Thematic Categories as well as measures of authoritativeness of documents and data. As an alternative encoding, JSON could be used following the Atom markup.

Finally, this section includes a crosswalk between the generalized Atom response and the Department of Defense Metadata Standard (DDMS), see Ref [1], as well as Dublin Core (DC), Ref[2], Ref[5].

7.2 Authority and Ranking

Authority and ranking covers a range of types, ranging from identification of reliability of an item, through to type classification (e.g. relevance to a given field) and through to open community-based questionnaire ranking, e.g. 'Rate how useful this dataset was to you'. The goal is to some degree to generate rankings automatically. A typical ranking might be based on how prevalent the given source is in all context documents. The concept being that if it is heavily used it has a high ranking. Ranking is expressed as:

- Category of the ranking
- Ranking classifications
- Ranking provenance

7.3 Categories, Code Lists, Tags and Controlled Vocabularies

An Authoritative Data Source Directory sorts datasets into one or more thematic categories. Categories can be prescribed, such as the NGA Human Terrain Entity Catalog Definition, or they can be organic and folksonomic in nature, and change as users create and apply new categories to the catalog.

The OWS-7 recommends utilizing the Atom <category /> tag to encode the thematic organization of datasets or features. The semantics of Atom already suggest that <category /> is a tag, or thematic organization, of the Atom entry. By use of the scheme attribute, the ADSD may specify a controlled vocabulary for the category, and a client can investigate the category's scheme, finding definitions, similar terms, or cross-walk through parent and sibling categories.

To request a code list, an HTTP GET is made to the data sources code list. This can vary between services depending on the types of categories or code lists that are employed. The typical response will be an Atom formatted XML document with an <entry /> for each category consisting of the following additional elements:

- title – the category name
- link@rel=self – the HTTP URI that links to the categories description

- link@rel=alternate – an HTTP URI that links to alternative format description of the category
- additional Atom elements as appropriate (updated, summary, et al.)

In addition, a category can itself belong to other categories. This provides for hierarchical and nested sets of categories. An example request may look like:

GET http://example.org/codelists/ows7DataSources

And the response would be:

```
<feed xmlns="http://www.w3.org/2005/Atom">
   <title>OWS7 Data Sources Categories</title>
   <link href="http://example.org/codelists/ows7DataSources" rel="self" />
   <id>urn:ogc:def:categories:ows7#DataCategories</id>
   <updated>2010-02-18T18:30:02Z</updated>

   <entry>
      <title>Communication / Media Preferences</title>
      <link rel="self" href="http://example.org/codelists/ows7DataSources/CommMedia" />
      <link rel="alternate" type="application/json"
href="http://example.org/codelists/ows7DataSources/CommMedia.json"/>
<id>urn:ogc:def:categories:ows7#CommunicationMediaPreferences</id>
      <updated>2010-02-18T18:30:02Z</updated>
   </entry>

   <entry>
      <title>Language</title>
      <link rel="self" href="http://example.org/codelists/ows7DataSources/LANG" />
      <link rel="alternate" type="application/json"
href="http://example.org/codelists/ows7DataSources/LANG.json"/>
      <category term="CommMedia" scheme="http://example.com/codelists/ows7DataSources"/>
      <id>urn:ogc:def:categories:ows7#Language</id>
      <updated>2010-02-18T18:30:02Z</updated>
   </entry>

   <entry>
      <title>Demographics</title>
      <link rel="self" href="http://example.org/codelists/ows7DataSources/DEM" />
      <link rel="alternate" type="application/json"
href="http://example.org/codelists/ows7DataSources/DEM.json"/>
      <id>urn:ogc:def:categories:ows7#Demographics</id>
      <updated>2010-02-18T18:30:02Z</updated>
   </entry>

   <entry>
      <title>Ethnicity</title>
      <link rel="self" href="http://example.org/codelists/ows7DataSources/ETH" />
      <link rel="alternate" type="application/json"
href="http://example.org/codelists/ows7DataSources/ETH.json"/>
      <category term="DEM" scheme="http://example.com/codelists/ows7DataSources"/>
      <id>urn:ogc:def:categories:ows7#Ethnicity</id>
      <updated>2010-02-18T18:30:02Z</updated>
   </entry>

   <entry>
      <title>Economy</title>
      <link rel="self" href="http://example.org/codelists/ows7DataSources/ECON" />
      <link rel="alternate" type="application/json"
href="http://example.org/codelists/ows7DataSources/ECON.json"/>
      <category term="ECON" scheme="http://example.com/codelists/ows7DataSources"/>
      <id>urn:ogc:def:categories:ows7#Economy</id>
      <updated>2010-02-18T18:30:02Z</updated>
   </entry>

</feed>
```

7.3.1 Categories in OpenSearch

In addition to utilizing the existing Atom category element, it is recommended that an equivalent OpenSearch description is added that allows for querying a catalog and specifying category or categories, and also for requesting the code list, or directory of categories.

OpenSearch provides for a simplified, discoverable templated query description to a search service that can implement any number of query specifications. By using a combination of OpenSearch and relevant extensions, ADSD's can be easily discovered and searched from a broad range of clients, as well as discover and consume data from upstream services. OpenSearch is described in depth on its web site http://opensearch.org.

Specifically, the Authoritative Data Source Directory needs to have the capability to convey the 'authoritativeness', or relative quality and utility, of data cataloged and hosted in the directory. This draft specification outlines an Atom encoding of a directory search including specifying the ranking algorithm to use in sorting and filtering the returned results.

The current version of the specification recommends the use of the Atom Ranking [http://tools.ietf.org/html/draft-snell-atompub-feed-index-10] draft extension. This extension provides for an encoding of enumerated ranking schemes, values, and human readable labels.

In addition to using the Atom Ranking extension, it is suggested that an equivalent OpenSearch extension encompassing the OGC OWS-7 ADSD recommendations is developed and provided to catalog and client developers.

7.3.2 OpenSearch Semantic extension

The Opensearch Semantic Extension (http://www.opensearch.org/Community/Proposal/Specifications/OpenSearch/Extensions/Semantic/1.0/Draft_1) offers a set of search operators that are suitable to search records that are classified with terms defined in authoritative dictionaries, or taxonomies. In particular, the "classifiedAs" operator can match.

In the scope of OWS-7 requirements , this extension has been extended to provide a way to retrieve taxonomies that are supported by the ADSD.

Therefore, adding this extension to the set of supported OpenSearch extensions is a possible solution to fulfill the usecase where the client first retrieves the set of supported authoritative dictionaries from the ADSD, and then requests all records classified by terms from these dictionaries.

7.3.3 Directory Atom Response

To demonstrate how the ADSD would encode a response, the following example highlights the query and response capabilities in searching for "Weather" data in an ADSD.

The resulting output is Atom entries for each data source with additional OpenSearch elements to include the query elements, pagination. The <category /> and <re:rank /> elements are used within an entry to provide category tags as well as relative rankings across two quantitative metrics: reliability and popularity. The actual calculation of these values are beyond the scope of the OWS-7 testbed, but each ADSD can provide sufficient metrics and methods as appropriate for the specific domain.


```
<feed xmlns="http://www.w3.org/2005/Atom" xmlns:re="http://purl.org/atompub/rank/1.0"
          xmlns:opensearch="http://a9.com/-/spec/opensearch/1.1/"
          xmlns:georss="http://www.georss.org/georss">
    <title>OWS7 ADSD Search by Category</title>
    <link href="http://example.org/search.atom?category=Weather" rel="self" type="application/atom+xml"
/>
    <id>urn:ogc:def:categories:ows7:search#Weather</id>
    <updated>2010-02-18T18:30:02Z</updated>
    <opensearch:totalResults>3200</opensearch:totalResults>
    <opensearch:startIndex>1</opensearch:startIndex>
    <opensearch:itemsPerPage>10</opensearch:itemsPerPage>
    <atom:link rel="search" type="application/opensearchdescription+xml"
href="http://example.com/opensearchdescription.xml"/>
    <opensearch:Query role="request" category="Weather" ogc:authority:ranking="reliability"
startPage="1" />

    <entry>
        <title>NOAA Cloud Cover</title>
        <guid>dataset:67</guid>
        <georss> ... </georss>
        <author>
            <name>USGS</name>
            <uri>http://weather.usgs.gov</uri>
        </author>
        <contributor>
            <name>Andrew</name>
            <uri>http://fortiusone.com</uri>
        </contributor>

        <!-- Theme definitions -->
        <category term="Weather" scheme="http://example.com/codelists/ows7DataSources"/>

        <!-- Authority Ranking -->
        <re:rank scheme="http://example.com/authority#reliability">5734</re:rank>
        <re:rank scheme="http://example.com/authority#popularity">27</re:rank>
    </entry>
</feed>
```

7.4 Suggested use of OpenSearch Ranking Extension

The previous example demonstrated how a simple term query returned ordered results based on rankings. The next step would be to provide an OpenSearch Extension in the ogc:authority namespace that allows for specifying the ranking algorithm as well as minimum and maximum filters on this ranking.

```
<Url type="application/atom+xml"
    template="http://example.com/search.atom?q={searchTerms}
        &ranking={ogc:authority:ranking}&minrated={ogc:authority:minranking}
        &maxrated={ogc:authority:maxranking}"/>
```

By providing a filter capability, ADSD clients can quickly discover and pivot through large sets of results to find relevant data sources based on necessary criteria.

1. "ogc:authority:ranking" - denotes the specific ranking algorithm to sort or filter by. If not minimum or maximum is included, catalog should return sorted by this ranking.

2. "ogc:authority:minranking" - denotes the minimum rating to filter using the ranking algorithm

3. "ogc:authority:maxranking" - denotes the maximum rating to filter using the ranking algorithm

Using the example from the previous section, an actual query may look like:

```
GET http://example.com/search.atom?q=Weather&ranking=reliability&minrated=5300
```

The result would be an Atom document with only sources that had a reliability rating higher than 5300. The ranking methods and specific quantitative measures are dependent on the implementation details of the ADSD and should follow standard organizational or industry metrics when possible.

7.5 Atom Categories and ebRIM classification schemes

As designed, an ADSD can utilize any schema of categories. In OWS-7, the *Human Terrain* classification scheme (see below) was created from research to be used for feature and statistical analysis (FSA). A sample of an ebRIM classification scheme as it is stored in the CSW registry looks like:

```
<rim:ClassificationScheme id="urn:ogc:def:categories:ows7:humanterrain" isInternal="true"
nodeType="urn:oasis:names:tc:ebxml-regrep:NodeType:UniqueCode" objectType="urn:oasis:names:tc:ebxml-
regrep:ObjectType:RegistryObject:ClassificationScheme" status="urn:oasis:names:tc:ebxml-
regrep:StatusType:Approved">
    <rim:Name>
        <rim:LocalizedString charset="ISO-8859-1" xml:lang="en" value="Human Terrain"/>
    </rim:Name>
    <rim:ClassificationNode code="REL" id="urn:ogc:def:categories:ows7:humanterrain:REL"
objectType="urn:oasis:names:tc:ebxml-regrep:ObjectType:RegistryObject:ClassificationNode"
parent="urn:ogc:def:categories:ows7:humanterrain" path="" status="urn:oasis:names:tc:ebxml-
regrep:StatusType:Approved">
        <rim:Name>
            <rim:LocalizedString charset="ISO-8859-1" xml:lang="en" value="Religion"/>
        </rim:Name>
         <rim:ClassificationNode code="ReligiousFacility"
id="urn:ogc:def:categories:ows7:humanterrain:REL:ReligiousFacility"
objectType="urn:oasis:names:tc:ebxml-regrep:ObjectType:RegistryObject:ClassificationNode"
parent="urn:ogc:def:categories:ows7:humanterrain:REL" path="" status="urn:oasis:names:tc:ebxml-
regrep:StatusType:Approved">
            <rim:Name>
                <rim:LocalizedString charset="ISO-8859-1" xml:lang="en" value="Religious Facility"/>
            </rim:Name>
            <rim:Description>
                 <rim:LocalizedString charset="ISO-8859-1" xml:lang="en" value="Facility where religious
events and practices take place."/>
            </rim:Description>
        </rim:ClassificationNode>
    </rim:ClassificationNode>
...
</rim:ClassificationScheme>
```

Other pre-loaded schemes include *Data Reliability*, *Source Reliability*, *Security Scheme and Restriction.*

These schemes will be made available from each catalog implementation in various formats. For example, to get the *Human Terrain* scheme back in *ebRIM* format users would execute something like:

```
GET http://ows-7.example.net/wes/serviceManagerCSW/csw?request=GetRecordById
      &id=urn:ogc:def:categories:ows7:humanterrain&service=CSW&elementSetName=full
      &outputSchema=urn:oasis:names:tc:ebxml-regrep:xsd:rim:3.0
```

To get the scheme back in *Atom* format users would execute something like:

```
GET http://ows-7.example.net/wes/opensearch/categories.atom
      &category=urn:ogc:def:categories:ows7:humanterrain
```


The basic idea here is that a user has one or more data sets / services they want publish. Using the ADSD client these schemes are displayed in such a way the user can *classify* their data with them. Using *OpenSearch Semantic* the user would then find these classifications as part of the search results. To do this OWS-7 recommends using the *category* tag.

Below is a sample Atom entry that utilizes this tag:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<feed xmlns="http://www.w3.org/2005/Atom" xmlns:opensearch="http://a9.com/-/spec/opensearch/1.1/"
xmlns:georss="http://www.georss.org/georss">
<title>CSW Search Results</title>
<updated>2010-04-20T00:00:00+00:00</updated>
<id>http://www.example.net</id>
<author>
<name>OGC Server</name>
</author>
<opensearch:totalResults>100</opensearch:totalResults>
<opensearch:itemsPerPage>100</opensearch:itemsPerPage>
<opensearch:startIndex>1</opensearch:startIndex>
<entry>
    <title>JRC Damage Assessment for Port-au-Prince, Haiti (200m grid cells)</title>
    <link rel="" href="http://Ows-7.example.net/wes/RecordSummaryPage.do?uuid=18C68D24-B05A-313B-12DF-
E2E72D11AA48&amp;recordLocale=en_US&amp;mode=mappable&amp;view=summary" />
    <link rel="related" href="http://finder.geocommons.com/overlays/20422.zip" title="downloadLink"/>
    <link rel="related" href="http://Ows-
7.example.net/wes/serviceManagerCSW/csw?request=GetRepositoryItem&amp;version=2.0.2&amp;service=CSW&amp;
id=18C68D24-B05A-313B-12DF-E2E72D11AA48" title="accessLink"/>
    <link rel="related" href="http://finder.geocommons.com/overlays/20422.csv" title="downloadLink"/>
    <category term="urn:ogc:def:categories:ows7:dataReliability:ProbablyTrue"
scheme="urn:ogc:def:categories:ows7:dataReliability" label="Probably True"/>
    <category term="urn:ogc:def:categories:ows7:humanterrain:Econ"
scheme="urn:ogc:def:categories:ows7:humanterrain" label="Economics"/>
    <category term="urn:ogc:def:categories:ows7:humanterrain:Econ:IndustrialBase"
scheme="urn:ogc:def:categories:ows7:humanterrain" label="Industrial Base"/>
    <category term="urn:ogc:def:categories:ows7:humanterrain:Lu"
scheme="urn:ogc:def:categories:ows7:humanterrain" label="Land Use Cover"/>
    <category term="urn:ogc:def:categories:ows7:securityClassification:Unclassified"
scheme="urn:ogc:def:categories:ows7:securityClassification" label="Unclassified"/>
    <category term="urn:ogc:def:categories:ows7:sourceReliability:Reliable"
scheme="urn:ogc:def:categories:ows7:sourceReliability" label="Reliable"/>
    <id>18C68D24-B05A-313B-12DF-E2E72D11AA48</id>
    <updated>2010-04-20T00:00:00+00:00</updated>
    <gml:LineString srsDimension="2" srsName="EPSG:4326" xmlns:gml="http://www.opengis.net/gml">
        <gml:posList>-74.792740983 20.183452923 -74.792740983 17.612662317 -71.69280421 17.612662317 -
71.69280421 20.183452923</gml:posList>
    </gml:LineString>
    <content type="text">
        Damage assessment based on visual comparison of pre- and post earthquake images, aggregated by
200m cells.
    </content>
</entry>...
```

The *category* tag above has three main attributes:

term: The "term" attribute is a string that identifies the category to which the entry belongs. In this case, the term is an identifier which is part of the scheme defined in the scheme attribute.

scheme: The "scheme" attribute is an URI that identifies a categorization scheme.

label: The "label" attribute provides a human-readable label for display in end-user applications.

OWS-7 also recommends that each *entry* above contain a link to DDMS metadata document so that a more detailed version of the metadata may be viewed.

For example:

```
...
<link rel="related" href="http://ows-
7.example.net/wes/serviceManagerCSW/csw?request=GetRecordById&amp;amp;id=12BFDC6E-8884-BC97-4BE6-
4318FC089455&amp;amp;service=CSW&amp;amp;elementSetName=full&amp;amp;outputSchema=http://metadata.dod.mil/mdr/ns/DDM
S/3.0" title="accessLink"/>
...
```

Notice how the *category* tag above resembles the *ddms:category* in the in the DDMS document below.

DDMS sample:

```
<!-- Methods for characterizing web service for discovery. Could include taxonomies used in the service
registry -->
<ddms:subjectCoverage>
    <ddms:Subject>
        <ddms:keyword ddms:value="XML Transformation"/>
        <ddms:keyword ddms:value="XSLT"/>
        <ddms:keyword ddms:value="data transformation"/>
        <ddms:keyword ddms:value="data fusion"/>
        <ddms:keyword ddms:value="mediation"/>
        <ddms:category ddms:qualifier="urn:ogc:def:categories:ows7:humanterrain"
ddms:code="urn:ogc:def:categories:ows7:humanterrain:DEM:Population:LifeExpectancyInYears"
ddms:label="Life Expectancy in Years" />
        <ddms:category ddms:qualifier="urn:ogc:def:categories:ows7:humanterrain"
ddms:code="urn:ogc:def:categories:ows7:humanterrain:MED:MedicalFacility:NumberOfMedicalStaff"
ddms:label="Number of Medical Staff"/>
        <ddms:category ddms:qualifier="urn:ogc:def:categories:ows7:humanterrain"
ddms:code="urn:ogc:def:categories:ows7:humanterrain:MED:Disease:DiseaseConditionsDescription"
ddms:label="Disease Conditions Description"/>
...
```

7.6 OWS Context Documents and the CSW Catalog

In OWS-7 various types of items should be able to be both published and discovered. These include services such as WMS (Web Map Service), WFS (Web Feature Service), WCS (Web Coverage Service), and SOS (Sensor Observation Service). As well as datasets such as KML/KMZ (Keyhole Markup Language), SLD (Styled-Layer Descriptor), and OWS Context Documents (OWC). Other types include web sites, books, images, and video. One of the main focuses of OWS-7 is to provide the ability to publish and discover OWS Context Documents (OWC).

7.7 Query/Response API for CSW and OpenSearch

Below is a sample CSW GetRecords request/response for an OWS Context Document stored in the registry.

7.7.1 CSW Request:

```
<?xml version="1.0" encoding="ISO-8859-1" standalone="no"?>
<csw:GetRecords xmlns:csw="http://www.opengis.net/cat/csw/2.0.2" xmlns:rim="urn:oasis:names:tc:ebxml-
regrep:xsd:rim:3.0" xmlns:ogc="http://www.opengis.net/ogc" xmlns:gml="http://www.opengis.net/gml"
service="CSW" version="2.0.2" resultType="results" startPosition="1" maxRecords="5"
outputFormat="application/xml" outputSchema="http://www.opengis.net/cat/csw/2.0.2"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.opengis.net/cat/csw/2.0.2 http://schemas.opengis.net/csw/2.0.2/CSW-
discovery.xsd">
    <csw:Query typeNames="csw:Record rim:Classification">
        <csw:ElementSetName typeNames='csw:Record'>full</csw:ElementSetName>
        <csw:Constraint version="1.1.0">
            <ogc:Filter>
                <ogc:And>
                    <ogc:PropertyIsEqualTo>
                        <ogc:PropertyName>csw:Record/dc:identifier</ogc:PropertyName>
                        <ogc:PropertyName>rim:Classification/@classifiedObject</ogc:PropertyName>
                    </ogc:PropertyIsEqualTo>
                    <ogc:PropertyIsEqualTo>
                        <ogc:PropertyName>rim:Classification/@classificationNode</ogc:PropertyName>
                        <ogc:Literal>urn:ogc:def:ebRIM-
ObjectType:OGC:Dataset:OWSContextDocument</ogc:Literal>
                    </ogc:PropertyIsEqualTo>
                    <ogc:PropertyIsLike wildCard="%" singleChar="_" escapeChar="\">
                        <ogc:PropertyName>AnyText</ogc:PropertyName>
                        <ogc:Literal>OWS Context%</ogc:Literal>
                    </ogc:PropertyIsLike>
                </ogc:And>
            </ogc:Filter>
        </csw:Constraint>
</csw:Query>
</csw:GetRecords>
```

7.7.2 CSW Response:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<csw:GetRecordsResponse xmlns:csw="http://www.opengis.net/cat/csw/2.0.2"
xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:dct="http://purl.org/dc/terms/"
xmlns:gml="http://www.opengis.net/gml" xmlns:ows="http://www.opengis.net/ows"
xmlns:rim="urn:oasis:names:tc:ebxml-regrep:xsd:rim:3.0" xmlns:wrs="http://www.opengis.net/cat/wrs/1.0"
xmlns:xlink="http://www.w3.org/1999/xlink" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:xml="http://www.w3.org/XML/1998/namespace"
xsi:schemaLocation="http://www.opengis.net/cat/csw/2.0.2 http://schemas.opengis.net/csw/2.0.2/CSW-
discovery.xsd">
    <csw:SearchStatus timestamp="2010-04-28T10:01:19"/>
    <csw:SearchResults elementSet="full" nextRecord="0" numberOfRecordsMatched="1"
numberOfRecordsReturned="1">
        <csw:Record>
            <dc:identifier>98E851DF-558F-2AF5-E4ED-A5AAE9324D8A</dc:identifier>
            <dc:title>OWC test</dc:title>
            <dc:type>urn:ogc:def:ebRIM-ObjectType:OGC:Dataset:OWSContextDocument</dc:type>
            <dc:subject>Data Resources: OWS Context Document (OWC)</dc:subject>
            <dc:subject>Data Resources: Dataset</dc:subject>
            <dct:modified>2010-04-22</dct:modified>
            <dct:abstract>OWS Context Example: GML Inline</dct:abstract>
            <dc:contributor>Open Geospatial Consortium, Inc. </dc:contributor>
            <dc:creator>Enter Affiliation Here</dc:creator>
            <dct:created>2010-04-22</dct:created>
            <dc:language>en</dc:language>
            <dct:references scheme="accessLink">http://Ows-
7.example.net/wes/serviceManagerCSW/csw?request=GetRepositoryItem&amp;version=2.0.2&amp;service=CSW&amp;
id=98E851DF-558F-2AF5-E4ED-A5AAE9324D8A</dct:references>
            <ows:BoundingBox crs="EPSG:4326">
                <ows:LowerCorner>18.366677977 -73.413137571</ows:LowerCorner>
                <ows:UpperCorner>18.487216361 -73.257454738</ows:UpperCorner>
            </ows:BoundingBox>
        </csw:Record>
    </csw:SearchResults>
</csw:GetRecordsResponse>
```

Notice how the 'dct:references' section contains a link to the actual OWC document that was uploaded and stored in the registry.

If we wanted to execute the same search using OpenSearch to get an Atom response we would make an Atom request like:

```
GET http://ows-7.example.net/wes/opensearch/search?q=OWSContext&type=atom
&classifiedAs=urn:ogc:def:ebRIM-ObjectType:OGC:Dataset:OWSContextDocument
```

Atom Response:


```
<?xml version="1.0" encoding="ISO-8859-1"?>
<feed xmlns="http://www.w3.org/2005/Atom" xmlns:opensearch="http://a9.com/-/spec/opensearch/1.1/"
xmlns:georss="http://www.georss.org/georss">
    <title>CSW Search Results</title>
    <updated>2010-04-22T00:00:00+00:00</updated>
    <id>http://www.example.net</id>
  <author>
        <name>OGC Server</name>
   </author>
    <opensearch:totalResults>1</opensearch:totalResults>
  <opensearch:itemsPerPage>100</opensearch:itemsPerPage>
    <opensearch:startIndex>1</opensearch:startIndex>
        <entry>
            <title>OWC test</title>
            <link rel="" href="http://Ows-7.example.net/wes/RecordSummaryPage.do?uuid=98E851DF-558F-
2AF5-E4ED-A5AAE9324D8A&amp;recordLocale=en_US&amp;mode=mappable&amp;view=summary" />
            <link rel="related" href="http://Ows-
7.example.net/wes/serviceManagerCSW/csw?request=GetRepositoryItem&amp;version=2.0.2&amp;service=CSW&amp;
id=98E851DF-558F-2AF5-E4ED-A5AAE9324D8A" title="accessLink"/>
            <id>98E851DF-558F-2AF5-E4ED-A5AAE9324D8A</id>
            <updated>2010-04-22T00:00:00+00:00</updated>
            <gml:LineString srsDimension="2" srsName="EPSG:4326" xmlns:gml="http://www.opengis.net/gml">
                <gml:posList>-73.413137571 18.487216361 -73.413137571 18.366677977 -73.257454738
18.366677977 -73.257454738 18.487216361</gml:posList>
            </gml:LineString>
                <content type="text">
                OWS Context Example: GML Inline
                </content>
        </entry>
</feed>
```

Notice that the link tag above provides the link to the actual OWC document that was uploaded and stored in the registry. Table 2 below shows a crosswalk between elements in Atom, DDMS, and Dublin Core.

Table 2: Atom Crosswalk with DDMS and DC

Atom Name	DDMS XML Tag Name	Dublin Core Name
title	Resource.title	title
link	Resource.relatedResources.RelatedResource.link	references
id	Resource.identifier	identifier
published	Resource.dates@posted	created
updated	Resource.dates@created Populated with the latest modification date.	modified
author	Resource.creator.Person	creator
contributor	Resource.contributor.Person	contributor
Content type	Resource.format.Media.mimType	type
Category	Resource.subjectCoverage.Subject.category	subject
Rights	Resource.rights	rights
Source	Resource.source	source
Summary	Resource.description	abstract

7.7.3 Role of the ADSD Results in Publication

The ADSD results will be returned as an Atom Response. That response could be thought of as a first draft OWS Context. The results should then be expanded into a more extensive OWS Context document.

7.7.4 Population of the ADSD

Population of the ADSD was agreed to be primarily a 'behind the scenes' activity within the OWS-7 testbed. The primary goal was to provide a better interface to identify authoritive data, and integrate the search with the process of exploitation and info-sharing. As such population was ad-hoc.

Population could be by open standards methods such as AtomPub or CSW-ebRIM Harvesting, or proprietary mechanisms. Population could also occur from either publication of Context Documents or identification of key sources accessed by context documents, but this area was also excluded from the OWS-7 Implementation.

There is a good potential through open submission of content to an ADSD that more data is found and determined to be relevant through usage and annotation. Common standards such as AtomPub allow for "bookmarking" and publishing relatively straightforward and encourage development of new clients that could be integrated into existing tools such as web browsers and desktop clients.

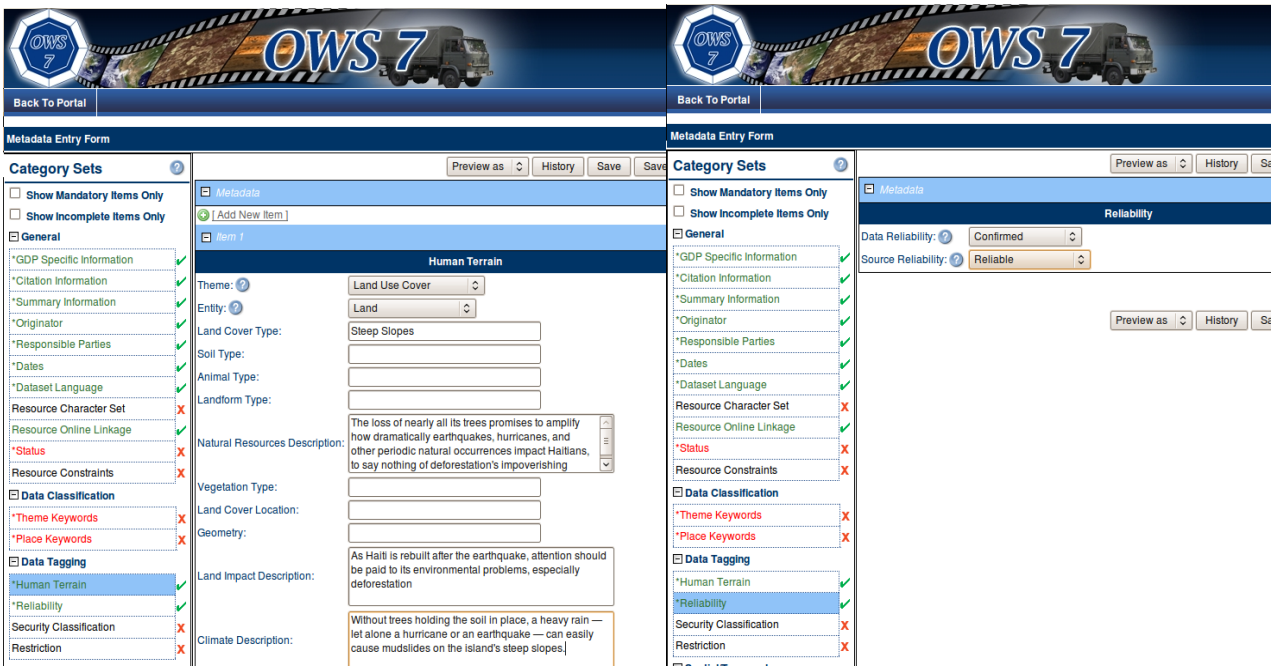


Figure 4: Compusult's Implementation of Human Terrain and Reliability Code Population

7.7.5 Reference Information

Information such as classification schemes/code-lists/type hierarchies need to be referenced by the ADSD. These may be stored in the registry which is supporting the ADSD but it may be appropriate to have these available as an independent service.

7.8 Summary

This section described the specific encodings of data source directories and rankings in Atom XML formats as exposed through OpenSearch-capable services. The code lists were referenced in ebRIM-XML as well as Atom entries to align with broadly accessible, existing tools. Future research should investigate the use of JSON and other encodings tht align with the Atom XML format encodings such that these other formats can be used by client services.

8 Engineering Viewpoint

The Engineering Viewpoint describes specific components linked by a communications network and is concerned with the interaction of distinct computational objects. This section describes the discovery of the OpenSearch Query endpoint of an ADSD server, formulation of the Query with an Atom response including authoritativeness measurements.

The more formal model, which is predicated on CSW (ebRIM), does allow clients with the capability to interact in a more precise way, relevant to more authoritative searches.

In addition there is a need to allow content to be passed back, and at present the premise is that this is relatively formal.

The overall model of the ADSD interactions as related to the use cases is shown below.

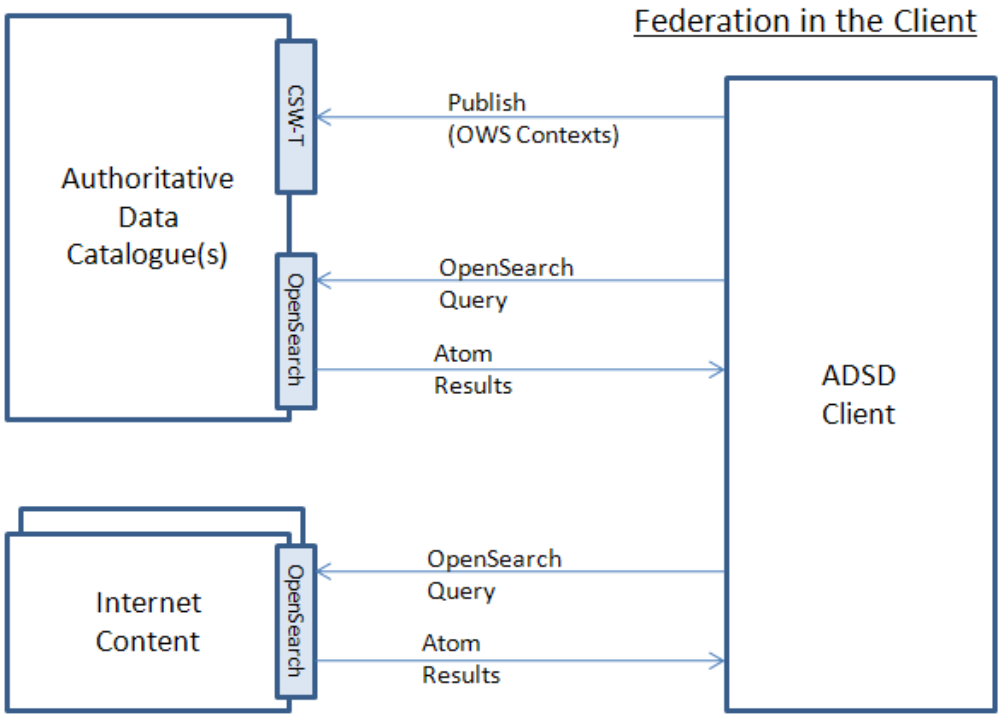


Figure 5: Search Architecture of an ADSD Catalog and Open Internet Sites

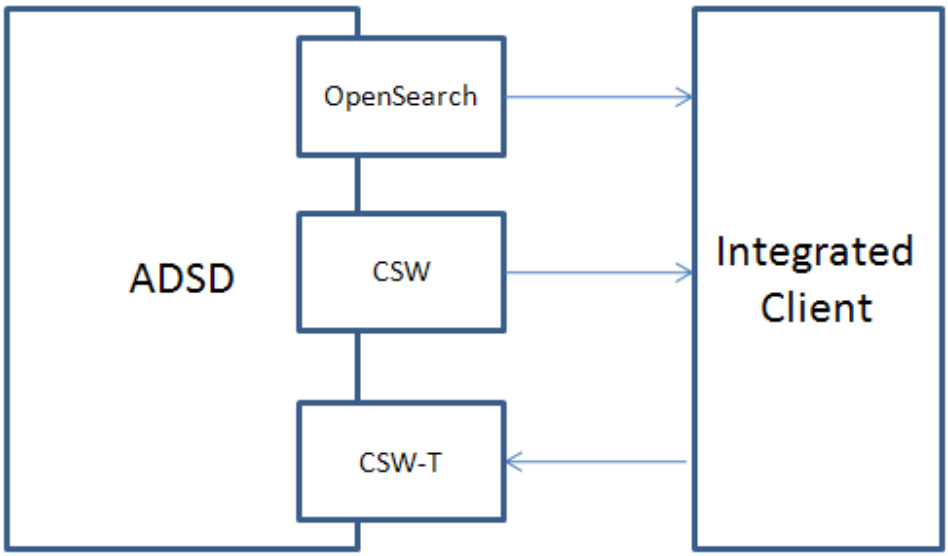


Figure 6: Overall Model of ADSD Interfaces

The arrows in the diagram show the predominant direction of dataflow, obviously transactions proceed in both directions. The transactional models for the above are shown below. The client potentially needs to deal with all of these, although the CSW and OpenSearch interfaces allow alternative clients to exploit the data even if they don't support both interfaces.

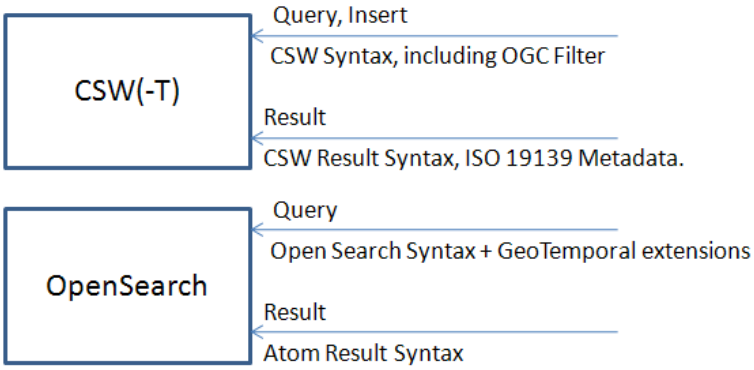


Figure 7: ADSD Interface Exchange Formats

A number of potential solution architectures exist to satisfy this interface. It is an open question as to whether to describe the ADSD as purely the above or whether to break this down into more components. It is possible that a single component (as is proposed in the OWS Technology Demonstration) could deliver the ADSD. And alternative view, with the ADSD being a broker not dependent on CSW is described below.

The premise of the diagram below is that the ADSD is orchestrated by a series of components, all complying with the OpenSearch interface, extended to support geo-temporal queries. The key issue is that while a CSW may be one way to deliver ADSD content it may be appropriate to integrate other services that are capable of delivering content.

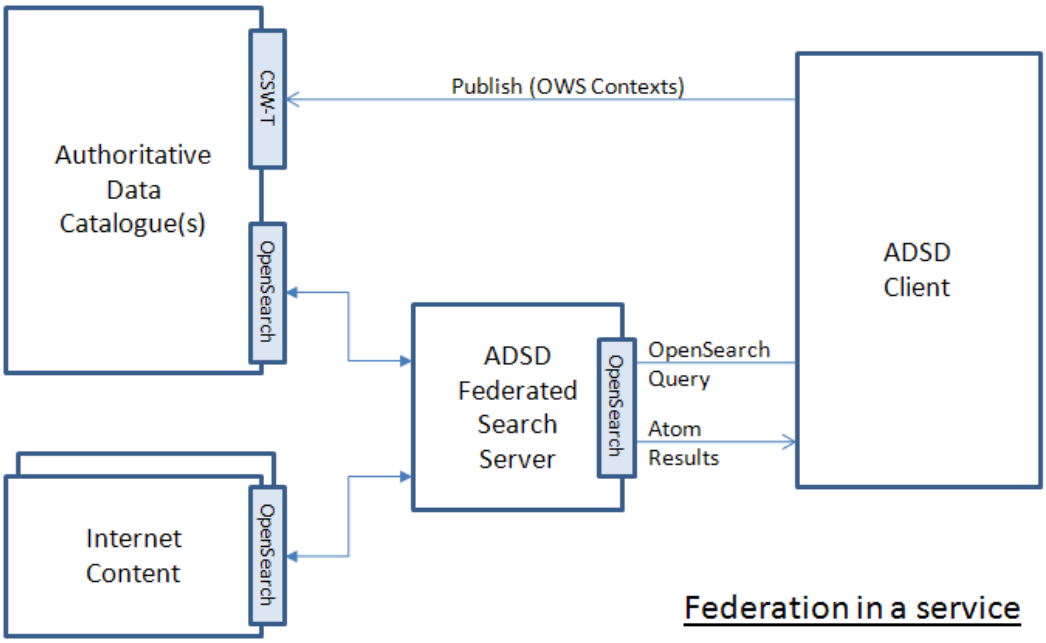


Figure 8: Component View of an ADSD Solution

In the OWS-7 test bed, these are the components that played a role in the ADSD architecture:

- Catalog Service for the Web
- OpenSearch
- OpenSearch extensions: Geo, Time, Suggestions, Ranking, Relevance, and Semantics
- An Integrated Client

8.1 OpenSearch Discovery

Reference the InfoSharing ER for specific details on discovering, searching, and results from an OpenSearch service.

8.2 CSW Catalogs

8.2.1 Storing OWS Context Documents (OWC) in the CSW Catalog

Storing the OWC document in the CSW registry requires using the CSW Transaction operation. For CSW implementors using ebRIM, OWS-7 recommends using an wrs:ExtrinsicObject to represent this document. The transaction below outlines what a sample Insert CSW transaction might look like.

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<csw:Transaction xmlns:csw="http://www.opengis.net/cat/csw/2.0.2"
xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:dct="http://purl.org/dc/terms/"
xmlns:gmd="http://www.isotc211.org/2005/gmd/" xmlns:gml="http://www.opengis.net/gml"
xmlns:ogc="http://www.opengis.net/ogc" xmlns:ows="http://www.opengis.net/ows"
xmlns:rim="urn:oasis:names:tc:ebxml-regrep:xsd:rim:3.0" xmlns:wrs="http://www.opengis.net/cat/wrs/1.0"
xmlns:xlink="http://www.w3.org/1999/xlink" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
service="CSW" version="2.0.2" xmlns:xml="http://www.w3.org/XML/1998/namespace"
xsi:schemaLocation="http://www.opengis.net/cat/csw/2.0.2 http://schemas.opengis.net/csw/2.0.2/CSW-
discovery.xsd urn:oasis:names:tc:ebxml-regrep:xsd:rim:3.0 http://docs.oasis-
open.org/regrep/v3.0/schema/rim.xsd http://www.opengis.net/cat/wrs/1.0
http://schemas.opengis.net/csw/2.0.2/profiles/ebrim/1.0/csw-ebrim.xsd">
    <csw:Insert>
        <wrs:ExtrinsicObject id="3C92068E-E4D8-D5A4-407A-856D199565B7" mimeType="text/xml"
objectType="urn:ogc:def:ebRIM-ObjectType:OGC:Dataset:OWSContextDocument">
            <rim:Slot name="http://purl.org/dc/terms/modified" slotType="urn:oasis:names:tc:ebxml-
regrep:DataType:DateTime">
                <wrs:ValueList>
                    <wrs:AnyValue>
                        <rim:LocalizedString charset="ISO-8859-1" xml:lang="en_US" value="2010-04-28"/>
                    </wrs:AnyValue>
                 </wrs:ValueList>
            </rim:Slot>
            <rim:Slot name="http://purl.org/dc/terms/created" slotType="urn:oasis:names:tc:ebxml-
regrep:DataType:DateTime">
                <wrs:ValueList>
                    <wrs:AnyValue>
                        <rim:LocalizedString charset="ISO-8859-1" xml:lang="en_US" value="2010-04-28"/>
                    </wrs:AnyValue>
                </wrs:ValueList>
            </rim:Slot>
            <rim:Slot name="http://purl.org/dc/elements/1.1/language"
slotType="urn:oasis:names:tc:ebxml-regrep:DataType:String">
                <wrs:ValueList>
                    <wrs:AnyValue>
                        <rim:LocalizedString charset="ISO-8859-1" xml:lang="en_US" value="en"/>
                    </wrs:AnyValue>
                </wrs:ValueList>
            </rim:Slot>
            <rim:Slot name="http://purl.org/dc/terms/dateSubmitted" slotType="urn:oasis:names:tc:ebxml-
regrep:DataType:DateTime">
                <wrs:ValueList>
                    <wrs:AnyValue>
                        <rim:LocalizedString charset="ISO-8859-1" xml:lang="en_US" value="2010-04-28"/>
                    </wrs:AnyValue>
                </wrs:ValueList>
            </rim:Slot>
            <rim:Slot name="http://purl.org/dc/elements/1.1/creator" slotType="urn:oasis:names:tc:ebxml-
regrep:DataType:String">
                <wrs:ValueList>
                    <wrs:AnyValue>
                        <rim:LocalizedString charset="ISO-8859-1" xml:lang="en_US" value="John Doe"/>
                    </wrs:AnyValue>
                </wrs:ValueList>
            </rim:Slot>
            <rim:Slot name="http://purl.org/dc/elements/1.1/contributor"
slotType="urn:oasis:names:tc:ebxml-regrep:DataType:String">
                <wrs:ValueList>
                    <wrs:AnyValue>
                        <rim:LocalizedString charset="ISO-8859-1" xml:lang="en_US" value="Open
Geospatial Consortium, Inc. "/>
                    </wrs:AnyValue>
                </wrs:ValueList>
            </rim:Slot>
            <rim:Slot name="http://purl.org/dc/terms/spatial" slotType="urn:ogc:def:dataType:ISO-
19107:GM_Envelope">
                <wrs:ValueList>
                    <wrs:AnyValue>
                        <gml:Envelope srsName="EPSG:4326">
                            <gml:lowerCorner>18.366677977 -73.413137571</gml:lowerCorner>
                            <gml:upperCorner>18.487216361 -73.257454738</gml:upperCorner>
                        </gml:Envelope>
```

```
                    </wrs:AnyValue>
                </wrs:ValueList>
            </rim:Slot>
            <rim:Name>
                <rim:LocalizedString charset="ISO-8859-1" xml:lang="en_US" value="OWS Context Example:
OGC Services"/>
            </rim:Name>
            <rim:Description>
                <rim:LocalizedString charset="ISO-8859-1" xml:lang="en_US" value="OWS Context Example:
OGC Services"/>
            </rim:Description>
            <wrs:repositoryItemRef xlink:href="http%3A%2F%2FOws-
7.example.net%2Fwes%2FserviceManagerCSW%2Fcsw%3Frequest%3DGetRepositoryItem%26id%3D3C92068E-E4D8-D5A4-
407A-856D199565B7"            xlink:type="simple"/>
        </wrs:ExtrinsicObject>
        <rim:Association associationType="urn:oasis:names:tc:ebxml-regrep:AssociationType:OffersService"
id="D8811E59-B9CE-5726-335B-17FEEC90D00F" sourceObject="1F21C457-DA59-1776-26B0-EBB8485116A2"
targetObject="3C92068E-E4D8-D5A4-407A-856D199565B7"/>
        <rim:ExternalLink externalURI="http://Ows-
7.example.net/wes/serviceManagerCSW/csw?request=GetRepositoryItem&amp;version=2.0.2&amp;service=CSW&amp;
id=3C92068E-E4D8-D5A4-407A-856D199565B7" id="12B5994E-92D8-86A4-8A20-23DC4A8C276B">
            <rim:Slot name="LinkageType" slotType="urn:oasis:names:tc:ebxml-regrep:DataType:String">
                <wrs:ValueList>
                    <wrs:AnyValue>
                        <rim:LocalizedString charset="ISO-8859-1" xml:lang="en_US" value="accessLink"/>
                    </wrs:AnyValue>
                </wrs:ValueList>
            </rim:Slot>
            <rim:Name>
                <rim:LocalizedString charset="ISO-8859-1" xml:lang="en" value="OWC OGC services.xml"/>
            </rim:Name>
        </rim:ExternalLink>
        <rim:Association associationType="urn:oasis:names:tc:ebxml-
regrep:AssociationType:ExternallyLinks" id="06369D00-FA11-4509-8B7F-AD463856FCF1"
sourceObject="12B5994E-92D8-86A4-8A20-23DC4A8C276B" targetObject="3C92068E-E4D8-D5A4-407A-
856D199565B7"/>
        <rim:Classification classificationNode="urn:ogc:def:ebRIM-ObjectType:OGC:Dataset"
classificationScheme="urn:cslt:def:DataResources" classifiedObject="3C92068E-E4D8-D5A4-407A-
856D199565B7" id="E329BA31-E336-143C-0974-C124E8524529"/>
        <rim:Classification classificationNode="urn:ogc:def:ebRIM-
ObjectType:OGC:Dataset:OWSContextDocument" classificationScheme="urn:cslt:def:DataResources"
classifiedObject="3C92068E-E4D8-D5A4-407A-856D199565B7" id="D9179EBA-F19F-BC1D-59E9-16E8756B5CB2"/>
    </csw:Insert>
</csw:Transaction>
```

The *wrs:ExtrinsicObject* extends the *rim:ExtrinsicObject* by adding the *repositoryItemRef* tag. This is used to store the URL of the actual OWC document which the object represents.

8.2.2 Defining the Required Elements

In the transaction above you will notice a number of slots defined. These slots are defined by the basic extension package for ebRIM and are directly related to the core queryables and presentables defined in the catalog services specification (see OGC 07-006r1). Notice how most of these Slots use the AnyValue element to accommodate multiple languages. Although this is not mandatory it is highly recommended.

Table 3 below shows the mappings between the various registry objects defined in the transaction and the common queryable and returnable elements defined in the OGC specification.

Table 3: ebRIM to OGC parameter cross-walk

ebRIM Object Element Name	Common queryable/returnable element	Type	Description
Extrinsic Object: id	Identifier	urn:oasis:names:tc:ebxml-regrep:DataType:String	An unique reference to the record within the catalogue
Extrinsic Object: objectType	Type	urn:oasis:names:tc:ebxml-regrep:DataType:String	The nature or genre of the content of the resource. Type can include general categories, genres or aggregation levelsofcontent.
Extrinsic Object: mimeType	Format	urn:oasis:names:tc:ebxml-regrep:DataType:String	The physical or digital manifestation of the resource
Extrinsic Object: Name	Title	urn:oasis:names:tc:ebxml-regrep:DataType:String	A name given to the resource
Extrinsic Object: Description	Abstract	urn:oasis:names:tc:ebxml-regrep:DataType:String	A summary of the content of the resource
Extrinsic Object Slot Name: http://purl.org/dc/elements/1.1/contributor	contributor	urn:oasis:names:tc:ebxml-regrep:DataType:String	An entity responsible for making contributions to the resource.
Extrinsic Object Slot Name: http://purl.org/dc/terms/spatial	BoundingBox	urn:oasis:names:tc:ebxml-regrep:DataType:ObjectRef	Spatial characteristics of the intellectual content of the resource.
Extrinsic Object Slot Name: http://purl.org/dc/elements/1.1/creator	creator	An entity primarily responsible for making the resource	urn:oasis:names:tc:ebxml-regrep:DataType:String
Extrinsic Object Slot Name: http://purl.org/dc/elements/1.1/date		urn:oasis:names:tc:ebxml-regrep:DataType:Date	A point or period of time associated with an event in the lifecycle of the resource.
Extrinsic Object Slot Name: http://purl.org/dc/terms/modified	Modified	urn:oasis:names:tc:ebxml-regrep:DataType:DateTime	Date on which the resource was changed.
Extrinsic Object Slot Name: http://purl.org/dc/elements/1.1/language	language	urn:oasis:names:tc:ebxml-regrep:DataType:String	A language of the resource.
Classification: classificationNode	subject	urn:oasis:names:tc:ebxml-regrep:DataType:String	The topic of the content of the resource

8.2.3 Defining the Required Classifications

The classifications defined in the transaction above are mandatory.

```
<rim:Classification classificationNode="urn:ogc:def:ebRIM-ObjectType:OGC:Dataset"
classificationScheme="urn:cslt:def:DataResources" classifiedObject="3C92068E-E4D8-D5A4-407A-
856D199565B7" id="E329BA31-E336-143C-0974-C124E8524529"/>

<rim:Classification classificationNode="urn:ogc:def:ebRIM-ObjectType:OGC:Dataset:OWSContextDocument"
classificationScheme="urn:cslt:def:DataResources" classifiedObject="3C92068E-E4D8-D5A4-407A-
856D199565B7" id="D9179EBA-F19F-BC1D-59E9-16E8756B5CB2"/>
```

These statements tell the registry that this object is a OWC context document and will permit it to be searched via the OpenSearch semantic extension.

9 Technology Viewpoint

This Technology viewpoint presents the current technology capable of delivering the ADSD Capability from Services through to client and in particular focuses on the implementation of the Testbed.

9.1 Scenario

The OWS-7 ADSD scenario is modeled on data and collaboration in response to the January 12, 2010 earthquake in Haiti. Following the immediate earthquake and in the days and weeks following numerous organizations gathered, registered, generated, and analyzed large amounts of geospatial data.

The workflow in this scenario highlights the use of multiple ADSDs as well as the general internet for an analyst to discover and access this data.

This viewpoint is intended to describe the state of current technology in being able to support the evolving ADSD standards. Within OWS-7 the primary evaluation of this is the demonstration infrastructure itself. This is shown in figure 1 below. The primary goal is to demonstrate how a client can interact with an OpenSearch service with Geo-support and exploit the results effectively.

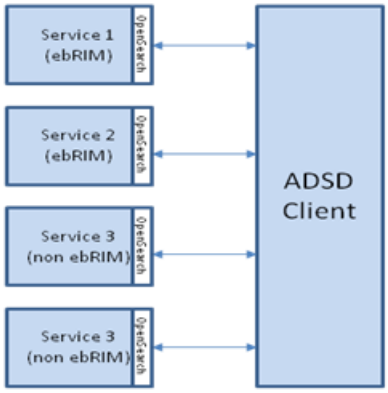


Figure 9: OWS-7 Technology Implementation

9.2 Technology Deployment

The demonstration in OWS-7 in most cases did not try to amalgamate the results from the various OpenSearch feeds, but just allowed the ADSD Client User to connect to each of them, return results to the user, and allow them to explore. One participant federated the OpenSearch query to ADSDs, YouTube, and Wikipedia.

The technology provided by different vendors to support this is as follows:

Table 4: ADSD Implementors

Technology Component	Implementor/Provider	Contribution
ADSD Client Application (OpenSearch)	Intergraph	Providing client capable of exploiting taxonomies in generating an OpenSearch Query and then displaying the results
ADSD Client Application (OpenSearch)	ESRI	Providing client capable of exploiting taxonomies in generating an OpenSearch Query and then displaying the results. Tested distributed search against OpenSeach endpoints, namely ERDAS and Compusult, YouTube and Wikipedia (included spatial constraints where applicable but not the classifiedAs).
OpenSearch Enabled ebRIM Registry with Publication	ESRI	CSW catalog containing the key taxonomies (e.g. human terrain)
OpenSearch Enabled ebRIM Registry with Publication	Compusult	ebRIM Registry containing the key taxonomies (e.g. human terrain)
OpenSearch Enabled ebRIM Registry with Publication	ERDAS	ebRIM Registry containing the key taxonomies (e.g. human terrain)
OpenSearch Enabled (non-ebRIM) Publish, Sharing, Discovery	FortiusOne	OpenSearch-enabled catalog and repository of open-source data. Geoprocessing for geojoin and geocoding.
Other (non-ebRIM) OpenSearch Services	Envitia	Gazetteer based general search component with semantic integration.

10 Summary Of Issues Identified/Further Work Suggestions

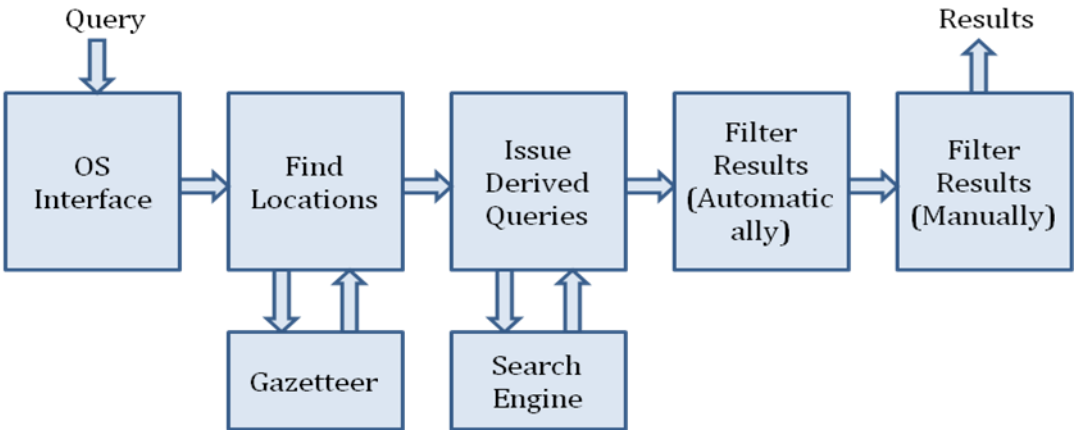
The ADSD sub-thread explored the ability of a decision support system to provide a unified environment for searching and reviewing all these different types of information resources through a single interface (while other sub-threads of the Feature and Decision Fusion thread of OWS-7 provided a unified environment for storing, and analyzing these information resources). For this work we adopted the OpenSearch metaphor for search, since it has become the most universal interface for searching content on the Web. If we wanted to have a search paradigm that was able to interface with all the content available on public and private networks, OpenSearch seemed the most promising technology available.

OpenSearch describes the search capabilities of a service, and we extended it to support some common functionality, such as the ability to constrain a search based on geographic and temporal parameters. This was quite straightforward, and has been done many times before. What was not straightforward was how we would bring the concept of "authoritativeness" to this mass-market search interface. Conceptually, we felt that authoritativeness is a combination of many factors. The most common factor is how much trust one has in the data source and the data source's publisher.

Organizations such as the Geoconnections, USGS, NGA, the UN, and Landgate to name a few, expend great amounts of time and effort to create comprehensive and updated data sets with excellent metadata. Any data set acquired from these organizations is almost certainly "authoritative". But one can rarely rely on organizations like these for all one's data needs. Identifying authoritativeness in data sets from less structured sources is more complex, and involves being able to account for authoritative concepts built up using "the wisdom of crowds" theory. OWS-7 was able to delve into these concepts in some depth, but there is still a wealth of research and prototyping experience required to fully explore the agenda set out here. This section identifies some areas that participants have identified for future work.

10.1 How to deal with non-authoritive sources (Non CSW)

Some experimentation was done in this area . A model that attempted to provide some structure to unstructured (and not explicitly geocoded) information. The concept was to take a general query (of the form 'Morbidity in Haiti' and translate 'in Haiti' into specific queries (using the gazetteer to identify relevant points in Haiti, e.g. Adminstrative areas and issuing each of these as specific queries). The results were associated with the points used to query. So the outcome of the OpenSearch ADSD request would be results that included the geo-tagging. Experimentation was done in delivering this in KML that allowed it to be easily displayed in a simple OpenLayers client. The model used is shown below (although steps 4 and 5 were not implemented and are really necessary to reduce the 'noise' in the results.



This search model hit two specific services (a gazetteer and a public search engine.

While not full functional this demonstrated that an approach like this might offer valuable input to the ADSD and results in an 'authoritative' document which references useful non-authoritive sources. These sources are only vaguely related to a map but offer potentially information that can be exploited. These can be easily accessed via an OpenSearch Query and returned via an Atom result set (and potentially as a context document.

10.2 Community based Ranking

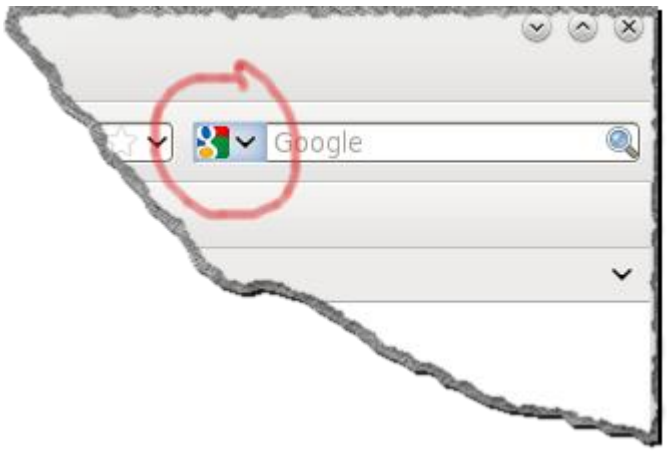
One area considered was community based ranking (potentially derived from either a well-defined community, from explicit inputs (*e.g.* Amazon or eBay rankings) or from examining the number of context documents a dataset was used in. While no specific experimentation was done here, the general approach is believed to be valid and useful and can easily be accommodated in both OpenSearch/Atom and ebRIM.

10.3 Population of ADSD using open standards

The issue of population of ADSD and harvesting was considered but was not a major focus in OWS-7. In the end all population was undertaken by proprietary or existing methods, *e.g.* CSW harvesting. And the general view was that at present there is no specific need to address this in OpenSearch. A viable option would be to use the Atom Publishing Protocol, AtomPub, in order to provide easy discovery and interfaces for populating the ADSD as well as annotations and updates. However this area might be considered for a future OWS Testbed.

10.4 Delivering to truly general clients (web browsers)

OpenSearch does allow users to access a specific search model from the displayed web site using the browser.



The experimentation in OWS-7 didn't try to experiment with this as the primary interface was an integrated client. The question as to how easily the ADSD service can be used in 'generic' clients (such as browsers) may ultimately be the test of inter-operability. This is a topic for a future OWS Testbed.

10.5 Inclusion of semantic interface options (interoperability)

The idea of semantic extensions to OpenSearch was discussed. As with some of the other approaches the general conclusion was that these extensions might simply overload the OpenSearch interface with complexity that is unnecessary. A concept that is considered relevant is to use semantic information (available in say CSW or another underlying support paradigm) to support providing hints to the user (related searches etc). This seems to offer the most benefit. This requires further experimentation.

10.6 Publishing Context Documents into ADSD and their 'Harvesting' for sources

Context Documents can be considered as search result outputs, either from a response from an ADSD or generated from a client application. There are potentially a number of ways a Context could be published. These are:

1. Addition using a proprietary technology to add them to the catalogue (as was accepted for other objects in OWS-7).
2. Add them using the transactional CSW (CSW-T) interface. This is the approach assumed in figure 7.
3. Add them by using AtomPub. This would mean the ADSD would need to support this additional interface.
4. Simply publishing them to an end-point via HTTP POST.

The primary merit of the approach shown in figure 7 (option 2 above) is that publishing to the registry allows firstly the context to be version controlled and authoritively referenced. It also allows the registry to trigger harvest operations on contexts that are added. This allows them to be automatically published and also links in the context to be extracted and used to rank the entries they reference (the approach of ranking based on usage). The merits of the other options have not been assessed and do require consideration in future work.

10.7 The merits of CSW versus OpenSearch and Why both are needed.

The general conclusion in OWS-7 was that these two models offer distinct advantages and have distinctly different use cases.

- CSW (particularly supported by the ebRIM model, although also to some degree relevant to CSW-ISO) provides a formalized repository and highly structured querying on formalized metadata. This is particularly important and relevant to machine access (i.e. the exploiter is automated).

- OpenSearch on the other hand is much more flexible in providing templated query interfaces. As a result it is particularly appropriate to human interactions.

A further conclusion was that mixing up these roles may actually result in a poorer overall result (something that doesn't do either job well. For example loading OpenSearch with well-defined formalized search patterns seems likely to simply obfuscate the interface. On the other hand making CSW 'loose' is not necessarily helpful either. Several implementations in OWS-7 indicated that it is possible to deliver an OpenSearch interface from an ebRIM CSW registry/repository and so a specific choice is not needed (apart from possibly considering how this mapping can be formalized and whether an OpenSearch interface is made mandatory or recommended as an interface to a registry/repository.

11 Annex A: Human Terrain Theme Entity Catalog Definitions

Table 5: Human Terrain Theme Entity Catalog Definitions (DRAFT)

Theme	Entity Code	Entity Name	Scale Needed	Scale likely available	Available Internally?	Source Suggested the Attribute is Needed	Likely Source of Data	Constraints and/or Notes	ID
		Scale Note: all DEM data requested Per capita per km^2 for built-up areas				A=Academic Literature Review,B=exemplar Binder,K=285+,O=SME review meetings,I=Pilot Interviews (Phase-2),C=Concept Map,S=NGA SME,IP=IPB			
Demo–graphics	DEM (1.0)	Birth Rate	R	G	N	A,O,C	Factbook		1.1
		Mortality Rate	R	G	N	A,O,C	Factbook		1.2
		Rate of Natural Increase	R	G	N	A,O	Factbook	Derived	1.3
		Carrying Capacity	R	G	N	A	?	Derived from Survey	1.4
		Demographic Transition Phase	G	G	N	A	Derived from Census IDB	Derived	1.5
		Life Expectancy	G	G	N	A,O,C	Factbook	CIA Factbook	1.6
		Emigration Rate	G	G	N	A,O,C	Factbook	Can be derived from Census	1.7
		Immigration Rate	G	G	N	A,O,C	Factbook	Can be derived from Census	1.8
		Population Density	R,L	L	N	A,O,I,C	Landscan	Landscan or Census	1.9
		Racial composition	L	G	N	A,O,I,C	Census	Need Survey Data or Census	1.10
		Urban/Rural population	L	G	N	I	Census or Img analysis	Could be derived from Population Density	1.11
		Gender	G	G	N	A	Factbook	Need Census	1.12
		Birth Control practices	R	R	N	O	?	Need Survey Data	1.13
		Age Distribution (cohorts)	L	G	N	B,K,O,I	Factbook, Census	Need Census	1.14
		Census Data	R,L	Varies	N	B,O, I	Individ. Country		1.15
		Population Movement - temporary	R,L	R,L	N	B,O	UN, Academia		1.16
		Population Movement - seasonal	R,L	R,L	N	B,O	Academia, othr studies		1.17

Theme	Entity Code	Entity Name	Scale Needed	Scale likely available	Available Internally?	Source Suggested the Attribute is Needed	Likely Source of Data	Constraints and/or Notes	ID
		Population Movement - refugee	R,L,S	R,L	N	B,O	UN		1.18
Religion	REL (2.0)	Religious composition	R	R	N	A,K,O,I,C	?		2.1
		Sacred/holy locations (includes places of worship)	S	L	Y (partial)	A,K,O,I,C	?		2.2
		Holidays, ceremonies, services, celebrations (timeline)	G,R,L	G	?	A,K,O,C,G	?		2.3
		Taboos and customs	G,R,L	G	N	K,O	?		2.4
		Religious history (includes conflicts)	G,R		N	K,O,P	?		2.5
		Religious laws	G	G	N	K,O	?		2.6
Language	LANG (3.0)	ethno-linguistic groups,	R	G	?	A,C,S	Ethnologue	mannerisms (unspoken language)	3.1
		Linguistics groupings	R	G	?	A,C,S	Ethnologue		3.2
		Number of speakers	G,R	G	N	A,C,S	Surveys		3.3
		Languages used for: education, broadcast, publications, business, religion, social activities, political)	?	L,S - limited	N	A,C,S	Research - multiple sources		3.4
		Dialects, slang	G,R,L	R	N	A,C,S	Research - multiple sources		3.5
		Dual languages (primary secondary)	G	G	N	A,C,S	Research - multiple sources		3.6
		Official Language, date made official	G	G	N	S	Factbook, or…?		3.7
Ethnicity	ETH (4.0)	Ethnic, tribal, racial groups	L	?	N	S	Surveys		4.1
		Gender-based Practices of Ethnic Group	L	?	N	S	Studies		4.2
		Leadership, relationships, power structure, council of elders	R	?	N	B,K,O,I,IP	Studies		4.3
		Religion of Clan	R	?	N	S	Studies		4.4
		Youth-elder relationships & social norms for these relationships	R	?	N	S	Studies		4.5

Theme	Entity Code	Entity Name	Scale Needed	Scale likely available	Available Internally?	Source Suggested the Attribute is Needed	Likely Source of Data	Constraints and/or Notes	ID
		Descent System - (patrilineal, matrilineal) needed to track lineage of leaders	R	?	N	S	Studies		4.6
		Property Ownership & Inheritance System	R,L	?	N	S	Surveys		4.7
		Influential personalities (storytellers, healers)	R,L	?	N	S	Surveys		4.8
		Prejudices or characterizations of others (inferiority, outside actors, supervisory)	R	?	?	S	Studies		4.9
		Rivalries & alliances over time	R	?	N	S	Studies		4.10
		Ethnic and Tribal relationships	G,R	?	N	S	Studies		4.11
		Colonial Legacy (historical alliances with others & external powers)	R	?	N	A,S,O	Studies		4.12
		Customs & Traditions	R,L	?	?	S	Studies		4.13
		Laws, Rules, Governance	R,L	?	N	A,O,IP	UN, othr sources		4.14
		Livelihood, housing & terrain characteristics where appropriate (also found in economy)	R,L	?	N	A,B,K,O,I,S	Census if avail.		4.15
		~~Racial or ethnic composition where appropriate~~	~~G,R~~	?	~~N~~		~~Surveys~~	(already captured in 1st attribute)	4.16
		Acculturation and assimilation	R,L	?	N	A,O,S	Surveys		4.17
Economy	ECON (5.0)	Financial Base (banking, currency, GNP, GDP, Inflation, Remittances)	G	G	N	K,O,I,C	Factbook, other sources for remittances		5.1
		trading centers/financial institutions	R,L,S	urban centers only	Y - partial	K,O			5.2
		Standards of Living (Distribution of wealth, poverty & land ownership)	G,R,L	G,R	N	A,B,K,O,I,IP			5.3
		Goods & Services (tourism, imports, exports, prices)	G,R,L	G,R	N	A,K,O,C			5.4
		Energy / electricity	G,R	G	Y - partial	A,B,K,O,I,C	derived from MS		5.5

Theme	Entity Code	Entity Name	Scale Needed	Scale likely available	Available Internally?	Source Suggested the Attribute is Needed	Likely Source of Data	Constraints and/or Notes	ID
		Workforce characterization(wages, skills, occupations, livelihoods incl. agriculture/animal husbandry)	R,L,S	G,R	N	A,K,O,I,C			5.6
		Industrial base (raw materials, processing, distribution, construction)	G,R,L	G,R,L	N	A,O,C	research		5.7
		Markets (formal, black, informal, physical & non-physical)	G,R	G	N	A,K,O,C,IP			5.8
Education	EDU (6.0)	school levels (preschool, primary, secondary, university)			N	A,K,O,I			
		literacy			N	A,B,K,O,I	Factbook, derived		
		school locations			Y (partial)	A,K,O,I	Multiple Sources		
		school information			Y (partial)	A,K,O,I			
		School types (trade, religious, formal, informal)			N	A,K,O,I			
		school curricula			N	K,O			
		school-level attributes			N	K,O			
		education by gender			N	A,K,O			
		% illiterate by province			N	B,O,I			
		school enrollmnt as % of age group			N	B,O,I			
Land Use / Cover	LU (7.0)	Vegetation (types - illicit or agricultural, indexes, distribution)				A,B,K,O,I,IP			
		Biomes				A,O			
		seasonal impact & weather impacts on terrain				A,O,IP			
		soils				A,O,			
		climate				A,O,I,IP			
		landforms				A,O,I			

Theme	Entity Code	Entity Name	Scale Needed	Scale likely available	Available Internally?	Source Suggested the Attribute is Needed	Likely Source of Data	Constraints and/or Notes	ID
		natural resources				A,O,I			
		hydrography / hydrology (floods, water security,irrigation)				A,B,O,I,IP			
		agriculture & crops				A,K,O			
		Terrain (elevation, line of sight)				A,B,O,I,IP			
		Fauna				K			
Medical / Health	MED (8.0)	Disease conditions, incidence, and locations (endemic regions)							
		Disease vectors and spread (human, animal, insect, and plant)							
		Disease due to sanitation							
		Disease due to malnutrition							
		environmental hazards (pollution, industrial waste sites)							
		Medical facility attributes (beds, op rms, docs, staff, condition of facility.)							
		Docs per capita, skill levels, distribution							
		Health workers & skills (incl. nurses)							
		Access to health care							
		Unmet Basic Needs Index							
		Health beliefs							
		Diet & Nutrition							
		Medical facility locations							
Political / Ideology	POL (9.0)	Voting patterns or behavior	R,L	G	N	K,O,I,C,IP	Open Sources		
		Civic / Welfare / Political / Ideological Groups (NGO's, relief ops, religious, illegal or any other structured group)	G,R,L,S	R,L,S	Y (limited)	K,O,I,C,IP	Open Sources		
		Administrative Boundaries & regions (political units, cities, populated places)	G,R,L,S	G,R,L	Y	A,B,K,O,I,C,S,IP			
		~~Political Groups (redundant)~~	~~G,R,L,S~~	~~G,R~~					

Theme	Entity Code	Entity Name	Scale Needed	Scale likely available	Available Internally?	Source Suggested the Attribute is Needed	Likely Source of Data	Constraints and/or Notes	ID
		Borders (actual, perceived, crossings, checkpoints)	G,R,L,S	G,R,L	Y	K,O,C,IP			
		Government Method (elected, appointed, etc.)	G,R	G	N	A,B,K,O,I,C,S,IP	Factbook, Country Studies		
		government structures & facilities incl police, fire, rescue	L,S	L,S	Y				
		Type of Nation-State	G	G	N				
		Leadership	G,R,L,S	G,R,L	Y				
		Judicial System - rule of law	G,R	G	N				
		Military / Police	G,R,L	G,R	Y				
		Political / Ideological perceptions & attitudes			N				
		Political stability & instability	G,R	G,R	?				
Communications / Media	COM (10.0)								
		Media influence & control							
		Press freedom							
		cellular communications							
		print media communications							
		Communication coverage areas (Sat, TV, cell, emergency, radio, backbones, cable & fibre networks)							
		Media type & distribution (including snail mail)							
		Oral Traditions or Word-of-mouth							
		Social networks (groups, associations, impacts)							
		Internet access and coverage, providers							
		% access to Radios							
		% households with at least 1 phone							
		% households with at least 1 radio							

Theme	Entity Code	Entity Name	Scale Needed	Scale likely available	Available Internally?	Source Suggested the Attribute is Needed	Likely Source of Data	Constraints and/or Notes	ID
Significant Events	EVNT (11.0)	conflicts: (ethnic, religious, boundary, political)							
		History/timeline of events							
		Natural Disasters							
		Violent events							
		Humanitarian disasters							
Transportation	TRANS (12.0)	road network by type	R,L,S	R,L	Y	K,O,I,IP			
		rail network	R	R	Y	K,O,I,IP			
		waterways & ports	R,L	R,L	Y	K,O,I,IP			
		airports	G,R,L	G,R,L	Y	K,O,I,IP			
		vehicles driven (characteristics)							
		pipelines							
		forms of transportation (animal, vehicle, rental)							
		Road Network Access							
		bus depots							
		bridges							
		aircraft							
		navigation infrastructure							