Analysis of Options and Recommendations for Representing and Exchanging Geospatial Data on the Exchange Network

1. Background

EPA and States have identified various issues associated with the use and exchange of geospatial data on the Exchange Network. In late 2006, the Exchange Network Leadership Council (ENLC) convened an Exchange Network Geospatial Strategy Team (see Appendix A for membership) to examine the following issues:

- **What are the best options for encoding geospatial data for exchanges on the Exchange Network?**
- **What are the challenges and options associated with multiple coordinate reference systems used by EPA, States, and others to represent geospatial data of interest in exchanges on the Exchange Network?**

This paper is the result of Team deliberations. The Team consisted of representatives not only interested in the Exchange Network, but also those involved in development of other networks such as the Environmental Public Health Tracking Network (Tracking Network) and the National Spatial Data Infrastructure (NSDI). The Team recognized that solutions identified for the Exchange Network could also be applied to other network efforts.

The Team held several calls, convened for a face-to-face two day meeting in December 2006, developed recommendations, and reviewed drafts of papers. Many different aspects of the issues were identified, as well as options for addressing them. The following recommendations are offered by the Team to the Exchange Network Leadership Council (ENLC) for consideration. The remainder of this paper provides the documentation that led to the recommendations.

2. Team Recommendations

Recommendation #1: Adopt the Geography Markup Language (GML) as the preferred format for representing geospatial data in Exchange Network schema.

The Geography Markup Language (GML) is a widely adopted XML vocabulary (schema) now in its third major version. It is a formally recognized Open Geospatial Consortium (OGC) and ISO standard. Various Exchange Network projects have been experimenting with variations on its use in exchanges for the last several years. It is currently seen to be the best non-proprietary XML format for representing geospatial data. GML is a rich, complex language that is capable of expressing complex geometries and features and includes the ability to work with multiple coordinate reference systems (latitude/longitude and Universal Transverse Mercator - UTM).

While endorsing the use of GML, the Team recognizes that some technical issues remain to be addressed including:

- The appropriate version of GML to adopt. Version 3.1 is most current, but not yet supported (as of May 2007) by ESRI tools. ESRI plans to include GML support in ArcGIS version 9.3. This may require the Exchange Network to phase the implementation of GML to ensure adequate vendor support.
- The specific representations of GML to select. GML allows multiple ways to model the same information, therefore specific guidance on GML practices may be needed to encourage more compatibility within and across Exchange Network flow areas.

Recommendation #2: Explore the use and adoption of GeoRSS GML (a specific sub-set of GML) where expression of simple geospatial data in Exchange Network flows is required.

GML is a large, complex standard that can express much more complex geospatial information than exist/will exist in most current and future Exchange Network flows. To encourage broad adoption, and facilitate the ability to “mash up” (easily integrate) data and applications on-the-fly, the GML community has developed GeoRSS GML. This is a formally defined subset (“profile”) of GML. It is, in effect, GML “lite”. GeoRSS GML is a new standard under development by the Open Geospatial Consortium (OGC), but is being rapidly adopted by Google, Yahoo and other vendors. Adoption of full GML by these vendors had lagged due to its complexity. Expressing Exchange Network data in GeoRSS GML creates the opportunity to easily leverage the tools of these vendors.

GeoRSS GML was created expressly to simplify GML adoption for the typical simple features encountered in most Exchange Network exchanges. GeoRSS GML was so named because its initial intended use was to geo-enable RSS (news) feeds. Use of the GeoRSS GML profile in no way commits or precludes the Exchange Network’s use of RSS technology. The name is an historical artifact.

GeoRSS GML is an important alternative to GML that provides means to express simple geographic features (points, lines, and simple polygons) with less overhead than is inherent in GML. It does not support topology, multiple-interacting polygons, or complex features. But it will contribute to the ability to develop web-services for geospatial data and supports the ability to express data in multiple coordinate reference systems.

Several technical issues require further consideration:

- Validating the use-case for GeoRSS GML.
- Providing clear guidance on when GML rather than GeoRSS GML should be used.
- Confirming the inter-operability of GeoRSS GML and GML in commonly used tools.

Recommendation #3: Establish a policy and guidelines that support the ability to simultaneously represent multiple coordinate reference systems within Exchange Network flows.

EPA has adopted a latitude/longitude coordinate standard and expects States and others to submit data to the Agency in compliance with this standard. Current Exchange Network shared schema components (SSC) allow only latitude/longitude formatted locational information. Some of the EPA databases only support latitude/longitude (e.g., Facility Registry System - FRS), while others can accommodate other coordinates such at UTM (e.g., Air Quality System – AQS and National Emissions Inventory – NEI). Exchange Network schemas vary widely in how they format geo-data. Many support the representation of only one set of coordinates (generally latitude/longitude, but not exclusively). Many States do not locally represent or manage their data in latitude/longitude, although relying on a wide variety of public domain and proprietary tools, they are capable of converting and transforming data from other coordinate systems into latitude/longitude (and from latitude/longitude to locally used reference systems). Many currently do this to be able to provide data to EPA.

As use of the Exchange Network for flowing data within States and between States grows, a need has been identified to support a diversity of coordinate reference systems beyond latitude/longitude within Exchange Network schemas, and to potentially support multiple coordinates representing the same feature, in one flow. The Team recommends that schema be developed, using GML as appropriate, to allow the option of representing multiple coordinate systems.

Recommendation #4: If coordinates are converted or transformed as part of an Exchange Network flow, the process used to convert or transform the coordinates from one reference system to another should be documented in metadata associated with the schema

The current Exchange Network shared schema components include extensive metadata, for each point, based on the Environmental Data Standards Council (EDSC) locational standard. GML also contains multiple metadata components. These should be evaluated for use as part of the evaluation and adoption process. These conventions include a means of indicating that a coordinate conversion process has been used, and specifying which one.

Recommendation #5:   Convene/continue a Team (or Teams) for discussions about the use of geospatial data on the Exchange Network (and other networks). Team(s) may be oriented toward geospatial policy issues or technical needs to address implementation of the recommended GML schema.

During the course of development of this paper, the Geospatial Strategy Team commented on the benefit of sharing and learning from other agencies implementing major networks. Engagement of the environmental, health, and geospatial communities resulted in enhanced understanding of their shared interests and common needs.

Further, development of the geospatial exchange capacity on the Exchange Network will require individuals with technical knowledge about GML and the Exchange Network working together to resolve approaches.

3. Technical Discussion: What are the best options for encoding geospatial data for exchanges on the Exchange Network?

The relationship between traditional data on the Exchange Network and geographic data is complex. Some users, who traditionally deal with tabular files of facilities, pollutants, or ambient monitoring data, tend to see the "location" or geographic coordinates as an attribute of the data file. The GIS users of the world, tend to see the geographic coordinates as the data and what those data represent (e.g., a monitoring site or facility identifier) as attributes of the geospatial data.

The focus of the recommendations and the discussion below is primarily on the former view of the world – traditionally tabular data that may contain some specific location information or geographic coordinates. The recommendations, however, recognize the desire in supporting seamless data integration, so are also supportive of more complex geographic representations (e.g., GML).

Problem/Opportunity Statement

Many of the historical exchanges of geospatial data on the Exchange Network have been limited to point data (e.g., the location of a facility or feature of interest related to a facility). In some cases, there is a need to exchange more complex geographic features, including linear features such as streams, or polygons that may represent the outline of a facility, wetland, parcel, etc. Historically, Exchange Network flows have involved the exchange of data primarily stored in tabular databases with minimal geographic components (i.e. facility footprint or monitoring site). The challenge is finding a means to exchange these data, without necessarily engaging in or adopting the full range of standards used by the traditional "GIS" community.

At the same time, while the interest is in simple geographies, there is also a desire to accommodate more complex geospatial data (e.g., that managed within a GIS) and to ensure that the Exchange Network can use the burgeoning field internet based geospatial tools/services (e.g., web mapping services). There is a need and opportunity to establish more seamless inter-operability of Exchange Network data with other internet based geo tools/services. Several demonstration applications to query and display arbitrary Exchange Network data by detecting the various XML tags used for geospatial data, and invoking the Google Maps API, have made the potential clear and compelling. Web and web service based tools that work with standardized geospatial data are growing rapidly. Establishing a standard means in current Exchange Network schema to represent various geographies is a critical need.

Principles of Standards on the Exchange Network

There are several principles or assumptions about standards on the Exchange Network:

- The Exchange Network is “standards based”,
- The Exchange Network will develop needed standards only when options for using existing standards have been fully explored;
- All flows will conform to a standard (for that flow) data exchange template (now XML schema);
- Schema will conform (across the Exchange Network) to common data standards, now expressed as Shared Schema Components (SSC)[1] (see below);
- The Exchange Network will seek a balance between data uniformity and local terminology/ standards, where possible local terminology should be preserved in data “round-trips”;
- Changes in standards will be implemented at the next “business” driven revision of the flow. Members are not expected to change a flow just because a derivative standard has changed.

Current Exchange Network guidance for geospatial (and all other data) standardization consists of a set of SSC and conventions for their use. SSCs are element and type definitions which can be re-used across flows. Unfortunately, implementation of SSC, especially those covering geospatial data has been highly uneven. Current Exchange Network XML schemas use a wide variety of XML tags/structures to describe geospatial information. Reasons for this are:

- When SSC were developed, there were no widely implemented, broadly supported XML conventions for geospatial information that could be easily adapted. Early schema used proprietary tags for point data, until unique Exchange Network tags were developed.
- It is often easier for flow designers to copy existing legacy formats for data than to adopt, or adapt the SSC;
- The SSC for geospatial data are limited to simple points, and are somewhat cumbersome. They contain more data (in the form of additional metadata) than most users want, and it is difficult to pick and choose elements.

.
Thus, the goal for identifying and adopting standards to support geospatial data representation on the Exchange Network is to build on already existing standards, based on non-proprietary approaches using on XML and to build SSC from these to support multiple Exchange Network flows.

XML Vocabularies for Geospatial Data

Investigations of options to meet the above principles have identified the following XML vocabularies for geospatial data.

[1] SSC are reusable XML schemas that organize related data elements common to multiple environmental data flows. These “snippets” of XML code are hosted on the Exchange Network website where users developing a flow can access a pre-determined list of SSC necessary for their particular data flow. SSC consist of a set of nested element, and element group definitions. While many of the element groups are small and discrete, developers often find it easier to simply pick and choose from the elements themselves rather than adopting the element groups whole (and then adapting those). Further in many cases the element names themselves have been changed. These issues make it difficult for applications (or humans) to identify core geo data (like a lat/long pair) tags.

Table 1: XML Vocabularies for Geospatial Data		
Vocabulary	**Description**	**Considerations**
Existing EN SSC	XML tags for Lat/long and extensive metadata	Not widely or consistently implemented in current EN schema; formats entirely proprietary to EN
W3C Geo	A basic RDF vocabulary that provides the Semantic Web community with a namespace for representing latitude, longitude and other information about spatially-located things, using WGS84 as a reference datum	Very simple point definition (Point/lat/long/alt) designed specifically for RDF documents. It is gaining some momentum as a namespace used within non-RDF XML documents, such as RSS 2.0 but does not provide definitions for more complex geometry.[2]
ESRI XML Format	XML vocabulary used by ESRI products	Proprietary to ESRI but widely used within industry by people using ArcGIS and its APIs.
GML	An XML encoding for the modeling, transport and storage of geographic information including the spatial and non-spatial properties of geographic features.	Widely adopted, non-proprietary vocabulary. Now in its third major version. See extensive discussion below. Used in the widely adopted OGC WMS and WFS. [3]
GeoRSS GML	GeoRSS GML is a super-set of the GML Point Profile and a subset of GML Simple Features. It can support multiple coordinate reference systems and is GML compatible.	Provides the areas of GML most often used by EN participants, representing points, lines and polygons. May not be suitable for more complex geometry. Adoption increasing due to ease of use when compared to full GML - see discussion below.
KML	XML vocabulary developed by Google, contains expressions for basic geometries but its focus in on interface control (e.g. zoom, perspective, detail), not geography description	Proprietary to Google but adopted for use with Google APIs. See discussion below about compatibility/evolution of KML and GML

Based on the principles outlined in the previous section, and these available XML vocabularies, GML was identified as the leading candidate for future Exchange Network use. GML has developed over the last decade or more, using non-proprietary XML.

Implementing GML on the Exchange Network

Geospatial data are inherently complex due to the relationships that must be maintained among the coordinates representing features of interest. This complexity means that software such as GIS that may be used to manage, represent, or exchange geospatial data are also complex. GML itself is a very large, complex vocabulary, thus requiring that GML consuming tools be the functional equivalent of full fledged GIS. During development of early Exchange Network schema and SSC, GML was considered as an Exchange Network standard, but rejected, because of its complexity and immaturity. Consequently Exchange Network schema adopted a combination of legacy standards (Environmental Data Standards Council and SSC) based XML tags for geospatial data.

---

[2] http://www.w3.org/2003/01/geo/#nonrdf
[3] http://www.opengeospatial.org/standards/wfs

In response to the complexity of the full GML vocabulary, various sub-sets (termed "profiles" in GML-speak) of GML have been defined to allow users to more easily reference the portions of the GML vocabulary needed for a given application. At the same time, adoption of "full" GML has slowly grown as well.

As of 2007, KML (described above) and variants of the GML profiles are leading "de facto" standards for use in web-services supporting geospatial data. Further, Google has recently begun participating in discussions with the GML community to ensure that future versions of KML are fully compatible with the GeoRSS GML profile described below.

The table below characterizes the current GML Profiles:

TABLE 2: GML PROFILES

Profile /Subset	Upward Compatible with GML?	Geometry Allowed	Industry Adoption	Multiple Coordinate Reference Systems	Notes
GML -FULL	☑	Points, lines, polygons and multi-polygons or complex features	Growing, but limited to uses requiring full expressivity.	☑	Suitable for highly complex geospatial data.
GeoRSS Simple		Point, lines, simple polygons and bounding boxes	Extensive (>1 million Google refs)	No, lat/long only	Simplified version of GeoRSS GML
GML Simple Features	☑	Points, lines, polygons and bounding boxes	Strong, but limited to GIS applications	☑	Designed primarily for use to/from traditional GIS software.
GeoRSS GML	☑	Points, lines, and simple polygons, bounding box	New but growing rapidly	☑	See text for description

A review of these options and discussions with GML and Exchange Network schema experts suggests that the Exchange Network should strongly consider adoption of the full GML, as well as adoption of the GeoRSS GML. GeoRSS GML will be an option for simple geography exchanges, while GML will support more complex flows. The geographic representation of points, lines, and polygons will be identical in each case.

The rationale for this recommendation is as follows:

- The GeoRSS GML profile fully supports the key features identified by the Team as being necessary for exchanges:
    - Can be used with different Coordinate Reference Systems (CRS).
    - Supports the needed geometries (point, simple line, and simple polygon).

- Provides a more tractable means of using just those areas of the GML vocabulary needed by most Exchange Network users.

- As a formal profile, GeoRSS GML is fully compatible with all GML services: (e.g. a full GML aware tool can interpret a GeoRSS GML expression without modification) (See Table 3).

TABLE 3: Comparison of GML 3.1 and GeoRSS GML Functionality

	GML 3.1	GeoRSS GML
Geometries Supported		
Point	✓	✓
Line	✓	✓
Polygon	✓	✓
Envelope	✓	✓
Arc	✓	✓
Various Other	✓	✗
Support for:	✓	✗
Topology	✓	✗
Complex Geometries	✓	✗
Data Quality Information	✓	✗
Metadata	✓	✗
Data Dictionary	✓	✗
Direction	✓	✗
Temporal Reference	✓	✗
Can carry alternate Coordinate reference Systems	✓	✓
Compatibility path to emerging GeoRSS GML web mapping and other services	✗	✓
Compatible with GML Tools	✓	✓
Compatible with GML 2	✓	✗
✓ = YES, ✗ = NO		

- GeoRSS GML[4], especially given its growing adoption, and coordination with vendors such as Google, presents a tremendous opportunity to "mash up" Exchange Network data with the much broader world of internet geo-services and data. Many of these services will *never* implement full GML because it is not needed. There is recognition that such services are the future of "light" GIS, and

[4] Note: The argument presented here recognizes that GeoRSS GML is on the path of many important web-service based tools. These tools are crucial complements to the Exchange Network web services future. Broader than RSS is a technology based on a very simple set of XML tags that describe "news" items (e.g., tags such as date, title, author, description and link). RSS has been successful, and is adapted to a wide range of uses, far beyond "news". Many of the most successful "mash-ups" of the web rely heavily on RSS. GeoRSS GML was designed to provide a powerful but simple way to geo-enable RSS. RSS is based on the concept of an "entry", where each entry is a "headline" and, in GeoRSS, each entry can be tagged with one "where" tag which defines a location (per the tags in GeoRSS). Data providers use "RSS Feed Generators" to produce RSS "feeds", and consumers use "feed readers" to consume them. The EN does not currently use RSS, instead, flow developers develop schema (set of XML tags) and then use the Node/Client to move them around.

crucially important to the future of the EN. The recommendation is for adoption of the definitions established in GeoRSS GML as the preferred representations for geospatial without adopting the RSS part. The XML tags would be incorporated into existing EN schema.

Risk considerations for this recommendation are that adoption of GML, without further guidance may result in Exchange Network flows using constructs, which, because of stylistic choices, cannot natively be used with the growing world of GeoRSS GML services (e.g., they are GML but not GeoRSS GML compliant). Conversely, the decision to adopt GeoRSS GML as a profile means that these services would be natively accessible. In both cases, native GML services would be accessible.

Some use of GML is already underway in a variety of Challenge Grants and other projects relying on the Exchange Network. See Table 4.

TABLE 4: EXAMPLES OF USE OF GML

Project	Project Description	GML Version	Reason for using GML
MnRAD (Minnesota Reach Address Database)	MnRAD is using GML 2.1 to flow event updates to MN's NHD.	2.1	The MnRAD project chose to develop a full GML profile that could be utilized by other flows.
Wisconsin Pilot Project	The project is using components of GML 2.1 to support the exchange of Superfund polygons to the LRT (locational reference tables).	2.1	
NatureServe	Sharing locations of endangered species critical habitats between states and NatureServe.	2.1	Used GML 2.1 to exchange point, linear and polygonal features.

Implementation Approach for the Exchange Network

The recommendation to adopt GML and GeoRSS GML can be implemented and documented through the Exchange Network's existing schema policies via the Network Technology Group. The new schema would replace, by reference, existing SSC. In addition, GeoRSS GML could be used as a starting point to begin to include GML elements without implementation of the entire GML 2.0 (or 3.1) specification Accompanying schema guidance documentation would also be useful.

The Exchange Network has an existing policy mechanism for XML standard adoption using the SSC. GeoRSS GML can be introduced to and support the Exchange Network schema development process by adoption of GeoRSS GML based element definitions as the new (version 3.0) geo-SSC. The Exchange Network can use the existing infrastructure (schema review, documentation, tools, training) to support adoption. This approach is depicted in Exhibit 1. The two flows shown in the Exhibit are fully compatible.

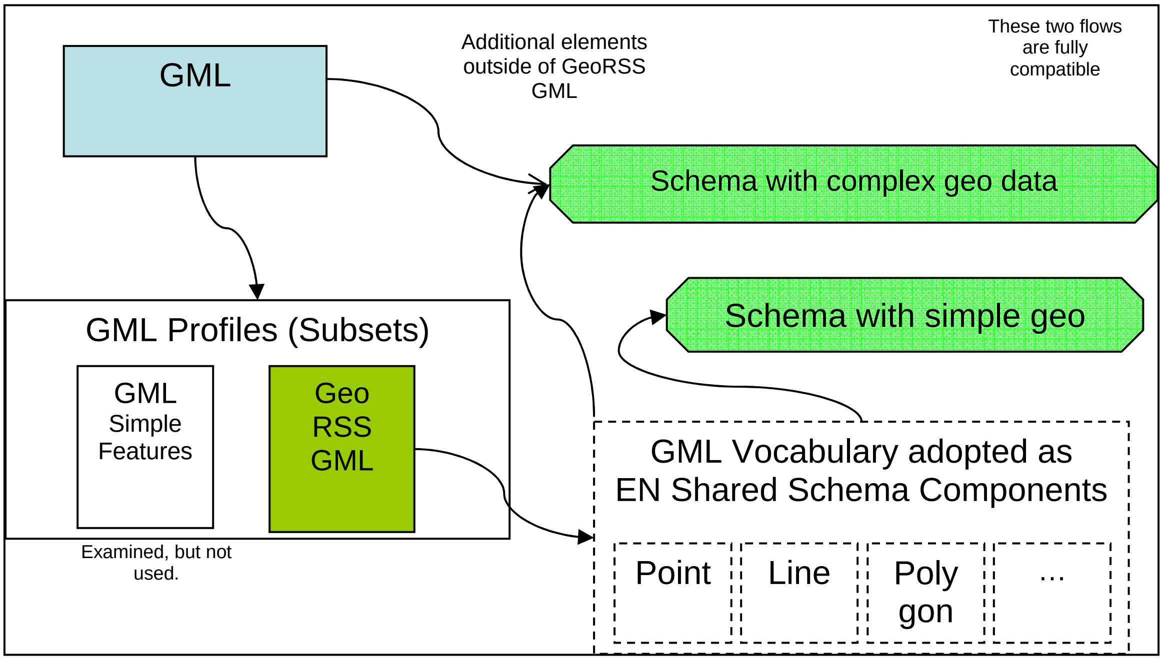

Exhibit 1: Relationship of GML, GeoRSS GML, and Exchange Network Schema

Once an SSC is created with GML or GeoRSS GML, schema developers requiring geospatial data formats would update their schema to include the components and data implementers would re-map their data to the new schema. While there is some cost associated with the initial SSC development, the required schema updates could be accomplished during regular updates that already occur, decreasing the overall cost of implementation. The incremental implementation costs of these changes, by themselves, are likely to be low for most users.

Implementation Issues

Given the challenges of fully implementing geospatial data exchanges, the ENLC should proceed incrementally, with clear charges to implementers. Issues to consider include:

- Schema architecture (technical XML implementation approach)
- Geo-metadata (how much/what geo-metadata should be carried)
- New terminology standardization
- Governance procedural questions (such as coordination with data standards efforts)

4. Technical Discussion: What are the challenges and options associated with multiple coordinate reference systems used by EPA, States, and others to represent geospatial data of interest in exchanges on the Exchange Network?

Problem/Opportunity Statement

Numerous coordinate reference systems (CRS) exist to link things to locations on the earth (see Appendix B for a brief discussion of CRS). Some CRS are useful for local applications such as a small state (e.g., State Plane Coordinates), while others provide a reference system that is usable globally (e.g., latitude/longitude). Organizations make decisions about the CRS they will primarily use based on a variety of factors. If data to be shared are represented in different CRS, organizations must address how to convert or transform the data such that they can be accurately associated with location and linked to other data.

EPA has adopted latitude/longitude as its CRS of preference. Most, but not all Programs at EPA require that locational data provided to the agency be represented with latitude/longitude coordinates. Data providers are not required to store their data in latitude/longitude, but to represent it with those coordinates to EPA. Several databases at EPA support only latitude/longitude (e.g., Facility Registry System – FRS). The Exchange Network has adopted the latitude/longitude standard and most (but again, not all flows) include latitude/longitude coordinates. Other national network efforts (e.g. CDC's Environmental Public Health Tracking Network[5] and the National Spatial Data Infrastructure) are silent on specific coordinate requirements.

The EPA policy poses challenges for a few States that do not represent or store data in latitude/longitude and do not have an interest or need to store those coordinates. Many States rely on State Plane or Universal Transverse Mercator (UTM) coordinates. As interest in and use of the Exchange Network among states and within States has increased, the ability to represent additional CRS within Exchange Network flows is desired. In some cases, this means the ability to support multiple, simultaneous CRS (e.g., a CRS used by the data sender and a CRS acceptable to the data receiver), in other cases, it means encoding CRS other than latitude/longitude.

Additionally, numerous programs and services exist to convert or transform coordinates between CRS (see Appendix C). There is no "standard" or quality assurance process that has been formally adopted by any entity to ensure that the process of conversion consistently results in accurate locations in another CRS. Frequently, the conversion process is not documented.

Previous discussions on the topic of coordinate transformation and latitude/longitude standards have been held by EPA and the States (some in groups formed under the ENLC). The challenge has not been clearly stated previously and no resolution has

[5] CDC is a participant in the Geospatial Strategy Team and is considering the need for guidance for geospatial features, but has not issued specific requirements.

been reached. An EPA-State Geospatial Data Standard Scoping Team was formed in 2004 to explore the feasibility of establishing geospatial data standards to facilitate transfer between EN partners. The Team discussed ways to develop shared definitions and recommended future work on developing a draft data standard to be addressed by geospatial and network experts. The Geospatial Strategy Team represents the group of experts addressing the issue.

Options for Coordinate Conversion/Transformation

The Geospatial Strategy Team originally identified the issue as a concern about effective means for performing coordinate conversions and transformations. During the December 2006 meeting, the following options for coordinate conversion/transformation were discussed. These are included here to document that conversation. Subsequently, during review of initial drafts, the issue was refined to be as stated above, not specifically concerned about approaches to coordinate conversion, but how to facilitate accurate sharing of information about location when different coordinate systems are in use. The recommendation specifically to address that issue is discussed in more detail in the next section.

- **Per current Exchange Network policy, conversions to latitude/longitude are done locally, by data providers. The list of open source coordinate conversion/transformation packages in Appendix C is an example of tools available. A comprehensive list could be developed, tested for quality/accuracy, and maintained by the Exchange Network.**

  The National Geodetic Survey in the Department of Commerce is the federal agency that oversees establishment of geodetic control nation-wide. They provide a service that includes both downloadable software and the means for performing coordinate operations interactively. They verify the ability of the software on their site to perform accurate coordinate operations. (see: http://www.ngs.noaa.gov/TOOLS/). This list can be used by any State or other Exchange Network partner to support coordinate operations. It also provides options for developing internal solutions if States have an interest and need, as source code is available. With a few exceptions, metadata are not generally available to describe coordinate operations.

- **Develop a centralized service, hosted by EPA or an outside organization (e.g. ECOS, FGDC, or OGC), based on commercially available software or free or "open source" code, to perform coordinate conversions.**

  The availability of such a service could accommodate changes in the evolution of datum or new coordinate systems such as the national grid more quickly than waiting for them to be developed as part of proprietary software. On the other hand, the National Geodetic Survey and agencies such as the U.S. Geological Survey work years in advance of such changes and frequent software updates provide ready means to accommodate these changes. A service could enhance the development of more complex web processing chains to integrate data from diverse sources, as well as applications. Discussions with entities traditionally responsible for coordinating and supporting national geospatial data efforts – the FGDC and OGC, specifically - have indicated that they have not identified a need for a centralized service to support coordinate operations (FGDC) or have not identified an effective means to cover the costs of providing such a service (OGC).

If users of the Exchange Network identified a need for a centralized service for coordinate operations, expertise and resources to support it would need to be identified.

- **Build coordinate operations into the Exchange Network nodes.**

Similar to the discussion under the previous bullet, there was no business need identified support this option as a recommendation. Software to conduct coordinate operations already exists. While building this type of operation into nodes on the EN is possible (could build on the OGC Coordinate Transformation Specification: http://www.opengeospatial.org/standards/ct) and potentially avoids the need to maintain a centralized service, such an approach could be costly, might entail licensing issues, and could result in a need to support multiple versions and "help desks" for local installations. Lack of expertise in geospatial transactions might impede local site or node administrators in implementing this option.

- **Data providers provide data in their native CRS and consumers perform the translation. EPA (or other data consumers) conducts transformations as needed.**

This option considers the possibility that data providers submit data in whatever format they use and that EPA, or other data consumers, performs the coordinate operations necessary to generate the CRS being used by the receiver. An advantage of this is that the burden of conversion is placed on the receiver.

Desirability and Feasibility of Providing for Additional Simultaneous Coordinate Reference Systems in Exchange Network Schema

The discussion below explores an option of establishing a provision (an optional "slot") in Exchange Network schema to allow for an additional CRS to be carried along with the Exchange Network standard latitude/longitude. This can be summarized as changing the debate from one of "latitude/longitude only" vs. "any CRC can be used", to a proposal for latitude/longitude AND any ONE other CRS.

As described above, many entities have an official CRS other than latitude/longitude. These entities may also have official policies requiring use of these CRS in official business. The Exchange Network policy of latitude/longitude as the primary CRS is upheld in the recommendations of this paper. This may leave some agencies in the dilemma of wanting to use the Exchange Network and having the means to perform transformations and conversions locally, but having concerns that their CRS of record is not "bound" to the included features. These concerns may include:

- That the official CRS is not made clear, and is not used;

- That there may be errors in the conversion used, and therefore a perceived or actual liability for errors;
- That the method of conversion is not specified, and may be critical to specialized users;
- That local users, expecting the native CRS cannot re-use the xml schema because it does not contain the native CRS

Team discussions, as well as debates in other geo-groups have indicated that the very obvious solution of allowing *an additional* CRS to "ride along" with the Exchange Network standard CRS may have many advantages, and impose only a very small additional overhead. Because GML already includes a rich vocabulary for clearly specifying the CRS in use (as well as conversions used between them) this change can be easily accommodated. This situation is very similar to other Exchange Network schema that carry both a "local" name for an element, e.g. the "official" name of a permit type issued by a state e.g., *and* an Exchange Network "standard" name for such permits. This allows local and remote users to have what they need, and it allows "lossless" conversions and round-trips of the data ---a record sent, processed and returned, will still have the names recognized by their respective users.

The Team therefore recommends that this option (already in ad hoc use by several Exchange Network schemas) be considered for general adoption. Under this option, schema (or the underlying shared schema components) would be designed with an additional slot for an alternative CRS; this could be optionally used by data producers. Data producers could fully express their "official" CRS as a form of additional documentation/disclaimer. Further, users who need the "official" CRC can pull it directly from the data, since it is included. Other users can ignore it (see below)

Additional design considerations for implementing this option include:

- ensuring this is done in a way that does not "break" compatibility with GeoRSS GML tools (see the example below for an illustration).
- ensuring that applications can find, use, and/or gracefully ignore the secondary CRS;

For example, one approach would be:

```
<location>
   <gml:Point>
       <gml:pos>45.256 -71.92</gml:pos>
   </gml:Point>
   <alternative CRS>
      <alternative CRS type="our agency official CRS"
      <gml:Point srsName="urn:ogc:def:crs:EPSG:6.6:26986">
            <gml:pos> -45.256 -110.45 </gml:pos>
      </gml:Point>
</location>
```

Note that the alternative CRS will be expressed in GML format and use that standard's list of CRS. Applications could ignore or process the `<alternative CRS>` element as needed. There is no "standard" way of providing these alternative CRS representations in GML, so some Exchange Network guidance will be required.

Coordinate Conversion Metadata

The ENLC Geospatial Strategy Team recognizes the critical importance of understanding the operations conducted on the data set, including transformations of coordinates. The way to do this is to capture detailed metadata on all aspects of the conversions/transformations. GML provides a means to do this, as do several other coordinate management packages.

References

Exchange Network Blueprint Team, 2000. Blueprint for a National Environmental Information Exchange Network. http://exchangenetwork.net/basics/blueprint_report.pdf

International Association of Oil and Gas Producers. Coordinate Conversions and Transformations including Formulas, Revised February 2007. http://www.epsg.org/guides/docs/G7-2.pdf

ISO 19111:2003 Geographic Referencing – Spatial Referencing by Coordinates http://www.iso.ch/iso/en/CatalogueDetailPage.CatalogueDetail?CSNUMBER=26016

National Geodetic Survey – Frequently Asked Questions http://www.ngs.noaa.gov/faq.shtml

Coordinate Systems Overview. Peter H. Dana, Department of Geography, University of Texas at Austin, USA. This unit is part of the *NCGIA Core Curriculum in Geographic Information Science*. These materials may be used for study, research, and education, but please credit the author, Peter H. Dana, and the project, *NCGIA Core Curriculum in GIScience*. http://www.ncgia.ucsb.edu/giscc/units/u013/u013.html

Standards and specifications
ISO 19111:2003, Geographic information - Spatial Referencing by Coordinates
ISO 19127, Geographic information - Geodetic Codes and Parameters
OpenGIS Implementation Specification: Coordinate Transformation Services, Revision 1.00
OGC Abstract Specification: Topic 2 - Spatial Referencing by Coordinates
OGC Web Coordinate Transformation Service (WCTS)
Recommended XML/GML 3.1.1 encoding of common CRS definitions (XML for CRS)
Recommended XML/GML 3.1.1 encoding of image CRS definitions (ImageCRS)

APPENDIX A: GEOSPATIAL STRATEGY TEAM MEMBERS

First Name	Last Name	Organization	Email	Phone
Tom	Aten	WI DNR	Thomas.Aten@Wisconsin.gov	608-267-7638
Andy	Battin	US EPA	Battin.Andrew@epamail.epa.gov	
Wendy	Blake-Coleman	US EPA	blake-coleman.wendy@epa.gov	202-566-1709
Mary	Blakeslee	ECOS	maryb@sso.org	202-624-3665
Paul	Caris	NJ DEP	Paul.Caris@dep.state.nj.us	609-633-1235
Alex	Charleston	CDC	aac4@cdc.gov	4044981028
Todd	Dabolt	US EPA	Dabolt.Thomas@epamail.epa.gov	202-564-1450
Ivan	DeLoatch	USGS	ideloatch@usgs.gov	703-648-5752
Sherry	Driber	NJ DEP	Sherry.Driber@dep.state.nj.us	
Chris	Duclos	FL DOH	Chris_Duclos@doh.state.fl.us	850-245-4264
Kristen	Durance	Ross & Associates	Kristen.Durance@ross-assoc.com	206-447-1805
David	Graves	Columbia River Intertribal Fish Commission	grad@critfc.org	503-736-3589
Julia	Harrell	NC DENR	julia.harrell@ncmail.net	919-715-0363
Craig	Kassinger	CDC	cak8@cdc.gov	404-498-0021
Kevin	Kirby	US EPA	kirby.kevin@epa.gov	202-566-1656
Roxanne	Lamb	FGDC/USGS	rhlamb@usgs.gov	
Matt	Leopard	US EPA	Leopard.Matthew@epa.gov	202-566-1698
Tom	McCarty	SAIC	mccartyt@saic.com	703-981-8943
Ron	McFarlane	NW Indian Fisheries	rmcfarlane@nwifc.org	360-528-4368
Mark	Olsen	MN PCA	mark.olsen@pca.state.mn.us	651-296-3412
Jeff	Patridge	MO DHSS	Jeff.Patridge@dhss.mo.gov	573-522-8330
Kurt	Rakouskas	ECOS	kurtr@sso.org	202-624-3684
Mark	Reichardt	Open GIS Consort.	mreichardt@opengis.org	301-840-1361
Dan	Saul	WA DOE	dsau461@ecy.wa.gov	360-407-6419
Harvey	Simon	US EPA	simon.harvey@epa.gov	212-637-3594
Raj	Singh	OGC	rsingh@opengeospatial.org	508-647-9385
Debbie	Stewart	WA DOE	Dste461@ecy.wa.gov	360-407-7048
Louis	Sweeny	Ross & Associates	louis.sweeny@ross-assoc.com	206-447-1805
John	Tooley	WA DOE	jtoo461@ecy.wa.gov	360-407-6418
Nancy	Tosta	Ross & Associates	nancy.tosta@ross-assoc.com	206-447-1805
Paul	Wiese	USGS	pmwiese@usgs.gov	303-202-4298
Paul	Will	CO DPHE	Paul.Will@state.co.us	303-692-3363

APPENDIX B: Background on Coordinate Systems and Coordinate Operations

Geographic data include references that relate features in the data to positions on the earth. These spatial references can be coordinates or other geographic identifiers such as addresses. A **coordinate system** is a set of mathematical rules for specifying how coordinates are assigned to points. The coordinate system is independent of the earth. A **coordinate reference system** (CRS) references coordinates to the earth. A CRS consists of a coordinate system and a **datum**. A datum describes the origin and orientation on the earth to establish the geodetic control for the coordinate system. Coordinate systems provide exact locational information when the details of the CRS are known. Geographic identifiers are frequently ambiguous. People often use the term "coordinate system" when they actually mean CRS.

A **coordinate operation** requires that two coordinate reference systems be defined prior to the operation. Coordinate operations are performed on coordinates and **not** on coordinate reference systems. This means that a coordinate reference system cannot be created from another coordinate reference system or modified by a coordinate operation. There are two types of coordinate operation. **Coordinate conversion** involves changes in the coordinate system, but not the datum. **Coordinate transformation** involves changes in both the coordinates and datum. As previously noted, transformations are subject to measurement errors, but conversions are not.

There are many different coordinate systems in use. Some of them are characterized as global, meaning they can be used anywhere in the world. Two common ones fitting this category are latitude/longitude and UTM as noted below. EPA relies on the use of latitude/longitude as previously noted:

- Latitude, longitude, and height system. The ***Prime Meridian*** and the ***Equator*** are the reference planes used to define latitude and longitude.
    - The ***geodetic latitude*** of a point is the angle between the equatorial plane and a line normal to the reference ellipsoid.
    - The ***geodetic longitude*** of a point is the angle between a reference plane and a plane passing through the point, both planes being perpendicular to the equatorial plane.
    - The ***geodetic height*** at a point is the distance from the reference ellipsoid to the point in a direction normal to the ellipsoid.

- **U**niversal **T**ransverse **M**ercator (UTM) coordinates. UTM defines two dimensional, horizontal, positions. UTM zones are identified by number.
    - Each zone has a central meridian
    - Locations within zones are measured in meters east from the meridian and north from the equator. Compensations are included so that all measures are expressed as positive numbers.

Other coordinate systems are considered regional; they are applied over limited areas such as states or regions. A key example of a regional coordinate reference system commonly used by States is the State Plane Coordinate System described below:

- State plane coordinate systems support a local reference system tied to a national datum.
    - The U.S. **State Plane System 1927** was based on the North American Datum 1927 (NAD-27) and encodes coordinates in feet
    - The **State Plane System 1983** is based on the North American Datum 1983 (NAD-83) and encodes coordinates in meters.
    - Each state has its own State Plane system with specific parameters and projections.
        - Smaller states use a single state plane zone while larger states are divided into several zones.
        - State plane zone boundaries often follow county boundaries.

APPENDIX C: NON-PROPRIETARY COORDINATE CONVERSION/ TRANSFORMATION SOFTWARE

Application Name	Website	Developer	Description
National Geodetic Toolkit	http://www.ngs.noaa.gov/TOOLS/	National Geodetic Survey	NGS has developed an on-line service that provides a library of conversion and translation tools and a service for interactively computing geodetic values. It is available as a web-based, single unit transformation service and downloadable desktop clients. NGS provides support and maintains updates and tests the software to ensure accuracy.
CORPSCON	http://crunch.tec.army.mil/software/corpscon/corpscon.html	US Army Corp of Engineers	A Windows-based program that allows a user to convert coordinates between Geographic, State Plane, UTM, and US National Grid systems on NAD 27, NAD 83 and HARNs. The program is based on the NADCON and VERTCON algorithms developed by the National Geodetic Survey (NGS)
GeoDLL	http://www.killetsoft.de/geodlle.htm	Killet Soft	GeoDLL is a function collection for datum and coordinate transformations between coordinates and reference systems, distance calculations, and geodetic functions that can be called from own applications. GeoDLL enables the program developer to perform professional grade coordinate transformations and supports all of the CRSs used in North America. Geodetic functions are available as source code to incorporate into alternative operating systems.
Proj4	http://proj.maptools.org/	USGS	Proj4 is an open source cartographic projections library that supports over 170 different projections including most of the ones used in national grid systems. All of the tools in this library require programming knowledge to implement and are not considered applications. The library contains two different transformation tools both based on Linux/Unix operating

			systems.
Gdal	**http://www.gdal.org/ogr/osr_tutorial.html**	Open Source	OGRCoordinateTransformation is an application within the GDAL translator library. This service is implemented on top of the Proj4 library noted above. It provides means to represent both projected and geographic coordinate reference systems and to transform between them. These services are loosely modeled on the OpenGIS Coordinate Transformations specification (OGC Spec), and use the same Well Known Text format for describing coordinate systems.
GeoTools	**http://docs.codehaus.org/display/GEOTOOLS/Home**	Open Source Project	An open source JAVA code library encompassing a variety of classes, object and methods required to implement geospatial applications. GeoTools is not an application but rather a suite of resources used to build both GIS server and desktop applications. Coordinate Transformation Services are included in GeoTools 2.0 and are able to transform between multiple CRS. Many Reference systems can be picked from pre-made static "objects" which make it easy for programmers to reference commonly used instances. It appears the NAD27 CRS are not included in the static objects menu but can be created by hand in the code.
Open Geospatial Consortium	**http://www.opengeospatial.org/standards/ct**	Open Geospatial Consortium	The Open Geospatial Consortium (OGC) provides a Coordinate Transformation Service Implementation Specification for coordinate transformations and standards for identifying coordinate systems and accessing coordinate transformation services. "When implemented, these specifications will ease data import; users of compliant applications will import data unaware of its coordinate system. If the application cannot import data in a given coordinate system, a compliant server will transform the coordinates to the native coordinate system."