# Essential Standards for Spatial Data Infrastructures that support environmental communication

by Trevor Taylor and Lance McKee

**Abstract.** As stated in the UN-GGIM's *Guide to the Role of Standards in Geospatial Information Management*, "In order to effectively leverage the value of geospatial information, the information must be easy to access and use." (UN-GGIM) Making geospatial information easy to access and use has always been the role of cartographers. This paper explains how information technology standards help cartographers fill that role, and how cartographers and others in the geospatial industry can help shape information technology for environmental communication.

**Keywords:** standards, environmental accounting, OGC

## Introduction

Cartography's purpose is to communicate information about the Earth and our activities on it. In this paper we consider environmental communication in particular and look at the ways in which modern cartography depends on – and can contribute to – the information technology standards framework.

In the pre-digital world, cartography was all about paper maps, charts, and globes – visual, graphical representations or metaphors for the surface of the Earth. This focus on visual communication persisted in early mapping and navigation software and in the "map layers" of Geographic Information Systems (GIS). However, GIS expanded the map metaphor into a universe of mathematical metaphors. Through the computer, the lines, polygons and colors suddenly became indicators not only of Earth features and phenomena, but also products of mathematical processes designed to further our understanding of spatial and temporal (not just geo*graphic*) relationships. 2 and 3 dimensional Earth features and phenomena that lend themselves to representation in map layers, and also phenomena in 4 or more dimensions (such as ocean currents and weather) that are better represented in complex equations, can now be modeled dynamically to help us understand the countless ways in which Earth features and phenomena interact.

The new powers of maps are a result of the mathematical metaphors that digital technology brings and of data's new independence from specific symbology and portrayal. Technology is not marginalizing cartography but rather expanding its scope. The tasks of cartographers are evolving with the advance of technology. Cartographers need to participate in the evolution of geospatial technology so that the technology meets its full potential as an enabler of sustainable development. Much of geospatial technology's evolution is determined by cooperative efforts to develop and advance standards.

## The Importance of "Green IT" Standards

Little need be said in this paper about humanity's precarious situation. World peace and prosperity, and perhaps our survival, now depend on how well we collaboratively manage our interactions with the Earth's atmosphere, water, soil, geology, energy resources and biota.

Information technology is all about the production, ordering, movement and use of bits and bytes. In a world of networked systems, bits and bytes must be ordered in ways that enable interaction among many systems and system components. Communication means "transmitting or exchanging through a common system of symbols, signs or behavior." Agreeing on a *common* system is the goal of standardization. In the information technology world, communication means "transmitting or exchanging through a common system of ordering bits and bytes."

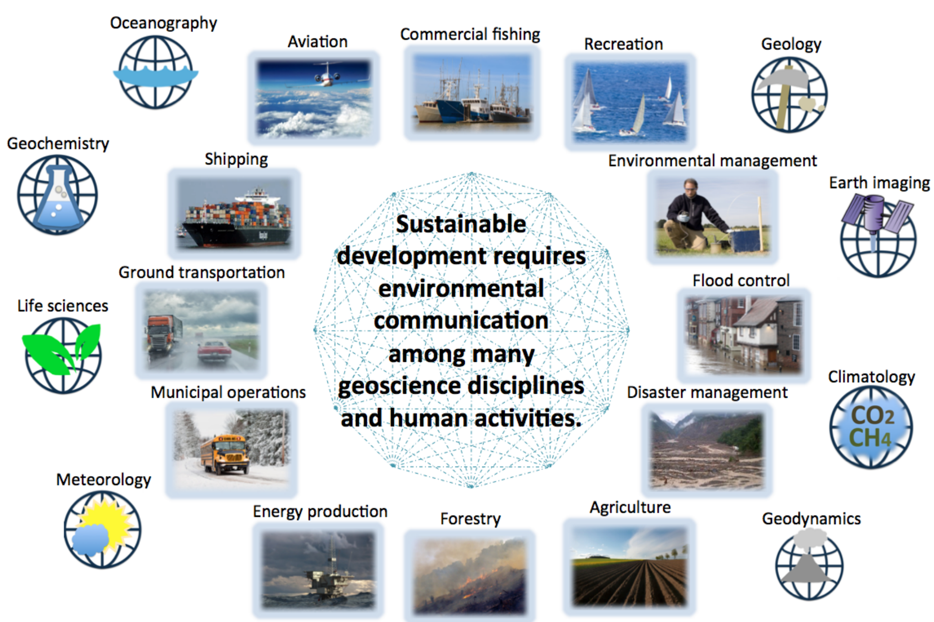


**Figure 1.** Measuring and managing our impact on the Earth requires frequent and widespread communication and integration of spatial environmental information.

In this paper we argue for the development and use of domain-specific but technically interoperable IT standards for communication and data integration within and between domains that focus on the environment.

These domains include each of the Earth sciences and many of the environmental response and management domains.

Not only do the standards apply to “green” domains, but the use of standards results in a reduction of waste in data collection and communication, which itself helps to “green” the economy.

Earth observations and measurements are not sufficient for understanding and managing our impacts on the Earth. What’s also necessary is *communicating and processing* those observations and measurements to turn them into information, knowledge, wisdom, policy and business. Maps, too, are not sufficient. We are visual creatures so maps are important, but flat, static digital maps are an artifact of geographic information in paper space. We need to put environmental geospatial information – micro, meso and macro, indoor and outdoor – into IT space. Geoscientists and environmental management stakeholders need to work with technologists so we can all benefit from the extraordinary and rapidly advancing capabilities of information technology.

Earth systems are all interrelated. This requires that the disciplines focused on different Earth systems – geology, hydrology, meteorology, etc. – use standard data encodings and geoprocessing software interfaces that are designed to be useful not only within those disciplines but also in cross-disciplinary studies and in non-academic activities that depend on geoscience. (Toth) In many cases the necessary standards don't exist and often almost no one is aware of their non-existence and the benefits they could bring.

Consider, for example, the data describing a city’s greenhouse gas emissions. The recently announced Global Protocol for Community-Scale Greenhouse Gas (GHG) Emission Inventories (GPC) (ICLEI) “uses a robust and clear framework to establish credible emissions accounting and reporting practices, thereby helping cities develop an emissions baseline, set mitigation goals, create more targeted climate action plans and track progress over time.” Such protocols – and there are many – are an important step toward describing what must be measured and communicated, but they don’t use and can’t by themselves provide an international standard way of encoding the data in a report. All the world’s carbon footprint calculators, GHG protocols and emissions standards are useful in local and limited ways but they do not communicate well. Many opportunities for collaboration and green business development are thus not realized.

Computers, in order to process data and communicate data, need to have data elements encoded using specific data types (integer, floating point, etc.). The data elements need to have identifiers, well-known and specific names that are used consistently across systems and reports. Ordering of elements is important. Relationships need to be articulated in standard ways. Attempts to verify, compare, update and aggregate GPC reports will require repetitive manual intervention until the GPC and other GHC protocols have a corresponding encoding standard or set of standards designed for international communication and integration of information about GHG emissions and offsets. Local systems don't necessarily need to use the international encoding standards, but they need to be able to translate local encodings to and from the standard encodings. Further, these standards must be designed to take full advantage of today's information technologies and spatial information technologies. The standards must also be designed to take advantage of emerging technologies, to the extent that this is possible. Once such standards are in place for all types of GHG reporting, preparing a Community-Scale Greenhouse Gas Emission Inventory will become a matter of automating the collection and aggregation of GHG for specific categories of emissions and offsets, such as methane emissions from landfills and carbon offsets provided by urban forests.

It is important to note that the environmental data communication standards we describe in this paper build on and add to the Spatial Data Infrastructures that nations and regions have been developing for last 20 years. The original SDI concept of "base data" layers shown in Figure 2 below has expanded to include "technical standards" that now enable easy movement of data and instructions between different GISs, Earth imaging systems, map browsers, alerting systems, sensor networks, and location services.

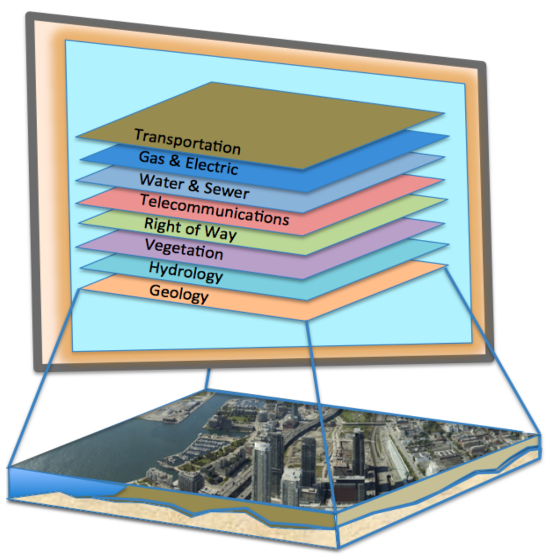


**Figure 2**. Local, national, and regional Spatial Data Infrastructures (SDIs) were originally conceived as a set of shared "GIS data layers" with common data models and metadata standards.

This paper is about further expanding SDIs by providing technical standards tailored specifically for these data layers and others. The OGC WaterML 2.0 Encoding Standard, for example, provides a common encoding for hydrology data. A candidate GeoSciML encoding standard for geology is in review. Others will be needed for the other base data layers and for other kinds of data used in environmental management.

The progress of environmental data communication standards is part of a larger progress toward "SDI maturity", as detailed in "A Guide to the Role of Standards in Geospatial Information" prepared for the UN Committee of Experts on Global Geospatial Information Management by the Open Geospatial Consortium (OGC), The International Organization for Standards (ISO) Technical Committee 211 Geographic information/Geomatics, and the International Hydrographic Organization (IHO). (UNGGIM)

In this paper we show examples of how environmental standardization work is being done and we outline a roadmap for how to build a unified set of environmental standards. We call on the cartography community to embrace this progress, and to help us add detail to the roadmap and enlist domain experts in building a unified standards platform for sustainable development.

## Environmental Accounting

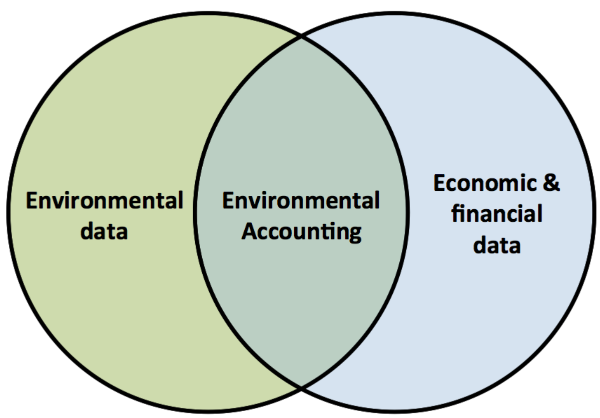


**Figure 3.** Environmental accounting is economic and financial accounting that takes into account data about environmental assets and liabilities.

Science-based analysis of environmental factors has little effect on the environment until the analyses impact the way business is done.

Environmental accounting and environmental auditing will depend on standards for environmental evidence, and these will require trustable measurements and trustable, transparent and digitally useful description and communication of measurements. Such communication will be necessary in all of the activities listed in Table 1 below.

Environmental IT standards can be used as a basis for indices that simplify and standardize environmental accounting tasks. Indices exist for other accounting purposes such as quantifying liquidity and summarizing changes in prices. Environmental data encoding and software interface standards can compress the complexity of environmental information in transparent ways so that accountants and auditors can use simply stated derived indices to produce, compare, aggregate and monetize results.

## What Needs to be Measured and Communicated?

Just as with creating a map, the work of developing communication standards begins with a clear and comprehensive statement of communication requirements. To develop environmental data communication standards, it is necessary to look at workflows and "use cases" involved in a wide array of present day environmental activities, such as those in Table 1:

| | | |
|---|---|---|
| Millennium goals | Cap and trade schemes | Oil spills |
| Ecodistricts | Environmental impact statements | Environmental regulation |
| Environmental taxes | Human Development Index | Environmental health |
| Sediment mgmt | Sustainable agriculture | Invasive species control |
| Microgrids | Environmental monitoring | Power plant monitoring |
| Carbon footprint calculators | City rating systems (STAR, C40, ISO 37120, etc.) | REDD (Reducing Emissions from Deforestation & Forest Degradation) |
| Resource/waste circulation | Integrated community energy systems | Triple bottom line accounting |

**Table 1**. A partial list of environmental activities that require efficient environmental data communication. This list will grow as environmental accounting becomes a factor in more and more human activities.

Analysis of spatial data communication requirements is the first step, because virtually all of this data is spatial and it needs to be integrated and used with other kinds of spatial data (jurisdictions, weather, transportation etc.). Next steps in standards development include:

2. Develop a conceptual model (most frequently using UML [Unified Modeling Language]) based on these requirements.
3. Work from the conceptual model to develop one or more candidate implementation standards based on targeted computing platforms such as XML, JSON and REST, or CSV.
4. Test prototype implementations of the standard in interoperability testbeds.
5. Vet the standard in a working group and call for public comments.
6. Implement the standard in products.
7. Promote the standard to encourage widespread use.

Tremendous societal and economic returns derive from relatively small strategic investments applied to such activity. As described later in this paper, with WaterML 2.0, GeoSciML, PipelineML and others, this work has begun.

## The IT Standards Foundation

As mentioned above, environmental IT standards must be designed to take full advantage of today's information technologies and their developers must anticipate emerging technologies.

TCP/IP, an IT standard, provides the communication foundation of the Internet. Other IT standards such as HTTP and XML (and the associated HTML standard) undergird the Web. Together, these fundamental IT standards and others provide a foundation for the common system for communicating geospatial data and geoprocessing instructions, which is specified mainly in ISO/TC 211 and OGC standards. Most environmental measurements have a location or area, so they are spatial data. (Because outdoor and indoor spaces are both part of our environment, and because technology is rapidly advancing to provide integrated description and measurement of indoor and outdoor spaces, we in the OGC increasingly talk about "spatial" instead of "geospatial" data.)

IT progress from 2000-2010 provided the Web services[1] foundation that supports OGC Web Services, which depend on XML and the OGC Geography Markup Language (GML), an XML grammar for encoding geospatial information. OGC Web Service standards[2] have brought geospatial technologies – GIS, remote sensing, navigation, facilities management etc. – out of their technology and vendor stovepipes into a much larger world of users and possibilities. Web services, however, are not the endpoint of IT evolution, and not all standards for sustainable development will be Web services. OGC Web Services standards will probably be with us for a very long time, but the TCP/IP and HTTP standards also provide a platform for new technology approaches not bounded by the established Web services paradigm. Currently, these include JSON[3], REST[4], Linked Data, apps and the Semantic Web, all of which offer new possibilities and a path to the future. These complement

[1] http://www.w3.org/2002/ws/Activity

[2] http://www.opengeospatial.org/standards

[3] http://json.org/

[4] http://en.wikipedia.org/wiki/Representational_state_transfer

the OGC Web Services standards with new approaches that can use and extend OGC Web Services.

Generally speaking, JSON, REST and apps are easier to implement than "heavy" Web services solutions. In the future, linked data and the Semantic Web will provide a quantum leap into a new level of IT-enhanced spatial awareness. Often these new approaches will involve simple point location data. Such content is unlikely to replace more complex spatial representations (polygons, grid arrays, 5D fluid models, triangulated irregular networks etc.), but simple point data will be useful in many situations, including Citizen Science[5] and the Internet of Things[6].

A concerted effort of collaboration is absolutely essential. Collaboration is necessary to create good standards that are widely implemented. Collaboration in developing standards also unveils deeper collaboration opportunities and requirements that cannot be met by technical standards. This is why standards development is best undertaken by both technical experts and policy experts.

The need to address specific spatial information requirements cuts across the missions of many Standards Development Organizations (SDOs). The OGC's established alliances[7] and history of success in forging alliances with other SDO's and professional organizations position the OGC as an SDO network hub for environmental standards development.

Most of the environmental standards we anticipate in this paper will be standards designed for particular domains of activity and particular "information communities." We explore this in the next sections.

## Semantic standards and technical standards

The domain-specific environmental IT standards development we call for in this paper involve semantic standards (metadata, data models, etc.) and technical standards (such as the OGC's software interfaces and encodings). Several key OGC standards were designed to be tailored to suit the needs of particular applications. In this section and the following section, we explain why and how environmental information communities (including hydrology, geology, weather/climate, pipelines and 3D urban models) have begun to do this tailoring by bringing their semantic standards experts together with technical standards experts in OGC working groups.

[5] http://www.nature.com/news/2010/100804/full/466685a.html

[6] http://www.mckinsey.com/insights/high_tech_telecoms_internet/the_internet_of_things

[7] http://www.opengeospatial.org/ogc/alliancepartners

## Semantic standards

In the geospatial world, an "information community" is an industry, profession, academic discipline or other domain that shares a set of spatial information communication requirements. The data model[8] used by an information community is their standard way of describing spatial information. It provides a data dictionary and related details necessary for the sharing, aggregation and comparison of data within the community. Metadata associated with a data set includes the data model along with other data about the data – date of collection, person or organization responsible for the collection, etc.

Most information communities who depend on sharing geospatial information have put in place data coordination committees and processes for creating and maintaining standard data models and metadata content standards. Such standards are often referred to as "semantic standards". Because of these standards, different information systems used within the community can "speak the same language" in terms of spatial feature dictionaries, ontologies and metadata. Different data sets that use the same data model can be aggregated or compared. Semantic standards also facilitate communication *between* information communities: When each community's data model is published and relatively stable, translation between different data models is easier and more precise, despite some inevitable loss of information. Data models necessarily evolve as information communities evolve, and so this data coordination process within and between domains is an ongoing activity.

[8] A deeper discussion of data models, ontologies, conceptual schemas etc. is beyond the scope of this paper. See "Ontologies and Data Models – are they the same?" http://topquadrantblog.blogspot.com/2011/09/ontologies-and-data-models-are-they.html.

# Technical standards

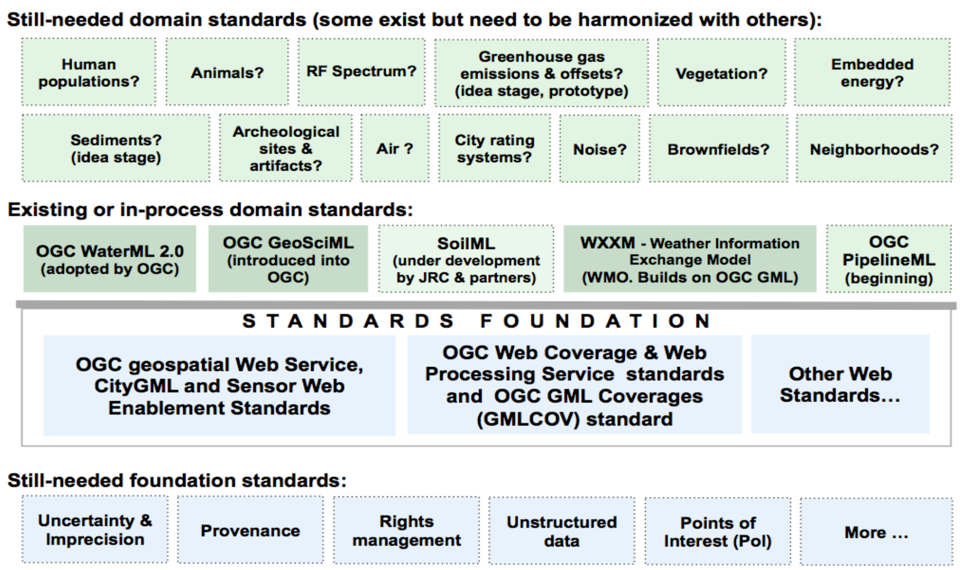


***Figure 4.*** *Much environmental standards work remains. Important foundational standards are in place for hydrology, geology, weather and soils. Progress depends on the commitment and participation of communities of interest who have a critical role to play in sustainable development.*

The members of the OGC have developed policies and procedures for working together to develop consensus-based open interface and encoding standards that provide a way for any two computer systems to request and return any kind of spatial data. These "technical standards" are broadly useful within all spatial data information communities, they support inter-community communication, and they are also essential for convergence and integration of different kinds of spatial technologies, such as 2D/3D/4D imaging, vector GIS, surveying, CAD, tracking, etc. The members of the OGC maximize new standards' viability by working together to promote widespread product implementation and market uptake of the standards.

Like semantic standards, technical standards evolve. The fundamental domain-neutral spatial technology standards framework is now in place, but rapid advances in technology make OGC members keenly aware that this foundation needs continual attention, as described above in the discussion of JSON, RESTful programming approaches, and linked data.

Technical standards are in place that can provide access control, security and certain privacy protections, but development also needs to address other issues such as geospatial data rights management and, as discussed above, data quality. Because technology is advancing so rapidly, much work remains in the broad area of technical standards for geospatial interoperability, despite the fact that a mature domain-neutral open spatial technology standards framework is already in place.

## Profiles and Application Schemas – Basis for Domain-specific Environmental Encoding Standards

Digital communications involving environmental data require a unitary system of communicating not only measurements but also descriptions of where, how and when the measurements were made.

The OGC Geography Markup Language (GML) Encoding Standard (also an ISO standard since 2007) provides the essential "where" and "when" components. An international open standard now widely implemented, GML is an XML grammar for expressing geographical features. Fundamental OGC Web Service Interface Standards such as the OGC Web Feature Service (WFS) Interface Standard are specifically designed to write and read GML-encoded data. The WFS standard is implemented in virtually all commercial GIS products, and therefore GML is the "lingua franca" for web-enabled geospatial content.

The OGC Observations and Measurements (O&M) Encoding Standard specifies an XML implementation of the OGC and ISO Observations and Measurements (O&M) conceptual model. It complements GML by providing essential information about "how" measurements were obtained. O&M provides a unitary system for encoding any type of observation or measurement, including volunteered geographic information (VGI), which may be anecdotal or photographic rather than the result of a measurement.

Both GML and O&M can be adapted to the particular needs of domains. They provide the means for domains to build their domain-consensus data models into an XML encoding that is recognized by any system that implements GML or O&M. The OGC WaterML 2.0 Encoding Standard, described below, provides a good example.

Any GML or O&M data can be used directly with other GML or O&M data. For example, an agronomist could do analysis using a hydrology layer and a crop type layer. Or an emergency response coordinator could use WaterML 2.0-encoded (see below) river flow data to trigger a Common Alert Protocol (CAP)[9] alert, because GML is embedded in CAP. Similarly, the US the National Information Exchange Model (NIEM), an XML encoding for exchanging information across state and local government bodies, uses GML and is interoperable with GML-based standards such as WaterML 2.0.

[9] http://www.galdosinc.com/archives/embedding-gml-in-non-gml-grammars

## Examples of Domain IT Standards

***WaterML 2.0*** Key players in the international hydrology community came together in the OGC to develop an international hydrologic data encoding standard. The work proceeds under the joint World Meteorological Organisation (WMO) and Open Geospatial Consortium (OGC) Hydrology Domain Working Group.

The OGC WaterML 2.0 Encoding Standard is implemented as an application schema of the Geography Markup Language version 3.2.1 and it makes use of the OGC Observations & Measurements standard. The core aspect of the model is the correct, precise description of time series hydrologic observations.

***WXXM*** The aeronautical community depends heavily on WMO weather data, but this community has special requirements for the sharing and use of the data. A new standard, WXXM – Weather Information Exchange Model, has been developed by the WMO, the International Civil Aviation Organisation ICAO, the US Federal Aviation Administration (FAA), EUROCONTROL and the OGC for the exchange of aeronautical weather information in the context of a net-centric and global interoperable Air Transport System (ATS).

WXXM uses GML tailored to the specific requirements of aeronautical meteorology and is based also on the OGC/ISO O&M. WXXM is a well-designed GML-based weather encoding model that it is consistent with standards from WMO and other organizations. For these reasons, other domains, such as the electric utility information system domain[10], are looking at using WXXM or harmonizing it with their standards.

***CF-netCDF*** The network Common Data Form (netCDF) is a data model and a collection of access libraries for array-oriented scientific data. Originally developed by the University Corporation for Atmospheric Research (UCAR), netCDF has been formally recognized by U.S. government standards bodies and has become a de facto standard used around the world, particularly in climate and ocean observation, analysis and modelling. For example, output datasets from climate models being used for the Fifth Assessment Report of the Intergovernmental Panel on Climate Change must be submitted in netCDF format, using the associated Climate and Forecast (CF) metadata conventions (CF-netCDF).

---

[10] International Electrotechnical Commission (IEC), US National Rural Electric Cooperative Association (NRECA), Smart Grid Interoperability Panel (SGIP) Another SDO whose remit requires weather information exchange is the American Society of Heating, Refrigerating and Air Conditioning Engineers (ASHRAE). ASHRAE SCP201P is a building/facility information model standard focused on loads; it includes a weather model based on WXXM..

NetCDF and its extensions provide an unprecedented degree of interoperability between complex Fluid Earth Systems (FES) (primarily involving oceans and the atmosphere) data and coverage-based data and systems (e.g. satellite observations); feature-based data and systems (e.g. GIS layers); and specimen measurements (e.g. sensor observations).

**GeoSciML** GeoSciML – Geoscience Markup Language – is a GML application schema used to support interoperability of geologic information provided by national Geology Survey organizations and other geological data custodians.

The GeoSciML project was initiated in 2003 under the auspices of the International Union of Geoscientists (UGS) Commission for the Management and Application of Geoscience Information (CGI)[29] working group on Data Model Collaboration - now the CGI Interoperability Working Group. In January 2013, a GeoSciML Standards Working Group was initiated in OGC, in collaboration with CGI, to develop a version 4 release as an OGC modular specification.

## PipelineML SWG

Pipelines are not natural resources, but OGC standards developed for energy, transportation, Smart Cities, emissions trading, urban modeling, indoor location, etc. are an important subset of sustainable development standards.

An OGC PipelineML Standards Working Group (SWG) was chartered in June 2014 to develop an open extensible standard intended to enable the interchange of pipeline data between parties, disparate systems and software applications without loss of accuracy, density or data resolution and without need for conversion between intermediate or proprietary formats.

## Other environmental domains in the OGC

The OGC City Geography Markup Language (CityGML), a GML application schema, provides an open encoding for multiple levels of 3D detail about the built environment. CityGML "Application Domain Extensions" (ADEs) have been developed for modeling noise, tunnels, bridges, Building Information Models (buildingSMART International's Industry Foundation Classes (IFCs)), water flow, utility networks, and immovable property taxation. OGC working groups have been chartered or are being chartered to address interoperability issues in urban planning, health, agriculture, and civil engineering. In addition, the OGC and the Electronic Commerce Code Management Association (ECCMA) aim to establish a joint working group to develop and promote implementation of a new standard under the name ePROP - electronic Property standardization. That proposed standard is intended to be a valuable support for workflows involving real estate and related financial dealings. This will include environmental workflows.

## Implications for cartography

OGC working groups will contribute to the essential evolution of environmental accounting, an evolution that we anticipate will benefit all aspects of the geospatial technology industry. Cartography can be expected to evolve in this context. Below are some of the possibilities:

- Cartography broadly defined as communication of spatial data will create opportunities for cartographers with interest and expertise in information technology and data modeling.
- As the technical boundary between indoor and outdoor spatial representation disappears, cartography will increase in scope.
- Cartography projects begin with understanding the client's requirements in detail. We see a growth in clients and a growth in the range of their requirements.
- With dynamic as well as static visualizations, there are design decisions to be made, editing and checking to be done.
- New technologies make older technologies and job skills obsolete. There's already far less hand-drafting, creation of negatives, overlays, lithography etc. But cartography will grow with the value of environmental data and thus there will be needs for people with basic understanding of cartography to manage cost estimating, scheduling, budgeting, and purchasing, supervising, planning, project coordination, public relations, and so on.
- As environmental data becomes more valuable to society, more attention will be paid to data custodianship, curation, cataloging, provenance checking, preservation etc. The data flow is increasing faster than our ability to store it, so decisions will need to be made about what to keep and how to publish it. These roles all involve cartography broadly defined. There will be far more people employed in these activities than are currently employed as map librarians.
- Currently much geoscience research data developed for research papers is lost after the papers are published. Society can ill afford this loss, so data curators are likely to be in demand. (McKee)
- Cartography research and education, like GIS research and education, will evolve to focus more on human communication of spatial data and system-to-system communication of spatial data.

The authors hope that this paper will help environmentally minded cartographers see how they and their organizations can advance their missions by advancing and embracing environmental IT standards.

## References

Parts of this paper have been excerpted from the OGC White paper, Information Technology Standards for Sustainable Development (2015), https://portal.opengeospatial.org/files/60920.

ICLEI "Global Protocol for Community-Scale Greenhouse Gas (GHG) Emission Inventories (GPC)" 2014 http://www.ghgprotocol.org/Release/GPC_launch

McKee "Standards enable open geoscience opportunities" (2011), Lance McKee, Earthzine http://earthzine.org/2011/03/30/standards-enable-open-geoscience-opportunities

Toth et al (2012) "A Conceptual Model for Developing Interoperability Specification in Spatial Data Infrastructures", Katalin Toth, Clemens Portele, Andreas Illert, Michael Lutz, Vanda Nunes de Lima, 2012, a JRC Reference Report. https://ec.europa.eu/jrc/sites/default/files/lbna25280enn_0.pdf

UN-GGIM (Committee of Experts on Global Geospatial Information Management) (2014), *A Guide to the Role of Standards in Geospatial Information Management*: http://ggim.un.org/docs/meetings/GGIM4/E-C20-2014-8_Essential%20Standards%20Guide%20for%20UNGGIM.pdf