Abstract—Geographic information is inherent to many application domains in various disciplines and constitutes an integral part of Earth sciences, including geology, geophysics, meteorology, hydrology, oceanography and soil science. Communication of sophisticated geographical data requires the use of complex technologies that enable interoperable geospatial information exchange channels. The primary authorities in geographic information standardization are ISO Technical Committee 211 (ISO/TC 211) and the Open Geospatial Consortium (OGC), which govern the abstract standards including information models for cross-domain concerns, architectures for distribution of geospatial services and implementation of ISO standards through service interfaces, data models and encodings.

The “language” that all participating parties should understand in the communication process is defined by Geography Markup Language (GML) Application Schema, which will conform to best practice guidelines and international standards if it is developed using the Hollow World modelling environment. Solid Ground toolkit may be especially useful during the modelling process in accomplishing some routine operations and refactorings. Application Schemas, which are defined in the conceptual terms of a particular domain, may be then easily transformed into a physical representation—a set of W3C XML Schemas [1] [2]. This transformation is achieved by using the FullMoon XML Processing framework.

This modelling approach, which places the structural definition of the information at the centre of the design process, is known as Model Driven Architecture [3], which makes the information model the only artefact that has to be maintained by the governing body. These technologies provide a complete set of tools required to design and implement an Application Schema using the ISO 19100 series of international standards.

Geoscience; computational geoscience; Application Schema; information modelling; conceptual modelling; data exchange; interoperability; UML; standardization; OGC; service interfaces; data models; encodings; GML; XML Schema.

I. INTRODUCTION

Geographic information (GI) is ubiquitous, so the communication of data between different domains or within the same community in an interoperable way necessitates the use of unified standards shared across disciplines that communities have agreed upon. The use of these standards enables the communication bridge between distributed applications and systems.

The primary authorities in GI standardization are ISO Technical Committee 211 (ISO/TC 211) and the Open Geospatial Consortium (OGC). While ISO/TC 211 focuses primarily on abstract standards including information models for cross-domain concerns, OGC concentrates on architectures for distribution of geospatial services and implementation of ISO standards through service interfaces, data models and encodings. Implementation of ISO 19136 [4], also known as Geography Markup Language (GML), provides a set of encoding patterns and auxiliary cross-domain XML components which serve as a framework for the implementation of specialized languages. These markup languages are based on models provided in a number of ISO/TC 211 standards, such that aspects that are common across domains are implemented in a common way. GML is an implementation of abstract artefacts, which is not intended for use “out of the box”. It is an XML grammar defined in W3C XML Schema (WXS) [1] [2] language that is intended to be used as the basis for concrete Application Schema implementation. Application Schemas detail the object types that are applicable to the specific community they are targeting. The GML specification includes rules for construction of Application Schemas by extending certain base classes such as gml:AbstractFeatureType.

Development of an Application Schema that is to be implemented using GML is complex. The ISO/TC 211 process requires that the information model is developed and formalized in Unified Markup Language (UML) [5] [6], which provides a formal notation for modelling software and business systems. Application Schemas follow a strict UML profile and define dependencies on ISO packages for any common elements (e.g. geometry, temporal primitives, spatial functions, etc.). This UML representation allows analysis and design to be undertaken in a graphical environment, proceeding seamlessly into implementation. This approach, known as Model Driven Architecture (MDA) [3], places the structural
The definition of the information at the centre of the design process and makes the information model the only artefact that has to be maintained by the governing body.

The Spatial Information Services Stack that we have developed provides three technologies (Hollow World, FullMoon, Solid Ground) described below that enable and support specialists throughout the Application Schema modelling process from conceptual information modelling to implementation. These technologies assist everything from the development and implementation of specialized GML-based Application Schemas for their specific domain to generation of WXS grammar documents. These schemas specify the data exchange format for transfer of domain data as a standard XML document, compatible with OGC Web Feature Service, which is the primary fine-grained GML data access interface.

## II. Conceptual Modelling

Design of an information system in MDA concentrates around the model, which defines abstract concepts, data entities, relationships, behaviour, operations and properties, which are intrinsic to those entities. Every modelling process begins with a conceptual definition of a model and its properties and then proceeds to the concrete physical implementation. ISO/TC 211 defines abstract standards for GML Application Schema modelling, which must be adhered to in all community domain models to ensure interoperability across disciplines. In the approach presented in this paper interoperability works by agreeing a common conceptual modelling framework together with canonical serialisations.

Community GML Application Schemas are developed by domain specialists based on abstract standards governed by ISO/TC 211 and OGC as described earlier. Fig. 1 illustrates this governance model and the modelling process itself, which consists of conceptual information modelling and model conversion into physical WXS-based implementation.

### A. Hollow World

Hollow World is a unified modelling environment that enables specialists in domains that utilize geospatial information to develop an information model for their application domain which conforms to international standards for interoperable GI. Instead of having developers code their Application Schemas manually, Hollow World offers an accessible and industry-standardized methodology for creating and editing Application Schema UML models, making the boundary between modelling and implementation processes more explicit. This allows domain experts to focus purely on the conceptual side of modelling that significantly simplifies the process and doesn’t require any specific knowledge of underlying implementation technologies (e.g. WXS). A model developed in this framework can be then easily transformed into a GML-conformant XML Schema.

Hollow World framework is provided by the ISO 19100 series of standards from the ISO/TC 211 “Harmonized Model" and is augmented by some additional components peculiarly useful for natural sciences and environmental monitoring applications. It provides a UML template which includes the components described in ISO 19100 standards, pre-loaded in a consistent manner. Users of Sparx Systems Enterprise Architect™ (EA) CASE tool may also leverage from the provided “UML Profile” containing the standard stereotypes and tagged values from ISO 19136.

### Figure 1. GML Application Schema modelling process.

<table>
<thead>
<tr>
<th>ISO/TC 211</th>
<th>Abstract standards for information modelling</th>
</tr>
</thead>
<tbody>
<tr>
<td>OGC</td>
<td>ISO implementations</td>
</tr>
<tr>
<td></td>
<td>Service interfaces</td>
</tr>
<tr>
<td></td>
<td>Data models</td>
</tr>
<tr>
<td></td>
<td>Encodings</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Community Application Schemas</th>
<th>Conceptual model</th>
<th>Physical model</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISO 19100 series of standards from the ISO/TC 211 “Harmonized Model”</td>
<td>UML Profile</td>
<td>FullMoon XML Processing Framework</td>
</tr>
<tr>
<td></td>
<td>Hollow World</td>
<td>Conformance testing</td>
</tr>
<tr>
<td></td>
<td>Solid Ground</td>
<td>XML Schemas</td>
</tr>
<tr>
<td></td>
<td>Conceptual UML model</td>
<td>Auxiliary documents</td>
</tr>
<tr>
<td></td>
<td>XML Metadata Interchange (XMI)</td>
<td>GML Application Schema</td>
</tr>
</tbody>
</table>
Scoping of the domain of interest takes an important aspect of conceptual modelling. A domain model or Application Schema should describe all the feature-types involved in transactions and operations in the domain. Since different aspects of a feature may be required for operations in different domains of discourse, the characterization of a feature-type may not be universal or comprehensive. It is merely necessary for the definition of a feature-type to accommodate the properties relevant to operations within the domain. Thus, scoping the domain is an important pre-requisite to make the domain-modelling exercise tractable, and the domain model only concerns the views of the feature-types required for operations within that domain.

The framework defines rules and information modelling guidelines for the formalization of a domain model, which includes packaging, model organization and dependency management. Following ISO 19109, the quantum of analysis and design is the “Application Schema”. Consistent with this, FullMoon XML Processing Framework complements Hollow World by being able to convert the UML model it produces directly into an XML Application Schema.

B. Solid Ground

Solid Ground is a toolset provided for the users of EA CASE tools that facilitates creation, management and exploitation of information models to drive intersystem interoperability. The toolset is implemented as an add-in and extends the MDA concepts by providing extensive support for Application Schema modelling in Enterprise Architect and Hollow World environments. A set of specially designed helper functions for the Hollow World environment assists in ensuring conformance to the UML profile, enforces the use of mandatory tagged values on classes and associations, and makes it easier to develop conformant models with best practice diagram views. Other features of the Solid Ground toolset include support for the ISO 19103 modelling idiom used within spatio-temporal application domains, information model import from various sources, version management, model mapping (i.e. establishes mapping between two separate UML models) and packaging.

III. FULLMOON XML PROCESSING FRAMEWORK

As required by the ISO/TC 211 process, the information model developed and formalized in UML which conforms to a strict UML profile can be then converted to an XML Schema according to a set of rules defined in ISO/DIS 19136:2007 Annex E [7]. The GML standard provides a WXS-based implementation of many common components and base-types for the development of Application Schemas. Integration with the ISO 19100 framework is achieved through rules for conversion of a domain model described using the UML profile and components to an XML Schema that imports the GML components and base types. This includes some additional package stereotypes («ApplicationSchema», «Leaf»), class stereotypes (e.g. «FeatureType») and tagged values on various elements of the model that fine-tune the XML implementation of the model.

FullMoon framework was originally designed for processing large UML models using XML mapping rules defined in ISO 19118, 19136 and 19139 standards. It processes the XML Metadata Interchange (XMI) [8] format representation of a model, generating XML schemas and some other views, with the mapping rules maintained as separate XQuery scripts. The framework architecture is schematically illustrated on Fig. 2. Models may be large and their XMI representation is highly verbose, so using traditional DOM and SAX parsers can be problematic. Efficient performance in FullMoon is achieved by using eXist Native XML Database engine to cache the model, which also provides XQuery [9] access to the XML infoset.

Apart from XML encodings of conceptual classes defined in the ISO 19100 series, Application Schemas may also re-use externally governed packages from other sources which may have existing canonical WXS representations. FullMoon supports this through a register for the Application Schema (which associates each externally governed package with an XML namespace and schema location) and a mapping table for each package (which associates each UML class with its representation as WXS element declarations and type definitions). The registers and tables are accessed by URI, so may be web-hosted at authoritative locations. All dependencies that the model has on components from other packages must be resolved in order for FullMoon to be able to perform a transformation, such as generating an encoding.

![Schematic architecture of the Fullmoon Processing Framework.](image-url)
The FullMoon framework is a “rules-driven” application that facilitates introducing, maintaining and enhancing of existent rules within rule sets. XQuery scripts implement the business logic used to transform input models into other views. Rule sets have been implemented for XML Schema and HTML documentation generation. A rule set has also been developed to test model validity and conformance to OGC standards. Detailed reports of non-conformities identify and locate errors in the UML Application Schema. The use of rule sets makes the FullMoon framework flexible to upgrade and easy to maintain, allowing the introduction of new conformance tests and/or processing rules.

IV. CONCLUSION

This Application Schema modelling approach and provided set of technologies guarantee a robust process for conceptual model design in strict accordance to ISO/TC 211 standards and guidelines, and provide a basis for smooth integration with OGC Web Feature Services. WXS schemas generated for an Application Schema using FullMoon framework define the grammar for XML documents used in communication between services and their consumers, and guarantees data and format integrity in the process. This plays an important role in Service-Oriented Architectures (SOA) [10] [11] and distributed information systems. Complex feature enabled OGC Web Feature Services bring information exchange interfaces to a qualitative new level that allows interaction with domain-specific services in an interoperable way.

REFERENCES