INSPIRE
Infrastructure for Spatial Information in Europe

Guidelines for the encoding of spatial data

Title	D2.7: Guidelines for the encoding of spatial data, Version 3.2
Creator	INSPIRE Drafting Team "Data Specifications"
Date	2010-05-19
Subject	Guidelines for the encoding of spatial data
Publisher	INSPIRE Drafting Team "Data Specifications"
Type	Text
Description	Baseline version of the guidelines for the encoding of spatial data
Contributor	Members of the INSPIRE Drafting Team "Data Specifications", INSPIRE Spatial Data Interest Communities & Legally Mandated Organisations, INSPIRE Consolidation Teams and other Drafting Teams
Format	MS Word (doc)
Source	INSPIRE Drafting Team "Data Specifications"
Rights	Public
Identifier	D2.7_v3.2.doc
Language	En
Relation	n/a
Coverage	Project duration

Table of contents

Foreword

INSPIRE is a Directive proposed by the European Commission in July 2004 setting the legal framework for the establishment of the Infrastructure for Spatial Information in the European Community, for the purposes of Community environmental policies and policies or activities which may have an impact on the environment.

INSPIRE should be based on the infrastructures for spatial information that are created and maintained by the Member States. The components of those infrastructures include: metadata, spatial data themes (as described in Annexes I, II, III of the Directive), spatial data services; network services and technologies; agreements on data and service sharing, access and use; coordination and monitoring mechanisms, processes and procedures.

The guiding principles of INSPIRE are that the infrastructures for spatial information in the Member States will be designed to ensure that spatial data are stored, made available and maintained at the most appropriate level; that it is possible to combine spatial data and services from different sources across the Community in a consistent way and share them between several users and applications; that it is possible for spatial data collected at one level of public authority to be shared between all the different levels of public authorities; that spatial data and services are made available under conditions that do not restrict their extensive use; that it is easy to discover available spatial data, to evaluate their fitness for purpose and to know the conditions applicable to their use.

The text of the INSPIRE Directive is available from the INSPIRE web site (http://inspire.jrc.ec.europa.eu/). The Directive identifies what needs to be achieved, and Member States had two years from the date of adoption (15th May 2007) to bring into force national legislation, regulations, and administrative procedures that define how the agreed objectives will be met taking into account the specific situation of each Member State. To ensure that the spatial data infrastructures of the Member States are compatible and usable in a Community and transboundary context, the Directive requires that common Implementing Rules (IR) are adopted in a number of specific areas. Implementing Rules should be adopted as Commission legal act and are binding in their entirety. The Commission is assisted in the process of adopting such rules by a regulatory committee composed by representatives of the Member States and European Parliament[1]. The committee is chaired by a representative of the Commission (this is known as the Comitology procedure). The committee will be established within three months from the entry in force of the Directive.

The IR will be shaped in their legal structure and form by the Commission legal services on the basis of technical documents prepared by especially convened Drafting Teams, for each of the main components of INSPIRE: metadata, data specifications, network services, data and service sharing, and monitoring procedures. For data specifications, the technical documents for each spatial data theme will be prepared by especially convened Thematic Working Groups.

This document represents a contribution of the Data Specification Drafting Team.

The previous major version of this document (version 2.0) was published on the INSPIRE web site for public view and commenting by registered SDICs and LMOs. 217 comments were received and have been resolved to produce this version. The comment resolution process included a workshop with representatives of SDICs and LMOs. Based on the discussions, the Drafting Team "Data Specifications" proposed comment resolutions that were reviewed by the Consolidation Team. The table containing the comments and the resolution is available on the INSPIRE web-site at http://inspire.jrc.ec.europa.eu/implementingRulesDocs_ds.cfm.

This baseline version (version 3.2) is published on the INSPIRE web site and has been improved in close co-operation with the Thematic Working Groups for the Annex I themes. The Thematic Working Groups have used the requirements and recommendations in this version to prepare the data

[1] The implementing rules are formally adopted through the comitology procedure that has been amended by Council Decision of 17 July 2006 (2006/512/EC). Under the new regulation, the Parliament and the Council are on equal footing for all comitology procedures related to co-decision acts. As a consequence, all measures must be ratified by all three institutions to come into force.

specifications as input for for the IR for the interoperability for spatial data sets and services. It is expected that this document will be updated during the next step of the data specification development process, if requirements for changes are identified.

It is important to note that this document is not a draft Implementing Rule, but a document that is a basis for the development of the thematic data specifications that will serve as technical basis for the legal text of the INSPIRE Implementing Rules.

The document will be publicly available as a ‘non-paper’, as it does not represent an official position of the Commission, and as such can not be invoked in the context of legal procedures.

Introduction

This document contains the baseline version of the guidelines for the encoding of spatial data (document identifier: D2.7).

One of the main tasks of the INSPIRE programme is to enable the interoperability and, where practicable, harmonisation of spatial data sets and services within Europe. Here, it is important to note that interoperability has to go beyond any particular community and has to take the various cross-community information needs into account. If one takes a look at the huge difference in the scope of the different themes (from reference systems to hydrography and from cadastral parcel to atmospheric conditions), the question does arise about the specific requirements of and for interoperability and harmonisation of the geographic information. These were also the questions faced by the Drafting Team "Data Specification" and one of the contributions of the Drafting Team is the identification of a set of *interoperability components*, which make the concepts of interoperability and harmonisation more tangible. Examples of interoperability components addressed in this document are: rules for application schemas, coordinate referencing and units model, identifier management, multi-lingual text and cultural adaptability, object referencing modelling, multiple representations (levels of detail) and consistency, and more. All these components do apply to (nearly) all themes identified within INSPIRE and this document together with the Generic Conceptual Model describes approaches to these shared components. Using this framework across the different themes will therefore result in a first level of interoperability.

It is important to note that “interoperability” is understood as providing access to spatial data sets as specified in Article 4 of the Directive through network services in a representation that allows for combining them with other such spatial data sets in a coherent way. This includes agreements about the different interoperability components. In other words, by enabling interoperability data can be used coherently, independent of whether the existing data set is actually changed (harmonised) or “just” transformed by a download service for publication in INSPIRE depending on the approach taken by the Member State. It is expected that these agreements will be based on existing data interoperability or harmonisation activities, whenever feasible and in-line with the environmental requirements.

The starting point for the development of INSPIRE data specifications is the input delivered by the LMOs and SDICs with their reference material and domain knowledge. Further and more specifically the foundation is formed by the internationally accepted standards reflecting the collective state-of-the-art knowledge (such as the reference model described in ISO 19101).

The individual themes (as defined in the Annexes I, II and III of the Directive and refined in document D2.3 'Definition of Annex Themes and Scope') have been and will be modelled based on document D2.5 ‘Generic Conceptual Model’. The result are data product specifications for the individual themes, i.e. conceptual information models that describe the relevant classes, their attributes, relationships, constraints, and possibly also operations as well as other appropriate information like data capturing information or data quality requirements. Care has to be taken that common or shared spatial object types relevant in multiple themes are identified and modelled in a consistent manner. This could then be considered a second level of interoperability: agreement on the shared (formal) semantics between the different themes. Note that the spatial characteristics of a spatial object will be represented by vector geometries, coverage functions and/or references to gazetteer entries.

The methodology (document D2.6) specifies how individual INSPIRE spatial data themes will be modelled based on the user requirements, the INSPIRE Generic Conceptual Model (document D2.5) and the relevant international standards. It provides a process model and tools to assist in the process. The document is applicable for INSPIRE data specifications. It is not required that it will be applied for the modelling of data specifications at the national level. What is important is that each Member State is able to transform existing data sets to the INSPIRE data specifications and publish the transformed data via network services. On the other hand, this methodology is expected to influence modelling activities at the national level, because it adds value to the national spatial data infrastructure and simplifies synchronisation with the INSPIRE data specifications.

How the geographic information will actually be encoded for the transfer process involved in download and other services giving access to data will be described in this document, the 'guidelines for the encoding of spatial data' - the third level of geographic information interoperability.

Besides the partner documents D2.3, D2.5, D2.6 and D2.8.m.n[2], this document is also related to other INSPIRE documents and registers:

- The terms used in this document are drawn from the "INSPIRE Glossary".
- INSPIRE application schemas are based on the Generic Conceptual Model and will continue to be maintained in the "Consolidated INSPIRE UML model" that also includes the external schemas, for example, the harmonised model of the ISO 19100 series published by ISO/TC 211. INSPIRE application schemas are in the process of being developed for every theme listed in the annexes of the INSPIRE Directive.
- The "INSPIRE Feature Concept Dictionary Register" is used to manage the names, definitions and descriptions of all spatial object types used in INSPIRE application schemas. In the future, the register may be extended to manage properties, too.
- Other registers include a coordinate reference system register, a feature catalogue register and a code list register.
- The implementing rule on metadata and associated guidance documents.
- The implementing rules on network services and associated guidance documents.

Figure 1 below illustrates relationships from the point of view of the data specifications. The boxes denote INSPIRE Implementing Rule documents or supporting documents, the cylinders registries. The arrows denote dependencies, the areas with dashed boundaries denote areas of responsibility.

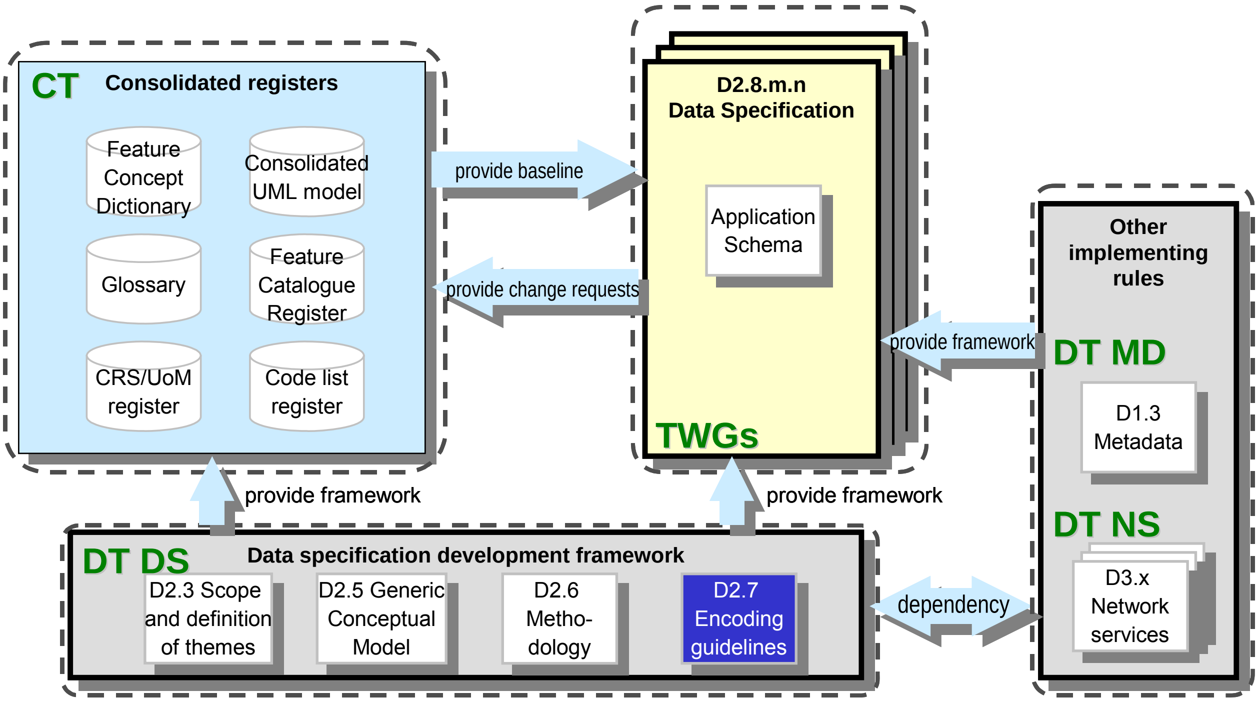

Figure 1 – The encoding guidelines as part of the data specification development framework

Since the conceptual modelling framework of INSPIRE is based in the ISO 19100 series of International Standards, in-depth knowledge about this series is required in every team developing an INSPIRE data specification.

The approach to encoding specified in this document can be summarised as follows:

- Conformant spatial data in INSPIRE conforms to the INSPIRE data specifications.
- The encoding of the spatial data in INSPIRE depends on the encoding rule chosen for the data.
- All valid encoding rules conform to ISO 19118.

[2] "m" is the number of the annex and "n" the sequential number of the theme within the annex.

- A default ISO 19136 and ISO/TS 19139-based encoding is specified. Additional or alternative encoding rules may be specified for each application schema.
- The encoding rules and application-schema-specific data structure schemas used within INSPIRE are published in guidance documents.

Two open issues remain and require a decision or further discussion with the Member States implementing the requirements as well as the future users of the data:

- The specification of URI schemes for persistent identifiers in the infrastructure.
- The encoding of void values in XML.

1 Scope

This document specifies requirements and recommendations for the encoding of spatial data for the purpose of data interchange between systems in INSPIRE. The focus of "data interchange" is understood by this document primarily as "access to data via services" which includes but is not limited to a download of a complete spatial data set.

The requirements and recommendations specified by this document are requirements and recommendations for encoding rules as specified by ISO 19118. These encoding rules will be specified in or referenced from INSPIRE data specifications.

NOTE 1 As a result, this document supports the development of the Implementing Rule on the interoperability of spatial data sets and associated guidance documents. The relevant requirements and recommendations stated in this document will be used to specify the necessary technical arrangements in those documents.

This document does not specify a mandatory encoding rule. The appropriate encoding rule(s) for each INSPIRE application schema will be determined based on the specific characteristics of the data.

For spatial data that belongs to a theme listed in Annex I or Annex II of the Directive, this document also addresses the interchange of updates of spatial data.

NOTE 2 Dependencies exist to the Implementing Rule with regard to download services, which is not yet available.

2 Conformance

This document specifies requirements INSPIRE data specifications. Every INSPIRE data specification shall pass the conformance test in Annex A.

3 Normative references

D2.5 v3.3, Generic Conceptual Model, April 2010

ISO 19118:-[3], Geographic Information – Encoding

ISO 19136:2007, Geographic Information – Geography Markup Language (GML)

ISO/TS 19139:2007, Geographic Information – Metadata – XML Schema implementation

IETF RFC 3986 (January 2005), Uniform Resource Identifier (URI): Generic Syntax

4 Terms and abbreviations

4.1 Terms

(1) application schema
conceptual schema for data required by one or more applications [ISO 19101]

(2) conversion rule
rule for converting instances in the input data structure to instances in the output data structure [ISO 19118]

(3) data interchange

[3] to be published

delivery, receipt and interpretation of data [ISO 19118]

(4) data transfer
movement of data from one point to another over a medium [ISO 19118]

NOTE Transfer of information implies transfer of data.

(5) encoding
conversion of data into a series of codes [ISO 19118]

(6) encoding rule
identifiable collection of conversion rules that define the **encoding** for a particular data structure [ISO 19118]

NOTE An encoding rule specifies the types of data to be converted as well as the syntax, structure and codes used in the resulting data structure.

(7) open standard
standard that is adopted and will be maintained by a not-for-profit organization in an open decision-making procedure available to all interested parties, that is available, distributable and usable either freely or at a nominal fee, and where the intellectual property is made irrevocably available on a royalty-free basis.

(8) transfer format
structured representation of data in a file for transfer between systems

NOTE Typically, a machine readable schema will document the structure of the data in the transfer file.

EXAMPLE GML (ISO 19136) encodes the application schema in XML Schema.

(9) spatial data
data with a direct or indirect reference to a specific location or geographic area [INSPIRE Directive]

(10) spatial data set
identifiable collection of spatial data [INSPIRE Directive]

(11) transfer protocol
common set of rules for defining interactions between distributed systems [ISO 19118]

(12) update <spatial data>
modification of parts of a **spatial data set**

NOTE The interchange of a complete spatial data set to replace and older version of the spatial data set is not considered an update to distinguish the different mechanisms required to create and process partial updates.

4.2 Abbreviations

CEN	European Committee for Standardization
D2.3	INSPIRE document "Definition and scoping of the Annex themes"
D2.5	INSPIRE document "Generic Conceptual Model"
D2.6	INSPIRE document "Methodology for the development of data specifications"
D2.7	INSPIRE document "Guidelines for the encoding of spatial data"
GI	Geographic Information
GML	Geography Markup Language
IANA	Internet Assigned Numbers Authority
IETF	Internet Engineering Task Force
INSPIRE	INfrastructure for SPatial InfoRmation in Europe
IR	Implementing Rule

ISO	International Organisation for Standardisation
KML	(formerly: Keyhole Markup Language)
LMO	Legally Mandated Organisation
OGC	Open Geospatial Consortium
SDI	Spatial Data Infrastructure
SDIC	Spatial Data Interest Community
TR	Technical Report
TS	Technical Specification
UML	Unified Modelling Language
URI	Uniform Resource Identifier
URN	Uniform Resource Name
UTF	Unicode Transformation Format
XML	eXtensible Markup Language

4.3 Verbal forms for the expression of provisions

In accordance with the ISO rules for drafting, the following verbal forms shall be interpreted in the given way:

- "shall" / "shall not": a requirement, mandatory for every data specification
- "should" / "should not": a recommendation, but an alternative approach may be chosen for a specific case if there are reasons to do so
- "may" / "need not": a permission

To make it easier to identify the mandatory requirements and the recommendations for INSPIRE data specifications in the text, they are highlighted and numbered.

Requirements are shown using this style.

Recommendations are shown using this style.

4.4 References within the document

In accordance with the ISO rules for drafting, references to highest level of the document structure include the word "Clause" (or "Annex" in case of an annex).

EXAMPLE "Clause 2", "Annex A"

References to lower levels within the document structure are given without this qualifier.

EXAMPLE 7.1, 7.1.8.4, A.1

References to ISO standards are given without the full title.

EXAMPLE "ISO 19101" instead of "ISO 19101 – Geographic Information – Reference Model" or "ISO 19101 (Reference Model)"

5 Background and principles

5.1 Requirements as stated in the INSPIRE Directive

5.1.1 Articles of the Directive

5.1.1.1 General remarks

This sub-clause provides an overview of the articles in the Directive which are addressed by this proposal and describes how they are addressed. To make this sub-clause easier to read, the relevant paragraphs from Chapter III “Interoperability of spatial data sets and services” are repeated in the text in italics.

5.1.1.2 Article 7(1)

Implementing rules laying down technical arrangements for the interoperability and, where practicable, harmonisation of spatial data sets and services, designed to amend nonessential elements of this Directive by supplementing it, shall be adopted in accordance with the regulatory procedure with scrutiny referred to in Article 22(3). Relevant user requirements, existing initiatives and international standards for the harmonisation of spatial data sets, as well as feasibility and cost-benefit considerations shall be taken into account in the development of the implementing rules. Where organisations established under international law have adopted relevant standards to ensure interoperability or harmonisation of spatial data sets and services, these standards shall be integrated, and the existing technical means shall be referred to, if appropriate, in the implementing rules mentioned in this paragraph.

This document, together with documents D2.3, D2.5 and D2.6, is intended to facilitate the drafting process of the implementing rules referenced above. In particular, the documents D2.5 and D2.6 provide a common framework for developing the various INSPIRE data specifications in a harmonised way. Access to data in an interoperable way is understood to occur by invoking network services that return data in a representation that allows for combining it with other INSPIRE data in a coherent way. The guidelines specified in this document are intended to support an approach to encoding data to enable interoperability between systems.

Inline with the standards-based approach taken in the common framework, encoding of data will be based on existing standards whenever possible.

This topic is discussed in more detail in D2.5 4.3.

5.1.1.3 Article 8(2)

The implementing rules shall address the following aspects of spatial data:
…
(e) updates of the data.

Article 8(2)(e) is relevant only for the Annexes I and II of the Directive and is addressed by this document (see clause 8).

5.1.2 Recitals in the Directive

Of the 35 recitals of the Directive, recital (6) is partially relevant for the technical specification of implementing rules on data specifications:

"The infrastructures for spatial information in the Member States should be designed to ensure

- *that spatial data are stored, made available and maintained at the most appropriate level;*
- *that it is possible to combine spatial data from different sources across the Community in a consistent way and share them between several users and applications;*
- *that it is possible for spatial data collected at one level of public authority to be shared between other public authorities […]."*

This points out that the encoding guidelines in particular needs to address the topic of references to spatial data maintained, stored and made available by other parties. See B.4 for requirements on the encoding of references.

Also, recital (16) and (28) state that *"implementing rules should be based, where possible, on international standards […]"* and that *"in order to benefit from the state of the art and actual experience of information infrastructures, it is appropriate that the measures necessary for the implementation of this Directive should be supported by international standards and standards adopted by European standardisation bodies.*"

5.2 A standards-based approach

The statements in D2.5, sub-clause 4.2 apply to the encoding guidelines, too. This document extends this to the implementation level, i.e. in the context of data the supported encoding rules that will be specified will conform to international or European standards whenever possible.

5.3 Data interoperability components

The work on INSPIRE data specifications is based on a framework that identifies the components relevant to the interoperability and harmonisation of data. These components are introduced and described in the Generic Conceptual Model, sub-clause 4.3. The different components cover different aspects that need to be addressed in the process. For each of the components, a separate clause in document D2.5 specifies how this component is addressed in the Generic Conceptual Model.

This document in particular addresses the harmonisation component "Data Transfer" which is described as follows:

“This component will describe methods for encoding spatial data as well as information products.

“The encoding of spatial objects will in general be model-driven, i.e. fully determined by the application schema in UML. Where appropriate, existing encodings will continue to be used.

“To support network services that are implemented as web services, spatial objects are expected to be primarily encoded in GML and metadata according to ISO/TS 19139. Coverage data is expected to use existing encodings for the range part, e.g. for the pixels of an orthophoto.”

5.4 Supported use cases

The following use cases for the download of spatial data are supported by these guidelines:

- The download of a complete spatial data set, spatial data set series or a named subset of a spatial data set. In this case, the spatial data to be transferred is fully defined by the supplier. Access to the data occurs by identifier (of the spatial data set) via a download service.

- The download of a single spatial object. Access to the spatial object occurs by identifier (of the spatial object) via a download service.

- The download of a user-defined subset of a spatial data set. The subset is selected by a set of queries created by the client of a download service and submitted to the download service. Queries select spatial objects based on their spatial, temporal, thematic and metadata properties as specified in the INSPIRE data specifications. Spatial data meeting the selection criteria are then retrieved from the spatial data set and provided to the user.

  EXAMPLE In the OGC Web Feature Service standard (ISO 19142), the GetFeature operation allows to query spatial objects based on selection criteria encoded using the Filter Encoding standard (ISO 19143).

- The download of updates of data since a user-specified date and time – based on life-cycle information in the data, if available.

In all cases, the download of the spatial data may potentially occur via the following modes:
- access to spatial data in form of a file by http URI
- synchronous or asynchronous online transfer of a response from a web service
- transfer of spatial data via physical media[4]

NOTE The use cases and download modes discussed in this clause include all modes identified by ISO 19109 for the data interchange between information systems.

The details of the modes supported by download services are specified in the Implementing Rule on the download service and associated guidance documents. INSPIRE data specifications may specify additional requirements and recommendations regarding the supported download modes and use cases for the data of a spatial data theme.

6 Encoding concepts

6.1 General remarks

The encoding process is discussed in detail in ISO 19118. This clause provides a description of the fundamental concepts of data interchange from this standard. The content of this clause is derived from Clause 6 of ISO 19118.

6.2 Data interchange

An overview of a data interchange is shown in Figure 2. System A wants to send spatial data to system B. Both systems, A and B, store data in an internal database according to an internal schema, but the schemas are usually different. The following logical steps are taken in order to transfer spatial data from A's internal database to B's internal database.

1. The first step for system A is to translate its internal data into a data structure that is according to the common application schema I. This is done by defining a transformation from the concepts of the internal schema to the concepts of the application schema and by applying appropriate software to transform the data. In Figure 1 this mapping is denoted M_{AI}. The result is an application schema specific, but system dependent data structure i_A.

2. The next step is to use an encoding service, which applies the encoding rule R to create a data structure that is system independent and therefore suitable for transfer. This encoded dataset is called d and may be stored in a file system or transferred using a transfer service.

3. System A then invokes a transfer service to send the encoded dataset d to system B. The transfer service follows a transfer protocol for how to do packaging and how the actual transportation over an online or offline communication medium should take place.

4. The transfer service on system B receives the transferred data, and according to the protocol the dataset is unpacked and stored as an encoded dataset d, e.g. on an intermediate file.

5. In order to get an application schema specific data structure i_B, system B applies the inverse encoding rule R^{-1} to decode the encoded data.

6. To use the dataset, B translates the application schema specific data structure i_B into its internal database. This is done by defining a transformation from the concepts of the application schema to the concepts of the internal schema and by applying appropriate software to transform the data. In Figure 1 this mapping is denoted M_{IB}.

[4] The reference material analysis indicates that delivery via mass storage media is still and will continue to be used in practice.

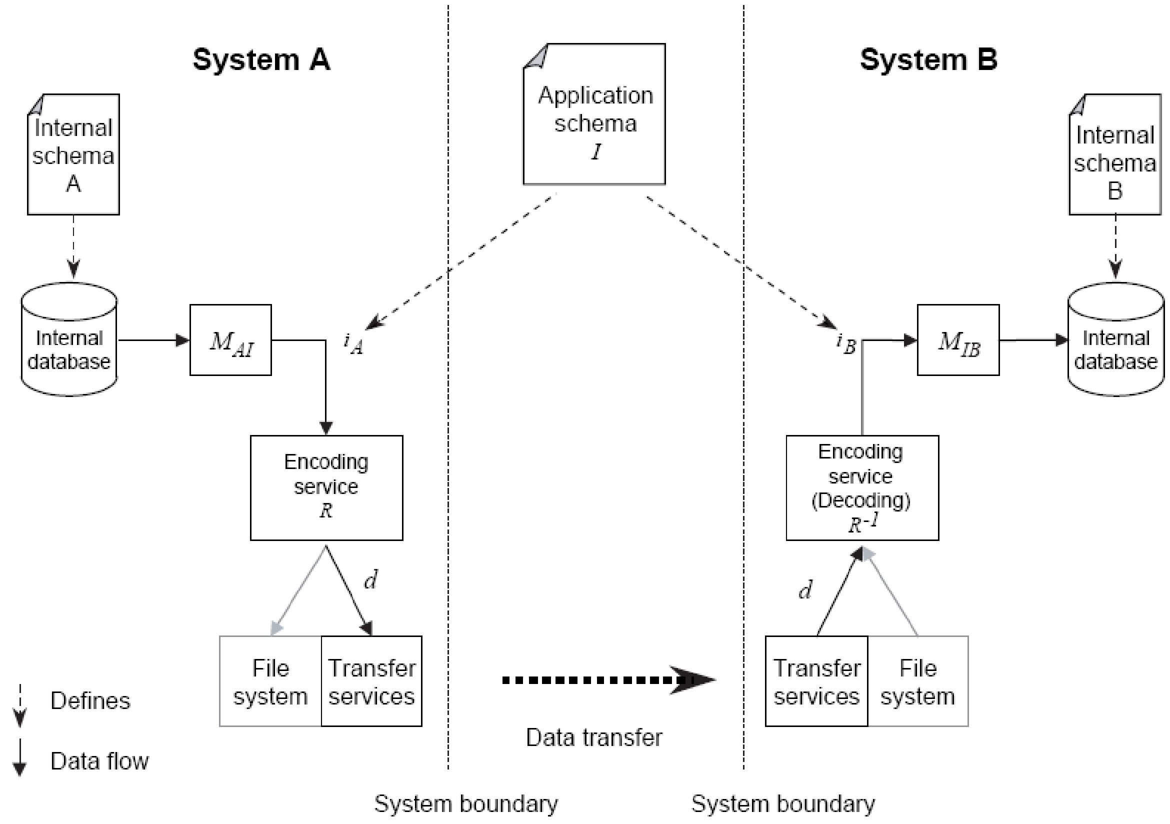

Figure 2 – Overview of data interchange between two systems [ISO 19118]

I.e., to ensure a successful interchange, A and B must agree on three aspects:

- An application schema I that defines the semantics of the content and the logical structures of the spatial data. Application schemas in INSPIRE conform to the requirements stated in the Generic Conceptual Model and are specified in INSPIRE data specifications.
- An encoding rule R defining the conversion rules for how to code spatial data corresponding to an application schema into a system-independent data structure. This document specifies guidelines related to such encoding rules (see steps 2 and 5).
- A transfer protocol for the transfer of the data between the two systems. Transfer protocols applicable in INSPIRE are specified in the Implementing Rule on Network Services and associated technical guidance documents.

6.3 Encoding rules

6.3.1 Concept

An encoding rule is an identifiable collection of conversion rules that defines the encoding for a particular data structure. The encoding rule specifies the data types to be converted, as well as the syntax, structure and coding schemes used in the resulting data structure. An encoding rule is applied to application schema specific data structures to produce system-independent data structures suitable for transport or storage. In order to define an encoding rule three important aspects are specified: the input data structure, the output data structure and the conversion rules between the elements of the input and the output data structures. Both the input and output data structures are written using a conceptual schema language and the concepts in the languages are used to define the encoding rule.

6.3.2 Input data structure

The input data structure is an application schema specific data structure. The data structure can be thought of as a set of instances, i.e. $\boldsymbol{i} = \{ i_1, ..., i_p \}$, where each instance i_k is an instance of a concept I_l defined in the application schema I. The application schema defines a set of concepts defined in the application schema $\boldsymbol{I} = \{ I_1, ..., I_m \}$.

6.3.3 Output data structure

The output data structure is defined by a data structure schema $\boldsymbol{D} = \{ D_1, ..., D_s \}$. *D* is the schema for the output structure and is not shown in Figure 2. The output data structure can be thought of as a set of instances, i.e. $\boldsymbol{d} = \{ d_1, ..., d_q \}$ where each instance d_k is an instance of a concept D_l. The schema D defines the syntax, structure and coding schemes of the output data structure.

The data structure schema is described with a schema language.

EXAMPLE An example of a output data structure schema is XML Schema.

6.3.4 Conversion rules

A conversion rule specifies how an instance in the input data structure shall be converted to zero, one, or more instances in the output data structure. The conversion rules are defined and based on the concepts of the conceptual schema language (i.e. UML and the ISO 19109 General Feature Model as referenced in the Generic Conceptual Model) on one hand and on the concepts of the output data structure schema D on the other hand. We need to specify a conversion rule R_i for each of the legal combinations of concepts in the conceptual schema language. The set of conversion rules are $\boldsymbol{R} = \{ R_1, ..., R_n \}$, where R_i is the i-th conversion rule.

NOTE The conversion rules are defined based on the two schema languages and not on any particular application schema. This is a generic approach that allows developers to write application-schema-independent encoding services, which can be used for different application schemas as long as the schemas are defined in the same conceptual schema language.

6.3.5 Encoding service

An encoding service is a software component that implements an encoding rule and provides an interface to encoding and/or decoding functionality. It is an integral part of data interchange.

Figure 3 presents the details of an encoding service and its relationships to important specification schemas (application schema *I* and data structure schema *D*). The encoding service reads the input data structure *i* and convert the instances to an output data structure *d* or vice versa.

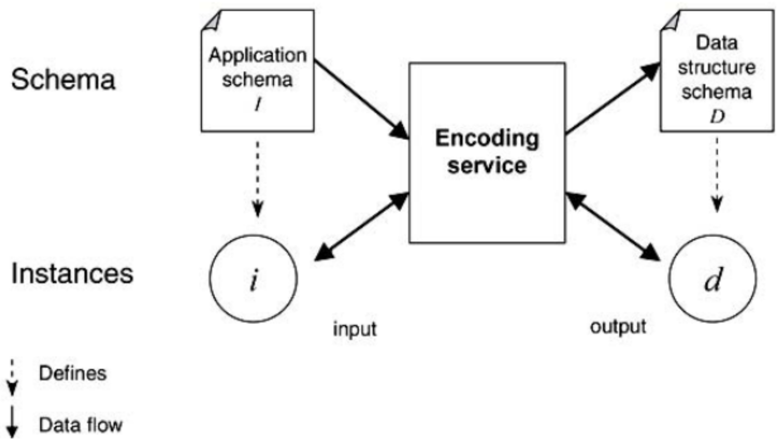

Figure 1 — Overview of the encoding process [ISO 19118]

NOTE It has to be understood that the encoding rule converts from a structure that is already application-schema-specific to a representation suitable for transfer. I.e. the internal schemas A and B will usually specify more information than it is represented in the application schema *I*. As a result, the internal databases in systems A and B will usually contain more information than it can be transferred based on the application schema.

6.3.6 Schema and instance level

Figure 3 also highlights that an encoding rule specifies conversion rules at both the schema and the instance level:

- At the schema level, the conversion rules define a mapping for each of the concepts defined in the application schema to corresponding concepts in the data structure schema.

- At the instance level the conversion rules define a mapping for each of the instances in the input data structure to corresponding instances in the output data structure. The instance conversion rules may be deduced from the schema conversion rules.

7 Encoding rules in INSPIRE

This document defines requirements and recommendations on encoding rules that are the basis for the interoperability of spatial data.

Requirement 1	Every encoding rule in INSPIRE shall conform to ISO 19118. In particular, it shall specify schema conversion rules for all elements of the conceptual schema language that are used in the INSPIRE application schemas to which the rule is applied.

NOTE 1 This requirement is intended to be part of the implementing rules for the interoperability of spatial data sets and services. Clear and unambiguous mappings from the concepts to the implementation level are considered a minimum requirement, otherwise the encoded data would no longer reflect the agreed application schema.

While it is not expected that specific encoding rules and output data structure schemas will be mandated as part of the implementing rules, interoperability between systems will require support for common encoding in all systems interchanging data. Thus, one or more recommended encodings will be specified for each INSPIRE application schema in guidance documents.

Requirement 2	Every data specification shall specify a mandatory encoding rule that has to be supported for the spatial data of that theme.

NOTE 2 The intention of this requirement is to ensure that data for the theme is available in a common encoding across Europe.

NOTE 3 Different encoding rules may apply to different "download use cases", see 5.4, although this should be avoided, if possible.

GML (ISO 19136) and ISO/TS 19139 are promoted as the default encoding in INSPIRE. The main reasons for this are:
- GML and ISO/TS 19139 cover encoding rules for large parts of the INSPIRE application schemas. This is not the case for any other commonly used encoding.
- GML specifies a XML based encoding rule for ISO 19109 conformant application schemas specifying spatial object types that can be represented using a restricted profile of UML that allows for a conversion to XML Schema[5]. In addition, GML provides a standardised encoding for many

[5] XML Schema is chosen to encode XML-based encoding schemas, because of its usage in the geo-community. This is also reflected in its usage in OGC (e.g. OGC Web Services, Sensor Web Enablement standards, SLD, GML) and ISO/TC 211 (e.g. ISO 19110, ISO 19118, ISO 19136, ISO/TS 19139). XML Schema is much more

commonly used types from core standards of the ISO 19100 series (in particular ISO 19107, ISO 19108, ISO 19111, and ISO 19123) that form the foundation of the Generic Conceptual Model (D2.5).
- ISO/TS 19139 specifies a XML based encoding rule for conceptual schemas specifying types that describe geographic resources, e.g. metadata according to ISO 19115 and ISO 19115-2 as well as feature catalogues according to ISO 19110.
- The reference material provided by SDICs and LMOs shows that GML is increasingly used in Member States and international communities to represent and transfer geographic information.
- GML and ISO/TS 19139 are well integrated with the current candidate standards of the network services.
- The use of these standards is inline with the recommendations of CEN/TR 15449 on encoding which promotes GML as the encoding method when transferring spatial objects and ISO/TS 19139 as the encoding method when transferring information related to spatial data such as metadata, feature catalogues and data dictionaries.
- A default encoding rule allows for a coherent encoding approach inline with the overall interoperability requirements of the Directive.

NOTE 4 There are open issues with this default encoding in the context of file-based data as it is used in some Annex II and Annex III themes. Existing work on this topic, a description of the issues and proposals for addressing them are documented in Annex C.

NOTE 5 The scope of this document is restricted to the encoding of spatial data. For the encoding of register items, it is expected that standardised encodings will be adopted where they exist (e.g. in the case of ISO 19110).

Therefore, guidance documents specifying an XML-based encoding on the basis of GML and ISO/TS 19139 will be developed for all INSPIRE data specifications. Clause 0 specifies recommendations for the encoding of specific aspects, for example, references. That clause with the recommendations is expected to eventually form part of the GML guidance document.

Recommendation 1	Every data specification should use the default encoding rule specified in Annex B as the mandatory encoding rule.

The default encoding is applicable to all use cases described in 5.4. For coverage data and other file-based data, open issues and proposals for addressing them are documented in Annex C.

Recommendation 2	If the default encoding rule is not a mandatory encoding rule in a data specification, the reasons for this should be explained and the default encoding rule should be supported as an additional encoding rule.

Reasons for the use of an alternative encoding rules are:
- a specific encoding for certain data has to be used to meet performance requirements
- existing file-based data (binary or text) has to be integrated as-is
- the application schema for a spatial data theme cannot be provided in accordance with the Generic Conceptual Model
- the application of the default encoding rule to an application schema is impossible or creates an impractical GML application schema

While a default encoding rule is provided, the diversity of themes and practice in the communities is recognised by allowing the use of additional encoding rules.

Recommendation 3	Encoding rules should be based on international, preferably open, standards.

widely used in current SDIs and SDI-related international standards than any other language to describe schemas for XML-based encodings like RelaxNG.

While flexibility to support additional encoding rules is important, harmonisation and reduction of the spread of encoding rules is important, too. The list of recognised encoding rules and output data structure schemas will be maintained in a register as part of the data specification process.

Recommendation 4	Additional encoding rules should only be added, if the new encoding rule has unique characteristics required by the data that are not fulfilled by an encoding rule that has already been endorsed.

EXAMPLE An encoding rule to support geographic visualisation in a number of commonly-used clients could be KML-based. Unlike GML, KML is an XML language focused on geographic visualisation, including annotation of maps and images. Geographic visualization includes not only the presentation of graphical data on the globe, but also the control of the user's navigation in the sense of where to go and where to look.

Recommendation 5	If the GetFeatureInfo operation is offered by a view service, it should use the same encoding as the download service.

8 Guidelines for the encoding of updates

The download of updates of data since a user-specified date and time is one of the supported use cases.

In this version of the encoding guidelines, the selection of the spatial data for the update is supported as long as the corresponding application schema supports the relevant temporal information.

For example, if the application schema supports life-cycle information for the spatial objects documenting the time when new spatial objects were inserted, or existing spatial objects were updated or retired. In this case, the life-cycle information can be used in queries to select only those spatial objects that were affected by changes since a user-specified point in time in the past. I.e. such queries can be used to download updates of a spatial data set. Other examples are modelling spatial objects as time series, etc.

Some reference material uses a specific encoding for updates of data, e.g. the encoding of the individual update actions like
- insert of a new spatial object,
- update to a property of a spatial object,
- replacement of a spatial object by a new version,
- supersession of a spatial object by a new version,
- retirement of a spatial object, or
- deletion of a spatial object.

However, since no user requirements for such an encoding had been identified, such an encoding is not supported by this version of this document.

In addition, data specifications may specify additional technical arrangements for the encoding of updates.

9 Rules for exchange metadata

Exchange metadata is metadata about the encoded data that provides the necessary metadata about the transferred data to the receiver of the data.

Recommendation 6	For the download of a complete spatial data set, the data set should include the data set metadata for evaluation (MD_Metadata as specified in the INSPIRE data specification) and use (the INSPIRE data specification itself).

NOTE The motivation for this recommendation is that where complete datasets are exchanged the data set metadata is provided, too, as it is a full delivery of a data set and users can expect to receive a complete "package". This is a recommendation only and may be adapted where needed, as it may be more appropriate for some data sets (where metadata is a usually a small fraction of the complete spatial data) than for others (where metadata may be quite large).

EXAMPLE 1 In an XML instance the metadata elements will be encoded as part of the data set.

```
<i:DataSet
   xmlns:i="http://inspire.jrc.ec.europa.eu/schemas/base/3.3"
   xmlns:ad="http://inspire.jrc.ec.europa.eu/schemas/ad/3.1"
   xmlns:gmd="http://www.isotc211.org/2005/gmd"
   xmlns:gml="http://www.opengis.net/gml/3.2"
   xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
   xsi:schemaLocation=" http://inspire.jrc.ec.europa.eu/schemas/base/3.3
       http://schemas.inspire.jrc.it/baseTypes/3.3/all.xsd
       http://inspire.jrc.ec.europa.eu/schemas/ad/3.1
       http://schemas.inspire.jrc.it/addresses/3.1/all.xsd">
   <i:metadata>
      <gmd:MD_Metadata>
         <!-- metadata elements encoded according to ISO/TS 19139 -->
      </gmd:MD_Metadata>
   </i:metadata>
   <!-- address features in the data set -->
</i:DataSet>
```

Recommendation 7 For the download of a user-selected part of a spatial data set (download of spatial objects based on queries), the response of the download service should not include the data set metadata, but should provide a reference to the data set or data set series metadata in a discovery or registry service.

EXAMPLE 2 In an XML instance the reference will be encoded using an Xlink:

```
<wfs:FeatureCollection
   xmlns:ad="http://inspire.jrc.ec.europa.eu/schemas/ad/3.1"
   xmlns:wfs="http://www.opengis.net/wfs/2.0"
   xmlns:gml="http://www.opengis.net/gml/3.2"
   xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
   xsi:schemaLocation="http://www.opengis.net/wfs/2.0
       http://schemas.opengis.net/wfs/2.0/wfs.xsd
       http://inspire.jrc.ec.europa.eu/schemas/ad/3.1
       http://schemas.inspire.jrc.it/addresses/3.1/all.xsd">
   <wfs:metadata
    xlink:href="http://service.xyz.org/discovery?REQUEST=GetRepositoryItem&ID=5325423"/>
   <!-- requested address features -->
</wfs:FeatureCollection>
```

Annex A
(normative)

Abstract Test Suite

A.1 Existence of a mandatory encoding rule

a) Test Purpose: Verify that a mandatory encoding exists.

b) Test Method: Inspect the data specification to verify that at least one mandatory encoding rule is specified for each application schema.

c) Reference: Requirement 2

d) Test Type: Basic Test

A.2 Completeness of each mandatory encoding rule

a) Test Purpose: Verify that each mandatory encoding rule is complete.

b) Test Method: Inspect each encoding rule to verify that for every concept in the application schemas a conversion rule exists that maps it to the output data structure schema.

c) Reference: Requirements 1 and 2

d) Test Type: Basic Test

A.3 Conformance of each mandatory encoding rule

a) Test Purpose: Verify that each mandatory encoding conforms to ISO 19118.

b) Test Method: Inspect each encoding rule to verify that it conforms to ISO 19118.

c) Reference: Requirement 1, ISO 19118 B.1-B.3

d) Test Type: Basic Test

Annex B
(normative)

Default encoding rule

B.1 General remark

For the avoidance of doubt, the requirements stated in this annex are only applicable for INSPIRE application schemas that reference the default encoding rule as a mandatory encoding rule.

B.2 Use of URIs

Within INSPIRE specifications, metadata and spatial data require access to resources, such as schemas, documents, reference systems, spatial objects, definitions, code lists, spatial data services, etc. In order for the larger community of users of spatial data and developers of software components, a persistent mechanism to uniquely reference resource –independent of their physical location in the network.

Uniform Resource Identifiers (URIs) as defined in IETF RFC 3986 are intended to serve as these persistent, location-independent resource identifiers.

Recommendation 8 Identifiers of resources should be URIs in the "http" scheme.

Recommendation 9 While it is not strictly a requirement that the resource can be accessed via the HTTP protocol using its http URI, it is strongly recommended that this is the case.

The use of HTTP includes support for content negotiation to access a resource in different languages and/or different encodings, if supported.

Recommendation 10 Every Member State and the European Commission should develop, document and maintain a URI scheme for their resources.

EXAMPLE The URI scheme below is taken from the paper "Designing URI Sets for Location" developed within a holistic framework for all kinds of data in the United Kingdom. The paper employs the W3C definition of "spatial things":

Spatial Thing	"Anything with spatial extent, i.e. size, shape, or position. e.g. people, places, bowling balls, as well as abstract areas like cubes."	W3C "WGS84 Geo Positioning: an RDF vocabulary"

HTTP URI for spatial-things follows the standard pattern given above:

http://{sector}.data.gov.uk/id/{concept}/{reference}[/{version}]

and carries that structure through to associated reference document URIs of the form:

http://{sector}.data.gov.uk/doc/{concept}/{reference}[/{version}][[{rendition}]

Example Spatial Thing URI	Description
http://statistics.data.gov.uk/id/local-education-authority-area/201	City of London LEA area
http://statistics.data.gov.uk/id/local-authority-district/00HD	South-Gloucester Local Authority Area
http://transport.data.gov.uk/id/road/M5	The M5 motorway.
http://reference.data.gov.uk/id/postcode/SO164GU	The SO16 4GU UK Postcode area.
http://land.data.gov.uk/id/title-plan/cph-de15-27	A UK "cadastral" parcel.

Ideally, INSPIRE thematic references from INSPIRE spatial-objects to spatial-things will take the form of references made using URI assigned to designate the relevant spatial-thing. In this way reference data about the referenced spatial-thing is readily accessible from the web.

INSPIRE spatial-objects with an assigned INSPIRE Unique Object Identifier are assigned permanent a “location.data.gov.uk” URI of the following form:

http://location.data.gov.uk/doc/{namespace}/{local-id}[/{version-id}][{rendition}]

where the **{namespace}**, **{local-id}** and **{version-id}** fields carry values that are lexically identical to the corresponding components of a native INSPIRE unique object identifier. Note that the **{version-id}** field is optional. References made without a **{version-id}** refer to the most recent version of a spatial-object at the time that the reference is de-referenced (followed).

Example Spatial Object URI Data.gov.uk URI	Owning Authority URI	Description
http://location.data.gov.uk/doc/osgb/1234567890123456	http://data.ordnancesurvey.co.uk/doc/osgb/1234567890123456	A TOID based identifier from Ordnance Survey
http://location.data.gov.uk/doc/00BH/123456789012	http://location.walthamforest.gov.uk/doc/00BH/123456789012	A UPRN defined by Waltham Forest Borough Council

Issue 1 The URI scheme for the resources in the INSPIRE registers has not been developed. For now we use a pseudo scheme in this document.

B.3 Schema conversion

B.3.1 General rules

Requirement 3 A GML application shall be specified for the application schema.

NOTE The application schemas specified in the INSPIRE Consolidated UML Model – either in the Generic Conceptual Model or as part of a data specification for a spatial data theme – are managed independently. As a result there must be separate GML application schemas using different XML target namespaces.

Requirement 4 The encoding rule specified in ISO 19136 Annex E shall be applied with the additional rules stated in this Annex. For types within the scope of the ISO/TS 19139 encoding rule, the encoding rule of ISO/TS 19139 shall be applied.

Recommendation 11 To identify the applicable encoding rule, a tagged value “xsdEncodingRule” should be provided for packages and classifiers. A value “iso19136_2007_INSPIRE_Extensions” indicates the GML-based encoding rule, “iso19139_2007” indicates the ISO/TS 19139 encoding rule. If no value is provided “iso19136_2007_INSPIRE_Extensions” is the default.

B.3.2 Conversion of the application schema to a GML-specific implementation schema

Since the UML profile supported by the Generic Conceptual Model is more general than the UML profile specified by ISO 19136 Annex E, it may be required to derive first a GML-specific implementation schema in UML from the relevant INSPIRE application schema. This process is illustrated in Figure 3.

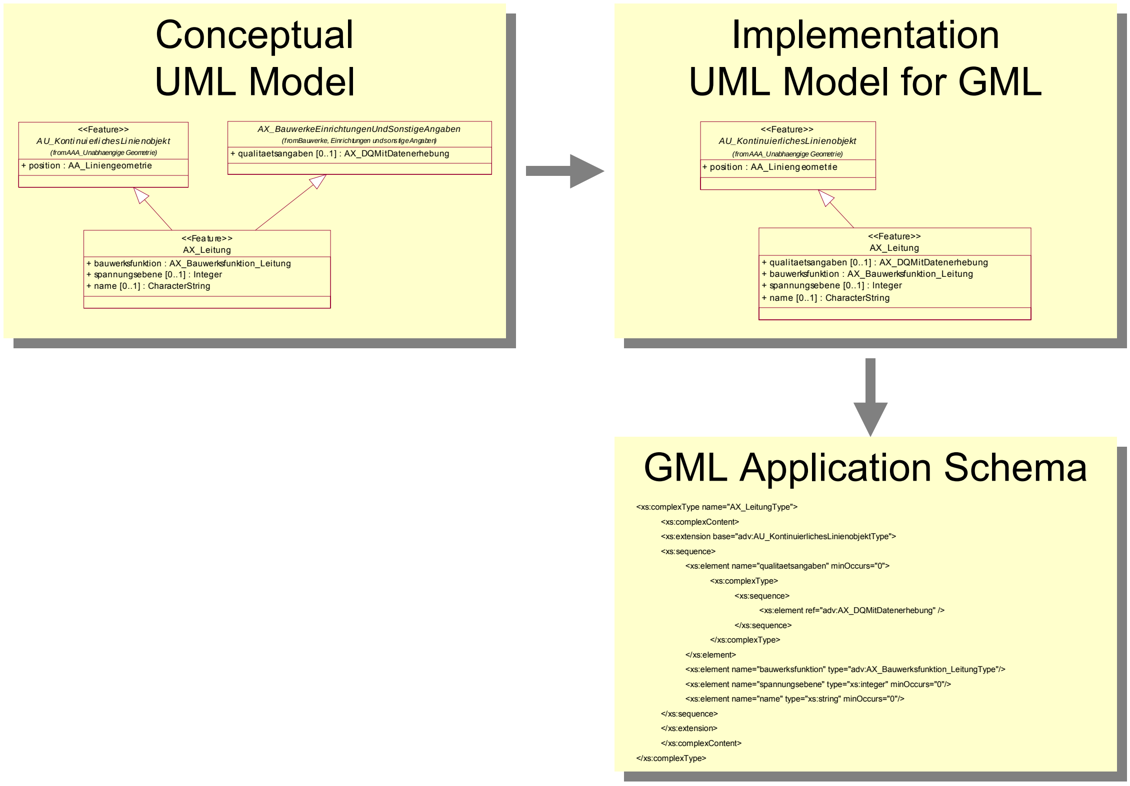

Figure 3 – Process of creating the GML application schema

Recommendation 12	If a transformation from the application schema on the conceptual level to the UML implementation profile from which the GML application schema is automatically derived is needed, it should follow a common set of rules across all themes.

The Generic Conceptual Model defines a UML profile which includes elements that are outside of the UML profile specified by ISO 19136 Annex E and ISO/TS 19139, for which those standards provide conversion rules to GML/XML.

For Annex I data specifications, no requirements for such conversions from the conceptual to an implementation model were identified. This may change in future revisions of these application schemas or in the data specifications for the Annex II and Annex III spatial data themes.

B.3.3 Extensions to the GML encoding rule

B.3.3.1 Properties with stereotype <<voidable>>

Requirement 5 Attributes and association roles with the stereotype <<voidable>> shall be converted to XML Schema as if the stereotype were ignored - except that the content model of the property element shall receive two additional optional attributes:

- The global attribute “xsi:nil” (specified by XML Schema); in the schema this is expressed by an attribute “nillable” with the value “true”.
- A local, unqualified attribute “nilReason” with the type “gml:NilReasonType”.

See GML 8.2.3.1 and 8.2.3.2 for more details.

- It has been commented in the stakeholder consultation of the Annex I data specifications, that this encoding of <<voidable>> causes substantial overhead in the GML encoding as most types have several voidable properties. In existing data, these will often not be available (this is why they were tagged as voidable), i.e. in the XML encoding several property elements per spatial object will often be represented with xsi:nil="true". This increases the size of encoded spatial object by a substantial amount.

The following options were considered:
- In the encoding rule, change Requirement 5 as follows: "Attributes and association roles with the stereotype <<voidable>> shall be converted to XML Schema as if the stereotype were ignored and the minimum multiplicity of the property shall be '0'." This approach would drop the distinction in this encoding between the two types of "no data" described in the Generic Conceptual Model 9.4.5: The characteristic is not present or not applicable in the real world or the characteristic is not present in the spatial object, but may be present or applicable in the real world.
- As an extension to the previous option, it could be considered to provide information about properties that are void throughout the spatial data set in the data set metadata. This also recognises that often the same properties will be void throughout the spatial data set.

In the comment resolution process it was decided to keep the existing rule, but to flag this issue as an open issue that should be addressed in collaboration with the Member States during the preparation of the implementation of the implementing rule on the interoperability of spatial data sets and services in the Member States.

The reasons for this decision are:
- The current requirement uses the XML encoding for such void values as foreseen in XML Schema. Before deviating from the standards, the practical need for a special rule should be established first objectively. The Drafting Team did not have sufficient information available to justify a change.
- As long as there is an agreement that the different cases of "no value" should be distinguishable also in the data, the first option described above is insufficient. If, however, it were considered more important to have a more compact encoding than being able to distinguish the cases, then this might be an option to simplify the encoding in such a way.
- Adding information about void properties in the dataset metadata is not a good idea either as it adds additional out-of-band information that most applications will not access or not even know where to find. Making even more relevant information undiscoverable to a standard client is not considered the right direction for making spatial data available on the internet. Furthermore, it is unclear how such an extension to the spatial data set metadata could be brought inline with the metadata regulation of INSPIRE.

B.3.3.2 Properties with stereotype <<lifeCycleInfo>>

Requirement 6	The stereotype <<lifeCycleInfo>> shall be ignored.

NOTE The stereotype is ignored as it has no impact on the XML Schema.

B.3.3.3 Properties with stereotype <<version>>

Requirement 7	The stereotype <<version>> shall be ignored.

NOTE The stereotype is ignored as it has no impact on the XML Schema, only in the encoding of the reference to the target object (version) in the instances. See B.4.3.

B.3.3.4 Association classes

Requirement 8	Association classes shall be encoded according to the conversion rule proposed in GML Change Request specified in OGC document 08-109[6].

This example below illustrates the conversion of association classes to types with associations to both ends of the original association. Figure 4 shows the schema on the conceptual level, Figure 5 the schema on the implementation level.

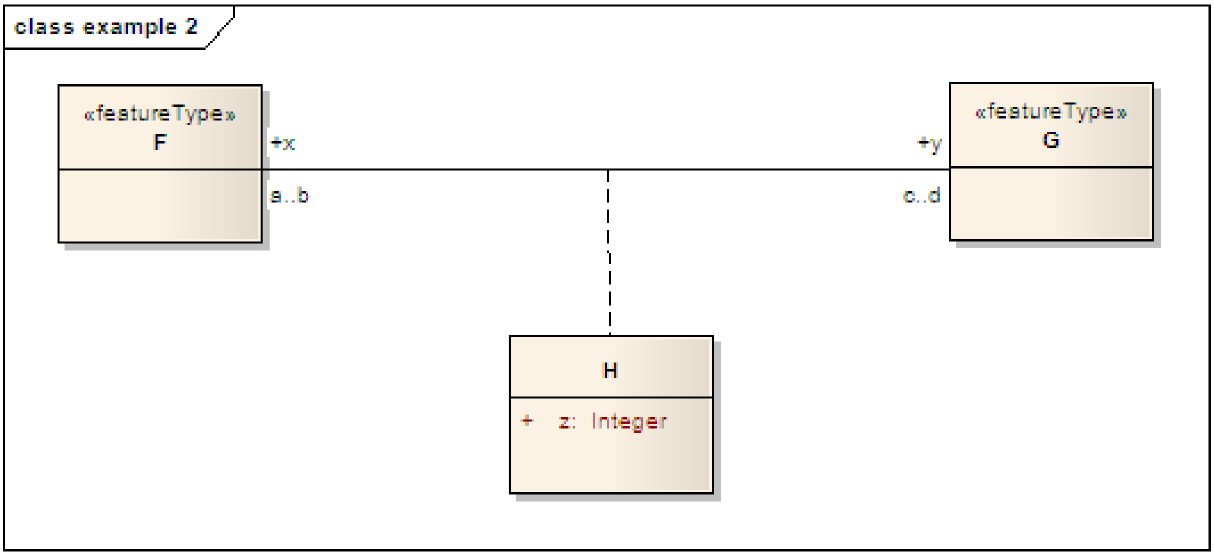

Figure 4 – Types in an application schema

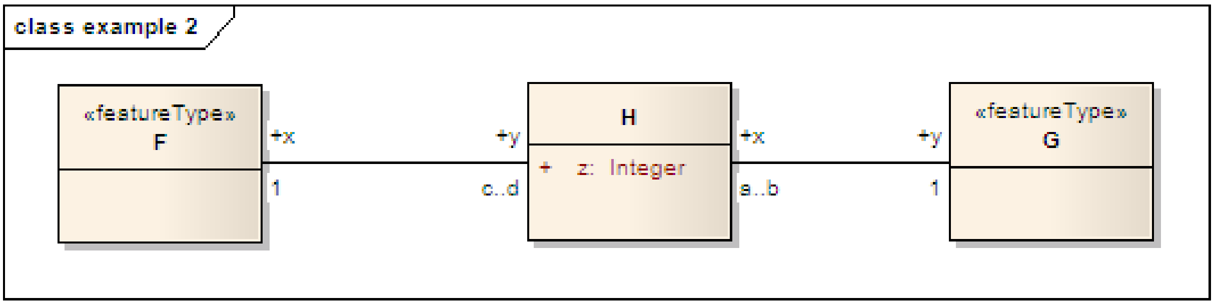

Figure 5 – Equivalent representation to be used as the basis for the encoding of the GML application schema

B.3.3.5 Mixin classes

Due to the fact that several implementation platforms including XML Schema supports only type derivation from a single base type (element substitutablity in XML Schema is restricted to a single element, too), the use of multiple inheritance is currently not supported by GML 3.2.1 Annex E.

However, for conceptual modelling, the ability to define abstract types which capture a set of properties that are associated with a concept is often appropriate and convenient.

[6] https://portal.opengeospatial.org/files/?artifact_id=29328

Requirement 9 Mixin classes shall have either the stereotype <<featureType>> or no stereotype and the tagged value “gmlMixin” with a value “true”.

Mixin classes shall not be encoded as object elements. Instead the properties of a mixin class shall be encoded as if they were specified locally in each direct subtype that is not a mixin class.

In addition, a property type shall be created if the tagged value “noPropertyType” is set to “false”. The content model of the property type shall use the standard pattern for property types with the difference that <choice> is used instead of <sequence> and the choice is a choice between all object elements of direct subtypes of the mixin class.

EXAMPLE In Figure 6 HydroObject is a mixin class.

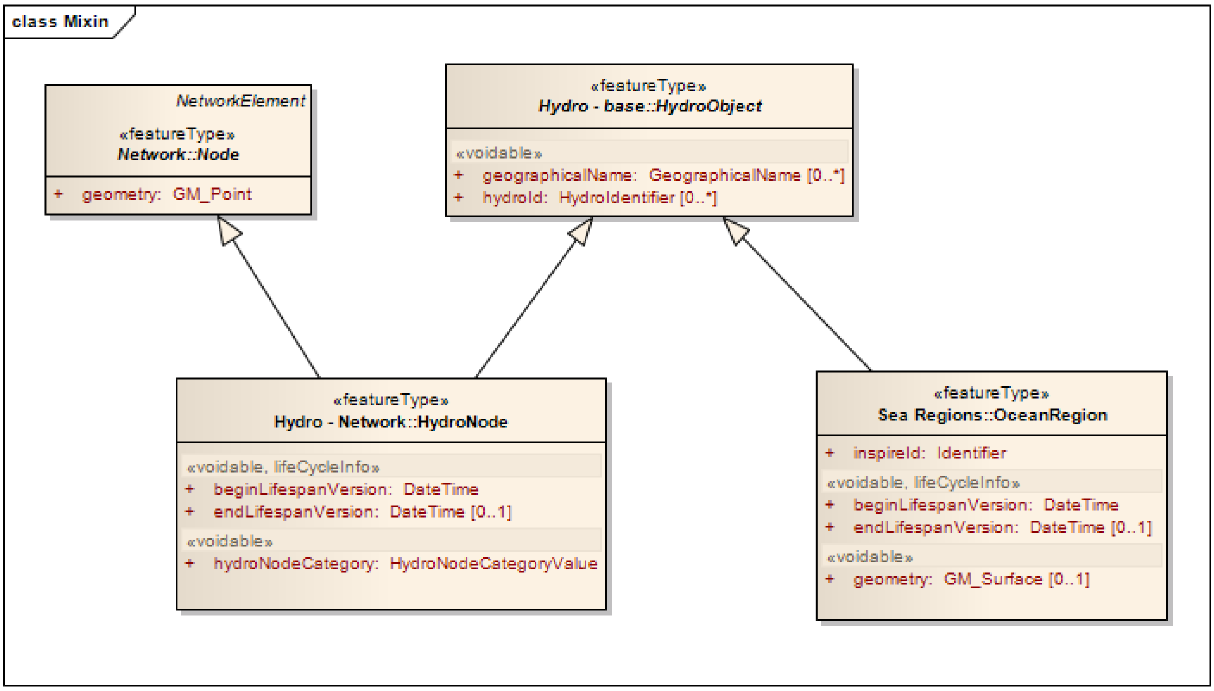

Figure 6 – Mixin example

The subtypes HydroNode and OceanRegion are encoded in XML Schema according to the rule as:

```
<element name="HydroNode" substitutionGroup="net:Node" type="hy-n:HydroNodeType"/>
<complexType name="HydroNodeType">
  <complexContent>
   <extension base="net:NodeType">
    <sequence>
     <element maxOccurs="unbounded" minOccurs="0" name="geographicalName" nillable="true">
      <complexType>
       <sequence>
        <element ref="gn:GeographicalName"/>
       </sequence>
       <attribute name="nilReason" type="gml:NilReasonType"/>
      </complexType>
     </element>
     <element maxOccurs="unbounded" minOccurs="0" name="hydroId" nillable="true">
      <complexType>
       <sequence>
        <element ref="hy:HydroIdentifier"/>
       </sequence>
       <attribute name="nilReason" type="gml:NilReasonType"/>
```

```
          </complexType>
        </element>
        <element maxOccurs="unbounded" minOccurs="0" name="relatedHydroObject" nillable="true"
type="gml:ReferenceType"/>
        <element name="beginLifespanVersion" nillable="true">
          <complexType>
            <simpleContent>
              <extension base="dateTime">
                <attribute name="nilReason" type="gml:NilReasonType"/>
              </extension>
            </simpleContent>
          </complexType>
        </element>
        <element minOccurs="0" name="endLifespanVersion" nillable="true">
          <complexType>
            <simpleContent>
              <extension base="dateTime">
                <attribute name="nilReason" type="gml:NilReasonType"/>
              </extension>
            </simpleContent>
          </complexType>
        </element>
        <element name="hydroNodeCategory" nillable="true">
          <complexType>
            <simpleContent>
              <extension base="gml:CodeType">
                <attribute name="nilReason" type="gml:NilReasonType"/>
              </extension>
            </simpleContent>
          </complexType>
        </element>
      </sequence>
    </extension>
  </complexContent>
 </complexType>
 <complexType name="HydroNodePropertyType">
   <sequence minOccurs="0">
     <element ref="hy-n:HydroNode"/>
   </sequence>
   <attributeGroup ref="gml:AssociationAttributeGroup"/>
   <attributeGroup ref="gml:OwnershipAttributeGroup"/>
 </complexType>

<element name="OceanRegion" substitutionGroup="gml:AbstractFeature"
 type="sr:OceanRegionType"/>
<complexType name="OceanRegionType">
   <complexContent>
     <extension base="gml:AbstractFeatureType">
       <sequence>
        <element maxOccurs="unbounded" minOccurs="0" name="geographicalName" nillable="true">
          <complexType>
            <sequence>
              <element ref="gn:GeographicalName"/>
            </sequence>
            <attribute name="nilReason" type="gml:NilReasonType"/>
          </complexType>
        </element>
        <element maxOccurs="unbounded" minOccurs="0" name="hydroId" nillable="true">
          <complexType>
            <sequence>
```

```
            <element ref="hy:HydroIdentifier"/>
          </sequence>
          <attribute name="nilReason" type="gml:NilReasonType"/>
        </complexType>
      </element>
      <element maxOccurs="unbounded" minOccurs="0" name="relatedHydroObject" nillable="true"
type="gml:ReferenceType">
      </element>
      <element name="beginLifespanVersion" nillable="true">
        <complexType>
          <simpleContent>
            <extension base="dateTime">
              <attribute name="nilReason" type="gml:NilReasonType"/>
            </extension>
          </simpleContent>
        </complexType>
      </element>
      <element minOccurs="0" name="endLifespanVersion" nillable="true">
        <complexType>
          <simpleContent>
            <extension base="dateTime">
              <attribute name="nilReason" type="gml:NilReasonType"/>
            </extension>
          </simpleContent>
        </complexType>
      </element>
      <element minOccurs="0" name="geometry" nillable="true" type="gml:SurfacePropertyType"/>
      <element name="inspireId" type="base:IdentifierPropertyType"/>
      <element maxOccurs="unbounded" minOccurs="0" name="foreshore" nillable="true"
type="gml:ReferenceType"/>
    </sequence>
   </extension>
  </complexContent>
 </complexType>
 <complexType name="OceanRegionPropertyType">
  <sequence minOccurs="0">
   <element ref="sr:OceanRegion"/>
  </sequence>
  <attributeGroup ref="gml:AssociationAttributeGroup"/>
  <attributeGroup ref="gml:OwnershipAttributeGroup"/>
 </complexType>
```

In addition, a property type is created for HydroObject:

```
<complexType name="HydroObjectPropertyType">
  <choice minOccurs="0">
   <element ref="hy-n:HydroNode"/>
   <element ref="sr:OceanRegion"/>
   ... elements of other subtypes
  </choice>
  <attributeGroup ref="gml:AssociationAttributeGroup"/>
  <attributeGroup ref="gml:OwnershipAttributeGroup"/>
 </complexType>
```

B.3.4 Restrictions on options in the GML encoding rule

The intention of the following recommendations is to harmonise the encoding style across the different data specifications where the GML encoding rule leaves encoding options.

Recommendation 13 All navigable feature association roles should be assigned a tagged value "inlineOrByReference" with the value "byReference".

NOTE 1 The result of this particular recommendation is that features are not embedded in other features in XML documents but that they are all first level objects in a feature collection. An example where this recommendation would in general be ignored are complex spatial objects that own their parts.

Requirement 10 All INSPIRE code lists shall be assigned a tagged value "asDictionary" with the value "true". Instance data shall reference a GML dictionary that encodes the valid value in the code list.

NOTE 2 As a result, the code list values are managed only in registries outside of the application schema. All instances have to reference the regsitry to allow applications to evaluate the code list value and its validity, its title in the official languages or its definition.

NOTE 3 If a code list relevant for an INSPIRE data specification is already maintained by an international organisation, then INSPIRE will simply adopt that code list and reference it. However, if the international organisation does not provide a representation of the code lists as a GML dictionary it may be required to maintain such a representation in an INSPIRE registry.

B.3.5 XML namespaces

Recommendation 14 The target namespace of the GML application schema should be a URI of the form

http://inspire.jrc.ec.europa.eu/schemas/<code>/<version>

where <code> is a unique, persistent short code of the GML application schema in INSPIRE and <version> identifies the particular version. The <code> value shall not include a colon.

See also Issue 1, for now this is a preliminary URI pattern.

EXAMPLE Version 3.1 of the GML application schema for the INSPIRE application schema for the spatial data theme “Addresses” will use “http://inspire.jrc.ec.europa.eu/schemas/ad/3.1” as the target namespace.

NOTE 1 This is required to provide a common approach to the use of XML namespaces across the different data specifications.

NOTE 2 For cases where existing XML encodings are adopted by INSPIRE this recommendation does not apply. However, where application schemas are originally specified in INSPIRE, the namespaces will follow the recommendation above.

B.4 Instance conversion

B.4.1 Character encoding

Requirement 11 XML documents shall be encoded using UTF-8 as character encoding.

NOTE Following this recommendation ensures that all linguistic texts can be encoded in any language – which in turn simplifies processing of data. The use of UTF also aligns with common practice and is the default character encoding for XML documents.

B.4.2 Encoding of an external object identifier

Recommendation 15 URIs of spatial objects should be persistent http URIs and include the namespace and the local identifier part of the INSPIRE identifier.

EXAMPLE 1 See the example in B.2.

NOTE The Generic Conceptual Model specifies constraints on the lexical space of the namespace, local identifier and version values.

Recommendation 16 The external object identifier should be encoded in a gml:identifier property of the feature with the codeSpace attribute set to http://inspire.jrc.ec.europa.eu/ids.

Issue 2 A description of identifiers in INSPIRE, and links to descriptions of the URI schemes of Member States and the European Commission should be available at the URI from the recommendation above.

EXAMPLE 2

```
<gml:identifier codeSpace="http://inspire.jrc.ec.europa.eu/ids">
        http://location.data.gov.uk/doc/00BH/123456789012
</gml:identifier>
```

The information from the external object identifier may also be used for the gml:id qualifier.

However, while the XML 1.0 standard allows colons in XML ID, this is prohibited in the Namespaces in XML 1.0 standard. Since XML namespaces are used in INSPIRE XML documents, a different separator between the identifier parts has to be used.

In addition, since XML ID values only have to be unique within a document, or in the case of an OGC Web Feature Service within the opaque datastore of the service, the identifier namespace may be omitted if only spatial objects from a single identifier namespaces are provided by a download service.

EXAMPLE 3 In this example, the identifier namespace from the identifier [namespace=”DEAAA”, local id=”DENW123412345678”, version=”2008-09-12T12:34:56Z”] is dropped and since the naming system uses a fixed length for the local identifier the version is appended without separator:

```
<cp:CadastralParcel gml:id=”DENW12341234567820080912T123456Z”>
        ...
</cp:CadastralParcel>
```

XML ID does not allow digits, points and dashes as start characters. In naming systems where this may occur this may be circumvented, for example, by adding a leading underscore character.

NOTE 4 This will not occur in cases where the identifier namespace is included as it will always start with a letter or an underscore.

EXAMPLE 4 With namespace (if the Web Feature Service hosts spatial objects from multiple namespaces) and a separator “__” the spatial object from example 3 would be encoded as

```
<cp:CadastralParcel gml:id=”DEAAA__DENW123412345678__20080912T123456Z”>
        ...
</cp:CadastralParcel>
```

B.4.3 Encoding of a reference to a spatial object

Recommendation 17 To reference a spatial object or a specific version of a spatial object its URI (see Recommendation 15) should be used.

EXAMPLE A reference to a cadastral parcel object with the identifier used in the example in B.4.2 could be encoded as (using a hypothetical URI scheme)

```
<parcel xlink:href="http://location.nrw.de/doc/DEAAA/NW123412345678"/>
```

This assumes that the reference is to the spatial object and not a specific version of the spatial object. In case of a reference to the specific version, the reference would be encoded as

```
<parcel xlink:href="http://location.nrw.de/doc/DEAAA/NW123412345678/20080912T123456Z"/>
```

B.4.4 Encoding of a reference to a registered item

Recommendation 18 URIs should be used to encode item identifiers of items in ISO 19135-compliant registers and to reference such items.

The URIs should use the following structure:

http://inspire-registry.jrc.ec.europa.eu/<register>/<item class>/<item identifier>

where
- <item class> is the name of the item class (ISO 19135: RE_ItemClass) of the registered item;
- <register> is the name of the register (ISO 19135: RE_Register);
- <item identifier> is the item identifier of the registered item (ISO 19135: RE_RegisterItem).

Colons should not be used in <item class>, <register> or <item identifier> values.

See also Issue 1, for now this is a preliminary URI pattern and implemented at this time.

Other URIs may be used, too, for items managed in external registers. For registers with stable URIs as identifiers, these URIs may be used instead. For external registers that do not provide stable URIs, the “inspire”-URIs will be used, but the <register> will identify the external register and authority.

In principle, INSPIRE will adopt existing registers and their items whenever possible as long as they are applicable to INSPIRE, managed in a structured way by a competent international organisation, and are accessible from the INSPIRE SDI. In this context, several aspects will have to be considered/addressed:

- Organisational: The adoption of a third-party register creates a dependency and requires an agreement between INSPIRE and the other organisation. This includes intellectual property rights aspects, notification of changes, etc.
- Operational: To address the interoperability requirements there must be ways how users of INSPIRE data will be able to understand items managed in the external registers. This may require setting up registries for such registers to make these items available within INSPIRE. This will require further work.
- Technical: To allow referencing items in external registers, the URI scheme used in D2.7 will be adapted to allow this.

EXAMPLE 1 http://inspire-registry.jrc.ec.europa.eu/IFCD/featureConcept/125 would be a reference to a feature concept in the INSPIRE Feature Concept Dictionary. The item identifier 125 is unique and identifies the item within the register. It should be noted that the name of the feature concept cannot be used as the item identifier as over time the same name may be associated with different concepts; e.g., when the definition of a feature concept is amended.

EXAMPLE 2 OGC provides a URN namespace to reference coordinate reference systems, e.g., urn:ogc:def:crs:EPSG::4258 would be a reference to the geographic coordinate reference system ETRS89 in the EPSG coordinate reference system register.

This applies to all items that will be managed in registers: terms, feature types, coordinate reference systems, coordinate operations, units of measurements, identifier namespaces, application schemas, etc. The list of item classes will be compiled during the data specification process.

NOTE 1 The intention of this recommendation is to harmonise the encoding style across the different data specifications where the GML encoding rule leaves freedom.

NOTE 2 The use of other URIs is intended to support the usage of URIs that are already commonly used and supported by software components or specified in other standards.

B.5 Implementation aspects <informative>

In encoding services, conversion rules have to be implemented at both the schema and the instance level (see 6.3.6):

- At the schema level, the conversion rules derive a GML application schema from the INSPIRE application schema in UML. This is done once - the resulting GML application schema is specified as part of the INSPIRE data specification.

- At the instance level the implementation will usually be system-specific, usually integrated into download services. The required software components encoding data structured according to an application schema, e.g. in a database, in XML according to the corresponding GML application schemas and vice versa are available in the market.

Annex C
(informative)

Encoding of file-based data

C.1 Overview

For many applications and INSPIRE themes, data typically are persisted in relational databases. Numerous implementations have demonstrated the feasibility of transforming such data to standardised encodings corresponding to common GML application schemas. However, this is not universally the case. Especially for coverage-based data[7], and in some environmental themes (e.g. ‘Orthoimagery’, ‘Oceanographic geographical features’, ‘Meteorological geographical features’), file-based data storage is more common.

For file-based data, there is no best practice solution for integration within a spatial data infrastructure based on the standards referenced from INSPIRE guidance documents. Further research is needed, but this annex outlines an approach that is proposed as the starting point for such work.

C.2 Requirements

C.2.1 Need to retain existing efficient file formats

In many cases, the reason for file-based storage is data volume – space-borne remote-sensing imagery, or computer simulation (e.g. weather forecast) datasets may be many terabytes in size. It is impractical to transform such data to an XML-based encoding[8]. For operational reasons, it is necessary to retain existing efficient file-based encodings for such data.

C.2.2 Need for interoperable representation

However, in order to establish cross-theme and cross-dataset interoperability for sharing these data, as required in the INSPIRE Directive, it is necessary also to provide a representation conforming to the model-driven approach outlined in the “Generic Conceptual Model” (D2.5).

C.2.3 Conceptual modelling unconstrained by file formats

For relational data, it is already best practice to develop conceptual models independent of existing storage structures (application schema design is independent of database schema). Indeed, this approach is encapsulated by the INSPIRE “Methodology for the development of data specifications” (D2.6). This should also be true for file-based data. For instance, an application schema should not merely reflect the physical structure of file formats – spatial object types should rather reflect more conceptual real-world objects.

C.2.4 Ability to model file extracts

A corollary of this requirement is that a conceptual model and interoperable representation should be able to describe *extracts* of files where necessary.

[7] The rules for the use of coverages in INSPIRE application schemas are discussed in the Generic Conceptual Model, in particular in sub-clause 10.4 "Rules for use of coverage functions" and the last paragraph in sub-clause 9.2.3.

[8] Work underway by W3C on an *Efficient XML Interchange Format* (http://www.w3.org/TR/2008/WD-exi-20080919/) may change this in the future.

EXAMPLE The HDF-EOS file format widely used in the earth sciences supports multiple spatial object types (point, grid, and swath) – a single file may contain any of these objects, in any number. However, an application schema may need to define only swath-type spatial objects (with the interoperable representation thereby constrained to refer only to swath objects within a file).

C.2.5 Ability to integrate logical structures across files

Similarly, individual files may contain only a partial view of a complete dataset. It is necessary for an application schema and interoperable representation to be able to describe logical structures that are integrated *across* multiple individual files.

EXAMPLE 1 Separate files may contain northern and southern hemispheres of a coverage dataset with global extent.

EXAMPLE 2 Individual calendar months of a two-year long tidal time series dataset may be stored in separate files.

EXAMPLE 3 The WMO 'GRIB' format provides independent two-dimensional horizontal slices of atmospheric parameters that, in the real world, may be three-dimensional (or four-dimensional, if the temporal dimensional is also taken into account).

C.3 Proposed interoperability model – joint representation

The ISO/DTS 19129 ("Geographic information - Imagery, gridded and coverage data framework") distinguishes between the logical *content* of a coverage or imagery dataset, the *encoding* defined by an exchange file format, and the need for mappings between them:

> *Most of the existing interchange standards relating to imagery, gridded and coverage data describe information in terms of its representation in an interchange format. The format defines data fields and describes the contents and meaning of these data fields. This implicitly defines the information content that can be carried by this interchange format. Some of the existing standards even separate their "information" from the encoding within the description of the standard, but in the end it is the encoding that defines these standards. The common content model defined in this Technical Specification allows for a mapping to the structures defined in the encoding standards.* (ISO/TC 211 N 2547, DTS 19129, "Geographic information - Imagery, gridded and coverage data framework")

Moreover, it recognises the role that GML can play in mediating between the logical content and the physical exchange format, Figure 1.

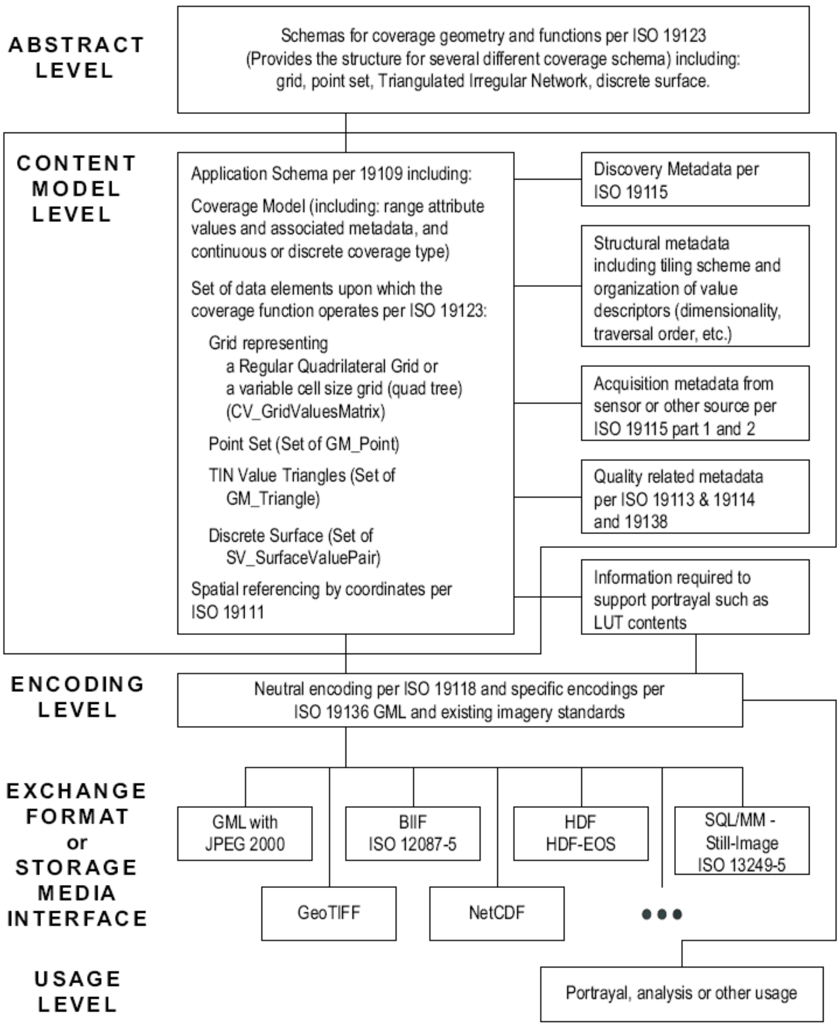

Figure 1: GML may provide a neutral *encoding* to mediate between logical *content* and physical *exchange* formats for file-based datasets [ISO/DTS 19129]

The interoperability model for INSPIRE file-based datasets requires an application schema for describing the logical structure and semantic content of a dataset, while retaining efficient data files themselves during exchange.

The INSPIRE default encoding for file-based datasets therefore consists of three components:

1. a GML application schema in accordance with Annex B of these Guidelines (providing a conceptual dataset 'skeleton')
2. one or more 'legacy' data files (containing the 'flesh' of the dataset)
3. a mapping from the GML representation to those files or file extracts providing the logical content

This approach effectively provides a 'GML wrapper' to data files.

EXAMPLE This joint representation is analogous to that used in the GMLJP2 specification.

C.4 Implementation

The general interoperability model described above does not imply a particular implementation of the mapping between a GML application schema and data files. However, GML itself provides a 'by-reference' pattern for object properties that may be used for a level of indirection between inline encoding of a property value and an external resource:

> *the value of the property is available elsewhere, and is identified by the value of an xlink:href attribute on the property element* (ISO 19136, §7.2.3.1)

Using this mechanism, the full richness of the Xlink specification may be exploited for referencing file-based content from within GML. This is described in GML §8.1, though the normative Xlink specification is available from W3C. Two additional xlink attributes provide important context:

> xlink:role *description of the nature of the target resource, given as a URI*
> xlink:arcrole *description of the role or purpose of the target resource in relation to the present resource, given as a URI*
> (ISO 19136, §8.1)

Thus, an implementation may use xlink attributes as follows:

- *href*: for referencing one or more files, file extracts, or aggregated file structures providing the logical content for a GML property element. (E.g. a file name or URI)
- *role*: for specifying the nature of the referenced file resource(s). (E.g. specifying the file format)
- *arcrole*: for asserting that the file resource(s) provide logical content for the referring GML element

C.5 Issues

While the above approach provides a general framework for interoperability of file-based data within INSPIRE, various elements of the implementation require further work.

C.5.1 Complex transformation semantics

In general, there may be a many-to-many relationship between spatial objects and file instances (C.2.4, C.2.5). The simplest case is where an entire file provides directly the logical content for a single GML property. An example might be an image file providing the range attributes for a gridded coverage. In this case, the xlink:href may refer directly to a single file. However, there is likely a need for richer mediation semantics in order to refer to file extracts, or to logical structures aggregated across multiple files. Third-party schemas may play a role as 'virtual' file formats[9].

C.5.2 GML properties with simple content

Not all GML properties support the 'by-reference' pattern. In particular, GML properties of 'simple XML type' generally use inline encoding. It needs to be determined whether such properties also require an indirection mechanism for file-based content as described above, and how it could be implemented.

C.5.3 Pilot project

The proposed mechanism has been explored only briefly (e.g. OGC 07-083), and there is need for a pilot project to investigate the adequacy and feasibility of this approach. Such a pilot should involve interested stakeholders from one or more relevant INSPIRE themes, and aim to develop best-practice guidelines, perhaps in collaboration also with OGC.

---

[9] Examples include CDML (http://esg.llnl.gov/cdat/cdms_html/cdms-6.htm), NcML (http://www.unidata.ucar.edu/software/netcdf/ncml/), and the CSML (http://ndg.nerc.ac.uk/csml, now http://csml.badc.rl.ac.uk) StorageDescriptor.

Bibliography

This clause lists documents that were used in the drafting of this document, but which are not normative.

CEN/TR 15449:2006, Geographic information — Standards, specifications, technical reports and guidelines, required to implement Spatial Data Infrastructure

EN ISO 19101:2005, Geographic information — Reference model

ISO/TS 19103:2005, Geographic Information — Conceptual schema language

NOTE A revision of the standard is in progress and currently has a target date of December 2009 for reaching "Draft International Standard" status. To the extent possible, alignment of this document with the future version has been taken into consideration.

EN ISO 19109:2006, Geographic Information — Rules for application schemas

EN ISO 19115:2005, Geographic information — Metadata

ISO 19115:2003/Cor 1:2006, Geographic information — Metadata — Technical Corrigendum 1

ISO/TS 19129:2009, Geographic information — Imagery, gridded and coverage data framework

EN ISO 19135:2007, Geographic information — Procedures for item registration

UML 2.1.2, Unified Modelling Language (UML) Superstructure and Infrastructure, Version 2.1.2

UK Location Council & Public Sector Information Domain of the CTO Council s cross Government Enterprise Architecture, Designing URI Sets for Location, Version 0.4, April 2010