INTERNATIONAL STANDARD

ISO 19118

First edition
2005-07-15

Geographic information — Encoding

Information géographique — Codage

Reference number
ISO 19118:2005(E)

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

Contents

Page

Figures

Tables

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

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

ISO 19118 was prepared by Technical Committee ISO/TC 211, *Geographic information/Geomatics*.

Introduction

This International Standard specifies the requirements for defining encoding rules to be used for interchange of geographic data within the ISO 19100 series of International Standards. An encoding rule allows geographic information defined by application schemas and standardized schemas to be coded into a system-independent data structure suitable for transport and storage. The encoding rule specifies the types of data to be coded and the syntax, structure and coding schemes used in the resulting data structure. The resulting data structure may be stored on digital media or transferred using transfer protocols. It is intended to be read and interpreted by computers, but may be in a form that is human readable.

The choice of one encoding rule for neutral data interchange does not exclude application domains and individual nations from defining and using their own encoding rules that can be platform dependent or more effective with regards to data size or processing complexity. XML is a subset of ISO/IEC 8879 and has been chosen because it is independent of computing platform and interoperable with the World Wide Web.

This International Standard is divided into three logical sections. The requirements for creating encoding rules based on UML schemas are specified in Clauses 6 to 8. The requirements for creating encoding service are specified in Clause 9, and the XML based encoding rule is specified in Annex A.

The XML based encoding rule is intended to be used for neutral data interchange. It relies on the Extensible Markup Language (XML) and the ISO/IEC 10646 character set standards. Introductions to XML and ISO/IEC 10646 are given in Annexes C and D, respectively. Annex E contains examples of the application of this International Standard.

The geographic information standards are organized in the ISO 19100 series of International Standards. The background, the overall structure of this series of International Standards and the fundamental description techniques are defined in ISO 19101, ISO/TS 19103 and ISO 19104.

Users of this International Standard will develop application schemas to capture the semantics of geographic information. An application schema is compiled by integrating elements from a set of standardized conceptual schemas developed in ISO 19107, ISO 19108, ISO 19110, ISO 19111, ISO 19112, ISO 19113, ISO 19115 and ISO 19117, including eventually new standardized conceptual schemas. How this integration will take place is described in ISO 19109. The ISO 19100 series of International Standards also defines a set of common services that shall be available when developing geographic information applications. The common services are generally defined in ISO 19119 and will cover access to and processing of geographic information according to the common information model. Two service areas are defined more closely in ISO 19116 and ISO 19117. ISO 19105, ISO 19106, ISO 19114 and this International Standard cover implementation issues.

Geographic information — Encoding

1 Scope

This International Standard specifies the requirements for defining encoding rules to be used for interchange of geographic data within the ISO 19100 series of International Standards.

This International Standard specifies

— requirements for creating encoding rules based on UML schemas,

— requirements for creating encoding services,

— an informative XML based encoding rule for neutral interchange of geographic data.

This International Standard does not specify any digital media, it does not define any transfer services or transfer protocols, nor does it specify how to encode inline large images.

2 Conformance

Two conformance levels are defined for this International Standard. The conformance levels are defined in the abstract test suite in Annex B.

3 Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC 10646, *Information technology — Universal Multiple-Octet Coded Character Set (UCS)*

ISO 8601, *Data elements and interchange formats — Information interchange — Representation of dates and times*

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

ISO 19109:2005, *Geographic information — Rules for application schema*

Extensible Markup Language (XML) 1.0 (Second Edition), W3C Recommendation 6 October 2000. Available at <http://www.w3.org/TR/REC-xml>

XML Schema Part 1: Structures, W3C Recommendation 2, May 2001. Available at <http://www.w3.org/TR/xmlschema-1/>

XML Schema Part 2: Datatypes, W3C Recommendation 2, May 2001. Available at <http://www.w3.org/TR/xmlschema-2/>

4 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

4.1
application schema
conceptual schema for **data** required by one or more applications

[ISO 19101]

NOTE An application schema describes the content, the structure and the constraints applicable to information in a specific application domain.

4.2
character
member of a set of elements that is used for the representation, organization, or control of **data**

[ISO/IEC 2382-1]

4.3
code
representation of a label according to a specified scheme

4.4
conceptual model
model that defines concepts of a **universe of discourse**

[ISO 19101]

4.5
conceptual schema
formal description of a **conceptual model**

[ISO 19101]

4.6
conceptual schema language
formal language based on a conceptual formalism for the purpose of representing **conceptual schemas**

[ISO 19101]

EXAMPLE UML, EXPRESS, IDEF1X.

NOTE A conceptual schema language may be lexical or graphical.

4.7
conversion rule
rule for converting instances in the input **data** structure to instances in the output **data** structure

4.8
data
reinterpretable representation of information in a formalized manner suitable for communication, interpretation, or processing

[ISO/IEC 2382-1]

4.9
data element
unit of **data** that, in a certain context, is considered indivisible

4.10
data interchange
delivery, receipt and interpretation of **data**

4.11
data transfer
movement of **data** from one point to another over a **medium**

NOTE Transfer of information implies transfer of data.

4.12
data type
specification of a **value domain** with operations allowed on values in this domain

[ISO/TS 19103]

EXAMPLE Integer, Real, Boolean, String and Date.

NOTE A data type is identified by a term, e.g. Integer. Values of the data types are of the specified value domain, e.g. all integer numbers between –65537 and 65536. The set of operations can be +, –, ÷ and × and is semantically well defined. A data type can be simple or complex. A simple data type defines a value domain where values are considered atomic in a certain context, e.g. Integer. A complex data type is a collection of data types which are grouped together. A complex data type may represent an object and can thus have identity.

4.13
dataset
identifiable collection of **data**

[ISO 19115]

4.14
encoding
conversion of **data** into a series of **codes**

4.15
encoding rule
identifiable collection of **conversion rules** that define the **encoding** for a particular **data** structure

EXAMPLE XML, ISO 10303-21, ISO/IEC 8211.

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.

4.16
encoding service
software component that has an **encoding rule** implemented

4.17
feature
abstraction of real world phenomena

[ISO 19101]

NOTE A feature may occur as a type or an instance. Feature type or feature instance is used when only one is meant.

4.18
file
named set of records stored or processed as a unit

[ISO/IEC 2382-1]

4.19
geographic data
data with implicit or explicit reference to a location relative to the Earth

[ISO 19109]

4.20
geographic information
information concerning phenomena implicitly or explicitly associated with a location relative to the Earth

[ISO 19101]

4.21
identifier
label that uniquely identifies an item or group of items

4.22
information
knowledge concerning objects, such as facts, events, things, processes, or ideas, including concepts, that within a certain context has a particular meaning

[ISO/IEC 2382-1]

4.23
instance model
representation **model** for storing **data** according to an **application schema**

4.24
interface <UML>
named set of operations that characterize the behaviour of an element

[ISO/IEC 19501]

4.25
interoperability
capability to communicate, execute programs, or transfer **data** among various functional units in a manner that requires the user to have little or no knowledge of the unique characteristics of those units

[ISO/IEC 2382-1]

4.26
medium
substance or agency for storing or transmitting **data**

EXAMPLE Compact disc, internet [1], radio waves, etc.

4.27
model
abstraction of some aspects of reality

[ISO 19109]

4.28
schema
formal description of a **model**

[ISO 19101]

4.29
schema model
representation **model** for storing **schemas**

EXAMPLE Representation model for a schema repository.

4.30
stereotype <UML>
new type of modelling element that extends the semantics of the metamodel

[ISO/IEC 19501]

NOTE Stereotypes must be based on certain existing types or classes in the metamodel. Stereotypes may extend the semantics, but not the structure of pre-existing types and classes. Certain stereotypes are predefined in the UML, others may be user defined. Stereotypes are one of three extensibility mechanisms in UML; the others are constraint and tagged value.

4.31
transfer protocol
common set of rules for defining interactions between distributed systems

4.32
universe of discourse
view of the real or hypothetical world that includes everything of interest

[ISO 19101]

4.33
value domain
set of accepted values

[ISO/TS 19103]

EXAMPLE The range 3-28, all integers, any ASCII character, enumeration of all accepted values (green, blue, white).

5 Symbols and abbreviated terms

API	Application Programming Interface
DTD	Document Type Definition
ID	Identifier
IDREF	An XML ID reference type
OMG	Object Management Group
UCS	Universal Multiple-Octet Coded Character Set
UML	Unified Modelling language
URI	Uniform Resource Identifier
UTF	UCS Transfer Format
UUID	Universal Unique Identifier
XMI	XML Metadata Interchange

XML	Extensible Markup Language
XPointer	XML Pointer language
XSD	XML Schema Document

6 Fundamental concepts and assumptions

6.1 Concepts

The purpose of the ISO 19100 series of International Standards is to enable interoperability between heterogeneous geographic information systems. To achieve interoperability between heterogeneous systems two fundamental issues need to be determined. The first issue is to define the semantics of the content and the logical structures of geographic data. This shall be done in an application schema. The second issue is to define a system and platform-independent data structure that can represent data corresponding to the application schema.

Subclauses 6.2 to 6.6 describe the fundamental concepts of data interchange, i.e. the procedure based on the application schema for encoding, delivery, receipt and interpretation of geographic data. 6.2 describes an overview of the data interchange process; 6.3 introduces application schemas that allow interpretation of geographic data; 6.4 describes the importance of the encoding rule for producing system-independent data structures; 6.5 describes a software component, called the encoding service, for executing the encoding rule; and 6.6 describes the procedure for delivery and receipt called the transfer service.

6.2 Data interchange

An overview of a data interchange is shown in Figure 1. System A wants to send a dataset to system B. To ensure a successful interchange, A and B must decide on three things, i.e. a common application schema I, which encoding rule R to apply, and what kind of transfer protocol to use. The application schema is the basis of a successful data transfer and defines the possible content and structure of the transferred data, whereas the encoding rule defines the conversion rules for how to code the data into a system-independent data structure.

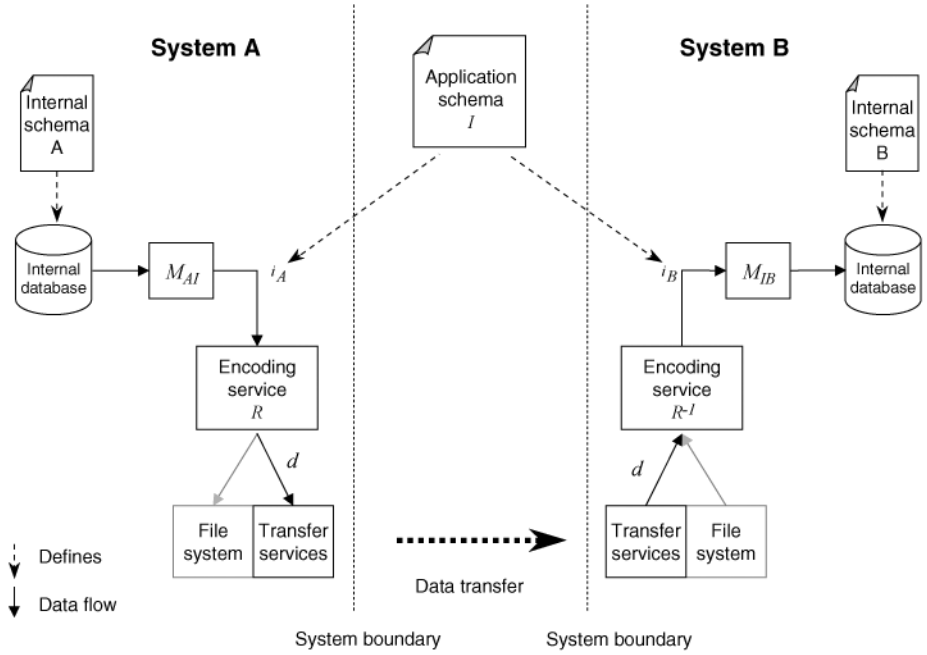

Figure 1 — Overview of data interchange between two systems

Both systems, A and B, store data in an internal database according to an internal schema, but the schemas are usually different, i.e. schema A ≠ schema B. The following logical steps must be taken in order to transfer a dataset from A's internal database to B's internal database.

a) 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. Here this is done by defining a mapping from the concepts of the internal schema to the concepts defined in the application schema and by writing appropriate mapping software to translate the data instances. In Figure 1 this mapping is denoted M_{AI}. The result is an application schema specific data structure i_A. The data structure is stored in memory or on an intermediate file and is system dependent and thus not suitable for transfer.

b) 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.

c) 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 on-line or off-line communication medium should take place. Both parties must agree upon the transfer protocol used.

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

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

f) To use the dataset, B must translate the application schema specific data structure i_B into its internal database. This is done by defining a mapping from the application schema into its internal schema and by writing software that does the actual translation. In Figure 1 this mapping is denoted M_{IB}.

This International Standard only specifies the requirements for creating encoding rules and the encoding services and not the whole data interchange process. Thus, only steps b) and e) are standardized. Steps c) and d) use general information technology transfer services.

6.3 Application schema

An application schema is a conceptual schema for applications with similar data requirements. The application schema is the basis of a successful data interchange and defines the possible content and structure of the data. It is also the basis for implementing application schema specific data structures for local storage of data.

The application schema used for encoding in compliance with this International Standard shall be written in the UML conceptual schema language, in accordance with ISO/TS 19103 and ISO 19109. These two standards specify a framework for how to write application schemas. The rules include specifications on how to use standardized schemas to define feature types. Both a sender and a receiver of data must have access to the application schema.

Implementation of application schema specific data structures may be semi-automated, e.g. by using a compiler for the lexical part or by using a graphical modelling tool with support for code generation.

The application schema shall be accessible to both ends of a data interchange to ensure a successful result. The application schema has to be transferred before data interchange takes place so that both the receiver and sender can prepare their systems by implementing mappings and data structures according to the application schema. It may be transferred together with the dataset, or it may be stored in a public place and referenced from the dataset.

The application schema may be interchanged by paper or electronic based methods. The XML Metadata Interchange (XMI) developed by the Object Management Group (OMG) is the recommended electronic method for data interchange of an application schema.