SLOVENSKI STANDARD
SIST CR 13568:1999

01-junij-1999

Geografske informacije - Opis podatkov - Jezik za konceptualno shemo

Geographic information - Data description - Conceptual schema language

Information géographique - Description de données - Langage de schéma conceptuel

Ta slovenski standard je istoveten z: **CR 13568:1999**

ICS:

07.040	Astronomija. Geodezija. Geografija	Astronomy. Geodesy. Geography
35.240.70	Uporabniške rešitve IT v znanosti	IT applications in science

SIST CR 13568:1999 **en**

CEN REPORT

RAPPORT CEN

CEN BERICHT

CR 13568

April 1999

ICS

English version

Geographic information - Data description - Conceptual schema language

Information géographique - Description de données - Langage de schéma conceptuel

This CEN Report was approved by CEN on 27 November 1996. It has been drawn up by the Technical Committee CEN/TC 287.

CEN members are the national standards bodies of Austria, Belgium, Czech Republic, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland and United Kingdom.

EUROPEAN COMMITTEE FOR STANDARDIZATION
COMITÉ EUROPÉEN DE NORMALISATION
EUROPÄISCHES KOMITEE FÜR NORMUNG

Central Secretariat: rue de Stassart, 36 B-1050 Brussels

Ref. No. CR 13568:1999 E

Foreword

This European Report has been drawn up by the Technical Committee CEN/TC 287 "Geographic Information" (secretariat AFNOR).

1 Scope

Conceptual schemas can be expressed in many ways. This report identifies and evaluates languages suitable to describe geometric information on the conceptual level. This includes both lexical and graphical languages.

The aim of this report is to propose the use of a single language that can be applied to all fields of geographic information.

2 References

This European Report incorporates by dated or undated reference, provisions from other publications. These references are cited at the appropriate places in the text and the publications are listed hereafter. For dated references, subsequent amendments to or revisions of any of these publications apply to this European Report only when incorporated in it by amendment or revision. For undated references the latest edition of the publication referred to applies.

prEN 12009	1995	Geographic Information - Reference Model
ISO TR 9007		Concepts and Terminology for the Conceptual Schema

3 Definitions

For the purposes of this report, the terms defined or explained in prEN 12009 „Geographic Information - Reference Model" apply. The following concepts have been used :

3.1 geographic information : Information concerning phenomena directly or indirectly associated with a location relative to the surface of the Earth.

3.2 geographic data : Data describing phenomena directly or indirectly associated with a location relative to the surface of the Earth.

3.3 metadata : Data describing data.

3.4 data interchange : Transfer of predefined and structured geographic data from one party to another.

3.5 conceptual modelling : Identifying and defining types of phenomena and their relationships.

3.6 conceptual formalism : Methods for the creation of conceptual schemas

3.7 lexical language : Formal language using words and mathematical symbols; a lexical language can be used to present conceptual schemas.

3.8 graphical notation : Formal language using graphical symbols; a graphical notation can be used to present conceptual schemas.

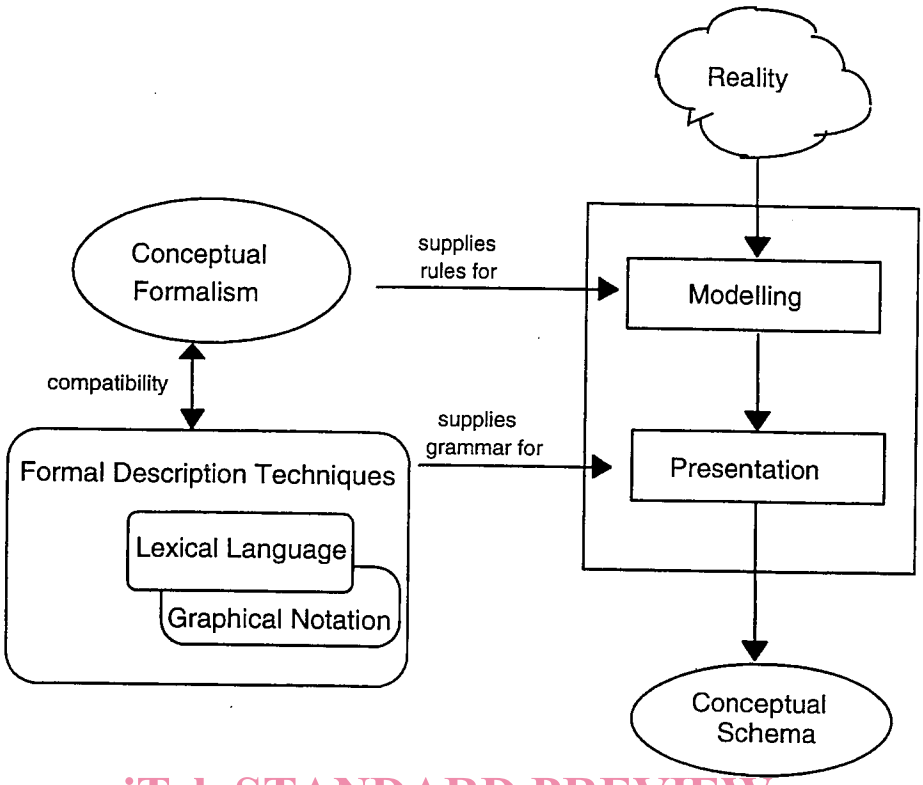

Figure 1 : Conceptual modelling process

Geographic Information Services must include facilities for searching, inquiring and delivering geographic data. Data descriptions appear in relation to these actions in the following way :

- search supported by metadata including the conceptual schema ;

- inquiry supported by the conceptual schema ;

- delivery supported by the transfer schema.

This report concentrates on choosing a conceptual schema language. The definition of techniques for mapping from a conceptual schema to a transfer schema is not part of this report.

4.2 Conceptual data description as a part of metadata

The business view introduces data as a deliverable product and aims at giving all necessary information to a customer for acquiring data. This information is not entirely bound to the semantics or the structure of the data and not necessarily visible when studying the data itself. Examples are the copyright and the pricing of the data. The description has to be in a form which is understandable to a human without any special knowledge or technical background. Data description from the business view is often called metadata.

The conceptual description of data shall be included in the metadata, but in many cases a generalised representation will be sufficient and a comprehensive description with all details will not be required.

An older grouping (Brodie et al., 1985) makes the following distinction :

- primitive data models (record structures) ;

- classical data models (hierarchical, network and relational structures) ;

- semantic data models.

The above mentioned groups of approaches all are semantic data models.

When considering existing conceptual schema languages they can often, but not always be mapped to a single formalism. For example, IDEF1X is based on an EAR approach and NIAM on a Binary Relationship approach, but EXPRESS is rather an „object based“[2] language and does not fit fully in any of the groups.

Many of the description languages appear only as a graphic notation without a lexical language. They are meant for visualising a conceptual schema during the modelling process but cannot provide complete computer processable descriptions of conceptual schemas.

For data transfer existing data are described, whereas in system design a conceptual schema is created for establishing a database. Because the difference does not affect the result, the terms schema and modelling can be used in describing existing data as well as in system design.

5.3 Elements of a conceptual schema

The conceptual schema introduces the semantics of the data. The elements recognised depend on the modelling method. Typically the data is described at the type level, i.e. the actual data is considered as instances or occurences of types. Each type has a definition which is common to all instances of that type. The primitive elements of data are identified and defined as well as associations between them. In the following, when necessary, the term *object type* is used for a unit of data at the type level and the term *relation* for an association among object types. If the properties of objects are identified they are termed *attributes*. The conceptual schema also defines constraints and rules within and between object types, attributes and relations. In addition, the conceptual schema defines the domains of attribute values without going into codings.

---

[2] The term „object based“ shall indicate that in general the object oriented formalism is used, only the dynamic aspects (and effectivly encapsulation) have been left out on purpose.

Req. 7 Mapping to the transfer schemas

For data transfer, the description language shall be supported by rules for encoding the data defined by this language.

Req. 8 Documentation

The description language shall be well documented as a basis for full understanding of the language.

Req. 9 International Standard

The description language shall be an official, preferably internationally accepted standard.

Req. 10 Supporting software

Software products shall be available to support the usage of the description language in respect of the requirements listed above.

Req. 11 Experience of and support for usage

The description language shall be well tested, preferably in practical use. The language should have been applied to geometric data. Supporting material and services should be available.

6.2 Potential future requirements

When object oriented techniques mature in data management, describing behaviour becomes an essential part of conceptual modelling. This would then have to be considered as a requirement for the conceptual schema language.

With the the spread of usage of multimedia, new forms of representing information are being introduced. Even if a conceptual schema is independent of the representation of data, multimedia will bring new information types (image, sound, video) and may put additional requirements on the conceptual schema language.

Conceptual description of data will have an important role when application programming interfaces (API) or service interfaces, as identified in the Reference Model, are to be defined. This may impose additional requirements on the conceptual schema language.

The standard conceptual schema language should include provisions for future developments to cope with these additional requirements.

6.3 Simplicity vs. complexity

The description language shall be as simple as possible. Simplicity would ensure wide usage of the language also among people who are not professionals in the technique in question.

However, this may be difficult to achieve since the structures and constraints in geographic data are typically complex. Complex matters cannot always be expressed by simple means or methods. Tools developed for the techniques provide user friendly

7.1.2 EDIFACT (ISO 9735) defines a syntax for messages. Application of it in describing conceptual schemas would be complicated, especially when expressing complex relations and constraints. In the context of geographic information transfer, EDIFACT could be used as a transfer syntax and the description methods could be applied as a transfer description language.

7.1.3 IRDS is an Information Resource Dictionary Standard Framework (ISO/IEC 10027). Description languages developed in the IRDS context were sought, and a language specified in ANSI X3.138-1988 was identified. It is a Command Line Interface, CLI, which is not suitable as a formal description language. A wide report on conceptual languages has been prepared in the IRDS context and published in ANSI X3H4.6/92 and in ISO/IEC JTC 1/SC 21 N 7486 by the name Draft IRDS Conceptual Schema. The development of IRDS standards is now under reconsideration and no relevant candidate can be found at the moment.

7.1.4 ISO 8211 defines a file which includes a data descriptive file and the actual data. The data descriptive file allows naming of records, fields and sub-fields but does not allow description of values of a domain. It is not a language for conceptual modelling. In the context of geographic information transfer, ISO 8211 could be used as a transfer description language when transferring files.

7.1.5 SPECtalk is a data definition language based on the object oriented approach. It has been used in specifying geographic data in SAIF, a draft national standard of Canada (C**/CGSB-171.1, 1991). The evaluators have not obtained a full documentation of this language. As the language is not a standard nor widely used and no information about the maintenance or software support is available the language is not evaluated any further.

7.1.6 STANLI method is a description technique used by the Swedish standardisation project in geographic information. It has been applied by CEN/TC 287/WG 2 in the definition of geometry sub-schemas. The graphic notation is easy to learn in reading and writing schemas but it does not provide a language in textual form. The expression power has not been sufficient for the work of WG 2, i.e. it was necessary to have written constraints as additional textual notes.

7.2 Evaluation method

The description languages have been evaluated on the basis of the requirements listed in clause 6. Information about the languages has been gathered from the documents and experts on those languages. A workshop on data description techniques was held in order to support the discussion about the languages.

In addition, a test case has been described by all the evaluated languages except SQL 3. The test case focusses on the requirements 2 (Elements of the conceptual schema) and 6 (Geometric aspects).

The test case described in English and its description in the evaluated languages are in annex A. The results are summarised in the evaluation of each language against requirement 2 and 6.

Schenk, D., Wilson, P., Information Modeling. The EXPRESS Way, Oxford University Press, New York, 1993.

Brief description :

EXPRESS is developed for the purpose of information modelling and description as a part of the STEP standard which supports the computer-interpretable representation and exchange of product data. The objective of EXPRESS is to provide a neutral mechanism capable of describing data throughout the life cycle of a product, independent of any particular system implementation. EXPRESS became an IS in 1994. Its development started in 1986 after both NIAM and IDEF1X had been applied in STEP work and were found inadequate.

The main components and features of EXPRESS are :

- object data types (in EXPRESS termed Entities), including subtypes and (multiple) supertypes ;

- attributes, including derived attributes; uniqueness and domain rules can be defined for attributes ;

- relationships, appearing as 1:n relationships except in aggregations ;

- global rules to specify constraints that have a wider scope than one object ;

- data types, which are grouped in simple data types, aggregation data types, constructed data types and generalised data types ;

- functions and procedures, including built-in functions and built-in procedures.

EXPRESS is an object based language. It includes, for example, multiple inheritance but does not have all characteristics, like encapsulation, typical of an object oriented language. The expression power of EXPRESS is high and it can be used in describing complex models. Large models can be handled as a set of sub-schemas.

Evaluation :

Req. 1 Formal language for conceptual modelling

EXPRESS is defined "to be a formal information requirements specification language". It is consistent with the requirements specified in ISO TR 9007.

Req. 2 Elements of the conceptual schema

All required elements are basic elements of EXPRESS.

Req. 3 Graphic notation

EXPRESS-G is a formal graphic notation documented in the normative annex D of ISO 10303-11. It supports a subset of EXPRESS (no functions and procedure, no rules but an indication of the existence of a rule).

EXPRESS-C, that introduces the idea of operations on objects and thus provides means for encapsulation. It is still an open question if it will become a standard in the future.

Part 22 of STEP, SDAI (short for: Standard Data Access Interface), will introduce the possibility not only to transfer whole dataset, but to access objects in a database. Language bindings, for example part 23 will be a C++ language binding, are being developed. It will enable a programmer to access EXPRESS-defined objects by using standard C++ features.

7.5 Evaluation of IDEF1X

Name :

IDEF1X (Integration Definition for Information Modelling)

References :

FIPS BUP 184, Integration Definition for Information Modelling (IDEF1X), 1993 December 21. For sale by National Technical Information Service, U.S: Department of Commerce, Springfield, VA 22161.

ISO/IEC JTC 1/SC 21 N 7486 (ANSI X3H4.6/92-001R3), Draft IRDS Conceptual Schema, 1992-12-03, pp. 93-112.

Anne L. Williams, Comparative analysis of the Modeling Languages: IDEF1X, EXPRESS and NIAM, 13 p., May 1989.

Brief description:

IDEF1X is a design tool which follows the EAR modelling approach. It is actually a graphic notation but also textual form(s) exist. Its primary objectives are to describe data resources in order to support the communication and understanding of data, the transformation into physical database design and the derivation of an integrated data definition from existing data resources.

The language was developed in the early 1980's by the US Air Force. In the USA, it is the most widely used information modelling methodology in the aerospace and defence industry and it also has wide use in commercial industry.

The components of IDEF1X are :

- objects (in IDEF1X termed Entities) :
  - identifier-Independent Entities,
  - identifier-Dependent Entities ;
- relationships :
  - identifying Connection Relationships,
  - non-Identifying Connection Relationships,