

Fødevaredata – Struktur og udvekslingsformat

Food data – Structure and interchange format

DS/EN 16104

København

DS projekt: M246493

ICS: 67.040

Første del af denne publikations betegnelse er:

DS/EN, hvilket betyder, at det er en europæisk standard, der har status som dansk standard.

Denne publikations overensstemmelse er:

IDT med: EN 16104:2012.

DS-publikationen er på engelsk.

DS-publikationstyper

Dansk Standard udgiver forskellige publikationstyper.

Typen på denne publikation fremgår af forsiden.

Der kan være tale om:

Dansk standard

- standard, der er udarbejdet på nationalt niveau, eller som er baseret på et andet lands nationale standard, eller
- standard, der er udarbejdet på internationalt og/eller europæisk niveau, og som har fået status som dansk standard

DS-information

- publikation, der er udarbejdet på nationalt niveau, og som ikke har opnået status som standard, eller
- publikation, der er udarbejdet på internationalt og/eller europæisk niveau, og som ikke har fået status som standard, fx en teknisk rapport, eller
- europæisk præstandard

DS-håndbog

- samling af standarder, eventuelt suppleret med informativt materiale

DS-hæfte

- publikation med informativt materiale

Til disse publikationstyper kan endvidere udgives

- tillæg og rettelserblade

DS-publikationsform

Publikationstyperne udgives i forskellig form som henholdsvis

- fuldtekstpublikation (publikationen er trykt i sin helhed)
- godkendelsesblad (publikationen leveres i kopi med et trykt DS-omslag)
- elektronisk (publikationen leveres på et elektronisk medie)

DS-betegnelse

Alle DS-publikationers betegnelse begynder med DS efterfulgt af et eller flere præfikser og et nr., fx **DS 383**, **DS/EN 5414** osv. Hvis der efter nr. er angivet et **A** eller **Cor**, betyder det, enten at det er et **tillæg** eller et **rettelsesblad** til hovedstandard, eller at det er indført i hovedstandard.

DS-betegnelse angives på forsiden.

Overensstemmelse med anden publikation:

Overensstemmelse kan enten være IDT, EQV, NEQ eller MOD

- **IDT:** Når publikationen er identisk med en given publikation.
- **EQV:** Når publikationen teknisk er i overensstemmelse med en given publikation, men præsentationen er ændret.
- **NEQ:** Når publikationen teknisk eller præsentationsmæssigt ikke er i overensstemmelse med en given standard, men udarbejdet på baggrund af denne.
- **MOD:** Når publikationen er modificeret i forhold til en given publikation.

ICS 35.240.60; 67.040

English Version

Food data - Structure and interchange format

Données sur les aliments - Structure et format d'échange

Lebensmitteldaten - Struktur und Austauschformat

This European Standard was approved by CEN on 3 November 2012.

CEN members are bound to comply with the CEN/CENELEC Internal Regulations which stipulate the conditions for giving this European Standard the status of a national standard without any alteration. Up-to-date lists and bibliographical references concerning such national standards may be obtained on application to the CEN-CENELEC Management Centre or to any CEN member.

This European Standard exists in three official versions (English, French, German). A version in any other language made by translation under the responsibility of a CEN member into its own language and notified to the CEN-CENELEC Management Centre has the same status as the official versions.

CEN members are the national standards bodies of Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, Former Yugoslav Republic of Macedonia, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey and United Kingdom.



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

Management Centre: Avenue Marnix 17, B-1000 Brussels

Contents

Page

Foreword.....	4
Introduction	5
1 Scope.....	7
2 Terms and definitions	8
3 Symbols (and abbreviated terms)	12
3.1 UML notation	12
3.2 Abbreviated terms and acronyms	12
4 Data structure.....	13
4.1 General	13
4.2 Main data structure	13
4.3 Bibliographic references.....	14
4.4 Classes for attribute specification	15
4.5 Classes for codes of controlled vocabularies	16
4.6 Classes for identifiers	17
5 Class descriptions	17
5.1 General	17
5.2 Classes for numerical and logical values	18
5.3 Classes for text	18
5.4 Class for date and time	19
5.5 Code_ classes, for attributes based on controlled vocabularies.....	20
5.6 Id_ classes, for identifiers.....	20
5.7 Aggregation class	21
5.8 Allergen class, AllergenSpec class, Code_Allergen classes	21
5.9 ArticleInfo class	22
5.10 Claim class, Code_ClaimType class, Code_ClaimCode class	23
5.11 Code_SciName class	23
5.12 Composite class, Code_Strategy class, Code_SamplingMethod class	23
5.13 Contact class.....	24
5.14 Contributor class	24
5.15 Descriptor class, Code_Descriptor class.....	24
5.16 Food class	25
5.17 Id_Food class	26
5.18 Image class.....	26
5.19 Ingredient class.....	26
5.20 Mandate class	27
5.21 Measure class, Id_Measure class, Code_Action class.....	27
5.22 Method class, Code_MethType class, Code_MethInd class	28
5.23 MethodStep class, Code_MethStep class	28
5.24 MethodValidation class, Code_MethValidation class.....	29
5.25 Performance class	29
5.26 Place class, Code_PlaceType class	31
5.27 Preparation class, Code_Preparation class.....	31
5.28 Programme class, Code_Programme class.....	31
5.29 Property class, Code_Property class.....	32
5.30 Quality class, Code_Quality class.....	32
5.31 Quantity class	33
5.32 Recipe class	33
5.33 Reference class, Code_RefType class.....	33
5.34 Sample class	34

5.35	Source.....	34
5.36	TypedDate class, Code_Date class	34
5.37	TypedValue class, Code_ValueType	35
5.38	UoM class, Code_UoM class.....	35
5.39	Vocabulary class	36
	Data encoding	36
6.1	General	36
6.2	Data instances with identity.....	36
6.3	Sequential order of XML elements	37
6.4	References to data instances	37
6.5	Data instances of class attributes	37
6.6	XML schema datatypes	38
6.7	Encoding of Decimal datatype class	38
6.8	Encoding of Date datatype class.....	38
6.9	Encoding of MultiText datatype class	38
6.10	Encoding of Vocabulary class	39
6.11	Encoding of Code_ classes	39
6.12	Encoding of list of valid values	39
6.13	Encoding of identifiers.....	41
	Annex A (informative) UML notation	43
	Annex B (informative) Examples of data instances	44
B.1	Overview.....	44
B.2	Agreements between actors	44
B.3	Data instances and dataset.....	48
B.4	Food description and food property measures	48
B.5	Aggregation of food property measures	52
B.6	Food composite and food samples	54
B.7	Food recipes.....	57
B.8	Food article with article information.....	60
	Annex C (informative) XML schema and example of XML encoding	63
C.1	General	63
C.2	XML schema	63
C.3	XML document example.....	64
	Annex D (informative) Examples of controlled vocabularies	66
D.1	General	66
D.2	Background	66
D.3	Food description	66
D.4	Food classification	67
D.5	Food property description	68
D.6	Analytical methods.....	68
D.7	Geographic places and languages	69
D.8	Units of measure	69
D.9	Other controlled vocabularies	70
	Bibliography.....	71

Foreword

This document (EN 16104:2012) has been prepared by Technical Committee CEN/TC 387 "Food data", the secretariat of which is held by SIS.

This European Standard shall be given the status of a national standard, either by publication of an identical text or by endorsement, at the latest by June 2013, and conflicting national standards shall be withdrawn at the latest by June 2013.

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

According to the CEN/CENELEC Internal Regulations, the national standards organisations of the following countries are bound to implement this European Standard: Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, Former Yugoslav Republic of Macedonia, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey and the United Kingdom.

Introduction

The term *food* generally refers to substances intended for human consumption, normally with exceptions for e.g. medicines, and includes raw or processed food products and substances used in the manufacture. The exact definition, however, may vary depending on legislation and cultural differences. This standard can be used regardless of such variations.

This standard uses *food properties* as a general term when describing food constituents such as nutrients, heavy metals, micro-organisms, but also when describing various physico-chemical properties of foods.

Food data address description and identification of foods and their food properties. They are needed and used for many purposes, e.g. labelling, product development, dietary treatment, nutritional treatment, consumer information, and research. Thus, there are many types of parties that need to generate, compile, interchange, or access detailed information about foods. These include:

- Food manufacturers
- Food analysis laboratories
- Authorities
- Researchers
- Resellers
- Retailers
- Nutritionists/dieticians
- Food distributors
- Consumers
- Restaurants/food service operators
- Software developers

The ability and need to manage food information vary between these parties. There are multiple instances of all parties mentioned, which means that the information is interchanged in a large number of relations between parties. For example, a food manufacturer may have the need to communicate food information with multiple resellers, multiple retailers, multiple distributors and multiple authorities in multiple countries, and so on.

Currently, there are differences among member states and parties in the way food data are expressed with respect e.g. food description, definition of nutrients and other food properties, and methods used to generate compositional values. A common European Standard, established within the CEN framework, is a key tool enabling unambiguous identification and description of food data and its quality in e.g. databases, for dissemination and interchange.

Several European and international initiatives have focused on improving and harmonising food data description and interchange. This standard is based on two initiatives: the EuroFIR project [11] (an EC Network of Excellence funded by the 6th Framework Programme for Research and technological Development) 2005-2010 and Food and Beverage Extension to the GS1 GDSN Trade Item standard [14].

The Eurofir project mainly concerned specifications for documentation and interchange of data on nutrients and bioactive substances in food composition databases, while the GS1 standard was intended for use by trading partners in both the food service as well as the food retail sector. In addition, a set of use cases were developed and analysed. This standard was also aligned with the EFSA Standard Sample Description [17], and certain elements and specifications were incorporated. As a result, this standard is more innovative and broader in scope - in so far as it should be fit for the purposes of all these parties.

The main aim of the standard is to provide a framework that facilitates and enables generation, compilation, dissemination and interchange of food data that are comparable and unambiguous with respect to the identity of foods, the description of foods and food property measures including their quality. The standard is structured to be robust and flexible enough to incorporate future extensions with respect to various types of data.

This standard will make it possible for any party in a community to send understandable food data to any other receiving party in that community. However, this standard does not include all definitions that are required. For example, the set of food properties that can be used, such as contents of various nutrients and heavy metals, is not included in the standard. These and all other so called controlled vocabularies will be agreed upon within the community. An annex of the standard provides examples of required controlled vocabularies.

The reasons for not including the controlled vocabularies are:

- Most controlled vocabularies, for example with new food properties, will be constantly updated.
- Communities around the world are maintaining and using their own controlled vocabularies.

The exchange of food data among different parties requires an agreement on not only what data to exchange but also on the encoding of the data. This standard includes data encoding rules based on XML which today is the most recognised general technique for data encoding.

Figure 1 illustrates a case where food data is exchanged between databases of partners (1) in some kind of community. They want to use this standard to set up a mutual agreement (2). Apart from selecting the standard (3), such an agreement will contain selections of controlled vocabularies and restrictions on data. Most of the data specified in the standard require a controlled vocabulary to be specified (5). Such controlled vocabularies are maintained by various organisations. An agreement will select the controlled vocabularies to use. Restrictions on data will be defined (4). For example, an agreement may state that a scientific name has to be provided for all foods, despite the fact that it is not required in the standard. In addition, an agreement may specify requirements on what food properties are to be exchanged or what language to use.

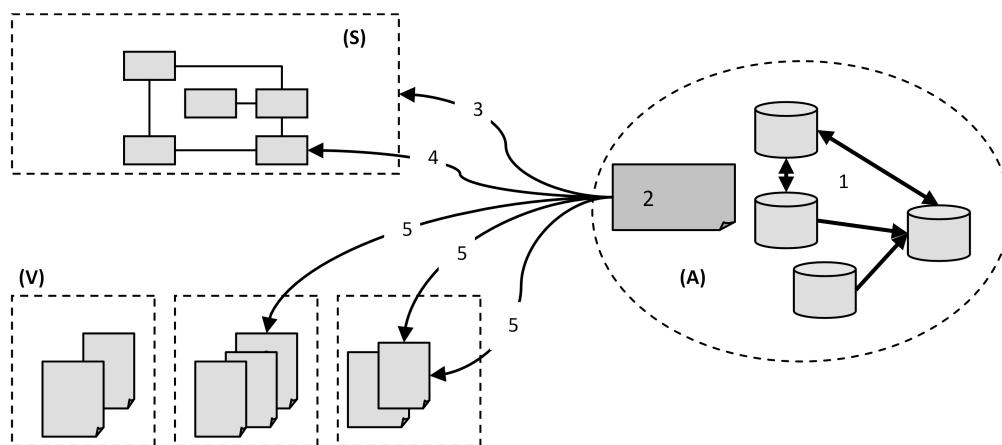


Figure 1 — Agreement for food data exchange (A) that are based on this standard (S) will also specify a set of controlled vocabularies (V).

1 Scope

This European Standard specifies requirements on the structure and semantics of food datasets and of interchange of food data for various applications.

Food data refers to information on various food properties and includes various steps in the generation and publication of such data, e.g. sampling, analysis, food description, food property and value description.

The standard regards food data as datasets covering:

- identification, description and classification of foods including food ingredients,
- qualitative and quantitative food properties that can be measured, calculated or estimated,
- data quality values and other metadata,
- specifications of methods used for obtaining these values,
- references to sources for the information reported.

This standard includes requirements on:

- semantics and data structure for food data,
- content of referenced controlled vocabularies,
- XML encoding for interchange of food data.

This standard does not include:

- food description methods,
- quality assessment methods,
- content of controlled vocabularies, for example controlled vocabularies for nutrients,
- database implementation.

2 Terms and definitions

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

2.1

association

relation between classes

Note 1 to entry: This term refers to ISO/IEC 19501:2005 (Unified Modelling Language) [1].

2.2

attribute

characteristics of a class described by the values it can have

Note 1 to entry: This term refers to ISO/IEC 19501:2005 (Unified Modelling Language) [1].

2.3

class

definition of data for description of a certain concept

Note 1 to entry: Classes as described in ISO/IEC 19501:2005 (Unified Modelling Language) [1] may also describe operations and methods. That possibility is not used in this standard.

Note 2 to entry: A class is used as a specification of a data instance. For example a data instance of an employee may be specified by a class that specifies two attributes: the name and the salary of the employee.

Note 3 to entry: A class can be used by other classes to specify an attribute. For example, a class for dates may be used to specify a delivery date in one class and a birthday in another class.

EXAMPLE 1 A class that defines a data structure with a code representing a language and a text.

EXAMPLE 2 A class that defines a set of codes for representation of languages.

EXAMPLE 3 A class that defines a text as a sequence of characters.

2.4

class diagram

schema of associations between classes

Note 1 to entry: This term refers to ISO/IEC 19501:2005 (Unified Modelling Language) [1].

2.5

controlled vocabulary

carefully selected set of terms such that each concept from the domain of discourse is described using only one term in the set and each term in the set describes only one concept

2.6

data instance

set of data that is specified by a class

2.7

datatype class

class for specification of attributes