

BS EN 16104:2012



BSI Standards Publication

Food data — Structure and interchange format

bsi.

...making excellence a habit.
AVA Documents™
Document Sharing Hub

National foreword

This British Standard is the UK implementation of EN 16104:2012.

The UK participation in its preparation was entrusted to Technical Committee AW/275, Food analysis - Horizontal methods.

A list of organizations represented on this committee can be obtained on request to its secretary.

This publication does not purport to include all the necessary provisions of a contract. Users are responsible for its correct application.

© The British Standards Institution 2013.
Published by BSI Standards Limited 2013

ISBN 978 0 580 70792 6

ICS 35.240.60; 67.040

Compliance with a British Standard cannot confer immunity from legal obligations.

This British Standard was published under the authority of the Standards Policy and Strategy Committee on 31 January 2013.

Amendments issued since publication

Date	Text affected
------	---------------

EUROPEAN STANDARD
NORME EUROPÉENNE
EUROPÄISCHE NORM

EN 16104

December 2012

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.

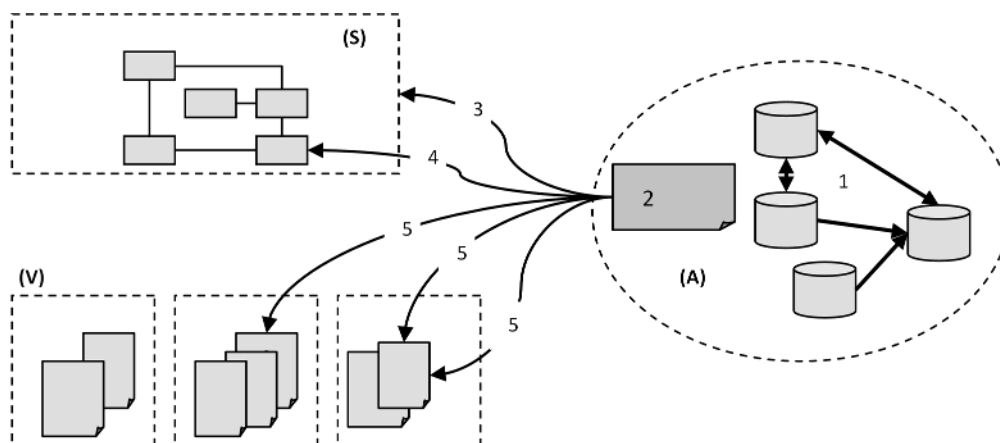


Figure 1 — Agreement for food data exchange (A) that are based on this standard (S) will also specify a set of controlled vocabularies (V).

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

2.8

final preparation instruction

instruction for how to prepare a real food product to be ready-to-eat

Note 1 to entry: A food article often includes a final preparation instruction that tells how to prepare it before you can eat it.

2.9

food

substance intended for human consumption

Note 1 to entry: The term food generally refers to substances intended for human consumption, normally with exceptions for e.g. medicines, feed, cosmetics, tobacco, 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.

Note 2 to entry: Codex definition of food: "Food" means any substance, whether processed, semi-processed or raw, which is intended for human consumption, and includes drinks, chewing gum and any substance which has been used in the manufacture, preparation or treatment of "food" but does not include cosmetics or tobacco or substances used only as drugs (CODEX STAN 1-1985 [20]).

Note 3 to entry: EC definition of food: Food means "Any substance or product, whether processed, partially processed or unprocessed, intended to be, or reasonably expected to be ingested by humans. 'Food' includes drink, chewing gum and any substance, including water, intentionally incorporated into the food during its manufacture, preparation or treatment. It includes water after the point of compliance as defined in Article 6 of Directive 98/83/EC and without prejudice to the requirements of Directives 80/778/EEC and 98/83/EC. 'Food' shall not include: (a) feed; (b) live animals unless they are prepared for placing on the market for human consumption; (c) plants prior to harvesting; (d) medicinal products within the meaning of Council Directives 65/65/EEC (1) and 92/73/EEC (2); (e) cosmetics within the meaning of Council Directive 76/768/EEC (3); (f) tobacco and tobacco products within the meaning of Council Directive 89/622/EEC (4); (g) narcotic or psychotropic substances within the meaning of the United Nations Single Convention on Narcotic Drugs, 1961, and the United Nations Convention on Psychotropic Substances, 1971; (h) residues and contaminants. (Regulation (EC) No 178/2002).

2.10

food article

real food product distributed in a specific packing

EXAMPLE 1 Canned tomato soup of a certain brand and type.

EXAMPLE 2 Corn flakes package of a certain brand and intended for a specific market.

EXAMPLE 3 Apples labelled with the apple variety and name of the producer.

2.11

food data

data identifying and describing properties of foods

2.12

food dataset

identified, structured collection of records of food data that can be stored or exchanged

2.13

food description system

method for systematic description of food as a combination of characteristics

Note 1 to entry: A food description system may include procedures for capturing and retrieval of information about the food.

Note 2 to entry: A food description system defines a set of aspects to be described and a set of food descriptors for each aspect.

2.14

food descriptor

single term in a controlled vocabulary for food description

2.15

food ingredient

food that is used in production of, and still present in, another food

EXAMPLE 1 According to EC directive 2000/13, an ingredient is any substance, including additives, that is used in, and still present after (even in an altered form), the manufacture or preparation of a food.

EXAMPLE 2 According to Codex Alimentarius "Ingredient" means any substance, including a food additive, used in the manufacture or preparation of a food and present in the final product although possibly in a modified form.

2.16

food ingredient list

list describing the absolute or relative quantity or the ranking of food ingredients in a food

2.17

food label information

information written on the package of a food article

2.18

food property

qualitative or quantitative characteristic for a food that can be measured, calculated or estimated

Note 1 to entry: Content of vitamins, minerals, allergens and micro-organisms as well as physico-chemical properties such as pH and specific gravity are examples of food properties.

Note 2 to entry: Sets of food properties are defined and maintained by organisations such as EuroFIR [11], INFOODS [19], and CODEX [16].

EXAMPLE Content of vitamin C; content of calcium; pH; viscosity; retention factor of vitamin C after boiling.

2.19

food property measure

value and unit of measure that quantifies or specifies a food property

Note 1 to entry: A food property measure may include statistical properties of the value.

2.20

food recipe

instructions on how to produce or prepare a food from a set of food ingredients

2.21

food sample

portion of a food that is assumed to represent the food

Note 1 to entry: UPAC [21] definition: portion of material selected from a larger quantity of material.

2.22

food sampling

procedure for the selection, withdrawal, preservation, transportation and preparation of the portions to be removed from a population as samples

2.23

food yield factor

value expressing the remaining compared to the original mass of a food after a certain processing or preparation

Note 1 to entry: This standard handles food yield factors as food properties.

2.24

generic food

abstraction of a food based on several occurrences of foods

Note 1 to entry: The description of a generic food is typically based on several occurrences, e.g. several food articles.

EXAMPLE 1 The food properties of a generic food "apple" may be based on samples of apples of different varieties and origin.

EXAMPLE 2 The food properties of a generic food "pasta" may be based on samples of pasta of different types and brands.

2.25

health claim

any representation that states, suggests or implies that a relationship exists between a food category, a food or one of its constituents and health

Note 1 to entry: The definition of health claim may differ somewhat between different legislation: EC 1924/2006 [24]: any claim that states suggests or implies that a relationship exists between a food category, a food or one of its constituents and health CAC/GL 23-1997 [22]: Health claim means any representation that states, suggests, or implies that a relationship exists between a food or a constituent of that food and health. Health claims include Nutrient function claims; Other function claims; Reduction of disease risk claims.

2.26

multi-ingredient food

food that is made from at least two food ingredients

2.27

nutrient claim

any representation which states, suggests or implies that a food has particular nutritional properties

Note 1 to entry: The definition of nutrient claim may differ somewhat between different legislation: EC 1924/2006 [24]: any claim which states suggests or implies that a food has particular beneficial nutritional properties. CAC/GL 23-1997 [22]: Nutrition claim means any representation which states, suggests or implies that a food has particular nutritional properties including but not limited to the energy value and to the content of protein, fat and carbohydrates, as well as the content of vitamins and minerals. The following do not constitute nutrition claims: (a) the mention of substances in the set of ingredients; (b) the mention of nutrients as a mandatory part of nutrition labelling; (c) quantitative or qualitative declaration of certain nutrients or ingredients on the label if required by national legislation.

2.28

primary food sample

portion of a food initially collected from a food

Note 1 to entry: Other definitions: IUPAC [21]: The collection of one or more increments or units initially taken from a population. The portions may be either combined (composited or bulked sample) or kept separate (gross sample). If combined and mixed to homogeneity, it is a blended bulk sample. CAC/GL 50-2004 [23]: A primary sample is the 'portion of product' collected from a lot during the first stage of the sampling process, and will normally be in the form of an item (if collected from a lot of prepacked products) or of an increment (if collected from a bulk lot).

2.29

real food product

food that is produced harvested or gathered in a certain way independent of package type

Note 1 to entry: A real food product may be produced using a recipe.

Note 2 to entry: A food article is a real food product that is packaged in a certain way.

Note 3 to entry: A food that is not a real food product, is a generic food.

EXAMPLE The production of a certain potato variety.

2.30

retention factor

value expressing the remaining compared to the original content of a food constituent after a certain processing or preparation

Note 1 to entry: This standard handles retention factors as food properties that are defined in a controlled vocabulary.

EXAMPLE 1 The retention factor of vitamin C after boiling.

EXAMPLE 2 The retention factor of fat after frying.

2.31

XML attribute

name-value pair within an XML element start tag

EXAMPLE The value-pair `currency="USD"` is an attribute in the following XML element:

```
<price currency="USD">12.50</price>
```

2.32

XML element

XML data encoding structure that is delimited by a start tag and an end tag

EXAMPLE

```
<employee>
  <name>John Smith</name>
  <salary>4711</salary>
</employee>
```

Note 1 to entry: In the example above there are three XML elements: *employee*, *name* and *salary*. `<name>` is a start tag and `</name>` is an end tag of the *name* element.

3 Symbols (and abbreviated terms)

3.1 UML notation

This standard uses graphical UML notation in Clause 5 and in Annex B. For a complete description of the UML notation, refer to ISO/IEC 19501:2005, *Information technology -- Open Distributed Processing -- Unified Modeling Language (UML) Version 1.4.2* [1]. For an informative explanation of the UML notation, refer to Annex A.

3.2 Abbreviated terms and acronyms

URI Uniform Resource Identifier

NOTE A URI consists of a string of characters used to identify or name a resource on the Internet. There are two types: URL (Uniform Resource Locator) which is a resources web address and URN (Uniform Resource Name) which is only a unique name of the resource.

UML Unified Modelling Language

XML Extensible Markup Language

4 Data structure

4.1 General

The data structure is specified by three UML class diagrams. The division into three class diagrams is only due to readability. Figure 2 specifies the main data structure, Figure 3 specifies the class for bibliographical references and Figure 4 contains datatype classes to be used when specifying attributes in other classes.

Each class is documented in Clause 5 by a description providing semantics and a table of the attributes and association roles that are specified for the class.

Classes with bold border lines are those that will specify data instances that have an identifier. These data instances can be referred to by other data instances and will be encoded in a special way, as described in Clause 7.

The notation of class diagrams is described in Annex A.

NOTE All names of classes, attributes and association roles are symbolic names to be maintained unchanged if this standard is translated.

4.2 Main data structure

The class diagram in Figure 2 specifies the main data structure of a food dataset.

NOTE Annex B provides data instance examples and an informal explanation of the class diagram.

Table 1 lists the classes covered by Figure 2 with a reference to the description of each class.

Table 1 — Classes in Figure 2

Class	Clause
Aggregation	5.7
ArticleInfo	5.9
Composite	5.12
Contributor	5.14
Food	5.16
Ingredient	5.19
Measure	5.21
Method	5.22
MethodStep	5.23
Property	5.29
Recipe	5.32
Sample	5.34
Source	5.35

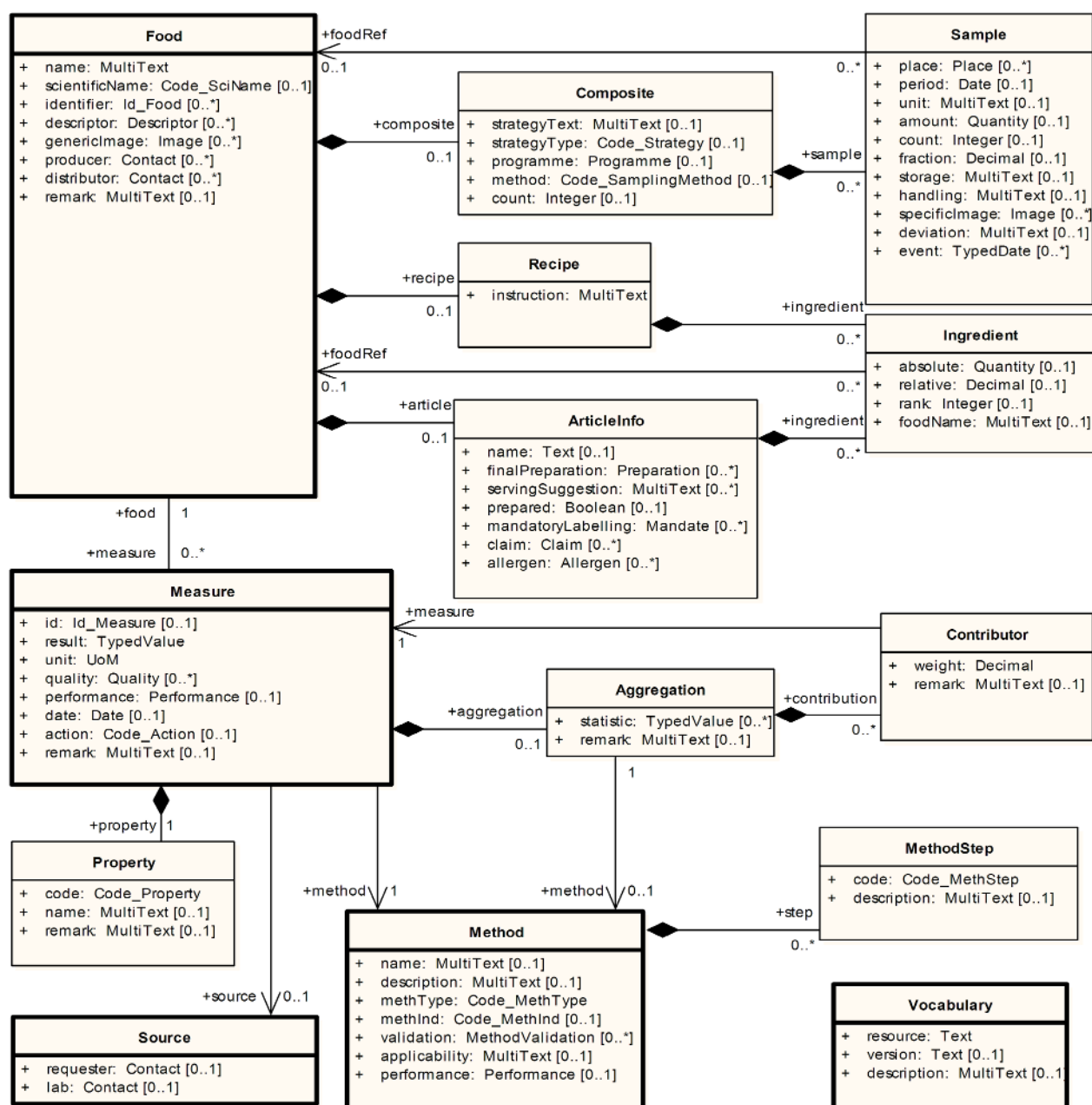


Figure 2 — Main data structure

4.3 Bibliographic references

The class diagram in Figure 3 shows the *Reference* class and the classes that have associations with it.

NOTE These associations are put in a separate class diagram to make the drawing of associations more clear.

Table 2 lists the classes covered by Figure 3 with a reference to the description of each class.

Table 2 — Classes in Figure 3

Class	Clause
Composite	5.12
Measure	5.21
Method	5.22
Recipe	5.32
Reference	5.33
Quality	5.30

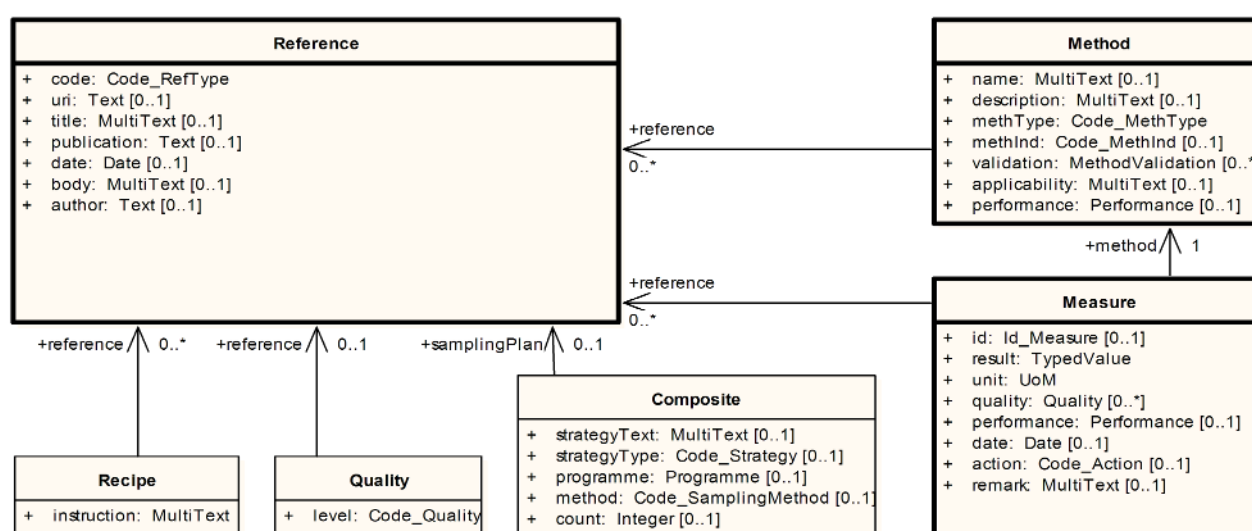


Figure 3 — Associations to the Reference class

4.4 Classes for attribute specification

The class diagram in Figure 4 describes datatype classes, that are classes used for specifying attributes. Table 3 lists the classes covered by Figure 4 with a reference to the description of each class. Classes for terms (see 5.5) and identifiers (see 5.6) are not shown in the diagram.

NOTE 1 These classes are put in a separate class diagram to make the drawing more clear.

NOTE 2 In a dataset, data instances of these classes will become parts of the data instances that are referring to them. Annex B contains many such examples, for example Figure B.5 where each *Property* data instance is part of a *Measure* data instance.

Table 3 — Classes in Figure 4

Class	Clause	Class	Clause
Allergen	5.8	MultiText	5.3
AllergenSpec	5.8	Performance	5.25
Boolean	5.2	Place	5.26
Claim	5.10	Preparation	5.27
Contact	5.13	Programme	5.28
Date	5.4	Quality	5.30
Decimal	5.2	Quantity	5.31
Descriptor	5.15	UoM	5.38
Image	5.18	Text	5.3
Integer	5.2	TypedDate	5.36
LangText	5.3	TypedValue	5.37
MethodValidation	5.24		

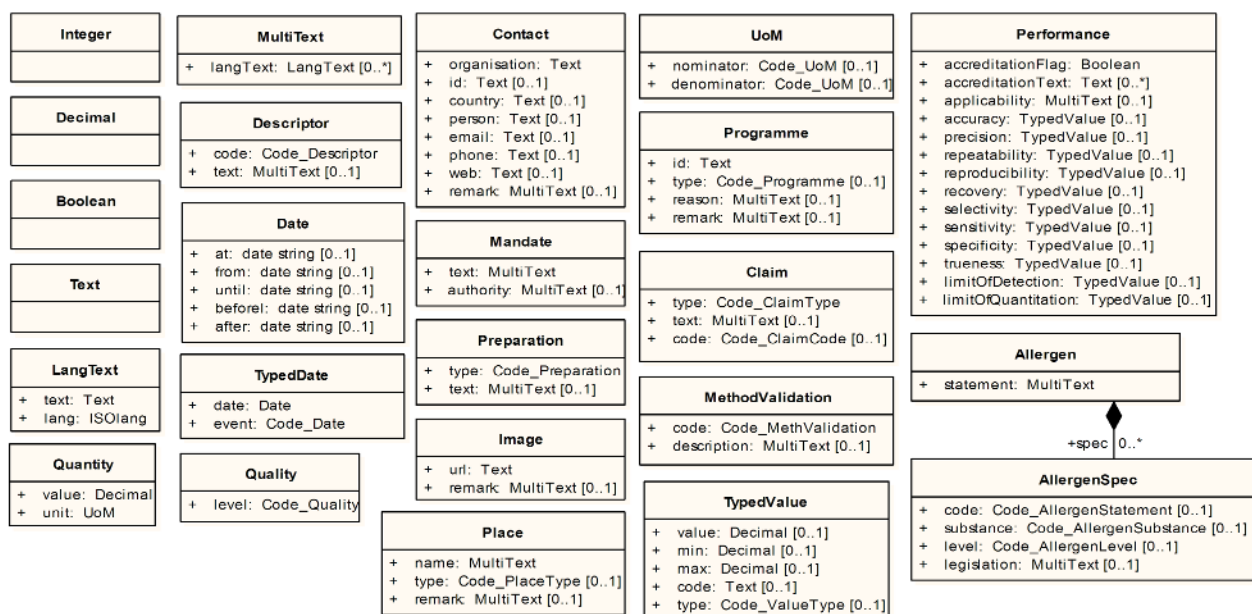


Figure 4 — Classes for attribute specification

4.5 Classes for codes of controlled vocabularies

Some attributes can be assigned codes that are defined as a term of a controlled vocabulary. When a code is assigned to such an attribute, the controlled vocabulary shall also be referred to by its identifier.

A food dataset shall specify the controlled vocabularies that data in the food dataset are referring. The *Vocabulary* class, described in 5.39, shall be used for specification of controlled vocabularies.

In the attribute specifications, each type of controlled vocabulary is a specific data type class with a name beginning with *Code_*.

EXAMPLE *Code_ValueType*, *Code_SamplingMethod* and *Code_Property* are examples of such classes.

NOTE 1 The *Code_*classes are not shown in the class diagrams.

NOTE 2 Annex D contains a list of controlled vocabularies.

4.6 Classes for identifiers

An identifier is a string of characters that is unique within a certain identifier system.

Identifiers are treated similar to codes. Each type of identifier is a specific data type class with a name beginning with *Id_*.

The *Vocabulary* class, described in 5.39, shall be used for specification of identifier systems.

EXAMPLE *Id_Food* and *Id_Measure* are examples of such classes.

NOTE 1 These classes are not shown in the class diagrams.

NOTE 2 This standard does not define any controlled vocabularies.

NOTE 3 This standard does not specify any methods for defining, publishing or accessing a controlled vocabulary. Annex D contains a list of controlled vocabularies.

5 Class descriptions

5.1 General

Clause 5 describes all classes. For most classes, attributes and association roles are described in a separate table with the columns described in Table 4.

Table 4 — Columns of tables for class description

Header	Explanation
Name	Name of the attribute or the role name of an associated class. Role names are written in italics.
Class	Name of the class to be used for the attribute or the name of the associated class. If a data instance of an associated class shall be referred to by its identifier, the class name is written in bold.
Conditions	Restrictions on number of occurrences in a dataset: Mandatory – there shall be one and only one occurrence Optional – there shall be no more than one occurrence Mandatory, repeatable – there shall be at least one occurrence Optional, repeatable – no restrictions Conditional – a more complex condition that is explained in the Semantics column In addition to this information, a condition below the table may state special constraints.
Semantics	Meaning, description and further explanations of the attribute or association.

NOTE 1 The optional condition means that you are conformant with the standard even if you don't provide data that could be provided.

NOTE 2 All names of classes, attributes and association roles are symbolic names to be maintained unchanged if this standard is translated.

5.2 Classes for numerical and logical values

Table 5 lists datatype classes for numerical and logical values.

A decimal number shall be stored and interchanged in a way that makes it possible to find out the number of significant digits.

NOTE 1 The XML encoding of a decimal number is described in Clause 6.

NOTE 2 See also TypedValue class (5.37) for a more general handling of values.

Table 5 — Classes for numerical and logical values

Class	Short description	Example
Decimal	decimal number with a certain number of significant digits	17, 3,45, 0,072
Integer	integer number	1 066
Boolean	true or false	True

NOTE 3 In this document, decimal numbers in the text and tables use a comma as a decimal separator according to ISO Directives Part 2 [7]. However, in XML encoding, a point is used for decimal separator. This is according to the W3C specification XML Schema Part 2: Datatypes [8]. Therefore, XML examples use a point for decimal separator.

5.3 Classes for text

The *Text* datatype class specifies a string of at least one character.

The *Text datatype class* shall be used for texts that are language independent. The *Text* datatype class is also used by the *LangText* datatype class.

EXAMPLE Identifiers are language independent texts.

The *LangText* datatype class, described in Table 6, shall be used to hold text in one specified language and is used by the *MultiText* datatype class only. For multi-lingual texts the *MultiText* datatype class, described in Table 7, shall be used.

Table 6 — LangText class attributes

Name	Class	Condition	Semantics
text	Text	Mandatory	String of characters forming a text in the language specified by the language attribute.
lang	ISOLang	Mandatory	A two-letter code according to ISO 639-1 [2] indicating the language, optionally followed by e dash and a two-letter country code according to ISO 3166-1 [3].

Table 7 — MultiText class attributes

Name	Class	Condition	Semantics
langText	LangText	Mandatory, repeatable	Text in different languages. One data instance of MultiText may only contain one text string in each language.

NOTE The possibilities of handling several languages may affect the selection of character encoding. It is essential that the encoding supports all the characters that may occur.

ISO 639-1 [2] lists codes for languages but does not differentiate between national variants, e.g. British and American English have the same code. Since the name for the same food may differ between language variants, it is recommended that codes for languages be combined with the country code, listed in ISO 3166-1 [3]. See also 6.9.

EXAMPLE "en" for English, "pt" for Portuguese, "en-CA" for Canadian English, "pt-BR" for Brazilian Portuguese.

5.4 Class for date and time

The *Date* class, described in Table 8, shall be used to represent points of time and periods of time. The use of attributes is dependent on what is to be expressed; see the Semantics column of Table 8.

A point of time shall be represented by a string that follows the pattern: CCYY-MM-DDThh:mm:ss where CCYY is the year, including century, MM is the month (01 to 12), DD is the day of the month (01 to 28, 29, 30 or 31), hh is the hour (00 to 23), mm is the minute (00 to 59) and ss is the second (00 to 59).

To give a point of time with less precision, one of the following patterns shall be used: CCYY-MM-DD, CCYY-MM or CCYY.

EXAMPLE 1 2009-12-24

EXAMPLE 2 2010-01

EXAMPLE 3 2012-01-13T17:45:00

Table 8 — Date class attributes

Name	Class	Condition	Semantics
at	string representing a point of time	Conditional	When the Date data type shall be used to specify a certain point of time, the <i>at</i> attribute shall hold the date string.
from	string representing a point of time	Conditional	When the Date data type shall be used to specify a limited period of time, the <i>from</i> and <i>until</i> attributes shall hold the date strings for the start and end of the period.
until	string representing a point of time	Conditional	When the Date data type shall be used to specify a limited period of time, the <i>from</i> and <i>until</i> attributes shall hold the date strings for the start and end of the period.
before	string representing a point of time	Conditional	When the Date data type shall be used to specify a period of time without a defined start, the <i>before</i> attribute shall hold the date string for the first point of time after the period.
after	string representing a point of time	Conditional	When the Date data type shall be used to specify a period of time without a defined end, the <i>after</i> attribute shall hold the date string for the last point of time before the period.

NOTE See 6.8 for more examples and 5.36 for the use of Date in TypedDate.

5.5 Code_ classes, for attributes based on controlled vocabularies

Attributes that may be assigned a code from a limited set of possible codes, have the datatype specified by a class with a name beginning with *Code_*. Such a class has attributes specifying the code, the controlled vocabulary and, optionally, the set of valid codes.

The specification of the controlled vocabulary shall be an identifier containing any printable characters (excluding space, tab, etc.). In the dataset, a data instance of the *Vocabulary* class shall bind the identifier to a source where the set of codes is specified. The source may be used by an application when validating a code.

However, since an XML schema is used for validation of a dataset, it is also possible to extend the XML schema by defining a set of valid codes to the dataset. This makes it possible to validate the codes prior to the processing of an application and may be suitable for small and stable controlled vocabularies, such as the set of value types (e.g. "as reported", "average", "maximum"). As an alternative, it is possible to specify, rather than the complete set of terms, a character pattern that can be used to check that a term is likely to be valid.

See 6.12 for how to make such an extension of the XML schema.

Table 9 — Code_ classes attributes and associations

Name	Class	Condition	Semantics
cv	Vocabulary	Mandatory	Identifier for the controlled vocabulary. Shall be bound by the use of the Vocabulary class to a source where the codes are specified.
code	Text	Mandatory	The code identifying the term.
type	See 6.12	Optional	Specification of valid codes.

NOTE 1 A controlled vocabulary contains a set of terms. The code provided is the identifier of the term.

NOTE 2 It is possible to use more than one controlled vocabulary for a certain datatype class.

NOTE 3 This standard does not specify the controlled vocabularies to use.

NOTE 4 Annex D lists examples of controlled vocabularies.

NOTE 5 The encoding of Code_ classes is different from the encoding of other classes; see 6.12.

EXAMPLE 1 Tables B.1 to B.12 in Annex B contain examples of data instances of the Vocabulary class.

EXAMPLE 2 A character pattern for LanguaL Facet A would be the letter A followed by four digits.

EXAMPLE 3 6.11 shows how a set of terms optionally can be bound to the Code_ValueType data type: To the right is the source for the controlled vocabulary, in this case a list of codes and explanations maintained by EuroFIR. The codes are listed in a new XML declaration (CV_EFV) that can be used instead of the original declaration (Code_ValueType). The dataset (to the left) contains data that refers to the controlled vocabulary as well as to the new XML declaration.

5.6 Id_ classes, for identifiers

Attributes that may be assigned an identifier that belongs to a certain identifier system, have the datatype specified by a class with a name beginning with *Id_*. Such a class has attributes specifying the code, the identifier system and, optionally, the character pattern of valid identifiers.

NOTE The encoding is described in 6.13.

EXAMPLE *Id_Food* is a datatype class for Food identifiers. An implementation may restrict the identifiers to be five digits by providing a datatype class named *Id_MyLab* based on *Id_Food*.

Table 10 — Id_ classes attributes and associations

Name	Class	Condition	Semantics
sys	Vocabulary	Mandatory	Identifier of the identifier system which is regarded as a controlled vocabulary. Shall be bound by the use of the Vocabulary class to a source where the identifier system is specified.
id	Text	Mandatory	The identifier string.
type	See 6.13	Optional	Specification of valid identifiers.

5.7 Aggregation class

The *Aggregation* class, described in Table 11, provides a way to give detailed information about the data that were used to produce a resulting food property measure. It shall be used to document the methods used for calculation, analysis and selection of a food property measure together with various statistical results.

To document all contributing and aggregated measures, associations to data instances of the *Contributor* class shall be made. This documentation is optional.

Table 11 — Aggregation class attributes and associations

Name	Class	Condition	Semantics
statistic	TypedValue	Optional, repeatable	Statistical values about the set of contributing measures in the aggregation, such as mean, median and standard deviation
remark	MultiText	Optional	Any remark of the aggregation.
<i>contribution</i>	Contributor	Optional, repeatable	Measures contributing to the resulting measure.
<i>method</i>	Method	Optional	Specification of the method that was used to produce the food property measure from the set of contributing measures in an aggregation.

5.8 Allergen class, AllergenSpec class, Code_Allergen classes

The *Allergen* class, described in Table 12, shall be used to specify an allergen statement.

The allergen statement shall be specified as a text and one or more allergens related to that statement.

Table 12 — Allergen class attributes

Name	Class	Condition	Semantics
statement	MultiText	Mandatory	Allergen statement
<i>spec</i>	AllergenSpec	Optional, repeatable	Specification of allergens