

This work was completed on behalf of the European Food Information Resource (EuroFIR) Consortium and funded under the EU 6th Framework Food Quality and Safety thematic priority. Contract FOOD-CT-2005-513944.

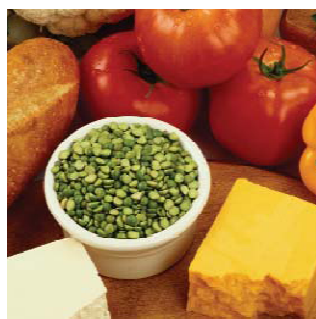
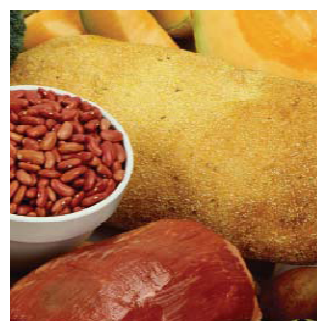
Proposal for structure and detail of a EuroFIR Standard on food composition data

II. Technical Annex

Version 2008

Prepared by
Wulf Becker, Anders Møller, Jayne Ireland,
Mark Roe, Ian Unwin, Heikki Pakkala

EuroFIR Technical Report D1.8.19



Disclaimer

This work was completed on behalf of the European Food Information Resource (EuroFIR) Consortium and funded under the EU 6th Framework Quality and Safety Programme, project number FP6-513944.

EuroFIR, the world leading European Network of Excellence on Food Composition Databank systems (<http://www.eurfir.net/>) is a partnership between 49 universities, research institutes and small-to-medium sized enterprises (SMEs) from 26 countries. EuroFIR aims to develop and integrate a comprehensive, coherent and validated databank providing a single, authoritative source of food composition data for Europe.

**EuroFIR Project Management Office
Institute of Food Research, Norwich Research Park
Norwich, Norfolk, NR4 7UA, UK**

**Proposal for structure and detail of a
EuroFIR Standard on food composition data**

II: TECHNICAL ANNEX.

VERSION 2008

**WULF BECKER
ANDERS MØLLER
JAYNE IRELAND
MARK ROE
IAN UNWIN
HEIKKI PAKKALA**

LEGAL NOTICE

Neither the EuroFIR Consortium nor any person acting on behalf of the EuroFIR Consortium is responsible for the use which might be made of the following information.

Information on the EuroFIR project is available on the Internet. It can be accessed through the EuroFIR server (<http://www.eurofir.net>).

Cataloguing information:

Proposal for structure and detail of a EuroFIR Standard on food composition data

II: Technical Annex. Version 2008.

Wulf Becker, Anders Møller, Jayne Ireland, Mark Roe, Ian Unwin and Heikki Pakkala

Denmark: Danish Food Information

2008 - pp. 35 - 21 x 29.7 cm.

ISBN 978-87-92125-10-1

EAN 9788792125101

© The EuroFIR Consortium, 2008

Reproduction is authorised provided the source is acknowledged.

Printed in Denmark.

Proposal for structure and detail of a EuroFIR standard on food composition data – version 2008

II: Technical Annex

Contents

1	Introduction	5
2	Data structure framework	5
	Data models	5
	Record identifiers	5
	Special modelling	6
	Properties	7
	Text encoding	8
	Priorities for properties	8
	Data Type Formats	9
3	Entity specifications	11
	Food	11
	FoodNames	16
	Recipe	18
	Ingredients	18
	Retention factors (entity awaiting final decision)	19
	Sample	21
	Component	22
	Method Specification	22
	Value	24
	Quality Assessment	27
	Contributing Values	28
	Reference	29
4	Standard Vocabularies (thesauri)	33
	Thesaurus language	33
	Concept description	33
5	XML template for food composition data interchange	35
	EuroFIR Food Transport Package Markup Language	35

1 Introduction

The Technical Annex brings together a description of the technical issues with listings of the data structures and thesauri. Section 3 gives the basic data structures, the form of previous descriptions is retained. The enhanced data structures section suggests the content of the specifications that may be required (e.g. including short and full element tags), as well as additional tables that may be required for full relational handling of the data, exemplified by Food Name and Food Descriptor tables.

2 Data structure framework

This section of the Technical Annex presents an overview of the proposed data structure specification for food composition data and defines some formal conventions used in the detailed description of the data tables and fields given in section 2.

Data models

Currently, the preferred data models for data management and data interchange are distinctly different. Commonly used database management software (DBMS) applies a relational model in data storage, in which a database table is defined for each entity type (and for any many-to-many links between entity tables) and relationships are established between tables through fields common to both. These usually involve a field (or set of fields) that acts as the unique identifier, or primary key, for the entity and "foreign key" fields in other tables that store the same identifier to form the relational link.

On the other hand, EuroFIR data interchange will use files in XML format. Although these can encapsulate a structure of tables linked through key fields, they more naturally follow a nested structure. This is demonstrated in Annex X, which provides an example of food composition data in the proposed EuroFIR format. In this, the Foods element holds separate Food elements that report the data for each individual food item. Within each Food element, together with the elements describing the food, are nested collections of Component records, each with its set of Value records. Thus the nesting accommodates one-to-many relationships between parent and child entities. Many-to-many relationships, such as that between the Value entity and the Reference entity, must link to separate sections or files for one of the entities, in this case with the detailed data for References being held elsewhere in the transport package or in a separate file (or Web-based resource).

This interchange format has some features that are similar to an object oriented model of food composition data, in which entities such as foods, components and values are considered as objects. Both the relational and object oriented data models have been described earlier (Unwin, 2005) and Annex Y reproduces an object oriented data model from that paper.

Record identifiers

In a relational DBMS, data are stored in tables for each distinct concept or entity type, which is a set of items that share common properties that thus can be recorded in the fields

(attributes¹) defined for that table. Each item, i.e. each record in the table, may be referred to as an entity instance, or simply an instance. In order to participate in relationships, each instance must have a unique identifier (consisting of one or more attributes) that distinguishes it from all other instances of the entity type.

Two types of identifier can be considered. A *generic ID*, often a positive integer assigned sequentially, may be added to each new instance when it is created. Alternatively, a *specific ID* may be used that is associated with the individual entity type, for example a food code, method code or reference code.

A field or group of fields that uniquely identifies each instance is known as a candidate key. The database designer selects one candidate key to be used as the record identifier, termed the primary key. An interchange package must also use a system of unique identifiers for records if the records are located in separate files or different parts of the same file (which may be the case in an XML-format file). However, unique identifiers are not essential for entity types where the records are always linked in a many-to-one relationship as the child records of a parent entity. An example is multiple food name records, for example synonyms or names in different languages, linked to a food item through the food identifier as foreign key. Similarly in such cases, identifiers are not required in an XML format, since in such a case the child records (e.g. separate food names) will always be nested under the food to which they relate.

Special modelling

The Eurofoods recommendations (Schlotke *et al.*, 2000) proposed that a special modelling of data should be applied to most of the major entities to provide greater flexibility in data management and interchange. In particular it would:

- allow multiple versions for textual descriptions to hold the data in multiple languages or other alternative versions such as synonyms
- allow free text and controlled language (thesaural) descriptors to be held in parallel
- allow attributes assigned more than one value (e.g. a LanguaL facet with more than one term for a food), referred to as set-valued attributes
- allow any attribute, or any value for a set-valued attribute, to be annotated with comments.

For entities modelled in this way, the database table for the entity type would not contain records for each entity instance, but a record for each property². These records would have the same field structure (i.e. set of attributes) for all such entities, based on pairing the property identifier with each value for the property, defined as a string field with maximum length 255 characters. In addition, the record would contain the identifier of the entity instance (as a *generic ID*), a Memo field (for the value in text format or when it is a memo-type property), a language code field, a field to flag the preferred alternative (e.g. of synonyms) and a fields for comments.

¹ In database terminology, field and attribute are synonymous. However, in XML an attribute is a property whose value is bound with the element tag, whereas the element content (in effect, the field) may include not only values but also nested elements.

² An attribute is considered the column heading of a database table, and therefore the more general term **property** is used when describing the special modelling. In XML, a property value may be reported through an attribute or as element content.

This proposal has been reviewed within the EuroFIR project and some possible difficulties have been identified (Unwin, 2005). Further, for an XML interchange specification, this modelling can be accomplished by nesting the additional information within the overall element for the property. Thus within the present document, the entity specifications will indicate where such additional information might be added, but consider the DBMS implementation a design matter for individual national databases. In the case of alternative text forms that can be considered part of the food composition data, for example the food name in various languages or its synonyms, these are specified as separate entities. In Annex Y, these and set-valued attributes, for example food description terms, are modelled as collections containing instances of the child entities.

Properties

In the entity specifications (section 3), all allowed properties for each entity are listed. Each property is given a descriptive name and its property identifiers to be used in interchange files. The first is the full element tag to be used in XML files and the second is a short identifier (max. 8-characters long), kept for backwards compatibility with the Eurofoods Recommendations. Also provided are the data type, priority information and Additional information containing further notes and explanations.

Properties are grouped together if they represent related information or only apply to a subset of the entity, for example when a FOOD instance is a food sample. In the specifications, group headers are formatted in bold-italics. These might be used in future implementations of editing or browsing software as headings for the display of the properties as groups or pages of related data.

The Eurofoods recommendations (Schlotke *et al.*, 2000) used an *isa-type relationship* to separate out for a given entity sub-types that require different sets of attributes. The example cited was bibliographic reference data, e.g. books, reports or journal articles. Different properties are appropriate for a reference, depending on its *publication type*. For example, a book will have an ISBN, a report may have an ISBN, but a journal article does not. Experience in the EPIC project indicated that the properties allowed in the Eurofoods recommendations were too restrictive, with more of the properties needed for the various publication types. In this case, it is more convenient to indicate the properties allowed by publication type through a tabular layout, although the underlying principle holds good.

As a default rule, a property *Remarks* of type memo is assigned to each entity within the database schema, allowing all additional information not covered elsewhere in the schema to be recorded as textual notes. However, the *Remarks* property should be used with care, for several reasons. If comments on several differing aspects of an entity are made, the user will need to read much irrelevant text in order to see if anything relevant is present³. This problem is exacerbated because plain text fields do not support paragraph markers (i.e. carriage return characters) to indicate that two comments are unrelated. This difficulty could be ameliorated if the *Remarks* property is treated as a set-valued attribute, allowing different comments to be kept separate. In the XML specification, multiple comments may be recorded in separate *Remark* sub-elements within the *Remarks* element or simpler by using a CR-LF (carriage return – line feed) marker in the text to separate paragraphs.

³ There is the corresponding system limitation on the ability to retrieve description of the entity that can only be records in the *Remarks* property.

Text encoding

All data are interchanged in textual form as this is the basis for the XML format. This may use the character set available through single-byte encoding (ISO, 1987), although the double-byte Unicode character set⁴ can also be used.

It is desirable that formatting of text should also be available. Text formatting includes:

- different fonts, font size and font styles such as bold and italic
- effects such as sub- and superscript, underline and small capitals
- layout indicators such as paragraph and tab markers
- tables

These features can be accommodated within XML markup, but are less easily handled within database systems. If they are to be included, consideration should be given to marking up according to meaning rather than display format. For example, the italic *S* used to denote stereochemical configuration might be coded as “<stereo>S</stereo>” rather than “<i>S</i>”.

These XML aspects need further consideration.

Priorities for properties

The priority for a property (reported in XML as an attribute or as element content) defines whether the information is required, recommended or optional. The priority will often depend on the database level (as described in Part I of this standard), with more documentation required at the earlier levels, closer to the original source of the data. The priorities defined in the specifications later in this Technical Annex are based on the information required by food composition data compilers, i.e. those in a level 3 aggregated and compiled database.

There are two priorities:

1. *Mandatory* (M) properties required as they build the core set of documented data that is needed to properly identify and describe the food composition data.
2. *Optional* (O) properties provide additional information describing the data and thereby assisting the user to assess them.

Priorities are also given for whole entity sets (i.e. database tables). If an optional entity set is used, the priorities for its properties apply as indicated in that entity set.

A property may not have the same priority in all circumstances or contexts. For example for the REFERENCE entity, the property ISBN might be defined as mandatory when the publication type indicates a book, but it is not relevant for a journal article. These variations can be noted in the specification either by using the *isa-type* relationship or in the Additional information.

On occasions, mandatory information may not be available, but in this case it should always be reported as such. Many properties have bespoke procedures for this, e.g. the use of “Anon.” for missing authors. Thesauri that are used for properties should always include the term “Unknown”. They may also include a term for “Other”, but the use of this normally indicates that the thesaurus needs to be upgraded.

⁴ The Unicode Consortium: <http://www.unicode.org/>

Data Type Formats

The Eurofoods recommendations (Schlotke *et al.*, 2000) specified data types in terms of the requirements for delimited and fixed-length text files. The constraints of an XML format differ and revised definitions of the data types available in the XML interchange files are given here. These definitions specify the information that can be incorporated into an interchange file and individual designers of food DBMS must decide how compatible data should be managed internally.

The two most fundamental changes relate to text fields. Firstly, as a text-based format, XML uses more reserved characters, but provides an alternative encoding of these. Thus characters such as ", ', &, ;, < and > must appear marked up in forms such as """. Import and export procedures for handling interchange files must apply the necessary translations. Secondly, the arbitrary distinction between strings of up to 255 characters in length and longer *memo* fields is removed. It may be that EuroFIR database managers decide that for some fields the 255-character limit should remain in place, in which case the field data type would be specified as *STR255*.

The data types given in the section 2 tables are specified below:

Data Type	Textual representation	Example
STR	Text String: only limited in length by the specification of the associated DBMS and including all properties previously specified as memo (MEM) fields. XML reserved characters must appear in interchange files as the appropriate entity reference.	
DAT	Date: generally in the form CCYY-MM-DD with leading zeros ⁵ . In case of reduced precision, days (DD) or months (MM) may be omitted starting from the extreme right-hand side ⁶ . Further, special formats may be specified for particular properties. For fields indicated as "DAT(TIM)", time is also allowed, using the format CCYY-MM-DD/hh:mm:ss	1999-01-21 1999-07 1984 1997-12-03/21:35:01
INT	Integer: in the range of +/- 2147483648 (= +/- 2 ³¹)	165
NUM	Decimal Numbers: use the point (. = ASCII 46) to separate decimals. All given decimals must be significant. Do not cut trailing zeros, i.e. trailing zeros should be used to indicate significant decimals.	3.472 5.0
FRC	Fraction: a decimal number between 0 and 1 (0 and 1 inclusive)	0.34
BLN	Boolean "True" or "False", which are the preferred encodings for interchange.	True
THS	Thesaurus Entry: use valid terms/codes in string format	B0123
FIL	Additional (multimedia) Files: Generally files are referenced as WWW URLs. If a leading "http://" or "ftp://" is omitted, a local file in the interchange package is the default, following standard	IMG123.JPG http://xyz.com/images/apple.gif

⁵ According to ISO 8601:1988

⁶ Note that in data management imprecise dates such as 1999-07 must be distinguished from the similar precise date of 1999-07-01.

Data Type	Textual representation	Example
	directory and filename conventions.	ftp://abc.org/docs/manual.doc
KEY / FKY	Keys and Foreign Keys: String format (<i>specific ID</i>) or positive integers (<i>generic ID</i>)	S012 136523
{ }	"Set-valued", i.e. multiple occurrences possible, e.g. {THS}. Each occurrence will occupy a separate instance of the element in the XML format.	Several descriptors from a same LanguageL facet. A book with more than one ISBN code.

3 Entity specifications

Food

The *Food* entity is required and used to describe the foods in the database and any interchange package. Further information on the sample(s) analysed may be reported through the *Sample* entity.

The original⁷ food names in the database language and in English are included in the *Food* entity to keep the *FoodNames* (synonyms) entity optional, but in data management the alternative approach of handling these names together with the synonyms and names in further languages in the *FoodNames* entity remains available. Scientific names may also be handled through the *FoodNames* entity. This is to keep backwards compatibility to former versions of the Technical Annex and its predecessors.

Properties of the Food entity

Property Name	Property ID	Data Type	Prio	Scope Note
Food Name and Identifiers				
Original Food Code	ORIGFDCD	KEY	M	The food code, ID, or abbreviation used to identify the food in the data set
Original Food Name ⁸	ORIGFDNM	STR	M	Food name in original language, as specified by a language attribute using the language code (ref. ISO 639), see FoodNames entity.
English Food Name ⁸	ENGFDNAM	STR	M	Food name in English, with preference given to British English., see FoodNames entity.
Scientific name ⁸	SCINAM	STR	O	Scientific food name. The scientific name should adhere to the following format: Genus species Author[, Year] e.g. <i>Gadus morhua</i> Linnaeus, 1758, see FoodNames entity.
Other food names ⁸				E.g. brand, product name, names in other languages than English, etc, see FoodNames entity
Other Food Identifiers				
CODEX Food Standards	CDXFDSTD	THS	O	Codex Alimentarius Food Standards code, see the Codex Alimentarius Standards list

⁷ The term “original” is used in this report to indicate the (often unique) name(s) and/or code(s) used in the present database/dataset. It does not necessarily bear any reference to eventual food names or food codes used in an eventual external source from which data on food information and values might be taken.

⁸ It is recommended that databases/datasets building on the specifications hold food names in the FoodNames entity, especially when many food names are registered in parallel in the database/datasets. However, for simplicity, when the number of different food names is low, the food names can be held in the food entity.

Property Name	Property ID	Data Type	Prio	Scope Note
				(http://www.codexalimentarius.net/web/standard_list.do?lang=en).
Global Trade Identification Number	GS1	STR	O	Bar codes for article numbers. Primarily applicable to identify specific samples http://www.gs1.org/productssolutions/barcodes/
Article Number				Version ?
E-Number	ENR	THS	O	If food is food additive, code according to the European E-Number system for additive standardisation (included in LanguaL facet A). For more information, see the European Commissions web site on food additives (http://ec.europa.eu/food/food/chemicalsafety/additives/index_en.htm)
INS-Code	INS	THS	O	If the food is a food additive, code according to the International Numbering System for food additives according to CODEX Alimentarius. For more information, see Codex Alimentarius General Standard for Food Additives (http://www.codexalimentarius.net/gsfaonline/index.html?lang=en)
Food Classifications and Description				
Original Food Group Code	ORIGGPCD	FKY	O	The code from the classification system used in the original dataset. In interchange, this system should be provided within the Content section of the file, within the <i>originalFoodClassification</i> element. If several classification systems are used in the database/dataset, and they are not included the current list
Product Type	PRODTYPE	{THS}	M	LanguaL facet A. Allows classification of foods according to multiple systems, such as CODEX, US-CFR, CIAA, EuroFIR, Eurocode, European Food Groups, see below. It is mandatory to have at least one classification system used consistently throughout a dataset. In EuroFIR, the EuroFIR classification must be used.
CODEX Food Categorization System for the General Standards for Food Additives	CDXFDADD	THS	O	Unless already indexed in LanguaL facet A . For more information, see Codex Alimentarius General Standard for Food Additives (http://www.codexalimentarius.net/gsfaonline/index.html?lang=en)
CODEX Classification for Foods and Feeds	CDXFDFD	THS	O	Unless already indexed in LanguaL facet A. For more information, see http://www.codexalimentarius.net/download/standards/41/CXA_004_1993e.pdf
CODEX Food Categorization System for Contaminants	CDXCONT	THS	O	Unless already indexed in LanguaL facet A. For more information see, http://www.fao.org/docrep/meeting/005/AC335E

Property Name	Property ID	Data Type	Prio	Scope Note
				/AC335E15.htm
FAO Food Balance Sheet Classification	FAOFBS	THS	O	Unless already indexed in LanguaL facet A. See faostat.fao.org/
CIAA Food Categorization	CIAA	THS	O	Unless already indexed in LanguaL facet A. The system is outdated and included in the Codex Food Categorization System for the General Standards for food Additives.
Eurocode2	EC2	THS	O	Unless already indexed in LanguaL facet A See http://eurofir.org/eurocode
Food Source	FOODSRCE	THS	O	LanguaL facet B. Individual plant or animal from which the food product or its major ingredient is derived; alternatively a chemical food source
Part of Plant or Animal	PARTPLAN	THS	O	LanguaL facet C. Anatomical part of the plant or animal from which the food product or its major ingredient is derived, e.g., *LEAF*, *ROOT OR TUBER*, *ORGAN MEAT*, *MILK* OR *EGG*; it also includes components of parts, such as *CREAM*, and extracts, concentrates or isolates, such as *PROTEIN EXTRACT* or *SUGAR*
Physical State Shape or Form	PHYSTATE	THS	O	LanguaL facet E. The physical state of the food product (liquid, semiliquid, semisolid, or solid). Solid food products are further subdivided by shape or form. Terms are provided for products that have both liquid and solid components or that incorporate air or other gases
Extent of Heat Treatment	HEATREAT	THS	O	LanguaL facet F. Used to broadly characterize a food product based on the extent of heat applied. Heat treatment affects the flavor and textural characteristics of a food and thus consumer preparation time. Heat treatment causes chemical changes and/or reduction of enzyme and of microbial activity and thus affects food safety and shelflife. Specifics of preparation are covered by *G. COOKING METHOD* and *H. TREATMENT APPLIED*.
Cooking Method	COOKMETH	{THS}	O	LanguaL facet G. E.g. cooked by dry heat, microwave, moist heat, cooked with fat or oil, reheated
Treatment Applied	TREATAPP	{THS}	O	LanguaL facet H. Used to specifically characterize a food product based on the treatment or processes applied to the product or any indexed ingredient. The processes include adding, substituting or removing components or modifying the food or component, e.g., through fermentation. Multiple values can be assigned.
Preservation Method	PRESMETH	{THS}	O	LanguaL facet J. The methods contributing to the prevention or retardation of microbial, enzymatic or oxidative spoilage and thus to the extension of shelf life. Index all methods for which information is available, even if a

Property Name	Property ID	Data Type	Prio	Scope Note
				<p>corresponding descriptor has already been used in *H. TREATMENT APPLIED*.</p> <p>Preservation descriptors refer to the finished food as a whole with these exceptions: (1) if the components of a multi-component food, such as cream pie, are preserved by different methods, index all methods; (2) if chemical preservatives are declared on the label, always index them even if it is known that the preservative was introduced through or is only present in a component or ingredient of the food; and (3) if the preservation method for an ingredient is declared on the label (such as brie cheese made from pasteurized milk) index it. Also use *INGREDIENT PRESERVED BY THERMAL PROCESSING* or *INGREDIENT PRESERVED BY IRRADIATION* when ingredients have been pasteurized, ultra pasteurized, sterilized or irradiated</p>
Packing Medium	PACKMED	{THS}	O	<p>LanguaL facet K. The medium in which the food is packed for preservation and handling or the medium surrounding homemade foods, e.g., peaches cooked in sugar syrup. The packing medium may provide a controlled environment for the food. It may also serve to improve palatability and consumer appeal.</p>
Container or Wrapping	CONTWRPG	{THS}	O	<p>LanguaL facet M. Type of container or wrapping defined by the main container material, the container form, and the material of the liner lids or ends. Also type of container or wrapping by form; prefer description by material first, then by form</p>
Food Contact Surface	FDCTSRFC	{THS}	O	<p>LanguaL facet N. The specific container or coating materials in direct contact with the food. Multiple values can be assigned.</p>
Consumer/Dietary use or Label Claim	LBLCLAIM	{THS}	O	<p>LanguaL facet P. Consumer group, human or animal, for which the food product is produced and marketed; dietary use, where the food has special characteristics, claims or uses or is intended for individuals with particular dietary needs; and label or labeling claims, used when special dietary use factor terms were derived from actual food labels. Multiple descriptors may be used from all three categories</p>
Area of Origin	AREAORIG	{THS}	O	<p>LanguaL facet R. Origin of main raw material or area where food was produced if food is a mixed product. Contains names of the continents, regions and countries of the world, as well as worldwide fishing regions, climate zones and geopolitical designations.</p> <p><i>Country/region most important.</i></p> <p><i>Instructions for use of the three "area" properties?</i></p> <p><i>Specific regulations for beef. - Traceability!</i></p>

Property Name	Property ID	Data Type	Prio	Scope Note
				<i>To be expanded in the next version.</i>
Area of Processing	AREAPROC	{THS}	O	LanguaL facet R. See above
Area of Marketing	AREACONS	{THS}	O	LanguaL facet R. See above
				Area(s) of consumption, sampling, etc. to be expanded in the next version.
Production Environment	PRODENVIR	THS	O	Included in LanguaL facet Z. Agricultural production environment (domestic, semi-domestic, wild), Growing conditions (conventional/organic, outdoor/under glass/hydroponic), Preparation establishment (e.g. food industry, restaurant, street vendor).
Additional Agricultural Production Conditions	AGRICOND	STR	O	To be used for information in addition to that already captured using LanguaL facet Z concerning e.g. soil conditions, watering schemes, feeding, harvesting, slaughtering, ripeness, etc.
Growing Condition	GROWCOND	THS	O	Included in LanguaL facet Z. Conventional/organic, outdoor/under glass/hydroponic.
Degree of Plant Maturity	MATURITY	THS	O	Included in LanguaL facet Z. E.g. ripe, unripe.
Genetically Modified	GENMOD	THS	O	Included in LanguaL facet Z as a sub-item of biotechnologically derived food
Cuisine	CUISINE	STR	O	Considered for future LanguaL facet Q. For example, a food may be described as belonging to Mediterranean cuisine.
Other characteristics of food	ADJUNCT	{THS}	O	Other characteristics included in LanguaL facet Z., e.g. Adjunct characteristics of meat/fish/poultry (colour, cuts), Alcohol content range, Fat content range, Formulated mix, HACCP guide, Material of contact prior to packaging, Presence of casing or rind, Presence of packing medium, Status of food name, etc.
Labelling Information				
Ingredient List	INGLIST	MEMO	O	List of ingredients as given on the package. More detail on the percentage or fraction of a specific ingredient may be given in the Ingredients entity.
Final Preparation	FINLPREP	MEMO	O	Final recommended preparation of food before consumption, e.g. heating a frozen dinner or canned food.
Serving Size	SERVSIZE	STR	O	Refer to amounts, in grams, dl, etc., as specified by the producer.
Servings per pack		NUM	O	This can only be filled in, if you are dealing with a single article/package.

Property Name	Property ID	Data Type	Prio	Scope Note
Serving suggestions		STR	O	
State of Preparation		BLN	STR O	Prepared or unprepared
Product yield after preparation / serving (cooking, dilution, draining)		STR	O	Quantity after preparation (GS1).
Allergen information: allergen type, allergen level		STR	O	Claim(s) concerning possible allergen(s) in the food. Description included in LanguaL facet P.
Dietary claim or use		STR	O	Health, nutrition, or other claim(s), like religious claim(s). Description included in LanguaL facet P.
Other properties of food				
Nature of Edible Portion	EDPORTNAT	STR	O	Which parts of the food are included in Edible Proportion. Use terms in LanguaL facet C
Nature of Waste	WASTENAT	STR	O	Which parts of the food are not edible and have been removed as waste, e.g. rind, bone, stone, peel, liquid from can, outer leaves, etc.
Colour	COLOR	STR	O	Colour values are currently not further specified. More detailed recommendations are planned for further versions.
Generic Image	GENIMAGE	{FIL}	O	The file names of generic images showing foods similar to the food or sample in question.
Specific Image	SPCIMAGE	{FIL}	O	The file names of specific images of the food sample, i.e. the food that was actually analysed.
Producer	PRODUCER	STR	O	Names the manufacturer(s) or producer(s) of the food(s) e.g. a farmer is considered a producer.
Distributor	DISTRIB	STR	O	Use when e.g. producer not stated. .
Remarks	REMARKS	STR	O	Any further remarks, or free text description of food item.

FoodNames

The *FoodNames* entity is optional. The basic *Food* entity, as described above, is required and includes properties for the original and English names of the food, as well as the preferred scientific name. However, if more than a few food names are held in the database/dataset, it

is recommended to use the *FoodNames* entity instead. In addition to the original, English and eventual scientific names – names in other languages, synonyms, regional names, alternative scientific names – may be reported as a collection in this related entity *FoodName*.

As they are a "set-valued" (multiply occurring) property, the food names associated with the food are reported as a collection within the *Food* entity of the data interchange package.

Properties of the *FoodNames* entity

Property Name	Property ID	Data Type	Prio	Scope Note
Original Food Code	ORIGFDCD	FKY	M	Original food code, providing the link to the parent <i>Food</i> instance
Food Name	FNAME	STR	M	A name for the food, excepting those given in the <i>Food</i> entity (preferred original food name, English name and scientific name). The name may be a synonym, a name (which may be the same as that in the original language or in English) in a further language, or a taxonomic (scientific) name. A scientific name should adhere to the following format: Genus species Author[, Year] e.g. <i>Gadus morhua</i> Linnaeus, 1758
Language/nomenclature	LANG	THS	M	Language code of the food name, according to ISO 639, e.g., "fr" for French, "en_uk" for British English, etc. For more information, see the EuroFIR thesauri page on the EuroFIR Technical website, http://www.eurofir.org/eurofir/EuroFIRThesauri.asp . For taxonomic names, use 'tx' for taxonomic nomenclature.
Name Role	PREF	THS	M	Used to distinguish between preferred and further names in the language, including other taxonomic names. Equals 'P' (preferred) or 'S' (synonym), with only one preferred name allowed in a given language. The preferred original food name, English name and scientific name are recorded directly in the <i>Food</i> entity. Brand, product/marketing name, etc. can also be listed in this property, see also attributes in the XML Food Data Transport Package (http://www.eurofir.org/eurofir/Downloads/XML%20Food%20Transport%20Package/EuroFIR_Food_Data_Transport_Package_1_3.pdf)
Name Reference	FNREF	FKY	O	Link to Reference table (RefID) for the publication from which the name was taken.
Name Remarks	REMARKS	STR	O	

Recipe

The *Recipe* entity is optional and can be used to record information about a recipe in connection with a food item in the data set. In database management, recipe information often belongs to a specific entity, but in the data interchange package will be included in *Food* description. However, for sake of clarity, the *Recipe* entity is described separately in the table below.

Property Name	Property ID	Data Type	Prio	Scope Note
Original Food Code	ORIGFDCD	FKY	M	Original food code, providing the link to the <i>Food</i> entity
Original Recipe Code	RECID	STR	O	Recipe ID in database, if different from food code
Recipe Reference	RECREF	{FKY}	M	Link to <i>Reference</i> entity (RefID). Describes the publication(s) on which the recipe was based.
Recipe Procedure	RECPROC	STR	O	Description of recipe in text format. Ingredient list and quantities are stored in the <i>Ingredients</i> entity.
Yield Factor for Water	YIELDWAT	NUM	O	Yield factor describes proportion of the total weight retained after water loss or gain after preparation. Also used for total yield, i.e. no information on the type of yield (water, fat, etc.) is available, if only this is recorded.
Yield Factor for Fat	YIELDFAT	NUM	O	The proportion of the total weight retained after fat loss or gain during preparation
Yield Factor for Alcohol	YIELDALC	NUM	O	The proportion of the total weight retained after alcohol loss during preparation
Reference for retention factors	RETREF	STR	O	The bibliographic reference for retention factors
Remarks	REMARKS	STR	O	Any further remarks in the recipe.

Ingredients

The *Ingredients* entity is optional and can be used to link a derived or aggregated food, food label information, a recipe dish or a composite (pooled) sample to all its contributing foods and their description.

This entity table corresponds to the Ingredient list in proposed recipe documentation (Reinivuo & Laitinen 2007) <http://www.eurofir.org/eurofir/RecipeCalculation.asp>.

Quantities are recorded as % or in grams weight to allow calculations. Recipe documentation (bibliographic reference, recipe procedure and yields) is provided in the *Recipe* entity.

The *Ingredients* entity allows food description by *Full Ingredient Indexing*, when the contributing foods (ingredients) are each described in the *Food* entity. This requires that all

ingredients are linked to items reported through the *Food* entity. *Full Ingredient Indexing* allows for better tracking of specific ingredients in foods and can facilitate intake/exposure studies. (Møller and Ireland, 2000).

The link to the *Food* table is optional for each contributing food (ingredient). This allows food label information (the ingredient list) to be recorded even though the ingredient may not be recorded as a food in the data set. Either the *Food ID* for the ingredient or the *Ingredient Name* is therefore mandatory. However, if the *Food ID* is not given, then the table cannot be used for *Full ingredient indexing* of the composite food.

Properties of the Ingredients entity

Property Name	Property ID	Data Type	Prio	Additional information
Recipe ID	RECID	FKY	M	Link to the <i>Recipe</i> entity
Food ID	FOODID	FKY	M/O	Link to the <i>Food</i> entity. The food record corresponding to the ingredient.
Ingredient Name	INGNAM	STR	M/O	Name of the contributing food or ingredient, e.g. on product package. <i>Ingredient Name</i> is mandatory if <i>Food ID</i> is not given.
Amount of Ingredient	INGAMOUNT	NUM	O	The amount of an ingredient (contributing food) given in weight, or as a fraction or percentage of total recipe weight.
Unit	INGUNIT	THS	O	Mandatory if the amount is given
Rank	RANK	INT	O	Often, the amount of ingredients is not known, only their order (e.g. ingredient lists on packaged foods). In this case, the rank of each ingredient could be given according to the order in the label information (i.e. 1,2,3,...).
Remarks	REMARKS	STR	O	Any further remarks.

Retention factors (entity awaiting final decision)

The retention factor entity lists the retention factors used in recipe calculation. Recommended factors will be defined by EuroFIR, to be used in the preferred calculation method, which includes yield factors at recipe level and retention factors at the ingredient level.

Properties of the Retention Factor entity

Property Name	Property ID	Data Type	Prio	Scope note
ID	FOODID	FKY	M	Link to either a specific EuroFIR food group or <i>Food</i> entity. To be decided.
Component ID	COMPID	FKY	M	Link to <i>Component</i> entity
Preparation Method	PREPMET	STR	M	The preferred content is a Languag facet G descriptor. Currently this property allows the use

of more specific terms defined locally.

Retention Factor	RETFACT	FRC	M
Reference	RETREF	FKY	M
Remark	REMARK	STR	O

Sample

The *Sample* entity is optional and can be used to record information about primary samples, sampling plan and sample handling. In database management, sample information often belongs to a specific entity, but in the data interchange package will be included in *Value* description for reporting original data, as it depends on both food and component. However, for sake of clarity, the *Sample* entity is described separately in the table below.

Property Name	Property ID	Data Type	Prio	Scope Note
Sample ID	SAMPID	KEY	M	Primary key of <i>Sample</i> entity
Sampling Reference	SAMPREF	FKY	O	Link to the Publication table (Publication ID), i.e. a publication describing the Sampling plan
Reason for Sampling	REASON	STR	O	Context of investigation e.g. for clinical, comprehensive, control, or contamination study
Sampling Strategy	SAMPSRAT	STR	O	Brief description of the sampling strategy.
Place of Sampling	PLCECOLL	STR	O	Where the sample were obtained, purchased, harvested, etc.
Date of Sampling	DATSAMPL	STR	O	Date/period when sample(s) was obtained (purchased, harvested, etc). The date should follow the ISO format CCYY-MM-DD but may also include ranges (e.g. 1999 - 2005, 2005-10 – 2005-11)
Primary Sample Unit Size	SPLEUNSIZE	STR	O	E.g. piece of apple, loaf of bread, can
No. of Primary Sample Units	SPLEUNNUM	NUM	O	
Composite sample	COMPOSITE	BLN	O	Defined as 'True' where different samples of the same food have been combined to produce a composite sample for analysis and 'False' where different samples have been analysed separately.
Primary Sample Handling	SPLEHANDL	STR	O	General handling of sample before arrival at laboratory, e.g. sample transport, storage conditions and duration.
Date of Arrival at Laboratory	LABARRIV	DAT	O	
Laboratory Storage	LABSTORE	STR	O	Storage conditions and duration in the laboratory before the start of the analytical process.

Component

The *Component* entity is mandatory.

Properties of the Component entity

Property Name	Property ID	Data Type	Prio	Additional information
EuroFIR Component Identifier	ECOMPID	THS	M	See Component thesaurus. The IDs are linked to other identifiers such as EUROFOODS, INFOODS and ChEBI
Original Component Code	ORIGPCPD	STR	M	The component code, ID, or abbreviation used to identify the component in the original dataset.
Original Component Name	ORIGCPNM	STR	M	The component name in the original language
English Component Name	ENGCPNAM	STR	O	Component name, preferably in British/UK English
Remarks	REMARKS	STR	O	Any further remarks.

Method Specification

In proposals on food composition database management, analytical or computational methods have often been described in a *Method* entity⁹. In data interchange, method documentation is placed at the *Value* level. However, the previous proposals for a *Method* entity have been unclear on how the entity fitted into the overall model of food composition data and thus how it should be linked to related entities. In order to clarify these issues, the entity has been renamed the *Method Specification* entity and covers analytical methods.

The *Method Specification* entity reports detailed information about an analytical method used for one or more values, including descriptions of the key steps applied and other aspects of the method. Many of the analytical method properties have been taken from a CODEX committee report on criteria for evaluation of acceptable methods of analysis for CODEX purposes (ref).

One *Method Specification* record differs from another when any of the information reported differs. However, it will not always be practical to identify an earlier *Method Specification* record that corresponds precisely to the information available for new analyses and thus more pragmatic guidelines are required. Normally the same *Method Specification* will be limited to the analysis performed within a single reference, such as an analytical report or a journal article, or a series of such references, for example when the same laboratory reports similar analyses for several projects following a single analytical specification, hence the entity name.

The *Method Specification* entity represents an optional, but preferred, approach to the handling of detailed method description and assessment. In the basic specification for data

⁹ see also *Schlotke et al*, 2000

interchange the data are output for each value. However, it is likely to form part of the specifications to be used in the future interchange of data resulting from the one-time evaluation of component values reported in a bibliographic reference.

Properties of the Method Specification entity

Property Name	Property ID	Data Type	Prio	Scope Note
Method ID	METHID	KEY	M	Primary key of <i>Method Specification</i> entity
Method Name	METHNAME	STR	M	The name of the analytical/calculation method in English
Original Method name	ORGMNAM	STR	M	The name of the analytical/calculation method in the original language
Official Method	OFFICIAL	STR	O	Gives abbreviation for official method or accreditation organisation/system, e.g. AOAC 985.29 method, NMKL (Nordic) system, COFRAC 60 accreditation (France). This may be multiple occurring data, if more than one organisations's standard is listed.
General Description	GENDESC	STR	O	Free-text field describing the method.
Method Reference	METHREF	FKY	O	Link to <i>Reference</i> entity providing information about the method used by the laboratory
Analytical key steps				
Extraction			O	
Separation			O	
Identification			O	
Detection			O	
Quantification			O	
Other method key steps			O	
Additional descriptors			O	Further method details, including quality index descriptors (to be defined by WP1.3 TG1)
Laboratory Performance				
Analytical Performance Details	ANDETAIL	STR	O	Quantification procedure, confirmation procedure, quality control, use of reference material and methods, etc.
Accuracy	ACCURACY	STR	O	The closeness of the agreement between the result of a measurement and a true value of the measure. It may be assessed by the use of reference materials.
Applicability	APPLICABILITY	STR	O	Specify the matrix, concentration range and, for Codex purposes, the preference to be given to

Property Name	Property ID	Data Type	Prio	Scope Note
				"general" methods.
Precision	PRECISION	NUM	O	The closeness of the agreement between independent test results obtained under prescribed conditions. The values obtained normally encompass both repeatability intra-laboratory and reproducibility inter-laboratory.
Repeatability (intra-laboratory)	REPEAT	NUM	O	The value r below which the absolute difference between two single test results obtained under repeatability conditions (i.e. same sample, same operator, same apparatus, same laboratory, and short interval of time) may be expected to lie within a specific probability (typical 95% and hence $r = 2.8 \times sd$, where sd = standard deviation, calculated from results generated under repeatability conditions.
Reproducibility (inter-laboratory)	REPRODUC	NUM	O	The value r below which the absolute difference between single test results obtained under reproducibility conditions (i.e. on identical material obtained by operators in different laboratories, using standardised test method) may be expected to lie within a specific probability (typical 95% and hence $r = 2.8 \times sd$, where sd = standard deviation, calculated from results generated under reproducibility conditions.
Recovery	RECOVERY	NUM	O	Proportion of the amount of analyte present or added to the test material which is extracted and presented for measurement.
Selectivity	SELECTIV	NUM	O	
Sensitivity	SENSITIV	NUM	O	
Specificity	SPECIFIC	NUM	O	The freedom of the analytical procedure from interference effects. It reflects the ability of the instrumentation to measure only the signal of the determined element.
Remarks	REMARKS	STR	O	Any further remarks.

Value

The *Value* entity is mandatory. A number of statistical parameters can be included (e.g. n, mean, median, minimum, maximum, standard deviation), either as reported or based on raw data if this information is available. A field called "Selected Value" is provided to store a single figure as the best representation of the statistic or other compositional information (e.g. logical zero, a standard fat content in milk or estimated from a related food), based on the decision of a data compiler. For some components such as Energy and Vitamin A activity, the value will have been calculated from the values of contributing components.

Method documentation is often recorded as a separate *Method Specification* entity in a food composition database, but will be documented at the *Value* level in the data interchange format.

Documentation of sampling plans and sample handling is also usually recorded as a separate *Sample* entity in a food composition database, but will be documented at the *Value* level in the data interchange format.

Properties of the Value entity

Property Name	Property ID	Data Type	Prio	Scope Note
Food Identifier	ORIGFDCD	FKY	M	Link to <i>Food</i> entity
Component Identifier	COMPID	FKY	M	Link to <i>Component</i> entity's ORIGPCD
Value ID	VALUEID	KEY	O	Required only if individual sample results are also interchanged as <i>Contributing Values</i>
Selected Value	SELVAL	NUM	M/O	The value that is considered the best representative according to the decision of the data compiler, previously referred to as "Best Location". This value may be based on an evaluation of analytical results from one or more samples, a composite sample or derived from literature data. The characteristics of the value are given by the Value Type. Additional properties, e.g. statistical, can be given (see Statistical Values). Generally, this property is required. In some cases, however, it might not be possible to assign a Selected Value (e.g. value unknown, value undecideable because of distribution showing more than one cluster of values). Selected Value may then be left empty with Statistical values given (e.g. minimum, maximum).
Unit	UNIT	THS	M	From the list given in EuroFIR Unit thesaurus
Matrix unit	MATRIX	THS	M	From the list given in EuroFIR Matrix unit thesaurus
Value Type	VALTYPE	THS	M	The Value Type is designed to further describe the figure in <i>Selected Value</i> or to give a qualitative description of the value when no <i>Selected Value</i> can be given. From the list given in EuroFIR ValueType thesaurus.
Acquisition Type	ACQTYPE	THS	M	Gives categories for the origin of a value, e.g. an evaluated food composition database or table, a scientific publication, analytical results commissioned by the compiler or results calculated by the compiler. Some Acquisition Types usually relate to the Reference associated with the value. From the list given in EuroFIR Acquisition Type thesaurus
Value reference	SOURCEID	{FKY}	M	One or more Reference, using the Reference Code(s) of the associated <i>References</i> record(s).
Date of Generation	DATEGEN	DAT	O	The date when this particular value was generated, e.g. date of analysis or compilation.
Date of Evaluation	DATEEVAL	DAT	O	The most recent date the value in question was evaluated or validated

Property Name	Property ID	Data Type	Prio	Scope Note
Quality Assessment ID	QAID	FKY	O	Link to QA of the value. Editors note! - This may be redundant information, if the Quality Assessment table exists.
Analytical Statistics				Use to provide additional description of the value given in Selected Value.
N	N	INT	O	Number of values contributing to the value given as Selected Value. Defined as the number of analytical portions of the food or the number of contributing values (e.g. values taken from food composition tables).
Analytical Portion Size	ANPORT	STR	O	The size of the sample, prepared from the laboratory sample, from which test portions are removed for testing or for analysis.
No. of Analytical Portion Replicates	NOANPORT	INT	O	The number of times (sub-) samples of a specific analytical sample is being analysed.
Other Statistical results				
Mean	MEAN	NUM	O	The arithmetic mean value of the statistic.
Median	MEDIAN	NUM	O	The median value of the statistic.
Minimum	MIN	NUM	O	The minimal value within the statistic.
Maximum	MAX	NUM	O	The maximal value within the statistic.
Standard Deviation	STDV	NUM	O	Should be used for normal distributions only. Standard Deviation = $\sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}$
Standard Error	STERR	NUM	O	Standard Error = $\frac{\text{Standard Deviation}}{\sqrt{n}}$
Method				Documentation of method is mandatory
Method Type	METHTYPE	THS	M	Type of Method used to generate the value. From the list given in EuroFIR Method Type thesaurus
Method Indicator	METHIND	THS	M	From the EuroFIR Method Indicator thesaurus
Method Parameter	METHPAR	NUM	M/O	Further method information for calculation methods, e.g. NCF and FACF. Mandatory only for calculated protein and fatty acid values.
Method ID	METHID	FKY	O	Link to <i>Method Specification</i> entity. In the data interchange package, information about the analytical or calculation method will be included in the <i>Value</i> entity.
Remarks	REMARKS	STR	O	Any remarks concerning the value.

Quality Assessment

Result of any systematic quality assessment applied by the data provider, usually for values taken from the scientific literature or analytical reports. A description of the quality assessment procedure should be given in the Content section of the exchanged data set. Scores for the quality categories should be recorded in the separate properties specified below, as well as a global Quality Index or Confidence Code for each value.

EuroFIR has developed a scheme for the systematic quality assessment of food composition analytical data reported in the scientific literature or analytical reports, i.e. at the initial database data level. The data in these initial level documents is assessed within seven categories relating to the food, sampling and analytical details reported. A further two categories are assessed when such data are compiled into an FCDB, relating to the matching and representativeness in relation to the FCDB food item.

A score is assigned for each quality category. These scores are reported through the Quality Assessment entity, together with a global Quality Index or Confidence Code for each value. The Content section of the exchanged data set should report the version(s) of the EuroFIR Quality Assurance system applied to the values or the details of any variant or alternative systems used as the basis for any information provided through the Quality Assessment entity.

Quality assessed initial level data may in future be available from an external EuroFIR source. This information associated with the literature articles or analytical reports would be reported non-redundantly, rather than duplicated for each value. For example, the assessment for each reported food and sample would be reported once for the document, using an interchange specification to be developed. Optionally, data management procedures may emulate this approach, with value-level details only being replicated for the EuroFIR Transport Package for compiled data.

The *Quality Assessment* entity is optional. Since at the value level there might be circumstances where more than one assessment record is linked to one value, this entity is defined as a many-to-one relationship with the Value ID as the foreign key.

Properties of the Quality Assessment entity

Property Name	Property ID	Data Type	Prio	Scope Note
Value ID	VALUEID	FKY	O	In cases where the Food ID-Component ID combination is unique, these can act as the unique value identifier. Further considerations are needed for the linking of this entity.
Quality Assessment ID	QAID	KEY	O	Key to QA of the value. Further considerations are needed for the linking of this entity.

Property Name	Property ID	Data Type	Prio	Scope Note
Quality System	QDESC	STR	O	Description of the quality assessment procedure. Decision on whether Quality assessment is done on a global level, on value level, or on both levels (e.g. one for scientific literature and one for manufacturers data) is needed. Further considerations are needed.
Quality Reference	QREF	FKY	O	Bibliographic Reference describing the Quality Assessment system used.
EuroFIR QA categories				
Quality of food description	QFOOD	NUM	O	Quality score for food description
Quality of component identification	QCOMP	NUM	O	Quality score for component identification
Quality of sampling plan	QSAMP	NUM	O	Quality score for food sampling
Quality of sample number	QNUMB	NUM	O	Quality score for number of analytical samples
Quality of sample handling	QHAND	NUM	O	Quality score for sample handling
Quality of method	QMETH	NUM	O	Quality score for method (analytical or calculation)
Quality of Analytical performance	QPERFORM	NUM	O	Quality score for Analytical performance/Quality control
Quality of food matching	QMATCH	NUM	O	Additional Quality score matching reported food to food in database
Quality of representativeness	QREPRESEN	NUM	O	Additional Quality score for representativeness food to national consumption
QI -EuroFIR	QIEUROFIR	STR	O	Global Quality Index or Confidence Code for a value, resulting from quality assessment according to the QI rating system procedure. See separate description.
CC-EuroFIR	CCEUROFIR	STR	O	Global Confidence Code for a value, resulting from quality assessment according to the CC rating system procedure. See separate description.
Remarks	REMARKS	STR	O	Remarks concerning scores obtained for the above Quality categories. This field can also report individual points given for various criteria within the above Quality categories, resulting from a quality assessment procedure.

Contributing Values

The *ContributingValues* entity is optional and holds the individual values used to create a *Selected Value* in the *Value* entity.

Properties of the Contributing Values entity (to be revised in next version)

Property Name	Property ID	Data Type	Prio	Additional information
Value ID	VALUEID	FKY	M	Link to the <i>Value</i> entity, i.e. the Selected value the single value belongs to.
Value	VALUE	NUM	M	A contributing value
Weight	WEIGHT	FRC	O	In case of weighted aggregations, a weight can be stored for every single value.
Remarks	REMARKS	STR	O	Any further remarks.

Reference

The *Reference* entity is required and holds bibliographical information of various publication types, including that for the interchange package itself. Properties linking to the *Reference* entity are included in various entities such as the *Food*, *Value* and *Method* entities.

Properties of the Reference entity

Property Name	Property ID	Data Type	Prio	Scope Note
Original Reference Code		KEY	M	Reference code identifying the Reference record in original data bank, used when linking e.g. an analytical report, journal, book, food composition database, recipe etc. to this reference
Standard Reference Code		{STR}	O	Reference code identifying the Reference record, i.e. DOI (Document Object Identifier) or PMI (PubMed Identifier). The common EuroFIR Bibliographic Reference repository is proposed under CiteXplore (http://www.ebi.ac.uk/citexplore).
Acquisition Type	ACQTYPE	THS	M	Nature/origin of reference (e.g. food composition table, food label, independent laboratory). From the list given in Acquisition Type thesaurus.
Reference Type	REFTYPE	THS	M	Gives categories for types of documents or 'publication media' according to the list given in Reference Type thesaurus (e.g. Book, Journal, Software). The reference type triggers further metadata (see below). From the list given in EuroFIR Reference Type thesaurus
Citation	CITATION	STR	M	Reference citation in a format currently used for scientific publications. This and the above fields are what will be included in the interchange package, not the fields listed below
Title	TITLE	{STR}	M	The title of the reference. Use this property multiple times to provide the title in the original language and in English.

Property Name	Property ID	Data Type	Prio	Scope Note
Authors	AUTHORS	STR	M	Use Vancouver system: Separate all multiple authors by semi-colon (;). For personal names, write the forename or initials after the last name, separated by comma. The attribute may be used for the name of an organisation where this is considered a corporate author, for example "AOAC", or for the abbreviation "Anon." where the authorship is anonymous.
Publication Date	PUBDATE	DAT	M	The year or more precise date when the reference was issued.
Version	VERSION	STR	O	Use this attribute for any versioning system other than publication date or edition number. This attribute is helpful for frequent updates.
Original Language	ORIGLANG	THS	M	The language that the reference was originally written in. According to ISO 639: a 2 character standard ISO language code plus an optional 2 character standard ISO country code separated by a blank character, e.g. "en" for English or "en UK" for British English.
Is a Book				
ISBN	ISBN	{STR}	O	International Standard Book Number
First Edition Date	FSTEDDAT	DAT	O	When was the first edition published?
Edition Number	EDNR	INT	O	What is the current edition?
Number of Pages	NRPAGES	STR	O	Total number of pages
Is a Article in Book				
Book Title	BKTITLE	STR	M	The title of the book in which the article appears. The title of the article is given in the TITLE property.
Editors	EDITORS	STR	M	The names of the editors of the book.
ISBN	ISBN	{STR}	O	International Standard Book Number of the book.
Pages	PAGES	STR	O	The book pages covered by the article, e.g. 45-67
Is a Journal (Issue)				
Long Journal Name	LGJRNAME	STR	O	
Abbreviated Journal Name	ABJRNAME	STR	M	
ISSN	ISSN	{STR}	O	
Volume	VOLUME	STR	M	
Issue	ISSUE	STR	M	
Is a Journal Article				
Long Journal Name	LGJRNAME	STR	O	

Property Name	Property ID	Data Type	Prio	Scope Note
Abbreviated Journal Name	ABJRNAME	STR	M	
ISSN	ISSN	{STR}	M	
Pages	PAGES	STR	O	The pages covered by the article, e.g. 375-383
Volume	VOLUME	STR	M	
Issue	ISSUE	STR	O	
Is a Report				
Series Name	SERINAME	STR	O	Use this property if the report is published within a series of other reports.
Series Number	SERINR	STR	O	The number of the report within the series.
ISBN	ISBN	{STR}	O	
Is a Article in Report				
Editors	EDITORS	STR	M	The names of the editors of the report.
Report Title	RPRTITLE	STR	M	The title of the report. The title of the article is given in the TITLE property.
Series Name	SERINAME	STR	O	
Series Number	SERINR	STR	O	
ISBN	ISBN	{STR}	O	
Pages	PAGES	STR	O	The pages of the report covered by the article, e.g. 45-67
Is a File or Database				
File Format	FILEFRMT	STR	M	Give information about the platform or computer system, the file is compatible to. Also mention the software needed to interpret the file.
WWW	WWW	STR	O	The internet address (URL) of the file (WWW or FTP)
Publication Medium	MEDIUM	STR	O	How is the file distributed: e.g. diskette, CD-ROM, tape, internet, etc.
Is a Software				
Operating System	OS	STR	O	Under which operating system (including version number) does the software run?
Primary Publication Media	MEDIA	STR	O	Storage media of the software, e.g. CD-ROM.
Is a Authoritative Document				
ISBN	ISBN	{STR}	O	

Property Name	Property ID	Data Type	Prio	Scope Note
ISSN	ISSN	{STR}	O	
Valid from	VALID	DAT	O	Since when is the document valid
<i>Is a Product Lable</i>				
Remarks	REMARKS	MEM	O	Any further remarks.

4 Standard Vocabularies (thesauri)

A set of standard vocabularies (thesauri) was defined within the COST Action 99 / EURO-FOODS recommendations for data interchange and management and further amended in the EPIC data interchange project. Each thesaurus consists of a set of concepts that may be arranged within a hierarchy. A concept is represented by a main descriptor – a term representing the concept – and is generally further described with a scope note, additional information, synonyms and related terms.

All thesauri are available on the EuroFIR technical website:

<http://eurofir.org/eurofir/EuroFIRThesauri.asp>

and updated regularly. Rules for updating will be developed.

Thesaurus language

The official thesauri will use English as their main language. It is up to each user to translate thesauri for local usage. However, it is recommended to establish a central authority within each country, or group of countries with the same language(s), to maintain and publish translations.

EuroFIR will keep track of existing translations and provide links to this information on the Internet. EuroFIR has created an eThesaurus server, which holds the EuroFIR and LanguaL thesauri.

Concept description

The following fields are given for each concept within a thesaurus. Code and Descriptor are required, while other fields are *optional*. However, in most cases it is necessary to give additional information (scope-note), in order to unambiguously define and describe a concept.

Concept property	Description
Code	A unique and short alphanumeric code identifying each concept. The code is mainly used in data interchange package and does not necessarily need to be self-explaining. Codes are <i>not</i> case sensitive. Codes are kept unchanged when translating a thesaurus.
Descriptor	A text-string describing the concept. This string, like the code, must be unique since it is the representation of the code to the user and is the preferred term for the concept.
Scope Note	A note attached to a descriptor to indicate its meaning within the indexing language, i.e. any specialities to be considered when applying the concept (e.g. exceptions, relation to other concepts, further clarifications and definitions).
Additional Information	A note attached to a descriptor to provide further definitions of encyclopaedic nature and links to authoritative sources (e.g. GRIN, Mansfeld, CAS, ChEBI).
Synonyms	Synonymous text strings that express exactly the same concept as the descriptor and help people to find a concept (e.g. vitamin B1 and thiamin). Synonyms are usually written lower case.
Related terms	Word or phrase that means nearly the same as the preferred term in the same language.

Dates of entry and changes (as well as "in-aktivation" when a term becomes non-preferred) are logged at the descriptor level.

The thesauri are displayed in four different 'displays'

- **Alphanumerical display**
Display of terms by the alphanumerical code (code, descriptor, scope note, additional information)
- **Alphabetical display**
Display of terms alphabetically (descriptor, code, broader term, scope note, additional information) including synonyms (non-descriptors) and related terms)
- **Systematic display ('tree display')**
Display of the tree structure of the thesaurus (descriptor and code)
- **XML format**
Display defined as EuroFIR Standard Vocabulary entity. This display is for use in IT-systems.

The EuroFIR thesauri are available from the EuroFIR Technical Website's thesauri section (<http://eurofir.org/eurofir/EuroFIRThesauri.asp>), where all displays are available.

The EuroFIR thesauri are also available in printed form as

Anders Møller, Ian D. Unwin, Jayne Ireland, Mark A. Roe, Wulf Becker, Paolo Colombani:
The EuroFIR Thesauri 2008.

Danish Food Information 2008. ISBN 978-87-92125-09-5.

The report can be downloaded from the EuroFIR Technical Website:

<http://www.eurofir.org/eurofir/Documents2008.asp>

5 XML template for food composition data interchange

EuroFIR Food Transport Package Markup Language

The EuroFIR Food Data Transport Package is described in detail in

Anders Møller and Tue Christensen:

EuroFIR Web Services - Food Data Transport Package, Version 1.3.

Danish Food Information 2008. ISBN 978-87-92125-08-8.

The report can be downloaded from the EuroFIR Technical Website:

<http://www.eurofir.org/eurofir/Documents2008.asp>

The term “food composition data” includes all information referring to the description and identification of foods and their content of components. One of the aims of EuroFIR Network of Excellence is to develop a standard that will be used as a framework for compiling and disseminating food composition data that are comparable and unambiguous with respect to the identity and description of foods, components and compositional values. Thus the standard should describe procedures for both data management and data interchange.

The present proposal is based on the COST99/Eurofoods recommendations intended for food composition database management and interchange (Schlotke et al., 2000) and, where appropriate, text from that document has been incorporated or adapted in the present document.

Experience from other projects, e.g. EPIC/ENDB (Charrondière et al., 2002), BASIS (Gry et al., 2002) and from existing standards (e.g. ISO standard on food safety management systems, ISO 22000:2005) is taken into account. The degrees of detail specified in the EuroFIR food composition recommendations will depend on what is feasible on a European level.

All documents are available on the EuroFIR technical website (<http://eurofir.org/>) and updated regularly.

ISBN 978-87-92125-10-1

EAN 9788792125101