

Proposed XML Schema for Nutrition Content of Ingredients (Foods) and Recipes

Introduction

Nutrition data (i.e., the nutrient content of foods) are a vital component in the analysis of dietary intake, recipes and the design and planning of menus. These data are generated by a diverse group of organisations, including food manufacturers and retailers, commodity suppliers, national food tables, research groups and so on. The consumers of these data include these same organisations, nutritional analysis software suppliers, health authorities, consultants and others. There is a noted demand amongst these consumer groups for up-to-date and accurate data.

At present there are no defined standards for the interchange of these data. Whilst some data are available in proprietary electronic formats (e.g., the UK food tables from the HMSO/FSA), these formats are not directly compatible with any standards. Other data suppliers (such as food manufacturers) either currently do not provide electronic formats of their product data, or supply this in other, incompatible formats. To make use of these data requires programming skills and a separate import process to translate the data into whatever format the end-user requires. This is both time-consuming and error-prone. Furthermore, data generators are often inconsistent even between issues of their own data. An example of this is the varying spreadsheet formats used by the FSA in their ‘catch-up’ data on their website.

This XML schema is proposed to provide a standard, extensible format with the flexibility to allow for a wide range of data transfer applications. In its V1-0b draft, it contains what we believe to be a reasonably complete information set concerning the nutritional composition of a food, recipe or other product. We are happy to receive comments and suggestions for data items which appear not to be covered by the schema in its present form.

Please send all comments to Tony_Johns@tinuvielsoftware.com

The Schema

The schema is contained in the file **NUTRITION-v1.0b.XSL**

An example of an instance can be found in the file **NUTTEST.XML**

An extensible stylesheet example can be found in the file **NUTRITION.XSLT**

The schema has a base type of **INGREDIENTS-RECIPES**. This contains individual foods, etc as the type **ITEM**. Each instance can contain multiple ITEMS.

Each ITEM contains the following elements :

FullName	This is the full, descriptive name for the item. For example “ <i>Bacon, rasher, lean and fat, fried, middle</i> ”. It can also be a manufacturer’s product name, e.g., “ <i>Kellogg’s Corn Flakes</i> ”.
Shortform	This is a shortened descriptor, as might be used in a label declaration of ingredients or a recipe. For example “ <i>Fried middle bacon</i> ” or “ <i>Cornflakes</i> ”.
OriginOrg	This is a string identifying the originating organisation. Ultimately it may be possible to enumerate this, but at the moment this may only be of relevance to the generating and consuming users.
OriginCode	A manufacturer’s or other identifying code, of relevance to the generating user.

Notes	Optional general string field for any notes about the product/item.	
FoodCode	If the product is part of the UK food tables, this code would be used to identify the edition and food number.	
FoodGroup	Where the food is part of the UK food tables, or is intended to be used within a software package which allows food searching by product group, this three-letter code identifies the group. The coding is as defined by HMSO/FSA in the UK food tables. These codes will ultimately be provided as an appendix to this proposal and may be enumerated.	
<i>Portions Data</i>	<i>These data provide portion sizes (in grams) and descriptors. As a minimum, a single, medium portion should be provided. Optional small and large portions can also be included.</i>	
Portion	The size of a medium portion in grams.	
PortionDesc	A string descriptor of the medium portion	
SmallPortion	The size of a small portion in grams.	
PortionDescriptor_S	A string descriptor of the small portion	
LargePortion	The size of a large portion in grams.	
PortionDescriptor_L	A string descriptor of the large portion	
<i>Base Nutrients</i>	<i>These data provide information on key macronutrients and energy and represent the minimum dataset for the item.</i>	
Protein	Value	The amount of protein in grams per 100g of the item.
	ValueQualifier	This is normally “none”. However, if the nutrient is present in a significant non-zero but unknown quantity, this item is “In-UnknownQuantity”. If the nutrient is present in trace amounts, this item is “trace”.
	Measure	Enumerated string values for the measure in which the nutrient Value is stated. The options are “grams”, “milligrams”, “micrograms”, “picograms”.
	FieldDescriptor	An optional item showing a field name from a database. This is an item which assists programmers in placing the data into the appropriate database field. It is likely that the string would be agreed between generator and consumer.
TotalNitrogen	Structure as for Protein	
TotalFat	Structure as for Protein	
Carbohydrate	Structure as for Protein	
Sugars	Optional. Structure as for Protein. Expressed as monosaccharide.	
Starch	Optional. Structure as for Protein. Expressed as monosaccharide equivalents.	
EnergyKcal	Value	The amount of kilocalories per 100g of the item.
	Measure	String value : “kcal”

	FieldDescriptor	An optional item showing a field name from a database. This is an item which assists programmers in placing the data into the appropriate database field. It is likely that the string would be agreed between generator and consumer.
EnergyKj	Value	The amount of kilojoules per 100g of the item.
	Measure	String value : "kJ"
	FieldDescriptor	An optional item showing a field name from a database. This is an item which assists programmers in placing the data into the appropriate database field. It is likely that the string would be agreed between generator and consumer.
<i>Other Nutrients</i>		<i>The schema allows for as many other nutrients as available from the generator. This can include vitamins, minerals, fatty acids, fibre, amino acids and other nutrients. There is a common element that can be used for these if required :</i>
Nutrient		
NutrientName.Name		The name of the nutrient. This should be recognisable to the consumer. Ultimately it is proposed to generate an enumeration list of defined names.
NutrientName.Language		The language code for the above name. By default en-GB .
<i>NutrientName.AltName</i>		<i>The name of the nutrient in another language.</i>
<i>NutrientName.AltLanguage</i>		<i>The alternate language code.</i>
Value		The amount of the nutrient in 100g of the item.
ValueQualifier		This is normally "none". However, if the nutrient is present in a significant non-zero but unknown quantity, this item is "In-UnknownQuantity". If the nutrient is present in trace amounts, this item is "trace".
Measure		Enumerated string values for the measure in which the nutrient Value is stated. The options are "grams", "milligrams", "micrograms", "picograms".
FieldDescriptor		An optional item showing a field name from a database. This is an item which assists programmers in placing the data into the appropriate database field. It is likely that the string would be agreed between generator and consumer.
EMS		Energy Content of nutrient. This element is used to identify the energy content of the nutrient. Macronutrients (Protein, Fat, Carbohydrate, Alcohol and their component parts) have an energy density which can be expressed in kcal/g or kJ/g. This element is made up of two parts : EMSValue : The value of the energy density, and EMSMeasure : The measure, either kcal/g or kJ/g.

Metadata

The following items are included in the Schema XSD file :

Creator	The document's creator organisation (Tinuviel Software)
Creator.contact	email contact for above organisation
Date.created	
Date.modified	
SchemaVersion	Currently 1-0b
Documentation	Description of the schema
Format.MediaType	text/xml
Format.syntax	http://www.w3.org/2001/XMLSchema
Format.description	XML Schema, W3C Recommendation 2001
Identifier	NUTRITION-001
Publisher	To be added
Relation.ConformsTo	http://www.w3.org/2001/XMLSchema
Status	draft
StatusDescription	Awaiting approved status
Category	Health
Title	Schema for nutritional content of foods
Type	message