NIH Public Access
Author Manuscript
J Food Compost Anal. Author manuscript; available in PMC 2013 March 01.

Published in final edited form as:
J Food Compost Anal. 2012 March 1; 25(2): 226–233. doi:10.1016/j.jfca.2011.10.003.

A structured vocabulary for indexing dietary supplements in databases in the United States

Leila G Saldanha[1,*,▣], **Johanna T Dwyer**[1,▣], **Joanne M Holden**[2,▣], **Jayne D. Ireland**[3], **Karen W Andrews**[2,▣], **Regan L Bailey**[1,▣], **Jaime J. Gahche**[4,▣], **Constance J Hardy**[5,▣], **Anders Møller**[3], **Susan M. Pilch**[6], and **Janet M Roseland**[2,▣]

[1]Office of Dietary Supplements, National Institutes of Health, 6100 Executive Blvd MSC 7517, Room 3B01, Bethesda, MD 20892, USA

[2]Nutrient Data Laboratory, Agricultural Research Service, US Department of Agriculture, 10300 Baltimore Avenue, Building 005, Beltsville, MD 20705, USA

[3]Danish Food Information, Borgediget 12, 4000 Roskilde, Denmark

[4]National Health and Nutrition Examination Survey/Planning Branch, National Center for Health Statistics/CDC, 3311 Toledo Road, Room 4332, Hyattsville, MD 20782, USA

[5]Center for Food Safety and Applied Nutrition, Food and Drug Administration, College Park MD 20740, USA

[6]MEDLINE/PubMed Database, National Institutes of Health Library, Bethesda, MD 20892, USA

Abstract

Food composition databases are critical to assess and plan dietary intakes. Dietary supplement databases are also needed because dietary supplements make significant contributions to total nutrient intakes. However, no uniform system exists for classifying dietary supplement products and indexing their ingredients in such databases. Differing approaches to classifying these products make it difficult to retrieve or link information effectively. A consistent approach to classifying information within food composition databases led to the development of LanguaL™, a structured vocabulary. LanguaL™ is being adapted as an interface tool for classifying and retrieving product information in dietary supplement databases. This paper outlines proposed changes to the LanguaL™ thesaurus for indexing dietary supplement products and ingredients in databases. The choice of 12 of the original 14 LanguaL™ facets pertinent to dietary supplements, modifications to their scopes, and applications are described. The 12 chosen facets are: Product Type; Source; Part of Source; Physical State, Shape or Form; Ingredients; Preservation Method, Packing Medium, Container or Wrapping; Contact Surface; Consumer Group/Dietary Use/Label Claim; Geographic Places and Regions; and Adjunct Characteristics of food.

[*]Corresponding author: Telephone: + 1 301-435-2920; Fax: +1 301-480-1845 saldanhl@mail.nih.gov (L.G. Saldanha).
[▣]Member of the Dietary Supplement Ingredient Database ad hoc Working Group

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

DISCLOSURE STATEMENT

No conflicts of interest reported by the authors.

Disclaimer: The findings and conclusions in this report are those of the author(s) and do not necessarily represent the views of the Office of Dietary Supplements, NIH, US FDA, CDC, the USDA, or any other entity of the US government.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Keywords

LanguaL; Government; Dietary supplements; Databases; Indexing; Structured vocabulary; Thesaurus; Food analysis; Food composition

1 Introduction

Several dietary supplement databases are available in the United States of America (Dwyer, et al 2008; Saldanha, et al., 2010). To use their data efficiently, it is necessary that both conventional food1 and dietary supplement databases use consistent descriptive systems for classifying them.

The concept of using a faceted thesaurus to index foods originated at the United States (US) Food and Drug Administration (FDA) in the mid-1970s, because of the need to overcome barriers in accessing and exchanging information about food products (McCann et al, 1988). Barriers included differences in food names, food descriptive terms, and nutrient names and units. Because of the enormity and the complexity of the data on a global level, much useful data had become isolated in different and incompatible files (Pennington et al., 1995). The LanguaL™ thesaurus was created to answer the need for a consistent cataloguing system.

The term LanguaL™ is derived from the Latin words *langua* (language or tongue) and *alimentaria* or “food”, so it represents the language for describing food. LanguaL™ is a structured, controlled vocabulary for describing foods, in a systematic organization that simplifies retrieval of information for data analysis. It is based on the principle that items within a database (whether they are dietary supplements or conventional food products) can be described by a combination of uniform terms chosen from “facets” that characterize various mutually exclusive attributes of these products. These facets include food groupings, main ingredient source, physical attributes, other ingredients and processing, packaging and packaging materials, dietary uses, and other miscellaneous characteristics. Within each facet of the thesaurus, descriptors are arranged in a hierarchical order from broader to narrower terms to facilitate indexing and retrieval (Hendricks, 1992; Pennington & Hendricks, 1992; Pennington & Butrum, 1991). “Scope notes” accompany the thesaurus descriptors to explain precisely when a particular descriptor should be used and to ensure uniform use of the terms by indexers and searchers. The thesaurus also provides additional information for many descriptors: these often refer to specific definitions, such as those in the US Code of Federal Regulations (CFR). Currently the European LanguaL™ Technical Committee administers LanguaL™ and the Danish Food Information hosts and maintains the LanguaL™ website (http://www.danfood.info/).

This paper outlines the development of a LanguaL™ Dietary Supplement Structured Vocabulary (LanguaL™ DS Thesaurus) for systematically indexing these products. Dietary supplement database compilers and managers can use this structured vocabulary to classify products for information retrieval and facilitate links to other databases. This paper provides modifications to the food product thesaurus, in order to capture the unique features of dietary supplement products and the preliminary schema for the LanguaL™ DS Thesaurus. The preliminary schema was developed by members from the US Federal Dietary Supplement Ingredient Database (DSID) ad hoc Working Group2 in collaboration with and

[1]Under the Federal Food, Drug and Cosmetic Act and its amendments, food labeling is required for most prepared foods, such as breads, cereals, canned and frozen foods, snacks, desserts, drinks, etc. Nutrition labeling for raw produce (fruits and vegetables) and fish is voluntary. Such products are referred to as "conventional" foods. Dietary supplements are a special category of products that comes under the general umbrella of foods, but which has separate labeling requirements.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

drawing heavily on the work of European experts. Expert opinion was also sought from participants at the post-conference workshop held in conjunction with the 34th National Nutrient Databank Conference in 2010.

Although the classification system described in this paper was developed for use in the US, it can be used in countries where dietary supplements and dietary ingredients are defined in a manner similar to DSHEA and it can be adapted to f it regulatory definitions in place in other countries. For example, in some European countries the terms used to describe dietary supplements are narrower than those provided for in the US. The "extra" terms can be dealt with in the "Scope Notes" that provide them, but clearly state that they are not used in the US classification of dietary supplements.

2 Schema for dietary supplement databases

2.1 Proposed Adaptations of the LanguaL™ Thesaurus for Use in Dietary Supplement Databases

The complete LanguaL™ Thesaurus can be viewed and downloaded through the Danish Food Information's website (http://www.danfood.info/). This thesaurus provides detailed information about each descriptor and the scope of each facet. Presently foods (other than dietary supplements) can be described using descriptors chosen from 14 facets. Of these, 12 are applicable to dietary supplements. These are: A. Product Type; B. Source; C. Part of Source; E. Physical State, Shape or Form; H. Treatment Applied/Added Ingredients; J. Preservation Method, K. Packing Medium, M. Container or Wrapping; N. Food Contact Surface; P. Consumer Group/Dietary Use/Label Claim; R. Geographic Places and Regions; and Z. Adjunct Characteristics of Food. Some LanguaL™ food facets were dropped, because of distinctive differences in the manufacturing processes for conventional foods versus dietary supplements. The two food facets that were dropped are **F**. Extent of Heat Treatment; and **G**. Cooking Method (see Table 1).

The titles of some facets also did not seem fitting for dietary supplements, so modifications to the descriptors were proposed. For example, Facet **H**, Treatment Applied, was provisionally changed to "Ingredients" for the LanguaL™ DS Thesaurus, a s under this facet all ingredients, other than major source, can be indexed.

The scope and descriptors in each facet were chosen to be consistent with the scope of dietary supplements outlined in the US Dietary Supplement Health Education Act of 1994 (DSHEA) (http://www.gpo.gov/fdsys/pkg/BILLS-103s784enr/pdf/BILLS-103s784enr.pdf). According to DSHEA, in the US a "dietary supplement" is defined as "*a product (other than tobacco) intended to supplement the diet that bears or contains one or more of the following dietary ingredients: (A) a vitamin; (B) a mineral; (C) an herb or other botanical; (D) an amino acid; (E) a dietary substance for use by man to supplement the diet by increasing the total dietary intake; or (F) a concentrate, metabolite, constituent, extract, or combination of any ingredient described in clause (A), (B), (C), (D), or (E).*" In addition, we considered the relevant regulations governing dietary supplements as published in the CFR, and US FDA guidance and practice in regulating dietary supplements (21CFR101; primarily 21 CFR 101.36, and 21 CFR 101.4 and 21 CFR 101.9) (http://ecfr.gpoaccess.gov/).

Only ingredients that meet the definition of dietary supplements according to DSHEA are described as "dietary (supplement) ingredients". All other ingredients are treated as "non-

[2]The Dietary Supplement Ingredient Database ad hoc Working Group was formed to assist with the development of the Dietary Supplement Ingredient Database (DSID). Information about DSID can be found at the DSID website: http://ars.usda.gov/Aboutus/docs.htm?docid=6255

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

dietary (supplement) ingredients". Examples of dietary ingredients include vitamins, minerals, botanicals, amino acids, and dietary fiber. Food additives, colors, preservatives and bulking agents, are examples of non-dietary ingredients. According to US FDA labeling regulations, all dietary ingredients must be listed within the Supplement Facts panel. Non-dietary ingredients are listed outside the Supplement Facts panel (see Figure 1). This separation of ingredients is necessary, as the regulations governing dietary and non-dietary ingredients differ. For example, structure/function claims can be made only for dietary ingredients in dietary supplement products. Regulations governing dietary and non-dietary ingredients may differ in other countries.

2.2 Indexing Dietary Supplements Using the LanguaL™ DS Thesaurus

Table 1 shows the 14 existing LanguaL™ facets and the proposed modification to allow indexation of dietary supplements. The table also indicates whether the corresponding information can be obtained from the product label. This point is relevant, because dietary supplement databases, such the National Health and Nutrition Examination Survey (NHANES) dietary supplement database, the National Library of Medicine's consumer oriented dietary supplement database and others, have been developed using information from product labels (Dwyer et al., 2008; Saldanha et al., 2010). In the US, the requirement that labels must reflect product composition is stated in the law according to DSHEA. The manufacturer is legally liable if false or misleading claims are made. A product is also considered a misbranded product if the label does not reflect the product contents. In contrast, DSID is a quantitative database. The information in DSID provides estimates of specific ingredient levels based upon chemical analysis of products.

The classification of dietary ingredients specified in DSHEA can be used to index Product Type under LanguaL™ Facet A. In addition, the primary ingredient by weight that characterizes a dietary supplement product can be indexed under Facet B (Product Source). Any additional ingredients in a dietary supplement product can be indexed under Facet H (Ingredients), where dietary and non-dietary ingredients are listed separately in accordance with FDA regulations.

Table 2 further details the proposed categorization scheme, the narrower (more descriptive) terms, and scope for each facet in the proposed LanguaL™ DS Thesaurus. The Narrow Terms are hierarchical. When the LanguaL ™ DS Thesaurus is operational, it is envisioned that there will be pop-up boxes to guide the indexer with the coding. Based on the level of detail available, the person indexing would check the appropriate boxes and code the information accordingly. For example, in Table 2, the first narrow term column indicates under which category the product would fall. If the product contains fiber and enzymes, it would fall under "5. Other dietary substance to supplement the diet" classification.

Every term in LanguaL™ is assigned an alphanumeric code, which is the indexing system. The letter refers to the facet. Specific terms/attributes that describe a product are assigned a numeric code. For example, the code H0311 refers to Facet H (an ingredient other than the primary ingredient) and 0311 is the code for niacin. Under Facet P, the exact language authorized by US FDA for health claims is included in the Scope Notes so that the indexer will not confuse a health claim with a structure/function claim.

Recently, codes have been added to LanguaL™ to classify dietary supplements at the highest level, i.e. Facet A. The Additional Information field for the new descriptor DIETARY SUPPLEMENT [A1298] contains reference to DSHEA, as well as to Codex Alimentarius and EC regulations. The new DS classification includes codes for all of the terms for indexing dietary supplements in databases in the US as well as narrower terms proposed in the European databases on dietary supplements (Finland, Netherlands, France,

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Denmark, and EPIC). These “extra” terms have Scope Notes that clearly state that they are not to be used in the US classification of dietary supplements. They can also be disabled in the US version for indexing dietary supplements.

2.3 Indexing a Generic Adult M ultivitamin/Mineral (MVM) Product using the LanguaL™ DS Thesaurus

Table 2 also shows how a generic adult MVM product like that shown in Figure 1, would be described using the LanguaL™ DS Thesaurus. Based on information provided in Supplement Facts panel, the product would be described as a MVM (combination) product under Facet A. Calcium carbonate would be identified as the “primary ingredient”, as it is the major ingredient by weight. Information described under Facets B and C is the “source” and “part of the source” of calcium carbonate. The “source” in this case is not specified and therefore assumed to be synthetic. All other ingredients captured in the Supplement Facts panel are dietary ingredients and are listed under Facet H “Ingredients”. All remaining ingredients listed outside the Supplement Facts panel, e.g. cornstarch, silicon dioxide and hydrogenated palm oil, are listed as non-dietary ingredients under Facet H.

In Figure 1, the source of the nutrient or dietary ingredient listed in the Supplement Facts panel is listed outside the Supplement Facts panel with all other ingredients in descending order by weight. The source of a nutrient or dietary ingredient can also be listed in parentheses after the name of the nutrient in the Supplement Facts panel. Both approaches are permitted under US FDA regulations. All other information entered in Table 2 to describe the MVM product, such as packaging, was obtained from the product as purchased.

Although it may appear that the amount of information that is needed to be compiled for each product is enormous, in practice only pertinent information is recorded. The indexer can also indicate that the information is “unknown”, i.e. simply not available, or that a certain facet is “not applicable” to a particular dietary supplement. For example, the geographic source of an ingredient is usually not available on a label (“unknown”). In contrast, Facet C (part of plant or animal) is not likely relevant (“not applicable”) when the main ingredient (facet B) is a chemical (e.g. calcium carbonate). Although a researcher may not search a database for “not known” or “not applicable” information, this distinction is useful information for researchers.

3 Conclusion

The proposed modifications of the LanguaL™ thesaurus described in this paper were used to help build the LanguaL™ DS Thesaurus. This Thesaurus will assist with classifying products and retrieving information about dietary supplements in databases. This tool should be of value to researchers and stakeholders, as it will improve the precision of matching dietary supplement products recorded in food consumption surveys with the composition of these products as indexed in dietary supplement databases. This more precise matching should result in more accurate estimates of nutrient intakes.

There are many advantages to using the LanguaL™ thesaurus to index dietary supplements in databases. Key applications and examples of these applications, some of which are also applicable to food composition databases, are summarized in Table 3.

The LanguaL™ DS Thesaurus enables database developers to catalogue and link their data and facilitate the sharing of databases. It should aid in better estimates of nutrient intakes from dietary supplements, because there will be a uniform approach to how data about dietary supplements products will be described, recorded, and retrieved. However, it is not

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

intended to and will not resolve all the challenges faced in the development and maintenance of dietary supplement databases.

Several considerations in the development of the s chema were discussed during preliminary meetings. A list of these considerations and how they will be addressed in the LanguaL™ DS Thesaurus are summarized in Table 4. The authors welcome comments and suggestions for improving the indexing system.

HIGHLIGHTS

- A shared tool that enables classifying dietary supplement products in databases.
- Aids in a uniform way to how supplements are described, recorded, and retrieved.
- Enables database developers to link their data and share supplement databases.

Acknowledgments

The Office of Dietary Supplements, National Institute of Health funded the development of the LanguaL™ Dietary Supplement Structured Vocabulary for use in the United States.

References

Dwyer JT, et al. Progress in developing analytical and label-based dietary supplement databases at NIH's Office of Dietary Supplements. Journal of Food Composition and Analysis. 2008; 21:S83–S93.

Eurofoods Working Group on Food Description, Terminology and Nomenclature, Report by the COST Action 99. [Retrieved October 9, 2011] LanguaL 2000 Thesaurus. 2000. Report No. EUR 19540. from Danish Food Information's Website: http://www.danfood.info/

Hendricks, TC. LanguaL: An automated method for describing and retrieving data about food. In: Sempolus, AP.; Butrum, RR., editors. International Food Data Bases and Information Exchange. Vol. 68. World Rev Nutr Diet. Basel, Karger; 1992. p. 94-103.

McCann A, et al. FDA's Factored Food Vocabulary for food product description. J Am Diet Assoc. 1988; 88:336–341. [PubMed: 3346493]

Pennington JA, Butrum RR. Food descriptions using taxonomy and the 'LanguaL" system. Trends in Food Science & Technology. 1991 Nov.:285–288.

Pennington JA, Hendricks TC. Proposal for an international interface standard for food databases. Food Additives Contaminants. 1992; 9(3):265–275.

Pennington JA, et al. International interface standard for food databases. Food Additives Contaminants. 1995; 12(6):809–820.

Saldanha LG, et al. Online dietary supplement resources. J Am Diet Assoc. 2010; 110(10):1426–1431. [PubMed: 20869478]

Dietary Supplement Health and Education Act of 1994. [Retrieved October 9, 2011] Pub Law 103–417, 108 Stat. 4325,Bill Number S.784. 1994 Oct 25. from US Government Printing Office Web Site: http://www.gpo.gov/fdsys/pkg/BILLS-103s784enr/pdf/BILLS-103s784enr.pdf:

Agriculture Research Service, US Department of Agriculture. [Retrieved October 9, 2011] Dietary Supplement Ingredient Database (DSID). 2011. from the Dietary Supplement Ingredient Database (DSID) home page: http://dietarysupplementdatabase.usda.nih.gov

US Food & and Drug Administration. [Retrieved October 9, 2011] A Dietary Supplement Labeling Guide. 2005. from the Dietary Supplements Website: http://www.fda.gov/Food/GuidanceComplianceRegulatoryInformation/GuidanceDocuments/DietarySupplements/DietarySupplementlabelingguide/default.htm

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

[Retrieved October 9, 2011] Code of Federal Regulations (CFR), US Government Printing Office, Title 21: Food and Drugs, - Part 101—Food Labeling. from the Government Printing Office Website: http://ecfr.gpoaccess.gov/

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Supplement Facts Serving Size 1 Tablet	
Each Tablet Contains	**% Daily Value**
Vitamin A 3,500 IU (43% as Beta-Carotene)	70%
Vitamin C 100 mg	167%
Vitamin D 800 IU	200%
Vitamin E 35 IU	117%
Vitamin K 50 mcg	63%
Thiamin 1.1 mg	73%
Riboflavin 1.1 mg	65%
Niacin 14 mg	70%
Vitamin B_6 5 mg	250%
Folic Acid 400 mcg	100%
Vitamin B_{12} 50 mcg	833%
Biotin 30 mcg	10%
Pantothenic Acid 5 mg	50%
Calcium 500 mg	50%
Iron 8 mg	44%
Phosphorus 20 mg	2%
Iodine 150 mcg	100%
Magnesium 50 mg	13%
Zinc 15 mg	100%
Selenium 55 mcg	79%
Copper 0.5 mg	25%
Manganese 2.3 mg	115%
Chromium 50 mcg	42%
Molybdenum 50 mcg	67%
Chloride 72 mg	2%
Potassium 80 mg	2%
Boron 150 mcg	*
Nickel 5 mcg	*
Silicon 2 mg	*

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Vanadium 10 mcg	*
Lutein 300 mcg	*
[redacted]	
* Daily Value not established.	

INGREDIENTS:
Calcium Carbonate, Potassium Chloride, Pregelatinized Corn Starch, Ascorbic Acid (Vit. C), Dibasic Calcium Phosphate, Magnesium Oxide, Microcrystalline Cellulose, Crospovidone. Contains < 2% of: Acacia, Ascorbyl Palmitate, Beta-Carotene, BHT, Biotin, Boric Acid, Calcium Pantothenate, Calcium Stearate, Cholecalciferol (Vit. D_3), Chromium Picolinate, Citric Acid, Corn Starch, Cupric Sulfate, Cyanocobalamin (Vit. B_{12}), dl-Alpha Tocopheryl Acetate (Vit. E), FD&C Blue No. 2 Aluminum Lake, FD&C Red No. 40 Aluminum Lake, Ferrous Fumarate, Folic Acid, Gelatin, Hydrogenated Palm Oil, Hypromellose, Lecithin, Lutein, Magnesium Stearate, Manganese Sulfate, Medium-Chain Triglycerides, Modified Food Starch, Niacinamide, Nickelous Sulfate, Phytonadione (Vit. K), Polyethylene Glycol, Polyvinyl Alcohol, Potassium Iodide, Pyridoxine Hydrochloride (Vit. B_6), Riboflavin (Vit. B_2), Silicon Dioxide, Sodium Ascorbate, Sodium Benzoate, Sodium Borate, Sodium Citrate, Sodium Metavanadate, Sodium Molybdate, Sodium Selenate, Sorbic Acid, Sucrose, Talc, Thiamine Mononitrate (Vit. B_1), Titanium Dioxide, Tocopherols, Vitamin A Acetate (Vit. A), Zinc Oxide.

Figure 1.
Generic multivitamin/mineral (MVM) Supplement Facts panel.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 1

Proposed LanguaL™ facets for dietary supplements

Characteristic	Facet letter	Food Product Descriptor	Dietary Supplement Descriptor	Information available from DS[1] labels - Yes/No/Partial
Food Group	**A**	Product Type		Yes
Food Origin	**B**	Food Source	Dietary Ingredient Source (*major or primary ingredient that characterizes product*).	Yes
	C	Part of Plant or Animal	Part of Source (*major or primary ingredient*)	Yes
Physical Attributes	**E**	Physical State, Shape or Form	Physical State, Shape or Form	Yes
Processing	**F**	Extent of Heat Treatment	*Not relevant to dietary supplements*	NA
	G	Cooking Method	*Not relevant to dietary supplements*	NA
	H	Treatment Applied	Suggest using "Ingredients" as alternate facet name for dietary supplement databases.	Yes
	J	Preservation Method	Preservation Method	Partial
Packaging	**K**	Packing Medium	Packing Medium	Partial
	M	Container or Wrapping	Outside Package Container or Wrapping (*physical container or package, wrapping of dietary supplements*).	No
	N	Food Contact Surface	Dietary Supplement Contact Surface: *(Material(s) in direct contact with product.)*	No
Dietary Uses	**P**	Consumer Group/Dietary Use/ Label Claim	Label Claims/Consumer Group/Dietary Use	Yes
Geographic Origin	**R**	Geographic Places and Regions	Geographic Places & Regions *(Place of Manufacture/Origin)*	Partial
Miscellaneous Characteristics	**Z**	Adjunct Characteristics of Food	Adjunct Characteristics of Dietary Supplements - *(Distribution Channels)*	Partial

[1] Dietary Supplements

[2] Dietary Supplement Health and Education Act of 1994

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Table 2

LanguaL™ Dietary Supplement Structured Vocabulary and with indexing of generic multivitamin/mineral (MVM) product shown in Figure 1

Facet	Broad Term (BT)	Narrow Term (NT)	Narrower Term (NT)	Narrower Term (NT)	MVM Example
Description of product					Generic MVM
A. Product Type	Dietary Supplements (USA[1])	1. Vitamin			
		2. Mineral			
		3. Herbal/botanical	Yeast, algae, fungus		
		4. Amino acid/protein			
		5. Other dietary substance supplement the diet			
			Fiber		
		6. Metabolite, constituent, extract, isolate, or combination of any of these.	Hormone precursors; steroid precursors; 7-dehydrocholesterol, lutein, omega-3's, CoQ10		
		7 Combination of any of the above ingredients listed in 1–6 above	MVM		x
			Mineral(s) & botanical		
			MVM & amino acids		
			MVM & botanical		
			MVM & fatty acids		
			Vitamin(s) & botanical		
			Vitamin(s) & fatty acids		
			Other not listed above		
B: Product Source (major or primary ingredient that characterizes product).		1. Algae, fungus, yeast (lower plant forms).			
		2. Animal (fish oil, bone meal, thyroid or glandulars, desiccated liver)			
		3. Chemical (synthetic source)			x
		4. Higher plant (according to botanical genera/species)			
		5. Bacteria (probiotics).			
		6. Not identified			
C: Part of Source (major or primary ingredient)		1. Process: extract, concentrate, metabolite, isolate or isolated part of plant or animal			x

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

Facet	Broad Term (BT)	Narrow Term (NT)	Narrower Term (NT)	Narrower Term (NT)	MVM Example
		2. Anatomical part of plant or animal	1. Lower plants: algae part or fungus part		
			2. Animal body part (such as gland, etc.)		
			3. Plant part (seed, root, stem, leaf, bark, or whole plant)	Flower	
				Leaf	
				Root	
				Bark	
				Seed	
				Whole plants	
				Stem & parts	
				Bulb	
			4. NA		
E: Physical State, Shape or Form		1.Tablet/caplet	Chewable		
			Non-chewable		x
			NA		
		2.Capsule			
		3. Soft gel			
		4. Liquid			
		5.Powder			
		6.Bars			
		7.Other forms.	Wafer		
			Lozenge (includes strips)		
			Gel		
			Teabags		
			Gummies (balls, etc.).		
		8. Encapsulated (coated/covered)			
		9. NA			
H: Ingredients (all other ingredients that are not included in Facet B)		1. Also contains dietary ingredient (as defined by DSHEA).	1. Vitamin	From LanguaL™ list of terms	x

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Facet	Broad Term (BT)	Narrow Term (NT)	Narrower Term (NT)	Narrower Term (NT)	MVM Example
			2. Mineral	From LanguaL™ list of terms	x
			3. Herbal/botanical	Herbs of Commerce as defined in CFR	
			4. Amino acid/protein	Insert list of amino acids	
			5. Other dietary substance to supplement the diet (fiber, enzymes).	Electrolytes	
				Enzyme	
				Fiber	
			6. Metabolite, constituent, extract, isolate or combination of any of these.		
		2. Also contains non-dietary ingredient (everything that is not a dietary ingredient; and **listed below Supplement Facts panel**)	Antioxidants		x
			Preservatives		
			Binding Agent		x
			Bulking agents (fillers, excipients)		
			Colors		
			Emulsifiers		
			Flavors		
			Sweeteners		
			Other		
			Coating materials	Vegetarian	
				Non-vegetarian (e.g. gelatin)	
				Halal/Kosher	
J. Preservation Method		1. Irradiation			
		2. By chemicals			X
		3. By gases, e.g. steam			
		4. Other			
K. Packing Medium		1. Alcohol			
		2. Gas, other than air			
		3. Other			

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

Facet	Broad Term (BT)	Narrow Term (NT)	Narrower Term (NT)	Narrower Term (NT)	MVM Example
M: Outside Package Container or Wrapping (physical container or package, wrapping of DS).	Container or wrapping by form	1. Bottle			x
		2. Blister packet			
		3. Packet/pouch			
		4. Dropper			
		5. Envelope			
		6. Multicontainer package			
		7. Can			
		8. Tetrapac (box)			
		Wrapper, material unspecified			
N: Dietary Supplement Contact Surface: (material(s) in contact with product.)		1. Glass			
		2. Metal			
		3. Paper or paperboard			
		4. Plastic			x
		5. Cotton			
		6. Metal foil			
P: Label Claims/Consumer Group/Dietary Use	1. Claims	1. Nutrient content claims (used FDA approved list for nutrient content claims appropriate for supplements	1. More of nutrient X		
			2. Percent of nutrient X (% of DV)		x
			3. Comparative percent : Nutrient X compared to Y		
			4. Good source of nutrient X (10–19% or more of DV)		
			5. High / Excellent/ Rich in nutrient X (20% or more of DV)		
		2.Health Claims (*NOTE: will use approved FDA language for health claims)*	1. Calcium and osteoporosis		
			2. Folate and neural tube defects		
			3. Soluble fiber from certain foods and risk of coronary heart disease		
		3. Qualified Health Claims*(NOTE: will use approved FDA language for QHC)*	1. 0.8 mg folic acid and neural tube birth defects		

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Facet	Broad Term (BT)	Narrow Term (NT)	Narrower Term (NT)	Narrower Term (NT)	MVM Example
			2. Vitamins and vascular disease		
			3. Selenium and cancer		
			4. Antioxidant vitamins and cancer		
			5. Phosphatidylserine and cognitive dysfunction and dementia		
			6. Omega 3 fatty acids and coronary heart disease		
			7. Monounsaturated fatty acids from olive oil and coronary heart disease		
			8. Green tea and cancer		
			9. Chromium picolinate and diabetes		
			10. Calcium and colon/rectal cancer		
			11. Calcium and recurrent colon/rectal polyps		
			12.Calcium and hypertension,		
			13. Calcium and pregnancy induced hypertension and preeclampsia		
			14. Unsaturated fatty acids from canola oil and reduced risk of coronary heart disease		
			15. Corn oil and corn oil containing products and a reduced risk of heart disease		
		4. Structure function claims. (*Note: Unlike health claims, there is no approved FDA language for these claims.*)	1."General health"		
			2."Immune function"		
			3. "Sports/performance enhancing"		
			4."Energy"		
			5."Weight loss"		
			6."Mental health /brain health"		

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

Facet	Broad Term (BT)	Narrow Term (NT)	Narrower Term (NT)	Narrower Term (NT)	MVM Example
			7."Occasional insomnia"		
			8. "Supports mood"		
			9. "Menopause"		
			10. "Cardiovascular heart health"		
			11. "Joint health"		
			12. "Sexual health"		
			13. "Bone health"		
			14. "Gastrointestinal health"		
			15. "Modulates blood glucose levels"		
			16. "Hair, skin and nails"		
			17. "Vision/eye health"		
			18. Other		
			Structure function claim disclaimer	a. Included	
				b. Not included	
		5. Other claims:	Statement of material facts		
			One or more ingredients is claimed to be standardized.		
	2. Intended User Group		1. Four years and above	All-ages	
				Weight loss	
				Body builders/Performance enhancing	
				Age-specific men	
				Age-specific women	
				Adult	x
			2. Infants		
			3. Children less than 4 years		
			4. Pregnant and lactating women		
			5. Blood type		
			6. Race		

J Food Compost Anal. Author manuscript; available in PMC 2013 March 01.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

Facet	Broad Term (BT)	Narrow Term (NT)	Narrower Term (NT)	Narrower Term (NT)	MVM Example
R: Geographic Places & Regions (Place of Manufacture/Origin)					
Z: Adjunct Characteristics of DS - (Distribution Channels).					

[1] As defined in Dietary Supplement Health and Education Act of 1994 (DSHEA). Pub L 103-417, 108 STAT. 4235, (Oct. 25, 1994)

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 3

Use of LanguaL™ in Dietary Supplement Databases

Advantages	Example
Facilitate user-specified searches that will enable users to combine appropriate descriptors and download relevant data.	A researcher investigating the association between intake of calcium salts and blood pressure can use LanguaL™ to identify sources of calcium in conventional foods and dietary supplements and then download the information of interest to create a specialized database of calcium sources for the research study
Ease in combining food and dietary supplement databases to more accurately estimate nutrient intakes and explain health outcomes in epidemiological studies. Pertinent since food composition databases use LanguaL™ to index foods.	This is an important functionality in determining nutrients of public health interest and in arriving at dietary recommendations such as national dietary guidelines. For example, nutrients, such as calcium, vitamin B_{12}, and folic acid in dietary supplements make important contributions to the diets of pregnant women and older adults.
Facilitate uniformity in reporting research when applied to the National Health and Nutrition Examination Survey (NHANES) dietary supplement database.	The treatment of dietary supplement products is not uniform across studies. Currently, researchers use their own unique approaches to defining and classifying dietary supplements for analysis, making it difficult to compare findings across studies. Because LanguaL™ is a structured, controlled vocabulary for describing products, there will be a consistent approach to entering and classifying products in databases, thus facilitating uniform reporting of research.
Help trace the source of an adulterant, pesticide, contaminant, or food borne illness when a safety problem is identified.	If a dietary supplement product contains an adulterant of interest and the source of that adulterant is known and described in the database, then products containing the adulterant of interest can be identified in the database, and traced in the food supply.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 4

Considerations in the development of LanguaL™ Indexing System for dietary supplements

Consideration	Comment
Populating and maintaining dietary supplement databases is a time-consuming task, due to the large number of products that are introduced, modified, or relabeled each year in the U.S.A.	Since LanguaL™ can facilitate a standardized data cataloguing system, it can promote seamless sharing of data among database developers once label data has been entered, thus avoiding duplicative efforts.
Product type (Facet A), is a very important facet for researchers, since it captures what type of dietary supplement was consumed from dietary intake records. The proposed dietary supplement classification could be expanded to fully describe products based on their intended or primary use, e.g. bone health.	The narrow terms for Facet A were developed to be consistent with the definition of dietary supplements outlined in DSHEA, which is based on product composition. Additional "narrow" terms based on intended use could be added to Facet P. Consumer Use/Label Claims. This facet can thus be used to group products based on their intended use. However, manufacturers often list several uses for a product and the primary intended use might not be indicated or obvious, or differ from consumer use.
Problems arise when indexing dietary supplements that have multiple ingredients grouped under "proprietary formula" in the supplement facts panel. As dietary ingredients can affect the bioavailability of nutrients, the ability to identify all dietary ingredients in a database is considered an important feature for some researchers.	The practice of grouping ingredients under a "proprietary formula" is permissible under existing U.S. FDA regulations. However, this practice will not affect indexing nutrients that are of primary interest to NHANES, since only dietary ingredients without a Daily Value (DV) can be included in a "proprietary formula". DVs are established by the U.S. FDA for labeling purposes.
It is difficult to include possibly significant details when indexing multi-ingredient dietary supplements.	There are two approaches to indexing multi-ingredient products: 1) Index the main ingredient by weight in facets B & C with the remaining ingredients listed in facet H. A drawback to this alternative is that only the main ingredient will have full information recorded. 2) Full ingredient indexing, with all ingredients indexed and listed in descending order by weight, so that full information retrieval can be achieved. The drawback to this alternative is that an additional system analogous to a recipe management system would be required.