Inclusion of Dietary Supplements as a Subcategory of Foods in Food Ontology Systems

Leila Saldanha[1]*, Johanna T Dwyer[1], Richard Bailen[1],

[1]Office of Dietary Supplements, National Institutes of Health (NIH), USA.

*Corresponding author (saldanhl@mail.nih.gov)

Background: In 1994, The Dietary Supplement Health and Education Act (DSHEA) created Dietary Supplements as a new category of regulated food products in the US (See: https://www.gpo.gov/fdsys/pkg/STATUTE-104/pdf/STATUTE-104-Pg2353.pdf). Since then, sales have grown from an estimated $3.7 billion to $46 billion. The Act also created the Office of Dietary Supplements within the National Institutes of Health (NIH) to encourage research on dietary supplement science.

Users of Dietary Supplements: According to the population-based 2015-16 National Health and Nutrition Examination Survey (NHANES), 56% of US adults ages 20+ years use dietary supplements. Use varied by age; among individuals 2-5 years 46%, 12-19 years 28%, and 60+ years 74%. Since half of the supplements used are micronutrient supplements, dietary supplements make a major contribution to nutrient intakes.

Users of the DSLD Database: According to a 2019 analysis, 117 publications in peer-reviewed and technical publications cited NIH's "Dietary Supplement Label Database" (DSLD). The DSLD provides information on the composition of supplements extracted from product labels. Of the 117 publications, 35 (32%) were from federal users in food and nutrition sciences and 82 (68%) were from first authors who were from outside the federal government. Of that 82, 23% were by non-US based authors, 55% were by authors from schools (or institutions) of informatics, medicine, pharmacy or toxicology, 13% from schools of nutrition or public health, and others were 33%. The primary citation use (48%) was in toxicology or chemistry research to identify products that contained a specific ingredient of interest.

These findings provide information about DSLD users to consider in strategies for future directions. Although one intent of DSLD is as a tool for estimating nutrient intakes, no publications used it for this purpose. This suggests the need for interoperability domestically with food composition databases, such as USDA's FoodData Central, and medical/drug

databases using systems such as the Unique Ingredient Identifier (UNII) and Global Substance Registration System (GSRS). Also, the unexpectedly large number of international users suggests the need to be interoperable with supplement, medicine/drug, and food composition databases in other countries.

Adoption of Food Vocabularies/Ontology Systems: The coding of Dietary Supplements provides some unique challenges. Although regulated as a subcategory of foods, DSHEA permitted the addition of ingredients that cannot be claimed within the Nutrition Facts box[1] on food labels, such as botanicals. The current version of FoodOn (which is largely based on the LanguaL Thesaurus) currently does not include botanicals and many other non-Daily Value ingredients that can be listed within the Supplement Facts box on dietary supplement labels. Another challenge is the lack of standardization of how these non-Daily Value ingredients are listed on labels[2]. For example, there were 95 "synonymous" for acai listed on labels in the DSLD. The listing included as the fruit, juice, powder, extract, or as a combination of these forms and their concentrations. US regulations also give dietary supplement manufacturers more flexibility on how they can list the amounts of dietary ingredients within the Supplement Facts box. Unlike food labels, non-Daily Value ingredients can be listed as Proprietary Blends, i.e., the name of the ingredient is listed, but not the amount of that ingredient. Thus, "dosing" information is not available for these ingredients. LanguaL was evaluated for indexing dietary supplements in databases in the US[3], but only three facets are used in the DSLD. These three are product type (Facet A1298), product form (Facet E0154), intended target group (Facet P0026), and type of claim made (Facet P0023). Since dietary ingredients listed on labels are not coded systematically, going forward the ingredients in the DSLD will be assigned UNII codes.

Using Unique Ingredient Identifier (UNII) and Global Substance Registration System (GSRS) in Ontology Formats: GSRS is a collaborative project between the US Food and Drug Administration (FDA), the NIH National Center for Advancing Translational Sciences (NCATS) and international stakeholders, and can be found at https://www.fda.gov/industry/fda-resources-data-standards/fdas-global-substance-registration-system. GSRS provides a system for the defining and identifying substances within medicinal products or substances used for medicinal purposes, including dietary supplements, foods, and cosmetics The GSRS effort is compared with the periodic table of elements in that it enables the use of common terminology that is accepted worldwide so that consumers, clinicians, and researchers can describe or identify the

ingredients contained in food and medicinal and other products that are both domestically and foreign-sourced. The GSRS provides the common identifier for ingredients by using a consistent definition of substances compliant with the ISO 11238 standard. GSRS is housed at NIH's NCATS as part of the Global Ingredient Archival System (GINAS) project and can be found at https://ncats.nih.gov/expertise/preclinical/ginas.

In the GSRS, substances are explicitly defined in a consistent, standardized, scientifically useful manner. First, the substance is classified into one of a few categories or domains (e.g., chemical, protein, plant substance). Then specified criteria (e.g., chemical structure, DNA sequence) are used to distinguish substances within each category. Finally, a substance is assigned a unique alphanumeric number (UNII code), which can be used as a rapid way to refer to that substance in the future. A 'substance' is defined as "any physical material that has a discrete existence, irrespective of origin." As of May 2020, there was information for 109,973 substances in the GSRS system. See https://fdasis.nlm.nih.gov/srs/.

While the GSRS and UNII are not structured as an ontology system, because the codes are rigorous scientific descriptions of substances, they do lend themselves to be transformed/mapped into an ontology format. The GSRS does not develop nomenclature for substances, it references existing nomenclatures. For example, the description of the nutrient iron (the active moiety) is based on WHO's Anatomical Therapeutic Chemical (ATC) classification system (See: https://www.whocc.no/atc_ddd_index/). In the case of iron, the active substance is classified in a hierarchy with five different levels.

Conclusion: We conclude that dietary supplement components should be treated as a subcategory of foods in ontology systems to permit interoperability between dietary supplement and food databases**.**

References:

1. Saldanha LG. *Changes That Matter to You: FDA's New Nutrition and Supplement Facts Labels.* Nutrition Today 2018;53(6):254-260
2. Saldanha LG, Dwyer JT, Bailen RA, et al. *Characteristics and Challenges of Dietary Supplement Databases Derived from Label Information.* J Nutr. 2018;148:1413S-21S.
3. Saldanha LG, Dwyer JT, Holden JM, et al. *A Structured Vocabulary for Indexing Dietary Supplements in Databases in The United States.* J Food Composition and Analysis. 2012 25:226–233.

Keywords:

- GSRS (Global Substance Registration System)
- UNII (Unique Ingredient Identifier)
- GINAS (Global Ingredient Archival System)
- LanguaL Thesaurus
- DSLD (Dietary Supplement Label Database)