The Crop Dietary Nutrition Ontology (CDNO): aligning the domains of production and consumption

Liliana Andrés-Hernández[1], Razlin Azman Halimi[1], Abdul Baten[2], Kai Blumberg[3], Damion Dooley[4] , Ben Liu[1,] Ramil Mauleon[1], Ramona Walls[3], Graham J. King[1]*

[1] Southern Cross University, Lismore, NSW 2480, Australia;
[2] AgResearch, Grasslands Research Centre, Palmerston North, 4410, New Zealand
[3] Bio5 Institute, the University of Arizona, Tucson, AZ 85719, USA.
[4] Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, BC, Canada.

*Corresponding author (graham.king@scu.edu.au).

Achieving food and nutritional security requires the accurate communication of data, information and knowledge, especially where they relate to dietary nutrition. As we have recently outlined (Andrés-Hernández et al., 2020), there are substantial limitations to comparing the relative nutritional value of crops, with poorly integrated and inconsistent data management systems. In particular, there is a mismatch in the ability of current vocabularies, ontological systems and data exchange mechanisms to reflect relationships between underlying plant genetic resources and crop cultivars, agricultural and horticultural crop production systems, post-harvest grading and the subsequent food processing and supply to human diet. This not only affects consumer choice, genetic improvement, agronomy and processing of major food crops, but generates a disproportionate barrier to the assessment, comparison and adoption of alternative food sources from minor or underutilised crops (Azman Halimi et al., 2019). Given that nutritional composition may be assessed at all stages leading to final dietary consumption, there are considerable benefits to making such relationships explicit (Andrés-Hernández et al., 2020).

We have proposed the Crop Dietary Nutrition Ontology (CDNO) as a structured controlled vocabulary for describing dietary nutritional composition and function of crop plants. In developing CDNO (Andrés-Hernández et al., 2020) we wished to ensure compatibility with vocabularies that describe environmental materials and relevant aspects of food supply, processing and consumption. Specifically, in order to facilitate interchange between ontologies pertinent to these different domains, we evaluated the representation of terminologies that

could be associated with nutritional composition within the Environmental Ontology (ENVO) (Buttigieg et al., 2013) and the Food Ontology (FoodOn) (Dooley et al., 2018).

We have proposed a number of revisions, which may result in addition of new terminologies within FoodOn. Our analysis has also highlighted the need to resolve issues relating to how organisms (crop and livestock) are currently represented within FoodOn, and consideration of issues relating to harvested materials. In order to increase compatibility with FoodOn, the class "harvested material", "manufactured product" and "crop plant" are re-used within CDNO. We believe this review of food source terminologies within FoodOn will more closely reflect classes and terminologies associated with the crop dietary nutrition domain.

The aim of CDNO is to increase the dialogue and language shared by stakeholders in agricultural production, plant scientists and breeders, as well as dieticians and nutritionists. We have proposed that CDNO initially be structured with three major classes. These include the class hierarchy [CDNO: dietary nutritional components], primarily to represent chemical constituents of crop harvested materials and their downstream use as food products. The relationships within this class are based initially on the data definitions described within the hierarchical nutrition schema developed by Azman Halimi et al. (2019). For the development of the dietary nutritional components, this class hierarchy will include build on existing design patterns within ENVO that reuse a subset of chemical terminologies from the Chemical Entities for Biological Interest (CHEBI) (Degtyarenko et al., 2008). In addition to those already describe from FoodOn, terms from the Plant Ontology (PO) (Jaiswal et al., 2002) are reused to describe the parts of crop plants being analysed.

The class hierarchy [CDNO: dietary function] will be developed to describe associated functional attributes, including those that may be of nutritional benefit such as anti-oxidants, fibre and low GI, as well as anti-nutritional factors and food toxins. Many of these components and attributes are analysed by plant scientists, either when characterising plant genetic resources, for market advantage in crop breeding, or post-harvest prior to food processing. The third class hierarchy of [CDNO: analytical methods] aims to represent methodologies and protocols used for sampling, extraction, and analysis of nutritional components. We propose reusing terms from the Ontology for Biomedical Investigations (OBI) and the Physico-Chemical Methods and Properties (FIX) ontology for the development of this class hierarchy.

Overall, we hope this initiative will ameliorate the representation and communication of information relating to crop dietary nutritional composition throughout the plant-based

production and food supply chains. We are working with rich datasets for several crop use-cases including comparative analysis of nutritional components within legume grains.