# FoodON: A Global Farm-to-Fork Food Ontology

## The Development of a Universal Food Vocabulary


Emma Griffiths, Fiona Brinkman
Department of Molecular Biology and Biochemistry
Simon Fraser University
Greater Vancouver, Canada

Damion Dooley, Will Hsiao
BC Public Health Laboratory
Vancouver, Canada

Pier Buttigieg
Alfred-Wegener-Institut, Helmholtz-Zentrum für Polar- und Meeresforschung
Bremerhaven, Germany

Robert Hoehndorf
Department of Computer Science
King Abdullah University
Thuwal, Saudi Arabia



***Abstract*— Several resources and standards for indexing food descriptors currently exist, but their content and interrelations are not semantically and logically coherent. Simultaneously, the need to represent knowledge about food is central to many fields including biomedicine and sustainable development. FoodON is a new ontology built to interoperate with the OBO Library and to represent entities which bear a "food role". It encompasses materials in natural ecosystems and food webs as well as human-centric categorization and handling of food. The latter will be the initial focus of the ontology, and we aim to develop semantics for food safety, food security, the agricultural and animal husbandry practices linked to food production, culinary, nutritional and chemical ingredients and processes. The scope of FoodON is ambitious and will require input from multiple domains. FoodON will import or map to material in existing ontologies and standards and will create content to cover gaps in the representation of food-related products and processes. As a robust food ontology can only be created by consensus and wide adoption, we are currently forming an international consortium to build partnerships, solicit domain expertise, and gather use cases to guide the ontology's development. The products of this work are being applied to research and clinical datasets such as those associated with the Canadian Healthy Infant Longitudinal Development (CHILD) study which examines the causal factors of asthma and allergy development in children, and the Integrated Rapid Infectious Disease Analysis (IRIDA) platform for genomic epidemiology and foodborne outbreak investigation.**



***Keywords—food ontology, food safety, food security, food allergy, nutrition***


## I. Introduction

All living things require food to survive. While some organisms can utilize sunlight and chemicals to synthesize their own nutrients, all animals require exogenous food sources. Food is such a critical resource for life that a substantial proportion of time and effort in an animal's life is spent procuring it. Humans have built intricate systems for the production, distribution, acquisition, regulation, treatment, consumption and disposal of food. The relationships between humans and food are complex, and have real consequences for policy, security, education and science. The ways in which we communicate about food can have major impacts in many aspects of life, from cultural practices to purchasing and economics to health and safety. Many food vocabularies have been developed to meet local needs, but differences in their content, structure and purpose limit their global adoption. As such, these resources conflict when information must be shared on a broader basis. A food ontology which seeks to harmonize the semantics behind these vocabularies would standardize information transfer between national and international organizations, and allow integration of diverse data entities.


This work was supported by Genome Canada and Genome BC.


## II. Development of the FOODON Food Ontology

### *A. Potential Use Cases for a Global Food Ontology*

Initial interviews with domain experts (public health agencies, food allergy strategy committee members, agriculture and agri-food specialists, US Food and Drug Administration officials etc) have suggested a number of use cases for which a global food ontology will have immediate and considerable impact. These use cases include food safety (source attribution during outbreaks, contamination traceback, risk assessment), food security (production, storage, processing, distribution, consumption), food allergy, chemical exposure and nutritional assessment, food regulation, trade (import/export/economics), cultural practices, cooking and culinary use, food webs/food cycles and animal husbandry. An example of how a food ontology would have immediate impact is in resolving discrepancies arising from food synonyms (e.g. cold cuts vs deli meats). Such discrepancies can obscure query results and have important consequences in real-time investigations. We will use these use cases to guide development and test the competencies of FoodON as it develops.

### *B. Harmonization of Existing Food Vocabularies*

Several resources and standards for indexing food descriptors currently exist. The Codex Alimentarius is a collection of internationally recognized standards, codes of practice, guidelines, and other recommendations relating to foods, food production, and food safety commissioned by the United

Nations Food and Agriculture Organization (http://www.fao.org/fao-who-codexalimentarius/codex-home/en/). LanguaL, or the Langua aLimentaria [1], was first developed in the late seventies by USFDA's Centre for Food Safety and Applied Nutrition, and is now hosted by a consultancy, Danish Food Informatics, on behalf of various European partners. LanguaL provides 14 main facets, or hierarchies of descriptive terms that have been used to describe around 35,000 foods. It has been internationalized with equivalent terms in Czech, Danish, English, French, German, Italian, Portuguese, Spanish and Hungarian. Europe is also home to FoodEx [2], a food dictionary constructed by the European Food Safety Authority (EFSA). The latest version (Food Ex2) provides a comprehensive classification of terms and is designed to facilitate food exposure assessment. Other related resources include the USDA National Nutrient Database for Standard Reference [3], and the Health Canada Compendium of Analytical Methods, designed to foster compliance of the food industry with standards and guidelines relative to microbiological and extraneous material in foods [4]. Ontological representation of food products is somewhat scattered and not comparably comprehensive. For example, following requests from its users due to the lack of an independent food ontology, the Environment Ontology (ENVO) contained representations of a range of food products [5]. The experimental FoodO ontology was designed to represent the FooDB database describing food items and chemical/nutritional composition (http://aber-owl.net/ontology/FOODO). While these resources are certainly useful, they have only been developed for specific uses and often include different facets and organizational principles. These features limit interoperability and reduce the range of application an integrative and semantically coherent food ontology would possess.

### *C. FoodON Structure and Content*

We are developing FoodON following the principles of the OBO Foundry and Library [6]. Its upper levels will be concerned with food composition, pre-collection (farming or animal husbandry) environment, collection features (season, age of plant or animal, harvesting technology), storage and distribution features as they pertain to food security, culinary preparation and packaging, food safety factors, and consumption patterns. Food-relevant terms are being imported from existing ontologies including NCBI Taxonomy for source animal / plant / fungi / bacterial organism, Uberon [7] for food source part and type descriptors, ENVO for food environmental context, and CHEBI [8] for chemical constituents and contaminants. LanguaL, a traditional database, will be used for food source common names, preservation methods, and a few other facets. The ontology is open and available on GitHub (https://github.com/FoodOntology) such that editors, each with their own allotment of term identifiers, can effectively contribute in parallel and the community can submit term requests and feedback.

## III. FORMATION OF THE FOODON CONSORTIUM

The creation of such a Food Ontology is a large undertaking and is much more tractable through an open-member consortium framework. As harmonization of the food ontology can only be achieved by consensus and wide adoption, we are currently forming an international consortium to build partnerships and solicit domain expertise. An open-membership consortium will be key to developing a sustainable and extensible FoodON. Priorities of the consortium include term contributions, integration of regionally-specific food vocabulary and strategies for international uptake. Subsets of FoodON will be utilized by the Integrated Rapid Infectious Disease Analysis (IRIDA) platform which supports real-time foodborne outbreak investigation, as part of the Genomic Epidemiology Application Ontology (GenEpiO, www.irida.ca). FoodON vocabulary will also be useful for standardizing food descriptions within the Canadian Healthy Infant Longitudinal Development (CHILD) study, which examines the causal factors of asthma and allergy development in children. Interested participants should contact the authors for further details.

## REFERENCES

[1] Ireland JD & Møller A (2010). LanguaL food description: a learning process. Eur J Clin Nutr. 64 Suppl 3:S44-8.

[2] European Food Safety Authority (2015). The food classification and description system FoodEx2 (revision 2). EFSA supporting publication 2015:EN-804. 90 pp.

[3] Ahuja JK, Moshfegh AJ, Holden JM & Harris E (2013). USDA food and nutrient databases provide the infrastructure for food and nutrition research, policy, and practice. J Nutr. 143(2):241S-9S.

[4] Health Canada (2011). "Guidelines for the Relative Validation of Indirect Qualitative Food Microbiological Methods" in The Compendium of Analytical Methods (Vol 1).

[5] Buttigieg PL, Morrison N, Smith B, Mungall CJ, Lewis SE & the ENVO Consortium (2013). The environment ontology: contextualising biological and biomedical entities. J Biomed Semantics. 4(1):43.

[6] Smith B, Ashburner M, Rosse C, Bard J, Bug W, Ceusters W, Goldberg LJ, Eilback K, Ireland A, Mungall CJ, The OBI Consortium, Leontis N, Rocca-Serra P, Ruttenberg A,Sansone S-A, Scheuermann RH, Shah N, Whetzel PL & Lewis S (2007). The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. Nat Biotechnol. 25(11): 1251-5. http://doi.org/10.1038/nbt1346

[7] Mungall CJ, Torniai C, Gkoutos GV, Lewis SE & Haendel MA (2012). Uberon, an integrative multi-species anatomy ontology. Genome Biol. 13(1), R5. http://doi.org/10.1186/gb-2012-13-1-r5

[8] Degtyarenko K, de Matos P, Ennis M, Hadtings J, Zbinden M, McNaught A, Alcantara R, Darsow M, Guedj M & Ashburner M (2008). ChEBI: a database ontology for chemical entities of biological interest. Nucleic Acids Res. 36 (Database issue), D344-50. http://doi.org/10.1093/nar/gkm791

This work was supported by Genome Canada and Genome BC.