USDA Global Branded Food Products Database
Enhancing Information Transparency and Safeguarding Data Quality
Updated August 2019

The utility and benefit of any database is found in the quantity and quality of the data it captures. The partners in the Public-Private Partnership for the USDA Global Branded Food Products Database (BFPD), work with food and beverage companies to deliver quality data for public transparency and use by the research community and others.

What is Data Quality?

Safeguarding data quality is critical to ensuring the integrity of the nutrient composition information of branded products being shared. Quality data is complete, standardized, time stamped, and accurate. These factors are so central to the use and exchange of data that they have been written into the operational standards and "validation rules" of the Data Partners that provide these data: Label Insight and 1WorldSync, leveraging GS1 Global Standards.

Checks for Completeness, Consistency, and Standards

The Data Partners, 1WorldSync and Label Insight, collect and assess product data provided by the product owner. They assess the data to ensure they are properly captured in a complete, consistent, and standardized manner to be used in the USDA Global Branded Food Products Database. This is achieved through rigorous quality assurance processes, including use of the following measures:

1. *"Hard" and "Soft" Data-entry Validations:* "*Hard*" validations are strict restrictions on the type of data entered. For example, a product's net weight must be entered as an exact numerical value and accompanied by a unit of measure, e.g., oz, gal, lbs. When hard validations are not met, further data entry is forbidden. "Soft" validations highlight information that is likely incorrect based on labeling regulations or entry. For example, the absence of a required nutrient(s) or an invalid Global Trade Item Number™ (GTIN® – number contained in the UPC barcode) are considered soft validations. Soft validations will allow data entry to continue to occur.
2. *Outliers and Profiling:* This measure accounts for data that are provided and fall outside of a typical product range (or standard deviation) and that are flagged for review. For example, a chocolate-covered cookie with less than 40 grams or greater than 90 grams of total carbohydrates per 100 grams would fall significantly outside a typical chocolate-covered cookie product and would merit a flag and review.
3. *Random Sampling*: In addition to the above quality assurance measures, thousands of products are pulled each month and sampled for overall data accuracy.

Checks for Timeliness

1. Product Identification: To help ensure product data are up-to-date (time-stamped), brand owners are advised to follow the GS1 GTIN Management Standard. This set of rules provides guidance for managing data for products that are reformulated and change over time. The following is a summary of the most common product changes that require a new GTIN (UPC) for the product provided in the USDA Global Branded Food Products Database. This enables continuous flow of new and updated product data:

   • Branding – company's primary brand name changes
   • Packaging – 20% or more dimensional or gross weight change, additional/removal of certification mark (e.g., Organic, Kosher), pack/case quantity change, pre-defined product assortment changes, or pricing on packaging changes
   • Product – formulation/functionality changes that affect legally required declared information on the packaging of a product and also where the brand owner expects the consumer/supply chain partner to distinguish the difference and changes to legally required net content
   • Promotion – that affect net content (bonus amount of product) or are time critical and affect supply chain handling

2. Product Dates: As stated above, each product included in the USDA Global Branded Food Products Database carries a date and the name of the data partner that was used for the submission of product data. These dates are defined as:

   • Date Available: This is the date when the product record was available for inclusion in the data type.
   • Date Published: This is the date the product record was added to the USDA Global Branded Food Products Database in FoodData Central.
   • Date Modified: This date reflects when the product data was last modified by the data partner

FoodData Central publishes updates periodically. These updates will publish the most recent data available from the data providers.

The goal of the Public-Private Partnership is to enhance public health and the sharing of open data by complementing USDA FoodData Central with nutrient composition and ingredient information of branded foods and private label data provided by the food industry. To this end, the partners work collaboratively to ensure the right validations are in place to maintain quality data. As the database continues to evolve, the partners may add to the current data quality processes to further enhance the product records.