CRITICAL REVIEW

Review of International Food Classification and Description

Jayne D. Ireland*† and Anders Møller‡

*Informatics Centre for Food Quality (AFSSA-Ciqual), Maisons-Alfort, France; and ‡Danish Food Administration, Søborg, Denmark

Received July 7, 1999, and in revised form March 29, 2000

The preparation of reliable data on food requires precise nomenclature and detailed description of foods. Even data of good quality can be a source of error if they are derived from foods that are not clearly defined. Moreover, it is difficult to exchange data on foods, or to understand and compare nutritional status for different countries or individuals, without a coherent description of foods in databases. The present paper reviews the existing international methods of identifying foods in the computerized databases: Codex Alimentarius, CIAA Food Categorization systems, Harmonized Commodity Description and Coding System, Procome, Eurocode-2, INFIC, LanguaL, INFOODS, IIS, COST Action 99 Recommendations. The paper describes and contrasts the various systems, to point out where the systems are complementary, where they are in conflict, and whether they can be linked. © 2000 Academic Press

Key Words: food classification; food description; Codex Alimentarius; Eurocode-2; INFOODS; IIS; LanguaL.

INTRODUCTION

There is a general agreement on the importance of food nomenclature and description. The preparation of reliable data on food requires precise identification of foods. Even data of good quality can be a source of error if they are derived from foods that are not clearly defined (Polacchi, 1987).

The simple food name can be inadequate or ambiguous to those who are not closely acquainted with the local language and culture (e.g., “sweetbread”). A common name may be misleading when the same name is used for different foods in different regions or when it is used for foods having different scientific names (e.g., “catfish”). Likewise, one may not recognize some terms used by people in other parts of the world or maybe even within the same country. The situation is further confused by homonyms, synonyms, identical brand names for different products, and culinary or technological terms.

As most databases employ different methods of identifying foods, it is difficult to exchange data between countries, between organizations within the same country, or even between workers in the same institution. This paper reviews existing food identification systems used in food composition and consumption databases with the

*† To whom correspondence and reprint requests should be addressed. Fax: 33 + 1-49772650. E-mail: j.ireland@dg.afssa.fr
view that an international understanding of standardized food identification could solve many of the problems arising from the mis-identification of foods.

**FOOD CLASSIFICATION SYSTEMS**

In the past, there have been two separate and seemingly opposing methods to solve the problem of food identification: classify foods in “universal” categories or add food description to foods in a database. In the first approach, a variety of food classification systems has been developed. Some of these classifications have been formulated to describe food habits, while others fulfil requirements set by regulatory bodies. Classification systems are often standardized, as they may be based on legal documents, the most standardized being “vocabularies”.

*National and Regional Food Classification Systems*

Most national and regional databases use country-specific food classification systems, based on national criteria, and the food groups may be very specific. This is mainly due to legal aspects and traditions, besides the economic and cultural importance of foods. For example, there is a separate group for coconut products in the food composition tables of the Pacific Islands, groups for bananas, maize, and cornbreads in the Central America and Panama database, and a group for edible insects in the Thai food composition database (Burlingame, 1998). National or regional classification systems are most often difficult to use on an international basis, as the food classes defined may not be applicable to all cultures.

*Codex Alimentarius Food Standards*

The Codex Alimentarius (Codex Alimentarius, 1989) is a comprehensive collection of food standards and related information prepared by the Food and Agriculture Organization of the United Nations (FAO)—World Health Organization (WHO) Codex Alimentarius Commission. The Commission was established in 1962 to implement the Joint FAO/WHO Food Standards Programme, the purpose of which is to protect the health of the consumer and to facilitate international trade in foods. The standards prepared by the Commission are formally submitted to Member Governments for acceptance and incorporation into national food legislation. Codex standards can be used as basic sources of information for the food industry, food technologists, universities, consumers and many other groups interested in the quality and safety of foods. The Codex Alimentarius includes standards for all the principal foods, whether processed, semi-processed or raw, for distribution to the consumer. Some examples are canned tomatoes, chocolate, gluten-free foods, and European Regional Standard for mayonnaise. On the other hand, the Codex Alimentarius is not a food classification system and there are no standards for all foods.

*CIAA Food Categorization System/Codex Food Categorization System*

The CIAA Food Categorization System (CIAA, 1994) is a European approved and accepted classification, developed by the Confederation of the Food and Drink Industries of the EEC. It is a hierarchical food classification system, designed to serve as an allocation tool for food additives as a basis for their authorization at the European Community level. The CIAA Food Categorization System lays the basis for the Codex Food Categorization System and the food classification used in the food additive “positive lists” in the Nordic countries.
The Codex Food Categorization System was presented by the FAO/WHO Codex Alimentarius Commission on Food Additives and Contaminants in the Codex General Standard on Food Additives (Codex Alimentarius, 1996). As in the CIAA system, foods are classified in 16 main food categories, then sub-categories, according to additive authorizations. When the use of an additive is permitted in a certain category, it is automatically permitted in all its sub-categories, unless otherwise stated.

Both classifications cover all foodstuffs, even those where additives are not allowed. They deal with foods as marketed, thus making the classifications interesting also in a food consumption context. Both classifications are food additive driven and hence emphasize processed foods, as the methods of preparation and conservation of foods condition the use of certain additives. There are some special categories for prepared dishes, but they do not cover all prepared foods. The latter are allowed to some extent to contain the additives allocated to the categories to which their ingredients belong.

Codex Classification of Foods and Animal Feeds/Codex General Standard for Contaminants and Toxins in Foods

The FAO/WHO Codex Classification of Foods and Animal Feeds (Codex Alimentarius, 1993) has been developed by the Codex Committee on Pesticide Residues (CCPR). The classification is intended to be as complete a listing of food commodities in trade as possible, classified into groups on the basis of the commodities’ similar potential for pesticide residues (Codex Alimentarius, 1968).

The food categorization system of the Codex Alimentarius General Standard for Contaminants and Toxins in Foods (Codex Alimentarius, GSC) uses the system that was developed in the framework of the CCPR, as it is also suitable for contaminants. It extends the CCPR classification to include processed products but goes no further than type or group level. The GSC is intended to promote harmonization of grouping commodities with a similar potential for residues for which a maximum residue limit can be set. It may also be appropriate for other purposes such as setting maximum levels for other types of residues or for other contaminants in food.

Multi-ingredient manufactured foods containing ingredients of both plant and animal origin are listed as of plant or animal origin depending upon the main ingredients. The food commodities selected for these classifications are mainly those having current or potential significance in international or national trade, but a limited number of commodities of regional importance have also been included.

Harmonized Commodity Description and Coding System

The World Trade Organization’s Harmonized Commodity Description and Coding System (World Trade Organization, 1996) is used for international trade and by national governments to generate trade statistics. The same trade coding system is agreed upon internationally and implemented in all countries. In Europe, it corresponds to the European Combined Nomenclature (European Commission, 1996).

The coding system comprises 20 sections, four of which apply to foods. The titles of the sections, chapters and sub-chapters are only for reference; for legal purposes, classification is determined according to the terms of the heading (e.g., “leeks and other alliaceous vegetables”). Products are listed in order to define customs tariffs, foods as bought, i.e., single, unprocessed foods. The list does not contain all the foods found in nutrient databases and especially lacks prepared foods and food products. To the best of our knowledge, the World Trade Organization’s food coding system is not used in the context of food composition data.
Food Balance Sheets and other Food Classification Systems

Food Balance Sheets are compiled by the FAO (http://www.fao.org/), OECD (http://www.oecd.org/) and EUROSTAT (http://europa.eu.int/eurostat.html) on an annual basis. A food balance sheet (FBS) presents a comprehensive picture of a country’s food supply during a specific reference period. It shows for each food item (i.e., primary commodity and a number of processed commodities) the areas of supply and its utilization. The food supply available for a specific period is the result of the total quantity of food stuff produced, added to the total quantity imported and adjusted to any change in stocks during this period. FBS expresses food consumption in kg/head/year for broadly defined food groups (e.g., “bovine meat”). They have the advantages of providing a harmonized set of data for all member states, being carried out on an annual basis, and providing information for member states who do not have national food consumption surveys.

All Household Budget Surveys (HBS) in the EU member countries use the PROCOME food classification scheme based on the European Combined Nomenclature and transformed by EUROSTAT. The EU DAFNE Project on HBS has enhanced the PROCOME classification scheme, placing more emphasis on goods such as food and improving comparability with other international food classifications (Trichopoulou and Lagiou, 1997). Even so, the HBS classification schemes only cover commodities in broadly defined food groups (e.g., “Other cereals and preparations”) and not foods as consumed.

Eurocode-2 Food Coding System

Eurocode-2 (Poortvliet and Kohlmeier, 1993) is a mono-hierarchical classification of foods according to groups and subgroups that are useful in dietary studies. The Eurocodes are not specific or detailed enough to replace national codes in computerized food composition or consumption databanks, especially for nutritional calculations, or for assessing exposure to contaminants and food additives. A problem with the system is the difficulty in defining logical rules for assigning a given food item to a specific main group (particularly for mixed foods). Anomalies are also created because the categorization policy differs between the main groups. Furthermore, Eurocode-2 contains several food definitions which are not in agreement with standards set by Codex Alimentarius and/or directives issued by the European Commission, thus making its use difficult for governmental bodies. A revision of the draft of the Eurocode-2 Food Coding System (http://www.vfd2.dk/eurocode/) has been undertaken in the framework of the European COST Action 99 (Unwin and Møller, 1998, 2000). This should enable its wider adoption for recording dietary surveys and for food classification/aggregation in composition databases.

Other Food Classification Systems

Other food classification systems worth mentioning are the Classification of Foods and Physical Properties (Jowitt, 1989) and the United Nations Common Coding System (United Nations, 1994) for goods, services, country and currency codes.

Comparison of Food Classification Systems

All of these food classification systems have been designed by and for people who know the foods involved and the uses that will be made of the data. The corollary of this is that information needed by outside users may be absent. Another common
characteristic observed is that their food codes are not specific or detailed enough to replace national codes in comprehensive food composition or consumption databases, especially for nutritional calculations. They can, however, be used to classify and/or aggregate foods in these databases.

Classification systems have been created for different purposes and reflect different legislations. For example, when classifying cheeses, the CIAA system (additive driven) first differentiates unripened, ripened, processed and analogue cheese; the second criterion is the conditioning, conservation and presence of rind. In Eurocode 2 (food consumption surveys), cheeses are first classed depending on their consistency (hard, soft, fresh), then according to their fat content. PROCOME (household budget surveys) simply classifies all cheeses under “cheese”, and CCPR (residue and contaminant driven) under “secondary milk products”. The classifications are often contradictory, and their very existence shows that there can be no single satisfactory international classification system. In other words, there is no single classification system that can serve all the needs of every food composition database compiler.

Another approach is to identify foods in databases by internal codes plus descriptions.

**FOOD DESCRIPTION SYSTEMS**

Food classification and food description may have very different goals, and this leads to very different appearances of the systems. A classification system tends to group or aggregate foods with similar characteristics; it is a tool of the “end-user” of data. A description system, on the other hand, is a tool of the data originator, who wants to give a description of the food, as precisely as possible, without the necessity of aggregating them.

**INFIC/ENFIC System**

Consistent indexing and retrieval can be attained using faceted thesauri (ISO, 1986), in which vocabulary control is achieved by deliberately restricting the scope of terms and through its display of hierarchical relationship. Due to its flexible structure, such a vocabulary can be amended by adding new viewpoints for food description or by including more details within facets. A faceted thesaurus is thus well adapted to describe the features of foods.

An example of a successful implementation of a faceted thesaurus is the INFIC/ENFIC System (Haendler, 1985; Harris et al., 1980). It is essentially a reference tool whenever an international level is to be considered for communicating about feeds (official documents and scientific publications) or for disseminating or exchanging feed composition data. However, its facets and terms (http://www.pi.net/ifs2.htm) are not adapted to human needs.

**LanguaL Thesaurus**

The LanguaL thesaurus is used in the U.S.A. and Europe for numeric databases on foods. Initially called Factored Food Vocabulary (McCann et al., 1988), the thesaurus was begun in the late 1970s by the Center for Food Safety and Applied Nutrition of the United States Food and Drug Administration (Hendricks, 1992). Since 1996, the European LanguaL Technical Committee has administered the thesaurus. Altogether, over 40 000 food products have been described in various countries using the thesaurus.
LanguaL is a multilingual thesaurus (ISO, 1985) organized in 14 facets characteristic of the nutritional and/or hygienic quality of foods. Each descriptor possesses an underlying code that points to equivalent terms in different languages, which renders the thesaurus both language-independent and suitable for use in numerical databases.

The thesaurus has been significantly modified over the last 2 years in order to provide links to international food categories and coding systems (Møller and Ireland, 2000a,b). An official international version of the thesaurus has been published on the LanguaL Internet site (http://food.ethz.ch/langual), where copies of the thesaurus are available upon request (Schlotke, 1996). A user interface allows the search of foods available in the American, Danish, French and Hungarian databases, in order to promote data interchange and provide a useful tool for persons looking for food composition data.

Conversely, many food and nutrition professionals find the thesaurus difficult to use (Deary, 1993). Some facets need further clarification, and it lacks some terms or specific food groups that may be used in national food composition tables. There is also a need for software to search for and index appropriate terms. The European LanguaL Technical Committee is currently addressing these issues.

INFOODS Nomenclature System

The INFOODS Guidelines for Describing Foods were prepared by the INFOODS Food Nomenclature and Terminology Committee in 1987 (Truswell et al., 1991). The purpose of the INFOODS nomenclature system was to provide a framework for the exchange of data between data sources and compilers of food composition databases. The system is a broad, multifaceted and open-ended mechanism designed to capture all information which might be available and which might be of some use to someone.

The INFOODS Guidelines propose criteria for deciding whether a food is “single” or “mixed” and provide different sets of descriptive facets for these two classes of foods. It is an open-ended, free-text food description system, a listing of features or entities that might influence the composition of food and that collectors of data should be encouraged to record. Still, it was always the intention to develop thesauri for the different facets of the INFOODS system, in order to provide the indexer/retriever with a list of possible terms for any aspect. A draft thesaurus of terms for food processing and preparation was produced (Truswell et al., 1986) but did not have extensive circulation for comment and has never been published.

The INFOODS System, or customized forms of it, is used in New Zealand, the South Pacific, several ASEAN countries, two African countries and 10 Latin American countries. It is also being incorporated into the working systems in Middle Asia and South Asia (Burlingame, 1998).

Other Food Description Systems

Several other food thesauri have been constructed in the past for managing bibliographic information: the CAB thesaurus (http://www.cabi.org/catalog/dbmanual/thesaur.htm) used by Nutrition Abstracts and Reviews; the IFIS thesaurus (http://www.ifis.co.uk/index.html) used by the Food Science and Technology Abstracts; the AGROVOC thesaurus (FAO, 1998; http://www.fao.org/agrovoc/) used by the FAO AGRIS and CARIS databanks. They were, however, designed for documentation purposes and do not have the specificity to describe foods in food composition databases. They are not precise enough for the task of food description and are therefore not suitable for identifying foods in nutrient databases.
Comparison of Food Description Systems

The two major systems used to describe foods in food composition databases are the LanguaL thesaurus, with well-defined terms, and the INFOODS system, in free-text dependent on national language. A comparison of these two systems was carried out by INFOODS regional data centre coordinators (Burlingame, 1998). The LanguaL thesaurus scored better in relation to addressing issues of barriers of language and culture, which is also the reason for its adoption in Europe. On the other hand, the candidates in the test judged the maintenance of LanguaL descriptors and codes in a database to be time consuming. However, it is our opinion that this task is a necessary procedure for all thesauri and one that is also used in the INFOODS (Klensin et al., 1989) food component nomenclature system.

The INFOODS System scored better in relation to “friendliness” to data compilers and local usefulness by conventional users of food composition data. As the answers to the questions are given in the free text, the INFOODS System is simpler and quicker to use and does not necessitate looking up terms and codes in lists. Free-text food descriptions allow information that is more detailed and thus not limited by insufficiencies in the choice of terms from a thesaurus.

The demand for a language-independent thesaurus (like LanguaL) and the requirement for a practical, in-the-field system (like the INFOODS System) for food description in databases has led to attempts to link complementary systems, to create a minimum set of standards and a harmonized approach for identifying foods world-wide. Examples of this combined approach are “system mapping” and the “International Interface Standard for Food Databases”, described in the next section.

COMBINED APPROACHES FOR IDENTIFYING FOODS

Mapping Systems

It is possible to create links between food identification systems, by “mapping” one system to another. This has been done in connection with Eurocode-2 and LanguaL, whereby a link or a set of keys define the LanguaL descriptors for each of all the defined Eurocode-2 codes (Moller et al., 1993). A similar mapping was established between the Italian, German and Codex Alimentarius food codes within the CARE Food Safety project (Weigert et al., 1994).

FDA “International Interface Standard for Food Databases”

The United States Food and Drug Administration developed the “International Interface Standard for Food Databases” (IIS) in order to facilitate retrieval of information from food databases and to improve and standardize food descriptions (Pennington et al., 1995). The IIS includes food names, LanguaL terms, recipe information, INFOODS facets and other classification systems. Although it has not yet obtained international acceptance, the IIS is an invaluable step towards the definition of the relevant types of meta-data in the domain of food databases.

COST Action 99 Recommendations for Food Data Interchange

More recently, a Working Group of the European COST Action 99 “Food consumption and composition data/Eurofoods” has prepared Recommendations for Food Composition Database Management and Data Interchange (Schlotke et al., 2000;
The COST Action 99 Recommendations define attributes for food description in more detail than the INFOODS Guidelines and in a more homogeneous structure than the IIS. In all, more than 50 properties that influence the nutritional value of a food have been listed. According to the nature of the food attribute, the description will be in free text or point to terms in a standardized thesaurus (e.g., LanguaL, ISO). Different national languages are accommodated by using multilingual thesauri and classification systems and by allowing the food name to be given in more than one language. Finally, the Recommendations incorporate already existing international standards for international acceptability. The COST Recommendations thus yield food identification that is detailed, structured, flexible and suitable for use in numeric database.

CONCLUSIONS AND FUTURE WORK

The first breakthrough in international food identification has been the clear recognition of the advantages of using a multifaceted approach for identifying foods in food-related databases (composition, consumption). The second breakthrough has been the recognition of the need to include alternate classification/description systems.

Criteria for such a combined system were laid down by the IIS. They were further stated at the workshop on “Food description, Nomenclature and Terminology” during the Second International Food Data Conference: encompassing several parallel complementary schemes, structured, robust to accommodate different national languages, flexible for use by all users and for all types of foods, specific enough to avoid misclassification, adequately documented, and internationally acceptable.

Work in the field of food identification in food composition databases will continue through an international IUNS/FAO Task Force, as was proposed at the Third International Food Data Conference. This Task Force will have the task of overviewing and focusing the work done on food classification and description in order to harmonize international use of these issues.

REFERENCES


Codex Alimentarius Commission, GSC. General Standard for Contaminants and Toxins in Foods. Codex Alimentarius Annexe V.


