DESIGNING A COMPUTERIZED NUTRIENT DATABASE

Nanci D. Garoon, M.S., R.D.

Kraft, Inc.

Nutrition has entered the computer age. While nutrient database systems have traditionally been used by institutions to plan diets, evaluate school lunch menus, and calculate an individual's daily nutrient intake, these systems have become increasingly popular in the food industry as tools for formulating new products or recipes, generating nutrition label information, comparing the nutrient content of competitor's products, determining marketing trends and supporting advertising claims.

Approximately four years ago, the Nutrition Research Laboratory at Kraft, Inc. identified the need to develop a novel computerized nutrient database to supplement the systems already in place. The strategic planning process began and by January, 1985 the development of an in-house computerized nutrient database was approved by management.

The purpose of this presentation is to describe how an on-line interactive computerized database system was created at Kraft, Inc. The presentation will:

1. identify the problem;
2. describe how the system was created; and
3. illustrate key functions of the system.

The first step toward creating the database was to identify exactly what the situation was and why something had to be done. The Nutrition Research Group at the Kraft Technology Center is not only a true research laboratory but also a very active service group. That is, the group is continuously asked to provide information on the nutrient composition of generic foods and food categories, on competitor's products and, of course, on the entire line of Kraft products.

Certainly, many of you have seen a Kraft-sponsored television special and heard a recipe from the famous Kraft Kitchens. Calculating the nutrient content of individual food items as well as the recipes from the Kraft Kitchens are routine activities in the group. The mechanism in place to accomplish these tasks was manual, slow and narrow in capability. There was no doubt that an interactive on-line computerized nutrient database with query capabilities would reduce the time spent searching for and calculating nutrition information.

After gathering the facts and analyzing the significance of each, a list of needs and requirements of the database was generated. At the very least, the database had to:

1. calculate the nutrient content of selected food items on a per serving and per 100 gram basis;
2. calculate the RDA or USRDA of a generic, branded, ethnic or fast food product;
3. create and store recipes using either common measuring units (cups, tablespoon) or metric units, and calculate a nutrient profile on a per serving basis;
4. provide an editing capability that would allow a user to add, modify or delete a food item; and
5. provide an updating feature.

The next step was to determine whether or not there was an existing system suitable for our purposes. If any of you have perused the 1986 (5th) edition of the Nutrient Data Bank Directory you will see that last year, 99 different systems were listed in the directory and (as stated in the preface) this listing is not comprehensive. A number of databases were screened and evaluated, and a decision was made to purchase the Michigan State University (MSU) Nutrient Database. This decision was based on the following attributes of the MSU database:

1. Approximately 5,000 foods including ingredients, brand-specific products and fast food items are listed.
2. Fields exist for 74 different nutrients.

3. The source of the data is documented and is extracted from USDA handbooks, periodical literature and food manufacturers.
4. Regular (yearly) updates are available.

Hardware and software issues were not important for the development of this Kraft system. We were not concerned about program languages or implementing outside computer services. The goal was to purchase an existing database and augment it to meet our specific needs in a Kraft environment. We defined our own processes and are maintaining our own database.

In August, 1985 the first MSU computerized data tape arrived at the Kraft Technology Center. The data was in a flat or sequential file format. Again, no software programs were purchased to access and manipulate the information.

Documentation accompanying the tape described how the data was blocked. The data was divided into 80 unique fields. This slide illustrates the layout form:

Note: 1. Group-item number (unique)
2. Field for source code
3. Weight relationship fields
4. Nutrient content fields
   - 10 characters in length
   - implied decimal at hundredth place

A comparable form was designed and Kraft data was transcribed into a structure analogous to that of the MSU data structure. This information was subsequently key-punched. At this stage, a software specialist was hired to mesh the Kraft data with the MSU data and create "The Kraft/MSU Hybrid Nutrient Database".

Identifying data requirements and designing informational processes took several months of work with a consultant. A relational database management system (VAX-Rdb) as opposed to a hierarchical or network (VAX-DBMS) type database structure was selected for these reasons:

1. Our system was small to medium in size. Hierarchical database management systems are usually applied to large systems.
2. Hierarchical or network database management systems require more technical expertise than relational databases. A less experienced staff person is needed to implement a relational database.
3. For a relational system, ad hoc queries make up the majority of database activity. Rdb forms relationships as they are requested, they are not rigidly built into the database. Conversely, in a hierarchical or network model the data items and their relationships must be hard coded making them inflexible. Therefore, a relational model was preferred because it allowed the user to redefine data and data relationships. A relational database is a better fit when you don't know ahead of time exactly how the data will be used.

In addition to defining the exact functions the system was to perform, many other issues surfaced during the design phase. Two of these were first, how users were going to query the system and second, how to identify Kraft data as separate from MSU data.

An example of the dilemma surrounding database query had to do with a user requesting information on Kraft Mild Cheddar Cheese? or was that Kraft Natural Mild Cheddar Cheese? or even Kraft Mild 100% Natural Cheddar Cheese? This problem was solved by adding three "new" fields to the database:

1. class;
2. group; and
3. sub-group.

A class, group and sub-group name was assigned to each of the over 5,000 food entries comprising the Kraft/MSU Hybrid Nutrient Database. Therefore, one way for a user to access

the system is by declaring a class, group and sub-group category. The system searches the database and generates a list of items that fit the specified categories. The user scrolls down the list and selects the item or items of choice.

The documentation issue was easily solved by using the established source field in the MSU database to specify "Kraft CA" (calculated) and "Kraft AN" (analytical) data. This delineation not only helps to maintain the integrity of Kraft data, but is also used as a distinguishing feature when the Kraft/MSU Hybrid Database is updated with new Michigan State University data.

Several months of programming passed before the system was finally ready for installation and testing. Today, the system is implemented and can perform a number of functions. This list illustrates a few key features:

1. Retrieve nutrition information for food items.
2. Rank selected database items by the nutritional content of a selected nutrient. The user specifies the food items or categories to rank and the nutrient to search.
   For example: User selects ice cream, yogurt and cheese and directs the computer to rank these items according to the amount of calcium in a serving.
3. List the food items that have a selected nutrient in an amount less than or greater than a specified RDA or USRDA.
   For example: User requests a list of all the ready-to-eat cereals that provide greater than 25% of the RDA per serving for children age four.
4. Determine the nutrient content of a recipe. The user must enter the ingredients, the quantity of each ingredient and the number of servings in the recipe. The system displays requested nutritional information on a per serving basis and prompts the user for any ingredient, nutrient or quantity changes.
5. Update the Kraft/MSU Hybrid Database from new MSU data. This function runs in batch mode. The system first determines if the record is a new or existing food item. If it is a new food item the record is added in the correct sequence to the database. If it is an existing food item, the system replaces the existing record with the new record.

There is one task yet to complete before use of the Kraft/MSU Hybrid Database is in full force. A "User's Manual" will be written to provide step by step instructions on each query option. This document will describe both the functional and operational characteristics of the database. All users will receive a copy of the manual and receive on-line training so that shortcuts and/or nuances in the query capabilities can be individually described.