INFOODS: A 1986 STATUS REPORT

William M. Rand
INFOODS Secretariat
Mass. Inst. Tech.
Cambridge, Massachusetts

INTRODUCTION

INFOODS stands for "International Network of Food Data Systems". It was put together a few years ago to serve as a focal point for those who are, or should be, interested in what is in the foods we consume. Its goal is to focus attention and effort on improving the quantity, quality, and accessability of food composition data.

In terms of the future, INFOODS is trying to create an environment -- in essence, the essential machinery -- so that anyone, anywhere in the world, can ask questions about what is in foods and get reasonable answers in a reasonable length of time. What we are working towards is the time when someone can write a letter, pick up a telephone, or sit down at a terminal and make a request for some aspect of food composition data and get a satisfactory response. In order to give this rather lofty goal more substance, let us consider some of the questions that might be asked (and indeed are asked now). This, unfortunately, leads directly into consideration of some of the problems with the current situation.

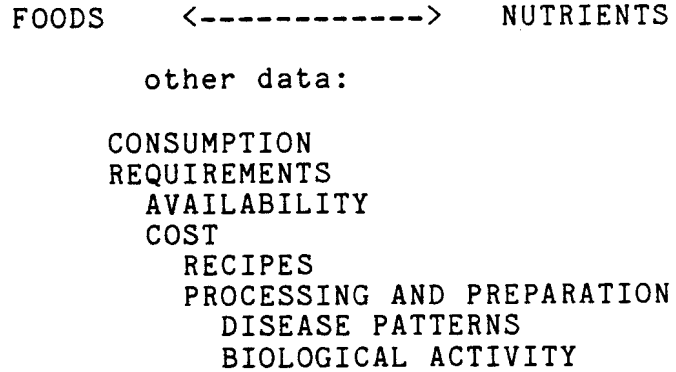
NEEDS FOR FOOD COMPOSITION DATA

For a start, the basic questions are those relating foods and nutrients. Someone wants to know what nutrients are in specific foods, or what foods contain what nutrients. For example one might want to know what is in fruit, and desire a matrix of data; or what is in a specific fruit, for example, an apple, and consider a row of data a satisfactory answer. Alternately, one might focus on a single nutrient, wanting to know the vitamin A content of fruit, or quite specifically, the vitamin A content of an apple.

Parallel to these questions are those essentially going the other way -- wanting to know, for example, a good source of vitamin A or what foods someone with a specific metabolic disorder should avoid, or what foods would make up an interesting, nutritious diet.

Beyond these fairly straighforward questions lie a host of more complex ones, which require data beyond what foods have what nutrients. [See Figure 1.]

Figure 1



Foods, being at the very foundation of our existance, are involved with an incredibly varied spectrum of interesting questions.

There are questions about consumption and requirement:

How do people in Keyna consume their iron requirement?

What foods consumed in the Far East are made from peanuts?

There are questions about the cost and availability of foods:

Are there good local substitutes for the wheat that so many countries import?

What foods on the international market could reasonably be fortified with vitamin A?

Then there are the questions that ask about how foods are prepared and processed:

What fried foods are consumed in Peru?

What foods in India are prepared in copper vessels?

Finally, there are those questions that require information on the geographic patterns of disease, or on the biological interactions between foods, nutrients, and other components of foods:

> Is there any correlation between fish oil consumption and atherosclerosis?
>
> Are any foods/components that are consumed in Latin America correlated with the incidence of gastric cancer?

This list, of course, could go on and on, in fact one of the problems that we continually face is that the questions are really more interesting than the data. I think that this is one of the reasons that the data have often been neglected. Of these questions, good data simply do not exist for many, and even where there are data our problems are not over. Let me consider what appears to be a simple question in a little more detail. [See Table 1.]

Table 1

HOW MUCH VITAMIN A IS IN AN APPLE?

US	5	RE	VITAMIN A
	53	IU	VITAMIN A
UK	0	MCG	RETINOL
	23	MCG	CAROTENE
GERMANY	0.000	MCG	VITAMINE A
	0.047	MG	CAROTENE
INDIA	0	MCG	CAROTENE
CHINA	.08,.4	MG	CAROTENE
JAPAN	0	MCG	RETINOL
	11	MCG	CAROTENE
	0	IU	RETINOL POTENCY
KOREA	10	IU	
LATIN AMERICA	10	MCG	VITAMIN A

"How much vitamin A is in an apple?" seems innocuous enough, however, if you are taking an international view, and go to specific country tables, you find that things are not at all straightforward. Different tables have different numbers. The next logical step is to ask why? Since there are two components of this simple question -- the food and the nutrient, we can break the question into two parts: what is an apple and what is vitamin A.

Table 2 shows one aspect of this -- and I chose apple as an example because I thought that it would be well-defined, given that you can move beyond wanting to know what is in the apple you ate yesterday, or the one you may eat tomorrow.

Table 2

WHAT IS AN APPLE?

USA	MALUS SYLVESTRIS
UK	MALUS PUMILA
GERMANY	PIRUS MALUS
INDIA	MALUS SYLVESTRIS
CHINA	MALUS PUMILA, M. DOMESTICA, M. DASYPHYLLA, M. COMMUNIS
	PYRUS MALUS
KOREA	MALUS PUMILA
LATIN AMERICA	MALUS SYLVESTRIS
CODEX ALIMENTARIUS (FAO/WHO)	MALUS DOMESTICUS

If you are interested in the vitamin A content of a generic apple you might want to look at how each table defines "apple". At first glance it would seem that the differences between the tabled entries might follow from the fact that each is looking at something different, even different genera, although all are called "apple". However, if you go into things a bit more deeply, you find that probably these are all the same species at least, the differing names reflecting taxonomic battles over priorities rather that real differences. Of course, there are most probably very great differences between the apples that were analyzed, depending on variety or cultivar and where and how grown, but the point I would like to emphasize is that the tables have introduced a new confusion into all this, probably quite independent of the data, and moreover, few give any information that would really help the user sort out why the data differ.

I could have prepared a similar table about what is vitamin A -- for example pointing out that many tables use 1/6 to convert from beta-carotene to retinol equivalents, while, at least in India, 1/4 is used, but this gets very complicated. So, to move from the particular to the general, considering the whole area of food composition data, Table 3 shows six major problem areas that INFOODS has identified at the moment.

Table 3

FOOD COMPOSITION DATA - PROBLEM AREAS

1. GETTING MORE AND BETTER DATA
2. CATALOGUING EXISTING DATA
3. DESCRIBING AND IDENTIFYING DATA
4. CONNECTING DATA AND USERS
5. LINKING TO OTHER DATA
6. USING THE DATA

PROBLEMS OF FOOD COMPOSITION DATA

First is the fact that we need more and better data. There are large gaps in our knowledge of what is in the foods that are consumed around the world. And also, we need much better information about how these foods and nutrients interact.

Second, we need better information on exactly what data do exist, where they are, and how good they are.

Third, we need a scheme for unambiguiously identifying and describing foods and the components of foods, as well as the data themselves.

Fourth, there needs to be some way of easily connecting the users with the data. Free enterprise is a wonderful institution, however, the software vendors have not really addressed the overall problem, and certainly not in the global context.

Fifth, we need to be able to link, and again easily, food data to other data sets, from consumption to costs. It is a real bother to have to solve the linkage problems anew each time anyone wants to do something.

And, sixth, as this group knows especially well, there are problems in the use of the data that do exist. Not, of course, that any of us would misuse data, but there are others. I happen to feel, on good days, that many of these problems stem from the other five problems, from the lack of appropriate data, or the difficulty of getting the right data.

ACTIVITIES OF INFOODS

In the area of data quality we are sponsoring the English chemist, David Southgate, in his production of a manual on the gathering of food composition data. We are trying to work with Wayne Wolfe, of the USDA, on developing some global scheme for food reference standards. We have worked out an arrangement with the United Nations University whereby it is now giving INFOODS fellowships to permit developing countries to send people for training in food analysis. Building on these activities we are starting to think about developing short courses on food analysis that can travel to parts of the developing world. In terms of the data that get into tables, one big area that we are trying to get organized now is how non-analytic data are estimated -- we are trying to codify methods of calculation and imputation.

In the area of nomenclature and terminology we have a task force, run by Stewart Truswell of Australia, that is developing a scheme for global food and data identification. As I indicated earlier, this is not a simple problem -- the first thing that must be realized that there is a critical distinction between the data and the food from which they come. Once you have defined an apple, you need to define what apple or apples were analysed, and how, and what happened to the data then. These data about the data can be called "metadata", and we are slowly moving into trying to set out some ground rules for these -- how does one define and measure data quality, for example, and how do you communicate to the user information on how well certain data might fit in with their needs.

The area of access to the data is being coordinated, and to a large extent done, by John Klensin of MIT, who is here today and will talk more to this topic this afternoon. Our directory of food composition data tables, an international counterpart to Loretta Hoover's one for the United States, is now in its first edition, and currently being revised and expanded. We have produced a scheme for interchanging data between different parts of the world. In a cooperative project with USDA, we are currently embarked on a exploration of how existing computer networks can be used to communicate betweeen individuals and groups in the field.

One of our next big projects is the design of a regional data center, where we are preparing to give advice to various groupings of developing countries as to how they can organize their resources to improve their food composition data situation.

The other major activity that we are trying to accomplish is that of linking together people in the field. Within the US there is this conference; however, little exists internationally. We have a very small secretariat where we try to coordinate things, with varying degrees of success. We put out a newsletter to tell everyone what we are doing. Early next year we will start an international Journal of Food Composition and Analysis with Kent Stewart of VPI as editor, publishing scientific articles covering the whole field from food analysis to data manipulation to usage.

Additionally, we are trying, again with varying degrees of success, to get the world split up into geographic regions of managable size, and within each region to get an organization set up to deal with food composition data. We have forged links with two existing groups -- EUROFOODS, based in Holland, and NORFOODS, based in Upsalla, and we are working hard now on getting an ASIAFOODS and a LATINFOODS going. We have plans to talk to some of the countries in Eastern Europe this fall and there are indications that the countries of the South Pacific may be gathering together to work on the problems of their area. In the future are plans for Africa and the Middle and Near East, but things are more difficult there.

This briefly surveys what INFOODS is all about, and I hope gives you some feel for its whats and whys. While our stated focus is international, this is misleading in that the problems that we are trying to deal with are the problems that exist in the United States. Our basic tenent is that because food and disease are growing more and more international, the problems of food composition data must be considered in their global context.