Compilation of food composition datasets- an analysis of user needs through the Use Case Approach.

Running title: Food composition data compilation-user needs

Bernadette Egan PhD[1*] Charo Hodgkins BSc[1], Anja Fragodt MSc[1] and Monique Raats PhD[1]

[1]Food Consumer Behaviour and Health Research Centre, Faculty of Arts and Human Sciences, University of Surrey, Guildford, Surrey UK

[*]Correspondence: Dr Bernadette Egan, Food, Consumer Behaviour and Health Research Centre, Faculty of Arts and Human Sciences, University of Surrey, Guildford, Surrey, GU2 7XH, UK;

Tel: +44(0)1483682888; Fax: +44(0)1483682888

E-mail: m.egan@surrey.ac.uk

This work was completed on behalf of the European Food Information Resource Consortium and funded under the EU 6th Framework Food quality and Safety programme, project number FP6-513944.

Abstract

Objectives: To identify the common requirements of users involved in the compilation of food composition datasets with a view to informing the development of a common access system to food composition data, within the European Food Information Resource (EuroFIR) project.

Methods: A number of examples of food composition dataset compilation have been examined employing the Use Case approach, namely the compilation of a dataset for a national nutrition survey, for a cross-national nutrition study and for a nutritional software programme.

Results: The key user requirement identified from the compilation step analysed by the Use Case approach is the increased availability of and access to more detailed food composition data on a wider range of foods and nutrients.

Conclusions: Food composition data serve a variety of purposes and different user groups will often have both common needs and more individual or specific needs of their datasets. The development of Use Cases for specific processes effectively identifies the needs of users, highlighting any similarities and/or differences in those needs. The application of the Use Case approach to support the software development activities within EuroFIR ensures that user needs are effectively identified and captured in a systematic and documented way.

Keywords compilation, food composition datasets, nutrition, Use Case, user needs,

1. Introduction

Data on the composition of foods are essential in key areas such as clinical practice, research, public health and the food industry .Within each of these areas, food composition data serve a variety of purposes (Egan, Fragodt, Raats, Hodgkins & Lumbers, 2007; Williamson, 2005) and different user groups will often have both common needs and more individual or specific needs of their datasets. For example estimating the nutrient content of foods would appear to be a common use of food composition data across the different groups mentioned but the use of data for food labelling and nutrient claims is primarily the preserve of the food industry.

1.1 Compilation of food composition data

The compilation of food composition datasets is a very important step within the overall production and management of food composition data. The general steps involved in the

compilation process were described by Greenfield and Southgate (2003) who highlighted the necessity of considering user needs in the process if it is to be 'fit for purpose'. More recently, work by Westenbrink *et al*. (2009) expanded the description of the process with a view to identifying the critical control points for ensuring data quality in the overall compilation process. This work highlighted once again the need to engage users in several key steps. Furthermore the shift from the generation and storage of food composition datasets in hard copy to more sophisticated, complex electronic databases has made it necessary to develop a more in-depth understanding of the compilation process from a user perspective. Previous work identified the key processes that involve food composition data and the associated primary users (Egan *et al.*2009). A key user group identified was that involved in the compilation of food composition datasets. Users from that group, National Nutrition Survey managers, researchers involved in cross-national studies and nutritional analysis software developers, were approached with a view to mapping their individual compilation processes and identifying their associated needs through the Use Case approach.

1.2 National Nutrition Surveys

National nutrition surveys measure the food and nutrient intake of a representative sample of the population and allow governments to develop, implement and monitor national food and nutrition policies. National food composition datasets are created and maintained as tools for analysing the results of national nutrition surveys and the conversion of food consumption data into nutrient intake requires a comprehensive food composition database that is accessible and representative of national food supplies.

1.3 Cross-national Nutrition Studies

Researchers working in nutritional epidemiology identify relationships between dietary factors and human health with a view to identifying potential causes of diet-related diseases. Collecting reliable dietary data is challenging as accuracy is important if true associations are to be detected and thus cohort or prospective studies are considered more reliable, often involving large numbers of participants and spanning a number of countries. A cross-national study such as the European Prospective Investigation into Cancer and Nutrition (EPIC) study involved participants from ten different European countries (Deharveng *et al.* 1999) and in order to establish relationships between food and nutrient intakes and disease a standardised

European food composition database (FCDB) was required (Charrondiere *et al.* 2002). An important issue in the development of such a system is consideration of user needs and specifically their functional requirements of the system in question.

1.4 Nutritional Analysis Software

Nutritional analysis software is used by a variety of users for many different purposes, including assessing food intake, planning diets and food labelling (Stumbo, 2003). In a report on the future development of The Composition of Foods in the UK, 52% of users reported using third-party nutritional analysis software as a means of accessing food composition data (Raats, Angus & Lumbers, 2004). A more recent inventory of UK nutritional analysis software highlighted the diverse and rapidly changing nature of the market for such software products (Church, 2006). Inevitably the delivery of food composition data to a wider audience will involve developers of nutritional software programmes and hence these represent another key user group.

2. Methods

Use Cases were originally developed to describe the functional requirements of a software system from a user perspective (Jacobson, Christerson, Jonsson & Övergaard, 1992). Use Case write-ups may be developed for all activities within a given process but often the primary focus is on those steps where the user is directly interacting with software or another system supplying data. Each compilation process may be described as a sequence of activities and a Use Case write up developed for each of those activities; individual activities are in turn described in terms of all the possible steps required for achieving a single goal. A successful flow of events details the sequence of steps which result in all the user's functional requirements being met. Inevitably some steps result in success; some result in failure or in the user needing to employ a 'coping strategy' by branching away from the successful flow in order to satisfy their needs. It is these exceptions or alternative flows that clearly identify where the user's needs are not being met by the system employed.

To ensure that the steps in the compilation process were accurately described expert users were identified from each of the three user groups. Draft outline processes were prepared using the template developed previously (Egan *et al*. 2009) and these were provided to the targeted users and subsequently modified to accurately reflect their individual processes. The

resultant activity steps (Table 1) were compared across the three user groups and whilst many were identified as being user specific, activity step 2 was clearly common across the groups. This step represented the fundamental part of the compilation process i.e. the consideration and selection of food composition data and was then progressed into a Use Case write-up with a view to highlighting where user needs were not being met.

3. Results

It is clear from comparing the compilation processes that the factors involved in Activity Step 1 i.e. the selection of foods to be included in the food composition database, vary between the user groups. For national nutrition surveys, new foods are added to the existing dataset on the basis that they have been included in the food diary/dietary recalls of the survey participants. For a cross-national study, foods and nutrients are selected based on their relevance to the study and their availability in the national food composition datasets. In contrast, a software developer will continually be attempting to extend their food composition dataset on the basis of relevance to their target customers.

With respect to Activity Step 2, namely the consideration and selection of the relevant food composition data for the food, there is a great deal of similarity across the three groups. Effectively once the compiler has decided on the food they wish to include in their dataset, they tend to follow a similar process to search for and select appropriate nutritional information to support that food. It is therefore possible in this situation to develop a generic Use Case for Step 2 which describes the successful flow of events and the exceptions for all three datasets (Table 2).The exceptions or alternative flows listed in Table 2 cover the range of coping strategies employed by these user groups to compile their particular datasets.

4. Discussion

A principal aim of the EuroFIR project is to develop sustainable systems that will increase the availability of European national food composition data through a common access system (Moller, Unwin, Becker & Ireland, 2007). Use Cases have been widely adopted as a means of identifying user requirements, in part due to their ease of comprehension and their description of system requirements in terms of user interaction (Kanyaru & Phalp, 2009; Siau & Lee, 2004). By developing Use Cases for key processes involving food composition data the functional requirements of common access systems can be identified and satisfied. The key

requirement identified from the compilation step analysed here is the increased availability of and access to more detailed food composition data on a wider range of foods and nutrients. This is a need identified previously through the Use Case approach for another user type, namely a dietician working in a non-clinical setting. (Egan *et al*, 2009.) and also reflects the findings of McCabe-Sellers &Chenard (2008) who reported the need for data on commercial ingredients and nutritional supplements by US dieticians. Clearly application of the Use Case approach within EuroFIR ensures that user needs are effectively identified and captured in a systematic and documented way. Furthermore this structured approach informs the development of usability testing protocols for the resulting software applications to ensure that user needs are met, which is crucial if EuroFIR is to be perceived as valuable by the widest possible audience.

Acknowledgements

This work was completed on behalf of the European Food Information Resource Consortium and funded under the EU 6th Framework Food quality and Safety programme, project number FP6-513944.

Conflict of interest

None

References:

Charrondière UR, Vignat J, Møller A, Ireland JD, Becker W, Church S *et al.* (2002). The European Nutrient Database (ENDB) for nutritional epidemiology. *J Food Comp Anal* **15**, 435-451.

Church S (2006) The history of food composition databases. *Nutrition Bulletin* **31**: 15–20.

Deharveng G, Charrondière UR, Slimani N, Southgate DAT and Riboli E (1999). Comparison of nutrients in the food composition tables available in the nine European countries participating in EPIC*. Eur J Clin Nutr*. **53**, 60-79.

Egan, M.B., Fragodt, A., Raats, M.M., Hodgkins, C. & Lumbers, M. (2007). The importance of harmonizing food composition data across Europe. *Eur J Clin Nutr,* **61***,* 813-821.

Egan MB, Hodgkins C, Fragodt A, Krines C, Raats MM (2009). User-centred food composition data - analysis of user needs through the Use Case Approach. *Food Chemistry,* **113**, 804-807.

Greenfield H and Southgate DAT (2003. Food Composition Data: Production, Management and Use 2nd Edition, FAO Rome.

Jacobson, I., Christerson, M., Jonsson, P. & Övergaard, G. (1992). Object-orientated software engineering: a Use Case driven approach. *Inform Software Tech* **34***,* 825-826

Kanyaru, JM & Phalp, K. (2009). Validating software requirements with enactable Use Case descriptions. *Requirements Eng,* **14***,* 1-14.

McCabe-Sellers, B.J. and Chenard, C.A. (2008). Meeting the needs of US dietitians for food composition data. *J Food Comp Anal,* **21***,* S27-S34.

Møller, A., Unwin, I.D., Becker, W. & Ireland, J. (2007). EuroFIR's food databank systems for nutrients and bioactives. *Trends Food Sci Tech,* **18***,* 428-433.

Raats MM, Angus F, Lumbers ML (2004). An assessment of issues pertinent to the future development of The Composition of Foods. Report prepared for the Food Standards Agency.

Siau, K & Lee, L (2004). Are Use Case and class diagrams complementary in requirements analysis? An experimental study on Use Case and class diagrams in UML. *Requirements Eng* **9***,* 229-237.

Stumbo, P.J. (2003) Novel software applications using nutrient databases. *J Food Comp Anal,* **16***,* 293-297

Westenbrink S, Oseredczuk M, Castanheira I, Roe M (2009). Food Composition Databases: the EuroFIR approach to develop tools to assure the quality of the data compilation process. *Food Chem* **113**, 759–767

Williamson, C. (2005). The different uses of food composition databases, Synthesis report No 2. Available at URL: www.eurofir.net/UsesspofspFCDBs.pdf

Table 1: Activity steps in the Compilation of a Food Composition database for three different user groups.

Activity Step	Description of principal tasks
1	Decision on which foods/.nutrients are relevant and need to be included in the database
2	Consideration and selection of relevant food composition data
3	Coding of food composition data
4-6	Steps may include • Collection of dietary information from survey population • Assessment and standardisation of national study database (NDS) • Imputation of missing values • Choice and application of algorithms to produce aggregated data.

Table 2 Generic Use Case - Activity Step 2: Consideration and selection of food composition data

Primary Users	:Nutritionist/database manager/software developer
Preconditions (what must be true before Use Case runs)	User has access to a software programme that is being used to create the database. User has electronic access to food composition data sources
Goal	To populate the database with up-to-date, food composition data on all required food items
Successful flow of events (*the case in which all the user's functional requirements are satisfied*):	1. User searches for food item alphabetically or based on food groups in national food composition database or in an already existing (study) database 2. User selects required food item. 3. User compares portion sizes determined from the survey with reference portion sizes. (national nutrition survey) 4 Use searches for nutrient composition in national food composition database 5. User manually enters nutrient composition of foods in study database (if necessary) 6. User proceeds to next activity
Exceptions/Alternative flows	1.a Food item is not found in either the national food composition database or the existing study database from previous projects *Alternative flows:* ➢ User chooses food item from another food composition database

- User commissions a laboratory analysis if resources are available
- User chooses food item from scientific literature
- User replaces food item of interest with an equivalent or similar food that is available
- User adds food item as a new one with nutrient information from a food product label
- User adds food item as a new one with estimated nutrient information
- User chooses an 'aggregated' food value

3.a Reference portion sizes are not available in national food composition database:

Alternative flows:

- identify portion size from alternative source and enter

4.a Food item available, missing nutrient values

Alternative flows:

- User chooses nutrient composition data from another food composition database
- User commissions a laboratory analysis if resources are available
- User chooses nutrient composition data from scientific literature

	➢ User chooses nutrient composition data from a food product label ➢ User replaces food item with an equivalent or similar food from database, which includes the required nutrient values ➢ User approximates nutrient composition by recipe calculation