A Proposal for an European Large Knowledge Repository in Advanced Food Composition Tables for Measuring the Diet Intake

Coltell O1,2, Madueño F1, Falomir Z1, Corella D2,3

1Languages and Computer Systems Department. Universitat Jaume I, Castellón, Spain.  
2CIBER Physiopathology of obesity and nutrition (CIBEROBN). Institute of Health Carlos III, Madrid, Spain.  
3Department of Preventive Medicine and Public Health, University of Valencia, Valencia, Spain.  
\{oscar.coltell, francisco.madueno, zfalomir\}@uji.es, dolores.corella@uv.es

Abstract. A proposal for designing and developing a European Repository of Knowledge on Advanced Food Composition Tables (FCTs) based on the existing national FCTs, is proposed in this paper. The requirements of the system, the interoperability strategies, and the cooperation of each national FCT for maintaining and updating the repository are discussed.

Keywords. Knowledge repositories, Food Composition Tables (FCTs), Minimal set of FCTs.

1 Introduction

The study of the interaction between diet and the genome is crucial to prevent and treat cardiovascular diseases, some cancers, type 2 diabetes, etc. The assessment of a person's diet is a tedious task, and in practice, a portion of the intake information is evaluated and then the habitual participants' intake is extrapolated. In order to obtain enough statistical power to avoid measurement errors and changes in diet, it is necessary to obtain repeated measures of dietary information from a large number of participants over time. For extracting information regarding participants' diet, nutritionist use Food Frequency Questionnaires (FFQ), 24 hour dietary recalls (24HDRs), dietary records or dietary histories (for more details see Falomir et al.[1]). These surveys collect consumed foods or dishes, which can be transformed into energy and nutrient intake using Food Composition Tables (FCTs).

Last June 14 in The Hague, the Joint Programming Initiative1 (JPI) in “A Healthy Diet for a Healthy Life” was held and the 2010-2020 roadmap for harmonising and structuring research efforts in the area of food, nutrition and health was presented. The goal of the JPI conference was to define the Strategic Research Agenda for the period 2011-2020 and beyond2, which main aims are to provide a holistic approach to: (i) identify the key factors that affect diet-related diseases, (ii) discover new relevant parameters and mechanisms and (iii) define strategies that contribute to the development of actions, policies and innovative products suitable to reduce the burden of diet-related diseases. The JPI Agenda developed the corresponding subroadmap for each one of the three key interacting research areas that were identified and described in the previous

1JPI: https://www.healthydietforhealthylife.eu/hdhlconference/ (Last acces June 30,2012)  
2The JPI Strategic Research Agenda for the period 2011-2020 and beyond.  
Vision Document\textsuperscript{1} of the JPI. The Research Areas (RA) are the following: RA1-Determinants of diet and physical activity; RA2-Diet and food production; and RA3-Diet-related chronic diseases.

Each research area roadmap in the Agenda presents two prime initiatives: for 2012-2014 and 2015-2019. The prime initiative for RA1 (2012-2014) is “Establish a European transdisciplinary research network on determinants of dietary and physical activity behaviours and the relation with health and best practice implementation strategies for sustainable changes”. This initiative is a research challenge which must starts with the collection, integration and assessment of monitoring systems, databases, determinants and outcome measures. And one of the research needs to face the challenge is to establish and maintain an integrated trans-disciplinary database, with potential for secondary analysis by interested researchers with specific research hypothesis, assuming the initial data are collected according to best practice in biological, behavioural, socio-economic and environmental science traditions.

2 The Problem of National Food Composition Tables

Technically speaking, the research challenge of creating a European FCT (EFCT) involves a technological challenge in the field of large databases and large repositories. The Scientific Advisory Board of the JPI, called DEDIPAC, claimed that the EFCT should not be a “data” or “information” database, but a knowledge network repository with contributions of at least 27 European countries. The specific challenge to face is to organize the existing knowledge, their supporting infrastructures and their associated management requirements of the databases containing national Food Composition Tables (FCT) and their integration in a large knowledge repository. Traditionally, FCTs were tables where a portion of each single food was decomposed in energy, macronutrients and other components that are not nutrients. The standard size of the portion is 100 g, but some FTCs take the edible part of the food (i.e. discarding the peel in oranges; in this case, 100 g of edible orange), and other FTCs take the whole food (i.e., the whole 100 g of orange, including the peel). Moreover, macronutrients are grouped in families, as lipids, proteins, carbohydrates; and the no nutrients are minerals, vitamins and aminoacids. Usually, each FTC register contains around 50 components. However, the number of components may vary in each FTC. Regarding national and private (academic or enterprise) FCT creation, although they can be standardized and biochemically proved, they are usually different from country to country (or depending on the academic organization or enterprise aims and resources).

With the evolution of the information and communication technologies, FCTs were converted in databases and, later, Web services were added to allow online access to them. But the drawbacks of the traditional FTC were inherited by the FCT databases and emerged some specific problems as, for example, the lack of service due to site saturation or network breakdowns, the restricted access, the lack of programmed access (a set of procedures to manage queries coming from applications), the native language, and so on. That is the situation of the European FCT provided by the FAO\textsuperscript{4} or EuroFIR\textsuperscript{5}.


\textsuperscript{3} EuroFIR. http://www.eurofir.net/. (Last access in June 30, 2012)
Therefore, the aim of this paper is to discuss a viable proposal for designing and developing a European Repository of Knowledge on Advanced FCTs based on the existing national FCTs, their system interoperability strategies, and the cooperation of each national FCT for maintaining and updating the repository.

3 The Knowledge on Advanced Food Composition Tables

Regarding the existing national FTC databases, the first step should be to define a Minimal Set of FCT data (MS-FCT). The MS-FCT is the common data that holds every FCT database in the same or approached format (no need of transformation or conversion). On the other hand, the Standard Set of FCT data (SS-FCT) must be defined. The SS-FCT is the standardized data that every FCT database should contain according strategic objectives of the knowledge repository (homogeneity, integration, interoperability).

The second step is to define the knowledge levels in the repository. Initially, we have defined the following levels:

1. **Level 1: Food Composition.** Basic knowledge about the composition of each food but with the following variations: national FCT source, determination methods for each component, local and regional variations of the food, and original language.

2. **Level 2: Dish Composition.** Knowledge about the composition of dishes in single food, the standard portions (in Metrical and English measures) and their corresponding images, the corresponding recipes (the same food mixture is different according the cooking process), and the local and regional variations in recipes and portions.

3. **Level 3: Diet patterns.** Knowledge about discovered diet patterns nutritional studies using data mining strategies. From diet patterns, it would be possible to generate diet models to apply in the kind of studies described in the JPI research areas prime initiatives.

4. **Level 4: Diet-disease effects.** Knowledge about associations and interactions between diet and disease (via genetic and phenotypic factors), recommendations for specific populations (i.e., celiac), high risk food for specific diseases, lowering risk food for specific diseases, etc.

All together should run in cooperation with every national security system for the FCT database, providing full access to authorized sources, level of service and frequent updates to guarantee the quality and accuracy of the provided knowledge in the repository.

Previous works by our group (see Fabregat el al. [2]), developed some medium scale projects in the area of medical informatics for automating nutritional questionnaires and calculating the nutritional composition of meals using several FCTs which used an ontology for translating the components in the different FCTs to a common name. Based on the success of this approach, we we propose a solution to the integration of all European FCTs based on ontologies and web services, and asynchronous web technologies for assuring the minimal response time in knowledge queries, and for providing modular services, and the maximal underlying data organization.

**Acknowledgments.** This work has been partially supported by: (1) Spanish Ministry of Science and Innovation (MICINN) under the project named GEWIMICS and referenced as SAF2009-12304/GEN; (2) the Cooperative Research Thematic Network (RETIC) on Computational Medicine (COMBIOMED) referenced as RD07/0067/0006 and funded by Carlos III Spanish Health Institute (ISCIII); (3) the Biomedical Research Centre Network
(CIBER) named Fisiopatología de la Obesidad y Nutrición (CIBERobn) as an initiative of ISCIII; (4) Generalitat Valenciana under grant numbers BEST/2010/211, BEST/2011/261, and ACOMP/2011/145; and (5) Spanish National Centre of Cardiovascular Research (CNIC) under project number CNIC06.

References

1. Falomir Z., Arregui M., Madueño F., Coltell C., Corella D., Automation of Food Questionnaires in Medical Studies: a state-of-the-art review and future prospects, Computers in Biology and Medicine, Accepted June 2012, in press.