Towards a WWW-accessible Knowledge Base on Breast Cancer Prognosis and Therapy

Rose DIENG,
INRIA, Acacia Project, 2004 route des Lucioles, BP 93, 06902 Sophia-Antipolis Cedex FRANCE

Carmelina RUGGIERO,
DIST, Universita di Genova, Via Opera Pia 13, 16145 Genova, ITALIA

1 Introduction

This paper presents a collaborative project between the ACACIA Project at INRIA-Sophia-Antipolis and DIST (Department of Informatics Systems and Telematics of the University of Genova), for the development of a WWW-accessible Knowledge Base on Breast Cancer Prognosis and Therapy. The ACACIA Project performs research on knowledge acquisition and expertise capitalization while DIST has a long experience in the field of Biomedical Engineering (cf. participation to ESPRIT, AIM, BRITE or TIDE projects, applications on breast cancer).

The collaborative project holds in the framework of the French-Italian Galileo program, and focuses on the determination of the proper prognosis for breast cancer - which is a disease extremely frequent in France and in Italy. The main prognostic factor is the lymph node status. But, all possible variabilities may be found depending on other prognostic factors which are continuously updated by cancer researchers, in particular biological factors (cf. figure 1).

The collaboration of both teams began in 1995:

• In the framework of his postdoctoral training in the ACACIA project, Dr Roberto Sacile (from Genoa University) performed interviews of an expert of IST (Scientific Institute for Cancer Study and Therapy) of Genoa and analysed some Web servers and some articles on the field. Thus he built a CommonKADS expertise model on breast cancer prognosis [Sacile95, SRD95, SRD96].

• He also used the knowledge acquisition tool COKACE developed by Dr Olivier Corby: this tool allows to build and validate CommonKADS expertise models) [CD96]. It enabled a validation of the proposed expertise model and its transformation in HTML format, accessible through Web. Moreover, access and interrogation of Wais bibliographic bases, and access to documents sources are provided.
2 Description of the collaborative project

The intended user of our further research is the student in oncology, the researcher specializing in oncology or the practitioner giving treatments on breast cancer or the developer of a knowledge-based system in this field. Such users will need access to updated knowledge on breast cancer diagnosis, prognosis and therapy. Therefore, we aim at building a knowledge base which should:

- be able to connect different guidelines used in different cancer research centres (i.e. Genoa-IST, Nice-Lacassagne...) and to join different points of view,
- be able to be updated with knowledge coming from newly published research (specifically those on biological aspects),
- be conceptual in the sense that it may be understood and even hand-used by experts.

Our teams will collaborate for:

- Capitalization of expertise on the breast cancer prognosis and therapy, from multiple expertise sources (practitioners, researchers, articles...).
- Building a multi-experts knowledge base accessible on the Web, and linked to documentary expertise sources.

The project aims at building a knowledge base on breast cancer prognosis and therapy. This knowledge base will be accessible via Internet and World Wide Web, in order to enable its consultation by the different kinds of end-users stressed above. The work is planned as follows:

- The knowledge will be elicited from multiple expertise sources: experts from IST (Genoa) or from Centre Lacassagne (Nice), articles and documents on the domain, accessible Web servers on breast cancer. It will be modelled by exploiting the knowledge acquisition method CommonKADS developed in the Esprit projects P1098 et P5248, and the language CML (Conceptual Modelling Language) offered by CommonKADS for knowledge modelling.
- We will study techniques for management of several CommonKADS models corresponding to several viewpoints (that may be different or complementary on the same domain). Such CommonKADS models will be stored in electronic documents, with hypertext links between the different models.
- Moreover, hypertext links will also be maintained between the documents considered as expertise sources (articles, etc.) and the base of CommonKADS models, which will allow indirectly a navigation between...
such expertise sources. Two types of electronic documents will then be exploited: the documents containing the CommonKADS models and the documents containing the articles used as expertise sources.

- The CommonKADS models will also be described with the format HTML (HyperText Markup Language) and set up on a Web server in order to allow their friendly consultation on the Web via Internet. We will also exploit a Wais server to manage bibliographic references to filter the access to articles.

So, the main interest of the knowledge base will be its structuration according to multiple viewpoints and its connections with articles written in natural language (see figure 2).

![Diagram](image)

**Figure 2 : Multiple Expertise Models**

### 3 Conclusions

This research should be interesting for:

- Knowledge Acquisition: from multiple expertise sources (cf. hypertext links between several CommonKADS expertise models corresponding to different viewpoints), from documents (cf. hypertext links with the articles) and by exploiting the Web.
- Telemedicine: by offering a kind of digital library (made of a knowledge base and documentary sources), available on the Web, and by allowing a comparison of multiple methods or viewpoints on breast cancer prognosis and therapy, our work should help to improve training and research on this field.

### References


[SRD96] Roberto SACILE, Carmelina RUGGIERO, Rose DIENG. Using CommonKADS to build an expertise model for breast cancer prognosis and therapy. Accepted at *Medical Informatics*, 1996.