

Conceptual document browsing and retrieval in *Kabiria*[°]

A. Celentano^{§†}, M.G. Fugini^{§‡}, S. Pozzi[†]

[§] Politecnico di Milano, Piazza Leonardo da Vinci 32, I-20133 Milano, Italy

[†] CEFRIEL, Via Emanuelli 15, I-20126 Milano, Italy

[‡] Università di Brescia, Via Valotti 9, I-25060 Brescia, Italy

E-mail: fugini@ipmel1.polimi.it

Kabiria is a document classification and retrieval system that integrates classical information retrieval facilities with features of knowledge representation for document retrieval. In fact, *Kabiria* is based on a *semantic model* of documents where knowledge about the procedures and the rules of the office environment drives the filing and retrieval activities. The model includes knowledge regarding the document structure and text, and *operational knowledge* of the environment where documents are used. Details about the model can be found in [1].

The knowledge is organized in *Kabiria* as a semantic network representing both the document structures and roles, and as a set of rules representing the laws, regulations, and processing habits that influence the document life cycle.

The implementation of *Kabiria* is based on a client-server architecture supporting three basic activities on documents: *classification*, according to the semantic types defined in the model; *filing*, both in a traditional document base and as classes of the knowledge net according to the document roles; *retrieval*, according both to the structure, contents, and role of documents and to the *domain dependencies* holding between the office environment and the domain context where the office is embedded.

Accordingly, *Kabiria* is composed of a *classification* module, to load into the system document classes and instances, and an *insertion* module, to feed the document base according to the knowledge about document roles. A *retrieval* module, including a browser and a query processor, supports the navigation in the knowledge base and the retrieval of documents. The browser and query facilities are described in [2].

A prototype implementation of *Kabiria* has been developed at CEFRIEL on 386-based Personal Computers running Microsoft Windows 3.0, and GoldWorks III, a Lisp based environment of Gold Hill Computers Inc., for the development of knowledge-based software systems supporting the object oriented and rule based paradigms. The architecture of *Kabiria* is illustrated in [3].

REFERENCES

- [1] A. Celentano, M.G. Fugini, S. Pozzi, "Knowledge-based retrieval of office documents", *SIGIR '90, Proc. 13th Intl. Conf. on Research and Development in Information Retrieval*, Bruxelles, September 5-7, 1990.
- [2] A. Celentano, M.G. Fugini, S. Pozzi, "Querying office systems about document roles", *SIGIR '91, Proc. 14th Intl. Conf. on Research and Development in Information Retrieval*, Chicago, October 13-16, 1991.
- [3] A. Celentano, M.G. Fugini, S. Pozzi, "Classification and retrieval of documents using office organization knowledge", *COCS '91, ACM/IEEE Conf. on Organizational Computing Systems*, Atlanta, November 5-8, 1991.

[°] The Project *Kabiria* has been developed at CEFRIEL with the support of BULL HN Italia