The ERCIM Technical Reference Digital Library

Stefania Biagioni[1], Josè Luis Borbinha[2], Reginald Ferber[3], Preben Hansen[4],
Sarantos Kapidakis[5], Laszlo Kovacs[6], Frank Roos[7], Anne-Marie Vercoustre[8]

[1]IEI-CNR, Via S. Maria 46, 56126 Pisa, Italy
[2]INESC, Rua Alves Redol, 9, Lisbon, Portugal
[3]GMD-IPSI, Dolivostr. 15, D-64293 Darmstadt, Germany
[4]SICS, S-164 28 Kista, Sweden
[5]ICS-FORTH, Vassilika Vouton, GR-71110 Heraklion, Crete, Greece
[6]SZTAKI, H-1518 Budapest, Hungary
[7]CWI, Kruislaan 413, NL-1098 SJ Amsterdam, The Netherlands
[8]INRIA, Domaine de Voluceau, Rocquencourt, France

1   Introduction

Within the context of the DELOS[1] Working Group, eight institutions of the European
Research Consortium for Informatics and Mathematics (ERCIM) are currently
collaborating on the installation of an ERCIM Technical Reference Digital Library
(ETRDL). The aim is to implement and test a prototype infrastructure for networked
access to a distributed multi-format collection of technical documents produced by
ERCIM members. The collection is managed by a set of interoperating servers, based
on the Dienst system developed by a US consortium led by Cornell University and
adopted by NCSTRL (Networked Computer Science Technical Reference Library).
Pilot server sites have already been set up at half of the 14 ERCIM national labs.
Servers are expected to be installed at the other centres soon. The aim is to assist
ERCIM scientists to make their research results immediately available world-wide
and provide them with appropriate on-line facilities to access the technical
documentation of others working in the same field. Public access to this reference
service is provided through Internet.

2   Common User Interface

In addition to the basic service provided by the DIENST system, some additional
functionalities are being implemented in the ETRDL common user interface in order
to meet the particular needs of the European IT scientific community. An author
submission form has been included to facilitate the insertion of new documents by the
users themselves. The service can be accessed through the DELOS Web site.

Extending the Metadata Set. Dienst provides services to store both documents and
their metadata. Retrieval is based on the registered metadata. Dienst accepts a small
set of metadata but can be configured to handle additional elements. We are now
implementing an extension to the current configuration of the Dienst system in order
to increase retrieval options. The additional data fields are compatible with the Dublin

---

[1]For more information on the activities of the DELOS Working Group on Digital Library
  related research issues (ESPRIT - LTR No. 21057), see http: www.iei.pi.cnr.it/DELOS/.

Core (DC) metadescription standard. The Dienst code has thus been modified to enable the new fields to be indexed and searched.

Common Classification Scheme. One set of extensions are fields for the ACM Computing Classification, for the AMS Mathematics Subject Classification, and for free keywords. For submission and retrieval, users can browse the classifications, mark selected keywords and insert them in the appropriate fields. Authors should enter terms from at least one classification; they can also use all three fields. Searches are performed on all three fields by default, but may be restricted to single fields. Problems caused by the adoption of multiple classification schemes are being studied.

3   Multilingual Interface

Multilinguality is an issue of strategic importance for the European scientific community. The first activities of the ETRDL in this area are aimed at (i) implementing an interface capable of handling multiple languages and (ii) providing very basic functionalities for cross-language querying.

Multilingual Access and Browsing. Each national site is responsible for localisation, i.e. implementation of local site user interfaces (also) in the national language as well as English: one of the tasks of the group will be to investigate problems involved in rendering the Dublin Core element set multilingual. Documents are tagged for language and character code switching mechanisms are provided for the local display and printing of non-Latin-1 languages (Hungarian and Greek in our collection). However, it is agreed that UNICODE must be adopted eventually in order to fully internationalise the system.

Cross-language Querying. In the short term a simple form of cross-language querying will use controlled keyword (ACM/AMS) terms. All documents in the ETRDL, in whatever language, classified using this scheme, can thus be searched. Authors are also requested to include an abstract in English, which makes English free term searching over documents in any language also possible. INESC has developed an LDAP service with a multilingual repository for the ACM and AMS classification systems, which is integrated in the ETRDL system. This multilingual service makes cross-language querying in local languages possible. In the longer term, other methods for cross-language querying which enable the user of the DL service to retrieve texts composed or indexed in one language via a query formulated in another will be investigated.

4   Gateway to Z39.50

The Z39.50 protocol enables access to library data; adopted by the EU Librarians Programme, it will be the official protocol for search and retrieval in European libraries. A Z39.50 access to Dienst has been developed at FORTH and will be integrated in the ETRDL service.