

Towards a Semantic Portal for Minority Rights in Europe*

Johann Gamper and Paolo Dongilli
Free University of Bozen/Bolzano
{johann.gamper,paolo.dongilli}@unibz.it

1 Motivation

Due to a new sensibility towards minority questions, the documentation and easy availability of data on ethnic, national, and linguistic minorities is increasingly significant. While documents of international organisations are mostly available on the Web, the gathering of extensive information on the legal standards in different countries, such as national legislation or case-law, is still a difficult task. The existing databases are either limited to a few countries (e.g. MINELRES and CEDIME-SE) or they only cover a few topics (e.g. Mercator). Furthermore, those databases are mostly unstructured, basically providing a simple list of documents without being organized by keywords or topics. Thus, a quick search or a comparison between different countries is often difficult.

2 The MIRIS System

Taking into account this gap, the ongoing research project MIRIS (Minority Rights Information System) aims to develop a comprehensive information system about ethnic minorities and autonomies in Europe. For the very first time it will offer an intelligent Web portal to relevant information on minorities of all countries which are member states of the Council of Europe, providing a coherent presentation and powerful query capabilities including content-based search. MIRIS adopts a knowledge-based approach similar to [Staab *et al.*, 2000], heavily relying on the use of domain ontologies and metadata. Current Web standards including XML and RDF are applied. The first version with limited functionality can be accessed at <http://miris.eurac.edu>.

Text Documents. Most of the relevant information is stored in human-readable documents, such as the state reports to which the member states are committed, minority reports, international law contracts, national laws, etc. The most important documents are physically stored in our system. These documents are encoded for structural information using XML, which allows for a better visualization and navigation. For the other documents we store the URL.

Metadata and Ontology. In order to support intelligent access capabilities, we enrich the original documents with metadata. The metadata include administrative information, such as the document type and the date of entry into force, but the

most important information to support content-based search is the annotation with topics, which explicitly describes the content of the documents on a semantic basis. The topics are organized in a domain ontology which represents common knowledge about ethnic minorities and provides the vocabulary for the annotation. Depending on the importance of the document, content annotation is done at the document-level or at the level of other structural units such as law articles.

Knowledge Base and Inference Engine. The knowledge base and inference engine serve two essential goals: content-based access to the text documents or parts of them as well as the possibility to ask complex queries retrieving factual knowledge in combination with documents. In addition to the metadata, the knowledge base contains factual knowledge about ethnic minorities, e.g. the minorities in a specific country, population numbers, and languages used. This allows to ask specific queries such as “What is the largest minority in Italy?”. The inference engine can deduce new knowledge which is not explicitly stored. For example, if a particular article from the Framework Convention is not referred to in a state report, one can conclude that the corresponding topic is not yet regulated in that country.

Query Capabilities. Various possibilities to access the information system will be provided, from a focused search for specific information in the knowledge and database up to the generation of reports to a certain topic. In the browsing mode the user can navigate via hypertext links through the entire information space. Each hyperlink contains a query which is dynamically evaluated. The ontology is used both to expand queries and to provide different views. The user can also formulate complex queries which might include one or more concepts from the topics ontology and yield a list of relevant documents possibly combined with factual knowledge and links to related documents. For example, the query “Which topics are regulated in a specific country?” might result in a number of topics together with the documents which cover these topics. Moreover, MIRIS offers also fulltext search.

References

[Staab *et al.*, 2000] S. Staab, J. Angele, S. Decker, M. Erdmann, A. Hotho, A. Maedche, H.-P. Schnurr, R. Studer, and Y. Sure. AI for the Web — ontology-based community Web portals. In *Proceedings of the AAAI'00*, pages 1034–1039, Austin, Texas USA, July 2000.

*In cooperation with the European Academy of Bozen/Bolzano