LEGAL KNOWLEDGE REPRESENTATION USING A FACETED SCHEME

Michelle Cumyn (Université Laval)
Günter Reiner (Universität der Bundeswehr Hamburg)
Sabine Mas (Université de Montréal)
David Lesieur (Université de Montréal)

PROJECT DESCRIPTION

CONTEXT. A database supports legal research by matching a user's request for information with documents of the database that contain it. Indexes are among the oldest tools to achieve that aim.

It is well-established that the full-text search is less effective than the use of a subject index. A query using indexed terms provides greater precision (results are more relevant) and better recall (fewer relevant results are missed). Searching with indexed terms is the traditional method still being taught in law schools. However, users seem to prefer the full-text search to other methods of document retrieval.

The performance of full-text searching is enhanced by the relevance sorting of document results. Relevance sorting algorithms are criticized because they are not transparent. The factors used to sort results and their weighting cannot be controlled by users.

In the legal domain, automatic indexing still does not provide the same quality as manual indexing, which is expensive. Legal data base providers find themselves at the crossroads: will they give up subject indexing altogether? Is the development of sophisticated relevance sorting algorithms the key to future developments in the field?

HYPOTHESES. Our hypothesis is that the subject index remains an essential tool for document retrieval in the legal domain. However, it has yet to be adapted to the possibilities offered by information technology. More specifically, a faceted indexing scheme provides a flexible structure that is more intuitive than traditional legal classifications. Moreover, the traditional approach to indexing, which favors the use of legal concepts, should be complemented by one that better represents the facts.

METHOD. We have designed a new model for indexing legal documents using facets. In order to test our model, we have created a prototype database with a sample of 2 500 judicial decisions in the fields of administrative law, labor law and the law of obligations. Our prototype will be tested over the coming year, in order to determine whether search results are improved by the use of facets.

Our interdisciplinary project brings together research in information science and legal methodology.

FUTURE DEVELOPMENTS. We expect that our model will support the full or partial automation of legal indexing. It may also provide a structure for the development of other AI applications in the legal domain.

RANGANATHAN'S THEORY OF FACETED CLASSIFICATION

S.R. Ranganathan, an Indian mathematician and librarian (1892-1972), conceived the theory and method of facet analysis. Every subject matter or unit of knowledge can be broken down according to five fundamental categories or "facets."

Ranganathan's Facets

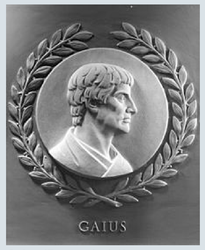
- **Personality** (entity or person, thing)
- **Matter** (substance, property, attribute, material)
- **Energy** (operation, action, activity)
- **Space**
- **Time**

Ranganathan's facets, which one can think of as dimensions of a topic, revolutionized the field of knowledge organization. The use of "facets" in the context of a Web interface is a distant relative of Ranganathan's theory.

GAIUS' TRIPARTITE SCHEME

Gaius was a famous Roman jurist who lived in the 2nd century. His tripartite division is familiar to all Western legal systems:

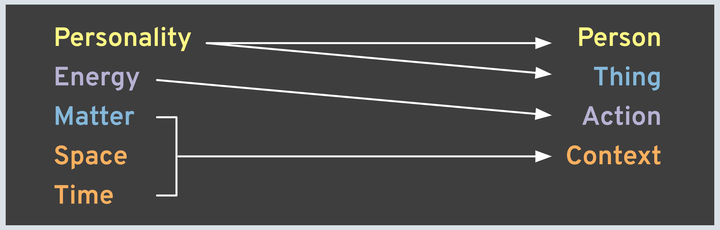
> "All our law is about persons, things and actions"
>
> — *Gaius, Institutiones, 1.8*



Gaius (Roman jurist)

OUR SYNTHESIS OF RANGANATHAN AND GAIUS' SCHEMES

Ranganathan's facet Personality comprises Persons and Things, which are always distinguished in law. Ranganathan's facets Matter, Space and Time, although relevant, are less important in the context of law. We treat them as elements of Context.



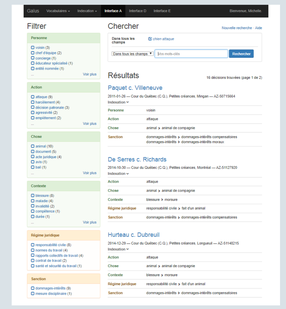
OUR FACETED SCHEME

Traditional classification and indexing methods favor the use of legal concepts. They rarely describe the facts. We have found that the categories drawn from Ranganathan and Gaius are most useful to describe the facts that are legally relevant. Our scheme represents the relationship between relevant facts and legal response. Legal categories and sanctions contain many of the terms currently found in subject indexing.

Facts	⟷	Legal Response
Person		Legal Category
Action		Sanction
Thing		
Context		

OUR PROTOTYPE DATABASE

As we could not rely on existing thesauri, we developed our own controlled vocabulary (in French) under the supervision of an expert librarian. We prepared a detailed indexing policy and manually indexed the 2 500 cases included in our prototype data base. Please feel free to try the prototype yourself!



ACKNOWLEDGMENTS

This research is funded by the Social Sciences and Humanities Research Council of Canada (2015–2020). We are grateful to Michèle Hudon, retired professor from the Université de Montréal, for her invaluable contribution to this project. We are also grateful for the research assistance of Christophe Achdjian, Dominique Boulanger, Cécile Gaiffe, Marjorie Gauchier, Joanie Gosselin-Lepage, Ariane Joncas, Béatrice Lecomte, Amélie Roy, Erin Sandberg, Forrest Wakarchuk and Iryna Zazulya. Thanks to Alexandre Fortier, our expert librarian and Charles Tremblay-Potvin, our project coordinator. Finally, we would like to thank the Société québécoise d'information juridique (SOQUIJ) for their cooperation and interest.