



Mapping classifications and linking related classes through SciGator, a DDC-based browsing library interface

Marcin Trzmielewski, Claudio Gnoli, Marco Lardera, Gaia Heidi Pallesstrini, Matea Sipic

► To cite this version:

Marcin Trzmielewski, Claudio Gnoli, Marco Lardera, Gaia Heidi Pallesstrini, Matea Sipic. Mapping classifications and linking related classes through SciGator, a DDC-based browsing library interface. Catalogue & Index, 2017, 188. hal-01886550

HAL Id: hal-01886550

<https://hal.science/hal-01886550>

Submitted on 2 Oct 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Mapping classifications and linking related classes through SciGator, a DDC-based browsing library interface

Marcin Trzmielewski, Claudio Gnoli, Marco Lardera, Gaia Heidi Pallesstrini and Matea Sipic, Library Service, University of Pavia

Introduction

The plurality and diversity of KOSs on the horizon of documentary institutions and the missing of index unification at different levels (local, national, international) disturbs the interoperability of bibliographic data and makes a documentary research less pertinent. Web 3.0 is still a vague promise of interconnection between machines and tools and of an easy exchange of bibliographic data. The first step into the interoperability of the records in the OPAC is the provision of tools allowing mapping of different index vocabularies in the same documentary area or institution. With the information explosion and new trends in science that prefer interdisciplinary studies, plenty of contemporary published documents now deal with several subjects touched on the same work. Because of the limitations of enumerative classification schemes (often chosen for indexing documents in university libraries) where the combination of concepts from several disciplines is complicated or impossible, interdisciplinary aspects are not made retrievable enough. Today, with relational database systems and PHP, Java or HTML languages, adequate online tools linking different disciplines can finally be proposed.

SciGator¹, one of the rare tools that join mapping of index vocabularies and linking of related subjects on one and the same interface, was created in December 2014 by the Library Service at University of Pavia, Italy. As a general reference KOS, SciGator adopts the Dewey Decimal Classification (DDC), which is also used as a basis for linking related classes and mapping other schemes. The DDC-based browsing library interface was presented on June 28, 2017 during the Symposium of the EDUG Annual Meeting held at the National Library of France in Paris². With the present paper, we invite you to discover this innovative tool and to navigate with us in the OPAC.

DDC in Italy

Nowadays, DDC is the most used classification scheme to index documents in Italian libraries, which increases the potential interoperability of the local catalogues with other information resources and tools. An Italian version of the online DDC schedules exists, called WebDewey Italiana, to which the scientific libraries of the University of Pavia are subscribed. These libraries were quite scattered until 2009, when they were reorganized into 8 libraries and three of those, The Science Library, the Science and Technology Library and the Medical Library, were converted to a DDC-base shelving system. Conversion of documents in these libraries still progresses, while other libraries are considering to join this classification network as well.

Origins of mapping and linking interface

Unfortunately, each of the Pavia scientific libraries is still physically divided into several sections, each with a different tradition of shelving based on local schemes. Moreover, they are in different positions in the town. So, books on related subjects or on the same subject are often shelved in different places, which is a potential source of confusion for users. In this situation, the Library Service staff realized that DDC may work as a virtual bridge between different local schemes and shelvings.

1. SciGator interface: <http://scigator.unipv.it/>

2. EDUG, Annual meetings, 27-28 June 2017: <http://edug.pansoft.de/tiki-index.php?page=2017+meeting>. A short report of the meeting is available on 025.431: *The Dewey Blog* at <http://ddc.typepad.com/025431/2017/07/2017-european-ddc-users-group-meeting.html>

Moreover, input of a subject heading in the bibliographic records during the cataloguing process is not an available option. So, bibliographic records in Pavia are only indexed by subject when they are derived from the SBN national bibliographic database. For the other records, shelfmarks are the only database field where subject information can be stored.

Furthermore, a well-known limitation of DDC is that it forces every item to be shelved or indexed under a specific discipline. That implies a loss of information on related subjects discussed in the same work. For example, many books owned by our libraries deal with mathematical subjects applied to physics, or with building subjects applied to architecture. This factor stimulated our staff to look for a tool allowing more interdisciplinary searches.

All the above considered, the Library Service came up with the idea of SciGator, an interface allowing to browse DDC classes adopted in Pavia libraries and at the same time to navigate from these to related DDC classes or to equivalent classes in different KOSs that are also used to shelve documents.

SciGator database structure

Like many of the modern search tools, SciGator is a relational MySQL data base structured in 6 main fields, as we may see in Fig. 1:

1. 'notation' – a DDC number, for example: 628.9
2. 'caption' - a WebDewey Italiana name of the DDC class, for example: *protezione antincendio*
3. 'captione' - English translation of the DDC class, for example: *fire prevention*
4. 'scopenote' - notes
5. 'seealso'- a DDC notation of a related class, for example: 363.1 (*public security*)
6. 'equivalent' - notation of an equivalent class in a local scheme, for example: AR16 (*fires*)

notation	varchar(20)	<input type="text" value="620.14"/>
caption	varchar(255)	<input type="text" value="ceramica e materiali affini"/>
captione	varchar(255)	<input type="text" value="ceramics and analogous materials"/>
scopenote	text	<input type="text"/>
seealso1	varchar(15)	<input type="text" value="666"/>
seealso2	varchar(15)	<input type="text"/>
seealso3	varchar(15)	<input type="text"/>
seealso4	varchar(15)	<input type="text"/>
equivalent1	varchar(20)	<input type="text" value="32H"/>

Fig. 1 The SciGator database fields

SciGator Web interface

SciGator interface is written in PHP language and managed by phpMyAdmin software tool and has a tree structure. Subjects can be found under their respective disciplinary class, or in a linked class in the network of cross references, as you may observe in Fig. 2. So, in this case 628.9 (fire prevention) is our root notation. Then, its related class 363.1 (public security) is connected by an arrow '→' and its equivalent in a local scheme AR16 is connected by around sign '≈'.

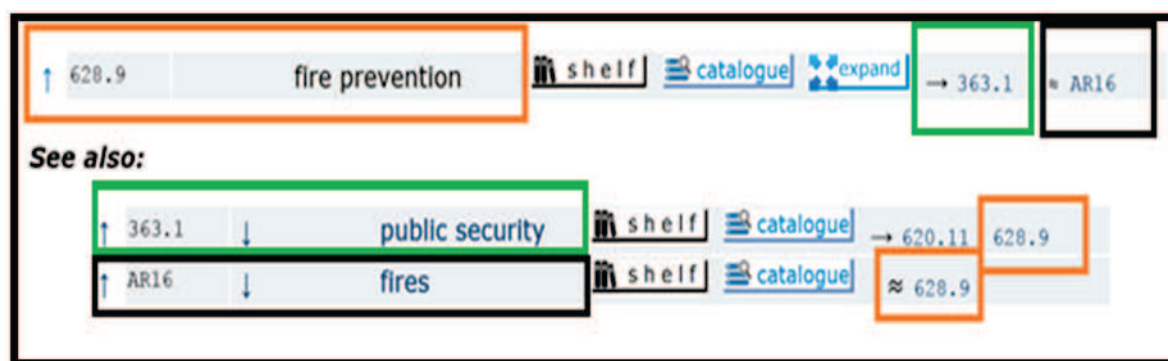


Fig. 2 The SciGator interface

Browsing functionalities

At present, SciGator interface proposes 3 browsing functionalities in the form of 3 buttons displayed between the root caption class and the related classes (Fig. 2).

The first one, 'shelf' (icon with black book spines), allows the user to retrieve all documents having a shelf mark that begins with the corresponding notation. Its subclasses will also be included because of the default right truncation. The second one, 'catalogue' (icon with a lens and a blue list of records) enables the user to retrieve all documents having the corresponding notation as their shelf mark or as a subject metadata coming from SBN - Italian National Catalogue. The last one, 'expand', (icon with four divergent arrows), enables query expansion by retrieving all documents in the previous options plus documents shelved or indexed by related classes, including both associated DDC classes and equivalent classes in local schemes. Then, after each browsing action, the user is redirected to the OPAC interface when he may consult the results of its research. All the obtained records are sorted by descending date of publication to answer better the scientific needs of users, asking generally for the most recently edited documents.

Some limitations

The limitations of this tool are those of the online catalogue in which searches are launched. For example, the final notation for all the learning and exercises books, especially used in mathematics textbooks, is 07. But in the catalogue only right truncation is possible. So, in SciGator as well, a search for only exercises books (of calculus or geometry ones for example) won't be allowed because the search by left truncation is not available.

Furthermore, all the 3 browsing options proposed by SciGator interface often produce a great number of results, but the last one especially ('expand in the catalogue') may also increase noise, due to the well-known inverse proportionality between recall and precision. So, each user should be informed about that and take it into consideration during her expanded search.

Future work

Ideas for possible developments of SciGator increase day after day. Among others, we plan to eliminate the 'catalogue' button in the case where a class is from a local classification, as in practice only DDC classes can be found in catalogue records besides shelfmarks.

Another need would be to collect statistical data about the use of SciGator, which may allow us to identify categories of users and estimate their number. We suppose that SciGator interface is used the most by librarians, but we would like to reach other categories of users.

We are also aware that the Web interface should be improved through more user-friendly graphic design. Moreover, adding buttons for searching the same DDC classes in the SBN national catalogue and/or WorldCat browse button seems to be especially useful for interlibrary loan services.

Conclusion

In short, SciGator seems to be an effective mapping tool between DDC classes and their equivalents in local schemes and makes interdisciplinary search possible. This experience also shows how classification schemes have the potential to provide more powerful search tools than is currently the case on most information services.

Cooperation around a single tool between staff who came from different disciplines: librarians, information computer specialists, ERASMUS+ interns and the boys and girls of Alternative Civilian Service, including the authors of the present paper, proved to be successful. Librarians working at the front desk found SciGator a useful tool. We are now waiting for users' feedback, which would allow us to ameliorate its interface and to make it more user friendly.

Similar tools are starting to emerge on the European field. Norwegian mapping software 'CCMapper'³ or German project 'coli-conc'⁴ give evidence of the need of this type of search tools. At the EDUG Annual Meeting 2017 the opportunity of a potential collaboration between developers of such tools has emerged.

3. About Mapping through CCMapper: <http://www.uio.no/for-ansatte/enhetssider/ub/prosjekter/mapping-for-sluttbrukertjenester/delte-dokumenter/brosjyre-engelsk-endelig.pdf>

4. About coli-conc project: <https://coli-conc.gbv.de>