

Guidelines and methodology for the creation of the SBN authority file

Cristina Magliano
ICCU

Introduction

This conference follows the Study days on Cataloguing and authority control devoted to the future evolution of RICA and to authority control in SBN, which were conducted by the ICCU on November 21 and 22, 2002 in Rome. Among the participants were Barbara Tillett and Françoise Bourdon and also Prof. Mauro Guerrini. These study days also provided an opportunity for reflection and study in anticipation of today's international seminar, which has also been sponsored by the Libraries Directorate of the MBAC.

I should like to outline the context in which the national authority file project developed.

The birth and development of the National Library Service (Servizio bibliotecario nazionale) and the formation of the SBN Index, a collective catalogue of considerable proportions, formed and enlarged through online shared cataloguing activity, has led rapidly to the growth of the national database and of the database of authors (personal and corporate), which at present numbers approximately 1,900,000 records. In November 2002 the database consisted of 1,893,403 authors (personal and corporate) and 6,266,074 titles (5,346,993 monographs and 237,672 serials). In all there were 17,240,313 locations, the results of shared cataloguing by the libraries with SBN *Poli* and the loading of several Italian retrospective bibliographies, and of retrospective conversion projects mounted by libraries participating in the SBN cooperative network (such as the *Bibliografia nazionale italiana* 1958-1984, the *Bollettino delle opere moderne straniere acquisite dalle biblioteche statali* 1958-1980, retrospectively converted data relating to the *Fondi meridionalistici*, the *CUBI* 1886-1957).

From these introductory remarks it is clear that, from the beginning, the ICCU has been committed to the dissemination of standards for cooperative cataloguing. This it has done by publicizing the national application of standards for description and also by giving authoritative interpretations of our cataloguing code (RICA), especially the part concerning forms of name, so that all additions to the catalogue will be accurate, particularly with regard to access points, which also provide for comparisons guarding against the duplication of bibliographic records.

Furthermore, given the different types of materials and differing time-frames co-existing in the online catalogue, one comes across some choices made by libraries, reinforced by years of cataloguing practice in a non-sharing environment, which are not always uniform. Hence the need to provide criteria and methods for ensuring uniformity of treatment, which seemed particularly necessary with regard to elements affecting access and the grouping of records.

Work on the standardization of access points involved a comparison with the major foreign catalogues and bibliographies. For instance there was the formalization of punctuation in personal and corporate names, which was absent from our cataloguing code but practical for the automatic processing of data and especially for an arrangement that is acceptable and clear to the user.

Maintenance structure of the SBN Archive

In March 1994 a Working Group was created for essential regular maintenance activities. This was made up of expert librarians from the parties cooperating in SBN: librarians who - through

institutional expertise, geographic distribution, representation of particular SBN software packages, or special nature of functions and usage - could bring to the cooperative context, coordinated by the ICCU, the professionalism and commitment that are indispensable for their difficult and onerous assignment.

At present 14 library institutions are taking part in activity to do with Structure: the Regional Polo of Piemonte at the National University Library, Torino; Polo of the Universities of Lombardia – CILEA, University Library of Pavia; Polo MBCA at the Marciana National Library, Venice, Polo of the National Central Library, Florence; Polo of the National Central Library Vittorio Emanuele II, Rome; Regional Polo of Sicily, Central Library of the Region of Sicily, Palermo; Polo of the Library Network of Romagna, Ravenna; Joint Polo of Bologna - the University Library and the Library of the Archiginnasio, Bologna; the SBN University Polo of the Veneto, Padova; the MBCA Polo at the National Central Library Vittorio Emanuele III, Naples; Polo of the Cultural Institutes of Rome – IEI, Rome.

The first phase of data control was that of eliminating duplications among both titles and authors, as well as removing inconsistent and incorrect headings and adjusting them to agree with the dictates of our national rules.

The ICCU has published a *Guide to cataloguing in accordance with SBN procedures* that incorporates the ISBD standards, which are not explicitly present in our cataloguing code (RICA) and the coded data of the SBN format. The *Guide* has gone through two revised editions in order to incorporate directions and some general procedures for shared cataloguing practice, as these became consolidated and different institutions became involved.

The need for precise instructions above all for the treatment of access points has involved the group on Structure for monitoring in developing a series of instructions which together form a guide to acceptable practice in the performance of cooperative activity - a sort of code of ethics in relation to the work of cataloguing, having regard also to the needs of the end-user. In particular regarding authors, in order not to link a title with the wrong author, or duplicate an author already present in the Index, a great deal of emphasis has been placed on the need for care in searching and preventive identification, by checking for the presence of homonyms to which it is obvious that the record should not be connected, or for the presence of the author's name in a form other than the one used in the search.

Such rules also cover indications of a general organizational character, methods of work and especially, directions for the cataloguing/capture and correction of data. Over the years we have thus helped to maintain control over all data that the work of Structure regards as priority, but which has also involved cataloguers in the *Poli*. Intervention has mainly been concerned with elimination of author/title duplications and with corrections.

The experience gained has made it clear that the SBN Index requires a permanent body for maintenance, working full-time on quality control and conformity of data for authors and titles.

The SBN prototype authority file (AF)

The issues regarding an AF Archive began to be examined back in 1989 by an ICCU-BNCF study group. Analyzing the international standards edited by IFLA available at the time (GARE, *Guidelines for authority and reference entries*, published in 1984 and UNIMARC/ A of 1991), the group identified the elements necessary for the formation of such an instrument and assessed its feasibility, particularly comparing international standards and national rules. The group concluded its work in 1991 with the production of a technical document of functional specifications and

guidelines, among which a structure for technical cooperation between libraries was outlined at operational level.

The aim was to create an instrument for control within the environment of cooperation and SBN usage so that, across the riches of bibliographic data that were being accumulated, it would be possible to guarantee uniformity and consistency among access points and to derive the National authority file from them.

Unfortunately, from the analysis and study of the creation of such a model until its reality, several years were to pass, partly because of inadequate funding, partly because of the difficulty of testing a new and complex tool in a broad context such as that of SBN.

Beginning in 1992, under resolution CIPE 12/8/1992 published in the G.U. S.g. N.222 21/9/92, the Institute obtained an initial financial allocation to launch a study, in collaboration with the National Central Library of Florence, on the 'Reorganization of bibliographic services in Italy'. This study provided for the development of the SBN network through primary services in support of cataloguing and with assistance for research. From all of this arose the requirement to build a National authority file within the SBN environment. The original project provided for whole undertaking to be divided into different segments, execution of which then had to be subdivided into two stages:

In the first stage the principal functions came into effect:

- formation of a single database of authorities for authors and for uniform titles (gathering titles) of types A and B;
- procedures for loading of author records from the SBN author archive file, and of title records from the title archives for ancient and modern material respectively;
- functions in the Authorities archive of online creation, cancellation and correction/merging of records;
- procedures for periodic downloading in batch mode of data, validated or directly produced, from the Authorities archive into the SBN author archive; and, for uniform titles, into the Index archive files for modern and ancient titles respectively. Also the necessary control functions to guard against possible duplications or unwarranted suppressions during the loading and merging of "validated" data, in the Index, with an authority level higher than that of cataloguing.

The format for entry took into account international standards and the past experience of the ICCU in the course of work carried out as a participant in the Working Groups of the IFLA Section on Cataloguing.

In particular the rules contained in our code were applied with regard to forms of heading (*RICA, Regole italiane di catalogazione per autori*. Roma : ICCU, 1979), while at international level reference was made to the first GARE (*Guidelines for authority and reference entries*), published by IFLA in 1984, translated into Italian in 1993, and subsequently to GARR (*Guidelines for authority and reference records*) of 2001. A uniform structure was adopted for each type of entry in the archive file (authority record, reference, explanatory reference), with the aim of facilitating the international communication of bibliographic records. The GARR standard recognizes the importance of maintaining national rules in determining "authorised forms for headings", in order to facilitate consultation and understanding on the part of users; and those related to the display of authority and reference information for all types of heading, including uniform titles of legal and religious works, musical compositions, publications in series. In this sense this requirement has been met by providing for a type of heading also for titles and by adopting the method of making separate records for an authorized heading and for reference headings.

From the phase of experimentation and testing of the model, it then became possible to move to an experimental operation of the system, thanks to new funding.

The SBN Index enhancement project

With the project for enhancement of the SBN Index – L. 662.96 – “Services of deduplication of the modern SBN Index and augmentation of the SBN and national authority file”, which lasted for 18 months and was completed last November, significant development occurred in the work of maintenance and quality control of the bibliographic data in the national Index. Above all, however, a consolidation of the gains made to that point in the prototype national authority file project in SBN. The immediate goals have been:

- a) To eliminate records, authors and titles, duplicated in the modern SBN Index;
- b) To get the national authority file up and running.

Point a) concerned the action launched in November 2000, to eliminate duplicate records in the modern SBN Index database and clearly also foreshadowed deduplication of authors' names as an activity consequent upon the work on titles.

The second point concerned the fuelling and start-up of the authority file operation. The file had its origins in the SBN database and so it has the dual purpose of being a support and control mechanism for national union catalogue activities and at the same time, of serving as a point of reference also for organizations not in SBN. The working file resides on the central system, but it is a file separate from the Index database, with which it interacts at present in batch mode; however, with the new Index Evolution project/program it will be totally integrated with the bibliographic database operation and will be able to be updated on line.

In it are carried out the creation/correction/augmentation of authority records, references and general explanatory records, the structure of which follows the data provided for by GARR; in particular, notes containing biographical or historical information and cataloguers' notes showing sources consulted. In anticipation of international instructions, each authority record has been assigned an identifying number, alphanumeric with 13 characters, in place of the eventual international ISADN number.

This activity requires complex and dedicated work by expert staff. Determining the authorized form of the name of a person or body, or of a title, requires careful verification based on the code of rules and on inventory and other kinds of sources from which to extract the various pieces of information that will bear out the choices made. Furthermore, besides agreeing with the rules, those choices must take account of the requirements of a usage applying to contexts that differ according to type of institution (local government libraries, state or university libraries, etc.)

Participants

Together with the ICCU which was in charge and coordinated the SBN-AF Enhancement project, other participants were the BNCf, the BNCr, the Biblioteca Nazionale Braidense (**Milan**) with the CILEA university consortium and the University Library of Bologna. One member of staff worked on the project in each institution, under the supervision of a designated librarian.

In our situation the production, management and distribution of a national authority file of necessity saw several kinds of institutional involvement. This comprised part of the technical body working on SBN, which until then had contributed to the quality of the catalogue, that is, the *Structure management and maintenance of the SBN online catalogue*, the ICCU as project coordinator and

responsible for the National Library Service, and the National Central Library of Florence as responsible for the Italian national bibliography. The two central institutions have assumed a special role in this project, in that both are responsible for the creation of national tools and services and also for international exchange and the maintenance of international formats.

Actually, also at an international level (see the introduction to the 2nd edition of *Guidelines for authority records and references* - GARR, September 2000) IFLA emphasized how in different countries all bibliographic activities, including those relating to bibliographic control, are increasingly taking effect within a cooperative environment.

Starting from the assumption that each country is responsible for its national authors, work began with the names of Italian personal authors. In all 34,388 records have been examined and approximately 9,900 personal author headings processed, i.e. raised to authority record status complete with reference apparatus, biographical notes and notes of sources. In many cases this work also involved the systematic arrangement of the bibliographic records linked to the name.

Methods of work and selection criteria

The first stages of the work were slow and included experimentation to determine the most productive methods for work which had never before been implemented on a cooperative basis. The ICCU arranged training courses for the operators, conducted by the librarians who were trainers for the project. Later a series of instructions was issued, in particular specifying the qualifications to be used in author entries for the purpose of differentiating between like headings, in the following order of priority: a) qualifications which usually accompany the name; b) in the absence of such a usual qualification, chronological details are to be used; c) in the absence of both a) and b), a phrase may be formulated by the cataloguer, in Italian where possible. It was emphasized that qualifications in the form of a word or phrase had to be drawn from reference works, since they were to agree with the principle of preferring the form "commonly used". There were instructions for forenames in the form of initials, in particular when it is necessary for them to be spelled out, and a series of instructions for forms of reference and for notices grouping several undifferentiated authors. Such notices and the related procedures were published on the ICCU's website in order to communicate the choices made to cooperating institutions.

A procedures manual has also been produced and a file commenced for bibliographic reference works. This file is available for consultation online and is updated centrally, similarly to the one already produced by the Laboratory for retrospective bibliography.

On the basis of the workflow developed, operators are made aware of the steps to be followed, always checking the Index first for consistency and then working on the Authority file database.

On the latter is the network of references and all information pertaining to any notes which are necessary to identify a person who is connected to the bibliographic entities present in the Index.

This work in some cases has extended to the systematic arrangement of records relating to linked titles, working on both files in succession.

The greatest difficulties encountered concerned the stages of controlling the batch return of AF data into the Index database, which had never been previously tested in an online cataloguing environment, and methods relating to the selection criteria for the entities on which to work.

The choice of exhaustive preliminary research proved from the beginning to be a lengthy business and not useful, especially for modern authors who were not cited in reference works and with few associated titles.

The selection of areas by subject and/or timeframe, a choice made by the ICCU that brought to a conclusion work on Italian literary authors of the 20th century (approximately 1800 headings), demonstrated the validity of that choice, even with the difficulties associated with the numerous cases encountered of identical names.

The national authority file project and the Committee for the revision of RICA

It is important to bear in mind that the national AF project was proceeding at the same time as code revision was being undertaken by the Ministerial Committee on RICA. In this situation it was often problematic, where the code was controversial or unclear, to tackle cases which over the years had produced customary interpretations and practices that were not always uniform. Furthermore, some rather impractical solutions had emerged in the light of the retrieval potential of online catalogues, in which strategies can be adopted that in some cases disregard the standards traditionally followed. The section regarding forms of name is that which received most attention from the Committee, for the simple fact that, more than others, it is in need of updating and simplification.

Cataloguing in SBN, however, has never departed from RICA standards: an example of this is the choice of using dates to qualify only those authors with identical names, whereas others, including some at an international level, declare more widespread use of dates to be obligatory in the AF archive.

Future development

In the project for the evolution of the national Index there is envisaged a complete integration of the AF database with extension of the present procedures active in version 3270 of the SBN system in order to supply services independently of the Polo software. In addition the present archive authority file will be extended to other entities (uniform titles of music, edited collections, subject headings, classification, place names and printers' marks for ancient materials).

The main aim of the project is to restructure the complex architecture of the National Library Service (SBN), reviewing both the technological and organizational aspects of cooperation, for the purpose of opening up the system to collaboration with other networks and public and private structures - with special mention of other LMS systems – allowing a major expansion of the services offered to users.

The new protocol, besides guaranteeing continuity of operation of the present SBN Polo software, within an architecture and on a database both completely different, is meant to build architectural, organizational and technological tools to achieve the fundamental objectives of growth and evolution of the SBN system.

The plan for the database eventually having the potential to process records coming from databases external to SBN will set in motion the supply of new services. Such an “opening-up” through a new SBN-MARC communications protocol must of necessity bring with it tighter controls against duplication of data but must also allow for integration of data with eventual local solutions that will be able to enrich the network of links referring to authorities.

The model has been designed so as to be able to take in the variety of materials currently being processed by Italian libraries and present in databases presently managed by the ICCU that must be absorbed into the new database:

- ❑ Ancient books
- ❑ Modern books
- ❑ Music

The system preparing for the management of “Authority Files” integrated in the Database, will facilitate the standardization of the form of access points for cataloguing descriptions.

The archive file of national work thus conceived must yield – as a result of the choices made and the corrections and checking carried out – a reliable product, consistent and uniform in the criteria applied, valid for all authors and titles, whatever the source of the data, with regard to the various

databases external or present on the Index system (Ancient and Modern Indexes, Music database, Manuscripts, Census data).

This is a sizeable undertaking, which will require an enormous amount of checking and management, in that it requires cooperation on a national scale, which of necessity must involve various structures at various levels and must provide for them to be taken in gradually.

From the moment of cataloguing, all biographical and bibliographical information must be provided for new authors entered by all cataloguers on SBN. The ICCU has commissioned a study on the organization of such a service and an organizational model has already been identified.

A system with a strong, central coordinating facility is planned, to which are connected other facilities that are specialized and local, also with particular subject expertise. The operators of the AF will receive information from all sources and will also be open to proposed corrections. The central facility will have responsibility for training the cooperating facilities in matters regarding standards, rule interpretations and advisable checks on the work carried out.

The need also to relate to Italian non-library systems (museums and archives) will serve to harmonize differing requirements and tend towards a sharing and definition of data that they have in common, and specification of that which is separate, providing for use of differentiated displays in their respective OPACs, for the differing usage requirements of the various types of institution, also in the light of future interoperability.

Finally, the process of analysis and gradual evolution must take account of models already existing within international projects (VIAF, LEAF, NACO) and of IFLA's new study, FRANAR (Functional Requirements and Numbering of Authority Records) which is under development.

Furthermore, national application is planned of the *UNIMARC /Authorities, Universal format for authorities*, published by IFLA in 2001, of which the Institute is about to published the Italian translation.

I should like to conclude by remembering that in this international and cooperative context the idea has also been advanced of a harmonization of cataloguing rules and interoperability between different sectors, which is also one of the priorities and objectives of the IFLA Section on Cataloguing, as we heard from Dr Barbara Tillett. The library community is therefore hopeful for the advancement of standards that are more widely accepted and are applicable in different contexts and areas.