Quality Evaluation of Online Library Catalogues, Advanced Discovery Tools and Linked Data Technologies

Maria Teresa Biagetti, Antonella Iacono, and Antonella Trombone

1 Limitations in Using Online Library Catalogues

Monographs constitute the main part of the scientific production in the Human and Social Sciences (SSH). During the assessment of research, it is essential for evaluators to have reliable sources with which to verify the bibliographical data of the monographs presented by scholars who must be evaluated. The aim of this part is to verify the possibility of effective use of online public access catalogues to check the bibliographical data of the monographs published and presented by scholars in the course of the evaluation process.

First, it is necessary to define the features of monographs in the SSH and their treatment in cataloguing. Scientific publications in the Human and Social Sciences exist in great variety, unknown in other scientific fields. The list includes monographs by an individual author or several co-authors, with merged or distinct responsibility; collected works; the publication of unpublished historical sources and of monographs offering a new explanation of sources; indexes and bibliographies; monographs devoted to the description of older editions; collections of unpublished works; sociological essays.

Considering the Italian environment, the document drawn up by the Osservatorio della ricerca dell'Università degli studi di Bologna, "Definizione e principali criteri

M.T. Biagetti (✉)
Sapienza University of Rome, Rome, Italy
e-mail: mariateresa.biagetti@uniroma1.it

A. Iacono
Biella Public Library, Biella, Italy
e-mail: antonella.iacono@fastwebnet.it

A. Trombone
University of Basilicata, Central University Library, Potenza, Italy
e-mail: antonella.trombone@gmail.com

di valutazione dei prodotti della ricerca" (Bologna, June 2013),[1] is the most useful in order to define the concept of monograph.[2] It clearly defines different kinds of publications in the SSH area and provides the main features and peculiarities for each of them. The document addresses quantitative and qualitative elements. In particular, it underlines that monographs must show a significant extension and must offer a deepened study, characterised by a critical approach. Moreover, it points out that authors must have direct responsibility for the overall content of the work.

The last is a very important point. It involves the necessity of clearly recognising the direct responsibility of authors and co-authors in chapters of books during an evaluation procedure. This point is directly connected with specific problems of library cataloguing. The cataloguing rules provide broad categories in order to establish a standardised way to attribute authorship, but they cannot foresee all borderline cases. Besides this basic difficulty, the personal interpretations of the single cataloguers actively contribute to create an environment corrupted by different cataloguing choices. As a result, bibliographic data are inconsistent.

In addition to these remarks, we can note that in SSH it is not unusual to find books formally attributed to a single author highlighted on the title page, who in reality was the author of the introduction alone, but rather written by several authors who are not formally mentioned on the title page. Moreover, an author may be often presented on the title page as the only person responsible for a book in which are issued or reprinted many works by other authors. The author responsible for the publication selected the works considering a particular scientific point of view, that is, he or she was the *editor* even if he or she claimed the role of author for the whole content. In these cases it is difficult to establish if the responsibility is for the overall work.

Now let us consider the quality and reliability of the data offered by library catalogues. The main purpose of this part of the chapter is to highlight caveats and limitations in the use of online library catalogues.

The assessment of the quality of catalogues can be carried out by considering the following factors:

- Identify authors and disambiguate personal names.
- Distinguish the different editions of the monographs and possible reprints.
- Mistakes or personal view in attributions of authorship.
- Distinguish co-authorship roles.
- Distinguish purchased books from donated books.

---

[1] http://www.catalogougov.unibo.it/_docs/Definizione%20e%20principali%20criteri%20di%20 valutazione%20dei%20prodotti%20della%20ricerca_08.pdf

[2] Other notable definitions of monograph have been offered in Italy by the National Agency for the evaluation of University and research (ANVUR): VQR 2004–2010 http://www.anvur.org/attachments/article/122/vqr_d.m._n._17_del_15_07_2011_firmato.pdf; VQR 2011–2014 http://www.anvur.org/attachments/article/825/Bando%20VQR%202011-2014_secon~.pdf and SUA-RD for university research http://www.anvur.org/attachments/article/26/Linee%20Guida%20SUA_RD%20ALLEGATO%20A.pdf

1.1 The Identification and the Disambiguation of Authors

To identify authors and disambiguate those having the same names, a reliable tool for authority control is critical. The *Functional Requirements for Authority Data: a Conceptual Model* (IFLA 2009a) is the conceptual model for authority data.[3] The model defines the requirements that authority data should meet to satisfy the need of consistency and the necessities of users. As an example of *Authority file*, let us consider the *Authorities* service offered by the Library of Congress,[4] a tool that enables cataloguers to use uniform accesses to names, titles, subjects, and to disambiguate authors of the same name. It is a tool for use by technical services and an authority source of bibliographical information for users supplied on a separate website available free of charge. The Library of Congress authority records adopt MARC 21 *Format for Authority Data*.[5]

The Library of Congress developed its *Authority file* in collaboration with several libraries that participated in a cooperative program, *Name Authority Cooperative Project* (NACO), a project for authority control on both theoretical and practical levels. The libraries that participated in NACO must follow the standards established to create uniform accesses (Byrum Jr. 2003; Ilik 2015).

As regards Italy, Casalini Press, the most important seller of Italian books abroad, participated in the *Program for Cooperative Cataloging*, the largest project of participated cataloguing with bibliographic control, launched in 1995.[6] After its participation in the *Shelf-ready Project*,[7] authority data provided by Casalini were made available to participants in the *Program for Cooperative Cataloging*. Casalini Press is the provider of the authority records for Italian editorial works.

WorldCat[8] is the cooperative union catalogue developed and managed since 1996 by Online Computer Library Center (OCLC). WorldCat is the result of the control and editing of records produced by tens of thousands of libraries, OCLC members (OCLC Annual Report 2014–2015). Due to its size, WorldCat offers a great opportunity to find bibliographic information. However, its main limit is its lack of disambiguation of authors of the same name. Without a suitable authority file, authors' identification is impossible and data retrieval is imprecise. WorldCat

---

[3] http://www.ifla.org/files/assets/cataloguing/frad/frad_2013.pdf

[4] https://catalog.loc.gov/

[5] http://www.loc.gov/marc/authority/ecadhome.html

[6] http://www.loc.gov/catdir/pcc/ Starting from 2005, Casalini contributes to both *Authority control* projects, descriptive and semantic: *Name Authority Cooperative Program* (NACO), which starts in 1977, and *Subject Authority Cooperative Program* (SACO), launched in 1993.

[7] The *Shelf-ready Project* prescribed the delivery of books along with catalographic records, Table of Contents and the LoC control number. Since 2006, Casalini has annually furnished the Library of Congress with about 4000 *shelf-ready* books, almost half of the acquired Italian publications, ready to be put on the shelves, already processed and provided with an anti-theft device. (Genetasio and Terravecchia 2009).

[8] www.worldcat.org, online since 2006. WorldCat offers more than 347 million records, representing 2.3 billion items owned by libraries, in 480 languages and dialects.

Identities,[9] in fact, is the simple list of headings of authors and subjects, not completely disambiguated. Moreover, author searching is allowed using single terms of headings, which involves the retrieval of all the records in which the term is present.

Focusing now on the Italian environment, it is worth mentioning that the Servizio Bibliotecario Nazionale (SBN) managed by ICCU[10] since 2001 has assured uniform access for personal authors, corporate bodies, and uniform titles thanks to a reliable authority file, organised according to the *Guidelines for authority records and references* (GARR)[11] and UNIMARC/A.

The *authority file* made by ICCU[12] is therefore the essential tool to improve quality and retrieval of information present in the SBN collective catalogue and is developed in cooperation with the SBN's partners, either through normal cataloguing activities or through the more specialised work of defining the authority entries. The methodological guidelines and rules for implementing the authors' archive using a controlled and uniform procedure are provided in the document recently published by ICCU: Linee guida (2015). The OPAC of SBN provides a reliable and authoritative tool of *authority control*, with disambiguation through the date of birth and creation of authoritative cards for one part of the personal authors, often supplemented with biographical data. All authority entries can be consulted in the SBN OPAC. SBN contributes with its authority records (about 50,000 authority records relative to personal authors) to the Virtual International Authority File (VIAF),[13] the large OCLC collaborative project that collects the authority records of 30 countries.

In Italy there is another online tool with public access that makes it possible to verify monographs.[14] The *I libri* database created by Casalini Press[15] is a bibliographic database that offers a good level of coverage of commercial publishers and of academic publishing from 1985 to date, and includes updated data. It also shows bibliographic advertisements of publications that are in the publishers' programme,

---

[9] http://worldcat.org/identities/

[10] http://www.iccu.sbn.it/opencms/opencms/it/main/attivita/naz/pagina_335.html. The Central Institute for the Union Catalogue of Italian Libraries and Bibliographic Information (ICCU) is a central institute coordinated by the General Department for libraries and cultural institutes of the Italian ministry for cultural heritage and tourism (MIBACT). It was established through Decree of the president of the republic no. 805 dated 3 December 1975, assuming the functions of the National Centre for the Union Catalogue and bibliographical information (established in 1951). ICCU coordinates and manages SBN and national and international cataloguing projects, also using new technologies. http://www.iccu.sbn.it/opencms/opencms/it/

[11] http://www.ifla.org/VII/s13/garr/garr.pdf

[12] http://www.iccu.sbn.it/opencms/opencms/it/main/attivita/naz/pagina_335.html

[13] http://viaf.org

[14] It is to be noted that, by contrast, the Bibliografia Nazionale Italiana, elaborated by the Florence national central library on the basis of legal deposit by law, has been online for public access only since 30 May 2016, limited to years 2012–2014. Previously, it was accessible by purchasing a license for viewing.

[15] http://www.casalini.it/ilibri/index.asp.

with preliminary information, but it does not declare that its state is not yet definitive. Data are updated at the moment of effective publication of the monograph. It does not seem to have a robust system of *authority control*. However, even though it does not systematically disambiguate authors of the same name, it often provides indications of the author's university of affiliation in the notes field.[16]

1.2   Different Editions and Reprints

As regards identification of different editions of monographs, for the purpose of building an effective information source useful for evaluators in the activity of scientific evaluation, the catalogue must provide an easy and intuitive way for identification of different editions of a monograph, of different authorial roles and relationships among publications. Yet, still today, catalogues do not efficiently carry out this important function, with an evident loss in quality when retrieving bibliographic information.

In fact, research interfaces of online catalogues do not adequately apply the model outlined in *Functional requirements for bibliographic records* (FRBR). This study by IFLA, (IFLA 1998; IFLA, FRBR Final Report 2009b), which played an important role in the development and renewal of cataloguing theory in the past 15 years, has the merit of modelling the bibliographic universe, clearly distinguishing between different entities, ideal (works and their expressions) and material (editions and copies).

The catalogue organised according to the FRBR model should account for the identity of a publication considering its various titles, differences of its editions (manifestations), of the relationships between publications (e.g., between derivative works and works whose subject is constituted by other publications). In most cases it concerns a simple FRBR-*isation*, a procedure to realise groups based on the *works*, using specific algorithms. These groups are realised *a posteriori*, without restructuring the original records.

Considering the Italian environment and the SBN OPAC, it is to be noted that it can display the records of the different editions held by the libraries, which cooperate in the realisation of the national online catalogue. However, the Italian online catalogue also presents some critical elements. In addition to the records describing editions, at times we can have records describing reprints. A specific procedure makes this possible. The *Guida alla catalogazione SBN*, published in 1995 (ICCU 1995, p. 50–51) indicates to identify reprints as occurrences in the central index without creating a new record. Nevertheless, the *Guida* allows the cataloguer to create a new record for a reprint when one is not able to find the data of the edition in the index. On the other hand, the Casalini Press *I libri* database does not record

---

[16] The mentioned features refer to years 2014–2016. Starting from the year 2017, Casalini Press changed the layout of the database.

reprints; this is one important and appreciable element, very important when checking authors' monographs for scientific evaluations.

It is worth mentioning here that the new Italian cataloguing rules, REICAT (ICCU 2009), comply with FRBR, and that, starting from the moment of their application in Italy (2010), it will be able to improve the identification of works, expressions and manifestations. Despite the fact that the FRBR model was not applied in the Italian records before the publication of the Italian REICAT, and the different editions were not precisely differentiated, Italian public catalogues online can in any case be used to discriminate new editions from the unchanged reprints. They can also be used to identify false new editions, even when the title page declares it as a new edition. Among bibliographic data in records, it is possible to verify the ISBN number (if present), which changes with every new edition and is the same in case of reprints. One can also verify the pagination and the change of formats. One can check this information in the material description area, along with the presence of illustrations and accompanying material not present in the previous edition.

The procedure to verify bibliographic data can be automated through the development of an *ad hoc* software. The system, for instance, could take the authors' name and titles of works as input and acquire from the SBN OPAC the UNIMARC records, and in particular the following fields: 200, 205, 210, 215, 410, 700, 010$a. These fields concern the title and responsibility block, the place and date of publication, the name of the publisher, the material description data, the editorial series, the author's access and the ISBN number. The next step might consist in analysing the correspondence between the 10-digit ISBN format and the 13-digit format using the ISBN converter[17] tool. In the conversion from the 10-digit format to 13-digit, in fact, the last digit, a check-digit, is replaced using an algorithm. Thanks to this comparison, apparently different ISBNs can turn out to be simply ISBNs converted into the new format and therefore identify a reprint of the same edition and not a new edition.

1.3 Mistakes in Authorial Attributions

With regard to the verification of mistakes in attributions to authors and entities, we can remark that in the OPAC of the Italian Servizio Bibliotecario Nazionale, in particular in the case of publications devoted to the reconstruction of ancient book collections and the identification of editions, generally carried out on the basis of handwritten inventories, it frequently happened that cataloguers attributed authorial responsibility to the library that owned the book collection and not to the author who reconstructed it, and identified, described and indexed editions.

The Casalini Press *I libri* database, however, presents bibliographic information gathered from title pages without the cataloguers' interpretation. This places the Casalini's database in a particular condition. On the one hand, it avoids the always

---

[17] http://www.isbn.org/ISBN_converter

possible error of authorial responsibility and the wrong attribution of principal responsibility to an entity on the basis of the personal cataloguer's interpretation. On the other hand, the loss of the cataloguer's working according to specific cataloguing rules, REICAT in Italy, prevents roles and responsibilities from being correctly redefined, if these have been mistakenly reported on the title pages intentionally, e.g., in the case in which an editor, or a translator, is explicitly presented as author.

1.4 Co-author Roles

As regards the field of bibliographic description, both in the national standard (REICAT) and in the international ones (RDA, ISBD) the monograph is defined as a resource in only one part complete, or subject to be completed in a defined number of distinct parts (ICCU 2009; JSC RDA 2010; IFLA 2011).

The activity of bibliographic description assigns the authorial responsibility to publications on the basis of specific rules, defined by national and international standards mostly considering the main part of the publication and the mode of formal presentation of responsibilities in the primary information sources. Thus, it can happen, for example, that an author presented with absolute prominence on the title page could be given the paternity of a publication of which he has in reality written only a brief introductory part consisting in a modest number of pages.

With regard to co-authors, under the national REICAT rules or RDA indications examinations are restricted to the formal presentation of responsibilities on the prescribed information source. In neither case is it therefore possible to detect the responsibility of each cooperating authors, each one of them having written a part, or the responsibility for only one of the volumes published under the responsibility of all authors present in the sources. It is impossible to identify authorial responsibilities in the case of co-authors presented as being at the same level in the publication, and each one of them having written one part, declaring, e.g., personal responsibility inside the publication. Responsibilities that are attributed inside the publication cannot be detected through cataloguing data.

1.5 Acquisition of Books and Gifts

One of the factors relevant to the purpose of assessing catalogue quality is constituted by the possibility to verify, using online public catalogues, the responsibility of libraries in buying monographs and in accepting donations. Oftentimes university teachers and researchers donate a copy of their publications to their department library, as well as to the libraries of other Italian or foreign universities that own collections congruent with the topics treated in the volumes, and also publishing

companies often donate a copy of a book published by them to the library where it is presented.

It is important to take into account that many libraries are equipped with regulations that govern the acceptance and management of donations. However, the acceptance of book donations entails the assessment of consistency of gifts with topics and scientific level of the recipient library. This is very different from the decision to buy a book, to spend money, especially at a moment in which public budgets are particularly reduced. Such purchases are indicative of the will of the library, reflecting the library's guidelines for the development of its collection, to offer its users that specific book as it is considered significant for them. It is to be noted, moreover, that university libraries, in particular, adopt automatic procedures of purchase and return of books – *approval plans* – in agreement with publishers. Libraries decide the criteria to select books, often using the classes from DDC or LC, in accordance with their scientific sphere. However, vendors and providers choose single items, and the returns may not exceed 5% of the books sent (but this can vary, according to the agreements) (Nardini 2003; Morriello 2006).

To conclude, it is worthwhile to consider that in online public access catalogues information about the type of acquisition, gift or legal deposit by law are not shown, as this information is not considered relevant to users. However, MARC formats permit the inclusion of information about the type of acquisition, and the Library of Congress OPAC, for instance, in the 925 and 955 MARC fields shows step-by-step information about the book, from the acquisition to the cataloguing procedures. Besides, UNIMARC (IFLA *UNIMARC* 2007) and MARC21 formats permit the inclusion of information about the type of acquisition, by gift, bequest, loan, purchase, or deposit.[18]

The Italian SBN OPAC can display the records in UNIMARC and MARC21 formats, but in both cases it does not enable users to check information regarding the type of acquisitions.

After shedding some light on the advantages and drawbacks of using online library catalogues, we can say that the Italian SBN OPAC offers the best coverage for Italian libraries and a good level of accuracy of bibliographic data to verify the information of monographs published in Italy. Moreover, it presents a rather good authority control and can count on the participative cataloguing of thousands of libraries of a good level. It can be proposed as the most convenient tool to verify the bibliographic data of Italian monographs when performing assessment. It can be used along with the bibliographical database produced by Casalini Press, which is useful to control the edition of works when it does not register the reprints.

---

[18] http://www.loc.gov/marc/holdings/hd541.html

2 Discovery Tools: Hybrid Research Tools Based on the Web

Planning an analysis of data based on search tools available through libraries means necessarily taking into account the profound changes that have occurred – and are still in progress – in the search interfaces of electronic catalogues. Moreover, it is deemed important to highlight the increasing spread of research tools defined as "web scale discovery tools" that coexist with traditional online library catalogues. The analysis here proposed is aimed at examining the nature and operation of these research tools in order to understand if the new discovery tools, together with the OPAC or in their place, especially in university libraries, can affect the qualitative and quantitative outcomes of research into monographs.

Since the late 1990s, on the heels of the diffusion of electronic catalogues and OPAC, we have seen constant innovation in the online search interfaces of catalogues. The difficulties in using these tools in the beginning were mainly due to the need to structure research using exact terms or keywords, also putting them in the correct fields (Borgman 1996; Marchitelli and Frigimelica 2012). These interfaces are suitable for the "known item search" function, which implies the knowledge of at least one of the basic bibliographic elements – the author or the title of the publication – if not all the bibliographic details of the specific resource. Less expert users might have needed help from the librarian to formulate research in terms relevant to the system, just as for researches in specialised databases during the same period.[19]

Initially, the online catalogues sought to simplify research techniques and also allow the user to carry out an exploratory search in the catalogue using natural languages and allowing greater discovery of the publications possessed by a library. A primitive form of the discovery function, realised through catalogues, was offered by the "find" research function. This possibility was an alternative to the scrolling of lists of indexes and, combined with the Boolean operators, allowed the search for terms within all of the catalogue's description fields.[20] Alongside these innovations, the use of search engines as a source of information caused them to become the model for the elaboration of the next generation of catalogue search interfaces. These latter, conventionally defined "next generation catalogues", offer the possibility of simultaneously interrogating all the databases the library has access to, including the catalogue. These research systems interrogate the entire bibliographic collection of a library, including the institutional archives, the collections of online resources and the specialised databases subscribed to by the library (Christensen 2013; Nagy 2011; NISO 2013; Marchitelli 2015). The newest discovery tools add to the federated search system[21] consisting of next generation catalogues, in local and remote databases, a centralised index of scientific contents, a package of online

---

[19] Regarding search methods see Sweet (2012), Petrucciani (1984), p. 1–12.

[20] On the loss of professional and scientific identity of librarian instruments, see Petrucciani (2006).

[21] This expression means a simultaneous search in multiple databases, in this regard see Wang, Mi (2012).

resources to which the library has access through an indirect subscription included in the discovery system's licence.

The discovery tools are technically defined "web scale discovery systems" (Way 2010; Raieli 2015) and allow, therefore, the querying of the different resources that the library makes available to users through a single search box. The three distinctive components of a discovery tool are: the search interface, the central index and the local index (Gong 2012; Breeding 2015). The interface recalls that of search engines, it is navigable without the need for special instructions, starting from a unique Google like string, or with the possibility to articulate more search keys in an advanced search box. Search results are then presented in a short or complete format.

The sorting of results by relevance is a feature common to the various discovery systems and is accompanied by other ordering criteria, which, however, vary depending on the software chosen by the library. The algorithms that allow sorting by relevance of search results are the property of discovery systems and they are not always made public. Due to the huge amount of resources available in the various indexes interrogated by a discovery, including its internal index, sorting by relevance should enable the results that best match the search criteria to be placed at the top of the list, or that the most interesting and important resources appear first. The factors that determine the relevance of a publication can have a library science origin, therefore can be connected to the correspondence between the search terms and the results coming from the catalogue. Moreover, they can be influenced by the data of the library loans. Even user habits, the frequency of access and the number of connections to electronic resources identify the search paths and help establish rankings based on user behaviours that influence the discovery relevance criterion (Biagetti 2010; Breeding 2013). Consequently, for the purpose of an conscious organisation of search results, the other sorting criteria offered by the discovery tools besides relevance take on a fundamental importance. The possibility of sorting results by author or title, as well as by date, allows the user to get a list that is verifiable through objective criteria. In this regard it is necessary to point out that some producers of discovery systems are also suppliers of electronic publishing, and that without clear, objective and verifiable sorting criteria they might be believed to be favouring their own content through the discovery ranking (Kelley 2012).

Another feature of the discovery web interface is the use of facets and limiters to restrict or define the search results. The categories of terms used as facets are derived from cataloguing fields, from metadata and document formats.

The second part and distinctive element of the discovery tool is the central index, whose contents are not part of the library collections but are connected to the license of use, owned by the company that produces it. The central index contains metadata and full text resources that are the result of agreements with commercial publishers, to which are also added metadata and open access publications contained in research repositories.

Local indexes of discovery tools, the third part of which they are composed, contain very different documents that are located in separate databases different from the catalogue, such as digital or digitised library collections and the research institutional repository. These discovery systems allow, in fact, to index and also search specialised metadata formats like those of archives and museum materials.

The discovery tools are, therefore, index-based systems: the contents of all databases, local and remote, are re-indexed by the discovery system, including those subscribed to by the library with commercial agreements and the catalogue data. In theory, during the indexing process the system should treat all contents equally. In actuality, it is unclear how the contents are indexed and therefore retrieved in the research phase. There are no standards regulating this process, which remains, along with the relevance criteria, completely out of the librarians' control (Breeding 2015). Consequently, the results displayed by the interface of discovery tools are a secondary source, as they are the result of an activity of re-indexing of cataloguing data and of online resources' metadata searched by the discovery tool. To view the original descriptions, a connection must be made to the data source, always through the discovery interface, so to the catalogue or to the publisher's site, for example.

The problems that can be caused by the re-indexing method also derive from variations in the quality and completeness of metadata that these discovery systems receive from different sources. Metadata can vary in terms of quality and quantity (Somerville 2013) starting from the origin.

From what has been said so far, we can deduce that, in view of the possibility of simultaneous search in multiple, heterogeneous data silos, including proprietary indices of discovery tools, the critical issues detectable in these systems may be due to the lack of clarity about sorting criteria by relevance of search results. Another critical point is the re-indexing of all metadata originally attributed to resources and publications, an issue further complicated by the extreme variability and lack of clarity about the content of central indexes of discovery tools. The resources searchable through these systems are multiple, but, according to the considerations above it is clear that the accuracy cannot be the same as the precision achievable by querying the search interfaces of the original individual silos connected to the discovery tool. The discovery system offers less accuracy even compared to online catalogues, and a known item search carried out for a resource whose bibliographic data are partly known can be problematic. Similarly, researchers may prefer to query the native indexes of specialised databases directly because they need accurate answers, which do not require an exploratory research (Breeding 2015; Ellero 2013; Han 2012; Frederiksen 2015).

2.1 Checking the Availability of Monographs Through Discovery Tools

The effectiveness of discovery tools' search algorithms could encourage their application for quantitative and qualitative analysis of the presence of monographs in libraries. The limits currently recognised for such an application are the following:

- The re-indexing criteria of the resources implemented by the discovery producers are not made publicly available.
- The queried resources are heterogeneous and subject to ongoing negotiations, so the catalogue is only one of them.

- In almost all cases, the algorithms organising the search results cannot be negotiated with software houses. Therefore you can choose to query only the catalogues linked to the discovery tool, but the sorting order of the results is almost always not comparable to that of online library catalogues.

Moreover, it is necessary to examine carefully the websites that contain library search instruments, since many online catalogues have already replaced their OPAC (Online Public Access Catalogue), namely the search interface of the catalogue, with the discovery instrument adopted. If such a complete replacement has not occurred, there will be a coexistence of two search tools: on the one side the online library catalogue and on the other the discovery tool. The first stage of analysis must therefore include a distinction between websites wherein the catalogue and discovery tool coexist and those where the discovery tool completely replaces the catalogue search interface.

A second stage of analysis has to consider the search interface of the discovery tool, and, in particular, the sorting criteria of search results that it allows, and the requirements of facets and limiters applicable to search results. To search library catalogue data using discovery tools there must necessarily be a function that limits the query only to the catalogue. Moreover, success in terms of accuracy of this kind of search is linked to the type of classification of the results, or to the possibility of organisation of bibliographic records in short format that the discovery provides, beyond the relevance criterion. The possibility of organising the results through criteria characteristic of cataloguing indexing, in particular by author, title and date, allows data retrieval suitable for an activity of qualitative and quantitative verification of the presence of publications in a library.

For the purposes of verification to be carried out for our research project, a sample of monographs of 20% of the constructed database was selected for both Italian scientific areas surveyed.[22] Among the catalogues of libraries selected to test the presence of the monographs, for the purposes of this project the ones also queried through a discovery tool were chosen. The subsequent step was to choose two academic library systems as standard, an Italian one and a foreign one, which had not adopted the same discovery system.

Given these guidelines, the choice was the following: the library system of "Sapienza" University of Rome and the Oxford University Libraries.

The research therefore analysed only a sample of the monographs of both databases and the results are directly comparable with the searches for the same titles in their respective online catalogues.

The Oxford library discovery tool, SOLO (Search Oxford Libraries Online), offers two search options: "all libraries/collections" for a search in the local index of the discovery tool; "search everything" to search its local and central index.

---

[22] The survey was sponsored by the Italian Agency for the evaluation of the University and the research (ANVUR), and carried out in 2014 and 2015. See Section 4, Chapter 14 *Is the diffusion of books in library holdings a reliable indicator in research assessment?*, in particular Part 2.

Wanting to query only the catalogues of the Oxford libraries, the limiter "all libraries/collections" can always be selected before entering the search terms, namely the author and title of the monograph. In SOLO it is possible to select the following sorting criteria for the list of short results: relevance, date-newest, date-oldest, author, title, popularity.

The uniformly applied criteria for checking the presence of the monographs in SOLO (Search Oxford Libraries Online) are as follows:

- Search in "all libraries/collections".
- Search string composed by the author and the title of the monograph.
- Sorting of the list of short results by title or by author.

To perform a search in Discovery Sapienza it is possible to select the following sorting criteria of the short results: relevance, date-newest and date-oldest. The uniformly applied criteria to check the presence of the monographs in Discovery Sapienza are the following:

- Search using a keyword.
- Search string composed of the author and the title of the monograph.
- Use of the limiter "Available in library collections".
- Sorting of the list of short results by relevance, since the only alternative is a chronological order, ascending or descending.

The results recorded in the database show a perfect match of searches carried out first in the online catalogue and next in the discovery tool of Oxford University libraries: the search of the monographs selected as a sample provides, therefore, the same results in the catalogue and in the discovery tool.

The situation is different for the search done in "Discovery Sapienza": for monographs selected as a sample, and for both scientific areas, the searches in the online catalogue differ from those performed in the discovery system. The percentage of presence of monographs in the discovery tool amounts to about 30% less than the presence of these monographs in the online catalogue, for both scientific areas. However, removing the limiter "Available in library collections" when searching the monographs falling into cases described above, we find in 100% of cases the presence of a digital copy of the printed publications, resident in another local index or in central index owned by the discovery tool. In none of cases the digital copy is accompanied by the bibliographic record of the printed version, actually present in the online catalogue of the Sapienza libraries.

The case presented demonstrates the differences of use between the selected search systems and the differences in the results obtained in the case of a known item search carried out with the precise input of keywords, aimed to verify the presence of bibliographic resources in a database, not for an exploratory bibliographic research. In the case of "Discovery Sapienza", it is presumed that the presence in the indexes of a digital copy of the requested monograph prevails in the display settings over the presence of the same monograph in print format, partially hiding its presence in our search results.

2.2   Conclusions

The considered cases show that, with regard to the University of Oxford, there is a perfect match between the search done using the discovery tool and the search done through the library's online catalogue. As far as the results of Sapienza University library are concerned, there is a certain percentage of misalignment. Such a situation could depend on the settings adopted for the attribution of the relevance criteria to publications, for example. Actually, in the absence of selection criteria, or standard guidelines for the implementation of such discovery instruments, it can be stated that at present they are not suitable instruments for a quantitative analysis of the publications. They allow a federated search by querying all databases accessible through the library; they also make possible an exploratory research of resources indexed by the discovery tool but not owned by the library. However, searches through the online catalogue are to be preferred for quantitative and qualitative assessments, both for the clarity of the structure of the data and the comprehensibility of the document retrieval techniques, entirely managed in the field of Library and Information Science.

3   Application Perspectives of Linked Open Data in Research Assessment

At the end of this contribution, we felt it was appropriate to give an overview of the new opportunities that the application of *linked open data* to library catalogues could offer to scientific evaluation. In fact, the adoption of this technology to online library catalogues opens new and exciting perspectives that may also be of interest to this sector of inquiry.

First, we need to clarify what is meant by *linked data* and how this new approach can be used to enhance searches in online library catalogues.

As is known, Linked data, in the definition provided by the founder of the W3C[23] Tim Berners Lee, refers to a set of recommended best practices for publishing and connecting structured data on the Web, favouring the creation of a global information network whose contents are mainly exchanged and interpreted by machines, forming the basis for the realisation of the Semantic Web (Berners-Lee 2006).

Linked data can best be interpreted as a major paradigm shift in the way of understanding data, even of a bibliographic nature. The adoption of this technology offers the possibility to create "data" or structured information that is connected, interoperable and integrated with any other information found on the Web. It offers the possibility to integrate and connect the knowledge of the network in a global network of data connected to each other (the so-called "linked data cloud").

---

[23] Tim Berners-Lee invented the World Wide Web in 1989. He founded and directs the World Wide Consortium (W3C), the forum for technical development of the Web.

In the field of library OPACs this is a turning point: for years, libraries have entrusted their data to closed and exclusively *library-centric* bibliographic formats like MAchine Readable Cataloguing (MARC),[24] the main format used to store and exchange data for over 40 years. The adoption of this format has been identified by many scholars as the main cause of the slow evolution of library OPACs, since it directed their development towards solutions that have in fact prevented the use, exchange and discovery of bibliographic information on the Web (Yee 2009).

For decades bibliographic information was trapped in library OPACs and in various bibliographical archives, considered to be non-communicating "silos" (Naun 2010).

The most important paradigm shift that affects libraries is certainly that of the turn "from records to data". Creating "connected" bibliographic data in the new form of *linked data* means, in fact, first adopting the Resource description framework (RDF) as a new data model or a new logical model to express bibliographic data, modifying the concept of record as it has traditionally been conceived. The RDF model "breaks up" the information into "statements" or "triples" linking the data through qualified relationships.

In the Semantic Web it is expected that bibliographic data would adopt a form characterised by an increased granularity. As is the case for the Web, where the adoption of the LOD implies the turn from a network of HTML documents connected by links and designed as unique blocks of information to a new Web of linked data, in the field of libraries embracing the logic of linked data means building simple relationships between the elements of a bibliographic record and not between individual bibliographic records.

While the bibliographic record is currently formed by an indivisible block of bibliographic data, i.e. the traditional textual and monolithic record, with the application of linked data the record is deconstructed and broken down into a set of triples linked together and connected to other data on the Web.

This need for granularity of the bibliographic data is not new. With the formulation of the FRBR model (IFLA 1998) an irreversible process towards the conception of a more granular bibliographic record was begun. In this direction, the Library of Congress report *On The Record* (Library of Congress Working Group on the Future of Bibliographic Control 2008) was a turning point, and today an effective end point is represented by the drafting of a new BIBFRAME model (US Library of Congress and Zepheira 2016), currently under development, which should definitely replace the old MARC formats (Kroeger 2013).

This transition towards a new format of bibliographic data exchange will make the data fully compatible with the Web. The adoption of BIBFRAME is aimed to replacing MARC21 with a format fully suited to the Web, where there are both FRBR model entities, but also semantic annotations of various types and sources. Without going into the details of the new record structure predicted by BIBFRAME, it should be noted that this model aims to create a "bibliographic environment"

---

[24] https://www.loc.gov/marc/

(Miller et al. 2012)[25] that is built into the Web and may also contain information added by non-bibliographic sources.

Along with BIBFRAME as the primary means to encode a new widespread bibliographical environment in the Semantic Web, within the context of cataloguing theory the new RDA cataloguing rules (Resource Description and Access) published in 2010 created a new standard for access and description of information resources, "specifically designed for the digital world".

The cataloguing theory is oriented in this new direction: recent RDA (Joint Steering Committee for Development of RDA eds. 2010) cataloguing standards, thanks to the fruitful cooperation being developed with the Semantic Web community and following the logic of *linked data*, introduced in cataloguing a new scenario where every bibliographic record is formed by the extraction and orderly recomposition of "data" coming from various sources and made up by archives of names, works, expressions, events, places, concepts, etc. (Coyle 2010).

The new rules are clearly inspired by the Semantic Web, stating the need that bibliographic data be geared more closely to the data (being more "data-friendly"). As is known, the RDA is based on two fundamental objectives of identifying and connecting the resources. These goals come directly from user functions established in the FRBR/FRAD (IFLA 2009a) and ICP (IFLA 2008, 2009a, b) and are reflected in the articulation of the standards content. The growing role of authority records in the resource description clearly emerges in the new code.

The true new development of the RDA, as is clear from reading the guidelines, is its primary focus on content. It is concerned with the registration of the entity attributes and the recording of the relationships between the entities, i.e., the choice of the size of the catalogue and their attributes, and no longer the display or presentation of the elements. In this way RDA marks a sharp break from all previous codes, almost a Copernican revolution: from records management to entity management. We no longer produce a catalogue consisting of records, but rather define individual data formulated with terms extracted from ontologies and vocabularies on the Web.

The process of "identification" of the bibliographic data is conceived as a process in which each bibliographic authority data is accurately and uniquely identified and linked with other data. This systematic process of "precise identification of a resource on the Web' via a URI (Uniform Resource Identifier) allows you to build "dynamic links" to the "Web of data" that from the bibliographic data easily reach other types of data.

3.1   The Benefit for Scientific Research Evaluation

The evaluation of scientific research can greatly benefit from the new perspectives opened up by the application of linked data to bibliographic data and with the new cataloguing scenario we briefly mentioned. In this context it is clear that, in the

---

research work of testing the availability of an author's monographs in library catalogues, the use of systems that adopt this technology will have a strong impact in some key aspects that here we can only briefly list:

3.1.1 Authority Control

The authority control is exercised to accurately identify bibliographic entities (a person, a corporation, a work), through a set of useful information in order to avoid the formal inconsistences of the catalogue entries (IFLA 2009a).

The control of authority data is a crucial element in the evaluation of scientific research, and allows the catalogue user to clearly and precisely identify each entity of interest (the author and his/her publication). A critical aspect of bibliographic data created and currently present in the online catalogues is certainly the inaccurate management of identifiers of the entities registered in the bibliographic records.

In these archives, entities' identifiers (e.g., an author) are created and maintained locally and can hardly be of any use to the library software, which normally does not handle them but simply records their presence. In the future – with the application of RDA and Linked Open Data – it will be possible to use external repositories to identify these data and "qualify" the relationships between them, using attributes defined in specific domain ontologies (e.g., FRBR, RDA etc.).

In the context of bibliographic control, this process will become increasingly important and widespread.

RDA puts into large account the authority data, obtained by special controlled vocabularies, with a high presence of open vocabularies that can be enriched with new terms by the communities that will use them.

The use of dereferenceable URIs for bibliographic data ensures the correct identification of persons and uses the connection to international authority services such as VIAF, the Virtual International Authority File. VIAF[26] is a service implemented and maintained by OCLC in cooperation with 20 national libraries, which virtually includes the authority files in the catalogues in a single authority service. The aim of the project is the reduction of costs and language barriers and availability on the Web of authorities that are exposed in linked open data.

For a long time catalogues have kept their authority files not always shared or accessible. Cataloguers have created the authors names following the rules laid down in their own cataloguing guidelines. The ability to link a given authority to VIAF, which contains all forms (REICAT, RAK, RDA etc.) created by cataloguers, besides simplifying cataloguing work, avoiding errors and duplications, creates an advantage in the search process. Someone carrying out a search in an OPAC could find a bibliographic description starting from the desired form for the name and find the link to all the variant forms in other languages.

---

[26] http://viaf.org/viaf/data

3.1.2 Application of FRBR and Improvement of Bibliographic Information Retrieval

As part of the studies made in the last two decades, Functional Requirements for Bibliographic Records (FRBR) is undoubtedly the most important theoretical reflection of the objects of cataloguing and bibliographic record functions in light of new technological changes, but yet still suffers from the failure of the model's cataloguing application.

The application of the FRBR model to catalogues has a direct and profound impact on searching and retrieving information. The full application of the model would allow the user to perform "significant navigation" in the catalogue, making it possible to clearly understand the retrieved entities.

The catalogue structured according to the FRBR model should give an account of the identity of a work under its various titles, the differences between the editions (manifestations), between the works (e.g., reports, including derivative works and the works that have other works as subject).

It should also allow the proper identification of the authors (especially useful in cases of same names), and the different roles of responsibilities (viewing the works and expressions in which an author has responsibilities of different types e.g., all the works in which he/she has a primary responsibility, all those in which instead he/she has a secondary responsibility). In particular the implementation of FRBR is essential to allowing the catalogue to perform its aggregating function, or grouping all the expressions and manifestations of a work, thus presenting to the user the works and related editions and links with other works in an orderly manner. It is a natural consequence of a higher *bibliographic control* or the possibility to better account for all the variations of names and titles.

Although there is now a consistent set of cataloguing rules that are inspired by the model outlined in the FRBR study (including the Italian cataloguing rules REICAT), the views of the conceptual model are applied with poor results in the library OPACs, reducing the quality of the information retrieved in the catalogue. To date, the only library catalogues that apply the FRBR model, are the "new generation OPACs" or discovery tools. In these instruments, however, we speak of "FRBR-isation" as a technique to obtain groupings according to the work, made possible through special algorithms employed retrospectively by the software without changing the structure of the original bibliographic records. Precisely for this reason, these groups often prove ineffective. In conclusion, even when doing research in OPACs that declare explicitly to apply the FRBR model (e.g., WorldCat), you cannot always easily retrieve the individual works, editions and reprints, as you find yourself having to further analyse the results retrieved.

Searching and browsing in online catalogues today requires special attention, especially for the verification of the different editions of a work and the rebuilding of links between works that are not always adequately made explicit. This happens because we do not intervene directly on the bibliographic record that, as we said earlier, is expressed in a format (MARC) that is not suitable to express the richness of bibliographic relationships provided by the model.

The deconstruction of the bibliographic records in data that are linked through LOD technology and the ability to qualify the links between the data by reconstructing the relationships required by the IFLA[27] Functional Requirements conceptual models will allow for a hierarchical navigation in the future that groups related works with the individual works, expressions and manifestations, and the items linked to them. This will effectively perform "bibliographic functions" for the benefit of the user when using the catalogue, making it possible to identify the "significant" contents of the catalogue. There are already significant applications in electronic and digital library catalogues.

As is clear from what has been presented so far, the application of LOD increases the chance that the information retrieved from online library catalogues will provide some elements and indicators to assess the activity and value of the scientific production of an author.

These new elements can be added to others, providing effective feedback in the evaluation process, improving the quality and amount of data retrieved when querying an OPAC.

3.1.3 Convergence of Search Tools

The new perspective of the Semantic Web directly involves the development of information retrieval systems such as electronic catalogues and means of carrying out searches, moving beyond traditional means of information retrieval, facilitating the discovery of the contents sought as well as supporting documents at the same time, and integration with other documents in the Web by creating a widespread network of knowledge. There will be, therefore, an increasing number of large platforms where the bibliographic information will be integrated with other information from other areas of cultural heritage, such as archives and museums. At that point from a given piece of bibliographic data it will be possible to find other data that enrich the information found, drawn from cultural heritage data available on the Web.

3.1.4 Improved Search Interfaces and New Ways of Cataloguing Research

Some concrete projects for implementing linked open data in library catalogues – such as the recent case of the OPAC of the Swedish National Library LIBRIS,[28] or data.bnf.fr project, the great platform of the Bibliothèque Nationale de France,

---

[27] This is referring to the models that make up the so-called "family" of IFLA functional requirements FRBR, FRAD, FRSAD, now in the process of consolidation in the new FRBR-Library Reference Model (IFLA 2016).

[28] The OPAC of the Swedish National Library LIBRIS is available at URL: http://libris.kb.se. Since 2008, the library displays the data as Linked data using ontologies known as Dublin Core, Bibliontology and SKOS, and is preparing to adopt RDF (Resource Description Framework) as its

which brings together bibliographic data, archival and museum (then incorporating the OPAC data of BNF) – show the potential of linked open data applications in the field of bibliographic research.

The Swedish OPAC LIBRIS features an innovative interface that connects data from multiple sources and displays them while maintaining the logical sequence that leads from the "works" to their "expressions", to the individual "manifestations" as provided for by FRBR models. This resolves a problem that affects today's electronic catalogues, yet still based on the single "manifestation" and providing the user with an interface based on the works, the authors, on the subjects. They exploit the spaces to explore the data related to every entity of interest in the catalogue, greatly improving the retrieval. You can get richer results and smarter groupings (groupings of works with related issues or events) with great benefit for research.

In the French platform data.bnf.fr the query interface offers three main channels of access to the catalogue: the work, the author, the subject. For each entity is shown a special page made of linked and updated data. The platform has a page for each work, topic, and author of the catalogue, structured according to the levels provided by FRBR, and created dynamically using the authority record for the work provided by the traditional catalogue of the Bibliothèque Nationale de France. For each work the key information about the content, the related works and their relationships with its parts are provided. The query by author responds with a page for each author containing biographical information, a list of works and an indication of other works connected with the indication of the type of relationship (Wenz 2013). The query by subject, on the other hand, responds to those carrying out a semantic search with a page for each subject. These subject pages contain all the variant and associated forms and shows the authors who have treated the subject, providing the works of the catalogue on that subject. All pages make use of further information from external datasets, in particular from DBpedia, the largest interdisciplinary datasets and benchmark for the extraction of data in Linked Open Data. External connections are imported into the platform via additional information about the works, titles and subjects. All data is queried by search engines through the application of RDFa standard markup and the use of shared vocabularies like Schema.org, recognised by the major search engines.

The Italian Share Catalogue project[29] shows how it is possible to create an interface in LOD for unified access to the catalogues of a group of academic libraries, through the creation of an LOD substrate of people-works that greatly simplifies the research process, exploiting the enrichment coming from the Web of data and the modelling coming from the BIBframe specifications.

---

native format of the catalogue records. The introduction of the new version of the catalogue Libris XL expressed in the new linked data format includes both the new and the previous records, the purpose of replacing MARC 21 with a data format as close to the data model outlined in Bibframe. (Forssblad 2013). The data are published in JSON-LD (JSON for linking data, http://json-ld.org). Today, thanks to the transformation of data in RDF, in response to a question the OPAC is able to provide a list sorted according to and within the FRBR model and will be able to provide a page for each title, author and subject of the database.

[29] http://www.sharecampus.it/1/share_catalogue_969557.html

The main and most enlightening perspectives, which at the moment, however, do not find any real application – come from the ability to enrich the bibliographic authority data with data coming from universities and research centres.

Of particular interest – in terms of richness of content obtained from a search in the catalogue – are the contributions of research institutions and their data that will be made available and could usefully enrich the bibliographic data.

This data may include:

– Universities' registers appropriately exposed in linked open data.
– Research data conveyed and made available in open format.[30]
– The data captured by ontologies linking scholars and research products, products in the context of cooperation projects among universities[31] (see recent projects that connect the scientific production of universities like, for example, *Linked data for Libraries project*[32] and *Linked universities project*[33]).
– Data from the institutional repositories that will be exhibited in the form of linked open data (Konstantinou et al. 2014).

The connection of all these data could provide a network of information or a kind of mapping of scientific research, which could be exploited as a useful outline to bibliographic environment.

3.1.5   Interoperability with Web: Quantity and Quality of Bibliographic Data in the Open Web

What has been stated so far focuses on the importance of interoperability. This theme has always been at the heart of library issues: libraries and IFLA have long been concerned with ensuring interoperability between archives, libraries, and other components of the bibliographic circuit. Today, though, the question becomes crucial: it is no longer sufficient to produce data that is interoperable between the actors of the cultural sector, but it must be interoperable with the Web i.e., to make library data freely accessible online.

---

[30] On the issue of the scientific data in open format see Cassella (2013).

[31] Among these VIVO is an ontology standard that connects researchers and communities using LOD, but also a data set and an open community with strong international participation that connects researchers, publications, resources for research, financing, required courses, scientific activities. The ontology is based on other known ontologies (SIOC and FOAF). The major classes of this ontology are people, organisations and information resources.

[32] *Linked data for Libraries* is a project of Cornell University Library, Harvard Library Innovation Lab and Stanford to create a LOD platform to extract data from various existing silos, to improve access to and discovery of academic information in the three libraries

[33] *Linked universities* is a project to connect the university library resources in LOD format involving various universities in the UK, Germany, Finland, Greece, Turkey, Czech Republic. The project aims to create and support specific vocabularies for universities for common concepts like courses, qualifications, teaching materials and sharing experience, the tools to format data of the University as linked data.

The application of linked data to library catalogues will deliver the bibliographic data to the entire Web. Bibliographic information can then be retrieved by search engines (both traditional and semantic). OCLC has been moving in this direction, releasing a huge amount of data of the world's largest single catalogue, WorldCat, in the form of linked open data and incorporating into their pages the marking provided by RDFa (RDF in Attributes) which makes it possible to embed semantic annotations pages in the Web following a special scheme[34] or ontology that is recognised by the major search engines.[35]

In the emerging future, through exchange formats compatible with the Semantic Web, these data may be freely available on the Web in the new structure of the bibliographic records designed in BIBFRAME, which is destined to host new and valuable information. The new bibliographic model developed by the Library of Congress for the Semantic Web and the world of linked data, BIBFRAME also includes some important new characteristics that are worth noting.

Along with data on the work, its editions, and authority data, the model also includes the Annotation class. This class is designed to include information about a resource, such as administrative and management information of the resource's life cycle (as traceability and provenance are part of the workflow that characterises the creation of a bibliographic record), but also other data that are added by others (users, commercial actors, the web community).[36] Not to mention metadata created by cataloguers: data resource locations and holdings, and access policies (Mitchell 2013).

This will involve the online presence of management data for the item that are today "trapped" in library management systems (ILS) and are not present on the Web. We know how complex it is to extrapolate a number of management data for assessment purposes, such as whether a monograph was purchased or donated. In fact there is no trace of this data in the library OPAC, but they remain stored in the ILS, together with other data about the copy (the FRBR item).

Along with this information, annotations can also be created by third parties and library users, who will enrich the record with images, reviews and ratings obtained from reliable sources.

With the application of linked data bibliographic information disintegrates into the Semantic Web, therefore drawing upon data from different sources. It will be of primary importance to ensure the quality of cataloguing data, a problem that is both theoretical (what does quality of data mean) and practical (how to ensure quality

---

[34] https://schema.org/

[35] The OCLC data policy is presented in http://www.oclc.org/data.en.html. See as an example of enriched records with linked open data record in WorldCat. https://www.worldcat.org/title/information-a-very-short-introduction/oclc/743804876&referer=brief_results where you can see the RDFa encoding based on Schema.org (www.schema.org)

[36] The purpose of the Annotation class is to "Express opinions about a resource, for example a review. Attach institution specific information, for example holdings. Contribute enhancements to a resource description, for example cover art or summary descriptions". See BIBFRAME Annotation Model (2013).

through the certification of data provenances) and that greatly impacts the assessment practices. The commitment of the libraries, so it is hoped, will be to build certified and reliable data networks.

References

Berners-Lee, T. (2006). Linked data: Design issues [updated version 2009 06 18]. http://www.w3.org/DesignIssues/LinkedData.html. Accessed 17 July 2016.

Biagetti, M. T. (2010). Nuove funzionalità degli OPAC e relevance ranking. *Bollettino AIB, 50*(4), 339–356.

BIBFRAME Annotation Model. (2013). *BIBFRAME Community Draft, 26 August*. http://bibframe.org/documentation/annotations/#whatIsAnAnnotation. Accessed 17 July 2016.

Borgman, C. L. (1996). Why are online catalogs still hard to use? *Journal of the American Society for Information Science, 47*(7), 493–503.

Breeding, M. (2013). Next-generation discovery: An overview of the European scene. In S. Chambers (Ed.), *Catalogue 2.0. The future of library catalogue* (pp. 37–64). London: Facet Publishing.

Breeding, M. (2015). *The future of library resource discovery*. Baltimore: NISO.

Byrum, J. D., Jr. (2003). NACO: un modello di cooperazione per l'allestimento e la manutenzione di un authority database condiviso per i nomi. In *Authority control: definizione ed esperienze internazionali. Atti del Convegno internazionale (Firenze, 10–12 febbraio 2003)* (pp. 187–198). Firenze/Roma: Firenze University Press/Associazione Italiana Biblioteche.

Cassella, M. (2013). Dati aperti e ricerca scientifica: aspetti gestionali e normativi nel contesto dell'e-science. *AIB Studi, 53*(3) doi:https://doi.org/10.2426/aibstudi-9131.

Christensen, A. (2013). Next-generation catalogues: what do users think? In S. Chambers (Ed.), *Catalogue 2.0. The future of library catalogue* (pp. 1–15). London: Facet Publishing.

Coyle, K. (2010). RDA vocabularies for a twenty-first-century data environment. *Library Technology Reports, 46*(2) (monographic issue).

Ellero, N. P. (2013). Integration or disintegration: Where is discovery headed? *Journal of Library Metadata, 13*(4), 311–329.

Forssblad, H. (2013). *Uppdaterad information om Libris XL och nya katalogiseringsverktyget*, Librisbloggen.kb.se (blog).

Frederiksen, L. (2015). Exploring discovery. *Public Services Quarterly, 11*(1), 43–47.

Genetasio, G., & Terravecchia, E. (2009). Authority control at Casalini libri. *Biblioteche oggi, 27*(10), 42–52.

Gong, R. (2012). A review of planning and implementing resource discovery tools in academic libraries. *Journal of Electronic Resources Librarianship, 24*(4), 340–341.

Han, M. (2012). New discovery services and library bibliographic control. *Library Trends, 61*(1), 162–172.

ICCU. (1995). *Guida alla catalogazione in SBN. Pubblicazioni monografiche, pubblicazioni in serie* (2nd ed.). Roma: ICCU.

ICCU. (2009). *Regole italiane di catalogazione REICAT*. A cura della Commissione permanente per la revisione delle regole italiane di catalogazione. Roma: ICCU. http://www.iccu.sbn.it/opencms/export/sites/iccu/documenti/2015/REICAT-giugno2009.pdf.

ICCU. (2015). *Linee guida per la compilazione dell'Authority Control-Autore personale in SBN*. http://www.iccu.sbn.it/opencms/export/sites/iccu/documenti/2015/AF_Indicazioni_e_regole_riv_23_luglio2015.pdf. Accessed 18 Sept. 2016.

IFLA. (1998). *Functional requirements for bibliographic records: Final report*. München: K. G. Saur. http://www.ifla.org/VII/s13/frbr/frbr.pdf. Accessed 20 Sept 2016.

IFLA. (2007). *UNIMARC Manual for Holdings Format. Version 1. Recommended by the working group on the UNIMARC holdings format*. Approved by the Permanent UNIMARC Committee (PUC), pp. 36–86. http://www.ifla.org/files/assets/uca/unimarc-holdingsformat.pdf. Accessed 30 Dec. 2015.

IFLA. (2008). *Cataloguing principles: Statement of international cataloguing principles (ICP) and its glossary*. Edited by Barbara Tillett and Ana Lupe Cristán, München, K.G. Saur http://www.ifla.org/VII/s13/icc/imeicc-statement_of_principles-2008.pdf Italian version. IFLA (2009) Dichiarazione di principi internazionali di catalogazione, Roma, ICCU http://www.ifla.org/files/cataloguing/icp/icp_2009-it.pdf

IFLA. (2009a). In G. E. Patton (Ed.), *Functional requirements for authority data. A conceptual model*. München: K.G. Saur. Online version, as amended and corrected through July 2013, http://www.ifla.org/files/assets/cataloguing/frad/frad_2013.pdf.

IFLA. (2009b). *Study group on the functional requirements for bibliographic records, functional requirements for bibliographic records: Final report*. Approved by the Standing Committee of the IFLA Section on Cataloguing. September 1997. As amended and corrected through February 2009, http://www.ifla.org/VII/s13/frbr/frbr_2008.pdf.

IFLA. (2011). *ISBD International standard bibliographic description*, Consolidated edition. Edizione italiana, a cura dell'Istituto centrale per il catalogo unico delle biblioteche italiane e per le informazioni bibliografiche, Roma: ICCU 2012, p. 241. http://www.iccu.sbn.it/opencms/export/sites/iccu/documenti/2012/ISBD_NOV2012_online.pdf.

IFLA. (2016). *Library-reference model*. Pat Riva, Patrick Le Bœuf, and Maja Žumer Consolidation Editorial Group of the IFLA FRBR Review Group Draft for World-Wide Review 2016–02-21. Not yet endorsed by the IFLA Professional Committee or Governing Board, http://www.ifla.org/files/assets/cataloguing/frbr-lrm/frbr-lrm_20160225.pdf. Accessed 17 July 2016.

Ilik, V. (2015). Cataloger makeover: Creating non-MARC name authorities. *Cataloging & Classification Quarterly, 53*(3–4), 382–398.

Joint Steering Committee for Development of RDA (ed.) (2010). *Resource description & access*. RDA. American library association; Canadian library association; CILIP.

Kelley, M. (2012). Coming into focus: Web-scale discovery services face growing need for best practices. *Library Journal, 137*(17), 34–40.

Konstantinou, N., et al. (2014). Exposing scholarly information as linked open data: RDFizing DSpace contents. *Electronic Library, 32*(6), 834–851. Accessed 17 July 2016.

Kroeger, A. (2013). The road to BIBFRAME: The evolution of the idea of bibliographic transition into a post-MARC Future. *Cataloging & Classification Quarterly, 51*(8), 1–18.

Library of Congress Working Group On The Future Of Bibliographic Control. (2008). *On the record: Report of the library of congress working group on the future of bibliographic control*. Washington, DC: Library of Congress. http://www.loc.gov/bibliographic-future/news/lcwg-ontherecord-jan08-final.pdf. Accessed 17 July 2016.

Marchitelli, A. (2015). *Orientarsi tra le informazioni in biblioteca. Cataloghi, banche dati, motori di ricerca*. Milano: Editrice Bibliografica.

Marchitelli, A., & Frigimelica, G. (2012). *OPAC*. Roma: Associazione Italiana Biblioteche.

Miller, E., et al. (2012). *Bibliographic framework as a web of data: Linked data model and supporting services*. Washington, DC: Library of Congress.

Mitchell, E. (2013). Assessing the value of metadata in information services. *Technical Services Quarterly, 30*(2), 187–200.

Morriello, R. (2006). Acquisizioni tramite approval plan. Outsourcing o nuova opportunità per i bibliotecari? *Biblioteche oggi, 24*(1), 22–31.

Nagy, A. (2011). *Analyzing the next-generation catalog*. Chicago: American Library Association.

Nardini, R. F. (2003). Approval plans. In M. A. Drake (ed.), *Encyclopedia of library and information science*. New York: Dekker. 1, 133–138. Accessed 28 June 2016. doi:https://doi.org/10.1081/E-ELIS 120008874.

Naun, C. C. (2010). Next generation OPACs: A cataloging viewpoint. *Cataloging & Classification Quarterly, 48*(4), 330–342.

NISO. (2013). *ODI survey report: Reflections and perspectives on discovery services*. Available at http://www.niso.org/apps/group_public/download.php/9977/NISO%20ODI%20Survey%20Report%20Final.pdf. Accessed 16 July 2016.

OCLC Annual Report. (2014–2015). http://www.oclc.org/etc/designs/oclc/oclc-annualreport-15/images/215510-2014-2015OCLCAnnualReport.pdf.

Patton G (ed.) (2009) Functional requirements for authority data: A conceptual Model. Ed. by Glenn E. Patton. IFLA working group on functional requirements and numbering of authority records, final report December 2008. IFLA series on bibliographic control, 34. K.G. Saur, Munich.

Petrucciani, A. (1984). *Funzione e struttura del catalogo per autore*. Firenze: La Nuova Italia.

Petrucciani, A. (2006). La catalogazione, il mercato e la fiera dei luoghi comuni. *Bollettino AIB, 46*(3), 177–185.

Raieli, R. (2015). Limiti dell'information discovery e necessità dell'information literacy. *Nuovi annali della Scuola speciale per archivisti e bibliotecari, 29*, 179–194.

Somerville, M. M. (2013). Digital age discoverability: A collaborative organizational approach. *Serials Review, 39*(4), 234–239.

Sweet M (2012) There's nothing wrong with discovery services that can't be fixed by the reference layer. In: *The Fiesole collection development retreat*, April 12 2012. Available at http://www.casalini.it/retreat/2012_docs/sweet.pdf. Accessed 16 July 2016.

US Library of Congress and Zepheira. (2016). *Overview of the BIBFRAME 2.0 Model*. https://www.loc.gov/bibframe/docs/bibframe2-model.html. Accessed 17 July 2016.

Wang, Y., & Mi, J. (2012). Searchability and discoverability of library resources: Federated search and beyond. *College & Undergraduate Libraries, 19*(2–4), 229–245.

Way, D. (2010). The impact of web-scale discovery on the use of a library collection. *Serials Review, 36*(4), 214–220.

Wenz, R. (2013). Linked open data for new library services: The example of data.bnf.fr. *JLIS.it, 4*(1), 403–416. doi:https://doi.org/10.4403/jlis.it-5509. Accessed 17 July 2016.

Yee, M. (2009). Can bibliographic data be put directly onto the Semantic Web? *Information Technology and Libraries, 28*(2), 55–80. also available online: http://escholarship.org/uc/item/91b1830k.