A Turning Point for Catalogs: Ranganathan's Possible Point of View

Carlo Bianchini\textsuperscript{a} & Mauro Guerrini\textsuperscript{b}

\textsuperscript{a} Department of Musicology and Cultural Heritage, University of Pavia, Cremona, Italy
\textsuperscript{b} Department of History, Archaeology, Geography, Arts and Performance, University of Florence, Florence, Italy

Published online: 22 Dec 2014.
A Turning Point for Catalogs: Ranganathan’s Possible Point of View

CARLO BIANCHINI
Department of Musicology and Cultural Heritage, University of Pavia, Cremona, Italy

MAURO GUERRINI
Department of History, Archaeology, Geography, Arts and Performance, University of Florence, Florence, Italy

Since the end of the last century, catalogs have been changing more and more quickly. This change is following a recognizable course, beginning with the publication of Functional Requirements for Bibliographic Records, passing through the reorganization of international cataloging principles, the revision of international standards of the International Federation of Library Associations and Institutions (International Standard for Bibliographic Description), and the foundation of new cataloging codes, such as Resource Description and Access. While principles, models, and rules are well established, bibliographic formats seem to be a bottleneck and users seem far from libraries. This article aims to present an overview of current changes, potential convergences, developments, and weak points from Ranganathan’s point of view.

KEYWORDS cataloging, Semantic Web, BIBFRAME, RDA, users’ needs, S.R. Ranganathan, Five Laws of Library Science

INTRODUCTION

In Ranganathan’s view, a library “comes into existence only when readers, books, and staff function together. Readers, books, and staff form the trinity in a library.” He adds that “a collection of books becomes a library when,
and only when, a staff helps readers to find and use the books."¹ What does a trinity mean? It means that every change occurring to one of its elements implies necessary changes in the other two. A glance at our current scenario reveals deep changes in readers’ habits, in the idea and form of “book” collections, and in the way library services are put into effect. Any change in a library should begin by fulfilling the fifth law of library science: “A library is a growing organism.”²

So the core issue is: are libraries respecting the requirements of the fifth law of library science? Are changes in catalogs and search tools coherent with changes in users’ habits and in collections?

**Changes in Readers’ Habits and Expectations**

At the moment, readers seem to be the most important issue in the trinity; not only since it is for their benefit that a library exists, but because they seem to live more and more distant from libraries. They have moved onto the Internet, where they try to satisfy their informational needs as far as possible. In users’ habits

Search engines continue to dominate, topping the list of electronic sources most used to find online content (93%), followed closely by Wikipedia (88%). … Results show a decline in use of library Web sites, electronic journals and online databases since 2005.³

Most users (83%) begin their information searches from a search engine, even if a relevant and growing number of them (43%) feel that libraries are more trustworthy.⁴ In fact, while users describe trustworthiness and accuracy as the most relevant criteria for selecting sources, their behavior shows that the criteria of speed and convenience are more often applied. The fourth law of library science—Save the time of the reader⁵—seems to be more relevant today than in the past. In the chapter devoted to this fourth law in *The Five Laws of Library Science*, Ranganathan discussed catalogs and bibliographies extensively as bibliographic tools that are created to save the time of the reader. They must be faster, smarter, and richer.

Users need answers to two different kinds of searches:

1. *known item search*, where the user is familiar with keywords or relevant search terms, has specific information in mind, and knows where to start; and

2. *exploratory search*, where the user lacks familiarity with the subject, and needs guidance and general information.⁶
This approach was true in traditional library services, and it is still true in the digital environment. In addition, the classified organization of physical collections in libraries was able to also satisfy a third user need: serendipity. This approach enables the user to discover relevant resources physically shelved next to others they may discover.

To support all three approaches, libraries must be aware that users more and more are changing from local users to remote users. While this makes a serendipitous approach more difficult but not impossible to provide, if classified arrangements can be a “view” available to the remote user through the library system or catalog.

**CHANGES IN THE BIBLIOGRAPHIC UNIVERSE**

In the bibliographic universe, the nature, cycle of production, distribution, and fruition of publications is undergoing profound change in order to allow books to find their readers.

Bibliographic information has been a main goal of the library community for a long time, as it is a way to grant access to any recorded knowledge. Catalogs and bibliographies—even if different in function (the former being an index of books of a collection, the latter being a list of books sharing one or more characteristic)—were both means to realize Universal Bibliographic Control (UBC), which is based on the objective of

promoting a world-wide system for control and exchange of bibliographic information. The purpose of the system is to make universally and promptly available, in a form which is internationally acceptable, basic bibliographic data on all publications in all countries.8

UBC must be based on internationally adopted standards. Standards initially agreed on in the 1960s, such as MARC and International Standard for Bibliographic Description (ISBD), were questioned in the 1990s, when libraries were forced to take into account new developments in cataloging theory such as Functional Requirements for Bibliographic Records (FRBR)9 and later the International Cataloguing Principles (ICP),10 and in the bibliographic universe, which changed significantly with the introduction of electronic resources and later of the web.

Electronic resources eliminated one of the basic concepts of a book: its physical aspect, its concreteness. At the same time, electronic resources opened the way for a text to be created and/or distributed in different formats and carriers to contain equivalent content, but be useable by a large number of devices.

Books are trying to find their readers in many ways. A dematerialization of the text is in operation to meet users’ needs. The same text is produced and
published in many formats, suitable for different devices, to grant it the most widespread availability. A text can be available immediately for publication virtually; this influences the meaning of concepts relating to its bibliographic nature, such as “book” and “chapter” whose distinction was based on length, often measured by a specific carrier (e.g., number of pages).

Also the idea of a journal is touched by the dematerialization of the text; newspapers and journals were created to spread short and timely news, in the form of articles. After peer review and the editorial process, the content of articles can be made available rapidly by web technologies (e.g., Open Access repositories, and services like “online first”). An online first paper is a paper published before, and for this reason without, the context of a journal issue. An institutional repository is a database of articles and other bibliographic resources, including those recently published. So, apparently there is no longer a reason for maintaining “physically based” concepts as journals made of physical issues and volumes, or for a printed copy, and so apparently there is no reason for waiting for the completion of a structure in the form of issues of a journal.

What does “collection” mean, in this context? To grant users efficient access to resources in print, a shift from traditional to collective collections is needed. The collection has changed in focus, boundaries, and value. This evolution raises some questions. For example, when printed materials are required, how many libraries will make the physical item available for interlibrary loan? What number of libraries is necessary to assure efficient interlibrary loan?

This change questions whether catalogs are really able to represent such diffuse collections. So, just as with the issue of reader’s habits, the focus of changes in the bibliographic universe is on the catalog, too.

Even if changes to the nature of resources are of major relevance, the main change to the bibliographic universe is happening in the web. Outside the library community, the world of the web faced a deep change, shifting from a web of documents to the web of data, or Semantic Web. This evolution modified the people’s habits, including library users and their expectations of library services. The Semantic Web requires that libraries stay relevant in this new environment by profoundly changing their tools and services.

**CHANGES IN BIBLIOGRAPHIC TOOLS**

Traditional online public access catalogs (OPACs) evolved into OPACs 2.0 in two ways:

- extending the usefulness and search features of the catalogue by harnessing more bibliographic MARC and circulation data for searching, and
seamlessly incorporating data from other resources; social networking with personalization and user community tagging and reviewing to provide a richer discovery experience.\(^{13}\)

While users and collections changed and moved to the web, catalogs were not so quick; the result is a lack of balance in the library trinity. A distinction is needed between “catalog” in the strict sense and OPAC with its user friendly interface. The distinction is useful because it allows us to distinguish between two different lines of evolution in our tools: those relating to the catalog in the strict sense, and those relating to the OPAC.

During the last few decades, OPACs evolved considerably, moving from poor and scarcely engaging interfaces, often unable to answer both exploratory and known item searches, to more developed tools, enriched by many capabilities. One important development is linked to the Web 2.0 environment. Web 2.0 is characterized by social networking and participation on the Internet by user community involvement through tagging, blogs, or wikis, and by syndicated feeds or alerts.

Another line of evolution is defined by the increase in content offered to users. Traditionally in libraries there were as many catalogs as there were kinds of available resources. This approach forced users to search in each repository in order to get the complete range of answers. During the last decade, tools have been created to provide solutions, such as the integration of searching options relating to all types of information sources. This is the case, for example, of Primo, which offers a true one-stop shop for discovery and delivery, branded and customized to the individual institution’s needs, with a choice of a local or cloud-based implementation. With its impartial content coverage, advanced relevance-ranking algorithms—configurable by the institution—and groupings of similar search results, Primo eases information overload, helping users focus on the most relevant materials that meet their needs.\(^{14}\)

A “catalog” in a strict sense is the logical and physical structure of data recorded in an Integrated Library System (ILS). First catalogs were created to answer known item searches, as defined by Cutter’s objectives.\(^{15}\) However, they were largely unable to answer to exploratory searches, which still required the intervention of a reference librarian. As Ranganathan wrote, “The library catalogue presents a bundle of conventions. It is even treacherous. For, it appears to be in a familiar, natural language. But, in reality, the language of the catalogue is an artificial one.”\(^{16}\)

This is the reason why reference service was of the utmost importance in Ranganathan’s view; it was the main aid to help users in performing their exploratory searches in libraries.
Nevertheless, almost worldwide, the logical and physical structure of cataloging data has been built on Cutter's objectives and later on the *Paris Principles*. Interoperability among such cataloging data was based on updating descriptive standards and, above all, on the data exchange format that emerged from the Library of Congress–led initiative in the 1960s: MARC. MARC functioned very well for a long time, but the will to respect interoperability with MARC format (and later with MARC 21) stopped development on the logical structure of the catalog. In the meantime, users and collections changed and the result was a lack of balance in the library trinity. In an issue dedicated to the evolution of bibliographic data exchange, Ted Fons sums up the current context of the bibliographic format:

As librarians have increasingly professionalized and improved the core mandates of selection, acquisition, preservation, and description of library collections, there has been a corresponding fracturing and loss of effectiveness in another of our responsibilities: exposure. The user has generally moved away from the library catalog as the tool used early in the research process—it is now used, if at all, as a source for availability or fulfillment in the last mile of the research process. A companion theme ... is the widespread recognition that our current model for data exchange between library organizations has outlived its usefulness and is ripe for replacement with something with lower barriers to entry for library developers and partners. ... The key word ... is “effective.” Our goal should be to find methods that will maximize the full disclosure of unique and commodity library collections on the web. That includes taking risks with the formats and methods that the web search engines prefer.17

The 1960s approach to data format was based on a one-size-fits-all schema, and this proved to be wrong for the long term; as Gildas Illien, Director of the Bibliographic and Digital Information Department, Bibliothèque nationale de France, noted:

... we have experienced the limitations of trying to answer all functional and community requirements with a single format or implementation scheme. One size can’t fit all and doesn’t need to. ... I would say we are ideally looking for a scenario where we could meet the joint requirements of:

a) internal metadata management, including the management of legacy data not only for descriptive purposes, but also for digitization, rights management, and long term preservation of collections;

b) rich bibliographic data exchange services with no loss of granularity in description; and
c) standard data exchange and exposure on the web the people and search engines use.\textsuperscript{18}

So, a new approach is needed that might be termed an all-sizes-fit-one approach, as shown by the example of the Virtual International Authority File (VIAF).\textsuperscript{19} VIAF allows any form of name for the same entity produced by any national authority to fit one virtual record, identified by VIAF identifier. But this approach also requires a completely new data format. The library community will need to manage multiple exchange models and the next step is to develop these models in an orderly and efficient way.\textsuperscript{20} In which direction should we move?

**CHANGES IN BIBLIOGRAPHIC STANDARDS**

Many changes have occurred in the field of bibliographic standards; they are all well exemplified by the aims of *Resource Description and Access* (RDA). This new standard allows us to understand what we presently call cataloging in a deeply different way.

Innovations in RDA are many, and in many directions. The user is the focus of data produced by RDA; this simple assertion is easy to link to Ranganathan’s first law—Books are for use—and it is clearly in the direction of balancing the library trinity. Its effects are that bibliographic tools are moving where users are: on the web.

RDA stems from FRBR and ICP, and it aims to be a content standard, not a display standard nor an encoding standard. This approach means to distinguish two aspects traditionally seen as the same: (1) which data are recorded and (2) in which form and order data are presented. RDA provides for the first aspect, leaving the solution of the second to different technological environments and use contexts (e.g., defined by data producers, user services).

To identify and to relate are basic goals of RDA. Identifying each entity described in FRBR increases data granularity; relating granular data among them is necessary to create different kinds of descriptions starting from the same set of data (e.g., comprehensive, analytical, hierarchical, and composite descriptions). RDA devotes a large part of its guidelines to relationships to allow user navigation across the bibliographic universe, as described by Elaine Svenonius\textsuperscript{21} and required by ICP. As suggested by Barbara Tillett, RDA is intended to enable the creation of well-formed metadata about resources that can be used in any environment—whether a card-based catalog, an online catalog, or a web-based interactive resource discovery system.\textsuperscript{22}

The journey from a record-centered to a data-centered approach and the large use of vocabularies and relationships are fundamental features of
RDA, just as they are basic premises for integration of bibliographic data in the Semantic Web.

RDA inherits and develops two basic qualities of the Anglo-American Cataloguing Rules, Second Edition (AACR2): (1) the aim for worldwide diffusion and (2) a global approach to resources. This approach is evident from the title of RDA, in which the words “cataloguing” and “Anglo-American” were deleted because they were too tied to a narrow cultural context and to bibliographic tools and resources. The focus is instead on data, worldwide and reusable.

The user focus, granular data, navigability, worldwide diffusion, and global approach are aims fully oriented to the Semantic Web and necessary to counterbalance the gap between users and libraries tools.

**CHANGES IN BIBLIOGRAPHIC FORMAT**

The Bibliographic Framework Initiative (BIBFRAME) is a good example that shows us the course to be followed. BIBFRAME is a community effort, led by the Library of Congress, to start a transition from the MARC 21 communication format to a new Resource Description Framework (RDF)–based data model that embraces linked data practices in support of sharing and publishing metadata.

The initiative is presented as “Bibliographic Framework as a Linked Data Model” and its goal is to provide a pattern for modeling both future resources and bibliographic assets traditionally encoded in MARC 21. The main purpose of the initiative is to replace MARC 21, not only as an exchange format, but also as a cataloging format and as the internal format of integrated library systems.

Furthermore, the BIBFRAME report states that rather than a mere replacement for the library community’s current model/format, MARC, the new model is the foundation for the future of bibliographic description that happens on, in, and as part of the web and the networked world we live in. An interesting feature of BIBFRAME is its objective to be both “rule agnostic” and “model agnostic.” As Svensson noted:

Currently, however, the main discussion in the library community seems to focus more on the formats (e.g., MARC 21) than on an underlying model that can be expressed/serialized in different ways. This focus on the format is insofar counter-productive in that it tends to encourage the use of literals (strings) without analyzing what the information is about and how it relates to other pieces of information (things)—within or outside of a specific bibliographic description. Further, the preoccupation with data in the context of a particular format tends to prevent real innovation, since it is more focused on carrying the existing data forward than on analyzing which data would be necessary for what operation.
BIBFRAME is designed to find new ways to (1) differentiate clearly between conceptual content and its physical/digital manifestation(s), (2) unambiguously identify information entities (e.g., authorities), and (3) leverage and expose relationships between and among entities.27

The basic idea of the BIBFRAME project is that information recorded in MARC records can be referred to in three classes, with reference to which entity they convey information about. Data can be related to the intellectual essence of a work; to the actual instance of the work—that is, what a cataloger holds in his/her hand; and to the record (i.e., metadata such as control numbers, record handling codes, and other annotations).

From these classes, a new logic model was developed, consisting of the following main classes:

- **Creative Work**—a resource reflecting a conceptual essence of the cataloging item.
- **Instance**—a resource reflecting an individual, material embodiment of the Work.
- **Authority**—a resource reflecting key authority concepts that have defined relationships reflected in the Work and Instance. Examples of Authority Resources include People, Places, Topics, and Organizations.
- **Annotation**—a resource that decorates other BIBFRAME resources with additional information. Examples of such annotations include Library Holdings information, cover art, and reviews.28

Many demonstration datasets are available from the Library of Congress, the British Library, Deutsche National Bibliothek, and other institutions to show how MARC/EXtensible Markup Language (XML) records can be transformed to the BIBFRAME data model (http://bibframe.org/demos/).

BIBFRAME is more oriented to the Semantic Web and linked data than any previous format, as it relies on relationships between resources (Work-to-Work relationships; Work-to-Instance relationships; Work-to-Authority relationships; http://bibframe.org/), while the MARC format is focused on catalog records able to fully describe single entities (e.g., manifestations and agents). The BIBFRAME approach is more oriented to reuse data and to decrease redundancy.

In Ranganathan’s approach, work and instance would not be sufficient to properly describe bibliographic phenomena, as he organized documents conceptually in three levels, not two, based on Indian culture: work, expression, and physical medium.29 Compared with Ranganathan’s approach, BIBFRAME is a simplification, but the main point is that it is an expression of the willingness of libraries to achieve near-universal adoption of data exchange standards,30 and to respond to Semantic Web requirements.
CONCLUSIONS

Many phenomena are influencing one or more aspects of Ranganathan’s library trinity and they all press our catalogs for change: readers, collections, and services.

As to readers, they moved from libraries to the web and require more and more emphasis on fulfillment of the law “Save the time of the reader.” They are used to navigating, and want full navigation capabilities, such as following links from a resource description to its full text, and also links among similar resources. Above all, readers need links among concepts, to satisfy their need of knowledge in exploratory researches, and their unexpressed needs by serendipity.

As to collections, we note that the bibliographic universe is fundamentally changing the way single resources are produced, disseminated, found, and obtained, and this means we need to also change the way we represent them. Collections are changing in the way they can be accessed by users and created and controlled by libraries. Management of library collections is changing from the single library approach to a system-wide view of library collections. Lastly, content available online is changing in openness, both as documents in the Open Access context, and as data in the Linked Open Data (LOD) environment.

As to services, the move of data from bibliographic silos to the Semantic Web requires changes in bibliographic models and in the relationships between bibliographic entities and real world object entities described in the Semantic Web. The reuse of millions of MARC records affects the search for a new bibliographic format that must be compliant with past and future records, while data are moving from single database management systems (i.e., silos) to the Giant Global Graph envisioned by Tim Berners-Lee. It must be underscored that characteristics of new guidelines, such as RDA, point toward a growing global approach to the activity of cataloging and are leading to worldwide diffusion of their application.

Catalogs are in the middle of this development but are not yet at a turning point because, as we have seen, they are affected by all these changes that are underway. However, they are approaching a turning point; if catalogs do not change, they risk becoming further removed from readers and collections. Soon catalogs will “change or perish”; they must become an integral part of the Semantic Web, and they must restore the balance of Ranganathan’s library trinity.

NOTES


4. Ibid., 54.


7. By “fruition” we mean the changing ways that readers are using or consuming resources.


28. Ibid., 8.
