

Universal bibliographic control in the digital ecosystem: opportunities and challenges

Mauro Guerrini^(a)

a) Università degli Studi di Firenze

Contact: Mauro Guerrini, mauro.guerrini@unifi.it

Received: 16 April 2021; **Accepted:** 26 August 2021; **First Published:** 15 January 2022

ABSTRACT

The idea of universal bibliographic control (UBC) has been of interest for centuries in the history of cataloguing and is based on the humanistic ideal of sharing recorded knowledge produced anywhere in the world. In the contemporary era, IFLA has played a central role, stimulating national bibliographic agencies and other institutions to promote standards and collaborations that go beyond the national sphere, leading to multicenter and even more cooperative bibliographic control. The tradition of cataloguing also grows and is enriched by the dialogue with different communities and users' groups. The free reuse of data can take place in contexts very different from the original ones, multiplying for all the opportunities for universal access and the production of new knowledge: the UBC, therefore, looks at interoperability and flexibility in the dialogue with the various communities of stakeholders and with the cultural institutions.

KEYWORDS

Universal Bibliographic Control; UBC; Cataloging.

Culture is the only asset of humanity that, when divided between us all, becomes greater rather than smaller.

Hans-Georg Gadamer

As a “non-commercial public space” (IFLA Global Vision) – not only in a literal sense – libraries play a fundamental role also in the digital ecosystem

Conference BC2021

Bibliographic control: a central topic in LIS

The idea of universal bibliographic control has been of interest for centuries in the history of cataloguing, and it is based on the humanistic ideal of sharing collective knowledge in every part of the world. It probably began with Conrad Gesner’s *Bibliotheca Universalis* (1545–1549), the catalog of all printed books published up to that time in Latin, Greek, and Hebrew. Gesner called ‘Universalis’ his work, pursuing the goal of maximum bibliographic coverage in relation to the concrete literary reality of his time. His universal bibliography included a catalog for authors’ names, and a catalog for general as well as specific subjects (*loci*). Gesner established the connotations of the scientific and literary heritage and established the characteristics of indexing logic using four categorical levels: author, work, text, and edition.¹

In the contemporary era, IFLA has played a central role in the realm of Universal Bibliographic Control (UBC) by bringing together national bibliographic agencies and other institutions to promote standards and collaborations in this area. This also includes the work of promoting conferences and publishing texts and documents.² From 1990 through the 1st of March 2003, the Deutsche Bibliothek hosted the IFLA Universal Bibliographic Control and International MARC Core Activity (UBCIM),³ demonstrating the direct connection between UBC and technologies. For years IFLA has edited “IFLA Series on Bibliographic Control”. In particular, one book in that series entitled “National Bibliographies in the Digital Age: Guidance and New Directions”, edited by Maja Žumer in 2009,⁴ continues to be a fundamental reference text. A statement reaffirming IFLA’s commitment to UBC was endorsed by the Professional Committee in December 2012. Initiated by the Bibliography Section, that statement was also supported by the Cataloguing Section and the Classification and Indexing Section.⁵ The WLIC of Lyon in 2014, included in the programme a seminar entitled “Universal Bibliographic Control in the Digital Age: Golden Opportunity or Paradise Lost?”⁶ It was planned by the Cataloguing Section, with the Bibliography Section, the Classification Section, and the UNIMARC Strategic Programme.

¹ (Sabba 2012).

² (Anderson 1974); (Davinson 1975).

³ <https://archive.ifla.org/ubcim/>.

⁴ (IFLA 2009).

⁵ <https://www.ifla.org/publications/node/7468>.

⁶ Monday, 18 August 2014; see Session 86, <http://library.ifla.org/id/eprint/817/>.

Also, back in 2001, the Library of Congress organized the “Conference on Bibliographic Control for the New Millennium”,⁷ celebrating a significant anniversary precisely with this theme. The Library of Congress established an independent Working Group on the Future of Bibliographic Control that published the report entitled “On the record” in 2008.⁸

As we can see from these recent events, bibliographic control is central to the history of cataloguing and to the history of libraries themselves.

The concept of Bibliographic Control has changed and still changing radically, because the bibliographic universe and technologies are radically changed; and resources, actors, standards, and practices will presumably change further. It necessary, therefore, to explore the new boundaries of bibliographic control, in fact, the digital ecosystem.

Text and metadata as paradigm of bibliographic control

For centuries, a text (whether manuscript or printed) was identified by the physical volume. Today, ‘work’ is at the center, and increasingly its content can be presented and enjoyed in many forms. For example, a reader can choose between paper and e-books, based on his or her reading preferences. This content is now usually accompanied by a set of metadata. Metadata has become the protagonist of communication on the web; metadata is today the paradigm of bibliographic control. Some of the consequences are already evident. For example, the quality metadata of a resource contribute to its knowledge, enhancement, and success.⁹

The process of metadata creation for bibliographic resources starts with the creators of those resources – obviously providing the content –, and, in the modern era, usually providing the title, and some basic metadata; then, the publishers add their metadata, including some standard identifiers, an important step in the bibliographic control in the digital ecosystem. The process of metadata creation continues through the intellectual contribution of the cataloguers of the bibliographic agencies. Considerable is the initial investment in the creation of metadata based on authoritative sources.¹⁰

From the model of universal bibliographic control based on the centrality and exclusivity of the national bibliographic agencies, we are moving on to dynamic and shared bibliographic control. In the digital world, this is configured as a process of data reuse and enrichment, linking single data elements. In an evolving ecosystem, the international dimension is the virtual space where stakeholders meet. In this context, libraries, and in particular, the national libraries, no longer have the monopoly of bibliographic control. This poses an intellectual and operational challenge to library institutions. However, libraries, library networks and bibliographic agencies still play an important role, in particular, through strong collaborations among themselves, through their role as true protagonists of the standards of bibliographic control, standards flexible and at the same time binding and reliable. Still, libraries remain an essential part of the digital ecosystem.

⁷ Library of Congress, “Proceedings of the Bicentennial Conference on Bibliographic Control for the New Millennium” <https://www.loc.gov/catdir/bibcontrol/>.

⁸ <https://www.loc.gov/bibliographic-future/>.

⁹ (Guatelli 2020).

¹⁰ As an added aspect, metadata can serve as an antidote to even fake news; cfr. (Bredemeier 2019, 384 and so on).

What are the consequences of digital transformation for library catalogues, and work processes in metadata creation? What is the function of repositioned and reconfigured catalogues on the web? Understanding how texts are conveyed today requires cultural awareness and professional training: this is the basis of the process of literary and conceptual analyzing the resource. These two aspects – awareness and training – should be common to the training of other actors involved in the process, who serve as mediators of the knowledge process.

Beyond tradition

The data models and the semantic web paradigm invite us to go beyond that aspect of the cataloging tradition that entrusted only the bibliographic agencies with the role of authoritative producers of quality registration. Data models and the semantic web paradigm invite us to go beyond the cataloging tradition. That tradition provided for homogeneous descriptions for all the libraries. The contemporary perspective foresees the participation of new and different actors. In addition to libraries and librarians, other institutions (publishers, distributors, private agencies, universities), and professionals (archivists, museum professionals) are contributing to the recording and enrichment of metadata and authority files. In those context, libraries still play the role of intermediary with the other major producers of metadata. The participation of several actors is very positive, and everyone is invited to find a new balance between their different methodological and cultural traditions to pursue a common goal: the cooperative editing of quality metadata, possibly in open access. The best cataloging tradition in the completely new collaborative context is therefore maintained and indeed enhanced.

Another consequence is that the relationship between libraries, publishers and distributors becomes more strategic, because the publishers are the first, after the creators themselves (in the modern era), who should create the metadata of a resource, and later, that metadata is enhanced by libraries for the part that concerns libraries. Libraries feel, with particular responsibility, the issue of the shared construction of quality data, by virtue of the principles of precision, accuracy, and social sharing of the cultural heritage that have characterized their history.

Bibliographic control today is, therefore, multicentric, and even more cooperative than in the past. National bibliographic agencies maintain and reinforce their role in quality control of metadata and authority control, through the maintenance of fundamental tools, such as VIAF (Virtual International Authority File) and through support of international identifiers such as ISBN (International Standard Book Number), ISSN (International Standard Serial Number) and ISNI (International Standard Name Identifier), that are part of broader international cooperation and authority control projects.

VIAF and ISNI are different projects: VIAF is an international collaboration that supports a shared authority file; ISNI is a name identifier and a system for recording those numbers that define it. VIAF, in particular, provides authoritative services that reliably identify agents, places etc., and the works associated with them in the global registered knowledge network. Its philosophy is inspired by promoting all cultural perspectives equally, including all languages and scripts, and simplifying the work of bibliographic agencies and libraries. Many libraries and bibliographic agencies collaborate in sustaining these authoritative resources for the benefit of users everywhere.

The greater the accuracy of the data, the greater the benefits of using those authoritative sources. By aggregating and linking data, these sources for authority control can bring greater interoperability to the galleries, library, archival, and museum community (GLAM) as well as the publishing and book dealership industries.

The form of the name: conditioned by the cultural and linguistic context

The choice of form of a name associated to an entity is always culturally founded, but the selection of the preferred form of a name is, in many cases, complex, and depends upon the cultural and linguistic context in which that name is used. In the past, the bibliographic traditions of the Western world were privileged, but now the global dimension of communication changes all parameters. In the global cultural environment (as opposed to a single library's catalogue), there has been the important acknowledgement that there is no single form of an author's name that must be used by everyone. The choice of the form of a name to be displayed is conditioned by the cultural and linguistic context within which the dataset for that name is placed. IFLA LRM recalls that a named entity can have different *nomen*, all valid (e.g., Léonard de Vinci in France and Leonardo da Vinci in Italy; Cicero in a specializing library in Latin literature and Cicerone in a public library). The goal is to overcome the geography and dominance of a cultural area, and to respect the cultural and linguistic traditions of each Country, and of each individual cultural community in the solutions adopted.

The mechanism of "reconciliation" of the different forms with which an entity is known and identified in a global context (for example, the creator of a work), brought together in a group of variant forms, all recognized, becomes the principle for new ways of sharing information. The entity reconciliation process produces a cluster: it is a grouping of the different variant forms referable to the same entity; this entity is known in various *nomen* in different cultural, linguistic, geographical, domain contexts; all valid, usable and actually used variants. Linking various identifiers is of strategic importance. In all entity identification projects that make use of the reconciliation (or clustering) mechanism, it is customary to assign an identification to the recognized entity; identifier that connects to other identifiers assigned to the same entity in different contexts, and all valid. The clustering mechanism starts from the assumption that all forms of a name used in the global context have equal dignity; there is no particular preference for one or the other form. The context of belonging (the source from which that variant form of the name comes) and the need for use (the target that recalls that name) define each time the choice of the form to be considered the preferred "conditioned" form of the name. This is motivated by the desire to enrich the dataset, and to offer the reader as many channels as possible to reach the goal; this is the pragmatic and functional purpose of being able to identify, select and obtain the resource. The identifiers allow both the explication of the equivalence function of the forms of the cluster and the connection of the cluster to other clusters relating to the same entity. The choice of the preferred form of the name, the structuring of the string (according to syntactic rules known in the past only to catalogers), lose importance in the face of the practical need to create multiple and equivalent retrieval channels for the same resource. In the context of Universal Bibliographic Control, there remains the need to offer a form as a result of a national or cultural or linguistic choice; this is also achieved

through information presentation mechanisms linked to the cluster: the data on the “provenance” of the information (given on the source that produced the information) can be used in a double meaning:

- the more traditional: source that generated the information and that defines, within a cluster, which form is to be presented as preferred in a given context;
- that of the applicant target (Provenance of the applicant) which, on the basis of its own specific research need, guides the selection of the preferred form (also in this case, therefore, preferred in the specific context of the research).

Therefore, the cluster of variant forms is fundamental passage from Bibliographic Control intended as control of strings and access points to the more complex concept of entity identification, through different and variant identities with which it can be expressed. The choice of the form of the name and the linking of variants in clusters enhances the concept of universal bibliographic control that respects cultural variations for the display of names.

The tradition of cataloguing grows and enriches in dialogue with different communities and groups of users. The free reuse of data can take place in very different contexts from the original ones, multiplying for all the opportunities for universal access and for the production of new knowledge. The concept of cultural heritage values is a living idea.

The great changes brought on by the use of metadata have led to new perspectives on bibliographic control. UBC now contemplates interoperability and flexibility in dialogue with the various communities and with institutions of registered memory.

Who knows what the future will bring us? Perhaps, we are still at the beginning of the digital revolution. Precisely in the field of metadata and authority control, we could expect developments and surprises from alternative technologies on machine learning or artificial intelligence, a tool that promises to be very useful; a tool that takes nothing away from the cataloguer’s judgment, which remains a fundamental intellectual activity.

References

Anderson, Dorothy. 1974. *Universal Bibliographic Control: A long term policy, a plan for action*. Pullach/München: Verlag Dokumentation.

Bredemeier, Willi. 2019. *Zukunft der Informationswissenschaften: Hat die Informationswissenschaft eine Zukunft? Grundlagen und Perspektiven. Angebote in der Lehre. An den Fronten der Informationswissenschaft*. Simon Verlag für Bibliothekswissen.

Davinson, Donald Edward. 1975. *Bibliographic Control*. London: C. Bingley.

Guatelli, Fulvio. 2020. «FUP Scientific Cloud e l'editoria fatta dagli studiosi». *Società e storia* 167. doi:10.3280/SS2020167008.

<https://archive.ifla.org/ubcim/>.

<http://library.ifla.org/id/eprint/817/>.

<https://www.ifla.org/publications/node/7468>.

<https://www.loc.gov/bibliographic-future/>.

IFLA. 2009. *National Bibliographies in the Digital Age: Guidance and New Directions*. A cura di Maja Žumer. IFLA Series on Bibliographic Control 39. München: K.G. Saur.

Library of Congress, "Proceedings of the Bicentennial Conference on Bibliographic Control for the New Millennium" <<https://www.loc.gov/catdir/bibcontrol/>>.

Sabba, Fiammetta. 2012. *La 'Bibliotheca Universalis' di Conrad Gesner, monumento della cultura europea*. Roma: Bulzoni.