Knowl. Org. 50(2023)No.4
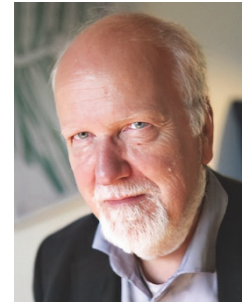B. Hjørland. Bibliographical Control

301

# Bibliographical Control

## Birger Hjørland

University of Copenhagen, Faculty of Humanities, Department of Communication, South Campus, building 14, 3rd. Floor, Karen Blixens Plads 8, 2300 Copenhagen S, Denmark, <birger.hjorland@hum.ku.dk>

Birger Hjørland holds an MA in psychology and PhD in library and information science. Since March 2020, he has been Professor Emeritus at the Department of Communication, University of Copenhagen and formerly professor in knowledge organization at the Royal School of Library and Information Science/Department of Information Studies (2001-2020) and at the University College in Borås (2000-2001). He is a member of the editorial boards of Knowledge Organization, Journal of the Association for Information Science and Technology and Journal of Documentation. His h-index on 2023-07-02 is 51 in Google Scholar and 35 in Web of Science.

**Abstract**: Section 1 of this article discusses the concept of bibliographical control and makes a distinction between this term, "bibliographical description," and related terms, which are often confused in the literature. It further discusses the function of bibliographical control and criticizes Patrick Wilson's distinction between "exploitative control" and "descriptive control." Section 2 presents projects for establishing bibliographic control from the *Library of Alexandria* to the Internet and Google, and it is found that these projects have often been dominated by a positivist dream to make all information in the world available to everybody. Section 3 discusses the theoretical problems of providing comprehensive coverage and retrieving documents represented in databases and argues that 100% coverage and retrievability is an unobtainable ideal. It is shown that bibliographical control has been taken very seriously in the field of medicine, where knowledge of the most important findings is of utmost importance. In principle, it is equally important in all other domains. The conclusion states that the alternative to a positivist dream of complete bibliographic control is a pragmatic philosophy aiming at optimizing bibliographic control supporting specific activities, perspectives, and interests.

† Derived from the article titled "Bibliographical Control" in the *ISKO Encyclopedia of Knowledge Organization*, Version 1.0 published 2023-03-28.

## 1.0 Introduction

The present article intends to clarify the concept of bibliographical control, illuminate its main functions, briefly describe its history in practice, and survey the criticism, debate, and different theoretical views associated with this practice. An underlying perspective throughout the article is to illuminate two conflicting philosophical approaches: "positivism" and "pragmatism," and their implicit influence on theory as well as practice.

### 1.1 The concept and idea of bibliographical control

Bibliographical control means making useful documents/information resources findable for those who might need them, but have no, or insufficient, knowledge about their identity. The present article is based on a view that comes close to a definition provided by Pearson (2010, 523):

> Bibliographical control. The concept of having a comprehensive and searchable listing of the entire published output in a particular field [or in a particular country or by a particular author or other formal demarcations].

> The term is today often used as a synonym for information organization (or bibliographical organization or knowledge organization, KO) and for the process of describing information resources (documents) and providing name, title, and subject access to the descriptions (Chan and Salaba 2016, 13-14; Joudrey, Taylor and Wisser 2018, 7; Library of Congress 2008; [1] Maxwell 2017; [2] Miksa 2013; Wikipedia 2023.[3])

302

Knowl. Org. 50(2023)No.4
B. Hjørland. Bibliographical Control

This is, however, an unfortunate way of definition the term as it focuses more on document description (or cataloging or organization) and less on the aim of providing control over available documents. As Wilson (1968, 3) wrote:

[W]hat we see as problems of control, others have seen as problems of organization [...] But not all the problems of bibliographical organization are equally important in a discussion of bibliographical control, and many discussions in the enormous literature of bibliographical organization will go unnoticed here.

Therefore, this article focuses on the aim of providing control over documents, which has much to do with the coverage of bibliographic databases. It will not go deeply into the issues of cataloging rules, document description standards, metadata, and the like, although this use of the term bibliographic control deviates somewhat from how the term is mostly used today. The two issues are related, and as we shall see later, a database like Google Scholar may have better coverage than MEDLINE, but may still be less useful, because documents are less well described and represented in Google Scholar.

What is bibliographic control, and what is the purpose of obtaining it? Wilson (1968, 4) also illuminated this, writing:

Bibliographical control is a form of power, and if knowledge itself is a form of power, as the familiar slogan claims,[4] bibliographical control is in a certain sense power over power, power to obtain the knowledge recorded in written form.

When academic students, professionals or researchers write papers, they are expected to build their papers on information in other papers, and to provide references to them. To do so, they need to know about the relevant sources, and ideally, they need to identify what Wilson (1968, 22) called "the best textual means to one's end," (i.e., the most important sources). In academia, this is almost a matter of course, but also for non-academic tasks available knowledge should be utilized, and thus also often depend on bibliographic control. The need is, however, most directly and explicit related to academic writing tasks.

Because relevant documents need not be textual, we rephrase Wilson's goal to "the best documents or the best information resources to one's end." As discussed further in this article, no library or information system can guarantee such an ideal performance, but they can serve users approaching this goal, and this is what bibliographical control is about: to help users approach the goal of finding the documents most relevant for solving their task.

Bibliographical control is often understood as control over what has been published, a control that makes it possible, in principle to overview what has been published. It may also be about unpublished documents, for example, unpublished theses and other so-called "gray literature." A new tendency is to include "data" (i.e., data documents)[5] . Higgins and Green (2009 and later editions) explicitly recommend including "gray literature" and unpublished clinical trials in searches for systematic reviews. In information science the term "document" includes museum objects and archival records, but these kinds of documents are in practice seldom (if ever) included in the literature about bibliographic control, although theoretically they might be. Bibliographic control is to make it possible to identify the documents of relevance for a given query. KO, cataloging, and document descriptions are tools for providing bibliographical control, but not in themselves bibliographical control. For some domains, it is of utmost importance to ensure bibliographical control, and especially in medicine, this is taken very seriously, with much research about the coverage and quality of different databases, as we shall see later in this article. But such control should, in principle, be equally relevant for all domains.

If one library in the world—whether a physical or a digital library—could have a copy of all documents ever produced, that library's catalog could be said to provide a kind of universal bibliographical control.[6] *The Library of Alexandria*, for example, seems to have had this ambition. There are alternative ways in which universal bibliographical control has been approached. They include attempts to make comprehensive universal bibliographies, such as *Bibliotheca Universalis*, the *Universal Bibliographical Repertory*, *Union catalogs* with *WorldCat*, and the *World Wide Web* with *Google*. A bibliography/ bibliographical database makes it possible to identify relevant documents but often leaves the task of locating those documents to libraries, digital archives, or publishers to library catalogs or related tools.

An alternative to one comprehensive bibliography represents IFLA's and UNESCO's program on universal bibliographical control to establish a system of national bibliographies that together could provide universal control. In practice, however, the most useful system has been the abstract journal and subject bibliographical databases, although this system is dependent on and varies in coverage and quality from discipline to discipline.

### 1.1 Two kinds of bibliographical control

Wilson (1968) made a distinction between two kinds of bibliographical control (1) "exploitative control" (2) "descriptive control":

Knowl. Org. 50(2023)No.4
B. Hjørland. Bibliographical Control

303

(1)  If one is reading for a purpose (and is a rational person) then one would like the best textual means to one's end. Therefore "the more important sort of bibliographical control is this: to have the power to produce the best textual means to one's ends." (Wilson 1968, 22)

(2)  The second sort of power and bibliographic control is to be able to provide a list of documents (or bibliographic records) fitting a certain description, that is evaluatively neutral. Wilson's examples are: all writings authored by Hobbes, all writings discussing the doctrine of eternal recurrence, all writings containing the word 'fatuity".

Wilson (1968, 69) considers that "*The Bibliographic Encyclopedia*" resembles (1), while "*The Catalog*" resembles (2). He considers the exhaustive national or subject bibliography as the typical instrument of descriptive control, whereas the typical means of exploitative control would be "the 'special' library serving a group of scholars accustomed to talking to each other, and staffed by persons approaching the type of bibliographical consultants rather than the bibliographical aid" (Wilson 1968, 150).

It might be argued, however, that even the second sort of bibliographic control is not evaluative neutral (at least not always). It is, for example, not always clear who authored a given document. The present author is not familiar with this issue in relation to the works attributed to Hobbes, but if we change the example to writings by Plato, the question "who wrote Plato's works" has a long history and even gave birth to a hermeneutic tradition.[7] In the contemporary environment, we have problems with ghostwriters, pseudonyms, and many guidelines for assigning authorship to documents ethically [8] . Therefore, even attributing author names to documents may sometime involve disagreements among experts and be interpretations and evaluations, and thus not be "evaluative neutral." However, the issue sticks even deeper because it has been questioned that any form of description can be neutral. Bateson (1977, 147), for example, wrote:

> You can never get away from theories of the nature of description whenever, wherever you have descriptions. All descriptions are based on theories of how to make descriptions. You cannot claim to have no epistemology. Those who so claim have nothing but a bad epistemology. And every description is based upon, and contains implicitly, a theory of how to describe. The Cartesian coordinates contain a theory of how to describe, and for many purposes, I believe, it is an inappropriate and dangerous theory.

Wilson (1968, 32) acknowledges that "description" and "evaluating" are not absolute opposites:

> Insofar as we can distinguish between evaluating and ('neutrally') describing, some element of each enters into every remark we ever make about anything, that there are no pure cases of evaluation and none of description, but at best a continuum extending indefinitely in both directions, the end points being unreachable.

Still, however, it seems that Wilson's distinction between the two kinds of bibliographical descriptions is problematic because "descriptions" in science and scholarship are influenced by theoretical assumptions and are developed to make identifications and selections possible, and the bibliographical descriptions as an ideal should reflect scientific concepts to make the literature retrievable. Consider an example:

In evidence-based medicine (EBM), the goal is to identify the part of the literature that best provides "evidence" for whether a given treatment is better than another (or better than no treatment). This corresponds to Wilson's first kind of bibliographic power: To identify the best textual means to an end — the end here being choosing the optimal medical treatments. The tradition known as EBM has developed criteria to distinguish between degrees of "evidence" in the medical literature. For example, evidence from a randomized controlled trial (RCT) is assigned greater weight than evidence from case studies, which again are given greater weight than the clinical experience of respected authorities. Therefore, norms of medical writings demand that the kind of evidence claimed in a medical study is described in the published record of the study (e.g., a journal article). At the bibliographical level, the kind of evidence is described by the assignment of subject headings to the journal article by the National Library of Medicine staff in the MEDLINE database. The description of the study as it appears in MEDLINE, therefore, reflects Wilson's second kind of bibliographical control. Because these descriptions are not, and cannot be, neutral (Hjørland 2011), Wilson's distinction between two kinds of bibliographical control collapses.

The collapse of Wilson's distinction can also be observed in "best match techniques," the ranking of research results used in modern search engines. Here the documents which are predicted of highest relevance to the user are ranked highest and presented first. However, This ranking is based on what must be characterized as descriptive elements instead of appraisals of the documents.

Alternatively, a more common distinction between two kinds of bibliographical control is suggested here (previously suggested in Hjørland 1997, 20):

304

Knowl. Org. 50(2023)No.4
B. Hjørland. Bibliographical Control

(1) Subject attribution [9] (subject cataloging, traditionally understood as classification and indexing) is the kind of bibliographical control that support users' identification of unknown documents of potential interest to their queries. (e.g., find documents about the doctrine of eternal recurrence or find studies about the relevance of psychotherapy for the treatment of ADHD). Subject assignment requires adequate subject knowledge.

(2) So-called descriptive cataloging [10] is the kind of bibliographical control that enables users to identify known items of which some characteristics, such as author, title, or publisher, are known. [11] Such characteristics can be searched in databases and help users find a record that allows them to get a copy of the document (e.g., from a library or a publisher). Descriptive cataloging is less dependent on subject knowledge and more dependent of knowledge associated with descriptive bibliography and traditions of document description.

## 2.0 Important historical examples of bibliographical control

This section presents a range of projects from the Library of Alexandria to the Internet, which has, in different ways, provided comprehensive bibliographical control of books and other kinds of documents. The examples have been chosen from what this author considers the most prominent and theoretically important ones in the history of bibliographical control. Each project has fulfilled important goals, and most of them also seem to reveal a dream of an unobtainable goal: to provide complete control of "information." This presentation of former projects, the assumptions on which they were based, and the criticism and debate related to them, may hopefully provide a basis for considering future projects. They are described very briefly, and with only a few references, many of these projects deserve an independent article in this Encyclopedia.

### 2.1 *Library of Alexandria* (c 285- BC)

Erskine (1995, 39) [12] wrote about the aims and acquisitions of this library:

It is clear from our evidence, scrappy as it is, that the Ptolemies made a determined effort to obtain as many books as possible for their library. Buying up books in the book markets of Athens and Rhodes was one way of increasing the collection [7], but the Ptolemies also turned to more extreme methods. According to Ga-

len, all books found on board ships that docked in Alexandria were seized, taken away, and copied. Then the copies, not the originals, were returned to the owners. The books acquired in this way were marked 'from the ships.' The Athenians, perhaps, should have known better than to lend one of the Ptolemies their precious official edition of the tragedies of Sophocles, Aeschylus, and Euripides, even if he did give them 15 talents as a security. The king kept the originals and returned the copies with the small consolation that they were produced on the very best papyrus available.[8] Whatever the truth of these stories, the view prevailed that the Library's appetite for books was voracious. Some even suggested that the Ptolemies wished to acquire copies of all books ever written, though translated into Greek first [9].

To have all written books within the library is an important step towards bibliographical control, but the next step: to make them retrievable, is also important. We do not have much knowledge about the degree to which retrievability was obtained . It is obvious, however, that this library deserves a prominent place in the history of bibliographic control [13].

### 2.2 Gessner's *Bibliotheca Universalis* (1545)

In 1548, the Swiss naturalist and bibliographer Conrad Gessner published *Bibliotheca Universalis*, the first bibliography of all known books in Hebrew, Greek and Latin after the invention of printing, in alphabetical order. Followed in 1548 by *Pandectae*, which was a systematic arrangement of the same bibliographical references. Although incomplete, [14] these works made Gessner a pioneer in the field of bibliography (sometimes called "the father of bibliography"), and *Bibliotheca Universalis* one of the main examples in the history of bibliographical control and an inspiration for many later attempts to produce comprehensive bibliographies [15].

### 2.3 The abstract journal and subject bibliographical databases (1790-)

In scholarly LIS literature about bibliographical control, bibliographical databases are seldom considered. This is here considered a shortcoming as they represent the most important element. The abstract journal understood as a journal providing abstracts of articles in other journals, became influential in the 19th century (Manzer 1977; Fyfe 2021), and abstract journals soon became a dominant element in scientific communication to establish bibliographical control over scientific literature in their respective disciplines. In the beginning, they were published as printed

Knowl. Org. 50(2023)No.4
B. Hjørland. Bibliographical Control

305

journals, in a period also as microfilm publications, but from the 1960s also published as electronic databases (paper editions typically ceased to be published about 2000).

The electronic databases made a huge difference in scholarly communication and gave rise to the field of information retrieval. The many innovations are described by Bourne and Hahn (2003). Another dimension of this revolution was the emergence of citation indexes [16].

The development of the citation indexes provides an important new, if not revolutionary, perspective on bibliographic control: the idea that documents are linked in networks by their citations and that the authors of these documents—prominent subject specialists—provide a kind of self-organizing system that provides a kind of bibliographic control (as all scientific papers are supposed to be based on the most important knowledge/documents, this must be reflected in their bibliographical references). In this perspective, we have two competing systems of bibliographical control: One made by bibliographic databases, and one made by the primary literature itself. These two systems supplement each other (with surprisingly little overlap), and their relative effectiveness is an important topic for research [17].

The system of separate subject bibliographies as an approach to bibliographic control may, in contrast to some alternatives, be considered what Taube (1951, 67) described as "separate, overlapping, duplicating bibliographical services—a chaos of conflicting organizations and purposes from which current national bibliography and current complete universal bibliography [...] were to rescue us?" But after a careful analysis of the issues involved, he found that systems for special disciplines, after all, is the best alternative because (71): "the categories of any discipline will reflect the basic interests and the purposes of those concerned with that discipline." It should be added that Taube's analysis explicitly was based on a "functional approach."

The revolution that started with scientific bibliographic databases continues today with Internet technologies, but databases focusing on scholarly communication should still be considered an extremely important element of bibliographic control over the information ecosystem (and not just something that is made obsolete by the WWW).

### 2.4 Universal Bibliographical Repertory (Répertoire Bibliographique Universel, RBU) (1895-)

The Universal Bibliographical Repertory (*Répertoire bibliographique universel*, RBU) or "the universal bibliography" was described by Paul Otlet (1897) as "an inventory of all that has been written at all times, in all languages, and on all subjects" (Wright 2014, 76). It is the best known among other attempts to provide bibliographical control about the *fin de siècle*. A rival to RBU was the London based *International Catalog of Scientific Literature*[18].

RBU was produced by the *International Institute of Bibliography* (*Institut International de Bibliographie*, IIB). By 1934 nearly 16 million cataloging cards had been collected and classified in the RBU. In the beginning, it was classified according to the *Dewey Decimal Classification* (DDC), but the very large extent of the bibliography made it necessary to develop a new system: the *Universal Decimal Classification* (UDC).

The Repertory received requests from all over the world and answers sent by mail in the form of copied cards relevant to the query. Otlet was very interested in applying new technologies. Unfortunately, the project lost its financial basis and much of the repertory was lost. It is a question, however, if the project could and should have continued or if a system of subject bibliographies (see Section 2.3) or of national bibliographies (see Section 2.6) would be better strategies. Wright (2014, 72) referred to some contemporary objections and criticism of it:

> Some objected to the project on its most basic premise, arguing that true universality across disciplines was an unachievable dream; better to have subject-specific schemes tailored to the nuances of each discipline rather than one shallow classification that tried to cover every subject under the sun.

RBU has been called a paper forerunner for Google and is today a part of the *Mundaneum* Museum supported by Google. [19] Csiszar (2013, 443; italics in original) wrote, however:

> [I]t is not obvious that the best way to deal with the literature problem should have been grandiose subject classification projects. Other more traditional—some might say more sensible—options existed: better specialized disciplinary bibliographies, expanded alphabetical subject indexes, or even relying more on the accumulated knowledge of recognized experts. (Indeed, given the ultimate failure of many of these projects, it might plausibly be argued that in hindsight universal and detailed classification *was* a misguided approach.)

See further in Otlet (1990), Heuvel (2009) and Wright (2014).

### 2.5 Union catalogs (1930s-)

Union catalogs are catalogs which list the holdings in more than one library or collection. They may be book catalogs, journal catalogs, incunabula catalogs etc., they may be national or international, universal, or subject-specific, etc. Although the earliest union catalogs go back to the 12th and 13th centuries, they have reached a much larger size in the

306

Knowl. Org. 50(2023)No.4
B. Hjørland. Bibliographical Control

20th century, in both print and electronic forms. We shall mention one example of printed union catalogs: the *National Union Catalog, Pre-1956 Imprints*, which is a 754-volume set containing books cataloged by the *Library of Congress* and other American and Canadian libraries in alphabetical order. This printed mammoth represents the apogee of printed catalogs and bibliographies which lost their relevance in the online era. [20]

### 2.5.2 WorldCat (1971)

*WorldCat* is a union catalog database provided by OCLC (formerly known as *Online Computer Library Center*), which is an American nonprofit cooperative organization. It is the largest database of its kind, and was established in 1971, but only got its present domain name in 1998. White and Zuccala (2018, 1502) wrote:

> WorldCat lists the holdings of more than 16,000 members; public, academic, and research libraries are major types. These libraries are mainly in North America, but they extend around the globe. Books in English dominate their collections, but many also have sizable holdings in other languages.

Compared to the subject bibliographical databases (see Section 2.3), union catalogs, including *WorldCat*, do not provide the same degree of bibliographic control, especially on article indexing. They serve, however, other functions, in particular (1) known-item verification of monographs and (2) locating of monographs in libraries [21].

### 2.6 IFLA's and UNESCO's program on universal bibliographic control (1970s-)

The *International Federation of Library Associations and Institutions* (IFLA) and the *United Nations Educational, Scientific and Cultural Organization* (UNESCO) had in the 1970s a program on "universal bibliographic control" which was based on the idea that each nation should produce a national bibliography and that international cooperation among national bibliographies should ensure "universal" bibliographic control. IFLA Professional Statement on Universal Bibliographic Control (2012) wrote:

> A National bibliographic agency (NBA) has the responsibility for providing the authoritative bibliographic data for publications of its own country and for making that data available to other NBAs, libraries, and other communities (for instance archives and museums) through appropriate and timely services with the goal of increasing open access to the bibliographic data.

National bibliographies themselves are much older than IFLA's program (although the present meaning of "national bibliography" only goes back to the 1950s).[22] For example, *Dansk bogfortegnelse* (Danish national bibliography: Books) has been published since 1851. Their preparation is often made by (and dependent on) national libraries and the concept of legal deposit of publications. The first law on legal deposit is from France in 1537, under which a copy of any published book had to be delivered to the king's library, for conservation purposes, and sometimes to facilitate censorship. Similar laws were passed in many other European countries. By contrast, in the U.S.A., the delivery of copies of printed books to the *Library of Congress* serves copyright purposes. This makes a difference in what books come under bibliographic control and exemplify how different interests influence the contents of national bibliographies.

It is a common misunderstanding, perhaps a kind of hype, that national bibliographies represent a prerequisite for all other bibliographies, and thereby have a foundational role in bibliographic control. This view was expressed by Downs (1954, 500):

> Viewing the question of bibliographical control in the perspective of history, there seems little doubt that effective national bibliographic organization must precede international or universal coverage. Starry-eyed bibliographers, who for generations have advocated a worldwide approach to bibliography, present an almost unbroken record of futility, frustration, and failure, except, perhaps, when they limit themselves to special aspects. If universal bibliography is ever to be achieved, it must be grounded upon the work of individual countries.[23]

However, scientific bibliographical databases such as MEDLINE, Web of Science, Scopus, PsycINFO etc. do not have their input from national bibliographies, and Taube (1951) argued that to base subject bibliographies on the representations of universal bibliographies and catalogs is also theoretically a bad idea because different contexts need different subject headings. Taube (66-67) exemplified his argument that the *Library of Congress Subject Headings* were too general to be used by the *Department of Agriculture* or by the *Army Medical Library*. And the other way round: the subject headings of a special library are too specific to be used by the *Library of Congress*. He wrote (67):

> But when the *Library of Congress* rejects a heading supplied by a special library, it prints the rejected heading along with its own heading because the *Library of Congress* officially recognizes that the purpose of a special library may require a degree of specificity in indexing not necessary or desirable for its

Knowl. Org. 50(2023)No.4
B. Hjørland. Bibliographical Control

307

own general purposes. In short, the distinction we discerned between general and special classification systems also exists between general and special collections of subject headings. And we conclude that there cannot be one subject key to bibliographic organization which would serve the librarian, the general reader, and all specialists.

IFLA's program focused mostly on cataloging standards (and this may be the reason for the widespread confusion of bibliographical control with KO and document description mentioned in Section 1). Critical attitudes have been expressed in relation to library cataloging, and by implication to this way of providing bibliographical control. Line (2005, 110-111) is one example of this criticism. He wrote:

> AACR2 [Anglo American Cataloging Rules 2nd ed.] is one of the most remarkable examples of trying to solve a problem by committee, with predictable results. The committee did not even tackle the right problem – what users surely want is not comprehensive or perfectly accurate bibliographic records, but far better subject access to books, comparable with that provided for scientific journal articles by the large international databases. No data on users' needs, whether for bibliographic information or subject access, were collected; instead, cataloguers discussed how to change the rules, rather as if hens were to gather together to discuss the design of eggs. I am however doing the committee an injustice in accusing them of not involving consumers in their discussions, because much of the use made of catalogues is in fact by cataloguers for the purpose of adding to them. Cataloguers would lose their status if it were shown that most cataloguing is a trivial job easily done by clerical staff or that the length of a catalogue entry was not a sign of virility.

Dunsire, Hillmann and Phipps (2012) suggested that bibliographic control in general, including the IFLA program[24], should consider the technologies of semantic web and linked data to provide better utilization of the metadata produced by many different groups and organizations. The discussion of these technologies cannot be done in the present article and is more an issue about document representation than genuine bibliographical control. However, the article's argument (164) to replace attempts based on "one-size-fits-all schema, rules and other international/global standards with what might be termed an all-sizes fit-one approach [... Which] can support a much richer ecology of bibliographic communities and their standards," seems to be an important idea.

## 2.7 The World Wide Web (WWW) and Google (1989-)

WWW was a hypertext information network proposed in 1989 by Tim Berners-Lee and Robert Cailliau at the *European Organization for Nuclear Research* (CERN). In 1990 it was named "World Wide Web" and launched by CERN on May 17, 1991. That system quickly revolutionized almost everything in the world. Its strength is the broad possibility of cooperating, that anybody with a computer can draw from and contribute to the network of information resources, saving a huge amount of labor. For example, *Wikipedia* was launched in 2001, stunned the world, and made times hard for established encyclopedias such as *Encyclopaedia Britannica*.

*Google Search* (also known as Google) was originally developed in 1996 by Larry Page, Sergey Brin, and Scott Hassan. It is a search engine designed to retrieve documents offered by web servers (WWW). It can find text, images and other media and provide access to free documents or to paywalls, where documents can be retrieved by payment. Of course, not all the information in the world is on the WWW, and there is information on the WWW that Google and other search engines do not reach - the so-called "deep web." [25]

In 2019, Google wrote about itself: "Our mission is to organize the world's information and make it universally accessible and useful." Google has achieved, among other things:

(1) To make the search engine with most comprehensive coverage of webpages (often including published full-text documents).

(2) Google Scholar, a database with a citation index and an impressive coverage of scientific literature. According to Gehanno, Rollin, and Darmoni (2013), it contains more relevant medical studies than MEDLINE! (of 738 original studies included in a gold standard database, all were found in Google Scholar, which is extraordinary). As stated later, even if this should be the case, MEDLINE may still provide better retrieval in subject searches.)

(3) Google Books has made many million books available on the Internet, full-text searchable. Many as free books, but many only with limited reading possibilities. In many cases, Google provides facilities to locate, for example, quotations, much better than the printed index in the books themselves.

In one way, Google represents the closest the world has ever come to universal bibliographic control. On the other hand,

308

Knowl. Org. 50(2023)No.4
B. Hjørland. Bibliographical Control

it could be said that Google does not really represent bibliographic control because links and documents often disappear (and the control thereby disappears). This is a problem that web archives such as the *Internet Archive* try to tackle. Apart from coverage, the issue is retrievability in subject searches. At this point classical bibliographical databases seem to perform better compared to Google's full-text search possibilities, due to better indexing and metadata. This problem is probably the most important one in information science today.

See further about the history of the web and the web in general in Gillies and Cailliau (2000), Gardiner and Musto (2010), and Tomer (2017). About coverage and retrieval effectiveness of Google see Lewandowski (2008), Gehanno, Rollin and Darmoni (2013), Bramer, Giustini and Kramer (2016), Gusenbauer and Haddaway (2020), and Hjørland (2021).

## 2. Theoretical problems in obtaining bibliographical control

It is obvious that the goal of obtaining comprehensive coverage in bibliographical databases meets many practical and financial problems. The cost of labor (and/or computer resources) to identify and describe documents has always been an obstacle. The idea of making a bibliography is relatively simple but requires much work to fulfill. However, such practical issues cannot be addressed in information science and knowledge organization, where the focus must be on the theoretical problems, including how to evaluate the effectiveness of different solutions. Two main theoretical issues concerns (1) coverage and (2) retrievability from the contents of a given database.

Concerning coverage, the first issue to consider is that despite the impression of completeness, any bibliographical database is always selective. The medical database MED-LINE, for example, contains more than 29 million references to the biomedical literature. Why isn't that enough? Experiences from systematic reviews show that more than this database is needed if a satisfactory identification of relevant findings should be obtained. [26] The problem lies in the selection of journals (or in general: information resources). Any selection must be based on criteria, and any set of criteria are only indicators, not a waterproof guarantee that all relevant items are included. Bradford (1953, 148) wrote about the scattering of articles on a given subject:

> It is, therefore, necessary to examine the extent to which articles on a given subject actually occur in periodicals devoted to quite other subjects: as, for instance, a paper on the mechanism of the heart, contributed to the *Proceedings of Physical Society*, or one on genetics, occurring in an agricultural magazine. In-

vestigation shows that this distribution follows a certain law, which can be deduced both theoretically from the principle of the unity of science and practically from examination of the references.

> According to this principle every scientific subject is related, more or less remotely, to every other scientific subject.
> It follows that from time to time, a periodical devoted to a special subject may contain an article of interest from the point of view of another subject. In other words, the articles of interest to a specialist must occur not only in the periodicals specializing on his subject, but also, from time to time, in other periodicals, which grow in number as the relation of their fields to that of his subject lessens and the number of articles on his subject in each periodical diminishes.

According to Bradford, this means that no subject database can ever be expected to have full bibliographical control. This problem is not relevant for universal databases such as *Google Scholar* (and less important in *Web of Science*, which is more selective), but still, complete coverage is an unobtainable dream. Following this principle, journals may be classified into "zones," where zone 1 represents the journals that have most articles on the subject, zone 2 includes the journals that have had an average amount of articles, and zone 3 comprises the long tail of journals that occasionally or seldom contain articles about that subject. Such classifications have been used as selection criteria for special libraries and databases, but this implicitly acknowledges the impossible of full theoretical coverage. There is, however, a perspective, that was neglected by Bradford, and has only been understood recently: Bradford realized that all documents are related to all others (due to the doctrine of unity of science). But how should we decide which are most closely related and important to include in a subject bibliography? The answer to this question is theory-dependent and may shift with "paradigm-shifts" in a field, and thus on interpretations rather than simple statistical figures about word-frequencies or related data. This is again an example of the difference between "positivist" and "pragmatist" philosophy.

Although Google has an extremely comprehensive index of the internet, not all pages are updated with the same frequency, and there are still problems like the dark web, although, as said in Section 2.7, Gehanno, Rollin, and Darmoni (2013) found that Google Scholar has better coverage than MEDLINE. More studies are needed to determine the relative coverage of the relevant literature, because other studies have indicated otherwise, [27] and because the databases seem to be competing and due to constant development. It is important to realize that studies of this kind pro-

Knowl. Org. 50(2023)No.4
B. Hjørland. Bibliographical Control

309

vide important information about whether bibliographical control is satisfactory or not.

Concerning retrieval from the contents covered by a given database, to have the relevant documents in a database (or in an internet-index like Google's) represented in a database is one thing. Another thing is how to select the document for a given query from this database/index. This problem is the subject of the whole field of information retrieval, and numerous approaches have been suggested and tried, which will not be presented in this article. Swanson (1986) expressed a problem, that sadly has been neglected in the literature of information science. He wrote:

> The search process, like a scientific theory, can be criticized and improved, but can never be verified as capable of retrieving all information relevant to a problem or theory. This essential incompleteness of search and retrieval therefore makes possible, and plausible, the existence of undiscovered public knowledge.

This is also a problem for both MEDLINE and Google (and Google Scholar). MEDLINE relies on *Medical Subject Headings* (MeSH) the effectiveness of which seems to be taken for granted by the *National Library of Medicine* but seems to need further research. About search engines like Google, it has frequently been pointed out in the research literature that they are not neutral, objective machines (Introna and Nissenbaum 2000; Granka 2010). Therefore, Lewandowski (2014) found that there is a need for alternatives, for many competing search engines. The problem of obtaining this is the extremely costly process of establishing a basic index of the Web on which the competing search engines could operate. About the cost of establishing such an index, he wrote (57):

> This type of project will require a considerable investment of funds. The total cost cannot be precisely forecast here. Several hundred million euros will likely be needed, however, when one considers the anecdotal reports provided by search engine operators. The losses Bing has reported on its search activities are one example. This may appear prohibitively expensive until one considers, for instance, that the German government invested roughly 100 million euros in the semantic technology developed as part of the Theseus program. Nevertheless, it is still clear that any one country alone cannot support this type of project. The only feasible option is a pan-European initiative.

Because of these issues, as stated in Section 1, bibliographical control in an absolute meaning is an unobtainable goal, even in principle. As pointed out by Wright (2014), all such ambitions throughout history represent a problematic pos-

itivist view — not the least Google's aim "to organize the world's information and make it universally accessible and useful [28]." Wright's theoretical analysis and critiques of approaches to bibliographical control are valuable, although one misses an analysis of what the opposite to positivism is, and how such an alternative would approach the issue of bibliographical control. The present author believes that this question is the key to solving the fundamental problems of information science. In the present article, the approach represented by scientific bibliographical databases comes closer to non-positivism (by avoiding universalist ideas), although also, at this next level, these opposite philosophies are at play in the way they select and represent documents.

But that does not mean that these efforts (including Google's) have not been extremely useful. What it means is that research needs to be based on a better theoretical and philosophical basis, and that further progress should be based on developing information systems optimal for specific purposes, interests, and disciplines rather than on mistaken philosophies. [18]

## 3. Conclusion

The idea of bibliographic control from the *Library of Alexandria* to the present date is an exciting story that demonstrates how different ages have used their available technologies and ideas to achieve the goal associated with this idea. The idea has often been associated with positivist dreams of universal, complete, all-purpose, and value-free collections, bibliographies, and retrieval systems. It is also often associated with the view that this goal will be provided by advances in information technology rather than by advances in understanding the goals to be fulfilled. An important exception from this common view is Wilson (1968, 5), who wrote:

> There is less of conceptual innovations [in relation to bibliographical control] in most remedial proposals than enthusiasts and promoters think, but the impressive technology is undeniable. But technology tells us neither what is worth aiming at (or away from), nor what is satisfactory degree of progress toward our elected goals. That we must discover or decide for ourselves; and that is best done with the greatest possible clarity about alternative goals and the intrinsic difficulties in their pursuit.

Among the suggestions found in the literature on bibliographical control are some important ideas, including (1) considering the network of citing papers a kind of self-organizing system for bibliographic control competing with bibliographies established by others, (2) establishing a pub-

310

Knowl. Org. 50(2023)No.4
B. Hjørland. Bibliographical Control

lic index of the Internet (Lewandowski 2014) on which competing search engines can be built to meet different goals and interests better and (3) an idea, that is not new but now suggested for the new internet technologies, is the idea by Dunsire, Hillmann and Phipps (2012) that bibliographical control must be adjusted to the needs of different purposes, groups, and subject. The present author finds that this last view also was the main message in Wilson (1968).

At the most general level, the conclusion in relation to bibliographical control is the same as for all other issues in information science and KO: there are no pure technological, neutral solutions, and ambitions to make all information in the world available for all purposes and interest is ill-conceived. At the philosophical level, the field should turn from (implicit) positivist assumptions to explicit functional/ pragmatic research strategies.

The problem of bibliographical control is important but seems to be neglected or redefined to be about document descriptions. This article has used some examples from the medical domain to illustrate how the goal of establishing bibliographical control is important and addressed by research about the coverage of databases and the retrievability of documents in the various databases. Hopefully, similar attention will develop in other domains, and come more in focus in information science with KO.

## Acknowledgements

## Endnotes

1. Library of Congress (2008, 6) defined: "Bibliographic control is the organization of library materials to facilitate discovery, management, identification, and access. It is as old as libraries themselves, and our current approaches to it are direct descendants of the librarianship of the 19th century."

2. Maxwell (2017, 447) wrote: "Bibliographic control is the process of creation, exchange, preservation, and use of data about information resources. Formal bibliographic control has been practiced for millennia, but modern techniques began to be developed and implemented in the nineteenth and twentieth centuries. A series of cataloging codes characterized this period. These codes governed the creation of library catalogs, first in book form, then on cards, and finally in electronic formats, including MAchine-Readable Cataloging (MARC)."

3. Another striking example is that Wikipedia redirects queries for "bibliographic control" to "Cataloging (li-

brary science)" and writing: "Bibliographic control provides the philosophical basis of cataloging, defining the rules that sufficiently describes information resources, to enable users find and select the most appropriate resource. A cataloger is an individual responsible for the processes of description, subject analysis, classification, and authority control of library materials." See https://en.wikipedia.org/w/index.php?title=Bibliographic_control&redirect=no and https://en.wikipedia.org/wiki/Cataloging_(library_science) (Visited January 4., 2023).

4. Wilson here has a footnote 5: "Francis Bacon, Novum Organum, Book 1, Aphorism 3: 'Human knowledge and human power meets in one, for where the cause is not known the effect cannot be produced'."

5. See Schöpfel et al. (2021).

6. The opposite of universal bibliographic control is partial bibliographic control. That part must be defined by formal criteria, it cannot just be, for example, the holdings of a single library, unless that library attempts to and have the means to survey the total output of publications within its field. If not formally defined, the word "control" probably loses its meaning.

7. Sevelsted and Tortzen (2010), unfortunately only available in Danish, is about this problem and may be considered an important humanistic background for a part of what in information science is called "authorship attribution" and "author name disambiguation."

8. See further in Osborne and Holland (2009).

9. "Subjects" are often considered inherent qualities of documents, and subject cataloging (classification and indexing) is often considered a part of the description of documents. This is, however, a problematic understanding according to the "request-oriented view" (or "policy-oriented view") of classification and indexing. Documents do not "have" subjects, but subjects are attributed or ascribed to documents to improve retrievability according to the purposes of the organization doing the indexing (see →Subject (of documents), Hjørland 2017).

10. One can sense a certain questioning of the term descriptive cataloging when Wilson (1989, 15) wrote about a needed reconceptualization "of the task of what hitherto has been called descriptive cataloging."

11. For a detailed examination of the term "known item search" see Lee, Renear and Smith (2006).

12. Notes [7], [8] and [9] from Erskine (1995, 39) are here omitted.

13. How the books were organized and retrieved in this library has been discussed by Olesen-Bagneux (2014). About the Library of Alexandria see further in El-Abbadi (1990).

Knowl. Org. 50(2023)No.4
B. Hjørland. Bibliographical Control

311

14. Downs (1954, 499) wrote: " 'The first bibliographer of the modern world,' Conrad Gesner of Zurich, in 1545, about a century after printing began, published his *Bibliotheca Universalis*, one of the monuments of early bibliography. His work fell far short of completeness, though, and, as Henry Bartlett Van Hoesen [1937] commented, '. . . if Gesner's bibliography was 'partial' and incomplete at a time when there were probably not more than 40,000 or 50,000 books in print, we may well despair of universality now'."

15. See further in Crippa and Araujo (2020) and Fischer (1966).

16. See Araújo, Castanha and Hjørland (2021).

17. See further in Pao and Worthen 1989; Pao 1993a, 1993b; Hjørland and Kyllesbech Nielsen 2001, 272-77; Hjørland 2013 and Hirt et al. 2022).

18. See Csiszar, 2013

19. Mundaneum is the overall center for RUB and other initiatives. Today it is a museum. Homepage: http://archives.mundaneum.org/en/

20. See further in Creider (2010) and Lass and Quandt (2004).

21. See further in Fowler (2010, 1269).

22. Madsen (2000, 46) wrote: "The national bibliography defined as a system or as a national bibliographic service is fairly recent. The term 'national bibliography' appears in the literature about the middle of the previous century- primarily as a designation of the national book list. It was not until the first half of the 20th century that the present broad definition was recognized. Important foundations for this 'new' perception are the international conferences arranged by UNESCO in 1950 and by IFLA and UNESCO in 1977."

23. Another example is Madsen (2000, 45), who wrote: "'The primary sources for all types of bibliography are the national bibliographies, i.e. each country's records of the literature that is published or has been published within the borders of that country. Only when this material has been obtained, can we say that it is possible to find the part of the whole world's production that is needed in the individual cases.'(from Danish) (Larsen, 1959, p. 299)." The same quote (and misunderstanding) is also given in Kungliga Biblioteket (2003, 150).

24. The IFLA program is further described in Anderson (1974 and 1984).

25. See further about "deep web" in Zheng et al. 2013. About "dark web" see Sobhan et al. 2022.

26. Hjørland (2015, 1568) wrote: "So far, all guidelines have considered MEDLINE the most important bibliographical medical database and expert searching based on Boolean retrieval the most important retrieval model. Higgins and Green (2009), for example, wrote: "It is recommended that for all Cochrane reviews, CENTRAL [The Cochrane Central Register of Controlled Trials, which is partly based on MEDLINE] and MEDLINE should be searched, as a minimum, together with EMBASE if it is available to either the CRG [Cochrane Review Group] or the review author" (electronic source, no page). Crumley et al. (2005) provide a systematic review examining the evidence concerning which sources should be used for the production of systematic reviews. They found that multiple source comprehensive searches are necessary to identify all randomized control trials for a systematic review and that indexing needs to be improved."

27. Gusenbauer and Haddaway (2020, 211): Google Scholar's search precision has been found to be significantly lower than 1% for systematic searches [Boeker, Vach and Motschall 2013]. This is not surprising, since our findings show that Google Scholar does not support many of the features required for systematic searches. Our findings support the criticism of Bramer et al. (2013), Bramer, Giustini and Kramer (2016) and Boeker, Vach and Motschall (2013) and indicate that Google Scholar's coverage and recall is an inadequate reason to use it as a principal search system in systematic searches.53 If a system such as Google Scholar fails to deliver retrieval capabilities that allow a reviewer to search systematically with high levels of recall, precision, transparency, and reproducibility, its coverage is irrelevant for query-based search. Google Scholar's extraordinary coverage acting as a multidisciplinary compendium of scientific world knowledge should not blind users to the fact that users' ability to access this compendium is severely limited, especially in terms of a systematic search."

28. For a criticism of the principles on which Google is based, see Hjørland 2021.

29. Kröber and Segeth (1976, 214, translated from German, italics in original) criticized the ideal that descriptions should be complete: "The mere description of facts, declared by positivism to be the sole task of science, defines the concern of science too narrowly in the sense of what has been said above. In addition, the *complete* description of a phenomenon that positivism demands is not only impossible, but it is also unnecessary. It is impossible because the phenomenon with its infinite variety of properties and relations to other phenomena would require an infinite number of descriptions; and it is unnecessary because scientific knowledge and the practical activity of people do not depend on an equally detailed description of the essential and inessential, the necessary and contingent, the general and individual properties and relations of the phenomenon. On the contrary, it depends on the knowledge of the essence, of the general in the individual, of the nec-

312

Knowl. Org. 50(2023)No.4
B. Hjørland. Bibliographical Control

essary and lawful in the contingent. The description can therefore only do justice to its function if it is not made absolute and detached from the other scientific knowledge processes and means, such as explanation, hypothesis, prognosis, etc., but is seen and practiced in unity with them."

## References

Anderson, Dorothy Pauline. 1984. "Universal Bibliographic Control." In *Encyclopedia of Library and Information Science*, edited by Allen Kent, Harold Lancour and Jay E Daily: New York: Dekker Vol. 37, Suppl. 2, 366–401.

Anderson, Dorothy Pauline. 1974. *Universal Bibliographic Control: A Long-Term Policy, A Plan for Action*. Munich, Germany: Verlag Dokumentation.

Araújo, Paula Carina de, Renata Cristina Gutierres Castanha and Birger Hjørland. 2021. "Citation Indexing and Indexes". *Knowledge Organization* 48, no.1: 72-101. Also available in *ISKO Encyclopedia of Knowledge Organization*, edited by Birger Hjørland and Claudio Gnoli, https://www.isko.org/cyclo/citation

Bateson, Gregory. 1977. "The Thing of It Is". In *Earth's Answer*, edited by Michael Katz, William P. Marsh and Gail Gordon Thompson. Explorations of Planetary Culture at the Lindisfarne Conferences. New York: Lindisfarne Books / Harper & Row, 143-54.

Boeker, Martin, Werner Vach and Edith Motschall. 2013. "Google Scholar as Replacement for Systematic Literature Searches: Good Relative Recall and Precision are not Enough." *BMC Medical Research Methodology* 13: 1-12. https://doi.org/10.1186/1471-2288-13-131

Bourne, Charles P. and Trudi Bellardo Hahn. 2003. *A History of Online Information Services 1963-1976*. Cambridge, Mass.: MIT Press.

Bradford, Samuel C. 1953. *Documentation.* 2nd ed. London, UK: Crosby Lockwood

Bramer, Wichor M., Dean Giustini and Bianca M. R. Kramer. 2016. "Comparing the Coverage, Recall, and Precision of Searches for 120 Systematic Reviews in Embase, MEDLINE, and Google Scholar: a Prospective Study." *Systematic Reviews* 5: 39. https://doi.org/10.1186/s13643-016-0215-7

Bramer, Wichor M., Dean Giustini, Bianca M. R. Kramer and P.F. Anderson. 2013. "The Comparative Recall of Google Scholar versus PubMed in Identical Searches for Biomedical Systematic Reviews: a Review of Searches used in Systematic Reviews." *Systematic Reviews* 2: 1-9. https://doi.org/10.1186/2046-4053-2-115

Chan, Lois Mai and Athena Salaba. 2016. *Cataloging and Classification: An Introduction*, 4th ed. Lanham, MD: Rowman & Littlefield Publishers.

Creider, Laurance C. 2010. "Union Catalog." In *The Oxford Companion to the Book*, edited by Michael Felix Suarez and H. R. Woudhuysen. Oxford, UK: Oxford University Press, volume 2: 1232.

Crippa, Giulia and Andre Vieira de Freitas Araujo. 2020. "Order of Knowledge, Selection and Bibliographical Tension in the 16th Century Between Gesnerian Universality and Possevinian Anti-Heretism." In *Knowledge organization at the interface: Proceedings of the Sixteenth International ISKO Conference 6-8 July 2020 Aalborg, Denmark*, edited by M. Lykke et al. Advances in Knowledge Organization 17: 105-14.

Crumley, Ellen T., Natasha Wiebe, Kristie Cramer, Terry P. Klassen and Lisa Hartling. 2005. "Which Resources Should be Used to Identify RCT/CCTs for Systematic Reviews: A Systematic Review." *BMC Medical Research Methodology* 5, no.24: 1-13. http://www.biomedcentral.com/content/pdf/1471-2288-5-24.pdf

Csiszar, Alex. 2013. "Bibliography as Anthropometry: Dreaming Scientific Order at the fin de siècle." *Library Trends* 62, n.2: 442-55. https://doi.org/10.1353/lib.2013.0041

Downs, Robert B. 1954. "Problems of Bibliographical Control." *Library Trends* 2: 498–508. https://core.ac.uk/download/pdf/4814884.pdf

Dunsire, Gordon, Diane Hillmann and Jon Phipps. 2012. "Reconsidering Universal Bibliographic Control in Light of the Semantic Web." *Journal of Library Metadata* 12, nos. 2/3: 164-76. https://doi.org/10.1080/19386389.2012.699831

El-Abbadi, Mostafa. 1990. *The Life and Fate of the Ancient Library of Alexandria*. Paris, France: Unesco/UNDP.

Erskine, Andrew. 1995. "Culture and Power in Ptolemaic Egypt: The Museum and Library of Alexandria". *Greece & Rome* 42, no. 1: 38-48. http://www.jstor.org/stable/643071.

Fischer, Hans. 1966. "Conrad Gessner (1516–1565) as Bibliographer and Encyclopedist." *Library* s5-XXI, no. 4: 269-281

Fowler, Chris. 2010. "WorldCat." In *The Oxford Companion to the Book*, edited by Michael Felix Suarez and H. R. Woudhuysen. Oxford, UK: Oxford University Press, volume 2: 1269.

Fyfe, Aileen. 2021. "Where Did the Practice of 'Abstracts' Come From?" In *A History of Scientific Journals: Lessons from the History of Royal Society Journal Publishing, 1665-2015* [Blog]. https://arts.st-andrews.ac.uk/philosophicaltransactions/where-did-the-practice-of-abstracts-come-from/

Gardiner, Eileen and Ronald G. Musto. 2010. "World Wide Web (WWW)." In *The Oxford Companion to the Book*, edited by Michael Felix Suarez and H. R. Woudhuysen. Oxford, UK: Oxford University Press, volume 2: 1269-70.

Knowl. Org. 50(2023)No.4
B. Hjørland. Bibliographical Control

313

Gehanno, Jean-François, Laetitia Rollin and Stefan J Darmoni. 2013. "Is the Coverage of Google Scholar Enough to be Used Alone for Systematic Reviews." *BMC Medical Informatics and Decision Making* 13, no.1: 7. http://www.biomedcentral.com/1472-6947/13/7

Gillies, James and Robert Cailliau. 2000. *How the Web Was Born: The Story of the World Wide Web*. New York: Oxford University Press.

Google. [2019]. "About Google". Visited April 28, 2019. Archived in the *Internet Archive* at http://web.archive.org/web/20190428121852/https://about.google/intl/da/

Granka, Laura A. 2010. "The Politics of Search. A Decade Retrospective." *The Information Society* 26, no. 5: 364-74. https://doi.org/10.1080/01972243.2010.511560

Gusenbauer, Michael and Neal R. Haddaway. 2020. "Which Academic Search Systems are Suitable for Systematic Reviews or Meta-analyses? Evaluating Retrieval Qualities of Google Scholar, PubMed, and 26 other Resources." *Research Synthesis Methods* 11: 181–217. https://doi.org/10.1002/jrsm.1378

Higgins, Julian P. T. and Sally Green (eds). 2009. *The Cochrane Handbook for Systematic Reviews of Interventions*, version 5.0.2. London: The Cochrane Collaboration. https://training.cochrane.org/handbook/archive/v5.0.2/

Hirt, Julian, Thomas Nordhausen, Christian Appenzeller-Herzog and Hannah Ewald. 2022. "Citation Tracking for Systematic Literature Searching: A Scoping Review." *MedRxiv: The Preprint Server for Health Sciences* 2022.09.29.22280494. https://doi.org/10.1101/2022.09.29.22280494

Hjørland, Birger and Lykke Kyllesbech Nielsen. 2001. "Subject Access Points in Electronic Retrieval". *Annual Review of Information Science and Technology* 35: 249-98. A version is available in *ISKO Encyclopedia of Knowledge Organization*, edited by Birger Hjørland and Claudio Gnoli: https://www.isko.org/cyclo/sap

Hjørland, Birger. 1997. *Information Seeking and Subject Representation: An Activity-Theoretical Approach to Information Science*. Westport, CO: Greenwood Press.

Hjørland, Birger. 2011. "Evidence Based Practice: An Analysis Based on the Philosophy of Science." *Journal of the American Society for Information Science and Technology* 62, no. 7: 1301-10. https://doi.org/10.1002/asi.21523

Hjørland, Birger. 2013. "Citation Analysis: A Social and Dynamic Approach to Knowledge Organization". *Information Processing and Management* 49, 1313-1325. DOI http://dx.doi.org/10.1016/j.ipm.2013.07.001. Also available in ISKO Encyclopedia of Knowledge Organization, edited by Birger Hjørland and Claudio Gnoli: https://www.isko.org/cyclo/citation_analysis

Hjørland, Birger. 2015. "Classical Databases and Knowledge Organization: A Case for Boolean Retrieval and Human Decision-Making During Searches." *Journal of the Association for Information Science and Technology* 66, n.8: 1559–75. https://doi.org/10.1002/asi.23250

Hjørland, Birger. 2017. "Subject (of documents)". *Knowledge Organization* 44, no. 1: 55-64. Also available in *ISKO Encyclopedia of Knowledge Organization*, eds. Birger Hjørland and Claudio Gnoli, https://www.isko.org/cyclo/subject

Hjørland, Birger. 2021. "Information Retrieval and Knowledge Organization: A Perspective from the Philosophy of Science". *Information* 12, no. 3: article 135; https://doi.org/10.3390/info12030135. Also published as "Information retrieval" in *ISKO Encyclopedia of Knowledge Organization*, eds. Birger Hjørland and Claudio Gnoli, https://www.isko.org/cyclo/ir

IFLA Professional Statement on Universal Bibliographic Control. 2012. https://www.ifla.org/wp-content/uploads/2019/05/assets/bibliography/Documents/ifla-professional-statement-on-ubc-en.pdf

Introna, Lucas D. and Helen Nissenbaum. Helen 2000. "Shaping the Web. Why the Politics of Search Engines Matters." *The Information Society* 16, no.3: 169-85. https://doi.org/10.1080/01972240050133634

Joudrey, Daniel N., Arlene G. Taylor and Katherine M. Wisser. 2018. *The Organization of Information* 4th ed. Santa Barbara, CA: Libraries Unlimited, an imprint of ABC-CLIO.

Kröber, Günter and Wolfgang Segeth. 1976. "Beschreibung". In *Philosophisches Wörterbuch*, edited by Georg Klaus and Manfred Buhr. Westberlin: Das europäische Buch, 1: 214.

Kungliga Biblioteket. 2003. *KB - ett nav i kunskapssamhället*: Kungl. Biblioteket - Sveriges Nationalbibliotek : Verksamhet Och Visioner : betänkande/av KB-utredningen. Stockholm, Fritzes offentliga publikationer. SOU 2003:129. https://www.regeringen.se/rattsliga-dokument/statens-offentliga-utredningar/2004/01/sou-2003129/

Larsen, Knud. 1959. *Bibliografiens opgaver og metoder*. [Bibliography, Its Aims and Methods], in *Lærebog i Biblioteksteknik*, 4th ed., edited by Erik Allerslev Jensen. Copenhagen: Dansk Bibliografisk Kontor, vol. 2: 293-314.

Lass, Andrew and Richard E. Quandt (eds.). 2004. *Union Catalogs at the Crossroad*. Hamburg, Germany: Hamburg University Press. https://d-nb.info/973052813/34

Lee, Jin Ha, Allen Renear and Linda C. Smith. 2006. "Known-Item Search: Variations on a Concept." *Proceedings of the American Society for Information Science and Technology* 43: 1-17. https://doi.org/10.1002/meet.14504301126 (Free access).

Lewandowski, Dirk. 2008. "The Retrieval Effectiveness of Web Search Engines: Considering Results Descrip-

314

Knowl. Org. 50(2023)No.4
B. Hjørland. Bibliographical Control

tions." *Journal of Documentation* 64, no.6: 915-37. https://doi.org/10.1108/00220410810912451

Lewandowski, Dirk. 2014. "Why We Need an Independent Index of the Web. In *Society of the Query Reader: Reflections on Web Search*, edited by René König and Miriam Rasch. Amsterdam: Institute of Network Cultures. http://networkcultures.org/query/wp-content/uploads/sites/4/2014/06/4.Dirk%5FLewandowski.pdf http://networkcultures.org/query/2013/11/11/dirk-lewandowski-why-we-need-an-independent-index-of-the-web/

Library of Congress. 2008. *On the Record; Report of the Library of Congress Working Group on the Future of Bibliographic Control*. https://www.loc.gov/bibliographic-future/news/lcwg-ontherecord-jan08-final.pdf

Line, Maurice B. 2005. "Librarianship as It is Practised: A Failure of Intellect, Imagination and Initiative." *Interlending & Document Supply* 33, no. 2: 109–13. https://doi.org/10.1108/02641610510602655

Madsen, Mona. 2000. "The National Bibliography in the Future: New Recommendations." *Alexandria* 12, no.1: 45-50. https://doi.org/10.1177/095574900001200104

Manzer, Bruce M. 1977. *The Abstract Journal, 1790-1920. Origin, Development and Diffusion*. Metuchen, N.J.: The Scarecrow Press.

Maxwell, Robert L. 2017. "Bibliographic Control." In *Encyclopedia of Library and Information Sciences*, 4th ed, edited by John D. McDonald and Michael Levine-Clark. Boca Raton, FL: CRC Press, volume 1: 447-55. https://doi.org/10.1081/E-ELIS4-120043092

Miksa, Shawne D. (ed.). 2013. *Functional Future for Bibliographic Control: Transitioning into New Communities of Practice and Awareness*. London: Routledge. https://doi.org/10.4324/9781315093802

Olesen-Bagneux, Ole. 2014. "The Memory Library: How the Library in Hellenistic Alexandria Worked." *Knowledge Organization* 41, no.1: 3-13.https://doi.org/ 10.5771/0943-7444-2014-1-3

Osborne, Jason W. and Abigail Holland. 2009. "What Is Authorship, and What Should It Be? A Survey of Prominent Guidelines for Determining Authorship in Scientific Publications." *Practical Assessment, Research, and Evaluation* 14, Article 15. https://doi.org/10.7275/25pe-ba85

Otlet, Paul. 1990. *International Organisation and Dissemination of Knowledge: Selected Essays of Paul Otlet*. Translated and Edited with an Introduction by W. Boyd Rayward. Amsterdam, The Netherlands: Elsevier. https://core.ac.uk/download/pdf/4814091.pdf

Pao, Miranda Lee and Dennis B. Worthen. 1989. "Retrieval Effectiveness by Semantic and Pragmatic Searching." *Journal of the American Society for Information Science* 40, no. 4: 226–235. https://doi.org/10.1002/(SICI) 1097-4571(198907)40:4<226::AID-ASI2>3.0.CO;2-6

Pao, Miranda Lee. 1993a. "Perusing the Literature via Citation Links." *Computers and Biomedical Research* 26, no.2: 143-56. https://doi.org/10.1006/cbmr.1993.1009

Pao, Miranda Lee. 1993b. "Term and Citation Retrieval: A Field Study". *Information Processing and Management* 29, n. 1: 95–112. https://doi.org/10.1016/0306-4573(93)90026-A

Pearson, David. 2010. "Bibliographical Control." In *The Oxford Companion to the Book*, edited by Michael Felix Suarez and H R Woudhuysen. Oxford, UK: Oxford University Press, volume 1: 523.

Schöpfel, Joachim, Dominic Farace, Hélène Prost, Antonella Zane and Birger Hjørland. 2021. "Data Documents." *Knowledge Organization* 48, no. 4: 307-328. DOI:10.5771/0943-7444-2021-4-307.

Sevelsted, Rasmus and Chr. Gorm Tortzen. 2010. "Hvem har skrevet Platons værker?" [Who Wrote the Works of Plato?]. In *Platon: samlede værker I ny oversættelse* II, edited by Jørgen Mejer and Chr. Gorm Thortzen. Copenhagen Denmark: Gyldendal, 9-41.

Sobhan, Shahriar, Timothy Williams, Md Jobair Hossain Faruk, Juan Rodriguez, Masrura Tasnim, Edwin Mathew, Jack Wright and Hossain Shahriar. 2022. "A Review of Dark Web: Trends and Future Directions." In *2022 IEEE 46th Annual Computers, Software, and Applications Conference* (COMPSAC). Los Alamitos, CA: The Institute of Electrical and Electronics Engineers (IEEE), 1780-5. https://10.1109/COMPSAC54236.2022.00283

Swanson, Don R. 1986. "Undiscovered Public Knowledge." *The Library Quarterly* 56, no. 2: 103-18. https://doi.org/10.1086/601720

Taube, Mortimer. 1951. "Functional Approach to Bibliographic Organization: A Critique and a Proposal." In *Bibliographic Organization. Papers Presented before the Fifteenth Annual Conference of the Graduate Library School July 24-29, 1950*, edited by Jesse H. Shera and Margaret E. Egan. Chicago, IL: The University of Chicago Press, 57-71.

Tomer, Christinger. 2017. "World Wide Web (WWW)." In *Encyclopedia of Library and Information Sciences*, 4th ed, edited by John D. McDonald and Michael Levine-Clark. Boca Raton, FL: CRC Press, volume 7, 5019-33. https://doi.org/ 10.1081/E-ELIS4-120053396

Van den Heuvel, Charles. 2009. "Web 2.0 and the Semantic Web in Research from a Historical Perspective: The Designs of Paul Otlet (1868-1944) for Telecommunication and Machine Readable Documentation to Organize Research and Society." *Knowledge Organization* 36, no. 4: 214-26. https://doi.org/10.5771/0943-7444-2009-4-214

Van Hoesen, Henry Bartlett and Frank Keller Walter. 1937. *Bibliography, Practical, Enumerative, Historical: An Introductory Manual*. New York, NY: C. Scribner's Sons.

White, Howard D. and Alesia Ann Zuccala. 2018. "Libcitations, WorldCat, Cultural Impact, and Fame." *Journal of the Association for Information Science and Technology* 69, n.12: 1502-12. https://doi.org/10.1002/asi.24064

Wilson, Patrick. 1968. *Two Kinds of Power: An Essay on Bibliographical Control*. Berkeley and Los Angeles: University of California Press.

Wilson, Patrick. 1989. "The Second Objective." In *The Conceptual Foundations of Descriptive Cataloging*, edited by Elaine Svenonius. San Diego, CA: Academic Press, 5-16.

Wright, Alex. 2014. *Cataloging the World: Paul Otlet and the Birth of the Information Age*. Oxford/New York: Oxford University Press.

Zheng, Qinghua, Zhaohui Wu, Xiaocheng Cheng, Lu Jiang and Jun Liu. 2013. "Learning to Crawl Deep Web." *Information Systems* 38, n.6: 801-19. https://doi.org/10.1016/j.is.2013.02.001