

THE LIBRARY QUARTERLY

Volume 58

APRIL 1988

Number 2

THE AACR2 AS A DESIGN SCHEMA FOR BIBLIOGRAPHIC DATABASES¹

Raya Fidel and Michael Crandall²

The Anglo-American Cataloguing Rules are actually a set of rules for a database—a bibliographic database. An examination of the rules from a generalized database approach, using the entity-relationship model, shows that each rule belongs to at least one of six types: content; establishing entities, relationships, or attributes; authorized sources; domain; format; and access points. The current arrangement of the rules and their individual composition ignores these types: it scatters rules of the same type and mixes a variety of types within a single rule. This results in confusion between the internal/external and conceptual levels of design and a poorly organized presentation of cataloging rules. A generalized database approach suggests an outline for a better organization of the rules. It also provides insight into a variety of issues, such as the role of the main entry. Further investigations using this approach would improve our understanding of the cataloging code.

Throughout their century-and-a-half-long development, cataloging rules have been “designed to facilitate the construction of files of records of documents” [1, p. 231]. As is evident in this age of automation, the cataloging code is actually a set of rules *essential* to the construction of a bibliographic database. Moreover, the cumulative history of the activities of the cataloging community is the largest pool of documented experi-

1. We thank Terry Brooks, Francis Miksa, Dagobert Soergel, Ellen Soper, Irene Travis, and Hans Wellisch for their insightful and useful comments on an earlier draft.
2. University of Washington, Graduate School of Library and Information Science, 133 Suzzallo Library, Seattle, Washington 98195.

[*Library Quarterly*, vol. 58, no. 2, pp. 123–142]

© 1988 by The University of Chicago. All rights reserved.

0024-2519/88/5802-0001\$01.00

ence in database design we have available in the modern world. To benefit both the code and database design in general, this experience should be tied to the general principles of database design.

The second edition of the *Anglo-American Cataloguing Rules* [2], or the AACR2, is the latest in a long line of cataloging rules, stretching back at least to Anthony Panizzi's ninety-one rules for cataloging works at the British Museum in 1841 [3]. Throughout this history, politics has played nearly as important a part in the development of the rules as has theory [4-6], resulting in an ever-growing collection of critiques and suggestions for improvements of the current set of rules (for example, [1, 7, 8]).

This paper investigates the rules that constitute AACR2 in a new light, using a generalized database point of view. While little has been written about AACR2 or its predecessors from this point of view, many problems discussed in the literature result from a lack of appreciation of this aspect of the rules. As Kathryn Weintraub wrote: "This new edition of the catalog code introduces a great many changes that are necessary because catalog records are now prepared for a shared data base. . . . However, many of the specific details and fine points have not been thought through as carefully as they could have been" [9, p. 442].

In this paper, we analyze the rules within the database framework to determine categories of rules and their arrangement, and we suggest changes in structure that might be helpful in future implementations of cataloging rules as well as in databases generally. As we shall see, a generalized analysis provides a wealth of useful information to database designers in all areas.

A Generalized Database Approach

Before we analyze the AACR2 from a generalized database approach, however, it will be useful to review some of the basic concepts in that approach.

Levels of Design

The design of a database requires many decisions at different levels. For example, it is clear that the decision about whether to include a general material designation in a bibliographic record is different in nature from the decision about the internal codes to use in a bibliographic record for each type of material, or from the decision about the terms to use for a specific designation when it is displayed to a user. The first type of decision determines what elements to *include* in a bibliographic record, while the second defines the actual codes to use in *storing* the informa-

tion, and the third selects the preferred terms to use when an element is *displayed*. In the database literature it is common to distinguish among three types of decisions, or levels of design: the *conceptual*, the *internal*, and the *external* levels [10].

The conceptual level relates to the representation of that part of the real world that the database is about. For example, the part of the real world that an online catalog is about is "all library materials commonly collected at the present time" [2, rule 0.1]. A designer of such a catalog must make decisions about matters such as what elements of information about this material to include (author, parallel title, title proper of series, and so forth), how to determine what a title of a published work is, whether an author's name should be recorded as it appears on the chief source of information or be checked against an authority list, and how many titles a published work may have.

The internal level relates to how the data are actually stored. The design of an online catalog, for example, would include decisions about the medium on which data are stored, the format in which data are stored, the indexes (for example, whether to create one index for all terms, subject headings, authors' names, and words in titles alike, or whether to create a separate index for each type of term), and the internal organization of records. All these decisions are on the internal level.

The external level relates to the particular views of the data, which are geared to specific purposes. Examples of decisions on the external level for an online catalog are: a means to express a request (for example, the complete title or just the first three words); a means to ask questions in a language that the computer "understands" (an interface language); a picture of the arrangement of the data (for example, bibliographic file, authority file); a picture of the possible manipulations of data (for example, a combination of author and title is allowed but not of title and subject heading); and formats for display of answers.

Roughly, the conceptual level deals with the theoretical aspects of the design, while the internal and external levels deal with the physical aspects. With these concepts we can now present two important principles of database design (fig. 1).

The principle of data independence requires that the decisions on the conceptual level should be independent of both the internal and the external levels [11, p. 25]. In other words, it requires that the theoretical decisions in database design be independent of the physical ones. Applying this principle to online catalog implies, for instance, that the decision whether to store general material designations in codes (the internal level) or the decision about how these designations will be displayed to

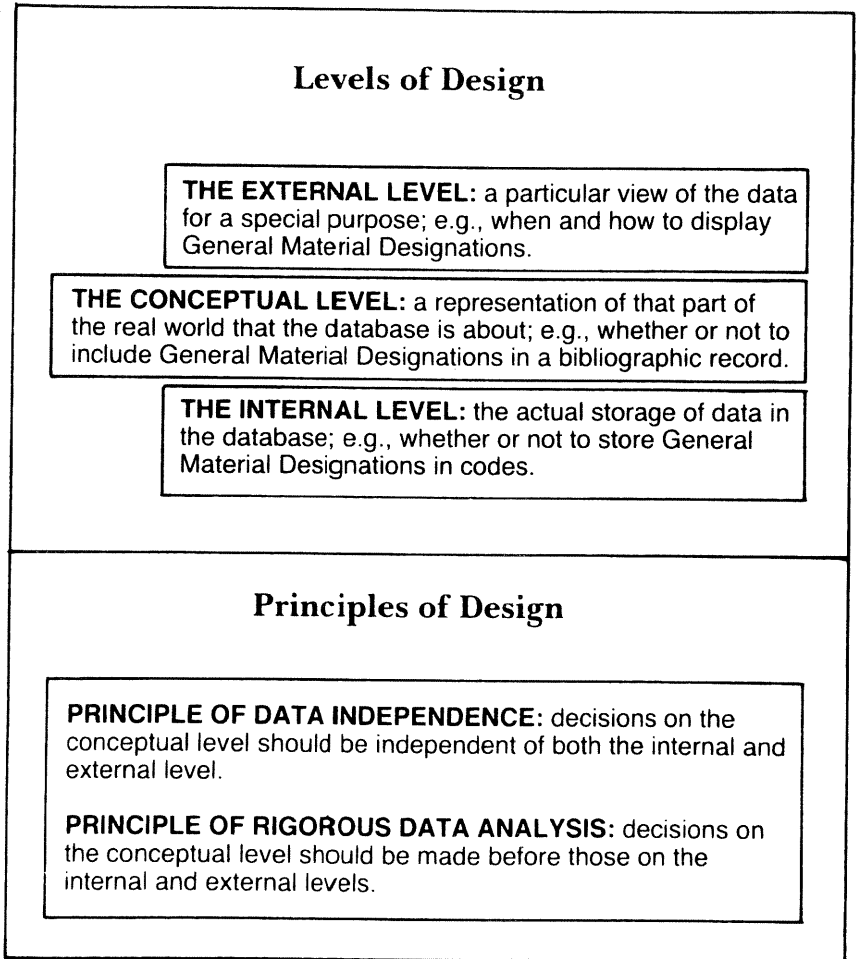


FIG. 1.—Levels and principles of database design

users (the external level) should not affect the decision about whether to include material designations in the catalog (the conceptual level).

The principle of rigorous data analysis requires that the decisions on the conceptual level should be made *before* those on the internal and external levels [11, p. 25]. It is on the conceptual level that one decides, for example, whether to include general material designations and what the authorized terms are to use for this purpose, and this in turn would determine how to store these designations (the internal level) and how to display them to users (the external level).

Designers of printed catalogs—whether in card, book, or microform—may at first glance find these principles unhelpful. Here, it is difficult to distinguish between the internal and the external levels of design. The internal and external records are almost one and the same, whatever the cataloger types on a card and whichever way the cards are organized for browsing and retrieval. However, when typing a card, catalogers may benefit from keeping in mind that their cataloging is performed on two distinct levels at the same time; at each decision point they can consider explicitly the means to store information efficiently (the internal level) as compared with the means to display information to users clearly (the external level). In addition, the distinction between the internal and the external levels is important for the design of any machine-aided catalog, whether printed or online. Thus, the separation between the conceptual level and the other levels is essential to the design of any kind of catalog.

The Entity-Relationship Model

Decisions on the conceptual level usually involve the development of a model that represents that part of the real world that the database is about. One of the tools that is used for this purpose is the entity-relationship model; it has been gaining recognition as a model for a proper expression of “pieces” of data and their relationships that are specific to a database. This model was first introduced in 1976 by Peter Pin-Shan Chen [12] and has since attracted much attention. It is an analytical tool for the representation of data in a database that is aided by a graphical display: the entity-relationship diagram. The model is based on three fundamental concepts.

Entity.—An entity is a “thing” that can be distinctly identified. Any distinguishable object, real or abstract, that is to be represented in a database is an entity. Thus, the book *Cataloging U.S.A.* is an entity; its title—*Cataloging U.S.A.*—is an entity, and its publisher—American Library Association—is yet another entity. The concept that is relevant to database design, however, is entity type, which is a set of entities of the same type, such as work, title, or publisher. For simplicity’s sake, the term *entity* is used here to express the concept of *entity type*. In the entity-relationship diagram (fig. 2), entities are represented by rectangular boxes.

Relationship.—A relationship is an association among entities. For example, “published by” is a relationship between the entities *work* and *publisher*. In the entity-relationship diagram, relationships are represented in diamond-shaped boxes.

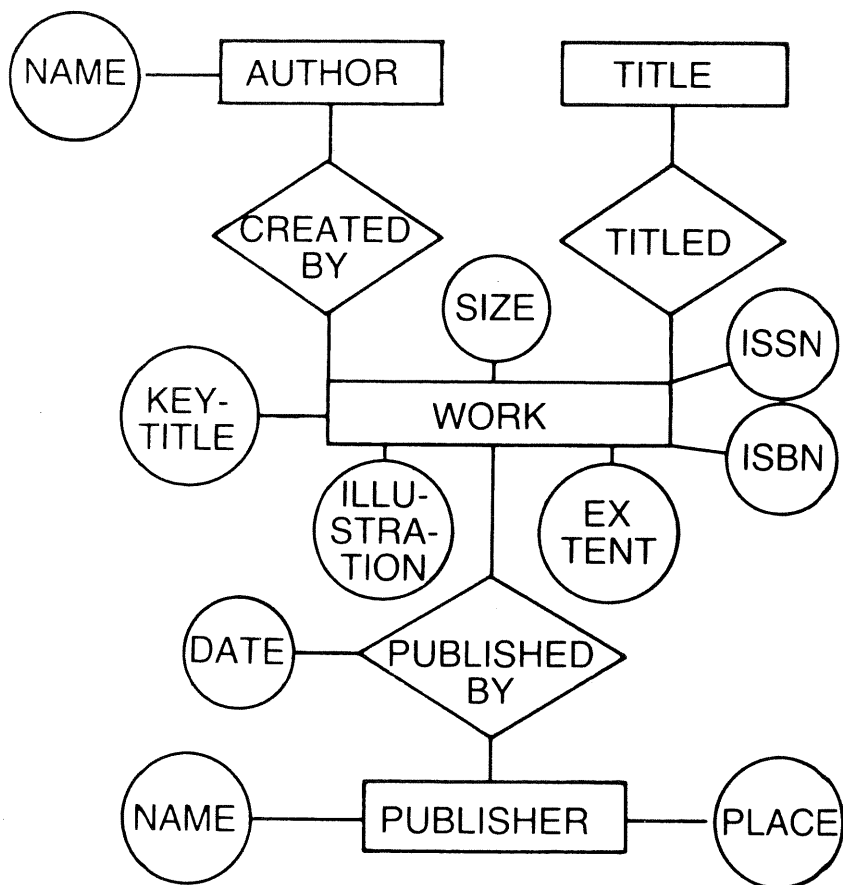


FIG. 2.—A graphic display of the entity-relationship model for some elements in a bibliographic record.

Attribute.—An attribute is a piece of information about an entity or about a relationship. *Name*, for example, is an attribute of the entity *publisher*, and so is *place*. *Date* of publication, however, is an attribute of the relationship “published by.” Attributes are represented by circles.

Figure 2 is an example of a graphic representation of a few entities from the bibliographic world with their relationships and attributes. Some readers may not agree with the specific configuration we have selected. Please bear in mind, however, that a complete representation of the bibliographic world requires extensive teamwork. This figure, on the other hand, is only an example to illustrate the nature of the entity-relationship diagram.

While the adequacy of the entity-relationship model as an analytical tool for database design is still controversial, the model has proven to be a useful means for the analysis of the structure of AACR2. For that purpose, two additional concepts need to be defined.

Attribute values and domain.—In building a database, one examines real-life objects and facts and represents them with a string of symbols in the database. In other words, attributes of, say, the book *Cataloging U.S.A.*—such as the size, the extent, or the nature of its illustrations—are each represented with a string of symbols. A string of symbols to represent an attribute of a particular entity is called a *value* of the attribute. Thus, “23 centimeters” is a value of the attribute *size*, and “xxii + 159” is a value of the attribute *extent*.

To facilitate the orderly and efficient construction and operation of a database, it is necessary to describe for each attribute what kind of values it can take, or, what its valid values are. The set of valid values for an attribute is called the *domain* of the attribute. This means that the values of an attribute are drawn from the domain of the attribute and from no other source. The instructions in rule 1.8B1 about how to record the International Standard Book Number (ISBN) and the International Standard Serial Number (ISSN) are examples of a definition of the domains of these attributes: “Give such numbers with the agreed abbreviation and with the standard spacing or hyphenation.”

With these basic concepts of a database approach explained, we now turn to our analysis of the structure of the AACR2.

A Typology of Rules

Using the entity-relationship model as a conceptual framework, we first analyzed the AACR2 to identify the types of rules that are included. In particular, we examined part 1, chapter 1 (General Rules for Description) and part 2, chapters 21–25 (Headings, Uniform Titles, and References). In part 1 we elected to limit our examination to chapter 1 because it is the framework on which most of the other chapters in this part are placed. We found that each rule (that is, each numbered section in AACR2) in these chapters belongs to one or more of the following categories or types of rule.

Content Rules

Content rules explicitly name the elements that should be included in a record, or in a part of it. One example of such a rule is rule 1.0D1: “For the first level of description, include at least the elements set out in this

schematic illustration." Other rules of this type may relate to more specific areas, such as rule 1.4B1, which is about the area that includes publication and distribution: "This area is used to record all information about the place, name, and date of all types of publishing, distributing, releasing, and issuing activities." Content rules actually determine which entity types, relationship types, or attribute types to establish in a database.

Rules for Establishing Entities, Relationships, or Attributes

The questions of whether a certain object qualifies as an entity, or whether a certain relationship indeed exists among two entities, are addressed by these rules. Rule 1.1B8 is an example of a rule to establish the entity *title*: "If the title proper appears in two or more languages or scripts, record as the title proper the one in the language or script of the main written, spoken, or sung content of the item." It instructs catalogers about how to decide whether a certain real-life string of characters is a title. In addition, most of the definitions in the AACR2 glossary [2, app. D] are examples of rules for establishing entities.

An example of a rule for establishing relationships is rule 21.1A1: "A personal author is the person chiefly responsible for the creation of the intellectual or artistic content of a work." It guides catalogers in making decisions about whether the relationship "created by" should be established between a given work and a given person.

Rules for establishing attributes are similar in their nature to rules for establishing entities or relationships. Rule 1.4C5 is an example of a rule to establish the attribute *place* of publication: "If a publisher, distributor, etc., has offices in more than one place and these are named in the item, always give the first named place." Specifically, the rule delineates a selection rule when more than one place can qualify as the place of publication. Rule 1.4F3, on the other hand, is a rule to establish an attribute of a relationship; it helps in determining the *date* on which a work was "published by" a *publisher*: "Give the date of a particular reissue of an edition as the date of publication only if it is specified in the edition area."

Rules for Authorized Sources

These rules determine what the authorized sources of information are. They may specify a particular source, such as in rule 2.0B1—"the chief source of information for printed monographs is the title page"—or they may allow for a flexible choice, such as in rule 1.5A2, which is about the physical description area: "Take information for this area from any source." Rules that provide instructions for cases in which the authorized source is lacking also belong to this category, such as rule

1.1B7: "Supply a title proper for an item lacking the prescribed chief source of information or its substitute from the rest of the item, or a reference source, or elsewhere."

Rules for Domain

Rules for domain *define* the valid values of an attribute and guide decisions in cases that are exceptions. A simple example of a rule for domain is rule 1.5B1: "Record the number of physical units of the item being described by giving the number of parts in arabic numerals." An example of a rule that resolves an exceptional case is rule 1.1F14: "Transcribe a statement of responsibility even if no person or body is named in that statement."

Rules for Format

The way data are actually recorded is determined by rules for format. Such rules may address a variety of issues: they may dictate a certain *sequence* in which data are recorded, they may provide a means of separating one element from another and thus provide *demarcation*, they may determine the *form* in which information is to be recorded, or they may designate the *location* in which the information about a certain element should be recorded.

An example of a format rule that determines a sequence is rule 1.7B: "Give notes in the order in which they are listed here." Another example of a format rule is rule 1.0C, which provides the demarcation for areas in a catalog card: "Precede each area, other than the first area, or each occurrence of a note or standard number, etc., by a full stop, space, dash, space (. —) unless the area begins a new paragraph."

Rule 1.1B6, on the other hand, is about the form of the title proper: "If a title proper includes separate letters or initials without full stops between them, record such letters without spaces between them." Last, an example of a rule about the location in which elements should be recorded is rule 1.1E5: "Transcribe other title information following the title proper or parallel title to which it pertains." It should be noted here that since the cataloging code is historically geared to printed catalogs, location is frequently designated by a preceding element of data. Therefore, the distinction between sequence and location is not always clear.

Rules for Access Points

The creation of indexes to a database is guided by rules for access points; these rules indicate the elements to be used as access points. Rule 21.1A2, for example, requires the creation of an author index: "Enter a work by one or more persons under the heading for the personal author."

The Appendix presents the rules in chapter 1 grouped according to their categories. The relationship between the levels of design and the types of rule is discussed later.

The Structure of AACR2

Using the typology of rules as an analytical tool, we examined the structure of the AACR2. Our objective was to determine how rules of the same type are distributed across chapters, within each chapter, and within each rule.

The Distribution of Types of Rules

Our first question in the analysis of the structure of AACR2 was whether there is a relationship between the types of rule and the chapter in which they are listed. We discovered a clear pattern.

Chapter 1 (General Rules for Description) includes rules of all types, except those for access points. Chapter 21 (Choice of Access Points) is about access points but is also composed of rules for establishing entities and relationships. More precisely, it includes rules for establishing the entities *author* and *title*, and the relationships "created by" and "titled." Chapter 22 (Headings for Persons) is primarily about rules for domain, with some "islands" of rules for establishing entities, such as rules 22.1A, 22.1B, and rules 22.2A to 22.2C4. Chapters 23 (Geographic Names) and 24 (Headings for Corporate Bodies) are clearly about domains, and so is chapter 25 (Uniform Titles), with a few exceptions.

The overall pattern is clear: a general section (chap. 1) includes most types of rule, while rules to establish specific entities, relationships, or attributes and rules for access points, as well as rules for domains of specific attributes, are collected in part 2 (chaps. 21–25).

Since chapter 1 includes rules of most types, we examined the concentration of types of rule in the chapter. Our analysis shows that as many as 100 AACR2 rules include statements that are content rules. It is important to note that 43 percent of these statements are about elements that should be included in the notes area. The next category is format rules; eighty-one rules include statements of that nature. Rules for domain appear in sixty-eight rules, and those for establishing entities, relationships, or attributes appear in forty-six rules. Last, thirteen rules include statements about authorized sources.

The Distinction among Types of Rules

Our second question in this analysis was whether the type of rule has an effect on the arrangement of the rules in chapter 1 and on the arrange-

ment of statements within each rule. We found no evidence of such an effect. Moreover, it seems that the current arrangement of the rules and their composition completely ignores the types of rule that are involved.

This finding is supported by two trends. First, rules of the same kind are often scattered. For instance, 1.1B8, 1.1B10, and 1.1G1 are rules about how to establish title proper. Second, it is not uncommon for one rule to relate to a variety of categories; 41 percent of the rules in chapter 1 are mixed rules. As an example of a mixed rule, consider rule 1.4F1: "Give the date of publication, distribution, etc., of the edition named in the edition area. If there is no edition statement, give the date of the first edition. Give dates in Western-style arabic numerals. If the date found in the item is not of the Gregorian or Julian calendar, give the date as found and follow it with year(s) of the Gregorian or Julian calendar." It includes four categories: authorized sources (first sentence), establishing entities and relationships (second sentence), domain (third sentence); and format (last sentence).

We discovered further that, at times, one statement within a rule appears as a rule of a definite type yet implicitly contains a rule of a different type. Rule 1.1F8 is an illustration of this phenomenon: "Add an explanatory word or short phrase to the statement of responsibility if the relationship between the title and the person(s) or body (bodies) named in the statement is not clear." On the surface this seems to be a content rule: it instructs catalogers to add an explanatory statement and thus introduces a new element to the record of a work. This rule, however, is important for establishing the relationship "created by"; it implies that even if the relationship between a work and an author is not clear, the statement of responsibility should appear—that is, the relationship should be established.

Discussion

What can be learned from these findings about the structure of the AACR2 and the arrangement of the rules? Our conclusion is that this arrangement is incompatible with the principles of database design. In addition, these principles can be used to suggest a preferred structure and to analyze a variety of issues, such as the role of the concept *main entry*. These observations are elucidated below.

Types of Rules and Levels of Design

To examine whether the structure of the AACR2 follows a database approach, we analyzed the arrangement of the rules in chapter 1 in light of the two principles of design mentioned above.

To do so, we looked for a correspondence between the levels of design and the types of rule. An operational criterion useful for identifying the level of design is to assign to the conceptual level types of rule that are not likely to change when moving from a printed catalog to an online catalog. Types of rule that depend on the nature of the catalog employed belong to the internal or the external levels. Using this criterion, it is easy to see that all the types of rule belong to the conceptual level, except for two: rules for format and for access points belong to the internal level of design. Content rules, rules for establishing entities, relationships, or attributes, and those for authorized sources, as well as rules for domain, are all answers to anticipated questions about how to represent library materials and are independent of the type of catalog used. In contrast, rules for format and for access points answer questions about how actually to record the elements that were *already selected*; printed catalogs require rules that are different from those for online catalogs.

In analyzing format rules in AACR2, it is often difficult to distinguish between the internal and the external levels. This is not a surprise because the code has been developed for printed catalogs in which the two levels are one and the same. For simplicity's sake we have assumed that these rules belong to the internal level. This in turn leads us to conclude that AACR2 does not include rules on the external level.

The first principle of database design, the principle of data independence, requires that decisions on the conceptual level must be separated from those on the internal and the external levels. The importance of this principle for cataloging is reinforced by the fact that the same set of rules is used for both printed and online catalogs. Surely, the internal record of, say, a card catalog is very different from an internal record of an online catalog. Moreover, internal records of online catalogs vary according to the system used. Therefore, for cataloging rules to be easily used by *all* catalogers, it is important to keep format and access points rules clearly separated from other rules. One may even envision future cataloging rules where format rules appear in distinct sections (whether among other types of rule or completely separated), each corresponding to the type of internal record employed.

The structure of AACR2 ignores this principle; format rules are not designated as such and are listed in the same sections with rules of another nature (see App.). Moreover, 60 percent of the format rules in chapter 1 appear in mixed rules. This issue is even further complicated in some rules because they include statements that appear to be format rules but implicitly include other types as well.

Rule 1.8C1 is an example of this complication. It is the main rule about key-titles, and it states: "Add the key-title of a serial, if it is found

on the item or is otherwise readily available, after the International Standard Serial Number (ISSN)." On the surface, this part of the rule is a location rule; it is about the place to record a key-title: after the ISSN. In essence, however, it informs the cataloger that a key-title can be included in the record, which is a content rule. It also indicates when it should be included ("if it is found on the item or is otherwise readily available"), which is a rule for establishing an attribute. In addition, it specifies the authorized sources for information about this attribute—the item itself or any other source.

The second principle of database design, the principle of rigorous data analysis, requires that one must first make decisions on the conceptual level and only then on the internal or external levels. Thus, while the first principle of database design points to a flaw in the structure of AACR2 and in the arrangement of rules, the second principle actually suggests a preferred arrangement of sections within a chapter, and of statements within a rule: list first the rules that belong to the conceptual level and then those that belong to the internal and external levels.

The practice of cataloging can also suggest the order in which rules *within* the conceptual level should be listed. Faced with a work to be cataloged, a cataloger first asks what elements about it should be included in a bibliographic record. Content rules answer such a question. Then, a cataloger decides which part of the information about a work is actually representing each element: who is responsible for the work, what its title or edition is, and so forth. Rules for establishing entities, relationships, or attributes—including the definitions of each element—have been formulated to guide such decisions. Next, one must find out what the authorized sources of information are for each element; rules for authorized sources help a cataloger to ascertain that the information available is indeed the "correct" information. Once an attribute has been established, rules for domain are consulted to determine how the attribute should be represented. This cycle—content, establishment, authorized sources, and domain rules—includes all the decisions that are necessary on the conceptual level.

Rule 1.8C1 can serve again as an example to demonstrate a preferred order. The last two sentences of this rule give instructions about how to establish a key-title: "Give the key-title even if it is identical with the title proper. If no ISSN is given, do not record the key-title." Thus, this rule starts with an apparent format statement that implicitly includes rules of other types and then provides rules for establishing the attribute *key-title*.

The database approach suggests that such a rule should be organized differently. First, the rule would explicitly state that a key-title could be included in a record. Next, it would mention the definition of key-title and the rules for its establishment—that is, when it is readily available,

even if it is identical with the title proper, but only if the ISSN is given. Then the sources of information would be cited—that is, the item or any other sources. The format rule (namely, that the key-title should be recorded after the ISSN) should be added in a separate section because catalogers who make no decisions about formats may find the section irrelevant for their cataloging, and others may require a completely different set of instructions to fit the format used in their system.

While rule 1.8C1 may not be typical, it illustrates clearly that the structure and the arrangement of the AACR2 do not follow a database approach and that such an approach proposes a preferred arrangement—one in which rules are stated explicitly and in a sequence that is useful for the practice of cataloging.

The Concept of Main Entry

The database approach can provide insights into general issues relating to the cataloging code. The concept of main entry is an example of such an issue.

Although the compilers of AACR2 considered abandoning the concept of main entry, the change was not made because of the “lack of time to explore the considerable implications of such a change” [2, rule 0.5]. However, the suggestion is made in the same rule that those libraries not wishing to use the concept of main entry may instead use the rules in chapter 21 as “guidelines in determining all the entries in particular instances.” While obviously meant as a compromise of sorts, this rule stirred up considerable reaction from those who felt it had gone too far (for example, [13]) or not far enough (for example, [14]).

Thus, while chapter 21 is explicitly dedicated to the concept of entry (as its title indicates), it is not clear whether it can keep its integrity when the concept of main entry is removed. In an attempt to clarify this issue by eliminating the main entry concept from chapter 21—as suggested in rule 0.5—Baughman and Svenonius found that not only did the rules become confusing, but the collocating objective of cataloging was violated [15]. In other words, the concept of main entry, whether or not explicitly stated, is central to chapter 21.

The analysis of AACR2 from a database approach sheds new light on this issue. It shows that the role of chapter 21 is central to the cataloging code and, therefore, the structure of AACR2 is strongly determined by the concept *entry* (or access points), and in particular by that of *main entry*.

As our findings show, chapter 1 includes a relatively small number of rules for establishing entities, relationships, and attributes: only forty-six rules, or 24 percent, include statements that are rules of that type. On the other hand, these rules provide the basic definitions for the elements to be included in a bibliographic record and are formulated to resolve

any "irregularity" that occurs in works to be cataloged. When formulating such rules, one anticipates all the problems that catalogers might encounter when they decide who an author is, what the title, edition, or publication date is, and so forth. In a way, such rules are most basic to cataloging of any kind. Moreover, experience shows that the large diversity in library materials requires a relatively large number of such rules. It stands to reason, therefore, that rules for establishing entities, relationships, or attributes would constitute a considerable portion of AACR2. Yet chapter 1 includes a relatively small number of these important rules.

In addition, if elements in bibliographic records were ranked by their "importance," *author* and *title* would rank the highest. After all, while several elements, such as ISBN, illustrations, or key-title, may be eliminated from a bibliographic description (as is the practice of many abstracting and indexing services), no such description can exist without author and title. Yet chapter 1 includes only six rules about the establishment of authors and six about the establishment of titles. Moreover, the chapter does not include the basic definitions about how to determine who is responsible for a work, or what a title proper is. Questions that catalogers might have when a work seems to have no author, when it is attributed to more than one author, or when a title changes between one edition and another are not answered in this general chapter. Most of the rules for establishing the entities *author* and *title*, and the relationships "created by" and "titled," are listed in chapter 21. Thus, the most basic rules for the most important elements in a bibliographic record are listed in chapter 21, the chapter about the choice of access points.

It is claimed that the structure of the AACR2 is based on normal cataloging procedure—describing a work in hand and then determining access points [14]. Yet, the inconvenience in going back and forth from chapter 1 to chapter 21 is apparent. The database approach, however, shows that this inconvenience is not a matter of a slight shortcoming but rather the result of a method that centers around the concept of main entry: a number of rules that address the most basic and frequent problems in cataloging are explicitly main-entry rules—they only *imply* how to establish entities or relationships.

Using database concepts, we observe that the main entry is viewed by AACR2 as yet another entity. This entity, however, is of a peculiar nature: it is not created to represent data from the real world but rather to provide access points, and it includes other entities in it, such as *author* or *title*. Moreover, the basic rules for the entities that might be included in the entity *main entry* are presented as rules for the main entry rather than as rules for the entities themselves. Consider, for example, the rule that determines who the author of a work is (rule 21.4A), the rule that

guides catalogers in their decision about an author when the responsibility for a work is erroneously attributed to a person (21.4C1), or the rules about selecting an author for works of shared responsibility (21.6B–21.6D). These are rules for main entry; they were formulated to instruct catalogers about the selection of a main entry and only imply how to establish the relationship “created by” for such works.

The centrality of the main entry approach stands out because, while rules for establishing entities, relationships, or attributes are on the conceptual level of design, rules for access points belong to the internal level—they indicate the type of indexes to be constructed. Burying rules from the conceptual level in internal rules indicates a strong inclination toward emphasizing these internal rules—the main-entry rules.

It is clear, therefore, that a new view of the concept of access points is required for AACR2 to be compatible with a database approach. To develop such a view would require a thorough investigation and a reorganization of the code—a major task for future research.

Suggestions for Further Research

It is useful to apply the database approach to the analysis of the AACR2 because in addition to its ability to point out flaws in the structure of AACR2, it suggests avenues for research that in turn could improve the cataloging code.

The first project in this direction would require a further refinement of the typology of rules. Such a refinement would be based on the classes of anticipated problem that rules of each type are supposed to solve.

Rules for authorized sources, for example, could be further subdivided into classes according to the following criteria. They all must include a general statement indicating the authorized source of information for each element in the bibliographic record. They also have to include, however, a rule for works that lack the authorized source and a rule for works that have more than one authorized source. In addition, if information about an element can be taken from a number of sources, those have to be ranked according to their preference.

While subdividing rules for authorized sources is a relatively straightforward task, rules of other types are more complex. One can identify subclasses in other types of rules, however, by analyzing individual rules in the AACR2 and explicitly stating the problem that each rule is supposed to solve.

Consider, for instance, rule 1.1F4: “Record a single statement of responsibility as such whether the two or more persons or corporate bodies named in it perform the same function or different function.”

This is a rule to establish the relationship "created by," and it anticipates a particular case: when two or more persons, with varying responsibilities, could qualify as authors. On a more general level, one can deduce that rules for establishing entities, relationships, or attributes include a class of rules that resolve a fuzzy situation: when several real-life objects or facts can qualify—and to varying degree—as an entity, a relationship, or an attribute.

Another example is rule 1.1F12: "Treat a noun phrase occurring in conjunction with a statement of responsibility as other title information if it is indicative of the nature of the work." This rule is again of the establishing type but it points to a borderline case, when a single real-life object or fact can be identified as either one entity (or relationship or attribute) or another. Thus, establishing rules include a class of rules for borderline cases.

A complete and detailed typology of rules can be created when a large enough number of rules is analyzed in this fashion. Such a typology can be used to insure the completeness of the cataloging code and to determine the structure that would be most useful for catalogers.

The ability of the cataloging code to provide a rule for most problems a cataloger might encounter is improved when the typology of rules is applied to each element in the bibliographic record. Thus, each element would have at least three rules about authorized sources: a general rule, a rule for lack of a source, and one for more than one source. Similarly, each element should have at least two types of establishing rule: for fuzzy cases and for borderline cases. Checking each element against a comprehensive checklist of possible problems would uncover "missing" rules and would enhance the completeness of the cataloging codes.

When all types of rule and their subclasses are explicitly listed, a preferred structure can be decided upon on a general level. This structure can then be applied to the code as a whole and to the rules for each element. To determine a useful structure would require some experimentation. It is not clear, for example, whether it is more useful to have one type of rule follow another for each element or to list first all the basic rules for an element—covering straightforward cases—and only then list rules that resolve special problematic cases.

The refined typology of rules, however, facilitates the testing of possible arrangements. Only when the arrangement of rules is based on a predefined and general structure, such as the typology, can one proceed systematically in testing for the most useful arrangement. Further, with a predefined structure, one can test the arrangement of rules for only a few elements and generalize the results to all the elements in the bibliographic record and to nonbibliographic databases.

Using a typology of rules based on anticipated problems to organize

REFERENCES

1. Shinebourne, J. A. "A Critique of AACR2." *Libri* 29 (October 1979): 231-59.
2. Gorman, Michael, and Winkler, Paul W., eds. *Anglo-American Cataloguing Rules*, 2d ed. Chicago: American Library Association, 1978.
3. Maxwell, Margaret F. "The Genesis of the Anglo-American Cataloging Rules." *Libri* 27 (September 1977): 238-62.
4. Tikku, Upinder Kumar. "Anglo-American Cataloging Rules, 1908-1978: A State of the Art." *Annals of Library Science and Documentation* 30 (September-December 1983): 151-65.
5. Hagler, Ronald. "Why AACR2?" In *AACR2 Seminar Papers*, edited by Ralph W. Manning. Ottawa: Canadian Library Association, 1981.
6. Lewis, Peter R. "The Politics of Catalog Code Revision and Future Considerations." In *The Making of a Code: The Issues Underlying AACR2*, edited by Doris Hargrett Clack. Chicago: American Library Association, 1980.
7. Musavi, Nurieh. "An Evaluation of AACR2." *Library Resources and Technical Services* 30 (April/June 1986): 137-48.
8. Rush, James E. "AACR2 from an Information Scientist's Point of View." *Cataloging & Classification Quarterly* 3 (Winter 1982/Spring 1983): 47-56.
9. Weintraub, Kathryn D. "AACR2: A Review Article." *Library Quarterly* 49 (October 1979): 435-43.
10. Tschritzis, D., and Klug, A. *The ANSI/X3/SPARC DBMS Framework Report of the Study Group on Data Base Management Systems*. Montvale, N.J.: AFIPS Press, 1977.
11. Martin, James, and McClure, Carma. *Structured Techniques for Computing*. Englewood Cliffs, N.J.: Prentice-Hall, 1985.
12. Chen, Peter Pin-Shan. "The Entity-Relationship Model—toward a Unified View of Data." *ACM Transactions on Database Systems* 1 (March 1976): 9-36.
13. Lubetzky, Seymour. "AACR2: Main Themes." In *The Making of a Code: The Issues Underlying AACR2*, edited by Doris Hargrett Clack. Chicago: American Library Association, 1980.
14. Gorman, Michael. "The Fundamentals of Bibliographic Cataloging and the AACR2." In *The Making of a Code: The Issues Underlying AACR2*, edited by Doris Hargrett Clack. Chicago: American Library Association, 1980.
15. Baughman, Betty, and Svenonius, Elaine. "AACR2: Main Entry Free?" *Cataloging & Classification Quarterly* 5 (Fall 1984): 1-15.
16. Shaw, M., et al. *Using AACR2: A Diagrammatic Approach to Part II of the Anglo-American Cataloguing Rules, Second Edition*. Phoenix, Ariz.: Oryx, 1980.
17. Hjerpe, Roland, and Olander, Birgitta. "Cataloging and Expert Systems: AACR2 as a Knowledge Base." *Journal of the American Society for Information Science*, May 1988.