Notes on Operations

FRBR Principles Applied to a Local Online Journal Finding Aid

By Chew Chiat Naun

This paper presents a case study in the development of an online journal finding aid at the University of Illinois at Urbana–Champaign (UIUC), with particular emphasis on cataloging issues. Although not consciously designed according to Functional Requirements for Bibliographic Records *(*FRBR*) principles, the Online Research Resources (ORR) system has proved amenable to* FRBR *analysis. The* FRBR *model was helpful in examining the user tasks to be served by the system, the appropriate data structure for the system, and the feasibility of mapping the required data from existing sources. The application of the* FRBR *model to serial publications, however, raises important questions for the model itself, particularly concerning the treatment of work-to-work relationships.*

The University of Illinois at Urbana-Champaign's (UIUC) Online Research Resources (ORR) registry (www.library.uiuc.edu/orr) is a database-driven, alphabetical list of online resources similar in principle to comparable lists provided by vendors such as Serials Solutions and TDNet. ORR is, in effect, an alternative or supplementary catalog for specialized access to online resources, especially electronic journals. Like other tools of its kind, it was designed partly to overcome some of the drawbacks of online catalogs in dealing with this class of material. Antelman says that such tools "are potential sources of innovation because they are amenable to experimentation in ways that our current integrated library systems are not."[1] UIUC's experience with ORR is a case study of a home-grown system built to local specifications.

ORR was not the first system of its kind developed by the UIUC library. An earlier electronic resources registry had been in existence for some years, but the acquisition of a data feed from TDNet in 2003 provided the impetus to redevelop the service. The TDNet service monitors a range of providers and notifies the library of any change either in content or location (URL, or uniform resource locator). Although TDNet normally supplies a public interface, the library chose to develop its own. The development work was undertaken by the library's systems office with the guidance of a committee comprised of staff from systems, public services, and technical services. The new version was built on a redesigned data structure capable of incorporating additional data from external sources. While the redevelopment was primarily intended to facilitate maintenance of the data by library staff, it also made possible significant improvements in the public interface.

ORR is not intended to be a comprehensive catalog of the library's electronic holdings. Its scope is limited to online article databases, journals, and reference works. The majority of electronic books were excluded on the principle that book-like objects were more appropriately represented in the library's online catalog. Each of the categories of resources covered by ORR presented its own metadata and interface design challenges. However, the most urgent—and in some ways

Chew Chiat Naun (naunc@uiuc.edu) is Senior Coordinating Cataloger, University of Illinois at Urbana–Champaign.

Submitted March 13, 2006; accepted for publication April 30, pending revision; revision submitted August 10, 2006, and accepted for publication.

the most complex—task for the developers of ORR was to facilitate access to online journals. Jones describes this class of publications as "the subset [of continuing resources] characterized by issues containing contributions by individual authors, the subset that is most often analyzed in abstracting and indexing services."[2] They will be the main focus of this paper.

At the time of writing, ORR listed 42,640 online journals, 1,344 reference works, and 439 article databases. These totals are for unique titles; the number of unique URLs is much higher. In each of the two years ORR has been operational, it has logged between four and five million hits, counting only links through to full-text content. ORR is a key resource for the library's patrons.

The technical and logistical aspects of ORR's development have been described by German, Shelburne, and Norman, and the reader is invited to consult their publications for additional information about this project.[3] The literature on the management of online journals is extensive, including the provision of access through online journal finding lists similar to ORR. Although several articles provide illuminating details about the database structures employed in these systems, relatively little appears to have been published dealing with the bibliographic relationships in particular.[4]

This paper examines the data structures employed in ORR with respect to bibliographic relationships among serial works, versions, and aggregates. These relationships are described in the International Federation of Library Association's (IFLA) *Functional Requirements for Bibliographic Records* (*FRBR*).[5] Although ORR was not designed with the *FRBR* model specifically in mind, its development was informed by many of the same considerations that underlie that model. The *FRBR* model provides a context to understand specific decisions made in creating ORR, including the compromises involved and areas where improvements may be sought in future.

This paper does not attempt to cover all aspects of ORR's design. For example, since its launch, ORR has been augmented with a rights management module and now more closely resembles a comprehensive electronic resources management system. These newer developments, and their relationship to the cataloging data in ORR, are beyond the scope of the present paper.

User Tasks

The *FRBR* report ascribes to the end-user the following tasks: to find documents matching a given set of criteria, to identify those that are relevant, to select the desired or available versions, and to obtain them.[6] *FRBR* also recognizes the need in some contexts to navigate between resources.[7] This breakdown is useful for understanding the purposes served by various elements of ORR's design. Before attempting an analysis, one must ask the question: exactly what is the user supposed to be trying to find, identify, select, and obtain?

ORR is primarily concerned with bibliographic control and access at the level of the serial publication, not at the level of the individual article. This emphasis reflects that of the traditional library catalog, where Tillett observes, "We cannot afford to always describe and identify every work although that may be the 'ideal'—(sometimes leaving such levels to abstracting and indexing services, sometimes to bibliographies, finding aids, and reference tools)."[8]

This point is well understood by librarians, but not self-evident to patrons. Most of the time, what a patron is interested in is the specific content of a journal article, and an inexperienced patron naturally approaches a tool like ORR with the expectation of finding individual articles directly. As Antelman puts it, "library users' sense of a serial work diverges significantly from the way it is currently implemented in library systems."[9] The identity of a serial work is not always a matter of indifference to the end user. To look no further than their utilitarian role, scholarly journals are an institutionalized part of the system for scholarly dissemination, review, and recognition. ORR includes at least two data elements at the serial-work level that reflect this role: the Institute for Scientific Information impact factor for each title and its peer review status. Nonetheless, the serial title is the primary unit of representation in ORR because it helps users obtain relevant documents. The serial publication is the vehicle of distribution for article content. Data relating to the manifestations and copies (or, in *FRBR* terminology, items) of the serial publication, including the URLs for available sources, coverage dates, and (for print holdings) location and call numbers, enable patrons to obtain copies of the articles they seek.

The role that ORR plays in supporting user tasks may be better understood in the context of concurrent plans at the UIUC library to introduce a broadcast search facility and link resolver. Broadcast searching will facilitate finding and identification tasks by enabling users to search for articles and citations in multiple databases simultaneously. The link resolver will act, where required, as a bridge between the results found in these databases, whether searched simultaneously or separately, and the selection and obtaining tasks jointly supported by the link-resolution knowledge base, serials management system, library catalog, and document delivery service. Part of the original plan for ORR was to serve as a knowledge base for reference linking. This plan was later modified when the library decided to acquire a commercial link resolver with its own knowledge base. Although not completely integrated,

these systems will provide mutually complementary access to the library's online collections.

ORR serves the core function of allowing users to find and identify known serial publications by title, view details of all the available sources (including coverage dates and access restrictions), and make an appropriate selection from among them. The difficulty is that no existing database contains all the requisite data. Even where the data are available, the necessary linkages between them do not always exist. For example, if a complete set of back files is not available for a given journal, the data necessary to locate a library's print holdings may not be obvious from a vendor's database. In order to support the desired user tasks, a system has to pull together disparate data from a range of otherwise unrelated sources and assemble them into a structure that will make displaying and navigating the relevant relationships possible. The data must be easily maintained so that keeping ORR complete and up to date is practical.

ORR is designed around a strategy that satisfies both of these requirements. The strategy is, where possible, to import data from existing sources and use them to populate ORR records according to a quality hierarchy at the level of each individual field. The quality hierarchy ranks the preferred sources for each data element, an approach that allows the database to be populated to the fullest extent possible, while ensuring that data in each field are drawn from the most authoritative, complete, or current source. Thus, while titles are available from both TDNet and the Voyager integrated library system (ILS) data feeds, the ILS source is preferred for title information. Conversely, the TDNet data receive priority for URLs. Automated processes alone cannot ensure the completeness and integrity of all the data; maintaining ORR still requires manual data entry and cleanup.

Functionality and Data

The biggest drawback of the UIUC online catalog as a discovery tool for the library's online collections is its size. At the time of writing, it contains just fewer than five million bibliographic records, compared to 45,000 titles listed in ORR. Although searches may be scoped in various ways, the proportion of unwanted hits inevitably remains high.

The online catalog's functionality has significant limitations, as well. Its proprietary design makes updating using external non-MARC data sources difficult, particularly with the degree of granularity needed. Certain entities conceptually important to the management of electronic resources, such as content providers, are difficult to represent adequately within the confines of the MARC format. The online catalog's ability to manipulate a variety of data into a desired Hypertext Markup Language (HTML) display format is strictly limited. It can collocate alternative versions of a title only to the limited degree that the *Anglo-American Cataloging Rules,* 2nd ed. (*AACR2*) record structures and the system's own relatively inflexible filing and display algorithms permit it to do so.[10] Although the MARC format has provisions for linking between alternative versions of a work and between successive titles in a journal's history, these linking mechanisms are only imperfectly implemented in the catalog's public interface.

The advantages of ORR as an alternative to the online catalog may be seen from a brief outline of the functionality of the system's public interface and some of its specific data elements and design features. Users consulting ORR may search all resources together, or scope their searches by resource type, the latter being recorded in a field in the ORR resource record (figure 1). Titles may be searched for an exact match with implied right truncation or by keyword, with a further option for implied truncation of each word within the title. This latter option is particularly useful for finding abbreviated titles. Searches match on variant titles as well as titles proper, thanks to cataloging data pulled in from the 24X title fields of MARC records.

The interface also allows the user to navigate between earlier and later titles in the serial work's history, providing linked title displays drawn again from MARC data, in this case from the 780 (previous title) and 785 (succeeding title) linking entry fields. Certain other work-level data elements are drawn from a variety of potential sources. International Standard Serial Number (ISSN) data, for example, are compiled opportunistically from TDNet, EBSCO, ILS, and Ulrich's *Periodicals Directory.* Ulrich's is also the usual source for the ISI impact factor and the peer-review status.

ORR's approach to subject access reflects the same priority given to access at the serial-work level. The decision was made very early not to offer generic keyword searching of the database, in spite of the prevalent practice of supplying a keyword option in almost any context. It was decided that keyword searching was suited mainly to the fine-grained subject access and article-level retrieval offered by article and citation databases. The broadcast search interface would be the appropriate place to encourage generic keyword searching. In contrast, the ORR interface was designed to allow very broad subject browsing using an in-house subject descriptor list. To assist in the assignment of these subject descriptors, the ORR database performs mappings from Library of Congress Subject Headings (LCSH) so that the UIUC descriptors are derived automatically on the basis of data in MARC records. UIUC reference librarians may change or add descriptors. They also may add

Figure 1. ORR's main search page

natural-language descriptions viewable by the public for some ORR resources (figure 2).

The database schema for ORR may be seen in figure 3. The data elements previously mentioned reside in the resource record, which is one of ORR's three main building blocks (the other two are the instance record and the interface record). These are invoked to display relevant information about the online sources or providers available for each title (including the dates covered by each source), any access conditions that may apply, and other related information such as current availability. ORR's public display groups information about sources directly under the entry for the relevant title. ORR thus offers a hierarchical display that is supported by a hierarchical record structure (figure 4). The user may toggle to a detailed display (figure 2). These displays are similar to the grouped catalog displays advocated by commentators such as Yee.[11]

The system has three building blocks but only two levels in the display. Most of the pertinent data at the level of the particular source or provider, such as the URL, are stored in the instance record. However, the existence of the interface record reflects the fact that online journals are typically acquired as part of a package that is licensed or purchased together and hosted on a common platform. The interface record often represents a provider, such as Wiley InterScience, but may sometimes represent instead a collection packaged by the provider, such as Wiley InterScience's chemistry back files. While a patron viewing ORR at the serial-title level seldom cares where the content comes from, this information is essential to a range of management tasks, including collection development and maintenance. The interface record makes providing an alternative view of the database possible, thus supporting these tasks and enabling librarians to view and deal with, for example, all the SilverPlatter databases or each of the various JSTOR collections together. In addition to data elements identifying the provider and, where applicable, the collection, the interface record also includes other information relevant to those entities, such as an identifier referencing the provider in the TDNet data feed. Data specific to a given title offered by the provider, such as URLs and coverage dates, are stored in the instance record. A few data elements, such as status (i.e., availability), are found in both the instance and the interface record. In these cases, the instance record supplies a default value that may be overridden or augmented for a particular title. The provider- or collection-level view of the database has a counterpart in the public interface where users can obtain similar listings by choosing the provider or collection from a drop-down list on the search page for all resource types. The user interface also displays and links to any print holdings that are available for each title in the local online catalog. This feature, and the structural issues it raises, will be discussed later in this paper.

Several commercial products are comparable to ORR in purpose, design, and functionality and some are highly innovative. Ex Libris' SFX-based journal list, for example, real-

Title Variations:	**[Main]** ***Applied organometallic chemistry*** **[Abbrev]** ***Appl. organomet. chem.*** **[Abbrev]** ***Applied organometallic chemistry (Print)***
ISSNs:	**Print ISSN: 0268-2605 Online ISSN: 0268-2605**
Publisher:	**John Wiley & Sons Ltd.**
Print Holdings:	**NONE**
Full Description:	**Applied Organometallic Chemistry publishes reviews, full papers, communications and the occasional reports on relevant conferences of applied work in the field of organometallics and metal/organic-ligand coordination chemistry. AOC publishes its content in three well-defined sections: (i) Environment, Biology and Toxicology; (ii) Speciation Analysis; (iii) Materials, Nanomaterials and Synthesis.**
Subject(s):	**Chemistry;**
Peer Reviewed:	**Y**
Primary Language:	**English**

Applied organometallic chemistry in Wiley InterScience	
Link:	**http://www.library.uiuc.edu/orr/get.php?instid=35752**
Access:	UIUC
Online Coverage:	**1996-Present**
Current Status:	✓ **Report an access problem with this specific resource.**
Added/Modified:	**Jan 15 2004 11:44AM; Feb 13 2006 02:47PM**

Applied organometallic chemistry in Wiley InterScience - Chemistry Backfiles	
Link:	**http://www.library.uiuc.edu/orr/get.php?instid=427180**
Access:	UIUC
Online Coverage:	**/1987-12/1995**
Current Status:	✓ **Report an access problem with this specific resource.**
Added/Modified:	**Jan 03 2005 02:16PM; Feb 13 2006 02:47PM**

Figure 2. Record display in ORR's public interface

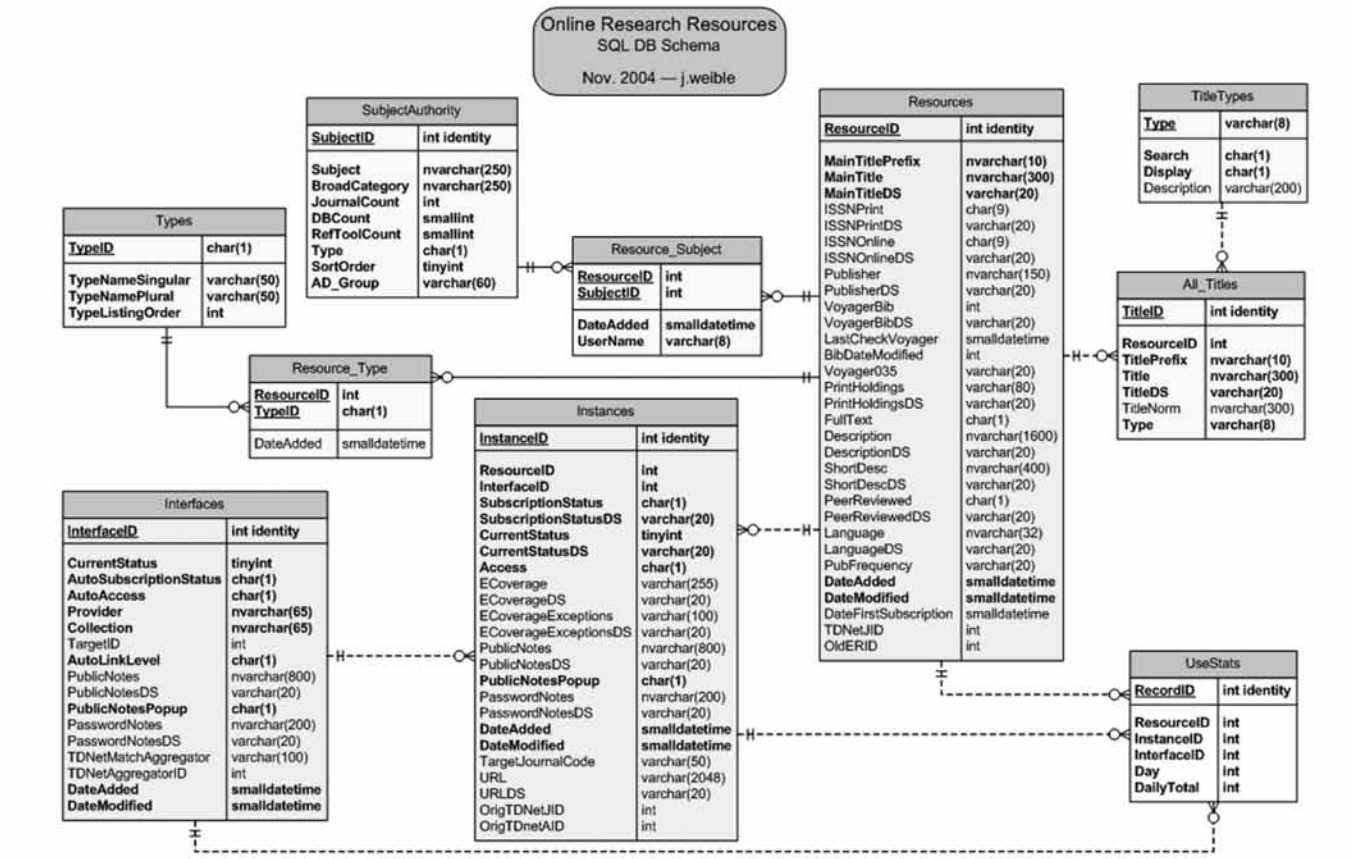

Figure 3. ORR's database schema

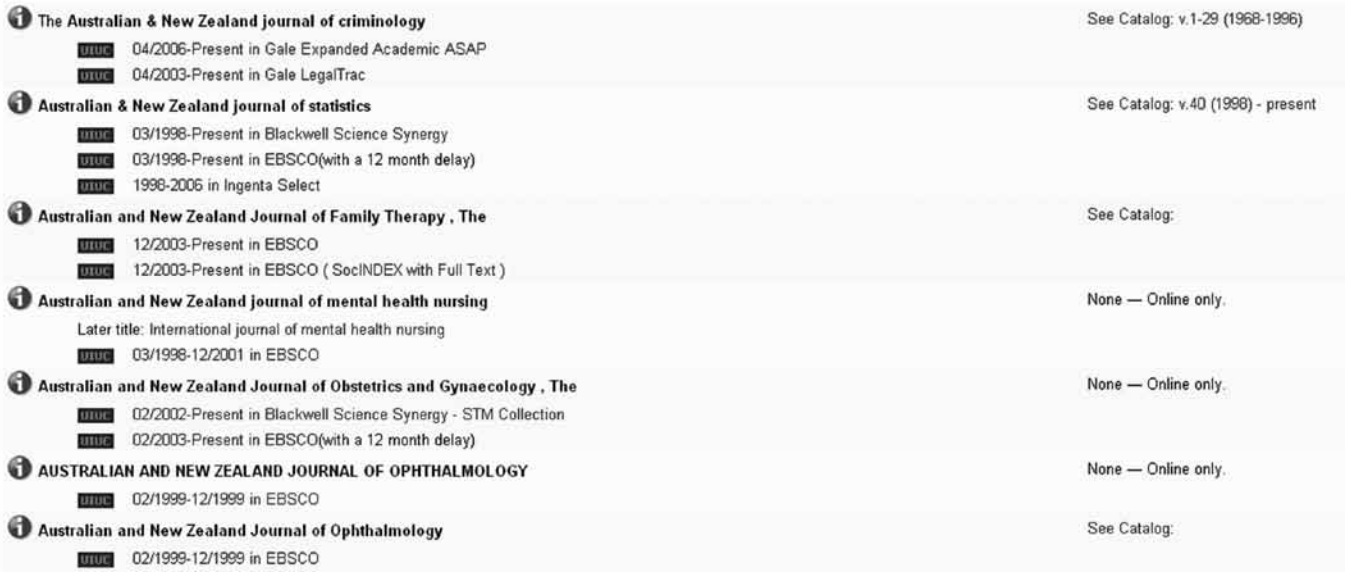

Figure 4. Title list in ORR's public interface

izes ORR's original design objective of driving a journal list and link resolver from the same knowledge base. ORR, however, offers a number of features not generally found elsewhere. While most commercial journal finding lists now provide a display of available sources grouped by title, and several also offer a link to the online catalog, most do not yet have the ability to link between earlier and later titles, or to search by provider or collection.

The *FRBR* Hierarchy

Librarians are most familiar with *FRBR* in its role as a lens through which to scrutinize a cataloging code, such as *AACR2*. The task in designing ORR was not to codify a set of cataloging rules, but to establish a data structure together with a set of procedures for populating it. These two tasks have the same objective of facilitating access to resources by creating a coherent and lucid representation of them.

The *FRBR* framework commends itself to the bibliographic management of online journals because of the overlapping needs to relate content from various providers to a common work; to link related content across different platforms; to associate holdings in different formats; and to trace a publication's identity through successive title changes, splits, or mergers. A model articulating these relationships within a comprehensive framework holds promise for guiding decisions about the appropriate record structure and content, database schema, and display format for representing these resources.

Most discussions of *FRBR* relationships focus on the hierarchy of Group 1 entities: work, expression, manifestation, and item. This hierarchy easily fits some aspects of ORR's design, even if not all of the entities and attributes implied by the latter are listed in the *FRBR* report's ontology. The resource record contains work-level data such as titles and subjects. The instance record corresponds to the manifestation, recording such attributes as the provider (which may be likened to a distributor for a print publication), access address, and source for access authorization.[12] The presentation format of the text is determined at the level of the individual provider each time it is viewed or printed, allowing for variations introduced by style sheets or other branding or customization features. Accordingly, each online viewing or printing may be considered an item (partially) instantiating the manifestation.

Discerning expression-level attributes in ORR is difficult. ORR elides attributes that could be modeled as distinguishing characteristics of expressions, such as whether accompanying graphics are provided.[13] In effect, ORR assimilates all electronic versions to the one expression. In this respect, its practice closely resembles the Cooperative Online Serial (CONSER) program's aggregator-neutral record, which similarly elides expression- and manifestation-level data to collocate all online versions under a single record.[14] Several recent analyses in the literature suggest that the appropriate treatment of expressions may be dependent on the nature of the works represented.[15] The *FRBR* report's statement, that "on a practical level, the degree to which bibliographic distinctions are made between variant expressions of a work will depend to some extent on the nature of the work itself, and on the

anticipated needs of users," supports this view.[16]

Finally, how does ORR's interface record fit into the *FRBR* model? The interface represents an aggregate entity that exists at an intersecting plane to the main hierarchy. Such aggregates are not peculiar to continuing resources. Structurally, they are like certain types of aggregates found in the monographic domain, such as "bound together" titles and collected editions, which share attributes at the item and the manifestation level respectively. Mimno, Crane, and Jones offer the following analysis of a collected edition: "At the work level, one play by Aeschylus is clearly a distinct entity from another play, but at the manifestation level, the publication information for every translated play in the volume is the same, and therefore should be kept in a single record."[17] They tentatively advocate linking a single manifestation-level record for the collected edition to multiple expression-level records for the individual titles. The interface record plays an analogous role in ORR, recording in one place manifestation-level data that apply to multiple serial works.

Identifiers and FRBRization

The ORR project combines aspects of two kinds of undertakings. It resembles certain *FRBR* implementations in that it populates its database by taking existing data and reconstructing them, via a predetermined algorithm, into a unified hierarchical structure. This is a process sometimes known as FRBRization. ORR also resembles a link resolver in that it is built around a massive consolidation of subscription- and holdings-related data, leveraged to facilitate effective linking and discovery among different content providers. In some ways, the ORR project resembles a specialized serial counterpart of OCLC projects like OpenWorldCat, which combine the two foregoing strategies by first FRBRizing an existing data set and then using a database of holdings to identify available copies.

In ORR, FRBRizing the ingested data organizes the links. Content from various providers is brought together under a single title. Links are supplied from the electronic versions to the print versions, and also between earlier and later titles. In one respect, the task of bringing together the relevant data is easier than with large-scale efforts, such as those undertaken by OCLC to FRBRize monograph records. Those projects rely on complex work keys such as author/title and author/uniform title combinations to create clusters of works, expressions, and manifestations with varying degrees of success.[18] By contrast, much of the desired clustering of data in ORR can be achieved through the simpler process of matching ISSNs from different sources and mapping selected data elements into the relevant fields in ORR records. ISSN is widely used and is assigned at the right level of granularity to serve adequately as a work identifier in most situations, at least relative to a given language and physical format. The availability of ISSN as a work key is fortunate because the uniform title headings (MARC field 130) characteristic of serial records are designed to distinguish titles rather than to collocate them, and are consequently of limited value as work keys.[19]

ISSN in its present form is far from ideal as a work identifier. Like other extant identifiers such as International Standard Book Numbers (ISBN) and Digital Object Identifiers (DOI), it addresses the need of publishers to identify distinct entities, but not the need of users to navigate between related ones. Knowing the ISSN of the print version does not help one find the online version, and vice versa, unless one has access to some kind of dictionary. This characteristic of identifiers is a consequence of the Principle of Functional Granularity, promulgated by the Indecs e-commerce body, which states that "it should be possible to identify an entity whenever it needs to be distinguished."[20] The principle says nothing about identifying related entities. At the time of this writing, a proposal was before the International Organization for Standardization (ISO) to introduce a mandatory Medium-Neutral ISSN (MNI) into the ISSN standard. This measure, if adopted, may ameliorate the existing difficulties considerably. Even in its present form, ISSN enjoys the important advantage of being uniform across different providers and through changes of publisher, something not true of ISBNs or DOIs. Each ORR record has two ISSN fields: one for the online version and one for the electronic version. These two fields jointly suffice to identify the title for most purposes. To establish the correspondence between print and online ISSNs, having sources of data that can associate the two, such as MARC records with ISSNs in their 022 and 776 fields, is valuable.

Serial Work Relationships

Seriality encompasses a much greater range of relationships than those of the *FRBR* Group 1 hierarchy. Serials change attributes such as titles, ISSNs, publishers, and physical format over time. Some changes give rise to new works or expressions bearing specific relationships to their immediate siblings or ancestors. Serials also break down into various kinds and levels of constituent subunits, such as issues, volumes, articles, indexes, and supplements. Some of these relationships are outlined in the *FRBR* report, including "successor" and "supplement" relationships between works, and whole-part relationships between serial works and their constituents.[21] These relationships define aggregates, and a comprehensive theory of serial aggregates would do much to put the design of serials-management systems on a sounder footing. Aggregates are a relatively undeveloped area in *FRBR*, receiving barely a page of direct dis-

cussion in the final report. Some progress has been made since the report's publication. A *FRBR* working group on aggregates now exists, and members of the *FRBR* community are developing general taxonomies of aggregates and their properties. For example, Albertsen and van Nuys identify a set of aggregate classes among which are several that are applicable to continuing resources: the "extension" class, which subsumes successively issued resources, including most conventional serial publications, the "update" class, which roughly corresponds to the notion of an integrating resource, and the "variant" class, which encompasses alternative versions of a publication.[22] In its present state, however, *FRBR* offers only limited guidance to the developers of a tool like ORR.

The question arises as to whether the aggregate is itself another work—a "super-work"—or indeed whether it is only the aggregate that may properly be identified as the work. Shadle advocates the latter position, arguing that the serial work should not necessarily be identified with any one record and, unless there is a merger or a split, the serial work should be considered to persist.[23] Although Shadle's position has strong intuitive appeal, Delsey's position, which allows the boundaries between works to be drawn by the prevailing cataloging code, can suffice.[24] Until a theory of aggregates is more fully developed, the position taken on this issue is not critical. The structure of ORR is compatible with both positions, and Delsey's approach has the advantage of simplifying the ontology. From a practical viewpoint, the important thing is less which title-level entities are called works and which are called manifestations, but more how well the relationships between the entities are captured. For the same reason, referring to aggregates as "super-works" is not crucial at this juncture, so long as works standing in specific relationships may form aggregates with definable properties.

The most important relationship that ORR must deal with is the successor relationship, or what serials librarians call title changes. This issue, and the related question of how ORR handles print holdings, highlights some unresolved issues with ORR's current data structure.

ORR follows the *AACR2* practice of successive entry cataloging. Each title change (or rather, each *major* title change) in a publication's history triggers the creation of a new record representing a related but distinct work. Each record contains links to the records for its predecessor and successor, but no structure represents the complete title history. In this respect, ORR exactly replicates the type of structure in *AACR2* catalogs. It also inherits one of the weaknesses of such catalogs, namely the fact that one's ability to reconstruct a complete title history is contingent upon the library owning a sufficiently unbroken run of holdings for that publication. If a serial publication has the title history S1, S2, S3, and the library owns issues of S1 and S3 but not S2, the bibliographic data in its catalog will not allow users to connect S1 with S3. This is the "missing link" problem. Although some feel that a full title history is not always desirable, it is invaluable in a distributed environment where complementary coverage may be available from different sources.[25]

This structural shortcoming overlaps with the problem of representing different formats. The *FRBR* report suggests that alternative formats are to be represented at the manifestation level.[26] That approach is not taken in ORR. The visual cues in the public display present any print holdings that are available, not as one version among others, but rather as a link to the online catalog. The display reflects the database schema, which locates print holdings data (as well as the ILS record identifier used to generate the link to the online catalog) not in the instance but in the resource record.

The differing treatment given to electronic and print formats can partly be explained by ORR's design objectives. The primary purpose of ORR is to represent available online content. For most users, the catalog link exists to provide a fallback should the desired full-text content not be available online—for example, if the issue sought predates the available back files. Accordingly, the instance record is optimized for online content. The library catalog remains the main source of information about print holdings and, rather than attempt to replicate its content in detail, ORR simply links to it.

This approach, however, equivocates between works and larger aggregates. The equivocation is often evident in the holdings data displayed in conjunction with the links to print and microform records in the online catalog. In some cases, holdings data are displayed for the specific journal title; in others, holdings data are displayed for the entire title history, including titles predating any available online content. In other words, the holdings data displayed in some cases represent another manifestation of the same work and in others represent a larger aggregate including that work and others. This inconsistency is partly the result of historical UIUC serials cataloging practice, which for a time followed latest entry, but it also reflects an unresolved tension in ORR's treatment of serial aggregates.

Locating the link at the work level does not solve the missing link problem. The problem arises in a particularly acute form in this setting. Returning to the example of a title history S1, S2, S3, consider a case where the S1 and S2 are issued in print, but the journal moves to an online-only format with S3. The ORR entry will naturally be for S3. In such a case, no print equivalent exists to which the ORR record can link. The same problem arises where a print version continues to be issued but the library

cancels its print subscription in favor of online access before a change of title. For example, the UIUC library cancelled its print subscription to *Archives of Otolaryngology* in 1975, but later regained access to this journal via an online subscription with coverage beginning in 1995. In the meantime, the journal had changed its title to *Archives of Otolaryngology—Head and Neck Surgery,* with a new ISSN. Again, the link between electronic and print holdings is lost. The problem compounds over time as each successive title change puts further distance between the latest online incarnation and its print predecessors. Until now, this problem has arisen only with a small number of titles, but ORR's developers will need to address it as more journals move toward online-only access.

The UIUC catalog uses a single record to represent print and online versions of a title. The problem would take a somewhat different form in catalogs that use multiple records. At UIUC, using the same record identifier to reference the bibliographic description for the work and to link to the print holdings is possible. Had UIUC used multiple records, the issues raised by aggregates would have been confronted at a much earlier stage of ORR's development. A library using multiple records would need to define explicitly an alternate relationship (defined in section 5.3.4 of the *FRBR* report) between the manifestations represented by the two records and, presumably, to enter both record identifiers in the database.

From Relationships to Families

Rules for title changes in the past may have been too strict. The cataloging rules may not adequately capture the notion that a work may persist through changes, even major changes, in title. The long-running debate in the serials cataloging community between successive and latest entry cataloging may reflect conflicting views about the identity of a serial work over time. Shadle's position on serial aggregates similarly gives expression to the desire to capture the nature of the serial work as a persisting entity.[27]

Why is there no record structure in ORR representing the aggregate's title history? A convenient source has yet to be found for the required data. To remain complete and current, ORR depends on external data sources and could not otherwise exist on its present scale. The same dependence also means that ORR is constrained by the quality of the available data, and by the data structure of the source. As with identifiers, ORR to some degree inherits the characteristics and underlying assumptions of existing standards. Had latest-entry rather than successive-entry cataloging been the norm, extracting the complete title history from the MARC record in hand would have been relatively easy.

A number of proposals coalesce to suggest a way forward. A 1993 study by Alan showed that more than 70 percent of "title-change record sets" within a sample of CONSER-authenticated MARC records were linked together by a combination of ISSNs, LC classification numbers (LCCN), and OCLC numbers in the 780 and 785 linking entry fields.[28] Antelman suggests that the same data could be used as the basis of a work-set algorithm that would create "bibliographic families" showing relationships between works.[29]

Tillett has advocated the use of authority records to show relationships among bibliographic entities, and Rosenberg and Hillman have proposed a structure for doing so with serial works.[30] Building authority structures based on data harvested using a strategy similar to Antelman's may be possible. Ideally, this authority file would be a large-scale shared enterprise, but even a local project within the limited context of ORR may be feasible. These authority records would record data—especially identifiers like ISSNs—relating to alternative formats, title changes, merges, splits, and other relationships. This approach would differ from the existing strategy used in ORR for linking title changes in that it would encompass a wider range of relationships and would allow all relationships to be shown to the user, overcoming the missing link problems. The same data would have other potential applications. It could be used to effect linkages between catalogs in a shared environment, for example, or to enhance link resolution.

The work-set algorithm suggested by Alan and Antelman could be supplemented by other sources of data, such as MARC 776 additional physical-form information and a subscription to the ISSN register. A proposed development by the ISSN International Centre promises an alternative model for implementing an authority structure.[31] The plan is to implement the ISSN database as a lookup and resolution service. A service of this kind would make possible the building of extremely powerful and flexible tools for discovering and accessing serial publications, and would allow the developers of systems such as ORR to overcome many current obstacles.

Given the pace of change in the current environment and the vagaries of journal publishing, a service resembling one of those outlined previously in this paper already may have been developed by the time this paper is published.

Sources and Targets: Other Issues

In ORR's distributed environment, many other issues arise with both the quality of the available data and with the characteristics of the resources to which ORR provides access. Data sources present particular problems.

- The TDNet data feed does not have a separate field for tracking title changes, instead giving this information in free text within the title field. Title changes have to be caught by library staff members, who then create a new record manually.
- Ulrich's *Periodicals Directory,* although it indicates if an electronic version is available, does not always provide the corresponding electronic ISSN. This information has to be supplied from other data feeds, or else by a human operator.
- The UIUC catalog uses the single record approach to represent the print and electronic versions of each journal. This approach can result in the omission of electronic ISSNs necessary for matching and linking. Best practice is to include 776 fields providing the ISSN in subfield x and other identifiers (OCLC number, LCCN) in subfield w.
- A single consolidated statement of print holdings is not usually available. Instead, the catalog breaks down the holdings for print copies by the various library locations. As already noted, the practice of successive entry is another obstacle to the provision of a single summary of holdings. ORR's print holdings field was initially populated partly with summary holdings data fortuitously available from another, unconnected project, but a different solution will need to be found for the longer term. In the future, data may be parsed from the 866 field of MARC-holdings records.

The targets to which ORR links can present a further layer of structural complexity. Just as each source of data has its own structure that must be mapped into ORR, each provider's manifestation of a title has its own implicit structure for presenting the constituent units of each work or group of works. Most examples fall into one of following categories:

- The entire history of a journal is entered on a single page under its current title alone. Earlier titles are not given, unless they happen to be reproduced on the scanned pages of the earlier issues themselves. An example, cited by Jones, is *Online Information Review,* which does not appear anywhere on the Emerald site under its earlier title, *Online and CD-ROM Review,* even though some of the issues available on the site were originally published under that title.[32] Because individual titles are searchable within ORR, it provides better title-level access than the vendor's own site. This is a decided advantage, since journal articles are cited using the title of the journal at the time of publication.
- All titles are accessed via a single page, with prominence given to the latest or current title. Individual titles are listed with their respective publication dates, but it may not be possible to retrieve them by a search within the native interface. Examples of providers following this format include Springer and the Royal Society of Chemistry. Again, title-level access is better in ORR than through the vendor's site, but with the further advantage, at least in the examples given, that a link to a page representing each distinct title is possible.
- Each title in the sequence is entered separately on its own page, with links provided between them. This is the most common arrangement, and most closely reflects successive entry practice. EBSCO, JSTOR, and many others follow this approach. In these cases, the ORR record for each title simply links to the corresponding page.
- No title-level page is given and content is available only by searching for articles by means of a search form. An example is OCLC FirstSearch, for its Wilson Select Plus collection. In these cases, ORR shows the user an icon indicating that a further search will be required after linking to the vendor page. Whether the icon is displayed is determined by a field called "AutoLinkLevel" in the interface record. This field indicates whether the link points to a page for the title or whether a further search will be required.

Conclusion

This paper has presented a case study of the cataloging issues involved in the creation of an online journal finding lists and serials management system. Although an a posteriori analysis, *FRBR* concepts are strongly applicable to this project. The Group 1 hierarchy is an obvious model for organizing content from different providers, while the application of the larger *FRBR* framework to serial relationships raises important issues regarding aggregates. The discussion also touched briefly on Group 2 and Group 3 entities—content providers and subjects respectively.

In hindsight, conducting a *FRBR* analysis in the early stages of the ORR project would have been advisable. Such an analysis might have helped to clarify some of the issues that emerged during ORR's development, especially the treatment of title histories and

print holdings. However, although the *FRBR* model provides a framework for conceptualizing the problems, it does not, at present, offer a complete blueprint for a solution. The challenge of applying *FRBR* to a serials system raises as many questions for the interpretation and future development of the *FRBR* model as it does for the design of the serials system itself.

This study suggests a number of possible topics for further consideration. The *FRBR* approach of relating user tasks to entity relationships may help to clarify what is needed to build interoperable services in a distributed environment. One potential line of inquiry, hinted at but not pursued in any depth here, is how *FRBR* may help to model algorithms for link resolution. Much of the effort in this project went into mapping data from outside sources into ORR. *FRBR* analysis should help rationalize the consolidation of data from various sources by ensuring that they map to entities at the right level of the *FRBR* hierarchy. More fundamentally, *FRBR* should be helpful in guiding the design of database structures for serials-management systems.

The emphasis of this paper has been largely conceptual. The creation of ORR has been, above all, a practical matter, and many aspects of its development are amenable to empirical study. This author and his UIUC colleagues hope to publish a more detailed examination of the process of populating the database and its outcomes.

References

1. Kristin Antelman, "Identifying the Serial Work as a Bibliographic Entity," *Library Resources & Technical Services* 48, no. 4 (Oct. 2004): 249.
2. Ed Jones, "The FRBR Model as Applied to Continuing Resources," *Library Resources & Technical Services* 49, no. 4 (Oct. 2005): 233.
3. Lisa German, Wendy Shelburne, and Michael Norman, "Creating the Ultimate Knowledge Base," (presentation at the 25th annual Charleston conference, Nov. 2004, Charleston, South Carolina). http://netfiles.uiuc.edu/manorman/CharlestonPresentation2005.ppt (accessed Jan. 27, 2006); Wendy Shelburne and Michael Norman, "Online Research Resources (ORR): University of Illinois at Urbana-Champaign's Integrated Management System for Electronic Resources," in *Charleston Conference Proceedings 2004,* ed. Rosann Bazirjian and Vicky Speck, 189–94 (Westport, Conn.: Libraries Unlimited, 2005).
4. Janis F. Brown, Janet L. Nelson, and Maggie Wineburgh-Freed, "Customized Electronic Resources Management System for a Multi-Library University: Viewpoint from One Library," *The Serials Librarian* 47, no. 4 (2005): 89–102; Laura Tull et al., "Integrating and Streamlining Electronic Resources Workflows via Innovative's Electronic Resource Management," *The Serials Librarian* 47, no. 4 (2005): 103–24.
5. IFLA Study Group on the Functional Requirements for Bibliographic Records, *Functional Requirements for Bibliographic Records: Final Report* (Munich: K. G. Saur, 1998).
6. Ibid., section 6.
7. Barbara Tillett, *What Is FRBR? A Conceptual Model for the Bibliographic Universe* (Washington, D.C.: Library of Congress Cataloging Distribution Service, 2004), 5. www.loc.gov/cds/FRBR.html (accessed Jan. 27, 2006).
8. Barbara Tillett, "Component Parts," online posting, Nov. 20, 2003, FRBR mailing list. www.ifla.org/VII/s13/wgfrbr/archive/FRBR_Listserv_Archive.pdf (accessed Jan. 27, 2006).
9. Antelman, "Identifying the Serial Work," 241.
10. *Anglo-American Cataloging Rules,* 2nd. ed., 1998 rev. (Ottawa: Canadian Library Assn.; London: Library Assn. Publishing; Chicago: ALA, 1998).
11. Martha Yee, "FRBRization: A Method for Turning Online Public Finding Lists into Online Public Catalogs," *Information Technology and Libraries* 24, no. 3 (Sept. 2005): 79–95.
12. IFLA Study Group on the Functional Requirements for Bibliographic Records, *Functional Requirements for Bibliographic Records,* section 4.4.
13. Ibid., section 4.3.
14. Naomi Kietzke Young, "The Aggregator-Neutral Record: New Procedures for Cataloging Continuing Resources," *The Serials Librarian* 45, no. 4 (2004): 37–42.
15. Mary-Louise Ayres, "Case Studies in Implementing Functional Requirements for Bibliographic Records (FRBR): AustLit and MusicAustralia," *Australian Library Journal* 54, no. 1 (Feb. 2005): 43–54; Gunilla Jonsson, "Cataloging of Hand Press Materials and the Concept of Expression in FRBR," *Cataloging & Classification Quarterly* 39, no. 3/4 (2005): 77–86; Yann Nicolas, "Folklore Requirements for Bibliographic Records: Oral Traditions and FRBR," *Cataloging & Classification Quarterly* 39, no. 3/4 (2005): 179–95.
16. IFLA Study Group on the Functional Requirements for Bibliographic Records, *Functional Requirements for Bibliographic Records,* 19.
17. David Mimno, Gregory Crane, and Alison Jones, "Hierarchical Catalog Records: Implementing a FRBR Catalog," *D-Lib Magazine* 11, no. 10 (Oct. 2005). www.dlib.org/dlib/october05/crane/10crane.html (accessed Aug. 4, 2006).
18. Thomas Hickey and Edward O'Neill, "FRBRizing OCLC's WorldCat," *Cataloging & Classification Quarterly* 39, no. 3/4 (2005): 239–51.
19. *CONSER Cataloging Manual*, 2002 ed. (Washington, D.C.: Library of Congress, 2002), module 5.1.
20. Indecs, *Putting Metadata to Rights: Summary Final Report* (2000). www.indecs.org/pdf/SummaryReport.pdf (accessed Jan. 27, 2006).
21. IFLA Study Group on the Functional Requirements for Bibliographic Records, *Functional Requirements for Bibliographic Records,* section 5.3.
22. Ketil Albertsen and Carol van Nuys, "Paradigma: FRBR and Digital Documents," *Cataloging &*

Classification Quarterly 39, no. 3/4 (2005): 125–49.
23. Steve Shadle, "FRBR and Serials" (presentation at the 2005 NASIG annual conference, May 2005, Minneapolis, Minn.). www.nasig.org/members/handouts/2005/Shadle.ppt (accessed Jan. 27, 2006).
24. Tom Delsey, *FRBR and Serials* (The Hague: IFLA, 2003). www.ifla.org/VII/s13/wgfrbr/papers/delsey.pdf (accessed Jan. 27, 2006).
25. Frieda Rosenberg and Diane Hillman, "An Approach to Serials with FRBR in Mind" (draft, revised Jan. 24, 2004). www.lib.unc.edu/cat/mfh/serials_approach_frbr.pdf (accessed Feb. 10, 2006).
26. IFLA Study Group on the Functional Requirements for Bibliographic Records, *Functional Requirements for Bibliographic Records,* section 5.3.4.
27. Shadle, "FRBR and Serials."
28. Robert Alan, "Linking Successive Entries Based upon the OCLC Control Number, ISSN, or LCCN," *Library Resources & Technical Services* 37, no. 4 (1993): 410.
29. Antelman, "Identifying the Serial Work," 244.
30. Barbara Tillett, "Bibliographic Universe Created by FRBR and FRAR" (presentation at the FRBR satellite meeting of the 2005 IFLA conference, Jarvenpaa, Finland, Aug. 2005). www.oclc.org/research/ events/frbr-workshop/presentations/tillett/FRBR_and_cat_rules.ppt (accessed Jan. 27, 2006); Rosenberg and Friedman, "An Approach to Serials with FRBR in Mind."
31. *CONSERLINE* no. 26 (Spring 2005). www.loc.gov/acq/conser/consln26.html (accessed Jan. 27, 2006).
32. Jones, "The FRBR Model as Applied to Continuing Resources," 231.