Building a "FRBR-Inspired" Catalog: The Perseus Digital Library Experience

Alison Babeu
Perseus Digital Library
1/31/2008

Table of Contents

1. Introduction

If one follows any of the major cataloging or library blogs these days, it is obvious that the topic of FRBR (Functional Requirements for Bibliographic Records) has increasingly become one of major significance for the library community. What began as a proposed conceptual entity-relationship model for improving the structure of bibliographic records has become a hotly debated topic with many tangled threads that have implications not just for cataloging but for many aspects of libraries and librarianship.[1]

In the fall of 2005, the Perseus Project experimented with creating a FRBRized catalog for its current online classics collection, a collection that consists of several hundred classical texts in Greek and Latin as well as reference works and scholarly commentaries regarding these works. In the last two years, with funding from the Mellon Foundation, Perseus has amassed and digitized a growing collection of classical texts (some as image books on our own servers that will eventually be made available through Fedora), and some available through the Open Content Alliance (OCA)[2], and created FRBRized cataloging data for these texts. This work was done largely as an experiment to see the potential of the FRBR model for creating a specialized catalog for classics.

Our catalog should not be called a FRBR catalog perhaps, but instead a "FRBR Inspired catalog." As such our main goal has been "practical findability," we are seeking to support the four identified user tasks of the FRBR model, or to "Search, Identify, Select, and Obtain," rather than to create a FRBR catalog, per se. By encoding as much information as possible in the MODS and MADS records we have created, we believe that useful searching will be supported, that by using unique identifiers for works and authors users will be able to identify that the entity they have located is the desired one, that by encoding expression level information (such as the language of the work, the translator, etc) users will be able to select which expression of a work they are interested in, and that by supplying links to different online manifestations that users will be able to obtain access to a digital copy of a work.

This white paper will discuss previous and current efforts by the Perseus Project in creating a FRBRized catalog, including the cataloging workflow, lessons learned during the process and will also seek to place this work in the larger context of research regarding FRBR, cataloging, Library 2.0 and the Semantic Web, and the growing importance of the FRBR model in the face of growing million book digital libraries.

2. What is FRBR?

Barbara Tillett, chief of cataloging policy at the Library of Congress (LC), has described FRBR, as a "conceptual model for the bibliographic universe." (Tillett 2004) The answer to the question "What is FRBR" and why it is important for libraries, both physical and digital, has received significant analysis, by authors far more eloquent than this one, so we will begin with several definitions.

"FRBR is not cataloging rules, a system architecture, a record format, a display standard or the end of the world! ….It is like the Bible. Many people will tell you what they think it says, but read the document if you really want to know" (Weiss and Shadle, 2007).

"FRBR is not a draft standard, nor is it intended to replace AACR2 or any other cataloging code. FRBR is a systematic, international examination of automated catalogs and the records that comprise them" (Allgood 2007).

---

[1] The International Federation of Library Associations (IFLA) first released the FRBR guidelines in 1998 which are available at http://www.ifla.org/VII/s13/frbr/frbr.pdf.
[2] http://www.archive.org/index.php

"In essence, FRBR is a model of a model, if one considers that a bibliographic record is a representation of a document and so, in its own way, is as much a model as FRBR. If one considers a title page or other chief source of information to be a representation of a document as well, and thus a model in its own right, FRBR is a model of a model of a model of a document. In the list of definitions above, the first and third fit FRBR most closely: FRBR is a representation and simplified description of the bibliographic universe" (Carlyle 2006).

As these varied definitions demonstrate, defining FRBR is not a simple task. The "launch" of the FRBR model has led to such a diverse body of research that the IFLA bibliography regarding FRBR now measures over 50 pages,[3] with two books dedicated entirely to FRBR also currently released in just the last few months (Taylor 2007, Maxwell 2007).

FRBR at its most basic is an entity relationship model that has attempted to create a conceptual framework that can assist in the creation of bibliographic records independent of any set of cataloging rules or encoding guidelines. IFLA designed the FRBR model to support what they believed to be the most important user tasks: to find materials, to identify an entity, to select an entity, and to obtain access to the entity desired; or, in other words, to be able to use a catalog to find materials, that the catalog records be sufficiently descriptive so that a user may successfully identify and then select the correct entity that they desire, and that the catalog will then facilitate assisting the user in retrieving the desired item. The FRBR guidelines include three groups of entities, Group 1: work, expression, manifestation and item, Group 2: person and corporate body, and Group 3: concept, object, event and place. The majority of current research has focused on the Group 1 entities, as has the work conducted by the Perseus Project.

The Group 1 entities are defined as following by IFLA:

Work—"a distinct intellectual or artistic creation."
Expression—"the intellectual or artistic realization of a work."
Manifestation—"the physical embodiment of an expression of a work."
Item—"a single exemplar of a manifestation." (IFLA 1998)

This list is frequently summarized by the statement, "a work is realized through an expression, an expression is embodied in a manifestation, and a manifestation is exemplified by an item" (IFLA 1998). To make this more concrete, Vergil's Aeneid is considered a work, Robert Fitzgerald's original English translation is viewed as one expression of that work, printings by different publishers of the same Fitzgerald translation are different manifestations of that expression, and my individual copy of one of those printings is an item. While all of these definitions (although principally that of the expression level) have received substantial consideration and further expansion, these most basic definitions helped guide the work conducted by the Perseus Project.

The remainder of this paper is as fellows: Section 3 reviews several areas of related work concerning FRBR and the issues raised relevant to the work considered here, Section 4 discusses our previous FRBR experimentation and how it differs from our current work, Section 5 explains our cataloging workflow and future work planned on the catalog, and Section 6 concludes with the lessons learned.

---

[3] Available at http://www.ifla.org/VII/s13/wgfrbr/bibliography.htm

3. Related Work

FRBR Research Overviews

A number of researchers have provided recent broad overviews of FRBR that have helped inform the work presented here. Most recently, the author participated in a study along with several dozen other "FRBR experts" conducted by Yin Zhang and Athena Salaba of the University of Kent, which was designed to determine the most critical issues and challenges facing FRBR research (Zhang and Salaba 2007a, Zhang and Salaba 2007b). Their research sought to identify critical issues within five areas "(1) the FRBR model, (2) FRBR-related standards, (3) FRBR applications, (4) FRBR system development, and (5) FRBR research" (Zhang and Salaba 2007b).

They learned that in terms of the FRBR model, the need to validate and verify the FRBR model with real world data and to harmonize it with other models were listed as top issues. For FRBR related standards, top issues included, the "need to develop cataloging rules in line with FRBR" and the "need to address FRBR-based record structures, record encoding standards and frameworks for FRBR implementations" (Zhang and Salaba 2007b). This second issue is of particular interest to the deliberations here, as one of the greatest challenges we faced was in how to create metadata that could support the basic FRBR tasks listed previously. Other issues in terms of FRBR standards included the need to promote interoperability for sharing of FRBR-based data, and to develop entity identifier standards and management mechanisms. In regards to FRBR applications, Zhang and Salaba learned that researchers wanted guidelines and examples for designing FRBR applications, and recognition that interpretations of the model might differ between communities. The issue of FRBR systems development raised a wide ranging list of concerns including the need to develop tools to facilitate the FRBRization process, the problem of FRBRizing legacy data cataloged under changing standards and in different formats, and the challenges of designing a user interface based on the FRBR model. Finally, Zhang and Salaba listed the top issues in regards to general FRBR research as including the need to conduct more user studies on FRBR systems and research into FRBR based displays, among many others.

The authors point out that across all of the five areas, three of the top ten issues, address "FRBR related standards." Zhang and Salaba posit that: "the most critical issue is developing cataloging rules in line with FRBR. More specifically, it is deemed important to address FRBR-based record structures and record encoding standards and frameworks, which are considered essential steps for FRBR implementations" (Zhang and Salaba 2007b). Similarly, they impart that, "the FRBRization of existing data created following other standards, such as Dublin Core (DC), Metadata Object Description Schema (MODS) and Metadata Authority Description Schema (MADS), causes similar concerns for digital collections" (Zhang and Salaba 2007b). Our research has also faced similar issues, especially in regards as to how to best utilize currently existing library metadata standards to support FRBRization.

FRBR, Catalogs and the Future of Cataloging

FRBR, RDA and Cataloging Rules

Many recent investigations of FRBR have raised questions about the nature of cataloging and library catalogs, and increasingly how this expensively created data can be both created more efficiently, shared, and moved onto the Internet for reuse in any number of automated applications. It seems now that any debate of the future of cataloging and metadata will also include at least a tacit nod to FRBR as well. The much anticipated and recently released Library of Congress Working Group on the Future of Bibliographic Control (LC WGFBC) has placed FRBR center stage in its top level recommendations in

terms of positioning the library community for the future, the reports states that libraries should "work to realize the potential of the FRBR framework for revealing and capitalizing on the various relationships that exist among information resources" (Amelung, et al. 2008). This is just one of many mentions of FRBR in this report, and an entire section is dedicated to the "realization of FRBR." Yet the report also recognizes the limitations of FRBR such as that there are as yet no standard ways to exchange work based data and no cataloging rules that support the creation of records based on the FRBR model, and thus includes the important caveat that "FRBR must be seen as a theoretical model whose practical implementation and its attendant costs are still unknown" (Amelung, et al. 2008).

This question of metadata interoperability and how the implementation of FRBR might affect previous metadata efforts is one of central importance. Salaba and Zhang report that, "One issue that arises is whether FRBR will make record-sharing easier or more difficult. Examples of issues to be resolved include what information should be part of the bibliographic record and what information part of the authority record. At what level should we create records: work-level, expression-level, manifestation-level or all?" (Salaba and Zhang 2007).

For many years, the standard for Anglo-American cataloging has been the Anglo-American Cataloging Rules (AACR), first published in 1967, with several major revisions to date.[4] This set of cataloging rules is set to be replaced by the Resource Description and Access (RDA) set of standards in 2009.[5] Although RDA has recently undergone some major revamping and the committee has released a number of documents illustrating how its elements have been mapped to the FRBR model (Resource Description and Access 2007), this evolving standard has been criticized as not going far enough by some librarians (Coyle and Hillmann 2007), while other librarians, most notably former American Library Association president Michael Gorman, have insisted it goes too far without accomplishing anything (Gorman 2007). Gorman has also criticized the FRBR model, stating that while "FRBR may have some merit as a way of looking at the theory of cataloguing—it has little as a foundational document for creating a cataloguing code" (Gorman 2007). In terms of RDA, Coyle and Hillmann assert that only "lip service" has been paid to the FRBR Group 1 entities, and are concerned that in RDA:

Preferences about identification of materials continue to focus on transcription in concert with rules for creating textual "uniform" titles by which related resources can be gathered together for display to users. Similarly, relationships between works or derivations have been expressed using textual citation-like forms in notes. These legacy practices fly in the face of the reality that in the digital world, identity is rarely expressed in a textual way, but instead standard linking technologies with Uniform Resource Identifiers (URIs) are preferred. Because most catalogers do not understand how these techniques can easily enable human readable displays, they tend to insist that cataloger-created textual notes are still the preferred methodology, and must be prescribed in the rules (Coyle and Hillmann 2007).

They fear that RDA's lack of support automated processes and the inability to integrate machine created metadata into bibliographic records will make RDA an untenable description standard for a digital future.

The LC WGFBC has also gone as far as to recommend that all work be suspended on RDA until its alignment with FRBR is better articulated and better business cases be made for how the new rules will actually be implemented in catalogs and align with existing metadata. As a final thought on RDA, the report suggests that, "Although RDA is being based on FRBR principles, FRBR itself is still evolving" (Amelung, et al. 2008).

---

[4] http://www.aacr2.org/
[5] http://www.collectionscanada.gc.ca/jsc/rda.html

FRBR Models and Fixing MARC Based Catalogs

Digital library consultant Karen Coyle offered a thorough assessment of the issues involved in how FRBR might improve library records and catalog systems in her seminal article, "Future Considerations: The Functional Library Systems Record," along with some examples of how FRBRized records might look. Coyle highlights the fact that, unfortunately, "the bibliographic view of what is functional and the system views of functional are not currently being discussed in concert. Bringing these two reform movements together would be a better formula for success than either of them would have on its own" (Coyle 2004). Indeed, many analyses of FRBR as a model often include little examination of how such models might be implemented in an actual library system. Another essential point Coyle makes is that FRBR shifts the focus away from the physical descriptions of individual publications to a concentration on their textual content and the relationships between them. The FRBR model, Coyle contends will move us "toward a view of a universe of inter-linked publications where users eventually will not need to be concerned with differences in formats or the vagaries of nearly identical printings of the same works." This is one aspect of FRBR that was especially important to our work explored here.

Perhaps the most thorough review of how current MARC records can be used to promote FBRRization can be found in a recent article by Martha Yee, the cataloging supervisor at the UCLA Film and Television Archive (Yee 2005). Yee believes that the most productive line of FRBR research may well be to thoroughly scrutinize how the attributes in already existing MARC bibliographic, authority and holdings records may be used to enable FRBRization of OPACs. In order to support her contention that much of the information needed to FRBRize catalogs is already present in MARC data, Yee gives a comprehensive list of MARC fields that can serve as identifiers for works and expressions in both bibliographic and authority records. Yee proposes that most OPACS are not truly catalogs, but instead serve as online finding lists, and rather grimly sums up the future for libraries, if they do not meet this pressing issue:

The rest of the world has become enamored of Google. Google cannot carry out the objectives of the catalog either. But if our choice is between OPACs, which are expensive but cannot carry out the objectives of the catalog, and Google, which is cheap and cannot carry out the objectives of the catalog, I know what the choice is likely to be (Yee 2005).

One basic solution to this problem in Yee's opinion is to "find and educate system designers who can grasp the fact that the complexity of our records is a direct result of the inherent complexity of the bibliographic universe" (Yee 2005). Both correcting legacy catalog data and developing adequate library systems that will be able to manage a "FRBR view" will likely be significant challenges.

The challenge of FRBRizing legacy data and the reality of current catalog systems were also addressed by a recent article by Maja Žumer, associate professor of library and information science at the University of Ljubljana. Žumer stresses that, "to make the transition to FRBR possible, it is necessary to extract the FRBR structure from existing data" (Žumer 2007). She also acknowledges, however, that because much legacy bibliographic data remains inconsistent and error-ridden, efforts to FRBRize such data will remain challenging for the foreseeable future. Furthermore, as Žumer points out, a great deal of important information within catalog records exists only in an unstructured format, such as notes fields, which makes automated processing of such information difficult. It is for this reason, that in our research, we sought to encode all information within each bibliographic record in a machine readable format, making limited use of notes fields.

Some of the earliest work in evaluating how legacy catalog data could be converted to the FRBR model was put forward by Peter Weinstein, who developed an ontology that expanded the entities included within FRBR: the entity of the work was remodeled as a "conception" and two new entities were added,

"materialization" or a "physical embodiment of a manifestation," and "instance" or "a particular copy of a materialization" (Weinstein 1998). Weinstein hoped to support the creation of a catalog based on "formal ontological model of bibliographic relations" and briefly looked at how MARC elements might be mapped to an ontological model.

In a similar vein, Hegna and Murtomaa have detailed their efforts in analyzing MARC records to determine what attributes could best be used for automatic conversion to FRBR. They determined that even though MARC records held attributes that could be used to identify work, expression and manifestation entities, inconsistencies in cataloging and other errors would prove a great hindrance to automated processing (Hegna and Murtomaa 2002).

Taking this work somewhat further, Aalberg, et al. have researched the creation of tools for converting MARC data to a FRBR model for a joint project between the Norwegian University of Science and Technology, the Norwegian bibliographic database BIBSYS and the National Library of Norway (Aalberg, et al. 2006, Mönch and Aalberg 2003). Their basic process involved identifying the different entities in a MARC record, selecting the MARC fields that describe each entity, and finding relationships between entities. The conversion tool they designed made extensive use of XSLT, with the input being MARC records in MarcXchange XML format, and the output "a record for each entity in a format that extends the MarcXchange with FRBR type attributes and a relationship element" (Aalberg, et al. 2006). They were able to generalize their tool by using a database to store variable data for the XSLT templates, including entity conditions, mapping of entities from MARC to FRBR, and the conditions for relationships. The tool they designed was used to convert 4,000,000 records in the BibSys database, and they learned during this conversion process that the major issues faced were inconsistent catalog data and scalability issues with XSLT. While the authors concede that it is as yet impractical for libraries to attempt full conversion to a FRBR based data model, they believe that conversion tools such as theirs can at least help create FRBRized views for catalogs.

FRBR and the Multiple Versions Problem

The greatest potential of the FRBR model is reasoned by many to be its potential ability to improve not just the quality of catalog records but the user experience of browsing and searching in online public access catalogs or OPACS. One of the major issues many users find when searching a library catalog is that most catalogs display multiple occurrences of a work not only through multiple records for all of its different manifestations but also through multiple records for each of those manifestations different formats, typically without clustering them in any sort of meaningful way. As Coyle and Hillmann explain:

Libraries' earliest experience with the proliferation of copies of resources in different physical formats was with the reproduction of printed materials, first in microformats, then in digital formats. Library cataloging rules required each new iteration in a different format to have its own entry in the catalog. Although seemingly efficient in allowing virtual "cloning" of catalog information from one version to another, in the end this practice proved to have a very negative impact on the usability of the catalog, causing an increase in catalog entries for what to many users is essentially the same resource (Coyle and Hillmann 2007).

One specific challenge of implementing FRBR then, is not only how but whether libraries should attempt to move from manifestation to expression level cataloging.

Jennifer Bowen, head of cataloging at River Campus Libraries at the University of Rochester, recently published a detailed overview of the attempts of the AACR's Format Variation Working Group (FVWG) to determine if expression level cataloging is something that can be practically achieved (Bowen 2005). Catalogers typically catalog at the manifestation level, because that is the book in hand, and as a practical

matter Bowen points out, libraries collect various manifestations of a work over time, not all at once. While the FVWG ultimately decided that libraries would need to continue the practice of cataloging manifestations, they also "reaffirmed the need to provide access to expressions within catalogs, but recommended that this be achieved instead through an exploration of expression-based collocation, rather than expression-based cataloging" (Bowen 2005). The FVWG suggested two forms for assigning expression level headings to records: cataloger created and system created (utilizing already existing bibliographic data in the record). In the cataloger created model, new rules will be created for the formation of uniform titles, specifically, "exploration of uniform title authority records as a means of distinguishing specific works and expressions within catalogs and of collocating manifestations of the same work and expression" (Allgood 2007). Yet one remaining issue is that uniform titles are in the end, language specific, and typically are not uniformly applied by all catalogers.

While Bowen believes that libraries will need to prepare their catalogs for expression level headings, in the mean time she maintains that libraries should still attempt to apply as much of the FRBR model locally as they can:

Instead of debating whether or not two resources represent the same expression or the same work, consider instead whether or not catalog users would benefit from having those resources grouped together under the same heading in the catalog. One of the underlying goals of FRBR's creators was to address the needs of catalog users, and these needs may outweigh the value of strictly adhering to the details of the FRBR model (Bowen 2005).

Libraries could begin by asking vendors for interfaces that will allow better FRBRization, but Bowen also points out that there is as yet little data demonstrating the increased usability of FRBR based interfaces for catalogs. Bowen concludes that while the potential benefits of implementing the FRBR model are widely accepted, including "better collocation, more efficient navigation of search results, and better bibliographic control in a global environment," the fact that it is not a formal standard for resource description may cause it to be years before its benefits make it uniformly into cataloging rules and consequently into library catalogs. At the same time, Bowen encourages libraries to attempt "do it yourself FRBR." Indeed, the University of Rochester where Bowen works is making significant inroads in this area with its Mellon Funded "Extensible Catalog Project."[6]

A recent article by Julian Everett Allgood appraises this "multiple version" or MulVer problem, and contends that solving the "MulVer" problem is of critical importance for libraries, since "users today have no patience for confusing OPAC displays with multiple hits for equivalent resources" (Allgood 2007). The FRBR model has great promise for libraries, Allgood believes, because it both encourages librarians to think about "catalogs rather than individual records" and also places "a renewed emphasis upon users and their needs." With a special focus on serial titles and records, Allgood lists a number of items that will be available in the "FRBR-aware catalog," including the ability to "index and retrieve elements or attributes present in both the authority file (i.e., works and expressions) and in the bibliographic/holdings file where manifestation and item data resides." While Allgood admits that creating a "FRBR-aware OPAC display" will not necessarily eliminate the multiple versions problem, he hopes that by clustering related works, expressions and manifestations more clearly, users will at least no longer need to consult multiple catalog records for equivalent versions. Rather than changing cataloging rules, Allgood observes that **"**FRBR's greater influence may be upon how ILMS system designers develop OPACs to cluster these manifestation-level descriptions into work and expression-level displays for users." He urges that libraries must either significantly change their cataloging practices or improve OPAC displays, arguing that the second option is more tenable in the immediate future.

---

[6] http://www.extensiblecatalog.info/

The multiple versions problem has also been evaluated by Zorana Ercegovac, who conducted experiments with FRBR through the study of Edwin A. Abbott's novel *Flatland*, a work with many expressions and manifestations (Ercegovac 2007). Through her survey of bibliographic records for *Flatland* and concurrent visualization of these records with the Library of Congress FRBR display tool[7] and other manual experiments, Ercegovac demonstrates that library catalogs need to supply users with more sophisticated browsing options, such as grouping bibliographic records into meaningful clusters. One important point made by Ercegovac is that many efforts to identify expressions of a work have focused only on the attribute of language, although many users may also be interested in types of derivative bibliographic relationships other than translations, such as amplified editions or other types of revised editions. Another major issue she listed was that a number of MARC fields did not allow enough granularity in terms of transcribing different statements of responsibility, such as illustrators, additional translators, authors of introductions, etc. She ultimately declares that in order for library catalogs to assist users that any given "metadata schema must be able to link multiple versions of the same resource, to show how these resources are interrelated explicitly, and how these are related to other similar resources in a digital collection" (Ercegovac 2007).

FRBR and the Future of Library Catalogs

The future of the library catalog and its future as a discovery system for users has received a great deal of attention lately in the library literature (Danskin 2006, Dempsey 2006, Marcum 2006, Mann 2007, Markey 2007, Miksa 2007, Medeiros 2007). Much of this work is beyond the scope of this paper, but this section will briefly review this issue as it relates to the FRBR model. This present level of debate regarding the current dismal state of library OPACs and library cataloging and data in general, has been inspired by a number of factors including the growth of mass digitization projects, the LC WGFBC report (Amelung, et al. 2008), and another influential report commissioned by the Library of Congress and written by Karen Calhoun, often simply referred to as the "Calhoun Report" (Calhoun 2006). This report contained an overview of the current state of library catalogs as well as recommendations and a blueprint for the future. In creating her report, Calhoun conducted an extensive survey of the literature and interviewed a number of prominent individuals in the cataloging community.

Calhoun's report lists a number of requirements to be met if library catalogs, and to some extent libraries, are to remain relevant in a digital world. Not only do libraries need to more successfully integrate library catalogs with web discovery tools, Calhoun contends, but they must find significant methods for cost reduction (including automating many processes) and research all possible means of sharing and repurposing catalog and authority data in new ways, such as through the supporting of mass digitization projects. The report also strongly calls for experimentation with FRBR, and includes some extensive commentary regarding FRBR drawn from her interviews (Calhoun 2006). While the interviewees had some issues with the FRBR model, they still strongly supported the ideas of aligning RDA with FRBR and finding ways to FRBRize catalog data.

The need to lift the catalog discovery experience to the "network level" was perhaps best articulated by Lorcan Dempsey, who has commented that **"**increasingly, we need to think of the catalogue, or catalogue services and data, making connections between users and relevant resources, and think of all the places where those connections should happen" (Dempsey 2006). In fact, Dempsey suggests the idea of FRBR bookmarklets as a way of drawing users back from other discovery environments such as Google to the catalog. For example, if a browser recognizes an ISBN in a Web resource, it could then use that data to query a library catalog to see if the book is available. This expanded notion of the catalog and thinking of catalog data as a way of making connections between resources and users in places outside of traditional library systems is an idea we hope to investigate in some of our future work.

---

[7] http://www.loc.gov/marc/marc-functional-analysis/tool.html

FRBRized Data and Metadata Models

As evidenced by Zhang and Salaba, many researchers are challenged by how to turn the FRBR model into workable cataloging rules or metadata structures that can be used in scalable real-world systems. The evaluation of the FRBR model and its implications for metadata has both a wealth of research literature and an active life in the blogosphere as well. One useful interpretation of the FRBR model as it relates to current metadata models and applications can be read at the blog "FutureLib"[8] co-created by several major players in the cataloging community. This blog includes a document "Framework for A Bibliographic Future" that sums up quite well some of the major challenges in working with FRBR, including that:

FRBR defines data elements in its attributes, but it needs to be restructured in a way that allows the development of different levels of granularity and that promotes extensibility of the schema, both over time and across communities. Ideally, the schema would be expressed in one or more machine-readable formats that facilitate its use by both people and computer applications. [9]

The idea of a machine readable FRBR schema that could be extended by different communities would certainly assist us in the cataloging work we are currently pursuing.

Maja Žumer has also emphasized that FRBR is a conceptual model, rather than a data model, which is why a number of researchers have implemented vastly different FRBRized systems (Žumer 2007). The LC WGFBC has also advocated for more practical modeling of FRBR, and has recommended the LC, the JSC, and the Dublin Core Metadata Initiative should, "work jointly to specify and commission exploratory work to model and represent a Bibliographic Description Vocabulary, drawing on the work of FRBR and RDA, the Dublin Core Abstract Model, and appropriate semantic Web technologies (e.g., SKOS)" (Amelung, et al. 2008). This section will consequently look at assorted interpretations and re-castings of the FRBR model, with a focus on digital library and cultural heritage research efforts.

FRBR as a Conceptual and Practical Model

Some recent work has probed the conceptual validity of the FRBR model rather than its practical cataloging and system applications. Work by Renear and Choi has analyzed the concept of inheritance in the FRBR model and claimed that the attributes of the work are not necessarily inherited by all of the lower levels such as the expression, manifestation and item entities (Renear and Choi, 2006). Allison Carlyle has also assessed the implications of the FRBR model for bibliographic description, and states that "FRBR is a conceptual model with the primary purpose of improving cataloging records (a product), cataloging (a process), and catalogs (a technology)" (Carlyle 2007). She ultimately posits that any evaluation of the FRBR model will depend on how successfully a FRBR implementation meets specified user needs declared at the outset.

Overall it is the expression level entity has caused the largest amount of confusion, to the point where the IFLA FRBR Working Group has directly addressed this issue by announcing a planned update to the expression entity that should be incorporated into the official model soon. As the original text of the model indicated that "any change in intellectual or artistic content constitutes a change in expression……no matter how minor the modification may be,"(IFLA 1998) the Working Group realized this had created a practical impossibility. As they recently acknowledged:

---

[8] http://futurelib.pbwiki.com/
[9] http://futurelib.pbwiki.com/Framework

Except for exact photographic reproductions, the only way to be absolutely positive that there are no tiny differences in the words contained in two different manifestations of the same textual work is to compare the two manifestations word by-word. This is obviously something that is not going to happen in any normal cataloguing situation, with the unfortunate result that under a strict interpretation of the definition, the entity *expression* could never actually be applied, thus completely loosing its potential for organizing the displays of those works with many manifestations (Riva and Cato 2007).

While the changed expression entity description has not as yet been added to the official FRBR document, this change will hopefully add further clarification.

A number of other challenges have also been issued to the FRBR model, including the argument that it fails to adequately support the cataloging of serials or other aggregate works. Kristin Antelman points out that the FRBR model has largely been developed with the traditional monograph in mind, and that mechanisms such as the main entry heading and uniform title prove to be weak identifiers for serials (Antelman 2004). Despite these issues, Maja Žumer notes that one positive development is the formation by IFLA of a FRBR working group on aggregates, which "will deal with anthologies, series, augmentations, and journals (all composites of individually created dependent/independent works), which are neither treated consistently nor in detail in the original model" (Žumer 2007). Since a great deal of our collection includes both anthologies and series, we will watch for this work with great interest.

The FRBR Model, RDF and Ontologies

The importance of testing the validity of the FRBR model with real world catalog data that was raised in the research by Zhang and Salaba has also recently been addressed by Martha Yee, who has created a set of cataloging rules specifically designed to implement the FRBR model with examples in RDF (Yee 2007). A preliminary model of FRBR concepts in RDF was also designed by Ian Davis, Richard Newman and Bruce D'Arcus.[10] At the same time, assessment of the FRBR model and its relevance beyond bibliographic data has also been conducted outside of the library community in recent years. Recent work by Pasin, et al. details how FRBR has been repurposed in the development of a ontology for philosophical knowledge (Pasin, et al. 2007). Perhaps the most significant effort in remodeling FRBR is the current effort to harmonize the CIDOC-CRM and FRBR through the creation of the ontology FRBRoo (Doerr and LeBoeuf 2007). A recent blog entry by Karen Coyle, presents an excellent overview of some of the differences between the Yee, Davis and FRBRoo models.[11]

The Perseus Project is currently evaluating the potential of the FRBRoo ontology in the shaping of our own catalog data.[12] A number of classes introduced by the most recent iteration of the FRBRoo ontology have been helpful, chiefly their expansion and refinement of the concept of the "Work" (Doerr and LeBoeuf 2007). Their initial class "Work" (F1) has been divided into two classes: an "Individual Work" (Class F46) or "Complex Work" (Class F21). While an individual work is said to be completely realized through a single "Self-Contained Expression" (Class F20), a "Complex Work" can consist of members that are either "Complex Works" themselves or "Individual Works." They furnish one example that is especially relevant for our purposes, that of an anthology of poems, which while a "Work" in its own right, also contains numerous expressions of other "Individual Works." This does not inherently make an anthology a "Complex Work," however, for although this class inherently "comprises works that have more than one work as members," the model specifically states that an anthology that does not exist in more than one version is not a "Complex Work" but rather an "Aggregation Work" (Class F48). An "Aggregation Work" "comprises works whose essence is the selection and/or arrangement of expressions

---

[10]http:/vocab.org/frbr/core
[11]Coyle, K. "Interpretations of FRBR Classes." Coyle's InFormation. Feb 12, 2007.
http://kcoyle.blogspot.com/2007/12/interpretations-of-frbr-classes.html. .
[12] For more information on these efforts, please see (Babeu, et al. 2007)

of other works. This does not make the contents of the aggregated expressions part of this work, but only parts of the resulting expression….," thus our collection contains many of this type of work such as *The Oxford Book of Latin Verse* or the *Anthologia Lyrica Graeca*.

The class "Complex Work" it seems is intended to better capture the nature of works such as Shakespeare's *Hamlet*, or Homer's *Iliad*, that are complex due to having not just many different editions and translations, but adaptations, derivations, commentaries, awful Hollywood film versions, etc. Consequently, "the boundaries of a Complex Work have nothing to do with the value of the intellectual achievement but only with the dominance of a concept. Thus, derivations such as translations are regarded as belonging to the same Complex Work, even though in addition they constitute an Individual Work themselves" (Doerr and LeBoeuf 2007). The vast majority of our collections thus constitute "Complex works" sometimes published individually and other times published as part of "Aggregation Works."

The FRBRoo ontology also helps to refine the entity "Expression" (Class F2) by dividing it into "Self-Contained Expression" (Class F20) and "Expression Fragment" (Class F23). While "Self-Contained Expressions" are regarded as a complete whole, one important aspect of this class is that "the quality of wholeness reflects the intention of its creator that this expression should convey the concept of the work. Such a "whole" can in turn be part of a larger 'whole'" (Doerr and LeBoeuf 2007). Indeed, many classical works that were considered as complete by their authors' only became known later through parts of larger works by different authors. Additionally, the "Expression Fragment" which is defined as a class that "comprises parts of Expressions and these parts are not Self-contained Expressions themselves" such as Sappho's poems, is very relevant to the nature of the cataloging we have been doing. FRBRoo also has a number of properties that may be useful in expanding the original FRBR-ER model including, R11 "is composed of", R12 "has member", R56 "is Realised in", R58 "is derivative of", R63 "incorporates" As this brief analysis illustrates, FRBRoo has a great deal of potential, principally in assisting those projects looking to clarify their data models before pursuing a FRBRization project.

While FRBRoo is still very much in its early stages, a recent article by Martin Doerr has looked at some of its potential for libraries, with the specific example of the European Digital Library[13] (Doerr 2007). Doerr advocates that the best strategy for creating a core ontology that will enable semantic interoperability in the library community is through the harmonization of various alternatives such as the CIDOC CRM and FRBR. In his summarization of the FRBR and CIDOC CRM harmonization efforts, Doerr related how the one of the groups' conversations led to the need to directly address one of the shortcomings of the initial FRBR model through expanding the class of the work:

Another important part of the discussion had to do with work containing other work, such as collections of poems. In the course of discussion however it was recognized, that virtually any book is composed of multiple, distinct works: the text, the illustrations, the editors work on lay-out, type phase etc. The latter was widely ignored in FRBR, and discussions tend to confuse the question of which contribution and work is most relevant with how to make the necessary distinctions in a model. This situation demanded for a general model explicating both the individual contribution and the unity of the integrated product (Doerr 2007).

The question of how best to model and efficiently structure catalog records for "Aggregation Works" remains an ongoing challenge for our current research project.

---

[13] http://www.theeuropeanlibrary.org/portal/index.html

FRBR and Complex Humanities Texts

The FRBR model has also proved inspirational for related modeling work within the domain of complex humanities text structure. The Canonical Text Services or (CTS protocol), which is currently in development by a team of classicists and digital humanists, is being designed to "define a network service enabling use of a distributed collection of texts according to notions that are traditional among classicists" (Porter, et al. 2006). The CTS protocol "extends the FRBR hierarchy both upwards and downwards, upwards by "grouping Works under a notional entity called 'TextGroup' " and "downwards, allowing identification and abstraction of citable chunks of text (Homer, *Iliad* Book 1, Line 123), or ranges of citable chunks (Hom. Il. 1.123-2.22)" (Babeu, et al. 2007). This extension of FRBR for classics makes clear how different communities may wish to adapt the FRBR model to meet their own individual needs.

As evidenced by the CTS and FRBRoo examples, the complicated structure of many humanities texts can require advanced models such as FRBR and sometimes beyond. Some recent related work has researched how complex aggregate works might best be modeled in current digital library systems (Buchanan, et al. 2007, Gow, et al. 2006). Buchanan, et al. reported on the difficulties of representing aggregate works such as encyclopedias, historical serials, or anthologies of collected poems in a typical digital library (DL) system architecture. For example, in systems such as DSpace, documents are represented as one or more binary files with associated metadata, meaning if a work is bound in two separate volumes, there are now two separate documents in the library which must then be linked for browsing and searching. To help implement better representation of humanities documents with DLs, they thus developed a more sophisticate model of an aggregate work:

We presuppose the existence of some 'document unit'. An ordered series of these may be collected together to form an aggregate, which can in turn form part of another aggregate. So, aggregate works are ordered trees with documents units at the leaves. … Units may themselves be documents, or parts of documents. Where aggregation stops and internal document structure commences is not necessarily clear, and often depends on library users' needs and the resources of the library administration. (Buchanan, et al. 2007)

In further developing their model they also impart a number of useful aggregate features to consider including: homogeneous aggregation (the articles in a journal), heterogeneous digital forms (an aggregate digitized over time), serial aggregation (a multi volume series of works), binding aggregation (a single work printed in multiple volumes), composite aggregation (a work published in parts with each part bound in a different aggregate, such as serialized fiction in a newspaper), containing aggregation (a small work never published on its own but contained within a larger work), heterogeneous aggregation ("a work created from units of diverse types," such as journals containing articles of different types), supplementary aggregation (augmented works or an original work supplemented with introduction, commentary by other authors), incomplete aggregation (series that were never finished), and finally variable aggregation (the New Testament).

As demonstrated by the above list, the types of aggregation found within humanities materials is quite complex. Of even greater issue, is developing a DL architecture that can deal with these types of aggregates. As the authors note:

The boundary between dealing with external and internal document structure is not fixed, and many of the issues discussed above as the aggregation issues may also occur *within* a given document… What is important, from the view of a DL system, is that the treatment of internal and external aggregation are treated consistently in the DL architecture and also in the user interface, to ease the task of readers and librarians alike. (Buchanan, et al. 2007)

After reviewing the major digital library softwares available, the authors conclude that support for aggregate works is limited at best, mainly for the more complicated forms (containing and composite aggregations), and outline several preliminary solutions developed for the Greenstone Digital Library.

Gow, et al. pursued similar explorations, with an expanded focus on what types of structural models would best support the encoding of complicated humanities texts (such as aggregate works) and what models are currently supported by digital libraries (Gow, et al. 2006). The authors submit that for document structure in the humanities domain: "there is a need for fine-grained navigation, search and referencing within documents, as well as an emphasis on preserving the structure of original source material" (Gow, et al. 2006). They maintain that the most basic model available in digital libraries is the document model (such as in DSPACE) where no internal divisions are allowed within a document. Another simple model is the classification model, where larger documents are separated into sub-units such as chapters, with each sub-unit then indexed as separate document and combined through classification in the library. The chapter model that treats each chapter as a separate work but integrates them together for searching is another simple model. The authors comment that while hierarchical models (such as METS files[14]) best serve the needs of humanities document and many DL systems can import these formats, few systems exploit the internal document structure. Gow, et al. believe the DL systems will need to support documents that have multiple internal hierarchical levels, a feature now partially implemented in Greenstone 3. While our project has not yet moved to the level of deciding what DL we may use to support our "FRBR Inspired" catalog, the lessons learned by these research projects have been helpful in refining our thinking of document models.

FRBR Metadata Models for Digital Libraries

The number of investigations of how FRBR can be used to assist in the modeling of metadata for digital library collections has also grown in recent years. Dunn, et al. have commented on using FRBR as a model for metadata in a musical digital library (Dunn, et al. 2006), while Weng and Jia examined how using the FRBR entity-relationship model might improve access to digital cultural materials such as digitized images (Weng and Jia 2006). Weng and Jia found that although many special collections have been digitized in recent years, few have been cataloged, which has created a situation where there is very limited intellectual access to these collections. After surveying numerous digital collection sites including American Memory, they put forward that: "the attributes of work and expression entities presented in the IFLA FRBR model should be applied differently for event-based digital cultural materials. Defining work- or expression-level entities under the event or theme will be more logical than under author and/or title for cultural materials. Doing expression-level cataloging might also work for this type of materials" (Weng and Jia 2006). The treatment of "events" as works whose expressions include various "related intellectual realizations depicting the event" (such as photographs) is an interesting interpretation of the FRBR model, one that as yet has not been evaluated by cultural digital heritage collections.

Recent work by Chamnogsri, et al. has focused on adapting the FRBR model into a metadata scheme that can be used to represent palm leaf manuscripts (PLM) and enhance their online retrieval in a digital library (Chamnogsri, et al. 2006). In their modified FRBR model, physical PLMs were represented as works, the language of the PLM was used to indicate the expression level, the manifestation level was applied to the formats in which each expression was available, and items were individual copies of a single format. The process of this modeling led the authors to realize that "creating successful document representations (metadata) for a digital library requires a useful model to help clarify what the digital library project is trying to do with metadata, what functions are required, how the metadata record should be structured, and what data elements it should contain." The FRBR model proved to be such a model.

Similarly, the ECHO Project (European CHronicles On-line), a group developing and testing a video digital library system for historical documentaries from four European archives, has also tailored the

---

[14] METS stands for "Metadata Encoding and Transmission Standards" and was developed by the LC, for more please see http://www.loc.gov/standards/mets/

FRBR model for its metadata purposes (Gennaro 2008). They have created a new series of entities AVDocument, Version, Media and Storage, which correspond to the Four FRBR Group 1 entities respectively. This metadata model has assisted them in the creation of Regia, an application for editing metadata for audiovisual documents cataloged as part of the ECHO project.

The FRBR model has also proved important in the development of cross-repository services for digital repositories/libraries, chiefly in how to model compound digital objects. The NSF funded Pathways project created an interoperable data model to support the exchange of data objects between repositories as well as to promote cross-repository services. Their data model included "entity elements" that model "the abstract aspects of digital objects and align with works and expressions in FRBR" and "datastream elements" that "model the concrete aspects of a digital object" and align with items in FRBR (Bekaert, et al. 2006). This early work has now been supplanted by the modeling efforts of the Open Archives Initiative-Object Reuse and Exchange (OAI-ORE) project which also shows some potential applicability for FRBR implementations. A recent outline of the model included an interesting example of how their concept of a "compound object" utilizing named graphs and "resource maps" could be used to model and link versions of the same book that has been digitized by several digitization projects, as well as link that book to related commentary and reviews (Lagoze and Van De Sompel 2007). The OAI-ORE has recently released alpha guidelines and how to both model and share "compound information objects."[15]

Other recent work in this area has been documented by Allinson, et al. in how the FRBR model has helped shape the creation of a Dublin Core application profile for scholarly works that are deposited in digital repositories (Allinson, et al. 2007). This application profile is being developed in the hopes of allowing the aggregation of content from multiple repositories and supporting different services. Their model is focused on Eprints, or scholarly research texts, and they modified a number of the FRBR entity and relationship labels. In contrast to the earlier IFLA FRBR model statement, their natural language model can be specified as: "A ScholarlyWork may be expressed as one or more Expressions. Each Expression may be manifested as one or more Manifestations. Each Manifestation may be made available as one or more Copies" (Allinson, et al. 2007). The authors believe that, "this application profile represents a relatively innovative approach to metadata, taking as it does the FRBR model and applying it to scholarly works. By making use of the benefits afforded by the DCMI Abstract Model, the profile is able to group descriptions of multiple entities into a single description set. Overall this approach is guided by the functional requirements identified above and the primary use case of richer, more functional, metadata." This research represents the important potential role FRBR plays in helping build richer models for expressing more advanced metadata needs.

FRBR Implementations & Experiments

The importance of experimenting with the FRBR model including designing real world scalable systems was recently brought to the fore by the initial recommendation of the LC WGFBC draft report, to suspend all work on RDA until further experimentation with FRBR was conducted (Amelung, et. al 2007). The language regarding RDA in the final report clarified but still reiterated this point, asserting that work on RDA should be suspended until "more, large-scale, comprehensive testing of FRBR as it relates to proposed provisions of RDA has been carried out against real cataloging data, and the results of those tests have been analyzed" (Amelung, et al. 2008).

---

[15] http://www.openarchives.org/ore/0.1/toc

These statements have caused a fair amount of controversy, as many responders pointed out that a great deal of FRBR experimentation has already been conducted.[16] Jonathan Rochkind, a systems librarian at Johns Hopkins University, commented on his blog that rather than abandoning RDA, all work on RDA should be viewed as part of the investigation of FRBR. He urges that "we need to get away from the idea that any kind of standard can somehow be produced in isolation from practice in one monolithic barrage and then never be returned to."[17] As the following section will establish by considering recent efforts to build FRBRized catalogs or systems, and as our own research has indicated, creating FRBRized catalogs and systems will take a process of exploration and iteration, rather than one monolithic effort.[18]

The terms FRBR catalog, FRBRized system, FRBR implementation, are often used interchangeably, with some confusion in the library literature. Salaba and Zhang have put forward that, "the majority of the current FRBR systems and prototypes are considered FRBRized systems; very few are new systems independent of older practices…. A main issue is that most FRBRization projects have used only bibliographic data for the identification of works, whereas authority data are also valuable sources for work representation and identification" (Salaba and Zhang 2007). Our current work involves determining how to make the best use of both authority and bibliographic data for work identification.

FRBR Experiments in Traditional Library Systems

Some of the earliest experiments with FRBRization were conducted by members of OCLC Research, who in a number of articles have related information on experiments with algorithms to group bibliographic records in WorldCat into works and expressions (Hickey, et al., 2002; Bennett 2003). The work conducted by Hickey, et al., focused on the identification of expressions. The authors first manually extracted from WorldCat a set of records representing the monograph *Humphrey Clinker*, and analyzing the ability of their algorithm to discover expressions within this set of records. They learned that their algorithm was able to identify 28 expressions in the set versus the 41 located through manual inspection, and was able to pull out 10 of 11 identified manifestations. The authors ultimately decided that due to the difficulty of algorithmically identifying expressions, their future work for the time being would concentrate on the identification of works rather than expressions (Hickey, et al., 2002).

Other early experiments with FRBR were also performed by the AustLit project based at the National Library of Australia (Ayres, et. al 2003). For their work, they augmented the FRBR Model with INDECS event modeling, in order to allow them to add temporal attributes to their data. Their research also introduced the concept of the "Superwork" a larger class that encompassed twelve different work types. A similar approach was detailed by (Vassallo 2006), who used FRBR and topic maps to enhance user navigation through cultural heritage materials. The Library of Congress has also conducted some preliminary work in experimenting with FRBR through the development of their FRBR Display Tool, which through the use of XSLT and MARC records allows libraries to visualize their bibliographic data into meaningful displays of works, manifestations, and expressions (Radebaugh and Keith 2005).

Recent research has illustrated how even partial implementations of a FRBR model might be of great assistance to catalog users. Simpson, et al. have reported on a project at the University Of Florida Smathers Library conducted in order to support better linking of catalog records for multiple manifestations of the same title (Simpson 2007). The library had acquired table of contents (toc) data to

---

[16] These responses have included a response from the RDA Committee, available at http://www.collectionscanada.gc.ca/jsc/cop-lcwgbibcontrol.html, and a formal response from OCLC, http://staff.oclc.org/~levan/LC%20WG%20Report%20Comments%20OCLC%2020071214.pdf

[17] Rothkind, J. "FRBR Imperfect? So Then?" Bibliographic Wilderness. Nov 28, 2007. http://bibwild.wordpress.com/2007/11/28/frbr-imperfect/

[18] This section will not attempt an exhaustive overview of all FRBR implementations, but for an extensive list of FRBR implementations, please see Chapter 3, "FRBR Implementations" in (Eden 2006)

enrich bibliographic records for many print books, a great number of which they already had in electronic format, yet only the record for the print book included this enhancement. The authors sought to find the best way to link the records for these two manifestations. Their article lists three important obstacles to creating FRBRized catalogs: 1) library management systems frequently cannot manage links that exist among bibliographic records 2) catalog records often lack the data necessary to reflect relationships 3) library management systems typically cannot make use of the data even if it is available. Despite these caveats, the authors agree with Jennifer Bowen that catalogers should take whatever small steps are possible in collaborating with both vendors and other libraries to implement at least some aspects of the FRBR model. Consequently, the authors developed a local solution to solve their problem through the utilization of a non-MARC system specific field that allowed direct connections among bibliographic records, holdings records and item records, an approach they describe as "guerilla cataloging." While Simpson, et al. recognized the limitations of their solution, they sought to highlight a creative solution to the linking problem, one that might inspire other institutions to evaluate similar approaches

Another "partial" implementation of FRBR principles to improve the user experience was conducted by Chew C. Naun of the University of Illinois at Urbana Champaign (UIUC), who recently described work utilizing the FRBR model in the design of online journal finding aid. Naun's research imparted that "the FRBR model was helpful in analyzing 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"(Naun 2007). In order to help their users more quickly find electronic resources, UIUC developed a tool called Online Research Resources (ORR), which provided extensive information on online article databases, journals and reference works. The title of the serial was used as the main representational unit. The original library plan was to use ORR as a knowledge base for reference linking, but they soon learned that no one database (serials management system, library catalog, link resolver database, etc.) contained all of the information they needed to allow ORR to support the four user tasks specified by FRBR: find, identify, select and obtain. To populate the ORR database the library had to import data from numerous different sources and create a quality hierarchy at levels of individual fields to indicate preferred sources for each data element.

Nonetheless, within the ORR system, it was still hard for their users to determine expression level attributes, as the system assimilated all electronic versions to one expression. The ORR system resembles many FRBRization projects according to Naun because it populates the database algorithmically by taking existing data, restructuring it and supporting hierarchical views of that data. Through FRBRizing their data, ORR was able to bring electronic serials content from different providers together under a single record listed under serial title. Links could then be supplied between electronic and print versions, and between earlier and later titles of a serial. While much of the clustering was achieved through matching of ISSNs, Chew declared that the ISSN alone does not function as a useful work identifier, as journals are typically assigned new ISSNs for electronic versions. The work presented by Naun demonstrates the importance of current work in improving the FRBR model for aggregate works.

Unfortunately due to the limitations of many current library systems, partial rather than fuller implementations of FRBR will likely continue for some time. The most comprehensive overview to date of FRBR implementations is available in (Yee 2005), and the interested reader is referred there for fuller looks at several implementations such as the Library of Congress FRBR Display Tool[19], OCLC Fiction Finder[20] and VTLS Virtua[21] and the now defunct RLG's RedLightGreen. Many FRBRization projects have also focused exclusively on the bibliographic record, whereas much of the data that represents work level data can be located only in the authority record. Another issue is that much of the data that could be

---

[19] http://www.loc.gov/marc/marc-functional-analysis/tool.html
[20] http://fictionfinder.oclc.org/
[21] http://www.vtls.com/Products/virtua.shtml

used to discriminate between different expressions that are found within catalog records is human but not machine readable, and thus not easily processed by algorithms.

A slightly more recent assessment of two prominent FRBR prototypes was conducted by Pisanski and Žumer, who reviewed OCLC's FictionFinder and an experimental prototype designed by the Australian National Library's called LibraryLabs (Pisanski and Žumer 2007). While FictionFinder focuses on the concept of the work, LibraryLabs experiments with FRBR as part of a larger experiment with library systems. Pisanski and Žumer explain that neither prototype fully follows the FRBR model, partially due to issues with both the FRBR model and cataloging data, but also emphasize that: "It has to be pointed out that FRBR is neither a standard nor a data model. In other words, FRBR in no way implies what implementation should be like" (Pisanski and Žumer 2007). The authors related that the LibraryLabs prototype was not limited to just books but included movies and other materials. This prototype grouped FRBR data at various levels, included a new group called "superwork" where top level records were grouped together, and used form and language attributes to differentiate between numerous works and expressions. In contrast, the OCLC FictionFinder prototype uses normalized title/author as the key for clustering records, and works were ranked according to the number of libraries that owned it. One major issue the authors had with FictionFinder was that it focused too exclusively on the concept of the work with no easy ability to sort results by manifestation level data, such as a specific publisher or illustrator name.

In a comparison of both prototypes, Pisanski and Žumer felt that both concentrated too much on equating the attribute of language with different expressions, since two different translations of a work in the same language would still be two different expressions. Both prototypes have issues the authors contend because "algorithms for eliciting FRBR structure" will only work as well as the bibliographic records on which they are based. For fully successful FRBR implementations, they believe that a number of things will be needed, including displays that involve better overviews of works, of expressions, as well as the relationships between connected or similar works, such as relationships between works in a series or the components parts of a larger work. The authors ultimately believe that, "based on these and other prototypes, it can be seen that the FRBR conceptual model has not really been accepted in practice. Not only is extracting FRBR structure difficult, in some cases it may even be impossible, particularly with limited resources. The model itself is lacking especially in terms of boundaries of an expression" (Pisanski and Žumer 2007).

Pisanski and Žumer also, however, tender one other interesting proposition, that the model of LibraryThing,[22] a social cataloging site that allows users to catalog their own books and tag them, including the ability for users themselves to "merge and separate wrongly identified groups of bibliographic entities" might be worth exploring, although the authors feel that the records should still be subject to librarian review. Social cataloging models such as those employed by LibraryThing represent an interesting way to combine the best of algorithms and distributed human labor in implementing basic levels of FRBR functionality.

The LibraryLabs prototype described by Pisanski and Žumer has also been explained extensively in several recent articles published by staff at the National Library of Australia (Dellitt and Fitch 2007, Dellitt and Boston 2007). As part of the general process of rethinking the library catalog, they have experimented with FRBR grouping of manifestations at the expression and work level. Dellitt and Fitch point out that both OCLC and LibraryThing maintain databases of titles grouped by ISBN, and that by incorporating this information into the catalog, the number of duplicates in results lists can be greatly reduced (Dellitt and Fitch 2007). Ultimately they hope to create a "FRBR view of bibliographic resources" but admit that there are current many challenges with merging and clustering MARC records

---

[22] http://www.librarything.com/

in order to accomplish such goals. Consequently the National Library of Australia has conducted a series of experiments with clustering MARC records by a variety of facets (e.g. author, title, subject, genre, etc.) in their union catalog LibrariesAustralia. These experiments have been conducted in order to allow more sophisticated relevance ranking (Dellitt and Boston 2007).

Yet another project that is experimenting with FRBR is the Extensible Catalog project based at the University of Rochester. They have recently announced that the Mellon Foundation will fund their second phase of development of the Extensible Catalog or XC, an open source system that will unify access to traditional and digital library resources and include support for FRBR. According to their recent report "XC will provide an extensible metadata platform supporting multiple schemas that can be searched simultaneously to support FRBR-like functionality and navigation" (Lindahl, et al. 2007).

Vendors who serve library systems have also begun to pay increased attention to the creation of FRBRized cataloging systems. The AquaBrowser software supplied by MediaLab is currently looking at different techniques to support FRBR views of library data.[23] Another library catalog software offering that is receiving a great deal of attention is the Endeca catalog software, which includes several sophisticated features such as faceted browsing, relevance ranked results, improved subject access, and auto-spelling correction. A recent implementation of Endeca at North Carolina State University has utilized not only these features but also examined how a "record rollup" feature provided by the software might be used to support the FRBR model in their catalog (Antelman, et al. 2006). Since this feature needs a "rollup key" they have opted to use the OCLC work identifier in bibliographic records to try and create "work-level record displays."

FRBR and Digital Library Systems

One of the fullest reviews of how FRBR might be utilized in digital library systems was related by George Buchanan, who developed a FRBR architecture implementation for the Greenstone Digital Library System (Buchanan 2006). The main focus of his research was on the construction of FRBR capabilities for a standard digital library architecture such as Greenstone. Two of the most significant challenges he documented were creating the underlying data structures and integrating the hierarchical concept of a work into the interface. After surveying some previous attempts at the automatic detection of FRBR entities, Buchanan advised that too many efforts focus exclusively on automatic solutions:

What this all demonstrates is that any expectation that the FRBR tree can be created without human support is gravely misplaced. Though frequently the matching different *items* of the same *manifestation* is readily achieved, abstracting *works* and *expressions* is fraught with problems. Given that the automatic identification of works is unreliable, using this approach for retrospectively adopting FRBR in a DL is deeply flawed. Thus, other approaches need to be investigated. In this paper, we present an approach that complements existing DL architectures, permitting the gradual adoption of full FRBR support for a digital collection (Buchanan 2006).

Buchanan believes that FRBR can best be used as a framework for linking content across different digital library systems. He made a number of changes to the Greenstone interface, including the ability to choose "other copies" or "other editions", allowing the user to view all copies of one expression or view copies of different expressions. He built tools that he hoped would help librarians construct "FRBR trees" and through his research has reached the conviction that "FRBR information will typically need to be encoded by a librarian." The "FRBR tree" contains information not just relating works to their manifestations and expressions but also to their authors and publishers and relates different works together, such as commentaries on original works.

---

[23] http://www.medialab.nl/blog/?p=30

To support his conclusions, Buchanan designed an editor interface for creating FRBR trees as a Java software component delivered as part of the Greenstone Librarian Interface. As part of this work, the question of unique identifiers became critical, and a URI schema was used to create DOIs for all works within the digital library. All FRBR data was stored within a MySQL database, and work, expression, manifestation and item elements were all encoded into separate tables in the database. The items table was used to link the FRBR database to actual works in the digital library. This approach also utilized the METS metadata framework and used URIS to store references to each copy. Buchanan learned that the nature of data storage in Greenstone easily supported the hierarchical nature of their FRBR data.

Similarly the Greenstone architecture that separates the user interface from underlying services proved to be important as well, because due to this architecture, the presence of a FRBR module made no change to the underlying DL services. Buchanan stated that it is wisest to implement FRBR content or databases as a separate module, not only to allow its use across different digital libraries but also to allow the aggregation of different FRBR databases within one digital library. He also noted some initial experiments in how FRBR might be used to integrate separate digital libraries, such as through the use of standard identifiers for works stored in a consistent format. One major finding of Buchanan's confirmed by our own work is that: "FRBR requires rich data, and a considerable investment in time and effort is needed for even limited, specialized domains" (Buchanan 2006).

FRBR, Identifiers, Authority Control and the Semantic Web

Unique Identifiers and Authority Control for Names

One major question regarding the practical implications of FRBR is whether or not unique identifiers are needed to identify not just works, but entities at every level of the FRBR hierarchy. A related issue is the debate regarding unique identifiers for individual authors, how these might be created and managed, and how current authority files might be better utilized in FRBRization efforts that tend to focus on bibliographic records only. Authority files are records created by libraries that specify the chosen or controlled heading for personal names, work titles or subject headings, and list variant forms that are linked to the authorized heading. They support the exact identification of entities so that catalogs can effectively collocate all works by one author, all books about one subject, etc. A good series of definitions have been offered by Jonathan Rochkind:

The purpose of authority control is to make sets of objects. The typical library examples are the set of all works written by author X; the set of all 'editions' (aka 'versions', aka expressions/manifestations in FRBR terminology) of a given work; the set of all works reprsenting (sic) a given subject.

We can also say that the purpose of authority control is in establishing un-ambiguous relationships between entities– or in practice, it's really more clear to say between our records for given entities.[24]

Rochkind explains that traditional methods of authority control often involved the use of headings (such as name strings, or name-title strings like "Shakespeare, William, 1564-1616. Antony and Cleopatra) to serve as work titles, and these headings could then be used in any record for an expression of this work. This is a problematic approach at best, as Rochkind points out, because it is language specific. But since some label needs to be displayed to a user, Rochkind advocates using "dumb identifiers" as unique identifiers for names, titles, etc. in authority records, while also using the headings stored in those records as needed to support context sensitive user displays in catalog systems.

---

[24] Rochkind, J. "The Purpose of Authority Control." Bibliographic Wilderness. August, 8, 2007.
 http:/bibwild.wordpress.com/2007/08/08/the-purpose-of-authority-control/

The limitations of catalogs using one authorized heading or label have been widely recognized with differing solutions proposed to alleviate this problem. Many systems are moving from authority control to the concept of access control, where the main goal is name disambiguation and all name variants are linked with no one name serving as an official heading, rather than strict enforcement of the use of one authorized heading only.[25] Nonetheless, the desire of librarians to be able to make better use of authority files and other library content standards and to be able to integrate them into their metadata tools has recently been recognized in a recent survey of metadata experts and the features they desired in automatic metadata generation applications (Greenberg, et al. 2006). This same survey also conveyed a desire for more FRBR support to be built into metadata creation tools, including that the "system should support automatic linking of metadata records, including referencing and cross referencing between related items."

The Library of Congress Name Authority File (LC NAF) is a heavily used resource in the area of authority control, and recently the LC WGFBC has stressed the importance of greater distributed creation of authority records and the need to create an international authority file. The report urges that all means to speed up and improve the process of finding, creating and linking authority records, including automated means such as name disambiguation algorithms, should be explored.[26] One important recommendation by the LC WGFBC is to make the LC NAF available as a Web resource, "for downloading or linking to through various Web service interfaces" (Amelung, et al. 2008). Our catalog work has made heavy use of the LC NAF, including linking our catalog records to an available version of it hosted by OCLC.[27]

One important project that is currently attempting to make better use of name authority files is OCLC's WorldCat Identities. Through this project, the majority of entities in the LC NAF now have their own linkable web page, which has been populated by data from WorldCat, including alternate names, lists of work titles, a publication timeline, related authors, and a tag cloud of subject headings. The authority records that we are currently creating are all being linked to their appropriate WorldCat Identities page.[28] WorldCat Identities does face some issues, however, with duplicate records and inconsistent names as its data was drawn from WorldCat and thus reflects years of varying cataloging practice and inconsistent data.

There have also been efforts by IFLA to develop a similar model to FRBR for authority records with the creation of a model initially called Functional Requirements and Numbering of Authority Records (FRANAR), but now called Functional Requirements for Authority Data or FRAD (Patton 2004, IFLA 2007). The FRAD working group by IFLA is assessing what basic level of data should be required in authority records that will be shared internationally, including studying the feasibility of an International Standard Authority Data Number or ISADN. A recent article by Glenn Patton relates how FRAD reflects the growing trend in library authority work, where "the focus of international authority control has been shifting away from a single form of entry universally used for an entity to the potential for multiple forms depending on the needs of the user. And, the focus is shifting from the physical sharing of records toward a sharing of the intellectual product of authority activities" (Patton 2005). Patton also related that the IFLA working group had temporarily suspended work on the ISADN until the FRAD model was further developed.

---

[25] A full analysis of the concept of authority control versus access control in libraries is beyond the scope of this paper, but for more, please see (Lam 2002), (Bennett and Williams 2006), and (Tillett 2003).

[26] Possibilities in combing automated name disambiguation techniques with authority files have been examined by (Patton, et al. 2004) and (Tudhope, et al. 2006)

[27] http://alcme.oclc.org/eprintsUK/index.html

[28] For example, see the page on Aristotle, http://orlabs.oclc.org/Identities/lccn-n79-4182

Another significant effort to develop an identifier for both personal and corporate names is the, International Standard Party Identifier, or ISPI (Paskin 2006). The ISPI is a proposed ISO standard and is defined as a "new international identification system for the parties (persons and corporate bodies) involved in the creation and production of content entities."[29] The archival community is also working at developing a common standard called Encoded Archival Context (EAC), which will be used to support the creation and exchange of information about the creators of archival content in standard way.[30]

In contrast to some of these efforts, Barbara Tillett has stressed that creating unique identifiers may be less important or at least less tractable than many proponents have claimed (Tillett 2007). She believes that the costs in assigning and maintaining unique identifiers along with the associated issues of registration and international management far outweigh the current need for them. Tillett lists the main authority tasks associated with the need for unique identifiers as avoiding duplication, international sharing, easing burden of controlled vocabulary maintenance and customization of displayed names/ terms, but insists that a unique identifier will not necessarily solve all these problems. She points out that a unique identifier is not entirely language independent if you remember that many countries do not use Arabic numerals. Tillett does acknowledge, however, that one benefit of a unique identifier might be system independence, or that is, an identifier that could be used across systems.

The traditional approach to library authority control of using a single authorized name is recognized by Tillett as increasingly untenable in a digital world. In contrast to using a unique identifier, however, Tillett believes that one alternative might be an "authority data cluster of variants forms of names possibly with links to sources locating where the variant name was found" (Tillett 2007). In the end, Tillett posits that international standard numbers should not be considered at the "name level" but could work at the "authority record level." Returning to her idea of an "authority data cluster," Tillett supports the idea of creating authority files where "all of the possible name variations in all languages and scripts could be brought together or linked in various ways (union authority file, linked authority files, etc.)." The Virtual International Authority File (VIAF) project, a joint project of the Library of Congress, the Deutsche National Bibliothek (DNB), the Bibliotheque national de France, and OCLC, follows this approach.[31]

For current practical purposes Tillett advises reusing the control numbers that already exist in authority files for personal names, an approach that we have followed. At the same time, Tillett and Corey Harper have urged that the VIAF needs to go one step further than simply linking the major authority files of the world. Authority records need to available in a machine readable format (such as in formats designed for the Semantic Web), and personal names and subject concepts must have their own unique URIs. They explain that: "identifying authority information with URIs allows those URIs to be re-used to tie other descriptions of people to authority records, which in turn link to their works" (Tillett and Harper 2007a).

Identifiers for Documents: Works, Expressions and Manifestations

A whole host of identifiers that might be used to identify the different FRBR entities exist in different flavors and for different uses in the current digital world. In particular, the International Standard Text Code (ISTC), a newly approved ISO standard expected to be released in 2008, has been presented as a possible identifier for FRBR at the work level.[32] According to Michael Holdsworth "the ISTC is an identifier of "textual works" (Holdsworth 2008). This standard, however, is not without its controversy (LeBoeuf 2005). Nonetheless even if the type of identifier is not always agreed on, a number of features all identifiers should have are generally agreed upon, including the need for the identifier to be unique, for

---

[29] http://www.collectionscanada.gc.ca/iso/tc46sc9/27729.htm
[30] http://www.iath.virginia.edu/eac/
[31] http://orlabs.oclc.org/viaf/
[32] For further discussion of the ISTC and other identifiers as they may be used within the book publishing community to represent digital content, please see (Holdsworth 2008).

it to be persistent, for it to be interoperable and extendable, and for it to be actionable or resolvable to a specific entity. Douglas Campbell makes the important point that "an identifier will only exist as long as anyone remembers the declaration of association. Persistence of identifiers is not so much about remembering the identifier itself, but what it is associated with"(Campbell 2007). Several recent publications present excellent overviews of the field of identifier systems currently available, so we shall touch only briefly on several more important ones as they may relate to FRBR in this section (Hilse and Kothe 2006, Paskin 2006, Vitiello 2004).

Many discussions involving identifiers at different levels of FRBR have advocated creating URIs for entities at each level of the FRBR model (Svensson 2007). Formally defined, "Uniform Resource Identifiers (URIs, aka URLs) are short strings that identify resources in the web: documents, images, downloadable files, services, electronic mailboxes, and other resources. They make resources available under a variety of naming schemes and access methods such as HTTP, FTP, and Internet mail addressable in the same simple way."[33] Typically URIs have been implemented either as URLs or URNs. While URLs are typically well understood, URNs "are intended to serve as persistent location-independent, resource identifiers."[34] URNs were created as a means of providing names for resources rather than addressing them. As URNs carry a Namespace Identifier (NID) from a defined list maintained by IANA, this allows the integration of multiple naming schemes (such as the ISBN) as subsets of the URN namespace. Many of the most important identifier schemes currently in use rely on URNs (Hilse and Kothe 2006).

In 2001, the Deutsche National Bibliothek (DNB) introduced persistent identifiers in the form of a URN within an international name space labeled the "National Bibliography Number" (NBN). This namespace was first developed and registered by the National Library of Finland (Hilse and Kothe 2006). The DNB then expanded the assignment and management of their NBN program through the EPICUR Project and created a production level system. They undertook this project because "new possibilities for the creation, presentation, description, and publication of digital content had caused an increasing demand for permanent addressing, citation and unique identification of digital objects" (Schroeder 2006). The DDB introduced NBN management for its own digital resources as well for online publications managed by other German institutions. NBNs had to be public identifiers so institutions could assign them independently and structure them according to local needs, but also so they could support "effective information exchange through a uniform integration in bibliographic metadata formats and bibliographic data transfer formats" (Schroeder 2006).

The main goals for NBNs are to provide permanent access to an individual digital object, and the DNB hopes to tie their identifiers efforts to those of other international URN projects. Participating institutions in the EPICUR Project receive sub-namespaces assigned by the DNB but assign their own NBNs which are then registered and resolved centrally at the DNB. Institutions have to manage their own URLS and update them within the NBN system. The DNB hopes to eventually support more granular digital objects by allowing NBNs to be assigned to digital object parts that have actionable URLs. While NBNs have been designed to accommodate existing library identifiers schemes, it seems so far the EPICUR project is the only large scale implementation.

Perhaps the best known identifier is the Digital Object Identifier (DOI), which is managed and controlled by the International DOI Foundation. As Hilse and Kothe explain, "the DOI provides administrative schemes and workflows for the management and persistent identification of digital objects" (Hilse and Kothe 2006). DOIs are based on the Handle system initially developed by the Coalition for National

---

[33] http://www.w3.org/Addressing/
[34] http://www.faqs.org/rfcs/rfc2141.html

Research Initiatives (CNRI)[35], but DOIs have created additional technical and administrative layers to support interoperability among the various DOI implementations as well as to integrate previous legacy identifier systems. DOIs have been very successful and tens of millions have been registered with the resolution mechanism used a few million times a month.

At the same time, however, registering and being able to assign DOIs has a number of costs depending on the level of participation. One issue from a FRBR standpoint is that the fundamental purpose of a DOI is to persistently identify a digital object as a piece of intellectual property, not as a bibliographic entity. Another major question is at what level of the FRBR entity hierarchy would a DOI be assigned? As Giuseppe Vitiello relates, according to the FRBR specification, "a DOI identifies neither a work, nor an expression, nor a manifestation. Or else, it can identify them all, as there is no distinction among the categories listed in FRBR. A "meta-identifier", the DOI interprets convergence literally on communication networks and applies to any digital object, making relevant only what is worth copyright protection" (Vitiello 2004).

Several persistent identifier schemes have also been developed by the library community. One of the more prominent is the ARK Scheme or Archival Resource Keys, a system maintained by the California Digital Library. ARKs focus on supplying protocols and software to support a framework for providing and resolving persistent identifiers, with the end goal of serving as a lightweight means of resolving and delivering the appropriate metadata. In contrast to other systems focused almost exclusively on promising persistent identifiers "the ARK scheme does not assert that the identifiers are persistent since that depends solely on the service(s) behind them" (Hilse and Kothe 2006). One interesting fact about ARKs is that they allow implementers to encode variants of a digital object (a feature that could be interesting at the item level of FRBR).

A comprehensive look at all the different levels of identification needed within the digital library world was recently presented by Juha Hakala of the Helsinki University Library, who developed a seven level identification framework to analyze what needs to be identified in digital libraries and what currently available identifiers there are to support such identification (Hakala 2006). He declares that "not only do traditional identifiers have serious challenges in adapting themselves to networked publishing; there are also new and diverse needs for developing novel identifier systems." His seven levels include: Organisations, People, Collections and Services, Works and Expressions, Manifestations, Component Parts, and Search Attributes, but we shall focus our analysis on the first six levels. Interestingly, Hakala mentions that much of the effort behind work level identifiers such as the ISTC has been driven by collecting societies rather than libraries. "Few library system vendors have implemented the Functional Requirements for Bibliographic Records (FRBR) model," Hakala explains, "and, since there are no work level metadata records, there is no need for work-level identifiers either for the time being."

Despite his assertion that there may be no need for work level identifiers, Hakala does list some possibilities for potential identifiers at each of his seven levels. For organizations, he lists one identifier that has received little adoption, the International Standard Identifier for Libraries and Related Organisations (ISIL). For people, he mentions both the ISADN and the ISPI. In contrast to Barbara Tillett, Hakala believes that, "In a global environment where preferred name forms differ a lot, the only simple means for matching the authority records properly is to use a standard identifier for authors. Merging based on name forms only will not always provide correct matches, and is likely to be cumbersome, since there are many data elements that need to be matched." At the level of collections and services, Hakala comments that although there is a need for such identifiers no viable systems currently exist. Hakala also groups together works and expressions in his framework, and asserts that as currently defined the ISTC, "a semantic identifier that contains indications of where and when and by whom it was

---

[35] http://www.handle.net/

assigned" could be used for either level. For non-textual works, there are two current identifiers, the ISAN (International Standard Audiovisual Number) and the ISWC (International Standard Musical Work Code), but no identifiers for images.

The largest number of available identifiers Hakala details is at the manifestation level and includes the ISBN, the ISSN and the NBN, all systems that he fears are fraught with difficulties. He offers the following opinion regarding these identifier systems:

To sum up, in the digital environment manifestation identifiers are challenged by both scale (there are a lot more things out there now that need to be identified than anticipated when the traditional identifier systems were developed) and scope (how to define the set of resources to which one can assign identifier X) …. One problem that many national libraries have faced is that only a fraction of resources in their collections have a traditional manifestation identifier. A web archive is a good example of a collection where most items lack identifiers (Hakala 2007).

Both the issues of identifier granularity and the lack of usable identifiers have been serious problems for our own FRBRization work.

Similar to the consideration of aggregates previously by Buchanan, et al. and Gow, et al., Hakala also adds some complexity to the manifestation level of FRBR by adding the level of "component parts." The difficulties in developing identifiers for an unending category of component parts from journal articles to book chapters is challenging Hakala contends due not just to scalability issues but to semantic complexities. As one example, he lists the Book Item and Component Identifier (BICI), an identifier that never received much adoption.[36] A related issue Hakala explains is that there is no agreement among the library community on to how build component part identifiers, just as there is little agreement on how to build manifestation or expression identifiers. Such lack of agreement, Hakala asserts, has led to the need for machine generated identifiers:

I believe that any reasonable component part identifier - at least as long as we talk about machine-readable and to some extent machine-understandable resources - should be self sustainable in the sense that it can be derived directly from the object itself. If automatic generation of these identifiers is not a viable option, then it will be difficult to reach satisfactory coverage in identifier assignment. Automatic identifier extraction, will of course only work if the resources are structured and contain, either in embedded metadata or otherwise, the data elements from which the identifier can be built (Hakala 2006).

This issue is one that we are struggling with as well; there are simply too many components in our texts needing identification (but lacking identifiers) for them to be assigned randomly. At the same time our XML catalog records are highly structured and will hopefully support automatic identifier extraction.

Creating identifiers for any level of the FRBR hierarchy will need to be an effort organized at the national or international level. While some proposals for work (ISTC) and manifestation identifiers (ISBN, ISSN, OCLC, LCCN) have been recommended the debate regarding expression level identifiers often falls back into recommendations for uniform titles. One relatively new identifier, the SICI (Serials Item and Contribution Identifier), seeks to provide identification of "either an issue of a serial title or a contribution (e.g., article) contained within a serial" and is believed by some to operate at the expression level (Vitiello 2004), but by others at the level of "component parts" (Hakala 2006).[37]

One advantage in working in the domain of classics is that many works already have work identifiers in canonical lists such as the Thesaurae Linguae Graecae (TLG) Canon[38], and the Packard Humanities

---

[36] http://www.niso.org/pdfs/BICI-DS.pdf
[37] For more on SICI, see http://www.niso.org/standards/resources/Z39-56.pdf
[38] http://www.tlg.uci.edu/

Institute Latin collections (PHI). Such work identifiers as are available have been included in all of our catalog records.

Some identifiers that have been examined closely at the manifestation level include the ISBN (for books) and the ISSN (for serials). One issue with ISBNs is that they only began being assigned to books in the 1970s, so they cannot work as a manifestation level identifier for older books. Two other major issues with ISBNs are that there is no centralized database and they lack actionability, in other words no ability to get to the resource by clicking on an ISBN (Vitiello 2004). Michael Holdsworth points out that even in the commercial realm of book identification ISBNs have their limitations, due to their "inability to act as a tool for collocating or linking together different manifestations of the same content, a function that is becoming critically important as multiple printed and digital versions of the same title proliferate in systems across the book industry" (Holdsworth 2008).

One FRBR experiment based on ISBNs is supported by OCLC, a web service entitled xISBN that allows libraries to retrieve a list of ISBNs (as well as related metadata) that are associated with a submitted ISBN. The service is designed to support applications such as web-based library catalogs and online booksellers and is based off of information in WorldCat. According to the website, "xISBN enables an end user to link to information about other versions of a source work" and is based on the OCLC FRBR Work-Set Algorithm.[39] Certain levels of free access are allowed but more robust levels of service require a subscription. Similarly, LibraryThing provides a service called thingISBN, a service allowing users to duplicate the same functionality.[40] Interestingly, the data powering thingISBN is powered by information gathered from its members, "who add, combine and separate editions by the thousands every day."

As listed above ISBNs have their limitations as manifestation identifiers, so there have been other proposals that have suggested using the OCLC accession number (a number listed within WorldCat catalog records) or the Library of Congress Control Number (LCCN) as a manifestation level identifier. While other countries have national libraries that could perhaps assign various levels of identifiers such as NBNs, the Library of Congress is not a national library and would be unlikely to serve this role. As the question of manifestation level identifiers is still entirely up for debate, in our catalog records we have encoded as many possible identifiers as we have found for each book, including ISBNs, OCLC numbers and LCCNs.

Whether or not any one identifier can serve the needs of FRBR will likely always be debatable, but this has not lessened the desire for work and expression level identifiers, a theme seen throughout the literature. As Kristin Antelman explains, "documents do not need to be described to be referenced in a networked world; they must be identified. An inherently descriptive element, such as title, cannot meet the requirements of a network identifier" (Antelman 2004). The challenges of creating standard identifiers include assignment responsibility, maintenance of identifier registries, as well as the provision of resolution services. As Antelman points out, "In a networked environment, the identifier associated with an object must not only be unique within the identifier namespace (a primary requirement of URNs), but also must operate within an unambiguous domain with unambiguous rules for identifier assignment" (Antelman 2004). She also fears that many current proposed identifier schemes such s DOIs are too concerned with describing objects from an intellectual property point of view and not a bibliographic one.

Similarly George Buchanan related in his development of a FRBR component for Greenstone, that unique identifiers for documents were a critical component:

---

[39] http://www.worldcat.org/affiliate/webservices/xisbn/apg.jsp
[40] This service by LibraryThing is offered free of charge, for more see http://www.librarything.com/thingology/2008/01/while-you-were-sleeping-thingisbn-got.php

Many digital library systems such as Greenstone and DSpace use unique, abstract identifiers for each document in its collection. These identifiers are commonly used by http requests and DL protocol calls. However, in the case of FRBR, we require identifiers for each type of entity (work, expression, etc.). These existing DL identifiers suffice for distinguishing between different instances of the same expression or work, but if used for the higher-level entities will cause needless confusion. Essentially, a unique DL identifier can merely be used as a reference to a particular item in FRBR (i.e. a single copy of a particular work) (Buchanan 2006).

After a brief evaluation of DOIs and some other identifiers, Buchanan agrees with earlier conclusions of the Perseus Digital Library that a canonical authority would need to assign work level identifiers (Mimno, et al. 2005).

Julian Everett Allgood has also proclaimed that "the MARC 21 authority format represents one possible medium for communicating and exchanging work and expression identifiers. Work and expression identifiers are critical for collocating manifestation-level descriptions, descriptions that multiply to create the MulVer problem" (Allgood 2007). Similarly, in their efforts to use the FRBR model to help them create an application profile for scholarly works, Allinson, et al. documented that "issues relating to normalised names, use of controlled subject vocabularies or other authority lists, dates and identifiers are common and many were identified in the course of our functional requirements gathering" (Allinson, et al. 2007).

As this review has demonstrated, the need for unique identifiers for entities at all level of the FRBR hierarchy is a highly debated topic with little agreement as yet as to which identifiers should be adopted for which level or what new identifiers may need to be created.

FRBR, Functional Identifiers and the Semantic Web

Debate regarding the need for unique identifiers to support the FRBR model inevitably brings up the questions of what exactly is being identified and what purpose unique identifiers might serve. Jonathan Rochkind believes that some of the confusion in the library community is caused by the fact that while identifiers typically serve two functions 1) as a "unique 'key' pointing to a foreign/entity record" typically constructed as a uniform string or number and, 2) as a label or textual identifier to include in an interface so that the user can identify the correct entity, e.g. a controlled personal name heading from an authority file, the library community typically uses only the second type of identifier to accomplish both purposes, when both types of identifiers are essential.[41] Systems need unique identifiers (numerical or strings) to support exchange across systems, but also need diverse labels (that are language and context specific) to be served up in different user interfaces.

This discussion of identifiers has led to some related conversations about FRBR within the Semantic Web community. David Weinberger submits that any implemented solution, will need to "take the semantics out of the identifier so that multiple semantics can be layered on top" (Weinberger 2005). He offers a brief consideration of FRBR and its potential for books but is doubtful that a universal unique identification scheme will ever be developed. In his recent analysis of identifiers, Douglas Campbell agrees that "ensuring uniqueness of identifiers within our own local context is feasible but is more difficult in a global context"(Campbell 2007).

Despite some doubts regarding the ability to create unique identifiers, FRBR is proving a key concept for creating possible intersections between the Semantic Web and library systems and catalogs. Stefan Gradmann recommends that library metadata implementations should be rethought in terms of the Semantic Web and that FRBR can be used as a "kind of pivot concept":

---

[41] Rochkind, J. "Two meanings of 'Identifier.'" Bibliographic Wilderness. April 4, 2007.
 http://bibwild.wordpress.com/2007/04/20/two-meanings-of-identifier/

In that sense, the proposal is not to view FRBR as a kind of ontology to be expressed in RDF, but rather to consider it a kind of specific meta-ontology in the field of librarian information objects, which would have to be expressed using RDF schema (or OWL) as a consequence and which, in turn, would be a suitable basis for catalogue implementation using RDF (Gradmann 2005).

Lars. G. Svensson of the Deutsche National Bibliothek (DNB) has made similar statements. He believes that the FRBR model will be play an important role in helping libraries move their data to the Semantic Web. Svensson believes that libraries first need to address three major tasks: 1) the creation of URIs for all library entities 2) moving all library vocabularies in a machine actionable format to the Web and 3) providing query interfaces against library data that are suitable for the Semantic Web (Svensson 2007). He contends that URIs will be of primary importance, because "URIs are one of the cornerstones of the Semantic Web, allowing for unique identification of each available resource, thus making it possible to refer to, annotate, or retrieve a representation of that resource. While all documents on the Web implicitly have a URI, this is only partly true for library holdings." A way in which a national library could begin, he advises, would be to start assigning URIs at the expression and manifestation levels for items the library owns, and for all modern literature added to their collections, the could create work level URIs. In addition to FRBR, he also recommends that the library community investigate such Semantic Web standards as FOAF[42] and SKOS.[43]

The DNB has already begun work in moving library standards and classification systems to the Web, and is researching how FRBR might be utilized to help them manage the large number of books they are placing online (Huther and Svensson 2007). This work also involves automating as much metadata production as possible and defining unambiguous identifiers for the digital objects that are created. For European and other countries with national libraries, the authors believe that a national library could be responsible for coining URIS at the expression and manifestation level of books. Some of the DNBs other current efforts have involved matching their authority name database the Personennamendatei with the German Wikipedia, which has further illuminated the need for persistent identifiers for individuals. They reported the important finding that: "The example Wikipedia shows, that today the bibliographic records linked to the authority files and provided with unambiguous identifiers can be seen as information entities, that have their own value that is no longer bound to the analog object, the book. In the world wide web, the digital metadata are used as a source of information that points on further information, but not necessarily on an analog object" (Huther and Svensson 2007).

Other librarians have also recommended moving library standards and data to the Web in a format where they can be repurposed, which includes the use of standard identifiers. Lorcan Dempsey has pointed out that, "libraries have made a major historical investment in structured data. We need to find good ways of releasing the value of that investment in productive use in these new services" (Dempsey 2006). The recent report by the LC WGFBC has made a similar finding:

The use of language strings such as personal or corporate names as identifiers for both display and data manipulation hinders data exchange across languages and across different information communities. Emphasis on textual strings as identifiers binds entries to a single language and thus hampers efforts to internationalize both authority files and bibliographic files that carry the authoritative heading forms. Text strings may change over time to reflect changes in display or access forms. The more that data are used by different applications, the more important it is that they be clearly identified using language-neutral identification schema (Amelung, et al. 2008).

---

[42] http://www.foaf-project.org/
[43] http://www.w3.org/2004/02/skos/intro: "SKOS is an area of work developing specifications and standards to support the use of knowledge organisation systems (KOS) such as thesauri, classification schemes, subject heading systems and taxonomies within the framework of the Semantic Web." For more on SKOS, please see (Miles 2006)

This statement led to the general recommendation that the LC should not only generate "standard Web-based identifiers for all data elements and vocabularies that LC maintains" but also that the entire library community should "work to include standard identifiers for individual data elements in bibliographic records, both prospectively and retrospectively, wherever such identifiers are defined, and work to identify changes in metadata carrier standards necessary to incorporate and use such identifiers" (Amelung, et al. 2008). The authors believe that this will likely lead to the need for URIs for personal names, subject headings, etc., with the desired outcome that, "the library community will share identifiers of authors, works, and other controlled elements of bibliographic data to enable interchange of data between different communities of use"(Amelung, et al. 2008).

Before the LC WGFBC formally recommended this idea, Corey Harper and Barbara Tillett had recognized the potential of "webifying" library vocabularies, thesauri and authority data into standards such as RDF, SKOS and OWL (Tillett and Harper 2007a, Tillett and Harper 2007b, Harper 2006). Libraries have spent decades creating valuable controlled vocabularies, but they are of limited utility in their current formats. Although many library collections have been exposed on the Web, the tools used to manage those collections have not typically been made available in a machine readable format. While standards such as MARCXML, MODS and MADS are a start, Harper and Tillett contend that the next step is to translate this metadata into the metadata of the Semantic Web. Recent research by Harper and Tillett has shown that both SKOS and OWL could be used to encode authority information while RDF shows promise for bibliographic information. Both communities could benefit from such effort they argue, because:

Libraries and the developers of the Semantic Web share goals for naming concepts, naming entities, and bringing different forms of those names together. Library tools have been developed over many decades and are very rich sources of connected data. We just need to now translate them into new tools to help the infrastructure of the Semantic Web (Tillett and Harper 2007a).

By using the common framework provided by SKOS, important library vocabularies could be made not only interoperable with other vocabularies but also be used to support Semantic Web and Web 2.0 applications. Rather than suspend RDA, they advise that a RDA Dublin Core application profile should be developed that is based on FRBR and FRAD, and is expressed as RDF or SKOS.

Another interesting idea raised by the authors is linking trusted authority data to FOAF statements, supporting users in their search for authoritative sources. They relate the following intriguing scenario:

As other resources start including metadata that uses identifiers or headings to link to a VIAF, the opportunity to connect more interesting bits of information can add significant value to any Web-based information resource. Wikipedia entries, journal articles, Who's Who biographical info, an individual's blog, their homepage, or the homepage of their place of work can all be interconnected, as well as linked to journal articles, bibliographic records in catalogs and in e-commerce sites, and a variety of other scholarly resources. These interconnections have extensive implications for research. Once there is a corpus of biographical information combined into a data store that is connected to authority data (as well as associated bibliographic data), the information can be used to make inferences about any document, article, Web page, or blog entry that turns up when searching for information. (Tillett and Harper 2007b)

As Tillett and Harper point out, by webifying library vocabularies and authority data, more sophisticated linking and reasoning will be possible for all sorts of Web applications. To test the practical application of these concepts, Corey Harper created a number of examples where he used XSLT to translate LCSH Authority records from MARCXML or MADS XML to SKOS RDF and listed the problems inherent therein (Harper 2006).

These ideas promoted by Svennson, the LC WGFBC and Tillett and Harper are in line with current efforts in the Semantic Web community to create URIs for everything under the sun, from concepts in a controlled vocabulary, to all of the entities described within both Wikipedia and the larger Web.[44] While many in the library community are probing the use of URIs in controlled vocabularies, URIs are not without their detractors. Harry Halpin has declared that, URIs, URLs and URNs all have "the lurking threat of ambiguity" for "there is no principled way to distinguish a URI for a web page versus a URI for a thing 'not on the Web'" (Halpin 2006). Nonetheless, by supporting the naming and linking of related concepts in a semantic way, which will in turn allow the creation of sophisticated and reusable ontologies, both the library and Semantic Web community will benefit. The varying efforts discussed here by both the library and Semantic Web communities in terms of the importance not only of the FRBR model but also of webifying all forms of library data have greatly informed the current work on the Perseus catalog.

FRBR and Million Book Libraries

As thousands of books have already begun to go online the importance of being able to find specific works or expressions and the question of how to manage relationships between their numerous manifestations has become more pressing than ever. This section will look at the increasing relevance of the FRBR model in the era of million book libraries.

A Need for FRBR and Cataloging

Some research has contended that only a small number of books exist in multiple manifestations, which might indicate that even as multiple copies of the same book are placed online, FRBR will be of limited utility. Research conducted by OCLC in regards to the existence of works in WorldCat determined that a majority of works have only one expression and manifestation, and only a small number of works have more than one expression (Bennett, et al. 2003). Bennett, et al. found that only about 20% of works in WorldCat have more than one expression and would likely benefit from FRBRizing efforts.

Maja Žumer, however, proposes looking at these results another way, "To some, this might suggest that only a relatively small percentage of a catalog would profit from FRBR, which could bring the viability of FRBR implementation as a whole into question. However, what is usually overlooked is the fact that the latter works have been published in many versions and editions, showing that there is demand for them, and they are central to the users" (Žumer 2007). A recent study by Richard Smiraglia also surveyed a number of popular works and their manifestations with a focus on a select set of 20th century best sellers (Smiraglia 2007). He established that "all but one work in the sample exists in multiple instantiations; many have large networks; and complex networks of instantiations have begun to appear in full text on the Web" (Smiraglia 2007). Smiraglia's research establishes the potential of using the FRBR concept of the work and how creating linkages between works may help users navigate in a complex information universe. Notwithstanding, the domain of classics can especially benefit from the FRBR model as most works exist not only in multiple translations into various languages but also have multiple manifestations of each translation. The classical domain also has a wealth of related and derivative works such as commentaries, grammars, lexica and numerous other related texts that should be linked to their original works.

Recent research has conveyed that a large number of duplicate copies are likely to be placed online, as it is often cheaper simply to scan a book than to determine if it has already been digitized. OCLC's analysis of the library holdings of the original Google Five libraries has illustrated that these five libraries hold

---

[44] This is a field with an extensive literature that goes beyond the scope of this paper but for some recent interesting efforts, please see (Bouquet, et al. 2007), and (Hepp, et al. 2007). For more on the potential relationship between the Semantic Web and the library community, see (Greenberg 2007).

10.5 million unique print books (out of 32 million in WorldCat), and that at least 44 percent of these books were held by more than two of the libraries. They relate that this "suggests that digitization of the full print book collections of the Google 5 would result in a little more than four out of every ten digitized books being redundant, assuming digitization of works (or titles), rather than manifestations, was the goal of the project" (Lavoie, et al. 2005). Thus it is likely that there will be multiple manifestations of all sorts of books online.

As the pace of digitization is staggering, some might question whether there is any purpose in attempting to create customized access to these massive digital libraries, or what role FRBR and cataloging might play in this new environment. Nonetheless, Karen Calhoun believes that legacy catalog data will be very important in helping to support discovery not just in library systems but within mass digitization projects:

While the scale of current mass digitization projects is impressive, even if digitization occurs at many times the speed calculated here, it may be safe to say that catalog records will have a role to play in discovery and retrieval of the world's library collections for at least a couple of decades and probably longer (Calhoun 2006).

On the other hand, Coyle and Hillmann have stipulated that library catalog data as currently created and simply migrated to the Internet will solve few problems, unless libraries more fully adopt the FRBR model and address the multiple versions problem listed previously:

As mass digitization projects go forward, catalogs are being swamped by these duplicate entries, and since there is little to distinguish the catalog entries for hard copies on the library's shelf and a full text digital copy that the user can access immediately, much confusion among users has ensued. Called the "multiple versions problem," it is one of the more glaring ways that current cataloging rules no longer serve the library's users, and even hinder the ability of systems designers to provide an efficient service for library catalog users (Coyle and Hillmann 2007).

Cataloging data that is not modeled after FRBR may prove to be of little help to users who are navigating through massive digital libraries. Nonetheless, the data contained within catalog records, particularly those records that are modeled after FRBR, might give users a more effective means of entrance into massive collections than keyword searching will allow.

The utility of FRBR in helping to manage large online text collections was detailed recently in a report from the California Digital Library (CDL) on the Melvyl Recommender Project (California Digital Library 2006). The CDL's original Melvyl Recommender Project sought to answer several different research questions, including whether a "full text indexing system (such as XTF) could be successfully adapted to serve a catalog containing millions of metadata records." Since they had learned that the answer to this question was yes, they decided to further probe how the XTF system would handle a mixed collection of millions of metadata records and thousands of full text objects. To find an answer to this second question they added thousands of full text objects from sources such as the Open Content Alliance (OCA) and from other CDL sources to the metadata index they had created for their original research. They soon discovered that the existence of "duplicate and near-duplicate records led to the need for dynamic FRBR grouping." Originally they had hoped to link OCA records to their corresponding records in the Melvyl catalog by using system (Gladys) identifiers, but they soon realized that only a small number of records had these identifiers. As the authors report, FRBR turned out to be the solution to their problem:

In the end, we had to give up on any kind of identifier linking, and our other simplistic linking ideas. Still, we needed to build a system with a unified, fairly easy to understand, merged display of query hits from all the data sources. The solution? FRBR. Our original thought was to investigate FRBR late in the project, but we now realized that we could use the technique of "work sets" to address our linking problems, and so FRBR suddenly took the front seat (California Digital Library 2006).

The authors soon learned that as the volume of data increased, FRBR become an increasingly crucial part of their solution, "it turned out to be the only feasible way we found to achieve a merged query display, as metadata inconsistency simply didn't allow simple identifier linking schemes to work."

In previous work on the Melvyl Recommender Project, they had dynamically FRBRized results using stylesheets that merged most relevant records by a score based on matching different record elements. They had also researched using the OCLC Workset algorithm that statically determines workset groupings at index time, but they found this static grouping made it to difficult to experiment with variations on the algorithm. Instead they attempted "dynamic grouping" where "groupings are only formed for documents which match the user's query." Dynamic grouping had its own related issues including the fact that a different query could provide different results and that a dynamic group might end up missing several relevant documents simply because the documents didn't match the query. They described their work as "a more deeply integrated version of dynamic FRBRization." Rather than embedding FRBR in the stylesheets they coded it into the Java and drew on their raw Lucene index, hoping this would allow entire result sets to be considered in grouping. They learned that the ultimate solution for speedy query processing and grouping lay in "cheap hardware" and "inexpensive RAM":

The basic idea was to create an in- memory random access table of just the essential elements from each record: title, author, date, and identifiers. The table would be loaded at start -up using the Lucene index as a source, and cached for subsequent requests. Using such a table, we were able to easily change and experiment with ways of grouping records using the metadata, only incurring the table load time at start -up. This strategy proved very successful and easy to implement (California Digital Library 2006).

The authors report that the code used to support dynamic FRBR grouping is an "extension of XTF's support for faceted browsing." Instead of drawing a facet directly from the index data, the FRBR code generates the facet data dynamically based on records drawn from the user's query. As the original FRBR Work Set algorithm depends on exact matches, they also sought to implement a less restrictive matching algorithm which decided if documents should be grouped together by utilizing a matching score calculated by comparing documents titles, authors, dates and identifiers. Since many documents had more than one title or author, this work involved matching lists rather than single items. After some manual sampling of records they determined appropriate weights for matching, and achieved good results after a number of experiments. For the time being, the FRBR algorithm only attempts to form groups that represent a single FRBR work, but the CDL plan to further experiment with decomposing works into items and to attempt to integrate LC NAF file for the production system.

The Challenges and Opportunities of Million Book Libraries

In our current catalog we have linked many of our catalog records to online manifestations located in the million book digital libraries created by Google and the OCA in particular. While we have attempted the creation of links to books digitized and made available through Microsoft Live Search,[45] it is very difficult to determine a URL from the complicated search results that will support a direct link. This raises the question of reliability and permanence of the links that we create, especially in terms of Google Books,[46] which has made no formal commitments to permanently host or supply access to the books it has digitized and has also placed a number of restrictions on what may be done with books it has digitized (Gayton 2006). A recent report by Oya Rieger serves as an excellent overview of the digital preservation and sustainability issues raised by these projects (Rieger 2007), and the preparations that libraries need to begin to make collectively to preserve these collections. While Google is currently providing access to the collections of the libraries that they have digitized, these same libraries are also creating their own

---

[45] http://search.live.com/results.aspx?q=&mkt=en-us&scope=books&FORM=LIVSOP
[46] http://books.google.com/

archives to maintain access to these books such as MBooks at the University of Michigan (Grogg and Ashmore 2007).

Of the variety of online digitized books projects, only the OCA has focused on the creation of a "permanent archive," one reason why we chose to have them host a number of the classical texts that we had scanned. They support open access to all of their texts, downloading of all public domain materials, and supply "collection and item-level metadata for all of its hosted collections in a variety of formats" (Johnson 2007). Similarly they encourage the creation of tools such as finding aids or catalogs that will increase access to the materials that they hold. A partner project of the OCA, the Open Library, seeks to create a virtual library catalog of every book that has ever existed with links to online copies of those books; they also plan to let users rewrite catalog records as necessary (Turnbull 2007). The lack of openness of Google Books, particularly computational access to the texts in the public domain, is a valid matter for concern, but one beyond the scope of this paper.[47]

Reactions among the library community to the mass digitization projects have been quite varied from despairing to pragmatic to wildly optimistic (Albanese 2007, Coyle 2006, Courant 2006, Jones 2006,). Similarly, reaction outside of the library community has also ran the gamut of opinions (Kelly 2006, Grafton 2007). Of more interest to us here, however, are the practical realities of attempting to create customized collections and linking catalog records for books to online manifestations that can be found within massive digital libraries. In fact, Kaufman and Ubois have stressed that one of the major issues of current digitization projects is that it does not allow the full potential of linking, including linking within books to be realized:

Technical barriers to achieving the full potential of text digitization are becoming more apparent. Scholars will increasingly wish to link to particular passages within digital books. This ability presupposes that URLs will remain stable and requires a deep link to a specific location within a book. Stable URLs in other types of media, such as sound recordings and moving images, are equally important and should be a consideration in any negotiation (Kaufman and Ubois 2007).

Despite some of these issues Mark Sandler, director of CIC Library Initiatives, hopes that libraries will look at the Google project as an opportunity, and recommends digitizing locally oriented collections and providing more in-depth cataloging for those resources that were previously ignored as libraries focused on mass-produced monographs. Several suggestions that were chiefly relevant to our work were Sandler's arguments that instead of mourning the loss of users to Google, that libraries should "be thinking of how this massive online collection can be optimized to better serve our users" and "curating communities of content that are responsive to the needs of specialized communities of users, and developing tools that would address the particular needs of these specialized communities of scholars" (Sandler 2005). Even more strongly in terms of the Google Books project, Sandler urges libraries to begin more fully contemplating the new realities this project might create:

If Google makes good on its plan to digitize and serve up fifteen million books, it is likely that users will enjoy searching through them for the items they believe they want. Librarians may be right that professionals are more efficient searchers and more discerning judges of the results. Nonetheless, the opinion that matters here is that of the end users, and they seem quite satisfied with their search strategies and the results they retrieve. (Sandler 2006).

Sandler urges libraries to cease defining themselves in terms of the "extent of their holdings" and instead concentrate on the "relevance of their services." As a final thought, he posits that since "different scholars mix and match content in different ways" our "specialists should be helping to build portals and virtual collections that reflect these widely varying understandings of a subject corpus" (Sandler 2006). We agree that the focus of libraries needs to shift in this transitional time to providing specialized services

---

[47] For a thorough investigation of these concerns, please see (Johnson 2007, Kaufman and Ubois 2007)

both for scholars and end users and more focused creation of specialized corpora, goals we are exploring through the Perseus FRBR-inspired catalog.

Nonetheless the online books available through the mass digitization projects, above all those provided by Google Books, have been criticized for a variety of factors.[48] Historian Robert Townsend has raised a number of issues, two of which specifically concern the work we are conducting, the limited or often incorrect metadata provided with each book, and the very narrow view of the public domain that Google has taken:

Beyond the fundamental quality of the scanning, a more significant problem is the incredibly poor descriptive information attached to many of the books on the site (the "metadata"). This is particularly evident in the serial publications, where having the proper name and date of a publication is especially important……These problems seemed fairly pervasive among serial publications on the site, which seem to take the acquisition date from the library catalog without any further review or input from those scanning in the text (Townsend 2007).

Paul Duguid has put foward a similar critique of Google Books, with a focus on the scanning quality of multiple editions of *Tristram Shandy* (Duguid 2007). Some of the major concerns he raises are about the opaque nature of Google Books which gives no indications of how many books it has scanned, and how the concentration on Google Books may diminish sources of funding for other digitization projects. Like Townshend, he notes how the sparse or incorrect metadata can make searching for even known items (such as a specific edition of a work like *Tristram Shandy*) a challenge:

Not a word is mentioned about multiple volumes or volume number. Indeed, a quick survey of the Google Book Project suggests that Google doesn't recognize volume numbers. Not only are the different editions (Harvard's from 1896, Stanford's from 1904) given exactly the same name, but also the different volumes of this Stanford's multivolume edition are labeled identically. Consequently, whatever algorithm Google uses to find the book, it is quite likely, as in this case, to offer volume II first. (Duguid 2007)

We have also experienced similar issues to those listed by Townsend and Duguid when searching for books in Google Books, particularly when we have been searching for multiple volumes in a set or separate volumes of a serial, as volume numbers are typically not included in the search results or even on the "About the Books Page."[49] This can make it quite time consuming to find the correct manifestation of a book, as you need to open the actual PDF to view the title page. Of greater concern, is what appears to be Google's limited interpretation of the public domain. We have often discovered books which are only available as a snippet view, with no discernible reason as to why, since the publication date is long before 1922, and even when considering the European copyright of the life of the author (or in this case typically the editor or translator) plus 70 years, the work seems to be clearly in the public domain. For example, a book that we had digitized and placed in the Open Content Alliance, *Cornuti theologiae Graecae compendium* which was published in 1881 and whose editor was born in 1841 and unlikely to have lived until 1938 is available only as a snippet view in Google Books, ironically the "About the Book" page links to the OCA version.[50] For another example, an English translation of several works by Cicero entitled *Cicero's Three Books of Offices or Moral Duties,* published in 1865, but only available as snippet view.[51]

---

[48] For a good overview of many of these issues, see "The Google Exchange" http://radar.oreilly.com/archives/2007/08/the_google_exch.html
[49] For example, see the "About the Book" page for the first volume of the five volume series *Poetae Latini Minores* published in 1879, the volume number can only be determined by looking at the actual PDF, http://books.google.com/books?id=I9QIAAAAQAAJ. An "About the Book" page for a book in the same series (Volume 3) has even less information available, http://books.google.com/books?id=a9QIAAAAQAAJ and the user needs to page through several pages in the PDF to find the title page.
[50] http://books.google.com/books?id=2kj1AAAACAAJ
[51] http://books.google.com/books?id=yZqbGwAACAAJ

On the other hand, some technologists have speculated that simply getting any digitized version online, even ones that are not perfect, is far better than not having a book online at all, an attitude with which we wholeheartedly agree. In a recent presentation to the RLG Programs group of OCLC, Dylan Tweney of *Wired Magazine,* emphasized that: "It doesn't matter if the digitized versions are particularly complete or even very good, because as long as you've got something, you can start augmenting it with catalog data, links to related works, citation analysis, and other information that you already have in electronic form" (Tweney 2007). This is the approach that we have taken in linking to Google Books.

While million book libraries may indeed have their limitations and difficulties, their importance in helping move us closer to compiling all the texts if not all the services of a "universal library" cannot be understated; we will continue linking our catalog records to all of the online manifestations that we can find in the hope that as these projects mature, these issues will became less prevalent.

Multiple Editions and Mass Digitization Projects

It would appear then that all of the mass digitization projects are likely to face the issue of multiple manifestations of different books. Both WorldCat.org (the freely available Internet version of OCLC's WorldCat catalog)[52] and Google Books both provide some attempts at allowing users to search for multiple editions or expressions of a work. Within basic catalog records in WorldCat, a clickable tab labeled "Editions" supplies a list of the different editions of a work that are available. Google Books has a similar feature, a link to see "More Editions" both from search results lists for books and for the "About This Book" page for any given book, though it is a feature that is not always available.

There has been little discussion as yet within the broader literature as to the potential of these specific features offered by WorldCat.org and Google Books and how well they actually implement FRBR. While there is a great deal of exploration in terms of how mass digitization projects may affect libraries in general, there seems to be little research literature that provides a systematic review of finding books within these projects or that details experimenting with some of the different features offered by such projects. Magda El-Sherbini has examined how libraries have responded to Open Worldcat and found that 46.2 percent of libraries felt that "Open WorldCat has nothing to do with cataloging their hidden collections or with the visibility of their library or their collections"(El-Sherbini 2006). Additionally, a recent overview of the ability to find books within these massive digital libraries (including Amazon.com and several others) was presented by Heather Christenson and Steve Toub of the California Digital Library (Christenson and Toub 2007). Some recent blog entries by Lisa Spiro, Director of the Digital Media Center in the Fondren Library at Rice University have also presented some initial examination of how the tools in Google Books might be used to help literary scholars, including briefly experimenting with the "More Editions" feature.[53]

4. "FRBR Catalog 1.0" and "FRBR-Inspired Catalog 2.0"

"FRBR Catalog 1.0"

The Perseus Project has for over two decades maintained a catalog of documents that we have collected, but our work with FRBR first began in the fall of 2005, after David Mimno, the then senior programmer at Perseus decided to see how FRBR might be utilized to structure the Perseus classical collection (Mimno, et al. 2005). Having previously taken a cataloging class, Mimno realized that FRBR provided a standard method by which we could express distinctions such as that between a generic work (the

---

[52] http://worldcat.org
[53] See for example, Spiro, L. "Literary DNA and Google Books." Digital Scholarship in the Humanities. Dec 8, 2007.
 http://digitalscholarship.wordpress.com/2007/12/08/literary-dna-and-google-books/

Agamemnon of Aeschylus) and its expressions (e.g., a specific edition of *Agamemnon*). Since the Perseus collection of digital texts was moderately sized (yet includes many works such as Homer's *Iliad* in multiple expressions), highly structured, previously cataloged and entirely digital, it was believed that it would be an ideal testbed to experiment with FRBR.

The initial catalog work also drew on the fact that Perseus already had many unique authorized identifiers for most of its works, these identifiers are called abstract bibliographic objects (ABOS), or unique work identifiers central to the Perseus XML document management system.

In very rough terms, an ABO is a book. More accurately, it is the unit of cataloguing. A single printed book may contain several different literary works, or a work may appear in several volumes. We create an ABO for each work that we expect will be requested by readers or cited by other works, and ABOs need not correspond to concrete documents presently in the DL (Smith, et al. 2001).

An ABO basically represents a "unit of intellectual content in the digital library" or in other words a work. The ABO identifier serves as the key in several metadata tables, and when given a distinct ABO the Perseus system can determine which individual XML documents or subdocuments "instantiate versions of the desired text." The same ABO can be used to represent different versions of text such as different translations and each version is "implemented as a separate XML document." ABOs are also used to determine when a text is a commentary on another text, as commentaries are regarded as belonging to a particular ABO rather than a specific version of a text.

After considering our different metadata options, it was decided to use a combination of the MODS[54] and MADS[55] standards created by the Library of Congress (LC) as the basis for the Perseus FRBR cataloging experiment. This decision was based largely on two key facts 1) these metadata standards are expressed in XML and 2) as standards maintained by the LC, our use of them would hopefully allow our data to be repurposed and support interoperability in general. MODS records that were available for different books in our collection were downloaded form the LC web service, and some other MODS records were created from bibliographic data listed in OCLC's WorldCat. While the Perseus online collection already had catalog records of a sort, these catalog records were not consistent with modern cataloging practices. The existing metadata database of Perseus was also used in the creation of the hierarchical catalog.

Since all of the books in the Perseus collection were entirely digital, our initial catalog implemented only the top three levels of the FRBR hierarchy. Work level records were MADS records, with both expression and manifestation levels implemented as MODS records. One difficulty with this approach is that all XPath queries against our database had to be formulated in both namespaces. The final step in creating the catalog involved dividing information in individual records into multiple hierarchical records. This was relatively simple for single-manifestation works (largely reference works available within the Perseus collection and represented as MODS records), with the record being passed through three different XSL filters. When works had more than one manifestation (these were largely primary texts in the classics collection), the work level records were generated from the existing metadata database at Perseus and made heavy use of ABOs.

We based our division of metadata fields between different FRBR levels by following guidelines initially developed by Sally McCallum (McCallum 2004). For the work level record, we used metadata fields that were common to all versions of a work such as author name and uniform title (where available), subject, classification, and genre. At the expression level, we included language, editor, translator, abstract, and table of contents fields, though here we discovered one issue with the initial MODS records, as they

---

[54] http://www.loc.gov/standards/mods/
[55] http://www.loc.gov/standards/mads/

tended to only encode "creator" rather than more granular roles. At the manifestation level, we included publication specific information including physical description information.

As we created this first catalog, we had to begin looking at the issue of the part-whole relationship among different works. Within our collection many manifestations of shorter works were found as part of larger volumes (such as volumes with the collected poems or orations of multiple Greek and Latin authors). The solution we implemented then involved "linking a single manifestation-level record for the multi-work volume to multiple expression-level works. This compromise works for our collection, but we cannot be certain that it will scale to larger collections, especially those for which analytical catalog data is not currently available"(Mimno, et al. 2005). As our catalog has grown in scale (it now includes several dozen volumes that contain hundreds of fragmentary works) this has become an increasingly challenging task, as shall be explained below.

The creation of a hierarchical catalog also involved a number of searching and scalability problems. Searching a hierarchical catalog can become extremely complicated as a query may need to draw on information from the work, expression and manifestation levels. The solution implemented for the initial catalog was to maintain two parallel versions of the catalog, both containing the same records. The first set was a collection of individual records, one for each work, expression and manifestation, these served as the "editable copies" or "source code" so to speak. The second set was a series of composite records or the "compiled" version, one for each work, where one XML document contained all of the expressions of the work and the manifestations of that expression. This compiled version essentially provided a "flat" catalog that was optimized for searching in the XML database eXist.[56] XSLT stylesheets were then used to control the display in response to user queries. Additionally, in this experiment we created custom tags <work> <expression> and <manifestation> to create the hierarchical structure of the composite records, but this was used largely as a stopgap measure, and other means of specifying relationships between blocks of XML will have to be researched.

"FRBR-Inspired Catalog 2.0"

The Perseus "FRBR Catalog 1.0" as it has affectionately become known internally was largely an experiment in which to explore what could be possible with FRBR. It was available briefly on our website, but a functioning online version of the initial catalog has not been available for a considerable amount of time. As "FRBR-Inspired Catalog 2.0" is still in its very initial stages and has grown drastically both in scope and in scale, it is premature to determine what technical solutions from the initial catalog will be utilized and what new ones may be developed.

What began as the catalog for the Perseus collection online has now grown to encompass a much wider collection of texts. While our more general purpose is to create a FRBRized catalog of both our collection and selected digital collections from million book libraries for the classical domain, we also hope to serve as a case study and to demonstrate what work can be done using already existing metadata standards and freely available online collections. The new collection of XML catalog records for our current catalog include:

·   Standard identifiers from library systems for author names and work titles (uniform names for authors, and uniform titles for works where possible), that have also been mapped to their corresponding names in classical canons.
·   Unique work identifiers drawn from existing canons such as the TLG, PHI, the STOA registry of Latin literature, and Perseus ABOS.

---

[56] http://exist.sourceforge.net/

· Analytical cataloging data (full lists of contained works and authors in structured XML, expanded manifestation level information such as addition of series information).
· Links to online bibliographic records (Worldcat.org)
· Links to online manifestations of books in the different mass digitization projects (including Google Books, the Open Content Alliance, Open Library), with page level linking to exact locations of works where supported (currently only Google Books).
· Enhanced authority records for classical authors that include longer lists of variant names, and lists of work identifiers so that the authority records can be effectively linked to our growing number of bibliographic records. All of these records are stored as MADS XML files.
· Newly created MADS XML authority records for several hundred classical authors of fragmentary works.

By including this expanded information, we hope to accomplish a number of goals:

· Support greater interoperability of traditional library and classical domain data through the use of standard author identifiers/names and uniform names for works.
· Allow a more sophisticated level of FRBRized searching through various manifestations by the use of standard work identifiers to identify the works contained therein.
· Provide greater intellectual access to large volumes of rare and standard classical works through analytical cataloging, which will support more granular identification and searching of digital resources.
· Explore the ease and ability of linking to multiple mass digitization projects, and investigate the idea of creating customized collections within million book digital libraries.
· Use existing library standards such as MODS and MADS to promote metadata interoperability.
· Repurpose already existing catalog and authority data to the largest extent possible (MODS records from the LC web service, MARCXML authority records from OCLC), and illustrate the importance of making both catalog and bibliographic data openly available for both human and machine processing.
· Make all catalog and authority files available as XML to promote reuse of the data.

Many of efforts are firmly aligned with the recent report by the LC WGFBC that has encouraged the library community to focus more efforts on the cataloging of rare materials, to reuse and enhance exiting catalog and authority data rather than create records from scratch, to increase participation in the creation of bibliographic and authority records, and most importantly to make all of this data available in a format such as XML so that it can be shared and reused by a variety of applications, including those outside of the library community (Amelung, et al. 2008). The LC WGFBC has a vision for the future strongly modeled on FRBR:

The Working Group envisions a bibliographic infrastructure wherein data about entities of interest (e.g., works, places, people, concepts, chronological periods) are encoded in agreed-upon ways and made available through agreed-upon Web protocols for ready and efficient use by other applications and services (Amelung, et al. 2008).

This goal is one that we are striving for in the catalog and authority data that we are creating.

The collections that will be included within the initial catalog can be divided into three categories 1) the current Perseus online collection of primary source and reference works that we have in fully TEI-compliant XML, 2) a collection of about 600 primary source and reference works in Greek and Latin that exist largely as image books with OCR but no XML transcriptions (we are still determining the best means of providing online access to these materials) 3) a collection of about 250 books that were scanned at the OCA and are all fully available online (they also have accompanying OCR). Further research has

also shown that the growing online mass digitization projects have a wealth of public domain classical texts which can be utilized as necessary to fill in gaps for particular authors or works.

A great deal of work on the current catalog has already been completed. The 600 primary and reference works that were acquired to add to the collection in the last year have all been cataloged, with full analytics, links to online authorities for names and with standard work identifiers. XML catalogue records for the current Perseus collection are also in the process of being created for the current catalog. Our online collection has grown substantially since the first FRBR experiment and the MODS schema has also changed significantly since we first experimented with creating a FRBRized catalogue for the Perseus collection two years ago. The initial MODS records that were downloaded for Perseus reference works were never corrected, and proved to contain a number of errors (such as the incorrect language, or listing an item as microfilm).   Similarly the catalog records created algorithmically for the Perseus primary sources contained substantially less information than is available in the current records being created for the Perseus collection. For example, we have expanded encoding of the role of "creator" to include "translator" and "editor"and linked all names to their authorized forms in the LC NAF. Additionally, all of the Perseus catalog records need to have standard work identifiers and links to the authoritative versions of names added so they are consistent with the records more recently created.

The OCA collection that has been digitized is in the process of being cataloged, with about 2/3 of the works having been added to our current catalogue. The full cataloging of this collection will add a significant number of new authors and works to our entire collection, as there are a number of composite works containing multiple authors (particularly "fragmentary authors") such as the three volume set of the *Poetae Lyrici Graeci*, the two volume set of the *Scaenicae Romanorum Poesis Fragment*a, the three volume *Comicorum Atticorum Fragmenta* and the four volume set of the *Stoicorum Veterum Fragmenta*.

The final step with all of the catalog records that we create will be to search for as many online manifestations of these texts as possible, such as within Google Books, the OCA, the Universal Digital Library,[57] and Microsoft Live Search books, and to link these catalog records to their various online manifestations. All of these catalog records will also be linked to their bibliographic records in the now freely available version of OCLC's Worldcat catalog.

As of late January 2008, the current catalog includes 988 authors for 2300 individual works (including over 150 anonymous works).   For many of the major authors and works, the collection has multiple editions and translations.   The cataloging work in its current stage departs from our previous work in a number of areas, such as through the inclusion of much more extensive information in the catalog records as was listed above. The majority of the work for the "FRBR Catalog 1.0" focused on the development of a hierarchical catalog model, with substantial time spent on the development of customized XSLT filters and XQuery expressions and little time spent on the actual catalog records. In addition, none of the "FRBR Catalog 1.0" records were linked to authority files for authors, editors and translators. While FRBR Catalog 1.0 created authority records only for works, our current work has also focused on the creation of authority records for each author, editor and translator as well. Additionally, while many of the authors included in the current collection such as Cicero already have existing authority records, these largely exist as MARC files. We have downloaded MARCXML files for all authors from an OCLC web service, and then used XSLT stylesheets created by LC to convert them to MADS.

A great bulk of the current cataloging work has involved analytical cataloging. While a number of the Perseus texts do include multiple works by multiple authors or multiple works by one author, few included the number of works found in some of the recent texts that have been added to the FRBR 2.0 Catalog. A large number of "Aggregation Works" that contain dozens of fragmentary works by many

---

[57] http://www.ulib.org/

different authors have been cataloged, including the 5 volume *Greek Anthology*, multiple editions of the *Anthologia Lyrica*, the multiple volume *Poetae Latini Minores*, the two volume *Historicorum Romanorum*, among many others. Additionally, there are a large number of volumes with multiple works by a single author such as multiple works by Aeschylus, Appian, Aristotle, Bede, Cicero, Demosthenes, Euripides, Lucian, Ovid, Plato, Plutarch, Plautus, Seneca the Younger, Sophocles, Statius, Suetonius, Tacitus, Terence, Varro, Vergil, and Xenophon.

Since the analytical cataloging process is not automated and is very time consuming, some difficult decisions have to be made as regards to at what level to stop cataloging or creating records. Currently, we have created single XML manifestation level MODS records for large volumes such as the *Greek Anthology* and *Historicorum Romanorum* (that contain all of the component records for the individual works) and then created separate expression level records for each individual work that is linked to its manifestation. Similarly, for those manifestations that contain works by two or more authors, we have also created individual linked expression level records. Without these expression level records we would not know the full range of authors available within the collection.

Nonetheless, the collection currently contains many volumes with multiple works by an individual author (such as Cicero) where only the manifestation level XML MODS records exists, but where component XML records for each individual work are nested within the one XML tree. Each of these component records, however, does have a unique work identifier, which will perhaps enable expression level records to be created algorithmically in the future. A related issue is that currently the component catalog records within some of these manifestation records (for example, manifestation records for the Loeb editions) list both the English and Latin or English and Greek expressions of a work in a single component record. In other words, these volumes do not yet have separate expression level records for each language. We are currently exploring how to adapt previous XSLT filters to automatically create expression level records for such volumes. The current plan is to use work identifiers as way of pulling expressions level records out of all of these volumes. Sample records and a more detailed explanation of the process of cataloging these works and the XML files that are created will be included in the cataloging workflow section that follows.

5. FRBR Cataloging Workflow

Finding Catalog Records for Books

Since we are using MODS records as the base for our catalog records, we first search the LC catalog[58] to determine if they own the book and thus have a downloadable MODS record from their web service.[59] Each catalog record for a book has a LCCN control number, which serves as a fairly unique identifier that we can then use to query the web service for the MODS record.   For example, the catalog record for the two volumes of Frank Justus Miller's Loeb's translations of Seneca's tragedies and published in 1917, has a LCCN of 17013966. Thus a sample query to the web service looks like:

http://z3950.loc.gov:7090/voyager?operation=searchRetrieve&recordSchema=mods&version=1.1&recordPacking=XML&query=bath.lccn=17013966&maximumRecords=1

If the LC does not own a particular manifestation, we typically took a similar edition of the work that we already had as a MODS record, and simply modified the bibliographic information as appropriate.[60] If

---

[58] http://catalog.loc.gov/
[59] http://z3950.loc.gov:7090/voyager?
[60] While currently all MODS records available at the web service are MOD version 1.1., we have been converting all downloaded records to MODS Version 3.2.

the LC had no similar editions or in a few cases did not own any versions of a work at all, we used bibliographic information from WorldCat to create a MODS record. The process of identifying the right edition (and even the correct manifestation of a specific edition) could be problematic, titles of the same classical work can vary greatly (predominantly for the more unique works which often had titles transcribed literally from Greek), and some catalog records were under older names for an author rather than the current authorized heading. Nonetheless, the ability to start with already existing catalog records has greatly enhanced the work we are currently doing.

Fixing Errors in the Catalog Record

As we have digitized works that are within the public domain many of the records for our collections came from the LC "Old Catalog," which means they often have not been modified for many years and reflected a variety of errors, including incorrectly coded languages, misspellings, statements of responsibility that were included within titles, etc. To return to our earlier example, of Frank Justus Miller's Loeb translations of Seneca, the original MODS record (Figure 1 in Appendix),[61] failed to encode the fact that this is a parallel text translation with both Latin and English, and also failed to encode a series statement reflecting that this is a Loeb edition.

Another common error was when the statement of responsibility or the statement describing the editor, translator, etc. was encoded along with the title. This was often due the fact that in many of these classical editions, the author's name, work title and editorial statement were all included in one long Latin or Greek line. For example, an edition of Terence's *Phormio*, edited by Karl Dziatzko and translated by Morris Hicky Morgan, includes the "recensuit Carolvs Dziatko" within the title statement (See Figure 3).

Another frequent issue was the misidentification of a language; the most common error was the encoding of Latin as Italian, or the failure to encode more than one language. For example, both the Teubner and Oxford editions of Greek authors typically included a Latin preface with a Greek text, often leading to these books being cataloged as being only in Latin or Italian, with no encoding of the Greek at all. For example, the original MODS record for the Oxford edition of Theophrastus *Characteres* edited by Hermann Diels, not only includes the statement of responsibility in the title, but also lists the sole language of this text as "ita" or Italian (See Figure 5). For this text, the preface is in Latin and the body of the text is in Greek with Latin notes. Similarly, the original MODS record for an 8 volume German edition of Thucydides *Historiae* also includes the statement of responsibility in the title statement, and encodes the only language as German (See Figure 7). While the title page, preface and notes are all in German, the main text is entirely in Greek.

Enhancing Catalog Records for Single Expression of a Work Manifestations

The simplest catalog records to enhance were those for single volume manifestations that included a single expression of a work by a single author. Basic enhancements included encoding additional language statements, linking names of authors and editors to the web pages for these names in the OCLC LC NAF web service, adding in standard work identifiers (TLG, ABO, PHI, STOA), and adding links to the manifestation record in Worldcat.org and to manifestations in the mass digitization projects, if any exist. Figure 4 illustrates the enhanced record for Terence's *Phormio* (in contrast to Figure 3). In addition, the statements of responsibility have been enhanced with much fuller descriptions, using standard terms drawn from the MARC Relator terms list.[62]

---

[61] All Figures referenced from this point can be found in the Appendix at the end of the paper.
[62] http://www.loc.gov/marc/relators/relaterm.html

The enhanced record for Theophrastus *Characteres* reflects the same basic enhancements (Figure 6), with the additional feature that the only online version of this book available is through a "snippet view" on Google Books. While some books have no online manifestations, other books have manifestations in several of the mass digitization projects. For example, the 1879 Teubner edition of a work once believed to be by Hyginus Gromaticus *Liber de munitionibus Castrorum*, not only has three online manifestations but also has two different work identifiers (Figure 9). We have also made use of the MARC Relator term of "Attributed Name" to indicate the now suspect nature of this authorship. For such single work manifestations, only a single MODS record needs to be created.

Other single work manifestations included many reference works, some that were published as single volumes such as the 1918 Allyn and Bacon edition of Charles Bennett's *New Latin Grammar*, and others that were published in multiple volumes such as John Edwin Sandys three volume *History of Classical Scholarship* or William Smith's monumental three volume *Dictionary of Greek and Roman Biography and Mythology.* For such multivolume works, an individual MODS record is created for each volume. Catalog records for works such as these do not have work identifiers as yet but do include a number of manifestation level identifiers, but still contain links to authorized names and online manifestations. For an example, please see Figure 10.

Enhancing Catalog Records for Multiple Work Manifestations

Far more time consuming is the analytical cataloging work or component cataloging, which involves detailed level cataloging of the individual authors and works that a larger work may contain. The range of works within this group includes five basic categories, each of which shall be considered in turn with sample XML records.

1) Single volume—Single author—Multiple expressions of multiple works
2) Single volume—Multiple authors—Multiple expressions of multiple works
3) Multiple volumes—Single author—Single expression of a single work
4) Multiple volumes—Single author—Multiple expressions of multiple works
5) Multiple volumes—Multiple authors—Multiple expressions of multiple works

Single Volume—Single Author-Multiple Expressions of Multiple Works

A number of individual volumes include the collected or the partially collected works of a classical author. Examples include the *Scripta Minora* of Arrian published in 1854 by Teubner (containing five different individual works), C. D. Yonge's English translation of three works by Cicero published in one volume by Bell in 1875, *The Academic Questions, Treatise De Finibus and Tusculan Disputations of M. R. Cicero,* and the 1909 Teubner edition of Euripides collected *Tragoediae* edited by Augustus Nauck. For volumes such as these we have not yet created individual expression level records for all of the contained works within each larger manifestation level record, so currently one MODS record has been created for each of these volumes. For the XML record of *Scripta Minora*, please see Figure 11. As this MODS record indicates, there are five <relatedItem type="constituent"> records for each individual work, along with a page level link to an online view in Google Books.

Single Volume—Multiple authors-Multiple Expressions of Multiple Works

Our current collection includes a number of thematic volumes that include collected fragments (orations, poems, histories, comedies, tragedies, etc.) of different classical authors. Examples include Teubner's *Orationes et fragmenta* published in 1892 and containing works by Antiphon, Gorgias of Leontini, Alcidamas, and Antisthenes, the 1883 Harper edition of *Sallust, Florus, and Velleius Paterculus* edited by J.S. Watson and containing works by Sallust, Florus and Velleius Paterculus, and the 1872 Lee and

Shepard printing of *Selections from Classic Latin Authors* edited by Francis Gardner and Buck Gay and containing works by Phaedrus, Justin and Nepos. Other significant texts in this category include anthologies of poems, such as *Latin Poets* by Francis B. Godolphin and *Greek Melic Poets* by Herbert Weir Smyth. In the case of the former work, there were often multiple expressions of the same work (e.g. a specific poem by Catullus or Ode by Horace) by different translators.

For volumes such as these both an individual manifestation level MODS record has been created as well as individual expression level MODS records for each constituent work. Within each constituent level record created in a single manifestation level record and then saved as an expression level record, the record was made fully recursive and included work title, author, editor, language, work identifiers, etc. We believe that these work identifiers will be essential for pulling together various expressions of works from different manifestation level records, chiefly those that have not as yet had individual expression level records created. The expression level record can then be linked back to its manifestation record, through the use of the <relatedItem type="host"> MODS element. For an example of the MODS record for *Orationes et Fragmenta* and an expression level record for this expression of Antiphons work *In Novercam* found within this text please see Figure 12 and Figure 13.

Multiple volumes—Single author—Single Expression of a Single Work

The current collection also contains a number of multiple volume sets that represent the single expression of a single work by a single author. Examples include the 8 volume edition of Thucydides *Historiae* edited by J. Classen and published by Weidmannsche between 1878 and 1884, the five volume *Historia Romana* of Cassius Dio published by Teubner between 1863 and 1865, several multiple volume printings of Livy's *Ab Urbe Condita*, and the four volume Loeb edition of Quintilian's *Institutio Oratoria* published between 1920 and 1922. For these types of series, we created a separate MODS record for each volume that contained the work identifier of the relevant work, and each of these MODS records are then linked through the use of common identifiers such as the LCCN or OCLC #. Each MODS record contains the same work\expression information but different manifestation level information such as different publication dates. We have not as yet created series level MODS records for these different volume sets, but are considering if this might be a worthwhile endeavor. For an example of an enhanced MODS record of a volume from the Classen edition of Thucydides *Historiae*, please see Figure 8.

Multiple volumes—Single author—Multiple Expressions of Multiple Works

A number of multiple volume series that include multiple expressions of multiple works by a single author such as the collected tragedies, orations or plays of a given author published in several volumes, are also present in the current collection. Examples include the two volume Loeb set of Seneca's *Tragedies* published in 1917 and translated by Frank Justus Miller, and the three volume set of Euripides *Fabulae* published by Oxford between 1902 and 1909. As with the above multiple volume series, an individual MODS level manifestation record was created for each volume, except in this case since each volume contained a number of different works, individual constituent records were created for each work within the larger manifestation level MODS records. At this point, individual expression level records were not created for each individual work found within a single volume. For an example of an enhanced MODS record for Volume One of the Miller translation of Seneca, please see Figure 2.

Multiple volumes---Multiple authors---Multiple Expressions of Multiple Works

The most time consuming series to catalog are the multiple volume series that contain multiple authors and multiple works, with some different series involving dozens of different authors. A number of texts fall into this category, including the five volumes of the *Greek Anthology*, multiple editions of the multi-volume *Anthologia Graeca*, the four volume *Stoicorum Veterum Fragmenta* published by Teubner in

1903, and the three volume *Comicorum Atticorum Fragment*. Some of these volumes have been cataloged but many of the larger ones have not as yet. Since most of these contain fragmentary works or "expression fragments" as they are known in FRBRoo (e.g. the poems of Sappho and Alcaeus), this work involved the creation of very large manifestation level MODS records and also hundreds of expression level records. As with single volumes that contain multiple expressions of multiple works by multiple authors, the individual expression level records are linked back to the main manifestation level record through the use of the <related item type="host"> MODS element. No specific MODS records are included for this section, as Figure 12 and Figure 13 demonstrate the types of XML records created.

Creating MADS authority records

A complementary part of this work has been the creation of MADS authority records for each author, editor, translator or other significant individual involved in the creation of any of our collected texts. While a large number of the better known classical authors and the majority of editors and translators have authority files available through the LC NAF, our collection contains many fragmentary and smaller authors who can be found in reference works regarding the classical world or in specialized bibliographies such as the PHI and TLG, but have never had official authority records created. We have created about 400 preliminary authority records for these authors, such as epigrammatists, fragmentary poets and fragmentary historians. Recently, a prototype for the VIAF project from OCLC, [63] which supports searching across the LC NAF and the authority files of the DNB and the Bibliothèque nationale de France (BnF). This has led to the discovery of authority records, mostly in the DNB, for some of these smaller authors. As we continue the enhancement and creation of authority records, we will now search this source as well for authority records. Any authority records that are discovered for authors for whom we have already created MADS records will be downloaded and merged with our current records.

The process of creating authority records involved a number of steps:

1) Identify the author of a work.
For many works this is a fairly straightforward process, such as for volumes that contain several authors at most, all of whom were identified in the original catalog record. This involved searching author names in the OCLC version of the LC NAF to find authorized headings to use in the MODS records. [64] Since many of our books had personal names labeled "from old catalog" we also modernized the names in the MODS records to reflect the most current authorized heading.

Nonetheless many books in our collection contain dozens of works by fragmentary authors. To identify these authors, we would utilize sources such as the TLG and the PHI to identify authors. Certain names represent many authors in the classical world (e.g. Dionysius, with various authors such as Dionysius of Halicanarssus, Dionysius Cato, Dionysius Chalcus, Dionysius Minor, Dionysius of Rhodes, Dionysius the Sophist, etc.), where some of the authors would have authority records and others would not. Often when cataloging texts in Greek, the text would simply list the name such as "Dionysius" since the geographic and other qualifiers have largely been added by scholars throughout the ages as means of disambiguating these names. In some cases there would be an authority record for an ambiguous name but the record would not be for the right author. For example, an astronomer named Maximus published several treatises on astronomy, and while there are over 12 authors named Maximus in the LC NAF, none of them represented this author, who is, however, identified in the TLG (#1487).

---

[63] http://orlabs.oclc.org/viaf/
[64] http://alcme.oclc.org/eprintsUK/index.html

2) Downloading and converting MARCXML records.

For those authors with existing LC NAF records, there is a link on the web page for each author record that allows you to download a MARCXML file.[65] This file would be downloaded and we would then use a XSLT stylesheet available from the LC to convert this file into MADS.[66] We have saved both the MARCXML files and the MADS files for each author, editor and translator name. All of this work has been done with the commercial XML editing tool Oxygen.[67]

3) Creating MADS records for authors with no authority files.

The creation of authority records for those authors who have no authority files is an ongoing process, and around 400 preliminary records have been created. Since we still have a large number of multi-volume series of fragmentary authors to catalog, the number of authors to be identified and who will likely need to have authority records created is still unknown. As authority records require both an authorized heading and listing any number of variant names as well as the sources used, we have chosen to use the name as listed in the TLG or PHI, or the author's name as listed in a reference work such as Brill's New Pauly Online (Brill), a commercially available online version of the monumental Pauly Wissowa classical encyclopedia.[68] We have also utilized other reference sources as well in the creation of these authority records including Oxford Reference Online which includes access to a number of classical dictionaries and encyclopedias, and the three volume Smith's *Dictionary of Greek and Roman Biography and Mythology*, available at the Making of America digital library.[69]

4) Enhancing all authority records with additional information.

All of the authority records for authors (including ones that had an original record in the LC NAF) are now being both standardized and enhanced with more information. Those authors who had LC NAF files have unique LCCN identifiers for their names, but the authors for whom we had to create authority records have no such identifiers. At the same time, many of these authors have identifiers in the TLG or PHI, so we have chosen to use those identifiers so each author will be represented by a unique identifier. All of these records include:

a) Additional variant names, such as differing forms of names listed in the TLG, PHI, STOA Registry of Latin Literature, Brill, as well as the abbreviated names found in the Liddle Scott Jones Lexicon.

b) All variant names are encoded with their language if it is known.

c) For each record, we have added in the <mads:fieldOfActivity> so that in the final catalog we can sort authors by their genres.

d) Multiple URLs have been added into records, such as links to an authors WorldCat Identities page, their Wikipedia page, and links to freely available reference works such as page views in the Smith's *Dictionary*.

e) Lists of all the author's work identifiers have been added so that these authority records can also be linked to the relevant bibliographic records in the catalog.

To show some of these features, we have included several MADS authority records in the Appendix. Figure 14 displays the enhanced MADS record for Aeschylus, Figure 15 shows the MADS record for

---

Adaeus that includes information from the authority record in the DNB, and Figure 16 shows an authority record we created for Acholius, who had no authority records anywhere.

Ongoing Cataloging Work

The original collection of primary and reference works that we have digitized as image books have been fully cataloged, but a great deal of cataloging remains to be done. Work is currently continuing on creating XML records for the current Perseus collection and integrating them within our current catalog structure. Similarly we are in the process of creating catalog records for the OCA works that we have had scanned, and creating authority records for those authors that need them. Additionally as we search the mass digitization projects we are discovering a number of other interesting classical texts that we may consider adding to the catalog as time allows. For many of the original MODS records that were created for the original collection of image books, page numbers were not originally encoded, nor were links located for online manifestations of works, mostly due to time constraints. As our work progresses, we plan to encode in all of our MODS records, page numbers for all of the texts that contain multiple works and links to any online manifestations that can be found (at the page level where this can be supported)

Exploring How to Model, Store and Link and Present the Catalog Data

One of the greatest challenges as we undertake this project is determining how the thousands of XML records that are being created will be stored, indexed, and linked to each other, as well as how the final catalog will be presented as part of the Perseus Digital Library online collection. Many of the original texts that we digitized currently are available only as large image books on our internal servers, with no way for us to link them to the catalog records. We are examining several options for where we can place these image books so they can be linked to, not just at the manifestation level, but also to allow linking the component records for individual works found within the manifestation level MODS records to these image books at the page level. Similarly, linking to the different mass digitization projects also can be a time consuming and difficult process. While we believe that our links to the Open Content Alliance and Open Library will remain persistent, we are less certain that our links to Google Books will remain viable in the long term. At the same time, Google Books is consistently the digital books collection where we are most likely to find an online manifestation of our texts. We have not currently linked any of our catalog records to books within Microsoft Live Search due to the inability to create persistent URLs to individual books.

Our collection has dozens of fragmentary works that can be as short as three lines by authors whose identities can only be found in a handful of classical reference works. In order to support the collation of works and individual authors, in particular those fragmentary authors who are found only on several pages buried in much larger texts, we need to be able to utilize unique identifiers as one potential linking mechanism. For the realm of classics this is a viable solution as we have a limited domain of texts and authors. We realize that many of the solutions we implemented here would not necessarily be practical for larger scale FRBR implementation.

When cataloging is completed, the next major steps will be to pick a XML database or other application that can support sophisticated indexing of both the MODS and MADS records. Additionally, deciding what elements should be made searchable or browsable will need to be considered carefully. At a minimum we hope to let users:

1) Browse a list of all authors.
2) Select a specific author.
3) Browse a list of all works.
4) Browse all works by a specific author.

5) Find all expressions of a particular work by a particular author.
6) Find a specific expression of a work (English translation of Herodotus by a particular translator).
7) Find all expressions of a specific work in a given language (find all the English translations of Herodotus in the catalog).

Determining both the backend and designing the interface to the FRBR catalog will also likely involve the examination of different digital library softwares and some of the "next generation catalog" open source applications that are becoming increasingly available.

6. Lessons Learned

Since work on the initial idea of a FRBRized catalog began back in the fall of 2005, a great deal has changed in the cataloging world. When we developed a literature review for the first paper back in 2005, FRBR was an issue that was talked about in the library literature but not with the same fervor and investigation that has been seen in the last three years.

As mass digitization projects have placed tens of thousands of books online, AACR2 is being remodeled into RDA, and even the Library of Congress is calling for a major rethinking of "bibliographic control" as it is currently practiced, we have come to think of the work we are doing as in the spirit of what Roy Tennant has described as "descriptive enrichment." Tennant points out that libraries need to expand their notion of the bibliographic (since they harvest metadata from web pages as well as online books), become more flexible about where they get their cataloging data (from libraries, publishers records, from free tagging sites such as Library Things, etc) and accept that they cannot control all of the data that is presented to users. He does not advocate the abandonment of traditional cataloging or controlled vocabularies, but argues:

My vision of our future is one in which for any item we care to describe we begin by taking whatever information about the item already exists. This could be the author's own metadata, or someone's tags, or the publisher's ONIX record. We then enrich this initial record as we can with controlled headings, authority work, and additional descriptive elements. But we don't ignore records that we've been unable to enrich in this way -- we simply treat them differently.[70]

Although we have made use of traditional library vocabularies and cataloging data, we have also extended these metadata with identifiers and names drawn from the classical world, and sought to add as much descriptive information to catalog records as is possible, from various non-library sources. At some point when the catalog becomes available online, we will be greatly interested in making all catalog and authority records not just downloadable but editable, so that as more books are placed online, our users can help us link to new online manifestations, correct any errors that they find, add new variant names, or link to more reference sources. The Perseus Digital Library solicits user inputs in several different venues, and the "FRBR-Inspired Catalog 2.0" will be no exception.

While we believe that this detailed level of analytical cataloging is helpful to users it is a time consuming process built mostly off of human labor, and as millions of books go online libraries will need to find new ways of automatically creating as much of this "analytical" metadata as possible. For many years now, libraries have added tables of contents (TOC) to catalog records as one way of enabling greater access to the contents of books. This solution may prove less useful for older books that do not have a TOC, or have an index that is in Latin and Greek, and thus does not lend themselves easily to automatic analysis or index generation. As we have seen, there is a lot of current debate about what level of cataloging detail

---

[70] Tennant, R. "The Future of Descriptive Enrichment." LJ Blog. Dec 10 2007.
http://www.libraryjournal.com/blog/1090000309/post/1920018592.html

users actually need, but since many of the more intellectually demanding tasks of catalog cannot yet be fully automated, exploration of how to best match human and machine labor are essential. There is a tradeoff between how much data you need in your records before a machine learning algorithm can start to learn from this "training data" and perhaps come to identify works and authors automatically. This will be a particular challenge for historical languages.

During our cataloging work, we soon realized that many older catalog records cannot be repurposed easily without some significant fixing. Additionally, as we sought to link to as much library data as possible, we began to realize that we needed both linkable and machine readable bibliographic and authority records (preferably identified by "URIs" and downloadable as XML). Although for our purposes, locating, downloading and linking to individual records within OCLC's WorldCat and online version of the LC NAF was possible, this would be a serious handicap for any automated processes seeking to make use of library data. Nonetheless, the importance of not reinventing the wheel through using existing library catalog records and library standards such as MODS and MADS, emphasized to us the importance of Roy Tennant's point of getting metadata wherever you can as well as sharing and linking that metadata.

Another issue we found that was our ability to link to online manifestations of our texts was very limited, and supported linking only to either a "splash page" about the book, or to an online PDF. While Google Books does support linking at the page level, it is unknown how long URLs created to link to pages might prove stable. Similarly, we on several occasions had a book that we linked into go from "full view" to "snippet view," although the books for all intents and purposes were in the public domain.[71]

As we move forward with the catalog, it will be interesting to see how the landscape of mass digitization projects and the world of cataloging continue to change. Additionally, as the FRBR model grows in importance, and if a FRBR schema or other metadata standard emerges, we will examine how our data can be reshaped around such a model.

7. References

Aalberg, T., Haugen, F., and Husby, O. (2006). "A Tool for Converting from Marc to FRBR." *Proceedings of 10th European Conference, ECDL 2006, Alicante, Spain*, pg. 453-456.

Albanese, A. R. (2007). "Scan This Book!" *Library Journal*, (August 15 2007). http://www.libraryjournal.com/article/CA6466634.html?industryid=47175.

Allgood, J. E. (2007). "Serials and Multiple Versions, or the Inexorable Trend Toward Work-Level Displays." *Library Resources & Technical Services*, 51(3), pg. 160-178.

Allinson, J., P. Johnston, and A. Powell. (2007) "A Dublin Core Application Profile for Scholarly Works." *Ariadne*, 50, January, http://www.ariadne.ac.uk/issue50/allinson-et-al/.

Amelung, R., et al. (2008). *On the Record: Report of the Library of Congress Working Group on the Future of Bibliographic Control*. Library of Congress, January 2008. http://www.loc.gov/bibliographic-future/news/lcwg-ontherecord-jan08-final.pdf

Amelung, R., et al. (2007). *Report on the Future of Bibliographic control: Draft Report for Comment*. Library of Congress. http://www.loc.gov/bibliographic-future/news/lcwg-report-draft-11-30-07-final.pdf

---

[71] For example, we had created a series of links to an online manifestation of the second edition of Volume 1 of *Historicorum Romanorum Reliquiae* published in 1914 (available at http://books.google.com/books?id=klGAxCkD06QC) only to discover that over night it had gone from full to snippet view. The "Corpus Scriptorum Latinorum" had linked to these texts at the page level, and all of their links thus became broken.

Antelman, K. (2004). "Identifying the Serial Work as a Bibliographic Entity." *Library Resources & Technical Services* 48 (4), pg. 238-255. http://eprints.rclis.org/archive/00002604/

Antelman, K, E. Lynema, and A.K. Pace. (2006). "Toward a Twenty-First Century Library Catalog." *Information Technology and Libraries*, 25 (3), pg. 128-139. http://eprints.rclis.org/archive/00007332/

Ayres, M. L., K. Kilner, K. Fitch, and A. Scarvell. (2003). "Report on the Successful AustLit: Australian Literature Gateway Implementation of the FRBR and INDECS Event Models, and Implications for Other FRBR Implementations." *International Cataloguing and Bibliographic Control*, 32, pg. 8-13.

Babeu, A., D. Bamman, G. Crane, R. Kummer, and G. Weaver. (2007). "Named Entity Identification and Cyberinfrastructure." *Research and Advanced Technology for Digital Libraries,* pg. 259-270. http://www.perseus.tufts.edu/~ababeu/ecdl2007.pdf

Bekaert, J., X. Liu, H V. de Sompel, C. Lagoze, S. Payette, and S. Warner. (2006). "Pathways Core: a Data Model for Cross-Repository Services." *JCDL '06: Proceedings of the 6th ACM/IEEE-CS Jjoint Conference on Digital Libraries*, pg. 368.

Bennett, D. B. and P. Williams. (2006). "Name Authority Challenges for Indexing and Abstracting Databases." *Evidence Based Library and Information Practice*, 1. http://ejournals.library.ualberta.ca/index.php/EBLIP/article/view/7

Bennett, R., B. Lavoie, B., and E. O'Neill. (2003). "The Concept of a Work in WorldCat." *Library Collections, Acquisitions and Technical Services*, 27(1), pg. 49-59.

Bouquet, P., H. Stoermer, and D. Giacomuzzi. (2007) "OKKAM: Enabling a Web of Entities." *Proceedings of the WWW2007 Workshop i3: Identity, Identifiers and Identification, Banff, Canada, May 8 2007*. http://ftp.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-249/submission_150.pdf

Bowen, J. (2005). "FRBR: Coming Soon to Your Library?" *Library Resources & Technical Services,* 49 (3), pg. 175-188. https://urresearch.rochester.edu/retrieve/4096/frbr_final.pdf

Buchanan, G. (2006). "FRBR: Enriching and Integrating Digital Libraries." *JCDL '06: Proceedings of the 6th ACM/IEEE-CS Joint Conference on Digital Libraries*, pg. 260-269.

Buchanan, G, J. Gow, A. Blandford, J. Rimmer, and C. Warwick. (2007). "Representing Aggregate Works in the Digital Library." *JCDL '07: Proceedings of the 2007 Conference on Digital Libraries,* pp. 247-56.

Calhoun, K. (2006). "The Changing Nature of the Catalog and its Integration with Other Discovery Tools." Report prepared for the Library of Congress. http://www.loc.gov/catdir/calhoun-report-final.pdf

California Digital Library. (2006). "The Melvyl Recommender Project Full Text Extension Supplementary Report." http://www.cdlib.org/inside/projects/melvyl_recommender/report_docs/mellon_extension.pdf

Campbell, D. (2007). "Identifying the Identifiers." *2007 Proceedings of the International Conference on Dublin Core and Metadata*, pg. 74-84. http://www.dcmipubs.org/ojs/index.php/pubs/article/view/34/16

Carlyle, A. (2006). "FRBR and the Bibliographic Universe, or, How to Read FRBR as a Model." *Library Resources & Technical Services*, 50 (4), pg. 264-273. Preprint available at: http://projects.ischool.washington.edu/acarlyle/Papers/Carlyle_FRBR_2006.htm

Chamnongsri, N., L. Manmart, V. Wuwongse, and E. Jacob. (2006) "Applying FRBR Model as a Conceptual Model in Development of Metadata for Digitized Thai Palm Leaf Manuscripts." *Proceedings of ICADL 2006*, pg. 254-263.

Christenson, H. and S. Toub. (2007). "Book Discovery in a Mass Digitized Environment: Comparing the Ability to Find Books in Mass Digitized Collections and in Best of Breed Book Discovery Interfaces." *DLF 2007 Fall Forum*, November, http://www.diglib.org/forums/fall2007/presentations/Christenson.pdf

Courant, P. (2006). "Scholarship and Academic Libraries (and their Kin) in the World of Google." *First Monday*, 11 (8), August, http://www.firstmonday.org/issues/issue11_8/courant/index.html

Coyle, K. (2004). "Future Considerations: The Functional Library Systems Record." *Library Hi-Tech*, 22 (2), pg. 166-174.

Coyle, K. (2006). "Mass Digitization of Books." *The Journal of Academic Librarianship*, 32 (6), November, pg. 641-645. http://www.kcoyle.net/jal-32-6.html

Coyle, K. and D. Hillmann. (2007). "Resource Description and Access (RDA): Cataloging Rules for the 20th Century." *D-Lib Magazine* 13 1/2, http://dlib.org/dlib/january07/coyle/01coyle.html

Danskin, A. (2006). "Tomorrow Never Knows"-The End of Cataloging.*" Proceedings of World Library And Information Congress: 72nd IFLA General Conference And Council,* http://www.ifla.org/IV/ifla72/papers/102-Danskin-en.pdf

Dellit, A. and T. Boston. (2007). "Relevance Ranking Of Results From MARC-Based Catalogues: From Guidelines To Implementation Exploiting Structured Metadata." Paper presented at Information Online 2007, http://ll01.nla.gov.au/InfoOnline2007Paper.html

Dellit, A. and K. Fitch. (2007). "Rethinking the Catalogue." Paper presented at Innovative Ideas Forum, National Library of Australia in April 2007. http://www.nla.gov.au/nla/staffpaper/2007/documents/Dellit-Fitch-Rethinkingthecatalogue.pdf

Dempsey, L. (2006). "The Library Catalogue in the New Discovery Environment: Some Thoughts." *Ariadne*, 48, July, http://www.ariadne.ac.uk/issue48/dempsey/

Doerr, M. (2007). "Modelling Intellectual Processes: The FRBR - CRM Harmonization." *DELOS 2007*. http://cidoc.ics.forth.gr/docs/doer_le_boeuf.pdf

Doerr, M. and P. LeBoeuf. (2007). "FRBR: Object-Oriented Definition and Mapping to the FRBR-ER. Version 0.8.3. Draft." *International Working Group on FRBR and CIDOC CRM Harmonisation*. http://cidoc.ics.forth.gr/docs/frbr_oo/frbr_docs/FRBR_oo_V0.8.3.pdf

Duguid, P. (2007). "Inheritance and Loss? A Brief Survey of Google Books." *First Monday,* 12, August, http://www.firstmonday.org/issues/issue12_8/duguid/index.html

Dunn, J.W., D. Byrd, M. Notess, J. Riley, and R. Scherle. (2006). "Variations2: Retrieving and Using Music in an Academic Setting." *Communications of the ACM,* 49 (8), August, pg. 53-58.

Eden, B. (2006). "FRBR Implementations (Functional Requirements for Bibliographic Records)." *Library Technology Reports*, 42 (6), Nov-Dec, pg. 32-42.

El-Sherbini, M. (2007). "Open WorldCat and Its Impact on Academic Libraries." *The Journal of Academic Librarianship* 33 (1), pg. 56-66.

Ercegovac, Z. (2006). "Multiple-Version Resources In Digital Libraries: Towards User-Centered Displays." *Journal of the American Society for Information Science and Technology*, 57 (8), pg. 1023-1032.

Gayton, C. M. (2006). "Alexandria Burned—Securing Knowledge Access in the Age of Google." *Vine: The Journal of Information and Knowledge Management Systems*, 36 (2), pg. 155-172.

Gennaro, C. (2008). "Regia: a Metadata Editor for Audiovisual Documents." *Multimedia Tools and Applications,* 36 (3), February, pg. 185-201.

Gorman, M. (2007). "RDA: The Coming Cataloging Debacle." http:/www.slc.bc.ca/rda1007.pdf

Gow, J., A. Blandford, G. Buchanan, C. Warwick, and J. Rimmer. (2006). "Document Structure in Humanities Collections." http://www.ergohci.ucl.ac.uk/people/j.gow/papers/DocumentStructure.pdf

Gradmann, S. (2005). "rdfs:frbr: Towards an Implementation Model for Library Catalogs Using Semantic Web Technology." *Cataloging and Classification Quarterly*, 39 (3/4), pg. 63-76. http://eprints.rclis.org/archive/00007064/

Grafton, A. (2007). "Future Reading: Digitization and its Discontents." *New Yorker*, Nov 5, http://www.newyorker.com/reporting/2007/11/05/071105fa_fact_grafton

Greenberg, J., K. Spurgin, and A. Crystal. (2006). "Functionalities for Automatic Metadata Generation Applications: a Survey of Metadata Experts' Opinions." *International Journal of Metadata, Semantics and Ontologies* 1 (1), pg. 3-20. http://www.inderscience.com/search/index.php?action=record&rec_id=8766

Greenberg, J. (2007). "Advancing the Semantic Web via Library Functions." *Cataloging & Classification Quarterly* 43 (3-4), pg. 203-225.

Grogg, J. E. and B. Ashmore. (2007). "Google Book Search Libraries and Their Digital Copies." *Searcher* (April 2007). http://www.infotoday.com/searcher/apr07/Grogg_Ashmore.shtml

Hakala, J. (2006). "The Seven Levels Of Identification: An Overview Of The Current State Of Identifying Objects Within Digital Libraries." *Program: Electronic Library and Information Systems*, 40 (4), pg. 361-371.

Halpin, H. (2006). "Identity, Reference and Meaning on the Web." *Proceedings of the Workshop on Identity, Meaning and the Web (IMW06) at WWW2006*, Edinburgh, Scotland. http://www.ibiblio.org/hhalpin/irw2006/hhalpin.html

Harper, C. A. (2006). "Encoding Library of Congress Subject Headings in SKOS: Authority Control for the Semantic Web." *Proceedings of the International Conference on Dublin Core and Metadata Applications, 3 - 6 October 2006,* https://scholarsbank.uoregon.edu/dspace/handle/1794/3268

Hegna, K. and E. Murtomaa. (2002). "Data Mining MARC to Find: FRBR?" March 13, 2002. http://folk.uio.no/knuthe/dok/frbr/datamining.pdf

Hepp, M., K. Siorpaes, and D. Bachlechner. (2007). "Harvesting Wiki Consensus: Using Wikipedia Entries as Vocabulary for Knowledge Management." *IEEE Internet Computing,* 11 (5), pg. 54-65.

Hickey, T B., E. O'Neill, and J. Toves. (2002). "Experiments with the IFLA Functional Requirements for Bibliographic Records (FRBR)." *D-Lib Magazine,* 8 (9), http://www.dlib.org/dlib/september02/hickey/09hickey.html

Hilse, H. and J. Kothe. (2006). *Implementing Persistent Identifiers*. Report commissioned by the Research and Development Department of the Goettingen State and University Library, November 2006. http://www.knaw.nl/ecpa/publ/pdf/2732.pdf

Holdsworth, M. (2008). *The Identification of Digital Book Content*. Report prepared for the Book Industry Study Group, January 2008. http://www.bisg.org/docs/DigitalIdentifiers_07Jan08.pdf

Huther, K.-E., and L. G. Svensson. (2007). "Finding Books Online." *6th International Semantic Web Conference and the 2nd Asian Semantic Web Conference, Workshop on Cultural Heritage on the Semantic Web*, pg. 69-75. http://www.cs.vu.nl/~laroyo/CH-SW/ISWC-wp9-proceedings.pdf

IFLA. (2007). *Functional Requirements for Authority Data: A Conceptual Model*. First Public Draft, April 2007. http://www.ifla.org/VII/d4/FRANAR-ConceptualModel-2ndReview.pdf

IFLA. (1998). *Functional Requirements for Bibliographic Records: Final Report*, Volume 19 of *UBCIM Publications-New Series*. München: K.G.Saur. http://www.ifla.org/VII/s13/frbr/frbr.pdf.

Johnson, R. K. (2007). "In Google's Broad Wake: Taking Responsibility for Shaping the Global Digital Library." *ARL: A Bimonthly Report on Research Library Issues and Actions from ARL, CNI, and SPARC* 250 (2007), pg. 1-17. http://www.arl.org/bm~doc/arlbr250digprinciples.pdf

Jones, E. (2006). "Pushing the Digitization Dialogue Forward: "Scholarship and Libraries in Transition" at the University of Michigan." *Journal of Electronic Publishing*, 9 (2), Summer, http://hdl.handle.net/2027/spo.3336451.0009.204

Kaufman, P. B. and J. Ubois. (2007). "Good Terms - Improving Commercial-Noncommercial Partnerships for Mass Digitization: A Report Prepared by Intelligent Television for RLG Programs, OCLC Programs and Research." *D-Lib Magazine,* 13 (11/12) http://www.dlib.org/dlib/november07/kaufman/11kaufman.html

Kelly, K. (2006). "Scan This Book!" *New York Times Magazine,* (May 2006): 42+. http://www.nytimes.com/2006/05/14/magazine/14publishing.html?ex=1305259200&en=c07443d368771bb8&ei=5090

Lagoze, C. and H. Van de Sompel. (2007). "Compound Information Objects: The OAI-ORE Perspective." Report to the Mellon Foundation, May. 2007. http://www.openarchives.org/ore/documents/CompoundObjects-200705.html

Lam, K. T. (2002). "XML and Global Name Access Control." *OCLC Systems and Services*, 18 (2), pg. 88-96. http://repository.ust.hk/dspace/handle/1783.1/443

Lavoie, B., L. Connaway, and L. Dempsey. (2005). "Anatomy of Aggregate Collections: The Example of Google Print for Libraries." *D-Lib Magazine*, 11 (9), http://www.dlib.org/dlib/september05/lavoie/09lavoie.html

Le Boeuf, P. (2005). "Identify 'textual works': ISTC: Controversy and Potential." Presentation at FRBR in 21st Century Catalogues: An Invitational Workshop, Dublin, Ohio, May 2-4, 2005. http://www.oclc.org/research/events/frbr-workshop/presentations/leboeuf/ISTC.ppt#3

Lindahl, D., J. Bowen, and N. F. Foster. (2007). "University of Rochester eXtensible Catalog Phase 1: Final Report." Report to the Mellon Foundation. River Campus Libraries, University of Rochester. http://docushare.lib.rochester.edu/docushare/dsweb/Get/Document-27534

Mann, T. (2007). "The Peloponnesian War and the Future of Reference, Cataloging, and Scholarship in Research Libraries." http://guild2910.org/Pelopponesian%20War%20June%2013%202007.pdf

Marcum, D. (2006). "The Future of Cataloging." *Library Resources & Technical Services*, 50 (1), January, pg. 5-10. http://www.loc.gov/library/reports/CatalogingSpeech.pdf

Markey, K. (2007). "The Online Library Catalog: Paradise Lost and Paradise Regained?" *D-Lib Magazine*, 13(1/2), http://dlib.org/dlib/january07/markey/01markey.html

Maxwell, R. L. (2008). *FRBR: A Guide for the Perplexed*. American Library Association. http://www.alastore.ala.org/SiteSolution.taf?_sn=catalog2&_pn=product_detail&_op=2340

McCallum, S. (2004). "An Introduction to the Metadata Object Description Schema (MODS)." *Library Hi-Tech*, 22 (1), pg. 82-88.

Medeiros, N. (2007). "The Catalog's Last Stand." *OCLC Systems & Services*, 23 (3), pg. 235-237.
http://eprints.rclis.org/archive/00011466/

Miksa, S. D. (2007). "You Need My Metadata: Demonstrating the Value of Library Cataloging (A Response to the Working Group on the Future of Bibliographic Control)."
http://courses.unt.edu/smiksa/documents/Miksa_response%20to%20WG_30July2007.pdf

Miles, A. (2006). "SKOS: Requirements for Standardization." *Proceedings of Dublin Core 2006*,
http://isegserv.itd.rl.ac.uk/public/skos/press/dc2006/camera-ready-paper.pdf

Mimno, D., A. Jones, and G. Crane. (2005). "Hierarchical Catalog Records: Implementing a FRBR Catalog." *D-Lib Magazine*, 11(10), http://www.dlib.org/dlib/october05/crane/10crane.html

Mönch, C. and T. Aalberg. (2003). "Automatic Conversion from MARC to FRBR." *Research and Advanced Technology for Digital Libraries, 7th European Conference, ECDL 2003, Trondheim, Norway* , pg. 405-411.

Naun, C. C. (2007). "FRBR Principles Applied to a Local Online Journal Finding Aid.*" Library Resources & Technical Services*, 51 (2), pg. 134-145.

Pasin, M., E. Motta, and Z. Zdrahal. (2007). "Capturing Knowledge about Philosophy." *K-CAP '07: Proceedings of the 4th International Conference on Knowledge Capture*, pg. 47-54.

Paskin, N. (2006). "Identifier Interoperability: A Report on Two Recent ISO Activities." *D-Lib Magazine,* 12 (4), http://www.dlib.org//dlib/april06/paskin/04paskin.html

Patton, G. (2004). "FRANAR: A Conceptual Model for Authority Data." *Cataloging & Classification Quarterly*, 38 (3/4), pg. 91-104.

Patton, G. (2005). "FRAR: Extending FRBR Concepts to Authority Data." Paper presented at 71th IFLA General Conference and Council. August 14-18th 2005, Oslo, Norway. http://www.ifla.org/IV/ifla71/papers/014e-Patton.pdf

Patton, M., D. Reynolds, S. G. Choudhury, and T. Dilauro. (2004). "Toward a Metadata Generation Framework: A Case Study at Johns Hopkins University." *D-Lib Magazine*, 10 (11), http://webdoc.sub.gwdg.de/edoc/aw/d-lib/dlib/november04/choudhury/11choudhury.html

Pisanski, J and M. Žumer. (2007). "Functional Requirements For Bibliographic Records: An Investigation Of Two Prototypes." *Program: Electronic Library & Information Systems*, 41 (4), pg. 400-417.

Porter, D., et al. (2006). "Creating CTS Collections." *Proceedings of Digital Humanities 2006*, pg. 269–274.

Radebaugh, J. and C. Keith. (2005). "FRBR Display Tool." *Cataloging and Classification Quarterly*, 39 (3/4), pg. 271-283.

Renear, A. H., and Y. Choi. (2006). "Modeling our Understanding, Understanding our Models: The Case of Inheritance in FRBR." *Proceedings 69th Annual Meeting of the American Society for Information Science and Technology (ASIST)*, 43, http://eprints.rclis.org/archive/00008158/

Riva, P. and A. Cato. (2007). "What's New with the FRBR Entity "Expression"?"
SCATNews: Newsletter of the Standing Committee of the IFLA Cataloging Section, December,
http://www.ifla.org/VII/s13/scatn/SCATNews28.pdf

Resource Description and Access. (2007). "RDA to FRBR Mapping."
http://www.collectionscanada.gc.ca/jsc/docs/5rda-frbrmappingrev.pdf

Rieger, O. Y. (2007). "Preservation in the Age of Large Scale Digitization: A White Paper." Paper published at request of the Council on Library and Information Resources, September 2007. http://www.clir.org/activities/details/lsdi.pdf

Salaba, A., and Y. Zhang. (2007). "From a Conceptual Model to Application and System Development." *ASIS&T Bulletin,* August/September,http:/www.asis.org/Bulletin/Aug-07/salaba_zhang.html

Sandler, M. (2005). "Collection Development in the Day of Google." *Library Resources and Technical Services*, 50 (4), pg. 239-243

Sandler, M. (2005). "Disruptive Beneficence: The Google Print Program and the Future of Libraries." *Internet Reference Services Quarterly* 10 (3/4), pg. 5-22.

Schroeder, K. (2006). "Persistent Identification for the Permanent Referencing of Digital Resources - The Activities of the EPICUR Project Enhanced Uniform Resource Name URN Management at Die Deutsche Bibliothek." *Serials Librarian*, 49 (3), pg. 75-87.

Simpson, B., J. Lundgren, and T. Barr. (2007). "Linking Print and Electronic Books: One Approach." *Library Resources & Technical Services*, 51 (2), pg. 146-152.

Smiraglia, R. P. (2007). "The "Works" Phenomenon and Best Selling Books." *Cataloging & Classification Quarterly* 44 (3/4), pg. 179-195.

Smith, D. A., A. Mahoney, and J. Rydberg-Cox. (2001). "Management of XML Documents in an Integrated Digital Library." *Extreme Markup Languages 2000,* http://xml.coverpages.org/perseus-hopperExtreme2000.pdf

Svensson, L. G. (2007). "National Libraries and the Semantic Web: Requirements and Applications." *ICSD Conference Proceedings*. http://drtc.isibang.ac.in:8080/handle/1849/326

Taylor, A. G. (2007). *Understanding FRBR: What It Is and How It Will Affect Our Retrieval Tools*. Libraries Unlimited.

Tillett, B. B. (2003). "Authority Control: State of the Art and New Perspectives." Paper presented at "Authority Control: Definition and International Experiences." http://www.sba.unifi.it/ac/relazioni/tillett_eng.pdf

Tillett, B. B. (2007). "Numbers to Identify Entities (ISADNs-International Standard Authority Data Numbers)." *Cataloging & Classification Quarterly*, 44 (3-4), pg. 343-361.

Tillett, B. B. (2004). "What is FRBR?" http://www.loc.gov/cds/FRBR.html

Tillett, B. B. and C. Harper. (2007a). "Library of Congress Controlled Vocabularies, the Virtual International Authority File, and their Application to the Semantic Web." Paper presented at *IFLA World Library and Information Congress: 73rd IFLA General Conference*. http://www.ifla.org/IV/ifla73/papers/147-Tillet_Harper-en.pdf

Tillett, B. B and C. Harper. (2007b) "Library of Congress Controlled Vocabularies and Their Application to the Semantic Web." *Cataloging & Classification Quarterly*, 43 (3-4), pg. 47-68. Preprint available at https://scholarsbank.uoregon.edu/dspace/bitstream/1794/3269/1/ccq_sem_web.pdf

Townsend, R.B. (2007) "Google Books: Is It Good for History?" *Perspectives,* 45 (6), http://www.historians.org/Perspectives/issues/2007/0709/0709vie1.cfm

Tudhope, D., T. Koch, and R. Heery. (2006). *Terminology Services and Technology: JISC State of the Art Review*. Report of the Joint Information Systems Committee, September 2006. http://www.ukoln.ac.uk/terminology/TSreview-jisc-final-Sept.html

Turnbull, G. (2007). "A Library Bigger Than Any Building." BBC News. 31 July 2007.
http://news.bbc.co.uk/2/hi/uk_news/magazine/6924022.stm

Tweney, D. (2007). "The Information Universe and What It Wants." Presentation made by Dylan Tweney to the RLG Programs Group of OCLC. http://dylan.tweney.com/presentations/rlg2007/rlg_talk_3_compatible.ppt

Vassallo, S. (2006). "Navigating Through Archives, Libraries and Museums: Topic Maps as a Harmonizing Instrument." *LNCS 3873*, pp. 231-240.

Vitiello, G. (2004). "Identifiers and Identification Systems an Informational Look at Policies and Roles from a Library Perspective." *D-Lib Magazine*, 10 (1), http://www.dlib.org/dlib/january04/vitiello/01vitiello.html.

Weinberger, D. (2005). "The Year of Unique IDS." *Journal of the Hyperlinked Organization*, December, http://www.hyperorg.com/backissues/joho-dec05-05.html.

Weinstein, P. C. (1998). "Ontology-Based Metadata: Transforming the MARC Legacy." *DL '98: Proceedings of the Third ACM Conference on Digital Libraries*, pg. 254-263.

Weiss, P. and S. Shadle. (2007). "FRBR in the Real World." *Serials Librarian*, 52(1/2), pg. 91-104.

Weng, C. and Mi, J. (2006.) "Towards Accessibility to Digital Cultural Materials: An FRBRized Approach." *OCLC Systems & Services,* 22 (3), pg. 217-232.

Yee, M. M. (2007). *Cataloging Rules*. http://myee.bol.ucla.edu/catrul.pdf

Yee, M. M. (2005) "FRBRization: a Method for Turning Online Public Finding Lists into Online Public Catalogs." *Information Technology and Libraries*, 24 (3), pg. 77-95. Postprint available free at: http://repositories.cdlib.org/postprints/715/

Zhang, Y., and A. Salaba. (2007a). "Major Issues Facing FRBR Research and Practice Identified in a Delphi Study." School of Library and Information Science, Kent University. Technical report. http://frbr.slis.kent.edu/publications/delphi_issues.pdf

Zhang, Y., and A. Salaba. (2007b). "Critical Issues and Challenges Facing FRBR Research and Practice." *ASIS&T Bulletin,* August/September, http://www.asis.org/Bulletin/Aug-07/zhang_salaba.html.

Žumer, M. (2007). "FRBR: The End of the Road or a New Beginning." *ASIS&T Bulletin*, August/September, http://www.asis.org/Bulletin/Aug-07/Žumer.html.

8. Appendix: Sample XML Records

Figure 1: Original MODS record for Miller's Seneca

```xml
<?xml version="1.0" encoding="UTF-8"?>
<mods xmlns="http://www.loc.gov/mods/v3" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
   xsi:schemaLocation="http://www.loc.gov/mods/v3 http://www.loc.gov/standards/mods/v3/mods-3-2.xsd">
   <titleInfo>
      <title>Seneca's Tragedies</title>
      <subTitle>with an English translation</subTitle>
   </titleInfo>
   <name type="personal">
      <namePart>Seneca, Lucius Annaeus</namePart>
      <namePart type="date">ca. 4 B.C.-65 A.D</namePart>
      <role>
         <roleTerm authority="marcrelator" type="text">creator</roleTerm>
      </role>
   </name>
   <name type="personal">
      <namePart>Miller, Frank Justus</namePart>
      <namePart type="date">1858- [from old catalog]</namePart>
      <role>
         <roleTerm type="text">tr.</roleTerm>
      </role>
   </name>
   <typeOfResource>text</typeOfResource>
   <originInfo>
      <place>
         <placeTerm type="code" authority="marccountry">enk</placeTerm>
      </place>
      <place>
         <placeTerm type="text">London</placeTerm>
      </place>
      <place>
         <placeTerm type="text">New York</placeTerm>
      </place>
      <publisher>W. Heinemann</publisher>
      <publisher>G.P. Putnam's sons</publisher>
      <dateIssued>1917</dateIssued>
      <issuance>monographic</issuance>
   </originInfo>
   <language>
      <languageTerm authority="iso639-2b" type="code">eng</languageTerm>
   </language>
   <physicalDescription>
      <form authority="marcform">print</form>
      <extent>2 v. 17 cm.</extent>
   </physicalDescription>
   <subject authority="lcsh">
      <name type="personal">
         <namePart>Seneca, Lucius Annaeus</namePart>
         <namePart type="date">ca. 4 B.C.-65 A.D</namePart>
      </name>
      <genre>Translations into English</genre>
   </subject>
   <subject authority="lcsh">
      <topic>Mythology, Classical</topic>
      <genre>Drama</genre>
   </subject>
   <classification authority="lcc">PA6156 .S4 1917</classification>
   <identifier type="lccn">17013966</identifier>
   <recordInfo>
      <recordContentSource authority="marcorg">DLC</recordContentSource>
      <recordCreationDate encoding="marc">830523</recordCreationDate>
      <recordChangeDate encoding="iso8601">20010514113227.0</recordChangeDate>
      <recordIdentifier>8713061</recordIdentifier>
   </recordInfo>
</mods>
```

Figure 2: Enhanced MODS record for a single volume of Miller's Seneca

```xml
<?xml version="1.0" encoding="UTF-8"?>
<modsCollection xmlns="http://www.loc.gov/mods/v3"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.loc.gov/mods/v3 http://www.loc.gov/standards/mods/v3/mods-3-2.xsd">
  <mods>
    <titleInfo>
      <title>Seneca's Tragedies</title>
      <subTitle>with an English translation</subTitle>
      <partNumber>Volume I</partNumber>
    </titleInfo>
    <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
      ns2:href="http://errol.oclc.org/laf/n82-166595.html">
      <namePart>Seneca, Lucius Annaeus</namePart>
      <namePart type="date">ca. 4 B.C.-65 A.D</namePart>
      <role>
        <roleTerm authority="marcrelator" type="text">creator</roleTerm>
      </role>
    </name>
    <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink" ns2:href="http://errol.oclc.org/laf/n88-232483.html">
      <namePart>Miller, Frank Justus</namePart>
      <role>
        <roleTerm authority="marcrelator" type="text">translator</roleTerm>
      </role>
    </name>
    <typeOfResource>text</typeOfResource>
    <originInfo>
      <place>
        <placeTerm type="code" authority="marccountry">enk</placeTerm>
      </place>
      <place>
        <placeTerm type="text">London</placeTerm>
      </place>
      <place>
        <placeTerm type="text">New York</placeTerm>
      </place>
      <publisher>W. Heinemann</publisher>
      <publisher>G.P. Putnam's sons</publisher>
      <dateIssued>1917</dateIssued>
      <issuance>monographic</issuance>
    </originInfo>
    <language>
      <languageTerm authority="iso639-2b" type="code">eng</languageTerm>
    </language>
    <language>
      <languageTerm authority="iso639-2b" type="code">lat</languageTerm>
    </language>
    <physicalDescription>
      <form authority="marcform">print</form>
      <extent>2 v. 17 cm.</extent>
      <note>Latin and English on opposite pages</note>
      <note>Bibliography: v. 1, p. xiii-xvi</note>
    </physicalDescription>
    <subject authority="lcsh">
      <name authority="" type="personal">
        <namePart>Seneca, Lucius Annaeus</namePart>
        <namePart type="date">ca. 4 B.C.-65 A.D</namePart>
      </name>
      <topic>Translations into English</topic>
    </subject>
    <subject authority="lcsh">
      <topic>Mythology, Classical</topic>
      <topic>Drama</topic>
    </subject>
    <classification authority="lcc">PA6156.S5</classification>
    <classification authority="ddc">871</classification>
```

```xml
<classification authority="ddc">870.8</classification>
<relatedItem type="series">
  <titleInfo>
    <title>The Loeb classical library</title>
  </titleInfo>
</relatedItem>
<identifier type="lccn">17013966</identifier>
<identifier type="oclc">1885902</identifier>
<location><url displayLabel="WorldCat">http://worldcat.org/oclc/1885902</url></location>
<location><url displayLabel="GoogleBooks">http://books.google.com/books?id=S6pM-L1KuD8C
</url></location>
<relatedItem type="constituent">
  <titleInfo>
    <title>Hercules Furens</title>
  </titleInfo>
  <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
    ns2:href="http://errol.oclc.org/laf/n82-166595.html">
    <namePart>Seneca, Lucius Annaeus</namePart>
    <namePart type="date">ca. 4 B.C.-65 A.D</namePart>
    <role>
      <roleTerm authority="marcrelator" type="text">creator</roleTerm>
    </role>
  </name>
  <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink" ns2:href="http://errol.oclc.org/laf/n88-
232483.html">
    <namePart>Miller, Frank Justus</namePart>
    <role>
      <roleTerm authority="marcrelator" type="text">translator</roleTerm>
    </role>
  </name>
  <language>
    <languageTerm authority="iso639-2b" type="code">eng</languageTerm>
  </language>
  <language>
    <languageTerm authority="iso639-2b" type="code">lat</languageTerm>
  </language>
  <identifier type="phi">1017.1</identifier>
  <identifier type="stoa author-text">stoa0255-stoa016</identifier>
  <part><extent unit="pages">
    <start>1</start>
    <end>119</end>
  </extent></part>
  <location><url
    displayLabel="GoogleBooks">http://books.google.com/books?id=S6pM-
L1KuD8C&amp;printsec=titlepage#PPA1,M1</url></location>
</relatedItem>
<relatedItem type="constituent">
  <titleInfo>
    <title>Troades</title>
  </titleInfo>
  <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
    ns2:href="http://errol.oclc.org/laf/n82-166595.html">
    <namePart>Seneca, Lucius Annaeus</namePart>
    <namePart type="date">ca. 4 B.C.-65 A.D</namePart>
    <role>
      <roleTerm authority="marcrelator" type="text">creator</roleTerm>
    </role>
  </name>
  <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink" ns2:href="http://errol.oclc.org/laf/n88-
232483.html">
    <namePart>Miller, Frank Justus</namePart>
    <role>
      <roleTerm authority="marcrelator" type="text">translator</roleTerm>
    </role>
  </name>
  <language>
    <languageTerm authority="iso639-2b" type="code">eng</languageTerm>
  </language>
  <language>
    <languageTerm authority="iso639-2b" type="code">lat</languageTerm>
```

```xml
        </language>
        <identifier type="phi">1017.2</identifier>
        <identifier type="stoa author-text">stoa0255-stoa024</identifier>
        <part><extent unit="pages">
          <start>121</start>
          <end>223</end>
        </extent></part>
        <location><url
          displayLabel="GoogleBooks">http://books.google.com/books?id=S6pM-
L1KuD8C&amp;printsec=titlepage#PPA121,M1</url></location>
      </relatedItem>
      <relatedItem type="constituent">
        <titleInfo>
          <title>Medea</title>
        </titleInfo>
        <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
          ns2:href="http://errol.oclc.org/laf/n82-166595.html">
          <namePart>Seneca, Lucius Annaeus</namePart>
          <namePart type="date">ca. 4 B.C.-65 A.D</namePart>
          <role>
            <roleTerm authority="marcrelator" type="text">creator</roleTerm>
          </role>
        </name>
        <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink" ns2:href="http://errol.oclc.org/laf/n88-
232483.html">
          <namePart>Miller, Frank Justus</namePart>
          <role>
            <roleTerm authority="marcrelator" type="text">translator</roleTerm>
          </role>
        </name>
        <language>
          <languageTerm authority="iso639-2b" type="code">eng</languageTerm>
        </language>
        <language>
          <languageTerm authority="iso639-2b" type="code">lat</languageTerm>
        </language>
        <identifier type="phi">1017.4</identifier>
        <identifier type="stoa author-text">stoa0255-stoa018</identifier>
        <part><extent unit="pages">
          <start>225</start>
          <end>315</end>
        </extent></part>
        <location><url
          displayLabel="GoogleBooks">http://books.google.com/books?id=S6pM-
L1KuD8C&amp;printsec=titlepage#PPA315,M1</url></location>
      </relatedItem>
      <relatedItem type="constituent">
        <titleInfo>
          <title>Hippolytus</title>
        </titleInfo>
        <titleInfo type="alternative">
          <title>Phaedra</title>
        </titleInfo>
        <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
          ns2:href="http://errol.oclc.org/laf/n82-166595.html">
          <namePart>Seneca, Lucius Annaeus</namePart>
          <namePart type="date">ca. 4 B.C.-65 A.D</namePart>
          <role>
            <roleTerm authority="marcrelator" type="text">creator</roleTerm>
          </role>
        </name>
        <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink" ns2:href="http://errol.oclc.org/laf/n88-
232483.html">
          <namePart>Miller, Frank Justus</namePart>
          <role>
            <roleTerm authority="marcrelator" type="text">translator</roleTerm>
          </role>
        </name>
        <language>
          <languageTerm authority="iso639-2b" type="code">eng</languageTerm>
```

```xml
            </language>
            <language>
               <languageTerm authority="iso639-2b" type="code">lat</languageTerm>
            </language>
            <identifier type="phi">1017.5</identifier>
            <identifier type="stoa author-text">stoa0255-stoa020</identifier>
            <part><extent unit="pages">
               <start>317</start>
               <end>423</end>
            </extent></part>
            <location><url
               displayLabel="GoogleBooks">http://books.google.com/books?id=S6pM-
L1KuD8C&amp;printsec=titlepage#PPA317,M1</url></location>
         </relatedItem>
         <relatedItem type="constituent">
            <titleInfo>
               <title>Oedipus</title>
            </titleInfo>
            <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
               ns2:href="http://errol.oclc.org/laf/n82-166595.html">
               <namePart>Seneca, Lucius Annaeus</namePart>
               <namePart type="date">ca. 4 B.C.-65 A.D</namePart>
               <role>
                  <roleTerm authority="marcrelator" type="text">creator</roleTerm>
               </role>
            </name>
            <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink" ns2:href="http://errol.oclc.org/laf/n88-
232483.html">
               <namePart>Miller, Frank Justus</namePart>
               <role>
                  <roleTerm authority="marcrelator" type="text">translator</roleTerm>
               </role>
            </name>
            <language>
               <languageTerm authority="iso639-2b" type="code">eng</languageTerm>
            </language>
            <language>
               <languageTerm authority="iso639-2b" type="code">lat</languageTerm>
            </language>
            <identifier type="phi">1017.6</identifier>
            <identifier type="stoa author-text">stoa0255-stoa019</identifier>
            <part><extent unit="pages">
               <start>425</start>
               <end>523</end>
            </extent></part>
            <location><url
               displayLabel="GoogleBooks">http://books.google.com/books?id=S6pM-
L1KuD8C&amp;printsec=titlepage#PPA425,M1</url></location>
         </relatedItem>
         <recordInfo>
            <recordContentSource authority="marcorg">DLC</recordContentSource>
            <recordCreationDate encoding="marc">830523</recordCreationDate>
            <recordChangeDate encoding="iso8601">20010514113227.0</recordChangeDate>
            <recordIdentifier>8713061</recordIdentifier>
         </recordInfo>
      </mods>
</modsCollection>
```

Figure 3: Original MODS record for Terence's *Phormio* Edited By Carolus Dziatzko

```xml
<?xml version="1.0" encoding="UTF-8"?>
<mods xmlns="http://www.loc.gov/mods/v3" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
   xsi:schemaLocation="http://www.loc.gov/mods/v3 http://www.loc.gov/standards/mods/v3/mods-3-2.xsd">
   <titleInfo>
      <title>Phormio; recensvit Carolvs Dziatzko</title>
   </titleInfo>
   <titleInfo type="alternative">
      <title>Phormio</title>
```

```xml
</titleInfo>
<name type="personal">
  <namePart>Terence.</namePart>
  <role>
    <roleTerm authority="marcrelator" type="text">creator</roleTerm>
  </role>
</name>
<name type="personal">
  <namePart>Dziatzko, Karl franz Otto</namePart>
  <namePart type="date">1842-1903, [from old catalog]</namePart>
  <role>
    <roleTerm type="text">ed.</roleTerm>
  </role>
</name>
<name type="personal">
  <namePart>Morgan, Morris Hicky</namePart>
  <namePart type="date">1859-1910, [from old catalog]</namePart>
  <role>
    <roleTerm type="text">tr.</roleTerm>
  </role>
</name>
<name type="personal">
  <namePart>Greenough, James Bradstreet</namePart>
  <namePart type="date">1833-1901. [from old catalog]</namePart>
</name>
<typeOfResource>text</typeOfResource>
<originInfo>
  <place>
    <placeTerm type="code" authority="marccountry">xx</placeTerm>
  </place>
  <place>
    <placeTerm type="text">Cantabrigiae</placeTerm>
  </place>
  <publisher>formis descripservnt I Wilson filivsqve</publisher>
  <dateIssued>1894</dateIssued>
  <issuance>monographic</issuance>
</originInfo>
<language>
  <languageTerm authority="iso639-2b" type="code">lat</languageTerm>
</language>
<physicalDescription>
  <form authority="marcform">print</form>
  <extent>xv, 101, [1] p. xxvi pl. 23 cm.</extent>
</physicalDescription>
<subject authority="lcsh">
  <topic>Master and servant</topic>
  <genre>Drama</genre>
</subject>
<subject authority="lcsh">
  <topic>Slaves</topic>
  <genre>Drama</genre>
</subject>
<subject authority="lcsh">
  <name type="personal">
    <namePart>Terence</namePart>
  </name>
  <topic>Illustrations</topic>
</subject>
<classification authority="lcc">PA6755 .P5 1894</classification>
<relatedItem>
  <titleInfo>
    <title>Mss. (Cod vat. lat. 3868) [from old catalog]</title>
  </titleInfo>
  <name type="personal">
    <namePart>Terence.</namePart>
  </name>
</relatedItem>
<identifier type="lccn">04031915</identifier>
<recordInfo>
  <recordContentSource authority="marcorg">DLC</recordContentSource>
```

```xml
        <recordCreationDate encoding="marc">840621</recordCreationDate>
        <recordChangeDate encoding="iso8601">20020424114132.0</recordChangeDate>
        <recordIdentifier>6328666</recordIdentifier>
    </recordInfo>
</mods>
```

Figure 4: Enhanced MODS record for Terence's *Phormio* Edited By Carolus Dziatzko

```xml
<?xml version="1.0" encoding="UTF-8"?>
<mods xmlns="http://www.loc.gov/mods/v3"
   xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
   xsi:schemaLocation="http://www.loc.gov/mods/v3 http://www.loc.gov/standards/mods/v3/mods-3-2.xsd">
    <titleInfo>
        <title>Phormio</title>
    </titleInfo>
    <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
        ns2:href="http://errol.oclc.org/laf/n79-66715.html">
        <namePart>Terence.</namePart>
        <role>
            <roleTerm authority="marcrelator" type="text">creator</roleTerm>
        </role>
    </name>
    <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
        ns2:href="http://errol.oclc.org/laf/nr94-44082.html">
        <namePart>Dziatzko, Karl franz Otto</namePart>
        <namePart type="date">1842-1903,</namePart>
        <role>
            <roleTerm authority="marcrelator" type="text">editor</roleTerm>
        </role>
    </name>
    <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
        ns2:href="http://errol.oclc.org/laf/no98-67006.html">
        <namePart>Morgan, Morris Hicky</namePart>
        <namePart type="date">1859-1910</namePart>
        <role>
            <roleTerm authority="marcrelator" type="text">translator</roleTerm>
        </role>
    </name>
    <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
        ns2:href="http://errol.oclc.org/laf/n80-3987.html
        ">
        <namePart>Greenough, James Bradstreet</namePart>
        <namePart type="date">1833-1901</namePart>
        <role>
            <roleTerm authority="marcrelator" type="text">author of introduction</roleTerm>
        </role>
    </name>
    <typeOfResource>text</typeOfResource>
    <originInfo>
        <place>
            <placeTerm type="code" authority="marccountry">xx</placeTerm>
        </place>
        <place>
            <placeTerm type="text">Cantabrigiae</placeTerm>
        </place>
        <publisher>formis descripservnt I Wilson filivsqve</publisher>
        <publisher>Harvard University</publisher>
        <dateModified>1900</dateModified>
        <dateCreated>1894</dateCreated>
        <issuance>monographic</issuance>
    </originInfo>
    <language objectPart="text">
        <languageTerm authority="iso639-2b" type="code">lat</languageTerm>
    </language>
    <language objectPart="text">
        <languageTerm authority="iso639-2b" type="code">eng</languageTerm>
    </language>
    <physicalDescription>
        <form authority="marcform">print</form>
```

```xml
        <extent>xv, 101, [1] p. xxvi pl. 23 cm.</extent>
    </physicalDescription>
    <note>Parallel translation of Dziatzko by Morgan, has two title pages</note>
    <note>Added title page in English: The Phormio of Terence; translated into English prose by
        M. H. Morgan, with a new prologue by J. B. Greenough and with the Vatican miniatures
        accurately reproduced for the first time. </note>
    <note>Version used in the production of Phormio at Harvard university, April 19, 1894.</note>
    <note type="statement of responsibility">recensvit Carolvs Dziatzko. Prologvm scripsit novom
        Iac. Br. Gronovivs. Accedvnt imagines codicis faticani nvnc primvm accvrate editae</note>
    <subject authority="lcsh">
        <topic>Master and servant</topic>
        <topic>Drama</topic>
    </subject>
    <subject authority="lcsh">
        <topic>Slaves</topic>
        <topic>Drama</topic>
    </subject>
    <subject authority="lcsh">
        <name authority="" type="personal">
            <namePart>Terence</namePart>
        </name>
        <topic>Illustrations</topic>
    </subject>
    <classification authority="lcc">PA6755</classification>
    <relatedItem>
        <titleInfo>
            <title>Mss. (Cod vat. lat. 3868) [from old catalog]</title>
        </titleInfo>
        <name authority="" type="personal">
            <namePart>Terence.</namePart>
        </name>
    </relatedItem>
    <identifier type="lccn">04031915</identifier>
  <identifier type="oclc">7884425</identifier>
    <location>
        <url displayLabel="WorldCat"> http://worldcat.org/oclc/7884425</url>
    </location>
  <location><url displayLabel="GoogleBooks" note="1894 Original
Edition">http://books.google.com/books?id=SIAMAAAAYAAJ</url></location>
        <identifier type="phi">134.4</identifier>
        <identifier type="stoa author-text">stoa0274-stoa006</identifier>
        <recordInfo>
            <recordContentSource authority="marcorg">DLC</recordContentSource>
            <recordCreationDate encoding="marc">840621</recordCreationDate>
            <recordChangeDate encoding="iso8601">20020424114132.0</recordChangeDate>
            <recordIdentifier>6328666</recordIdentifier>
        </recordInfo>
    </mods>
```

Figure 5: Original MODS Record for Theophrastus *Characteres*

```xml
<?xml version="1.0" encoding="UTF-8"?>
<mods xmlns="http://www.loc.gov/mods/v3" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
   xsi:schemaLocation="http://www.loc.gov/mods/v3 http://www.loc.gov/standards/mods/v3/mods-3-2.xsd">
    <titleInfo>
        <title>Characteres; recensvit Hermannvs Diels</title>
    </titleInfo>
    <name type="personal">
        <namePart>Theophrastus.</namePart>
        <role>
            <roleTerm authority="marcrelator" type="text">creator</roleTerm>
        </role>
    </name>
    <name type="personal">
        <namePart>Diels, Hermann</namePart>
        <namePart type="date">1848-1922</namePart>
        <role>
            <roleTerm type="text">ed.</roleTerm>
        </role>
```

```
      </name>
      <typeOfResource>text</typeOfResource>
      <originInfo>
        <place>
          <placeTerm type="code" authority="marccountry">xx</placeTerm>
        </place>
        <place>
          <placeTerm type="text">Oxonii</placeTerm>
        </place>
        <publisher>e typographeo Clarendoniano</publisher>
        <dateIssued>[1909]</dateIssued>
        <dateIssued encoding="marc">1909</dateIssued>
        <issuance>monographic</issuance>
      </originInfo>
      <language>
        <languageTerm authority="iso639-2b" type="code">ita</languageTerm>
      </language>
      <physicalDescription>
        <form authority="marcform">print</form>
        <extent>xxviii, [72] p. 19 cm.</extent>
      </physicalDescription>
      <subject authority="lcsh">
        <topic>Character sketches</topic>
        <genre>Early works to 1800</genre>
      </subject>
      <classification authority="lcc">PA3405.S8 T5 1909</classification>
      <identifier type="lccn">09017226</identifier>
      <recordInfo>
        <recordContentSource authority="marcorg">DLC</recordContentSource>
        <recordCreationDate encoding="marc">840512</recordCreationDate>
        <recordChangeDate encoding="iso8601">20031230154512.0</recordChangeDate>
        <recordIdentifier>8688424</recordIdentifier>
      </recordInfo>
</mods>
```

Figure 6: Enhanced MODS Record for Theophrastus *Characteres*

```
<?xml version="1.0" encoding="UTF-8"?>
<modsCollection xmlns="http://www.loc.gov/mods/v3"
   xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
   xsi:schemaLocation="http://www.loc.gov/mods/v3 http://www.loc.gov/standards/mods/v3/mods-3-2.xsd">
   <mods>
      <titleInfo>
         <title>Characteres</title>
      </titleInfo>
      <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
         ns2:href="http://errol.oclc.org/laf/n79-108364.html
         ">
         <namePart>Theophrastus.</namePart>
         <role>
            <roleTerm authority="marcrelator" type="text">creator</roleTerm>
         </role>
      </name>
      <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
         ns2:href="http://errol.oclc.org/laf/n85-257295.html
         ">
         <namePart>Diels, Hermann</namePart>
         <namePart type="date">1848-1922</namePart>
         <role>
            <roleTerm authority="marcrelator" type="text">editor.</roleTerm>
         </role>
      </name>
      <typeOfResource>text</typeOfResource>
      <originInfo>
         <place>
            <placeTerm type="code" authority="marccountry">xx</placeTerm>
         </place>
         <place>
```

```xml
        <placeTerm type="text">Oxonii</placeTerm>
      </place>
      <publisher>e typographeo Clarendoniano</publisher>
      <dateIssued>[1909]</dateIssued>
      <dateIssued encoding="marc">1909</dateIssued>
      <issuance>monographic</issuance>
    </originInfo>
    <language objectPart="preface">
      <languageTerm authority="iso639-2b" type="code">lat</languageTerm>
    </language>
    <language objectPart="text">
      <languageTerm authority="iso639-2b" type="code">grc</languageTerm>
    </language>
    <physicalDescription>
      <form authority="marcform">print</form>
      <extent>xxviii, [72] p. 19 cm.</extent>
    </physicalDescription>
    <note type="statement of responsibility">recensvit Hermannvs Diels</note>
    <subject authority="lcsh">
      <topic>Character sketches</topic>
      <topic>Early works to 1800</topic>
    </subject>
    <classification authority="lcc">PA3405.S8</classification>
    <classification authority="lcc">PA4448</classification>
    <relatedItem type="series">
      <titleInfo>
        <title>Scriptorum classicorum bibliotheca Oxoniensis</title>
      </titleInfo>
      <titleInfo type="translated">
        <title>Oxford Classical Texts</title>
      </titleInfo>
    </relatedItem>
    <identifier type="oclc"> 498452</identifier>
    <location>
      <url displayLabel="WorldCat">http://worldcat.org/oclc/498452</url>
    </location>
    <identifier type="tlg"> 0093.009</identifier>
    <identifier type="Perseus:abo">Perseus:abo:tlg,0093,009</identifier>
    <location>
      <url displayLabel="GoogleBooks" note="Snippet View"
        >http://books.google.com/books?id=qKoNAAAAIAAJ</url>
    </location>
    <recordInfo>
      <recordContentSource authority="marcorg">OCoLC</recordContentSource>
      <recordCreationDate encoding="marc"/>
      <recordChangeDate encoding="iso8601"/>
      <recordIdentifier/>
    </recordInfo>
  </mods>
</modsCollection>
```

Figure 7: Original MODS record for Thucydides *Historiae* (Classen Edition-8 volumes)

```xml
<?xml version="1.0" encoding="UTF-8"?>
<mods xmlns="http://www.loc.gov/mods/v3" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.loc.gov/mods/v3 http://www.loc.gov/standards/mods/v3/mods-3-2.xsd">
  <titleInfo>
    <title>Thukydides: erklärt von J. Classen</title>
  </titleInfo>
  <name type="personal">
    <namePart>Thucydides.</namePart>
    <role>
      <roleTerm authority="marcrelator" type="text">creator</roleTerm>
    </role>
  </name>
  <name type="personal">
    <namePart>Classen, Johannes</namePart>
    <namePart type="date">1805-1891, [from old catalog]</namePart>
    <role>
```

```xml
          <roleTerm type="text">ed.</roleTerm>
       </role>
    </name>
    <name type="personal">
       <namePart>Steup, Julius</namePart>
       <namePart type="date">1847-1925, [from old catalog]</namePart>
       <role>
          <roleTerm type="text">ed.</roleTerm>
       </role>
    </name>
    <typeOfResource>text</typeOfResource>
    <genre authority="marcgt">biography</genre>
    <originInfo>
       <place>
          <placeTerm type="code" authority="marccountry">gw</placeTerm>
       </place>
       <place>
          <placeTerm type="text">Berlin</placeTerm>
       </place>
       <publisher>Weidmann</publisher>
       <dateIssued>1882-1908</dateIssued>
       <dateIssued encoding="marc" point="start">1882</dateIssued>
       <dateIssued encoding="marc" point="end">1908</dateIssued>
       <issuance>monographic</issuance>
    </originInfo>
    <language>
       <languageTerm authority="iso639-2b" type="code">ger</languageTerm>
    </language>
    <physicalDescription>
       <form authority="marcform">print</form>
       <extent>8 v. front. (port.) maps. 21 cm.</extent>
    </physicalDescription>
    <subject authority="lcsh">
       <geographic>Greece</geographic>
       <topic>History</topic>
       <temporal>Peloponnesian War, 431-404 B.C</temporal>
    </subject>
    <classification authority="lcc">PA4452 .A2 1882</classification>
    <identifier type="lccn">g 01002886</identifier>
    <recordInfo>
       <recordContentSource authority="marcorg">DLC</recordContentSource>
       <recordCreationDate encoding="marc">840608</recordCreationDate>
       <recordChangeDate encoding="iso8601">20020708180715.0</recordChangeDate>
       <recordIdentifier>8659402</recordIdentifier>
    </recordInfo>
</mods>
```

Figure 8: Enhanced MODS record for Volume I from Thucydides *Historiae* (Classen Edition)

```xml
<?xml version="1.0" encoding="UTF-8"?>
<mods xmlns="http://www.loc.gov/mods/v3" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
   xsi:schemaLocation="http://www.loc.gov/mods/v3 http://www.loc.gov/standards/mods/v3/mods-3-2.xsd">
    <titleInfo>
       <title>Thukydides</title>
       <partNumber>Volume I</partNumber>
    </titleInfo>
    <titleInfo type="uniform">
       <title>History of the Peloponnesian War</title>
       <partNumber>Book I. Greek.</partNumber>
    </titleInfo>
    <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink" ns2:href="http://errol.oclc.org/laf/n79-21175.html">
       <namePart>Thucydides.</namePart>
       <role>
          <roleTerm authority="marcrelator" type="text">creator</roleTerm>
       </role>
```

```xml
      </name>
      <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink" ns2:href="http://errol.oclc.org/laf/n84-174325.html">
         <namePart>Classen, Johannes</namePart>
         <namePart type="date">1805-1891</namePart>
         <role>
            <roleTerm authority="marcrelator" type="text">editor</roleTerm>
         </role>
      </name>
      <typeOfResource>text</typeOfResource>
      <genre authority="marc">biography</genre>
      <originInfo>
         <place>
            <placeTerm type="code" authority="marccountry">gw</placeTerm>
         </place>
         <place>
            <placeTerm type="text">Berlin</placeTerm>
         </place>
         <publisher>Weidmann</publisher>
         <edition>3. Aufl. ...</edition>
         <dateIssued>1879</dateIssued>
         <issuance>monographic</issuance>
      </originInfo>
      <language objectPart="introduction">
         <languageTerm authority="iso639-2b" type="code">ger</languageTerm>
      </language>
      <language objectPart="notes">
         <languageTerm authority="iso639-2b" type="code">ger</languageTerm>
      </language>
      <language objectPart="text">
         <languageTerm authority="iso639-2b" type="code">grc</languageTerm>
      </language>
      <physicalDescription>
         <form authority="marcform">print</form>
         <extent>8 v.</extent>
      </physicalDescription>
      <note type="statement of responsbility">erklärt von J. Classen ...</note>
      <subject authority="lcsh">
         <geographic>Greece</geographic>
         <topic>History</topic>
         <temporal>Peloponnesian War, 431-404 B.C</temporal>
      </subject>
      <relatedItem type="series">
         <titleInfo>
            <title>Sammlung griechischer und lateinischer schriftsteller mit deutschen anmerkungen</title>
         </titleInfo>
      </relatedItem>
      <classification authority="lcc">PA4452</classification>
      <identifier type="oclc">24368548</identifier>
      <identifier type="tlg">0003.001</identifier>
      <recordInfo>
         <recordContentSource authority="marcorg">OCoLC</recordContentSource>
         <recordCreationDate encoding="marc"/>
         <recordChangeDate encoding="iso8601"/>
         <recordIdentifier/>
      </recordInfo>
   </mods>
```

Figure 9: Enhanced MODS record for Hyginus Gromaticus

```xml
<?xml version="1.0" encoding="UTF-8"?>
<mods xmlns="http://www.loc.gov/mods/v3" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
   xsi:schemaLocation="http://www.loc.gov/mods/v3 http://www.loc.gov/standards/mods/v3/mods-3-2.xsd">
   <titleInfo>
      <title>Hygini Gromatici Liber de munitionibus castrorum</title>
   </titleInfo>
   <titleInfo type="alternative">
      <title>De munitionibus castrorum</title>
   </titleInfo>
```

```xml
<name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
   ns2:href="http://errol.oclc.org/laf/n87-820217.html">
   <namePart>Hyginus, Gromaticus</namePart>
   <role>
      <roleTerm authority="marcrelator" type="text">creator</roleTerm>
   </role>
   <role>
      <roleTerm type="text">attributed name</roleTerm>
   </role>
</name>
<name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
   ns2:href="http://errol.oclc.org/laf/nr92-44962.html">
   <namePart>Gemoll, Wilhelm</namePart>
   <namePart type="date">b. 1850</namePart>
   <role>
      <roleTerm authority="marcrelator" type="text">editor</roleTerm>
   </role>
</name>
<typeOfResource>text</typeOfResource>
<originInfo>
   <place>
      <placeTerm type="code" authority="marccountry">gw</placeTerm>
   </place>
   <place>
      <placeTerm type="text">Lipsiae</placeTerm>
   </place>
   <publisher>in aedibus B. G. Teubneri</publisher>
   <dateIssued>1879</dateIssued>
   <issuance>monographic</issuance>
</originInfo>
<language>
   <languageTerm authority="iso639-2b" type="code">lat</languageTerm>
</language>
<physicalDescription>
   <form authority="marcform">print</form>
   <extent>1 p.l., 50 p. 18 cm.</extent>
</physicalDescription>
<note type="statement of responsibility">ex recensione Guilmeli Gemoll</note>
<subject authority="lcsh">
   <topic>Military camps</topic>
</subject>
<subject authority="lcsh">
   <geographic>Rome</geographic>
   <topic>Military antiquities.</topic>
</subject>
<classification authority="lcc">PA6104 .H9</classification>
<identifier type="lccn">34013813</identifier>
<identifier type="oclc">61607911 </identifier>
<identifier type="phi">1266.5</identifier>
<identifier type="stoa author-text">stoa0156-stoa005</identifier>
<identifier type="oca"> liberdemunitioni00hygiuoft </identifier>
<location>
   <url displayLabel="WorldCat">http://worldcat.org/oclc/61607911 </url>
</location>
<location>
   <url displayLabel="Open Content Alliance"
      >http://www.archive.org/details/liberdemunitioni00hygiuoft </url>
</location>
<location>
   <url displayLabel="Open Library"
      >http://www.openlibrary.org/details/liberdemunitioni00hygiuoft</url>
</location>
<location>
   <url displayLabel="GoogleBooks">http://books.google.com/books?id=4UUIAAAAQAAJ</url>
</location>
<recordInfo>
   <recordContentSource authority="marcorg">DLC</recordContentSource>
   <recordCreationDate encoding="marc">830808</recordCreationDate>
   <recordChangeDate encoding="iso8601">20041202141535.0</recordChangeDate>
   <recordIdentifier>9697759</recordIdentifier>
```

```xml
      </recordInfo>
</mods>
```

Figure 10: Enhanced MODS record for Bennett's *New Latin Grammar*

```xml
<?xml version="1.0" encoding="UTF-8"?>
<mods xmlns="http://www.loc.gov/mods/v3"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.loc.gov/mods/v3 http://www.loc.gov/standards/mods/v3/mods-3-2.xsd">
    <titleInfo>
       <title>New Latin grammar</title>
    </titleInfo>
    <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
       ns2:href="http://errol.oclc.org/laf/n87-136132.html">
       <namePart>Bennett, Charles E. (Charles Edwin)</namePart>
       <namePart type="date">1858-1921</namePart>
       <role>
          <roleTerm authority="marcrelator" type="text">creator</roleTerm>
       </role>
    </name>
    <typeOfResource>text</typeOfResource>
    <genre authority="marcgt">dictionary</genre>
    <originInfo>
       <place>
          <placeTerm type="code" authority="marccountry">mau</placeTerm>
       </place>
       <place>
          <placeTerm type="text">Boston</placeTerm>
       </place>
       <place>
          <placeTerm type="text">New York [etc.]</placeTerm>
       </place>
       <publisher>Allyn and Bacon</publisher>
       <dateIssued>[c1918]</dateIssued>
       <dateIssued encoding="marc">1918</dateIssued>
       <edition>[3d ed.]</edition>
       <issuance>monographic</issuance>
    </originInfo>
    <language>
       <languageTerm authority="iso639-2b" type="code">eng</languageTerm>
    </language>
    <physicalDescription>
       <form authority="marcform">print</form>
       <extent>xvi, 287 p. 20 cm.</extent>
    </physicalDescription>
    <note type="statement of responsibility">by Charles E. Bennett ...</note>
    <note>Includes index</note>
    <note> "A revision of ... [the author's] Latin grammar originally published in 1895."</note>
    <subject authority="lcsh">
       <topic>Latin language</topic>
       <topic>Grammar</topic>
    </subject>
    <classification authority="lcc">PA2087 .B5 1918</classification>
    <classification authority="ddc">475</classification>
    <identifier type="lccn">18019160</identifier>
    <identifier type="oclc">1246017</identifier>
    <location>
       <url displayLabel="WorldCat">http://worldcat.org/oclc/1246017 </url>
    </location>
    <location>
       <url displayLabel="GoogleBooks">http://books.google.com/books?id=nlLOLIBUkGcC</url>
    </location>
    <recordInfo>
       <recordContentSource authority="marcorg">DLC</recordContentSource>
       <recordCreationDate encoding="marc">830210</recordCreationDate>
       <recordChangeDate encoding="iso8601">20050920164809.0</recordChangeDate>
       <recordIdentifier>9184972</recordIdentifier>
```

```
    </recordInfo>
  </mods>
```

Figure 11: Enhanced MODS record for Arrian's *Scripta Minora*

```
<?xml version="1.0" encoding="UTF-8"?>
<mods xmlns="http://www.loc.gov/mods/v3"
   xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
   xsi:schemaLocation="http://www.loc.gov/mods/v3 http://www.loc.gov/standards/mods/v3/mods-3-2.xsd">
      <titleInfo lang="lat">
         <title>Arriani Nicomediensis Scripta Minora</title>
      </titleInfo>
      <titleInfo type="abbreviated">
         <title>Scripta Minora</title>
      </titleInfo>
      <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
         ns2:href="http://errol.oclc.org/laf/n79-23201.html">
         <namePart>Arrian.</namePart>
         <role>
            <roleTerm authority="marcrelator" type="text">creator</roleTerm>
         </role>
      </name>
      <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
         ns2:href="http://errol.oclc.org/laf/n93-57556.html">
         <namePart>Hercher, Rudolf</namePart>
         <namePart type="date">1821-1878</namePart>
         <role>
            <roleTerm authority="marcrelator" type="text">editor</roleTerm>
         </role>
      </name>
      <typeOfResource>text</typeOfResource>
      <originInfo>
         <place>
            <placeTerm type="code" authority="marccountry">gw</placeTerm>
         </place>
         <place>
            <placeTerm type="text">Lipsiae</placeTerm>
         </place>
         <publisher>Sumptibus et typis B.G. Teubneri</publisher>
         <dateIssued>1854</dateIssued>
         <issuance>monographic</issuance>
      </originInfo>
      <language objectPart="preface">
         <languageTerm authority="iso639-2b" type="code">lat</languageTerm>
      </language>
      <language objectPart="index">
         <languageTerm authority="iso639-2b" type="code">lat</languageTerm>
      </language>
      <language objectPart="text">
         <languageTerm authority="iso639-2b" type="code">grc</languageTerm>
      </language>
      <physicalDescription>
         <form authority="marcform">print</form>
         <extent>LXXVI, 156 p. ; 21 cm.</extent>
      </physicalDescription>
      <tableOfContents>Indica.--Cynegeticus.--Acies contra Alanos.--Periplus Ponti
         Euxini.--Tactica.</tableOfContents>
      <classification authority="lcc">PA3404</classification>
      <relatedItem type="series">
         <titleInfo>
            <title>Bibliotheca scriptorum graecorum et romanorum Teubneriana. [S.g.]</title>
         </titleInfo>
         <titleInfo type="abbreviated">
            <title>Teubner</title>
         </titleInfo>
      </relatedItem>
      <identifier type="oclc">44025324</identifier>
      <location>
```

```xml
        <url displayLabel="WorldCat">http://worldcat.org/oclc/44025324</url>
    </location>
    <location>
        <url displayLabel="GoogleBooks">http://books.google.com/books?id=XE0IAAAAQAAJ</url>
    </location>
    <relatedItem type="constituent">
        <titleInfo>
            <title>Indica</title>
        </titleInfo>
        <titleInfo type="alternative">
            <title>Historia Indica</title>
        </titleInfo>
        <titleInfo type="uniform">
            <title>Indica. Greek.</title>
        </titleInfo>
        <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
            ns2:href="http://errol.oclc.org/laf/n79-23201.html">
            <namePart>Arrian.</namePart>
            <role>
                <roleTerm authority="marcrelator" type="text">creator</roleTerm>
            </role>
        </name>
        <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
            ns2:href="http://errol.oclc.org/laf/n93-57556.html">
            <namePart>Hercher, Rudolf</namePart>
            <namePart type="date">1821-1878</namePart>
            <role>
                <roleTerm authority="marcrelator" type="text">editor</roleTerm>
            </role>
        </name>
        <language objectPart="text">
            <languageTerm authority="iso639-2b" type="code">grc</languageTerm>
        </language>
        <identifier type="tlg">0074.002</identifier>
        <part>
            <extent unit="pages">
                <start>1</start>
                <end>55</end>
            </extent>
        </part>
        <location>
            <url displayLabel="GoogleBooks"
                >http://books.google.com/books?id=XE0IAAAAQAAJ&amp;pg=PA1</url>
        </location>
    </relatedItem>
    <relatedItem type="constituent">
        <titleInfo>
            <title>Cynegeticus</title>
        </titleInfo>
        <titleInfo type="translated" lang="eng">
            <title>Book of the Chase</title>
        </titleInfo>
        <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
            ns2:href="http://errol.oclc.org/laf/n79-23201.html">
            <namePart>Arrian.</namePart>
            <role>
                <roleTerm authority="marcrelator" type="text">creator</roleTerm>
            </role>
        </name>
        <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
            ns2:href="http://errol.oclc.org/laf/n93-57556.html">
            <namePart>Hercher, Rudolf</namePart>
            <namePart type="date">1821-1878</namePart>
            <role>
                <roleTerm authority="marcrelator" type="text">editor</roleTerm>
            </role>
        </name>
        <language objectPart="text">
            <languageTerm authority="iso639-2b" type="code">grc</languageTerm>
        </language>
```

```xml
        <identifier type="tlg">0074.003</identifier>
        <part>
          <extent unit="pages">
            <start>56</start>
            <end>79</end>
          </extent>
        </part>
        <location>
          <url displayLabel="GoogleBooks"
            >http://books.google.com/books?id=XE0IAAAAQAAJ&amp;pg=PA56</url>
        </location>
    </relatedItem>
    <relatedItem type="constituent">
        <titleInfo>
          <title>Acies Contra Alanos</title>
        </titleInfo>
        <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
          ns2:href="http://errol.oclc.org/laf/n79-23201.html">
          <namePart>Arrian.</namePart>
          <role>
            <roleTerm authority="marcrelator" type="text">creator</roleTerm>
          </role>
        </name>
        <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
          ns2:href="http://errol.oclc.org/laf/n93-57556.html">
          <namePart>Hercher, Rudolf</namePart>
          <namePart type="date">1821-1878</namePart>
          <role>
            <roleTerm authority="marcrelator" type="text">editor</roleTerm>
          </role>
        </name>
        <language objectPart="text">
          <languageTerm authority="iso639-2b" type="code">grc</languageTerm>
        </language>
        <identifier type="tlg">0074.006</identifier>
        <part>
          <extent unit="pages">
            <start>80</start>
            <end>85</end>
          </extent>
        </part>
        <location>
          <url displayLabel="GoogleBooks"
            >http://books.google.com/books?id=XE0IAAAAQAAJ&amp;pg=PA80</url>
        </location>
    </relatedItem>
    <relatedItem type="constituent">
        <titleInfo>
          <title>Periplus Ponti Euxini</title>
        </titleInfo>
        <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
          ns2:href="http://errol.oclc.org/laf/n79-23201.html">
          <namePart>Arrian.</namePart>
          <role>
            <roleTerm authority="marcrelator" type="text">creator</roleTerm>
          </role>
        </name>
        <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
          ns2:href="http://errol.oclc.org/laf/n93-57556.html">
          <namePart>Hercher, Rudolf</namePart>
          <namePart type="date">1821-1878</namePart>
          <role>
            <roleTerm authority="marcrelator" type="text">editor</roleTerm>
          </role>
        </name>
        <language objectPart="text">
          <languageTerm authority="iso639-2b" type="code">grc</languageTerm>
        </language>
        <identifier type="tlg">0074.004</identifier>
        <part>
```

```xml
          <extent unit="pages">
             <start>86</start>
             <end>103</end>
          </extent>
       </part>
       <location>
          <url displayLabel="GoogleBooks"
             >http://books.google.com/books?id=XE0IAAAAQAAJ&amp;pg=PA86</url>
       </location>
    </relatedItem>
    <relatedItem type="constituent">
       <titleInfo>
          <title>Tactica</title>
       </titleInfo>
       <titleInfo type="alternative">
          <title>Ars Tactica</title>
       </titleInfo>
       <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
          ns2:href="http://errol.oclc.org/laf/n79-23201.html">
          <namePart>Arrian.</namePart>
          <role>
             <roleTerm authority="marcrelator" type="text">creator</roleTerm>
          </role>
       </name>
       <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
          ns2:href="http://errol.oclc.org/laf/n93-57556.html">
          <namePart>Hercher, Rudolf</namePart>
          <namePart type="date">1821-1878</namePart>
          <role>
             <roleTerm authority="marcrelator" type="text">editor</roleTerm>
          </role>
       </name>
       <language objectPart="text">
          <languageTerm authority="iso639-2b" type="code">grc</languageTerm>
       </language>
       <identifier type="tlg">0074.005</identifier>
       <part>
          <extent unit="pages">
             <start>104</start>
             <end>139</end>
          </extent>
       </part>
       <location>
          <url displayLabel="GoogleBooks"
             >http://books.google.com/books?id=XE0IAAAAQAAJ&amp;pg=PA104</url>
       </location>
    </relatedItem>
    <recordInfo>
       <recordContentSource authority="marcorg">OCoLC</recordContentSource>
       <recordCreationDate encoding="marc"/>
       <recordChangeDate encoding="iso8601"/>
       <recordIdentifier/>
    </recordInfo>
 </mods>
```

Figure 12: MODS record for *Orationes et Fragmenta*

```xml
<?xml version="1.0" encoding="UTF-8"?>
<modsCollection xmlns="http://www.loc.gov/mods/v3" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
   xsi:schemaLocation="http://www.loc.gov/mods/v3
   http://www.loc.gov/standards/mods/v3/mods-3-2.xsd">
   <mods>
      <titleInfo>
         <title>Orationes et fragmenta</title>
         <subTitle>adivnctis Gorgiae, Antisthenis, Alcidamantis, declamationibvs</subTitle>
      </titleInfo>
      <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink" ns2:href="http://errol.oclc.org/laf/n86-46008.html">
```

```xml
         <namePart>Blass, Friedrich</namePart>
         <namePart type="date">1843-1907</namePart>
         <role>
            <roleTerm authority="marcrelator" type="text">editor</roleTerm>
         </role>
      </name>
   <typeOfResource>text</typeOfResource>
   <originInfo>
      <place>
         <placeTerm type="code" authority="marccountry">gw</placeTerm>
      </place>
      <place>
         <placeTerm type="text">Lipsiae</placeTerm>
      </place>
      <publisher>Tevbneri</publisher>
      <dateIssued>1892</dateIssued>
      <edition>Editio altera correctior</edition>
      <issuance>monographic</issuance>
   </originInfo>
   <language objectPart="text">
      <languageTerm authority="iso639-2b" type="code">grc</languageTerm>
   </language>
   <language objectPart="preface">
      <languageTerm authority="iso639-2b" type="code">lat</languageTerm>
   </language>
   <language objectPart="notes">
      <languageTerm authority="iso639-2b" type="code">lat</languageTerm>
   </language>
   <physicalDescription>
      <form authority="marcform">print</form>
      <extent>xlvii, 212 p. ; 18 cm.</extent>
   </physicalDescription>
   <note type="statement of responsibility">edidit Fridericvs Blass.</note>
   <note>Text in Greek; editorial material in Latin.</note>
   <classification authority="lcc">PA3869</classification>
   <identifier type="oclc">62356548</identifier>
   <location><url displayLabel="WorldCat">http://worldcat.org/oclc/62356548</url></location>
   <location><url displayLabel="GoogleBooks">http://books.google.com/books?id=LvfjhLc5pUwC</url></location>
   <location><url displayLabel="Open Content Alliance">http://www.archive.org/details/antiphontisorati00antiuoft</url></location>
   <location><url displayLabel="Open Library">http://www.openlibrary.org/details/antiphontisorati00antiuoft</url></location>
   <relatedItem type="series">
      <titleInfo>
         <title>Bibliotheca scriptorum Graecorum et Romanorum Teubneriana. [S. g.]</title>
      </titleInfo>
      <titleInfo type="abbreviated">
         <title>Teubner</title>
      </titleInfo>
   </relatedItem>
   <relatedItem type="constituent">
      <titleInfo>
         <title>In Novercam</title>
      </titleInfo>
      <titleInfo type="translated" lang="eng">
         <title>Against the Stepmother for Poisoning</title>
      </titleInfo>
      <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink" ns2:href="http://errol.oclc.org/laf/n82-101412.html">
         <namePart>Antiphon</namePart>
         <namePart type="date">ca. 480-411 B.C</namePart>
         <role>
            <roleTerm authority="marcrelator" type="text">creator</roleTerm>
         </role>
      </name>
      <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink" ns2:href="http://errol.oclc.org/laf/n86-46008.html">
         <namePart>Blass, Friedrich</namePart>
         <namePart type="date">1843-1907</namePart>
         <role>
            <roleTerm authority="marcrelator" type="text">editor</roleTerm>
         </role>
      </name>
```

```xml
        <language objectPart="text">
          <languageTerm authority="iso639-2b" type="code">grc</languageTerm>
        </language>
        <identifier type="tlg">0028.001</identifier>
        <part>
          <extent unit="pages">
            <start>1</start>
            <end>11</end>
          </extent>
        </part>
        <location><url
          displayLabel="GoogleBooks">http://books.google.com/books?id=LvfjhLc5pUwC&amp;pg=PA1</url></location>
  </relatedItem>
  <relatedItem type="constituent">
      <titleInfo>
        <title>Tetralogia 1</title>
      </titleInfo>
      <titleInfo type="alternative" lang="eng">
        <title>First Tetralogy</title>
      </titleInfo>
      <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
        ns2:href="http://errol.oclc.org/laf/n82-101412.html          ">
        <namePart>Antiphon</namePart>
        <namePart type="date">ca. 480-411 B.C</namePart>
        <role>
          <roleTerm authority="marcrelator" type="text">creator</roleTerm>
        </role>
      </name>
      <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink" ns2:href="http://errol.oclc.org/laf/n86-46008.html">
        <namePart>Blass, Friedrich</namePart>
        <namePart type="date">1843-1907</namePart>
        <role>
          <roleTerm authority="marcrelator" type="text">editor</roleTerm>
        </role>
      </name>
      <identifier type="tlg">0028.002</identifier>
      <language objectPart="text">
        <languageTerm authority="iso639-2b" type="code">grc</languageTerm>
      </language>
      <part>
        <extent unit="pages">
          <start>12</start>
          <end>29</end>
        </extent>
      </part>
      <location><url
        displayLabel="GoogleBooks">http://books.google.com/books?id=LvfjhLc5pUwC&amp;pg=PA12</url></location>
  </relatedItem>
  <relatedItem type="constituent">
      <titleInfo>
        <title>Tetralogia 2</title>
      </titleInfo>
      <titleInfo type="alternative" lang="eng">
        <title>Second Tetralogy</title>
      </titleInfo>
      <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
        ns2:href="http://errol.oclc.org/laf/n82-101412.html          ">
        <namePart>Antiphon</namePart>
        <namePart type="date">ca. 480-411 B.C</namePart>
        <role>
          <roleTerm authority="marcrelator" type="text">creator</roleTerm>
        </role>
      </name>
      <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink" ns2:href="http://errol.oclc.org/laf/n86-46008.html">
        <namePart>Blass, Friedrich</namePart>
        <namePart type="date">1843-1907</namePart>
        <role>
          <roleTerm authority="marcrelator" type="text">editor</roleTerm>
        </role>
      </name>
```

```xml
        <identifier type="tlg">0028.003</identifier>
        <language objectPart="text">
            <languageTerm authority="iso639-2b" type="code">grc</languageTerm>
        </language>
        <part>
            <extent unit="pages">
                <start>29</start>
                <end>44</end>
            </extent>
        </part>
        <location><url
            displayLabel="GoogleBooks">http://books.google.com/books?id=LvfjhLc5pUwC&amp;pg=PA29</url></location>
            </relatedItem>
    <relatedItem type="constituent">
        <titleInfo>
            <title>Tetralogia 3</title>
        </titleInfo>
        <titleInfo type="alternative" lang="eng">
            <title>Third Tetralogy</title>
        </titleInfo>
        <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
            ns2:href="http://errol.oclc.org/laf/n82-101412.html              ">
            <namePart>Antiphon</namePart>
            <namePart type="date">ca. 480-411 B.C</namePart>
            <role>
                <roleTerm authority="marcrelator" type="text">creator</roleTerm>
            </role>
        </name>
        <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink" ns2:href="http://errol.oclc.org/laf/n86-46008.html">
            <namePart>Blass, Friedrich</namePart>
            <namePart type="date">1843-1907</namePart>
            <role>
                <roleTerm authority="marcrelator" type="text">editor</roleTerm>
            </role>
        </name>
        <identifier type="tlg">0028.004</identifier>
        <language objectPart="text">
            <languageTerm authority="iso639-2b" type="code">grc</languageTerm>
        </language>
        <part>
            <extent unit="pages">
                <start>44</start>
                <end>56</end>
            </extent>
        </part>
        <location><url
            displayLabel="GoogleBooks">http://books.google.com/books?id=LvfjhLc5pUwC&amp;pg=PA44</url></location>
            </relatedItem>
    <relatedItem type="constituent">
        <titleInfo>
            <title>De caede Herodis</title>
        </titleInfo>
        <titleInfo type="alternative" lang="eng">
            <title>On the murder of Herodes</title>
        </titleInfo>
        <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
            ns2:href="http://errol.oclc.org/laf/n82-101412.html              ">
            <namePart>Antiphon</namePart>
            <namePart type="date">ca. 480-411 B.C</namePart>
            <role>
                <roleTerm authority="marcrelator" type="text">creator</roleTerm>
            </role>
        </name>
        <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink" ns2:href="http://errol.oclc.org/laf/n86-46008.html">
            <namePart>Blass, Friedrich</namePart>
            <namePart type="date">1843-1907</namePart>
            <role>
                <roleTerm authority="marcrelator" type="text">editor</roleTerm>
            </role>
        </name>
```

```xml
          <identifier type="tlg">0028.005</identifier>
          <language objectPart="text">
              <languageTerm authority="iso639-2b" type="code">grc</languageTerm>
          </language>
          <part>
              <extent unit="pages">
                  <start>56</start>
                  <end>94</end>
              </extent>
          </part>
          <location><url
              displayLabel="GoogleBooks">http://books.google.com/books?id=LvfjhLc5pUwC&amp;pg=PA56</url></location>
      </relatedItem>
      <relatedItem type="constituent">
          <titleInfo>
              <title>De choreuta</title>
          </titleInfo>
          <titleInfo type="alternative" lang="eng">
              <title>On the choreutes</title>
          </titleInfo>
          <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
              ns2:href="http://errol.oclc.org/laf/n82-101412.html              ">
              <namePart>Antiphon</namePart>
              <namePart type="date">ca. 480-411 B.C</namePart>
              <role>
                  <roleTerm authority="marcrelator" type="text">creator</roleTerm>
              </role>
          </name>
          <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink" ns2:href="http://errol.oclc.org/laf/n86-46008.html">
              <namePart>Blass, Friedrich</namePart>
              <namePart type="date">1843-1907</namePart>
              <role>
                  <roleTerm authority="marcrelator" type="text">editor</roleTerm>
              </role>
          </name>
          <identifier type="tlg">0028.006</identifier>
          <language objectPart="text">
              <languageTerm authority="iso639-2b" type="code">grc</languageTerm>
          </language>
          <part>
              <extent unit="pages">
                  <start>94</start>
                  <end>115</end>
              </extent>
          </part>
          <location><url
              displayLabel="GoogleBooks">http://books.google.com/books?id=LvfjhLc5pUwC&amp;pg=PA94</url></location>
      </relatedItem>
      <relatedItem type="constituent">
          <titleInfo>
              <title>Fragmenta</title>
          </titleInfo>
          <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
              ns2:href="http://errol.oclc.org/laf/n82-101412.html          ">
              <namePart>Antiphon</namePart>
              <namePart type="date">ca. 480-411 B.C</namePart>
              <role>
                  <roleTerm authority="marcrelator" type="text">creator</roleTerm>
              </role>
          </name>
          <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink" ns2:href="http://errol.oclc.org/laf/n86-46008.html">
              <namePart>Blass, Friedrich</namePart>
              <namePart type="date">1843-1907</namePart>
              <role>
                  <roleTerm authority="marcrelator" type="text">editor</roleTerm>
              </role>
          </name>
          <identifier type="tlg">0028.007</identifier>
          <language objectPart="text">
              <languageTerm authority="iso639-2b" type="code">grc</languageTerm>
```

```xml
      </language>
      <part>
        <extent unit="pages">
          <start>116</start>
          <end>129</end>
        </extent>
      </part>
      <location><url
        displayLabel="GoogleBooks">http://books.google.com/books?id=LvfjhLc5pUwC&amp;pg=PA116</url></location>
  </relatedItem>
  <relatedItem type="constituent">
    <titleInfo>
      <title>Fragmenta</title>
    </titleInfo>
    <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
      ns2:href="http://errol.oclc.org/laf/n92-70218.html">
      <namePart>Antiphon</namePart>
      <namePart type="termsOfAddress">of Athens</namePart>
      <role>
        <roleTerm authority="marcrelator" type="text">creator</roleTerm>
      </role>
    </name>
    <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink" ns2:href="http://errol.oclc.org/laf/n86-46008.html">
      <namePart>Blass, Friedrich</namePart>
      <namePart type="date">1843-1907</namePart>
      <role>
        <roleTerm authority="marcrelator" type="text">editor</roleTerm>
      </role>
    </name>
    <identifier type="tlg">1147.001</identifier>
    <language objectPart="text">
      <languageTerm authority="iso639-2b" type="code">grc</languageTerm>
    </language>
    <part>
      <extent unit="pages">
        <start>130</start>
        <end>150</end>
      </extent>
    </part>
    <location><url
      displayLabel="GoogleBooks">http://books.google.com/books?id=LvfjhLc5pUwC&amp;pg=PA130</url></location>
  </relatedItem>
      <relatedItem type="constituent">
    <titleInfo>
      <title>Fragmenta</title>
    </titleInfo>
    <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
      ns2:href="http://errol.oclc.org/laf/n82-106049.html              ">
      <namePart>Gorgias</namePart>
      <namePart type="termsOfAddress">of Leontini</namePart>
      <role>
        <roleTerm authority="marcrelator" type="text">creator</roleTerm>
      </role>
    </name>
    <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink" ns2:href="http://errol.oclc.org/laf/n86-46008.html">
      <namePart>Blass, Friedrich</namePart>
      <namePart type="date">1843-1907</namePart>
      <role>
        <roleTerm authority="marcrelator" type="text">editor</roleTerm>
      </role>
    </name>
    <identifier type="tlg">0593.003</identifier>
    <note>These fragments contain various works such as "Encomium in Helenam" etc.</note>
    <language objectPart="text">
      <languageTerm authority="iso639-2b" type="code">grc</languageTerm>
    </language>
    <part>
      <extent unit="pages">
        <start>150</start>
        <end>174</end>
```

```xml
        </extent>
      </part>
      <location><url
        displayLabel="GoogleBooks">http://books.google.com/books?id=LvfjhLc5pUwC&amp;pg=PA150</url></location>
</relatedItem>
        <relatedItem type="constituent">
      <titleInfo>
        <title>Fragmenta</title>
      </titleInfo>
      <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
        ns2:href="http://errol.oclc.org/laf/n86-119012.html              ">
        <namePart>Antisthenes</namePart>
        <namePart type="date">ca. 445-ca. 360 B.C</namePart>
        <role>
          <roleTerm authority="marcrelator" type="text">creator</roleTerm>
        </role>
      </name>
      <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink" ns2:href="http://errol.oclc.org/laf/n86-46008.html">
        <namePart>Blass, Friedrich</namePart>
        <namePart type="date">1843-1907</namePart>
        <role>
          <roleTerm authority="marcrelator" type="text">editor</roleTerm>
        </role>
      </name>
      <identifier type="tlg">0591.002</identifier>
      <language objectPart="text">
        <languageTerm authority="iso639-2b" type="code">grc</languageTerm>
      </language>
      <part>
        <extent unit="pages">
          <start>175</start>
          <end>182</end>
        </extent>
      </part>
      <location><url
        displayLabel="GoogleBooks">http://books.google.com/books?id=LvfjhLc5pUwC&amp;pg=PA175</url></location>
</relatedItem>
<relatedItem type="constituent">
      <titleInfo>
        <title>Fragmenta</title>
      </titleInfo>
      <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
        ns2:href="http://errol.oclc.org/laf/n87-855622.html              ">
        <namePart>Alcidamas,</namePart>
        <namePart type="date">4th cent. B.C</namePart>
        <role>
          <roleTerm authority="marcrelator" type="text">creator</roleTerm>
        </role>
      </name>
      <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink" ns2:href="http://errol.oclc.org/laf/n86-46008.html">
        <namePart>Blass, Friedrich</namePart>
        <namePart type="date">1843-1907</namePart>
        <role>
          <roleTerm authority="marcrelator" type="text">editor</roleTerm>
        </role>
      </name>
      <identifier type="tlg">0610.001</identifier>
      <language objectPart="text">
        <languageTerm authority="iso639-2b" type="code">grc</languageTerm>
      </language>
      <part>
        <extent unit="pages">
          <start>183</start>
          <end>205</end>
        </extent>
      </part>
      <location><url
        displayLabel="GoogleBooks">http://books.google.com/books?id=LvfjhLc5pUwC&amp;pg=PA183</url></location>
</relatedItem>
<recordInfo>
```

```xml
            <recordContentSource authority="marcorg">OCoLC</recordContentSource>
            <recordCreationDate encoding="marc"/>
            <recordChangeDate encoding="iso8601"/>
            <recordIdentifier/>
        </recordInfo>
    </mods>
</modsCollection>
```

Figure 13: Expression Level MODS record for Antiphon's *In Novercam*

```xml
<?xml version="1.0" encoding="UTF-8"?>
<mods xmlns="http://www.loc.gov/mods/v3"
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xsi:schemaLocation="http://www.loc.gov/mods/v3 http://www.loc.gov/standards/mods/v3/mods-3-2.xsd">
    <titleInfo>
        <title>In Novercam</title>
    </titleInfo>
    <titleInfo type="translated" lang="eng">
        <title>Against the Stepmother for Poisoning</title>
    </titleInfo>
    <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
        ns2:href="http://errol.oclc.org/laf/n82-101412.html">
        <namePart>Antiphon</namePart>
        <namePart type="date">ca. 480-411 B.C</namePart>
        <role>
            <roleTerm authority="marcrelator" type="text">creator</roleTerm>
        </role>
    </name>
    <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
        ns2:href="http://errol.oclc.org/laf/n86-46008.html">
        <namePart>Blass, Friedrich</namePart>
        <namePart type="date">1843-1907</namePart>
        <role>
            <roleTerm authority="marcrelator" type="text">editor</roleTerm>
        </role>
    </name>
    <identifier type="tlg">0028.001</identifier>
    <language objectPart="text">
        <languageTerm authority="iso639-2b" type="code">grc</languageTerm>
    </language>
    <part>
        <extent unit="pages">
            <start>1</start>
            <end>11</end>
        </extent>
    </part>
    <location><url
        displayLabel="GoogleBooks">http://books.google.com/books?id=LvfjhLc5pUwC&amp;pg=PA1</url></location>
    <relatedItem type="host">
        <titleInfo>
            <title>Orationes et fragmenta</title>
            <subTitle>adivnctis Gorgiae, Antisthenis, Alcidamantis, declamationibvs</subTitle>
        </titleInfo>
        <name authority="naf" type="personal" xmlns:ns2="http://www.w3.org/1999/xlink"
            ns2:href="http://errol.oclc.org/laf/n86-46008.html">
            <namePart>Blass, Friedrich</namePart>
            <namePart type="date">1843-1907</namePart>
            <role>
                <roleTerm authority="marcrelator" type="text">editor</roleTerm>
            </role>
        </name>
        <typeOfResource>text</typeOfResource>
        <originInfo>
            <place>
                <placeTerm type="code" authority="marccountry">gw</placeTerm>
            </place>
            <place>
                <placeTerm type="text">Lipsiae</placeTerm>
```

```xml
        </place>
        <publisher>Tevbneri</publisher>
        <dateIssued>1892</dateIssued>
        <edition> Editio altera correctior</edition>
        <issuance>monographic</issuance>
      </originInfo>
      <language objectPart="text">
        <languageTerm authority="iso639-2b" type="code">grc</languageTerm>
      </language>
      <language objectPart="preface">
        <languageTerm authority="iso639-2b" type="code">lat</languageTerm>
      </language>
      <language objectPart="notes">
        <languageTerm authority="iso639-2b" type="code">lat</languageTerm>
      </language>
      <physicalDescription>
        <form authority="marcform">print</form>
        <extent>xlvii, 212 p. ; 18 cm.</extent>
      </physicalDescription>
      <note type="statement of responsibility">edidit Fridericvs Blass.</note>
      <note>Text in Greek; editorial material in Latin.</note>
      <classification authority="lcc">PA3869</classification>
      <identifier type="oclc">62356548</identifier>
      <location><url displayLabel="WorldCat">http://worldcat.org/oclc/62356548</url></location>
      <location><url displayLabel="GoogleBooks">http://books.google.com/books?id=LvfjhLc5pUwC</url></location>
      <location><url displayLabel="Open Content Alliance">http://www.archive.org/details/antiphontisorati00antiuoft</url></location>
      <location><url displayLabel="Open Library">http://www.openlibrary.org/details/antiphontisorati00antiuoft</url></location>
      <relatedItem type="series">
        <titleInfo>
          <title>Bibliotheca scriptorum Graecorum et Romanorum Teubneriana. [S.
            g.]</title>
        </titleInfo>
        <titleInfo type="abbreviated">
          <title>Teubner</title>
        </titleInfo>
      </relatedItem>
    </relatedItem>
  </mods>
```

Figure 14: Enhanced MADS record for Aeschylus

```xml
<?xml version="1.0" encoding="UTF-8"?>
<mads:mads xmlns:mads="http://www.loc.gov/mads/" xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:mods="http://www.loc.gov/mods/v3" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.loc.gov/mads/ http://www.loc.gov/standards/mads/mads.xsd http://www.loc.gov/mods/v3
http://www.loc.gov/standards/mods/v3/mods-3-2.xsd"
  version="1.0">
  <mads:authority>
    <mads:name type="personal" authority="naf">
      <mads:namePart>Aeschylus</mads:namePart>
    </mads:name>
  </mads:authority>
  <mads:variant type="other" lang="grc">
    <mads:name type="personal">
      <mads:namePart>Αισχύλος</mads:namePart>
    </mads:name>
  </mads:variant>
  <mads:variant type="other" lang="dut">
    <mads:name type="personal">
      <mads:namePart>Eschylus</mads:namePart>
    </mads:name>
  </mads:variant>
  <mads:variant type="other" lang="ger">
    <mads:name type="personal">
      <mads:namePart>Aischylos</mads:namePart>
      <mads:namePart type="termsOfAddress">aus Athen</mads:namePart>
    </mads:name>
```

```xml
    </mads:variant>
    <mads:variant type="other">
      <mads:name type="personal" authority="tlg">
        <mads:namePart>Aeschylus</mads:namePart>
        <mads:namePart type="termsOfAddress">Trag.</mads:namePart>
      </mads:name>
    </mads:variant>
    <mads:variant type="other">
      <mads:name type="personal" authority="lsj">
        <mads:namePart>Aeschylus</mads:namePart>
        <mads:namePart type="termsOfAddress">Tragicus</mads:namePart>
      </mads:name>
    </mads:variant>
    <mads:variant type="abbreviation">
      <mads:name type="personal" authority="lsj">
        <mads:namePart>A.</mads:namePart>
      </mads:name>
    </mads:variant>
    <mads:variant type="other" lang="spa">
      <mads:name type="personal">
        <mads:namePart>Esquilo</mads:namePart>
      </mads:name>
    </mads:variant>
    <mads:variant type="other">
      <mads:name type="personal">
        <mads:namePart>Eschilo</mads:namePart>
      </mads:name>
    </mads:variant>
    <mads:variant type="other">
      <mads:name type="personal">
        <mads:namePart>Eshil</mads:namePart>
      </mads:name>
    </mads:variant>
    <mads:variant type="other">
      <mads:name type="personal">
        <mads:namePart>Ã&#134;skiÌ&#129;los</mads:namePart>
      </mads:name>
    </mads:variant>
    <mads:variant type="other">
      <mads:name type="personal">
        <mads:namePart>Ajschylos</mads:namePart>
      </mads:name>
    </mads:variant>
    <mads:variant type="other">
      <mads:name type="personal">
        <mads:namePart>Eschil</mads:namePart>
      </mads:name>
    </mads:variant>
    <mads:variant type="other" lang="fre">
      <mads:name type="personal">
        <mads:namePart>Eschyle</mads:namePart>
      </mads:name>
    </mads:variant>
    <mads:variant type="other">
      <mads:name type="personal">
        <mads:namePart>AÌ&#136;schylos</mads:namePart>
      </mads:name>
    </mads:variant>
    <mads:variant type="other">
      <mads:name type="personal">
        <mads:namePart>Eskili</mads:namePart>
      </mads:name>
    </mads:variant>
    <mads:variant type="other">
      <mads:name type="personal">
        <mads:namePart>AiszkhuÌ&#136;losz</mads:namePart>
      </mads:name>
    </mads:variant>
    <mads:variant type="other">
      <mads:name type="personal">
```

```xml
        <mads:namePart>Eschylos</mads:namePart>
      </mads:name>
    </mads:variant>
    <mads:variant type="other">
      <mads:name type="personal">
        <mads:namePart>Iskilos</mads:namePart>
      </mads:name>
    </mads:variant>
    <mads:variant type="other">
      <mads:name type="personal">
        <mads:namePart>Aischúlos</mads:namePart>
      </mads:name>
    </mads:variant>
    <mads:note type="source">His Agamemnon, 1940</mads:note>
    <mads:note type="source">His The Oresteia, 1984: CIP t.p. (Aeschylus) data sheet (b. 525 B.C.)</mads:note>
    <mads:note type="source">His OÌ&#129;resteia, 1983: t.p. (Ã&#134;skiÌ&#129;los)</mads:note>
    <mads:note type="source">Chodkowski, R.R. "Agamemnon" Ajschylosa, 1985: t.p. (Ajschylosa)</mads:note>
    <mads:note type="source">Stroe, A. Orestia, 1984: t.p. (Eschil; Eschyle)</mads:note>
    <mads:note type="source">His Agamemnon des AÌ&#136;schylos, 1920</mads:note>
    <mads:note type="source">Kadare, I. Eschyle, ou, L'eÌ&#129;ternel perdant, c1988: t.p. verso
      (Eskili)</mads:note>
    <mads:note type="source">GoÌ&#136;roÌ&#136;g draÌ&#129;maÌ&#129;k, 1971: p. 2
      (AiszkhuÌ&#136;losz)</mads:note>
    <mads:note type="source">His Agamemnon, 1995: CIP t.p. (Aeschylus) pub. blurb (c. 525 BC-456 BC)</mads:note>
    <mads:note type="source">His The seven against Thebes, 1998: CIP galley (Aeschylus (c. 525
      BC-456 BC) Greek playwright, born at Eleusis, near Athens, generally considered to be the
      earliest important writer of the Western theatrical tradition, the first playwright to
      achieve official recognition in ancient Greece. Of his 90 plays only 7 survive)</mads:note>
    <mads:note type="source">Tragedye Eschylosa, 1873</mads:note>
    <mads:note type="source">Der geshmiedtÌ£er PrometÌ£heus, 19--?</mads:note>
    <mads:fieldOfActivity>tragic poet</mads:fieldOfActivity>
    <mads:identifier type="lccn">n 79055702 </mads:identifier>
    <mads:url>http://orlabs.oclc.org/Identities/key/lccn-n79-55702</mads:url>
    <mads:extension>
      <mads:description>List of related work identifiers</mads:description>
      <identifier type="tlg">0085.001</identifier>
      <identifier type="tlg">0085.002</identifier>
      <identifier type="tlg">0085.003</identifier>
      <identifier type="tlg">0085.004</identifier>
      <identifier type="tlg">0085.005</identifier>
      <identifier type="tlg">0085.006</identifier>
      <identifier type="tlg">0085.007</identifier>
      <identifier type="tlg">0085.008</identifier>
      <identifier type="tlg">0085.009</identifier>
      <identifier type="tlg">0085.010</identifier>
      <identifier type="tlg">0085.011</identifier>
      <identifier type="Perseus:abo">Perseus:abo:tlg,0085,001 </identifier>
      <identifier type="Perseus:abo">Perseus:abo:tlg,0085,002 </identifier>
      <identifier type="Perseus:abo">Perseus:abo:tlg,0085,003 </identifier>
      <identifier type="Perseus:abo">Perseus:abo:tlg,0085,004 </identifier>
      <identifier type="Perseus:abo">Perseus:abo:tlg,0085,005 </identifier>
      <identifier type="Perseus:abo">Perseus:abo:tlg,0085,006 </identifier>
      <identifier type="Perseus:abo">Perseus:abo:tlg,0085,007 </identifier>
    </mads:extension>
    <mads:recordInfo>
      <mods:recordContentSource authority="marcorg">DLC</mods:recordContentSource>
      <mods:recordCreationDate encoding="marc">790702</mods:recordCreationDate>
      <mods:recordChangeDate encoding="iso8601">20021210033154.0</mods:recordChangeDate>
      <mods:recordIdentifier>oca00288811 </mods:recordIdentifier>
      <mods:languageOfCataloging>
        <mods:languageTerm type="code" authority="iso639-2b">eng</mods:languageTerm>
      </mods:languageOfCataloging>
    </mads:recordInfo>
</mads:mads>
```

Figure 15: Perseus Created MADS Record for Adaeus (Includes info from merged DNB record)

```xml
<?xml version="1.0" encoding="UTF-8"?>
<mads:mads xmlns:mads="http://www.loc.gov/mads/"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:mods="http://www.loc.gov/mods/v3"
  xmlns:xlink="http://www.w3.org/1999/xlink" version="1.0"
  xsi:schemaLocation="http://www.loc.gov/mads/ http://www.loc.gov/standards/mads/mads.xsd
  http://www.loc.gov/mods/v3 http://www.loc.gov/standards/mods/v3/mods-3-2.xsd">
  <mads:authority>
    <mads:name type="personal" authority="tlg">
      <mads:namePart>Adaeus</mads:namePart>
      <mads:namePart type="termsOfAddress">Epigr.</mads:namePart>
      <mads:namePart type="date">A.D. 1</mads:namePart>
    </mads:name>
  </mads:authority>
  <mads:variant>
    <mads:name type="personal" authority="dnb">
      <mads:namePart>Adaeus</mads:namePart>
      <mads:namePart type="termsOfAddress">Macedo</mads:namePart>
      <mads:namePart type="date">1. Jh. n. Chr</mads:namePart>
    </mads:name>
  </mads:variant>
  <mads:variant type="other">
    <mads:name type="personal" authority="dnb">
      <mads:namePart>Adaeus</mads:namePart>
      <mads:namePart type="termsOfAddress">Rhetor</mads:namePart>
      <mads:namePart type="date">1. Jh. n. Chr</mads:namePart>
    </mads:name>
  </mads:variant>
  <mads:variant type="other" lang="ger">
    <mads:name type="personal" authority="dnb">
      <mads:namePart>Adaios</mads:namePart>
      <mads:namePart type="termsOfAddress">von Makedonien</mads:namePart>
      <mads:namePart type="date">1. Jh. n. Chr</mads:namePart>
    </mads:name>
  </mads:variant>
  <mads:variant type="other" lang="ger">
    <mads:name type="personal" authority="dnb">
      <mads:namePart>Adaios</mads:namePart>
      <mads:namePart type="termsOfAddress">Epigrammatiker des Philippos-Kranzes</mads:namePart>
      <mads:namePart type="date">1. Jh. n. Chr</mads:namePart>
    </mads:name>
  </mads:variant>
  <mads:variant type="other" lang="ger">
    <mads:name type="personal" authority="dnb">
      <mads:namePart>Adaios</mads:namePart>
      <mads:namePart type="termsOfAddress">Epigrammdichter</mads:namePart>
      <mads:namePart type="date">1. Jh. n. Chr</mads:namePart>
    </mads:name>
  </mads:variant>
  <mads:variant type="other">
    <mads:name type="personal">
      <mads:namePart>Adaios</mads:namePart>
      <mads:namePart type="termsOfAddress">Rhetor</mads:namePart>
      <mads:namePart type="date">1. Jh. n. Chr</mads:namePart>
    </mads:name>
  </mads:variant>
  <mads:variant type="other" lang="ger">
    <mads:name type="personal" authority="dnb">
      <mads:namePart>Adaeus</mads:namePart>
      <mads:namePart type="termsOfAddress">Lehrer des Arellius Fuscus</mads:namePart>
      <mads:namePart type="date">1. Jh. n. Chr</mads:namePart>
    </mads:name>
  </mads:variant>
  <mads:variant type="other">
    <mads:name type="personal" authority="Smith">
      <mads:namePart>Addaeus</mads:namePart>
```

```xml
        </mads:name>
      </mads:variant>
    <mads:variant type="other">
      <mads:name type="personal" authority="lsj">
        <mads:namePart>Adaeus </mads:namePart>
        <mads:namePart type="termsOfAddress">Epigrammaticus</mads:namePart>
      </mads:name>
    </mads:variant>
    <mads:variant type="other"
      xlink:href="http://www.brillonline.nl/subscriber/entry?entry=bnp_e103220">
      <mads:name type="personal" authority="Brill">
        <mads:namePart>Adaeus</mads:namePart>
      </mads:name>
    </mads:variant>
    <mads:variant type="other" lang="ger">
      <mads:name type="personal" authority="Brill">
        <mads:namePart>Adaios</mads:namePart>
      </mads:name>
    </mads:variant>
    <mads:variant type="other" lang="grc">
      <mads:name type="personal">
        <mads:namePart>&#x41;&#x64;&#x61;&#x69;&#x6f;&#x6a;</mads:namePart>
      </mads:name>
    </mads:variant>
    <mads:variant type="other" lang="grc">
      <mads:name type="personal">
        <mads:namePart>Αδαιος</mads:namePart>
      </mads:name>
    </mads:variant>
    <mads:fieldOfActivity>epigrammatist</mads:fieldOfActivity>
    <mads:fieldOfActivity>epigrammatiker</mads:fieldOfActivity>
    <mads:note type="source">TLG Canon of Greek Authors and Works, Third Edition, p. 4</mads:note>
    <mads:note type="source">Smith's Dictionary of Greek and Roman Biography and Mythology, Vol 1
      1867, p. 18: "ADAEUS, or ADDAEUS, a Greek epigrammatic poet, a native most probably of
      Macedonia.....The time when he lived cannot be fixed with certainty. Reiske, though on
      insufficient grounds, believes these two to be the same person. (Anthl. Graec. vi. 228, 258,
      vii. 51, 238, 240, 305, x. 20; Brunck, Anal. ii. p. 224; Jacobs, xiii. p. 831.)"</mads:note>
    <mads:note type="source">"Adaeus, Macedonian epigrammatist, mainly -- and without convincing
      reasons -- identified with the Asian rhetorician A. [3], a contemporary of Seneca the
      Elder's...." from "Degani, Enzo (Bologna). "Adaeus." Brill's New Pauly. Antiquity volumes
      edited by: Hubert Cancik and Helmuth Schneider . Brill, 2007. Brill Online. Tufts University
      Library. 20 July 2007."</mads:note>
    <mads:identifier type="tlg">0102</mads:identifier>
    <mads:extension>
      <mads:description>List of related work identifiers</mads:description>
      <identifier type="tlg">0102.001</identifier>
    </mads:extension>
    <mads:url>http://quod.lib.umich.edu/cgi/t/text/pageviewer-
idx?c=moa;cc=moa;rgn=full%20text;idno=ACL3129.0001.001;didno=ACL3129.0001.001;view=image;seq=00000033</mads:url>
    <mads:url>http://dispatch.opac.ddb.de/DB=4.1/PPN?PPN=102377995</mads:url>
    <mads:recordInfo>
      <mods:recordContentSource>Perseus Digital Library</mods:recordContentSource>
      <mods:recordCreationDate encoding="iso8601">20070627</mods:recordCreationDate>
      <mods:recordChangeDate encoding="iso8601">20070910</mods:recordChangeDate>
      <mods:recordIdentifier source="DNB">DNB|102377995</mods:recordIdentifier>
      <mods:languageOfCataloging>
        <mods:languageTerm authority="iso639-2b" type="code">ger</mods:languageTerm>
      </mods:languageOfCataloging>
      <mods:languageOfCataloging>
        <mods:languageTerm authority="iso639-2b" type="code">eng</mods:languageTerm>
      </mods:languageOfCataloging>
    </mads:recordInfo>
  </mads:mads>
```

Figure 16: Perseus Created MADS record for Acholius (No record in the DNB)

```xml
<?xml version="1.0" encoding="UTF-8"?>
<mads:mads xmlns:mads="http://www.loc.gov/mads/"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:mods="http://www.loc.gov/mods/v3"
  xmlns:xlink="http://www.w3.org/1999/xlink" version="1.0"
  xsi:schemaLocation="http://www.loc.gov/mads/ http://www.loc.gov/standards/mads/mads.xsd
  http://www.loc.gov/mods/v3 http://www.loc.gov/standards/mods/v3/mods-3-2.xsd">
  <mads:authority>
    <mads:name type="personal" authority="stoa">
      <mads:namePart>Acholius</mads:namePart>
    </mads:name>
  </mads:authority>
  <mads:fieldOfActivity>historian</mads:fieldOfActivity>
  <mads:note type="source"/>
  <mads:note type="source">Smith's Dictionary of Greek and Roman Biography and Mythology, Vol 1
    1867, p. 12: "Acholius held the office of Magister Admissionum in the reign of Valerian. (a.
    c. 253 -260.) One of his works was entitled Acta, and contained an account of the history of
    Aurelian. It was in nine books at least. (Vopisc. Aurel. 12.) He also wrote the life of
    Alexander Severus. (Lamprid. Alex. Sev. 14. 48. 68.)" </mads:note>
  <mads:identifier type="stoa">stoa0004</mads:identifier>
  <mads:url>http://quod.lib.umich.edu/cgi/t/text/pageviewer-
idx?c=moa;cc=moa;rgn=full%20text;idno=ACL3129.0001.001;didno=ACL3129.0001.001;view=image;seq=00000027</mads:url>
  <mads:extension>
    <mads:description>List of related work identifiers</mads:description>
    <identifier type="stoa author-text">stoa0004-stoa001</identifier>
  </mads:extension>
  <mads:recordInfo>
    <mods:recordContentSource>Perseus Digital Library</mods:recordContentSource>
    <mods:recordCreationDate encoding="iso8601">20070627</mods:recordCreationDate>
    <mods:recordChangeDate encoding="iso8601">20070911</mods:recordChangeDate>
    <mods:recordIdentifier/>
  </mads:recordInfo>
</mads:mads>
```