SUMMARY. Metadata schemes target resources as information-packages, without attention to the distinction between content and carrier. Most schema are derived without empirical understanding of the concepts that need to be represented, the ways in which terms representing the central concepts might best be derived, and how metadata descriptions will be used for retrieval. Research is required to resolve this dilemma, and much research will be required if the plethora of schemes that already exist are to be made efficacious for resource description and retrieval. A preliminary study was designed to see whether the bibliographic concept of “the work” could be of any relevance among artifacts held by a museum. The “works metaphor” is extended from the bibliographic to the artifactual domain by altering the terms of the definition slightly, thus: instantiation is understood as content genealogy.
Case studies of Etruscan artifacts from the University of Pennsylvania Museum of Archaeology and Anthropology are used to demonstrate the inherence of the work in non-documentary artifacts.

**KEYWORDS.** Artifacts, cultural information resources, instantiation, content genealogy, content metadata

**INTRODUCTION**

Even in the short life-span of the term “metadata,” many complex schemes have been developed rapidly to assist in the description of resources, so as to facilitate retrieval. Yet, like their bibliographic counterparts—cataloging rules—metadata schemes have emerged with two major (if traditional) difficulties. The first is that the schemes target resources as information-packages, without attention to the distinction between content and carrier, elsewhere described as item and work (Smiraglia 2001). The second difficulty is that most schema are derived without empirical understanding of the concepts that need to be represented, the ways in which terms representing the central concepts might best be derived, and how metadata descriptions will be used for retrieval. Regardless of the complexity or the modernity of metadata schemes, the historical problem is perpetuated—once again we are creating retrieval tools based on resources in hand, selected essentially by chance, without knowledge of the characteristics of other, similar resources. Research is required to resolve this dilemma, and much research will be required if the plethora of schemes that already exist are to be made efficacious for resource description and retrieval. Even more research will be required to facilitate cross-walks for knowledge sharing.

This paper represents one small step toward such research. I will report the results of a preliminary study, which was designed to see whether the bibliographic concept of “the work” could be of any relevance among artifacts held by a museum. To do so, I begin by posing “the work” as a metaphor for the content of an information-package. The empirical portion of the study is qualitative, involving case study
analysis of artifacts, and unstructured interviews of museum staff. A preliminary report appeared in Smiraglia (2004), from which much of the introductory material here is derived.

THE “WORKS” METAPHOR

“Works” are key entities in the universe of recorded knowledge. (Works are sometimes also described by these key terms: opera, oeuvres, Werke, etc.) Elsewhere (Smiraglia 2001, 2002) I have demonstrated that substantial proportions of works in the bibliographic universe exist in multiple instantiations, which can be understood epistemologically as falling into two categories: works whose instantiations are derivative, and those that are mutations from the progenitor. Taxonomies that can classify instantiations of works have been developed by Tillett (1987), IFLA (1998), Smiraglia (2001), and others, and several authors have contributed differing sets of language to describe works (see Smiraglia 2003). Smiraglia (2001), Vellucci (1994), Yee (1993), and others have demonstrated the vast potential for variant instantiation (i.e., mutation, derivation, etc.).

Works mutate and derive across time and culture in response to their entrance into a canon of cultural meaning. That is, works that become somehow associated with meaning beyond their own texts—such as the Bible, or Gone with the Wind—will be reproduced in a wide variety of editions, adaptations, performances, and so forth. Systems for information retrieval must be designed with this in mind, to facilitate gathering of the instantiations of a given work, as well as selection of one from among many. The metaphor of the work and its instantiations is not a uniquely documentary phenomenon; it can be extended into other domains of recorded knowledge. (For instance, Coleman (2003) has demonstrated the application of the “works” metaphor to scientific models.)

The present study represents an attempt to extend the works metaphor into the realm of natural artifacts that might be held by museum repositories.

The characteristics of bibliographic works are as follows (see Smiraglia 2001):

1. “Work” is an abstract concept (immaterial, conceptual);
2. A work is a new synthesis of knowledge that consists of ideational content and semantic content;
3. Once expressed, a work may take a variety of physical instantiations;
4. The expression of a work may change freely in either ideational or semantic content or both;
5. Relationships among works are complex; and
6. A work’s essential role is as a vehicle of communication between its creator and its consumers.

Instantiations (often referred to as “versions” or “editions” or even “manifestations” of a work) are the artifacts of the individual points in time when a given work is set down, or recorded either for preservation or for dissemination. A concrete example: an author creates a story, writes it out as a manuscript (first instantiation), has an edited version of it published as the first edition (second instantiation), and when it becomes a hit it is translated into another language (third instantiation, this time a mutation), and also sells it to be made into a screenplay (fourth instantiation, another mutation). The motion picture is also a hit (fifth instantiation) and is sold on DVD (sixth instantiation, or second instantiation of the motion picture). This fame encourages repeated new editions of the original novel both in expensive hardback editions and in paperback editions (seventh and eighth instantiations, these are considered derivations). A mutation is an instantiation in which the semantic or ideational content or both have been altered; a derivation is an instantiation, subsequent to the first, in which no semantic or ideational content has changed. IFLA’s FRBR model with some confusion refers to these as “expressions” and “manifestations.”

From Documents to Artifacts, from Instantiation to Content Genealogy

To carry the metaphor from the bibliographic domain to the artifactual domain we alter the terms of the definition slightly. What we have heretofore called “instantiation” is now understood as content genealogy, and we expect there might be many categories of mutation and derivation. A new term, “re-presentation” (hereafter rendered simply as “representation”), is used to describe images or descriptions of otherwise unique artifacts (fine art, natural science)—representation clearly precedes instantiation. Artifacts, real-world information objects, have distinct ideational and expressive content apart from their carriers as do bibliographic works. Artifacts are unique and immutable, but their representations are potentially many and are infinitely mutable. Further,
whereas in the documentary domain no particular instantiation can be claimed to be concurrent with the work, in the artifactual domain the unique artifact is itself concurrent with the work. We begin with a simple model (see Figure 1). The artifact is unique and immutable. Its representations, such as a photograph, can be contained on websites or in books. The book itself is subject to derivation, such as multiple editions, all of which will contain the representation of the artifact. For each representation, there can be many sites of reproduction (representations in print), all of which might be subject to multiple instantiation.

**CASE STUDY OF ETRUSCAN ARTIFACTS**

To date there is little or no evidence about the instance or extent of the representation of artifacts. Case study in the qualitative vein is an appro-
appropriate first step in unexplored territory. Essential research questions boiled down to whether multiple representations of an artifact might be extant and might therefore need to be represented in an information system, and whether the “works” metaphor might be useful for describing sets of representations. Therefore, a set of artifacts was sought that would be easily available to the researcher and that would be likely to have appeared in representations in print as well as in digital form. Etruscan artifacts from the University of Pennsylvania Museum of Archaeology and Anthropology were selected to provide qualitative evidence of the content genealogy of representations of non-documentary artifacts. These artifacts are housed in a popular exhibition that is also well-represented in digital form on the Web (http://www.museum.upenn.edu/new/worlds_intertwined/etruscan/main.shtml) and in a published Guide (2002). Eight artifacts (represented in Table 1) were selected based on their prominence in the exhibit, and with attention to their diversity.

### TABLE 1. The Eight Artifacts as Described by UPM

<table>
<thead>
<tr>
<th>Material</th>
<th>Artifact</th>
<th>Source</th>
<th>Date</th>
<th>URL*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Terracotta</td>
<td>Hut urn</td>
<td>Said to have come from the area between Albano and Genzano</td>
<td>8th c. BC</td>
<td>earliest.shtml</td>
</tr>
<tr>
<td>Impasto</td>
<td>Kotyle</td>
<td>Narce, Tomb I</td>
<td>7th c. BC</td>
<td>orientalizing.shtml</td>
</tr>
<tr>
<td>Impasto**</td>
<td>Footed bowl</td>
<td>Narce, Tomb 19M</td>
<td>early 7th c. BC</td>
<td>faliscan.shtml</td>
</tr>
<tr>
<td>Impasto**</td>
<td>Etrusco-Corinthian Olpe</td>
<td>Vulci, Tomb B</td>
<td>6th c. BC</td>
<td>daily.shtml</td>
</tr>
<tr>
<td></td>
<td>Nenfro Lintel***</td>
<td>Ovrieto, Crocifisso del Tufo Necropolis</td>
<td>ca. 550 BC</td>
<td>language.shtml</td>
</tr>
<tr>
<td>Terracotta</td>
<td>Kantharos</td>
<td>Vulci, Tomb B</td>
<td>6th c. BC</td>
<td>technology.shtml</td>
</tr>
<tr>
<td></td>
<td>Antefix with Satyr Head</td>
<td>Caere (Cerveteri)</td>
<td>4th c. BC</td>
<td>architecture.shtml</td>
</tr>
<tr>
<td></td>
<td>Alabaster</td>
<td>Cinerary urn</td>
<td>3rd-2nd c. BC</td>
<td>final.shtml</td>
</tr>
</tbody>
</table>

*http://www.museum.upenn.edu/new/worlds_intertwined/etruscan/main.shtml
** "Craftsmen often made ceramic copies of the bronze conical stands used by rulers for banquet wine. The more economical material allowed large scale, extravagant forms.”
*** “These ‘talking objects’ seem to express magically the power that ancient peoples felt came from the ability to read and write.”
Specifically, operational research questions were:

1. How many representations in-house can be found for these eight artifacts?
2. Does their inclusion in the exhibit indicate the possibility of greater representation or instantiation?
3. Are the representations instantiated outside the museum (in literature, on the Web, etc.)?
4. Is there a difference between representation of an artifact that is deliberately informative (one that bears text) and one that is merely persistent?

We can begin, then, with reference to our model as the base set. For each artifact, it is apparent that the artifact is represented by one photograph, which in turn has been digitized for use on the Web. This, therefore, is the minimum set (A, AR₁, AR₂) among our cases.

**Representation Taxonomy**

The Museum Archivist was the gatekeeper for this study, the one person with access to all internal records about the artifacts. We began with an interview, from which a meta-level taxonomy of representation emerged. At the top level the taxonomy resembles our initial model, except that representation now can take one of two categories—metadata (textual description) or representation (image or model):

Artifact
   Metadata
      Instantiation
   Representation
      Instantiation

For each artifact in the museum, both textual descriptions (metadata) and representations (images or models) exist. Both metadata sets and representations are commonly reproduced in a variety of ways, thus yielding instantiations. In fact, the two nodes—metadata and representation—can be broken down into more specific information-object sets. These are:
Metadata
  Finding aid
  Field notes
  Letters
  Conservation treatment notes
  Register descriptions
    Object cards
  Image order invoices
  Museum database records
  Catalog card records

Representation
  Field photos
  Drawings
  Working images
  3D models
  Exhibition color images
    Digitized images
  Conservation photos
  Photo archives
    Negatives
    Prints
    Transparencies
  Object reproductions

This, then, is the meta-level set of metadata and representation instantiations that might exist for any given artifact within the repository. It is important to note that each of these information-objects serves a specific purpose within the museum.
The Eight Artifacts

For each artifact, every record in-house was examined and noted. The taxonomy served to guide the research. That is, finding aids were examined first, then field notes and registers, and so forth. Ultimately several divisions of the museum were visited, as each holds its own unique records. Whenever possible records were copied, yielding (ironically) yet another set of instantiations. It was no surprise to discover that for each artifact both metadata and representations exist. Table 2 contains the total numbers of representations and metadata sets that were found for each artifact, both in-house and outside the repository. We compiled a bibliography of 150 texts on the Etruscans from the collections of research libraries in New York. Unfortunately, examination of those texts yielded no representations of these eight artifacts. Likewise, Web searches revealed no representations (not even those on the museum’s own Website). But internal records indicated the presence of several publications in which both metadata and images could be found. The symbols appearing in the columns for in-house metadata and representations correspond to terms in the taxonomy. The symbols for external sources correspond to publications on a short list:

The Gallery of Italic and Etruscan Art. 1933. The University Museum bulletin v. 4 no. 6 (Nov.): 6.
The list is in chronological order, which is itself informative. Note that there are periodic clusters around exhibits, catalogs, and journal articles. (Because the list was compiled from museum sources, some of the citations are incomplete and could not be verified independently.) Each of these publications appears to be unique—a search on OCLC’s Worldcat in July 2004 turned up no additional editions or other instantiations.

The chief feature of Table 2 is the fact that there is disparity among the artifacts in the quantity of metadata and image representations. That is, beyond the initial set of register description and exhibit image, certain of the artifacts have substantial sets of information-objects, both in-house and outside the museum. The Etrusco-Corinthian olpe has the most, and the lintel has the fewest. Because all eight artifacts were selected from the Web version of a popular exhibition, and all are included in the exhibition catalog, these factors must not contribute to the size of the representation set. The next factor checked was the presence of field notes—it seemed reasonable to consider whether the extent of notes created at the time of excavation might have any influence on the size of the representation set. But, as Table 2 demonstrates, no field notes were found for three of the artifacts—the Etrusco-Corinthian olpe, the hut urn, and the lintel. The olpe (with no field notes) has the largest representation set in the study, and the cinerary urn, which has the next largest set, has extensive field notes. So the presence of field notes was ruled out as a factor contributing to size of the representation set. Finally, whether the artifacts is a deliberately signifying object—one that has writing on it—also can be ruled out in this study. The lintel, with few representations, and the cinerary urn, with many, are the two signifying objects in the study.

Earlier I used the term ‘content genealogy’ to describe the succession of representations that occur along a chronological continuum. There are several rich examples in this study. The easiest to understand, perhaps, is the content of the museum’s database, which includes verbatim
the content of the object record cards; this creates a simple set: M, M'.
In addition to forming a linear sequence in time (a parallel to the bibliographic successive relationship), this relationship is one in which the original metadata set is embedded in a new set; this “embedding” parallels the amplification relationship demonstrated in the bibliographic arena. Another simple example is the digital image that appears on the website, which is the same as that which appears in the printed *Guide*; for each artifact there is a set: R', R'. The olpe yields some more complex examples. A lantern slide image, for which there is also a nitrate negative, a 35 mm. negative, and 3 copies of the 35 mm. negative, appears as a photograph in Dohan (1934); this yields the set: R'_ln, R'_nn, R'_{35n(35n1-3)}, R'_pr, R'_p2. There are two such (different) images of the olpe. And, in fact, all of the images listed in the column headed “representations in print” come from images held in the museum, with similar genealogies. The problem for metadata is to describe each representation accordingly and to allow linkage for collocation.

The last part of the analysis involved removing all of the functional metadata and representations—those created in-house—to see whether any difference might emerge in the size of the representation set. The reasoning here was that the in-house metadata sets seemed nearly identical across all eight artifacts. Given that, an obvious question is: what would the distribution look like without those sets? The answer, as before, is ambiguous. The olpe and the cinerary urn have the largest sets of representations in print, followed closely by the satyr head antefix. In sum, the olpe and the cinerary urn have the most representations and instantiations, and the satyr head is close behind. The other five artifacts have much smaller representation sets. Presence in the exhibit or on the Web, field notes, and signifying content all could be ruled out as contributors to the size of the representation set. However, the three artifacts with substantial representation sets all seem to have appeared in published volumes soon after their acquisition.

This observation was confirmed in an interview with the Conservator, who called these items “frequent flyers.” According to her, once an object has been published it gets requested for loan, photos of it are requested, and so forth, even if the museum has better or more interesting exemplars items in its collections. These items spend more time on loan than they do in the museum on exhibit. Each time the artifact is loaned an entirely new set of before- and after-metadata and photographs are generated to insure the continued integrity of the artifact. And, while on loan, the artifact is likely to have metadata and image rep-
resentations published in the loan-exhibition catalog, which will likely
generate more attention over time.

**Metadata Forms and Information Ecologies**

Earlier we noted that the artifacts all had more or less equivalent
metadata sets in-house. These sets can be summarized at the meta-level
in terms of their sources (field notes, accession ledgers, object records,
conservation records and reports, and online catalog records) and data
types (object entities, and activity-based descriptors). In fact, each divi-
sion of the museum that deals with the artifact generates its own
metadata set and might generate its own image representations as well.
Object-type metadata sets tend to include a common set of descriptors,
but activity-based descriptors differ from division to division depend-
ing on the work that has taken place. The accession ledger contains in-
formation about where and when the artifact was found, the object
record contains terms that locate the artifact in culture (people, time,
place), the conservation records related precise physical characteristics
and actions taken, loan records indicate the time and place to which the
item was loaned along with any special care requirements, and condi-
tion reports (used to note a breakage or other anomaly) include a narra-
tive about the incoming and outgoing condition of the artifact. The
museum’s online catalog uses a complex template that includes “tabs”
for all of the above, within which each division’s own metadata set can
reside.

But as noted before, each artifact also seems to carry its own unique
identifier set from division to division. This set, which I call the “object
entity,” includes:

- Object type
- Material
- Culture
- Source
- Collector
- Date acquired.

This, then, can be said to be the base metadata set uncovered in-house.
In this museum, at least, this set of identifiers is found useful across di-
visions and has the capability to uniquely identify each artifact.
However, there was little authority control—what was called “Etruscan” in one place might have been called “Etrurian” or “Italic” in another, for instance. These terms are not quite synonyms, but each was considered a useful identifier in one context or another. One artifact, the lintel, had two object records with two different numbers. The second record was uncovered during this research, but it was not the record used by the museum to represent the artifact in the exhibition or elsewhere.

The most interesting observation was the degree to which each division of the museum generates its own metadata and representations. Albrechtsen (2000) has written about what she calls “information-ecologies,” groups of people working together who generate their own domain-specific vocabulary and ontology. This museum is a perfect example. While a meta-level description exists in-house and is useful across the institution, each division of the museum is also encouraged to use its own information ecology to generate useful metadata, all of which eventually reside together in the online catalog.

CONCLUSIONS

The research questions for this study focused attention on whether there would be many representations for each artifact in-house (there were), whether there might also be many representations externally (and again there were), and whether mitigating factors concerning the size and diversity of the representation set could be identified. We found that these artifacts had in-house representations at a fairly predictable rate based on internal museum functions. Each artifact was represented in each division of the museum as it passed from accession to conservation to exhibit to loan, and so forth. Divergence in the size and diversity of the representation set appeared to be a function of the “popularity” of the artifact—popularity here defined as the aftermath of having appeared in print. Thus popularity, at least among these eight artifacts, was a function of publication and not of exhibit. We found that the object-entity has a stable and consistent metadata set. But we also discovered that specific ecologies within the museum contribute varying but related metadata sets as well.

All of this suggests that a meta-theory of content genealogy is possible. Such a meta-theory would harmonize the “works” metaphor from the bibliographic domain with a “representation” metaphor from the artifactual domain. Instantiation, or the evolution and reproduction of
representations is observed in both domains. Canonicity, which seems to contribute to the size of instantiation sets in the bibliographic domain, is mirrored in the artifactual domain by what we have here called “popularity.” At a meta-level, cultural acceptance of a work or an artifact creates public demand for more representations. This is in line with Eggert’s suggestion that works are collaborations over time—participatory entities that incorporate the “agencies” of those who encounter them (1994, 76).

Much more empirical evidence is necessary. The results here represent a preliminary analysis of a preliminary study—we have seen some interesting concepts emerge that might help formulate a larger-scale empirical study. The evolution, derivation, mutation, and dissemination of ideational and expressive content across time, culture, and linguistic boundaries must be the subject of further investigation. And any meta-theory must describe the interaction of signification (meaning) and reception (or perception).

Nash (2002) narrates the history of early Philadelphia, interpreting artifacts and archival evidence to compile the untold story of the first capital of the United States, which was also the largest and most racially diverse city of colonial America. He lends credence to our effort with this observation concerning a coffeepot (Nash 2002, 66-68):

Such a piece as a coffeepot . . . provides an example of how such an artifact can have multiple meanings. It can be viewed most directly as a handsome example of high-style eighteenth-century craftsmanship, as an intrinsically valuable work of decorative art. Through a second lens, the coffeepot can be seen as a crucial piece of evidence in tracing the new meaning of gentility in the eighteenth century. Amid rising consumerism, in both England and its colonies, genteel people developed a new sense of refinement, acted out in elegant manners, witty conversation, and graceful movements on occasions that depended on the importation of new beverages from exotic ports of call—in this case coffee beans from South America. Through a third lens, the Richardson coffeepot can be considered, although not actually seen, with regard to the organization of rhythms of work of the artisan who created the object. Behind the coffeepot lay several work processes involving African cultivation of the coffee beans, the sailors who shipped them to Philadelphia, and the small silversmith workshop production that linked together the labor of apprentices, journeymen, and master craftsmen. Finally, behind the coffeepot, absent from the
view of the lovely pot itself, resided the role of the crafts worker in the political and social life of a port town such as Philadelphia.

The richness of artifactual evidence belies the simplicity of the metadata that describe them. An artifact, like a “work,” is greater than the sum of its perceptions. Sitting in the focal point, an artifact can help a scholar tell a story that helps us better understand the world in which we live. The collectivity of the representations of the artifact, joined together in a virtual museum, offer the scholarly world a potentially much richer information resource. We will continue to need better empirical understanding to aid in the development of content metadata to meet that challenge.

WORKS CITED


Smiraglia, Richard P. 2004. Knowledge sharing and content genealogy: Extending the “works” model as a metaphor for non-documentary artifacts with case studies of Etruscan artifacts. In McIlwaine, Ia ed., Knowledge organization and the global in-
Part I: Intellectual Foundations


