Instantiation: Toward a Theory

Abstract: The concept of “instantiation,” the phenomenon of realization in time, is emerging in knowledge organization. Studies have demonstrated some consistent theoretical parameters. Epistemological and content analysis provide a background for meta-analysis. The result points the way to lacunae in the understanding of instantiation, including flaws in the FRBR model.


1.1 Instant, Instance, Instantiate, Instantiation

In the sphere of metadata, and in particular in the area of knowledge representation, the emerging concept of instantiation holds promise for the construction of increasingly sophisticated retrieval mechanisms. To move across the spectrum from raw data to information requires appropriate methods for the representation of knowledge. The complex phenomenon of instantiation must be understood from both theoretical and pragmatic perspectives to facilitate information retrieval and knowledge discovery.

Instantiation, essentially, is a generic term for the phenomenon of realization in time. Other terms are associated with the concept, but with more problematic overtones in their definitions. For instance, “version” is a term often used. Version implies deliberation in the creation of the phenomenon and also alteration. Another term used is “manifestation,” which implies physicality (manus being the Latin root for “hand”). Instantiation is a simpler term, used to signal a place in a sequence in time, but without these other implications of intellectual or physical detail. The term frees us to describe sets of multiply realized phenomena at an abstract level.

Bibliographically, “instantiation,” is the phenomenon addressed within the realm of bibliographic relationships by research into bibliographic ‘works,’ and more recently ‘content genealogy’ of artifactual representations. Specifically, an instantiation of a work exists whenever the work is realized in time (such as a performance or a reading), or when it is manifest in physical form (in a book, for example). A problem arises when multiple instantiations of a work (several editions, translations, etc.) exist and must be collocated in a retrieval system with sufficient information to assist in the selection of the instantiation of interest to a searcher. Similarly, unique artifacts can be represented by metadata or images (called representations), which can exist in multiple instantiations (a photographic negative, a print, its digital descendent, etc.). The same is true of the representations of archival documents, which might exist in paper photocopies, digital images, and so forth.
The *Oxford English Dictionary Online* defines instantiation as “the action or fact of instantiating: representation by an instance.” Both applications are useful for understanding the role of instantiation in knowledge representation. An instantiation exists empirically as a representative of an information object, which representation takes shape at a specific point in time as the result of an action. An instant is “an infinitely short space of time; a point of time; a moment.” That is, the locus of an instantiation is found along a temporal trajectory, ceding the possibility that a potential plethora of instantiations might emerge. Adverbially we understand an instant as a “pressing, urgent, or imminent happening at a time defined, succeeding or coming without interval.” In fact, it might even be something as trite as something hurriedly prepared for immediate use” (viz. a photocopy). This is borne out as well by an obsolete usage of “instant” to mean “that which is present, or is,” and the contemporary usage of “instance” to mean “an example, a case, a type, an indication, a cit[ation].”

According to the *Routledge Encyclopedia of Philosophy* instantiation takes the form of evidence:

- **Existential:** ‘o has P’—‘There is at least one [object] that has P’
- **Universal:** rule permits conclusion that any object o has the property P from the premise that everything has P
- **Inference:** signifying an argument or a step in an argument, or the process of passing from belief in the premises to belief in the conclusion
- **Quantifier:** syncategorematic operators such as ‘all,’ ‘some,’ ‘none.’

These epistemological usages enrich our definition by providing parameters around the phenomenon of instantiation, such that the instantiation is possessing the properties that define the object (its essence) and the set of instantiations of an object can be defined empirically by the truth or untruth of ‘is a’ relationships. Instantiation proceeds from an object—whether abstract in nature, like a bibliographic work, or concrete, like an extant artifact—which object provides a historicist anchor in its identity. These identity anchors constitute nodes in information retrieval systems.

The potential for instantiation is always present, but the action of some catalytic influence results in the real presence of what we call “an instantiation,” such that I₁ is a realization in time—an instantiation—of information object O₁ and the potential set of instantiations of that object might run from I₁–n. The catalytic influence, empirical evidence of which will be discussed below, lends the breathless quality of urgency to the creation of instantiations. That is they are not haphazard but rather they result from some demand close to the object of origin and at least hypothetically occur at more widely spaced points over time as the trajectory recedes temporally.

### 1.2 Instantiation in Information Science, Knowledge Organization, Semiotics

The term “instantiation” is used widely in the field of information science as well as in the domain of knowledge organization, in which its usage is related to a recent flurry of embrace of semiotic theory. To gain an understanding of the usage of the term in the field searches were conducted on Google Scholar™ and *Library and Information Science Abstracts* (*LISA*). Results were roughly comparable in interesting ways, and in all cases the term was used to mean (as I noted above) the concrete realization of something abstract, with a potential for multiple iterations.
The Google Scholar™ search passed through several iterations until eventually seven result sets were constituted. These ranged from 53,400 hits for “instantiation” alone, through various combinations of domain-specific keywords, to 340 hits for “instantiation semiotics.” The first ten results in each set were subjected to content analysis. Domains encountered included:

artificial intelligence, cognitive anthropology, computer science, engineering, experimental psychology, medicine, neuropsychology, semiotics, and systems analysis,

and within information science

bibliometrics, digital libraries, information systems, knowledge management, knowledge organization, music information retrieval, natural language processing, and ontology.

In a small number of cases (mostly my own work) the term was used in reference to sets of documentary entities. In a much larger number of cases the term was used to mean the coming to fruition of a system design. A more generic usage was also present, giving the sense of concretization of some abstract concept.

The search in LISA was, of course, much more focused. I looked for all hits on “instantiat*” and “instantiation.” The eventual result set had 70 records, representing publications in journals and proceedings ranging from 1991 to 2004. Here the result was more recognizably constitutive of information science as I recognize it. The following publication venues were represented:

AI magazine
Ariadne
Artificial intelligence
Artificial intelligence in medicine
Canadian Journal of Information and Library Science
Cataloging & classification quarterly
Computer communications
Computer networks
Computer networks and ISDN systems
Connection science
D-Lib magazine
Human computer interaction
Hypermedia
IEEE expert
Information and software technology
Information processing and management
Information systems
Information technology and libraries
Information technology and people
Interacting with computers
International journal of human-computer studies
Journal of communication
Journal of educational multimedia and hypermedia
Journal of the American Society for Information Science
Journal of the American Society for Information Science and Technology
Knowledge organization
Knowledge-based systems
Usage was similar to what was found in Google but more focused. In this set two usages predominated. First was the systems meaning, in which the term “instantiation” is used to mean the coming to fruition of a system design. Typical phrases were "system, which instantiates the architecture for name pronunciation," "novel instantiation of a digital photo library," and "we present the instantiation of our model for a real city." The second usage was the documentary usage, including most of my own publications about the “works” phenomenon, but also (notably) papers by Turner and Goodrum about videos, and descriptions of Tufts University’s Perseus Digital Library.

Thus in information science, ‘instantiation’ is conceptually a part of any theory of object representation, predominantly those of “the work,” “the artifact,” or ‘the document,” and also “the system.” The concept describes that point at which a given realization takes form, and, in the presence of an appropriate catalyst, around which a network of representatives might cluster. Such clusters has been called “instantiation networks.” Empirical evidence from several studies have demonstrated their existence and some of their parameters. Any theory of “instantiation” must begin with this empirical evidence.

2. Empirical Evidence

Studies to date have demonstrated some consistent theoretical parameters for the concept of instantiation, even across bibliographic and artifactual borders. Empirical evidence from several studies of bibliographic works (reported by Smiraglia 2001) was subjected to preliminary meta-analysis in Smiraglia (2002a). Three conclusions about documentary instantiation were drawn in that paper:

1. The majority of works exist in only one instantiation, but substantial proportions (associated in quantitative terms loosely with Lotka’s law) generate instantiation networks through mutation and derivation over instantiations over time.

Works are constituted by both ideational content (the ideas they convey) and semantic content (the strings, usually linguistic, by which they are conveyed), both of which are unique to the original iteration of the work (see Smiraglia 2001). Instantiations have been observed to be of two types: mutation and derivation (see Smiraglia 2002b). Instantiations produced by mutation are those in which semantic or ideational content have been altered. These are often described as translations, adaptations, performances, etc. A derivation is an instantiation, subsequent to the first, in which no semantic or ideational content has changed, as is assumed (perhaps unwisely) to be the case with editions, both simultaneous and subsequent. In fact, preliminary results suggested that approximately two-thirds of works have extensive networks of instantiations; the others are published in multiples (two or three simultaneous editions, for instance) but no new
instantiations appear over time. A tentative conclusion has been that some catalyst, probably a cultural influence, plays a role in generating instantiation networks. In the bibliographic realm the catalyst appears to be something called “canonicity”—acceptance of the work into a canon of literature creates demand for new and evolving instantiations. Furthermore, the demand is culturally sensitive and this can be observed empirically in the data sets.

2. Simultaneous and successive editions, and translations, predominate among the types of mutation and derivation observed. Works that are likely eventually to be at the center of large instantiation networks are published simultaneously. Evolution of the network begins with successive editions and translations. Other types of instantiation are much less prevalent. And,

3. Older progenitors are associated with the largest instantiation networks. Regression coefficients demonstrate a consistent but relatively weak growth rate among instantiation networks such that, for evolving networks, the longer the time-span the greater the number of instantiations will appear.

A more recent study (Smiraglia 2004a and 2004b) used a case study of Etruscan artifacts at The University of Pennsylvania Museum of Archeology and Anthropology. Attempting to discover whether the “works” metaphor, and therefore also the concept of instantiation, might be applicable beyond the world of bibliographic works, a set of eight artifacts were selected and all in-house and external representations were discovered. Analysis of the result sets demonstrated the efficacy of the metaphor—artifacts can constitute sets of representations, from which subsequent instantiation networks develop. In fact, all of the artifacts had in-house representations that were produced at a predictable rate based on internal museum functions. Each artifact was represented according to the policies of each division of the museum as it passed from acquisition to exhibition by way of conservation, loan, and other units. The original representations became centers of instantiation networks as copies of the representations (everything from photographs to three-dimensional models) were themselves copied, to be shared among divisions. Every representation was instantiated for the museum archives. Divergence in the instantiation networks, both their extent in number and the diversity of instantiation, appeared to be a function of the recognized popularity of the artifact. Specifically, artifacts that had photographs in print were more likely to attract public attention. This, in turn, was likely to generate more in-house instantiation of representations as well as external, published, instantiations.

Here we see a harmonization of the “works” metaphor from the bibliographic (or documentary) domain with a “representation” metaphor from the artifactual domain. Instantiation—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 a kinds of “popularity.” At a meta-level, cultural acceptance of a work or an artifact creates public demand for more representations.

These results are necessarily preliminary. It is clear that the phenomenon of instantiation has consistent elements, including: 1) the universality of instantiation in all domains of knowledge objects; 2) the concept of cultural catalyst as a predictor of instantiation; and, 3) the influence of time as a predictor of the degree of instantiation. For every information object $O_i$ the possibility of instantiation is present. The historic action of some catalytic influence results in the presence of an instantiation $I_i$. The simplest set, then, is the set
This is representative, for example, of an edition of a book, a photocopy of a document, and a photograph of an artifact. A set of instantiations—an instantiation network—might run from $I_{i-n}$. A common set, then is the set

$$O_i \rightarrow I_i$$

This is representative of a set in which, for example, many editions of a book appear over time, or many photocopies of a document are produced, or photographs of an artifact are produced repeatedly from an original negative image. The catalytic influence, which I said lends a breathless quality of urgency to the creation of instantiations, is a culturally sensitive phenomenon. That is they are not haphazard but rather they result from some demand close to the object of origin and at least hypothetically occur at more widely spaced points over time as the trajectory recedes temporally. Take for example a work that is popular in the time of its origin that appears in multiple editions, then recedes for centuries until a cultural shift “rediscover” it, and generates new sets of instantiations. In this case we have instantiation nodes within the set, from which instantiation networks can also proceed. Such a set looks like this

$$O_i \rightarrow I_{i-n} I_{i+n} I_{i+n}$$

Where each $I_i$ is a new instantiation node. At each node $n$ is the sum of the $i$s (hence the extent of the instantiation network), and $I_i$ is the representative of a node of instantiation. For knowledge organization it is important to have empirical evidence of the extent and depth of instantiation networks. The tenets of knowledge organization apply here, for there is no single, uniform, universal description of instantiation such that a book represents a work. Rather, the eventual constitution of an instantiation network is domain- and culturally-sensitive. History, past and future, determines the extent and depth of the network and the relationships, semantic and ideational, among the nodes. For information retrieval it is important to grasp the centrality of the historic nodes of instantiation within each set.

3. Conclusion: Lacunae in the theory of Instantiation

All of which brings me to the lacunae in this nascent theory of instantiation. We can group the lacunae into two sets—the empirical evidence of instantiation not yet gathered, and the pragmatic misinterpretation of instantiation in the FRBR model (IFLA 1998). The problem with the FRBR model is simply stated: FRBR refers to instantiation as “expression” and “manifestation,” and sequences them temporally, such that an expression must exist before its manifestation. Logically, this is the same action represented here as instantiation of an information object. And, empirically we see that there are two types of instantiation that are not particularly temporally distinct: mutation and derivation.

But our evidence is not yet sufficient to state fully a theory of instantiation. What we see here is that instantiation always refers to the concretization or realization of an abstraction. In our domain the result is an iteration of an information object. We have enough empirical evidence of the phenomenon in bibliographical contexts to project
hypotheses about the size and complexity of instantiation networks and the rate of their
growth over time. We need better analysis of the sorts of transitive relations that can be
observed at the nodes of instantiation—how exactly can we understand and model
changes in semantic and ideational content? Our evidence from the artifactual domain is
scant—better evidence is needed from the museum domain. And to date there is no
evidence from the archival domain, although the logical application of the phenomenon
of instantiation to the problem of the evolution of documentary content across physical
formats (say from paper to microform to digital form) is possible.

4. Works Cited

International Federation of Libray Associations. 1998. Functional Requirements for

Oxford English Dictionary online [OED online]. http://www.oed.com/ accessed April 9,
2005.

2005.

Smiraglia, Richard P. 2001. The nature of a work: implications for the organization of
knowledge. Lanham, MD: Scarecrow.

Canadian journal of information and library science 26 n2/3: 30-49.

Smiraglia, Richard P. 2002b. “Works as signs, symbols, and canons: the epistemology of