

## D-Lib Magazine June 1998

ISSN 1082-9873

# Authenticity of Digital Resources

## Towards a Statement of Requirements in the Research Process

---

David Bearman and Jennifer Trant

Archives & Museum Informatics

Pittsburgh, Pennsylvania USA

[www.archimuse.com](http://www.archimuse.com) [dbear@archimuse.com](mailto:dbear@archimuse.com), [jtrant@archimuse.com](mailto:jtrant@archimuse.com)

---

## Contents

### [I. Introduction](#)

### [II. Asserting and Assessing Authenticity](#)

### [III. Why Is Authenticity an Issue Now?](#)

### [IV. Types of Solutions Proposed](#)

### [V. Requirements for Authenticity and the Research Process](#)

#### [a. Information discovery](#)

#### [b. Information retrieval](#)

#### [c. Information Use: Collation, Analysis and Representation](#)

### [VI. Choices, Systems and Authenticity](#)

### [Notes](#)

## I. Introduction

Interpretation, and re-interpretation, of primary and secondary sources is the foundation of much humanistic scholarship. Construction of a convincing argument depends on an evaluation of the authenticity of source materials. Judgments about authenticity are based on assessments of the origins, completeness and internal integrity of a document. They may also draw from the

consistency and coherence that exists between a particular source and others in the same context or of the same type.

Traditionally, when scholars encountered original sources (artifacts, documents or works of art) many physical clues assisted in establishing their authenticity. If sources are studied in the surrogate (in photographs, microfilm or other reprographic form), all the questions concerning the authenticity of the original are overlaid with additional questions about the methods of representation. When the source is itself is distant from the original, such as a transcription of selections from a census or a collation of statistics, the methods employed and choices made in its creation are assessed by scholars along with the interpreted results; these methods impinge upon its subsequent utility for other kinds of analysis and interpretation.

Concerns about authenticity in sources, then, are not new. With the ubiquity of digital representations and the proliferation of source information on the Internet, these issues are further complicated. It is not unusual to find resources purporting to represent the same thing at many different sites. But we are without widely understood and employed methods of assessing and establishing authenticity of digital sources. It is crucial for the ongoing development of knowledge in many fields that it be possible to determine the relative authenticity of a number of different representations, through an analysis of the methods that have been chosen to transform the original into a digital form, or through an assessment of the methods used to capture original digital data. It is also essential that it be possible to establish the integrity of a particular digital copy.

The community of users and creators of scholarly digital documentation is not without proposed solutions to this problem. Many companies offer technologies, and a variety of institutions have established or proposed social mechanisms of control which make promises to satisfy some part of the requirements for digital authenticity. We feel the appropriateness of these solutions cannot be assessed until we are clearer about its needs in this area. Further articulation of requirements is a necessary precursor to systematic assessment of the degree to which different types of technical and social offerings satisfy scholarly needs. This note calls for both such further definition of requirements and the associated assessment of mechanisms being offered in order to hasten the development of trusted and widely adopted solutions.

## **II. Asserting and Assessing Authenticity**

At its extremes, authenticity carries with it all the philosophical problems of truth, but here we will try to confine the assertion that something is "authentic" to a number of more "provable" claims: that it is unaltered from the original; that it is what it purports to be; and/or that its representation is transparent (the rules are stated and, possibly, reversible). Establishing the full societal warrant for

authenticity will require both searches through relevant literature and involvement of specialists in many scholarly domains. A parallel process with a very similar purpose was pursued in definition of the societal warrant for "evidence" as part of an NHPRC funded research project on electronic evidence at the University of Pittsburgh.<sup>i</sup> Related questions have been asked about what makes a transformation authoritative in studies by Anne Kenney and others of the criteria for digital preservation quality.<sup>ii</sup>

This discussion of authenticity and integrity of information resources should not be confused with discussions of authentication,<sup>iii</sup> an equally important issue in networked information access. Authentication (or access control) addresses who has permission to do what with an information resource and is, therefore, grounded in methods of identification of the user. Authenticity (or content assurance) addresses what the resource purports to be and how it was created and is, therefore, grounded in methods of identification and verification of the resource.

Many cultural agencies are making substantial investments in transforming heritage information into digital form.<sup>iv</sup> For the scholars who study these works, the ability to distinguish an "authoritative" digital representation, from that made by someone who might have paid less attention to the requirements or impact of digital transformation, is important. This assessment is complicated, as knowledge of digital representation and modeling methods and access to appropriate capture technologies are not necessarily resident with those who have access to the original documents themselves. While a collecting institution may always have an apparent advantage in having access to "the real thing", a scholar may have far greater understanding of the content and context of a particular artifact, and/or greater sophistication about the implications of specific approaches to *re-presentation*.<sup>v</sup> Faithfulness in representation, a self-conscious awareness of the issues in digital conversion, and assurances of the authenticity of available resources will greatly influence the decision of a scholar to use a particular digital document or dataset.

Our own interest in authenticity, for instance, is partially related to trying to satisfy the requirements of art museums which are building a digital library of art for scholarly study and teaching.<sup>vi</sup> How will researchers locating digital representations of works of art on the Internet be able to distinguish between the reproductions they find and determine their authenticity and fidelity to the original work of art?<sup>vii</sup> Where will they find the self-conscious documentation by the creators of the representation which details the methods employed in making the representations? How will future researchers develop the "digital literacy" needed to assess digital documents, identify known artifacts introduced by particular processes, and correctly identify as yet unknown sources of distortion?<sup>viii</sup>

### **III. Why Is Authenticity an Issue Now?**

Forgeries and fakes have long been a concern of students of culture and the humanities, but fakery has not been a major issue for most researchers in the past, both because of the technical barriers to making plausible forgeries, and because of the difficulty with which such fakes entered an authoritative information stream. This changes with the broadening of the number of sources for information, the ease with which digital information can be altered, and the likelihood of the existence of many digital "copies" of the same work with slight differences. The underlying technology makes purposeful fakery easier and more tempting. In addition, lack of understanding of the distorting effects of specific methods of digitization makes the creation and distribution of copies that were designed to be faithful, but are, in fact, misleading in some manner, very common. <sup>ix</sup> New digital works are also challenging our concept of authentic representations, particularly in realms such as photography, where digital manipulation has extended the practice of retouching photographic prints beyond cosmetic to the misleading. Our faith in "visual documentation" is changing, as we learn of nature photographers cloning zebras, and moving pyramids;<sup>x</sup> in some venues, such as *Wired*, digital manipulation is expected.

Some of the most promising characteristics of digital objects are those which reveal the greatest authenticity problems. For example, objects tied to methods are programmed to behave in particular ways. Performance of works has generated an important literature on the nature of authenticity.<sup>xi</sup> We are without a parallel assessment of when and how digital performances can be considered faithful. Automatic "performance" of a work based on its symbolic representation is enabled by digital presentation and can have some benefits -- but its limitations must be understood. For example, a speech by a nineteenth century political figure can be "read" by a computer or a quartet can be "played" by computer simulated instruments, but both will be misleading in many respects. Even more complex re-enactments or re-creations can enable visualization of realities that are otherwise hard for people to understand, but it is important that those who are experiencing such manifestations understand the rules by which they were created.<sup>xii</sup>

It is interesting to note that the last time a significant shift took place in the technologies of copying, with the invention of printing, a similar concern about authenticity gave birth to the entire discipline of Diplomatics. Diplomats are again active in the search for methods of establishing digital authenticity.<sup>xiii</sup> Quite possibly, societal responses ultimately create mechanisms that are sufficient for most general assurances of authenticity as a basis for the widespread adoption of such new technologies -- if so, we are currently crafting such mechanisms for the 21st century.

## IV. Types of Solutions Proposed

Three quite distinct technical and social strategies for asserting authenticity can be

identified in proposed solutions: public, secret, and functionally dependent. These methods emphasize different problems and, not surprisingly, satisfy different aspects of our requirements.

Public methods for asserting the authenticity of sources include:

- the creation of copyright deposit "collections of record",
- certified deposits of original sources combined with record certification services,
- registration of unique document identifiers,
- publishing "key" data about documents which, when hashed, or otherwise calculated in a publicly available way, should match that of the document in hand, and
- defining metadata structures to carry document authentication declarations or proofs. [xiv](#)

Secret methods involve hiding data in the object to reveal its source. Techniques include:

- digital watermarking
- steganography [xv](#) and
- digital signatures.

Functionally dependent methods employ specific technologies that are bound together with the information source. Methods employing technical dependencies include:

- object encapsulation (whether physical or logical) [xvi](#)
- cryptotopes (TM) [xvii](#)
- encryption and
- embedded active agents.

Each of these technologies addresses problems beyond simple resource authentication. We are not always clear about which aspects of a technology are functionally most important. For example, as creators of resources do we apply visible "watermarks" as a deterrent against misuse? Or are we really looking for the potential evidentiary value of a hidden watermark or digital signature that can be cited in the case of a misuse or infringement? Or are we simply offering a "seal of approval" indicating a digital resource came from a trusted source? As users, are we interested in a secure environment that guarantees the integrity of a retrieved resource, or are we more wedded to the philosophy of open access, and the flexibility that results from the ability to move a digital object between systems and platforms? Are we therefore willing to invest in hashing and comparative routines to establish bit-level authentication when it is a research requirement? A clear understanding of our requirements in particular circumstances is the only

means to a comparative assessment of technological solutions.[xviii](#)

## V. Requirements for Authenticity and the Research Process

Questions of authenticity and assessment arise throughout the process of scholarly research, whenever a potential scholarly user finds, retrieves, or subsequently uses, a digital resource (for example, from the Internet). Because we believe that ultimately expressions of the requirements for authenticity will need to reflect both specific disciplines and their working methods, we cite examples below from the various stages in the research process: information discovery, retrieval, and use (collation, analysis, and representation).[xix](#)

### a. Information discovery

Information discovery, the identification and location of relevant resources, most often takes place through a query process. A user searches a database or index for a particular resource or type of resources (with more or less precision). Such a query most often results in a list of citations. The user, faced with the comparative assessment of this result, asks the question "What have I found?". Current lists of search results often seem highly redundant. Disambiguation of similar resources, simply from the information presented in a query result, is difficult if not impossible.

Unique identifiers offer one solution to the problem of disambiguation. If each version or edition of a resource has a distinct identifier, the researcher is then able to request a particular source. This source could be known to meet certain authenticity criteria. Or it could be that a specific source is sought because it was the basis for a previous analysis. In either of these cases, the unambiguous identification of a particular digital information source is key. For example, unique identification by a Digital Object Identifier ([www.doi.org](http://www.doi.org)) or Serial Item and Contribution Identifier ANSI/NISO Z39.56 - 1996 (version 2), registration with a copyright agency, or assignment of a Permanent URL (purl) by a trusted source, all identify a specific source as the particular one that provided the basis for a scholarly argument.

But unique identifiers only address one aspect of authenticity in discovery. Often, however, the researcher is faced with the difficulty of choosing between different sources, as a query identifies numerous versions of the same source. How then can one assess their desirability for retrieval? What knowledge of the source is necessary in order to assess its fitness for a particular scholarly purpose?[xx](#)

In making a decision to act upon the results of information discovery (i.e., to move on to the retrieval of an identified resource), the user wants further assurances of authenticity. Given that several representations of the same resource might have

been discovered, the scholar wants to know which representation is the most comprehensive, authorized, or complete. Given the multiple digitization strategies likely to have been employed, shouldn't the researcher have explicit knowledge of the rules of the transformation in order to be alert to the potential pitfalls of a given representation method?

On some level, the 'easy' answer to assessing authenticity in a large search result may derive from knowledge of the entity making the representation available. The perceived authority of a certain repository will come from respect for the methods they employed, and the care they take in maintaining digital versions of documents. But even with a known "pedigree," explicit declaration of representation methods, citation to discussions of the underlying representation problems, declaration of the credentials of the agency making the representation, and other similar demonstrations of an understanding of the issues are necessary to build confidence in the representation.

## **b. Information retrieval**

Once a resource is discovered, and judged to be of interest, the researcher makes a choice to retrieve it, consciously moving a copy of a digital source to a local workstation for future use. The user is still left with the question "how do I know that it is, in fact, what it purports to be?". Here the authenticity of a copy of a known resource might be assured by digital verification methods such as (public) hashing algorithms, (hidden) watermarks, or (functionally dependent) unlocking encryption. Such comparisons of a specific copy to a known version of integrity can assess and/or assert the integrity of the copy. Subsequent retrievals from a local source must, of course, be similarly verified.

The retrieval process itself is software enabled. If we trust the specific retrieval system to check the received copy against what was transmitted, this post-facto comparison may seem unnecessary. But methods we have constructed for convenience, to use data constructed in one environment with tools resident in our own environment, involve transformations between representations. End-users are unlikely to be aware of the types of approximations made in these "automatic" and behind-the-scenes translations, and will encounter "the original" in their own software environment complete with any such changes that have been introduced (which may include modifications that are significant to meaning, such as are caused by carrying floating points out to only three decimal places or changing millions of colors to thousands).

## **c. Information Use: Collation, Analysis, and Representation**

At the stage of analysis and use of resources retrieved, the scholar confronts questions of authenticity that are both similar to and distinct from the traditional assessments of primary or secondary source material. The analog world gives us

some models for comparative analysis. Reprints, changes of scale, and loss of color in black and white photography may either emphasize or obscure particular characteristics in the original source.

Scholars have historically relied on metadata provided by repositories to determine the meaning of otherwise ambiguous documents, to more or less success. A recent article in *The Record*, the journal of the US National Archives, illustrates the problem vividly.<sup>xxi</sup> The author discovered that a document in the Kennedy Library which has been used to support the claim that Kennedy ordered withdrawal of US troops from Vietnam in the month before his death actually relates to withdrawing US troops from the Mississippi school integration crisis in October 1962. The metadata for this record mistakenly indicated 11:16 am October 1, 1963 but other evidence conclusively demonstrates that it was actually 11:16 am October 1, 1962. The exercise of translating Mr. Stern's sleuthing into the digital realm invites some exacting requirements definition.

Representations may also implicitly impart a particular "point of view", either consciously or sub-consciously.<sup>xxii</sup> Since digital resources are often representations of information previously created in analog format, the verisimilitude of these representations are matters of considerable scholarly dispute. Different representations amplify different aspects of the "real thing"; for example, mark-up of a text may or may not identify particular linguistic elements; images taken with different wave lengths of light will show different features of an object; and images digitized with different degrees of attention to color will show different degrees of accuracy/fidelity. The issues of authenticity in digital transformations are, properly, about how well certain representations serve a desired purpose.

When we are using a representation (and all digital resources are representations), many of the issues which figure prominently in the field of translation become critical. We do not have to have quite the self-consciousness about transformation rules and formalism that Douglas Hofstadter displays in *Le Ton Beau de Marot*<sup>xxiii</sup> to realize that an ASCII text of the work of a typographically oriented concrete poet, or a word by word accurate transcription of a Shakespearean sonnet without proper line breaks, would significantly misrepresent the original work. Nor must we have long experience with the World Wide Web to know that it is populated by many such "copies" and that even more serious misrepresentations, often less easy to identify, are also common. Hofstadter's sophisticated discussion of these issues shows that all re-presentations are accompanied by decisions to abide with greater or lesser fidelity to certain constraints of the original. In some respects, then, each re-presentation is more suited to certain purposes than another. The specific degree to which all transformations respect some formal properties of the original and the responsibility of their creators to report on their methods will determine the extent of their usefulness for specific kinds of further analysis.<sup>xxiv</sup>

Issues of authenticity have always been critical in some areas of scholarship in which a resource is going to be used in a new way. For example, if we hope that by locating original data from an earlier study we will be able to collate some of its findings with new data we have collected, there are a multitude of questions about the original data and its representation (and intermediary analysis) which must be answered before we can safely combine the two sources. The apparent, relative ease of manipulation of digital datasets makes this research strategy seem more attractive; however, the lack of readily accessible repositories of documented research resources is still a major impediment.<sup>xxv</sup>

As the amount of primary research material that exists only in digital form increases, this issue will be exacerbated. Replicating results and verifying methods and methodologies are becoming more and more difficult. The assessment of authenticity in a digital resource can only be made by the scholar in conjunction with appropriately detailed documentation of the methods of representation used in the conversion, sampling or creation of a digital resource.<sup>xxvi</sup> Indeed, one subset of these analyses are historical debates about whether appropriate statistical measures, which are a kind of abstract representation, were employed. Others might include the assessment of the documentation of archaeological finds, and judgment about the incorporation of the results of a particular study into a larger analysis.

## **VI. Choices, Systems, and Authenticity**

Authenticity is, therefore, of concern for both information providers and information users throughout the entire research process. Researchers are concerned about identifying and assessing the integrity of a particular source of interest. Agencies may wish to identify copies of digital transformations they have made. This may be to assert their original care in the creation of a representation, or if they no longer physically control copies, to be able to identify them in the future, possibly replace them subsequently methods of representation change, or to ensure their continued integrity through association with an authoritative source. The creators of authentic digital representations may also feel that their investment in high quality authoritative knowledge representation requires compensation, requiring the ability to track copies of authoritative originals as part of licensing and protection mechanisms.

Many technical methods are being developed or offered that purport to address the issues of authenticity and integrity of information resources. To determine which methods are suited for what purposes, it is critical that we better understand the functional requirements for authenticity on the part of creators and potential users of digital resources, and appreciate where in the research process these requirements come into play.

Preliminary answers to all of these questions show that we are just beginning to

understand the social, economic and philosophical baggage that travels with our sense of the authentic. Rather than searching for a single solution, we should be developing our understanding of the various requirements for authenticity. Exploring these definitions, by looking at the relationship between humanistic research methodologies and aspects of authenticity, such as validity, originality, and credibility, will help us understand our assessment of a resource as genuine, certified, accurate, trustworthy or reliable. We might then be able to judge how certain technological or social systems might enable users to readily assess the fitness of a particular digital resource for a defined research purpose.

## Notes

<sup>i</sup>"Literary Warrant Supporting the Functional Requirements" *Functional Requirements for Evidence in Recordkeeping*, School of Information Sciences, University of Pittsburgh, <[http://www.sis.pitt.edu/~nhprc/warrant\\_audits.html](http://www.sis.pitt.edu/~nhprc/warrant_audits.html)>. See also Richard J. Cox and Wendy Duff, "Warrant and the Definition of Electronic Records; Questions Arising from the Pittsburgh Project", Proceedings from the Working Meeting on Electronic Records Research, Pittsburgh, PA, May 1997, published as a special issue of *Archives and Museum Informatics*, Vol. 11, no. 3-4, 1997, pp. 223-231. [Return to Text](#)

<sup>ii</sup>See for example Anne R. Kenney and Stephen Chapman, *Digital Resolution Requirements for Replacing Text-Based Material: Methods for Benchmarking Image Quality*. Washington, DC: Commission on Preservation and Access, April 1995. and *Digitization as a Means of Preservation?* European Commission on Preservation and Access, Amsterdam, October 1997. Final report of a working group of the Deutsche Forschungsgemeinschaft (German Research Association) by Hartmut Weber and Marianne Dörr, translated by Andrew Medlicott. <<http://www.clir.org/pubs/digpres.html>>. [Return to Text.](#)

<sup>iii</sup>See Coalition for Networked Information, "A White Paper on Authentication and Access Management Issues in Cross-organizational Use of Networked Information Resources", Clifford Lynch, editor. <<http://www.cni.org/projects/authentication/authentication-wp.html>>. [Return to text.](#)

<sup>iv</sup>Many examples can be cited here. Those that also report on the decision-making processes involved include *Towards the Digital Library, The British Library's Initiatives for Access Program*. edited by Liona Carpenter, Simon Shaw and Andrew Prescott, London: The British Library; 1998, and Hoyt N. Duggan "Creating an Electronic Archive of Piers Plowman" <<http://jefferson.village.virginia.edu/piers/report94.html>>. [Return to Text.](#) *Towards the Digital Library, The British Library's Initiatives for Access Program* is [reviewed](#) in this issue of *D-Lib Magazine*.

<sup>v</sup> Andrew Prescott reports on a successful collaboration between custodian and scholar in "Constructing the Electronic Beowulf," *Towards the Digital Library*, 1998, pp. 30-49. [Return to Text.](#)

<sup>vi</sup> The Art Museum Image Consortium <<http://www.amico.net>>. [Return to Text.](#)

<sup>vii</sup> For a discussion of the challenges facing museums in this area, see J. Trant, "When All You've Got's the Real Thing: Museums and Authenticity in a Networked World", presented at Qualität und Dokumentation, CIDOC Jahrestagung 1997, revised version to be published in *Archives and Museum Informatics*, Vol. 12, no. 2, 1998. [Return to Text.](#)

<sup>viii</sup> Margaret Headstrom identified this as a critical skill of future scholars at the Working Meeting on Electronic Records Research, Pittsburgh, June 1997. [Return to Text.](#)

<sup>ix</sup> "As we move from the actual to the reconstruction... we move into what are probably the most controversial, if the most technologically exciting, aspects of the Project.... In such an environment where damaged texts can be invisibly mended and the circumstances of actual performance replicated, a decorum for signalling fakes has to be carefully observed. There is a line that can be crossed ... whereby it is possible to 'improve' on the original, by increasing the contrast between darkened vellum and faded ink, by restoring old paste-downs to their original position, or by 'tidying up' a particularly crabbed hand. Some of these techniques have no place in an archive such as this. Other aspects of image manipulation, which involve speculative reconstruction, however well-founded, must always be scrupulously and prominently recorded...." Meg Twycross, Pamela King and Andrew Prescott, "The York Doomsday Project", *Towards the Digital Library*, p. 56. [Return to Text.](#)

<sup>x</sup> Kenneth Brower, "Photography in the Age of Falsification," *The Atlantic Monthly*, Vol. 281, no. 5, May 1998, pp. 92-111. [Return to Text.](#)

<sup>xi</sup> *Authenticity in Performance: Eighteenth Century Case Studies*, Cambridge University Press, Cambridge, 1990. [Return to Text.](#)

<sup>xii</sup> The lack of such explicit declaration of rules drew strong criticism of *Virtual Archaeology* from Harrison Eiteljorg III in a review to published in *Archives and Museum Informatics: the cultural heritage informatics quarterly*, (Vol. 12, no. 2, 1998). [Return to Text.](#)

<sup>xiii</sup>Luciana Duranti, "Diplomatics: New Uses for an Old Science. Parts I- VI," *Archivaria*, 28-33 (1989-92); For a investigation of the concept of authenticity in a variety of fields, see Eun G. Park, "The Nature of Authenticity in Multidisciplinary Fields and its Implementation in Electronic Records Management," Proceedings of the 6th National Conference of BAD (Librarians, Archivists and Documentalists Association, Associação de Bibliotecários, Arquivistas e Documentalistas, May 6-8, 1998, Aveiro, Portugal). [Return to Text.](#)

<sup>xiv</sup>Eric Miller, An Introduction to the Resource Description Framework, *D-Lib Magazine*, May 1998 <<http://www.dlib.org/dlib/may98/miller/05miller.html>>. "The RDF data model additionally provides for the description of other descriptions. For instance, often it is important to assess the credibility of a particular description (e.g., 'The Library of Congress told us that John Smith is the author of Document 1')." [Return to Text.](#)

<sup>xv</sup>Thanks here to Clifford Lynch for alerting us to this general term for "hidden writing" and calling our attention to Information Hiding, Steganography & Watermarking Information Hiding - An annotated bibliography, [28 May 1998] <<http://www.cl.cam.ac.uk/users/fapp2/steganography/bibliography/>>. [Return to Text.](#)

<sup>xvi</sup>For a discussion of the possibility of implementing "logical" encapsulation within relational or other environments for long-term protection of the integrity of digital objects, see David Bearman, "Towards a Reference Model for Business Acceptable Communication," (1994) <<http://www.lis.pitt.edu/~nhprc/prog6-5.html>>, and also David Bearman and Ken Sochats, "Metadata Requirements for Evidence" (1995) <<http://www.lis.pitt.edu/~nhprc/BACartic.html>>. [Return to Text.](#)

<sup>xvii</sup>Cryptolopes were the IBM trademark for an encapsulated object for authentic commerce. The technology was transferred to Lotus Corporation in December 1997, when IBM abandoned plans to offer it as an independent server site. PC Week, December 18, 1997, <<http://www.zdnet.com/pcweek/news/1215/18encrypt.html>>. [Return to Text.](#)

<sup>xviii</sup>Margaret Hedstrom's review of 'trusted systems' technologies in terms of requirements for managing electronic records provides an example of the deconstruction of a technology's salient characteristics in light of the requirements of a particular community. See "Building Record-Keeping Systems; Archivists Are Not Alone on the Wild Frontier," *Archivaria*, The Journal of the Association of Canadian Archivists, Number 44, Fall, 1997, pp. 44-71, particularly, pp. 56-63. [Return to Text.](#)

<sup>xix</sup>This outline of the research process was first developed in conjunction with discussions of metadata for information discovery -- the stated purpose of the Dublin Core <[http://purl.oclc.org/metadata/dublin\\_core](http://purl.oclc.org/metadata/dublin_core)>. At the "Image Metadata Workshop", Sept. 24-25, 1996, the authors realized that discussion of the Dublin Core Element Set would be clarified if its role in facilitating information discovery could be placed within the context of the entire Research Process, of which discovery is just the beginning. See David Bearman and Jennifer Trant, "Beyond Simple Resource Discovery: a framework for metadata declarations of disciplinary schema to support research in heterogeneous collections" presented at the International Symposium on Information Technology in Museums - Integrated Applications, Bonn, Germany, December 1, 1997. [Return to Text.](#)

<sup>xx</sup>A recent query on listserv of the Consortium of Art and Architectural Historians <caah@pucc.princeton.edu> provides an example: "Who are the people behind the two partial postings of Vasari's Lives?" asks Adrienne Deangelis in a posting on Fri, 10 Apr 1998 19:57:04 EDT. "There is one out of (apparently) the University of Baltimore: <<http://ubmail.ubalt.edu/~pfitz/ART/REN/VASARI.HTM>> and another (apparently) from the Australian National University <<http://online.anu.edu.au/ArtHistory>> ... Only some of the Lives have been done by each institution and they each have problems: the 1550 is of limited interest for art historians, and the ANU versions seem to have been done using some sort of text-scan which has not been proof-read." In asking for the editors to identify themselves, and citing institutional warrant for the distribution of these varying electronic versions, the questioner reveals two of her criteria for judging the authenticity of a source. [Return to Text.](#)

<sup>xxi</sup>Sheldon Stern, "Evidence! Evidence! All you people talk about is evidence!", *The Record*, May 1998 pp. 12-13. [Return to Text.](#)

<sup>xxii</sup>This was graphically illustrated during a recent lecture by Claudia Lazzaro at the Carnegie Museum of Art, Pittsburgh, about Italian Renaissance garden design. There, she contrasted her "documentary photographs" of these gardens with those of early pictorialist photographers. Where Lazzaro's compositions were balanced and regimented, emphasizing the symmetry and formalisms of the spaces depicted, the pictorialists chose obscured, angled views and partial vistas, depicting the mystery and romance of the same spaces. [Return to Text.](#)

<sup>xxiii</sup>Douglas Hofstadter, *Le Ton Beau de Marot*, Basic Books, NY, 1997. [Return to Text.](#)

<sup>xxiv</sup>The Relation Working Group of the Dublin Core has defined a number of possible relations between works, some of which offer insight into this problem.

See "A User Guide For Simple Dublin Core", Draft Version 5, linked to <[http://purl.oclc.org/docs/metadata/dublin\\_core/work\\_groups.html](http://purl.oclc.org/docs/metadata/dublin_core/work_groups.html)>. [Return to Text](#).

<sup>xxv</sup>The creation of such an archive of Humanities digital data for re-use is the stated mission of the Arts and Humanities Data Service. <<http://www.ahds.ac.uk>>. [Return to Text](#).

<sup>xxvi</sup>For example, the Dublin Core Format element has been extended into a series of sub-elements used to document the many different versions of a digital image created in the context of AMICO Library. <<http://www.amico.net/docs/dataspec.final.shtml>>. [Return to Text](#).

Copyright © 1998 David Bearman and Jennifer Trant

---

[Top](#) | [Magazine](#)  
[Search](#) | [Author Index](#) | [Title Index](#) | [Monthly Issues](#)  
[Previous Story](#) | [Next Story](#)  
[Comments](#) | [E-mail the Editor](#)

---

*hdl:cnri.dlib/june98-bearman*