

The Death of Correspondence? Interpersonal Communication and Special Collections in the Digital Age

IT IS HARD TO IMAGINE the organization of manuscript collections without the category of correspondence. Few classificatory entities are more ubiquitous, whether used as a records series in itself or as a modifying facet within a series. Most personal papers, including those drawn from literary, political, and commercial contexts, contain collected letters, and so do many organizational records. Collected letters, furthermore, represent the most likely aspect of a collection to lend themselves to publication after the death of the creator. No other literary form belongs so uniquely to libraries, archives, and special collections, whose compilation and safekeeping of scattered letters are prerequisites of the genre's very existence. Yet, if correspondence is intimately familiar, it is also elusive and protean. As David Barton and Nigel Hall have written, "while most people have an intuitive idea of what counts as letter writing, it is in fact a particularly difficult text object to define; after all, almost anything can be put in the form of a letter."¹ Personal and corporate archives are filled with items that occupy this liminal space: blank postcards, form letters, endorsements, printed invitations and announcements, business memos, advertisements—a whole menagerie of text objects that may or may not be regarded as correspondence.

While it may be tempting to say that, like obscenity, we recognize correspondence when we see it, finding guiding standards to identify it is not so simple. The major descriptive standards—DACS (*Describing Archives: a Content Standard*); ISAD(g) (*General International Standard Archival Description*); RAD (Rules for Archival Description); APPM (*Archives, Personal Papers and Manuscripts*) and AACR2 (*Anglo-American Cataloguing Rules*, 2nd ed.)—say little or nothing about what is considered to be correspondence. In their guide to archival description, Margaret Proctor and Michael Cook suggest that distinct terms like "correspondence, letter, postcard, note, card, telegram, letter-book entry, letterpress copy, draft/unsent" should be employed, but the terms remain undefined.² Manuscript collections rely on abbreviations like

1. David Barton and Nigel Hall, eds., *Letter Writing as a Social Practice* (Amsterdam and Philadelphia: John Benjamins Publishing Company, 2000), 1.

2. Margaret Proctor and Michael Cook, *Manual of Archival Description*, 3rd ed. (Aldershot, U.K.: Gower, 2000), 176.

ALS (autograph letter signed), TLS (typed letter signed), ANS (autograph note signed), ADS (autograph document signed), AMS (autograph manuscript signed), AQS (autograph quotation signed), AES (autograph endorsement signed), and so on, but again nowhere is official guidance on the distinction between a letter, a note, a document, and a manuscript to be found. Furthermore, which of these texts should be classified as correspondence and which should be assigned to another category is equally indeterminate.³

In the 21st century, when most communication is and will be taking place in digital environments, these elusive aspects become ever more pronounced. As paper letters are increasingly replaced by e-mail, blogs, social networks, and other technologies, the challenge of collecting and preserving an individual's or organization's correspondence is amplified. Distinguishing between genuine epistolary exchange and the casual banter that makes up most digital communication perhaps becomes a hopeless task. The situation calls for a thoroughgoing analysis of this once indispensable category that wrestles with the real possibility of its eventual obsolescence.

Elements in a Definition of Correspondence

In a range of works by literary critics, historians, and communication scholars that deal with the subject of letters, at least four elements appear to be potential signifiers of correspondence: the use of a postal system; formal characteristics like superscription and signature; privacy; and dialogue. While each of these is characteristic of correspondence, none of them is essential. It is certainly the case that letters historically have been dependent on a postal system. Though Egyptian and Greek leaders sent written messages on legal and military topics through courier systems, only oral testimony was considered strictly trustworthy. Hence, messengers were expected to recite the terms of the missives they carried and letters never moved through a postal network independent of these human carriers.⁴ Some early letters simply authenticated the oral message carried by the messenger. The first functioning postal system emerged under the Persian Emperor Cyrus in the 6th century B.C.E., and the idea was later taken up and expanded by the Romans. At this time, however, correspondence was entirely public and political. Use of the Roman *cursus publicus* (public roads) was limited to the Emperor, his family, state officials, and members of the Church and military.⁵ To otherwise send a letter was

3. See "Abbreviations Used by Archivists," available online at www.lib.uiowa.edu/spec-coll/MSC/Abbreviations/Abbreviations.html and "Finding Aid Abbreviations," available online at <http://steely.nku.edu/fa/abbr.html> [accessed 29 October 2009].

4. This prejudice continues to affect records management in the form of hearsay law. See Donald S. Skupsky, *Legal Requirements for Information Technology Systems: Evidence, Regulation, Government and International Requirements* (Denver: Information Requirements Clearinghouse, 1997), 11.

5. Klaus Beyrer, "Brieftransport in der Frühen Neuzeit: Entwicklung und Zäsuren" in Christina Antenhofer and Mario Müller, *Brief in politischer Kommunikation vom Alten Orient bis ins 20. Jahrhundert* (Göttingen: V&R unipress, 2008) p. 171.

illegal; under the Emperor Theodosius improper use of the system for private correspondence was, in fact, punishable by death.⁶

After the fall of the Empire, the remnants of the postal network were taken over by the Roman Church and used primarily by separate cloisters and orders to inform one another of the passing of a brother. Following the economic revival that began around 1000 C.E., postal systems were used more often for commercial purposes in the free cities. Medieval states still regulated use of their posts until 1600, when a liberalization of postal practices set in. With the ensuing rise of literacy, the publication of letter-writing manuals, and the dilution of exclusive aristocratic cultures, a universal craze for correspondence began that, depending on one's opinion of e-mail, has not really abated since. Thus, from the wax tablets passed between Abelard and Heloise by messenger's pouch to the *pneumatique* of the Parisian 19th century, to the bits of data floating through the ether today, correspondence has always depended on some sort of postal system. Yet we probably would not want to say that the use of such a system is a precondition of any act of correspondence since, after all, letters can also be hand-delivered, or transmitted aquatically in glass bottles, or remain unsent.

Historically, correspondence has been associated with particular formal features that, according to Alain Boureau, evolved during the Middle Ages. While letters had been used in Christian liturgy and ritual since Paul's epistles, only in the 11th and 12th centuries were rules delineated for proper letter-writing technique. At first, in the manual of the Benedictine monk Alberic of Monte Cassino, the focus was on the use of "rhythmic prose" as a way to authenticate epistolary (in other words, clerical) authority.⁷ But with the publication of texts like Adalbert of Samaria's *Praecepta dictaminum* (ca.1120) and especially Boncampagno's *Boncampagnonus* (1215), master epistoliers focused on the proper use of sub- and superscriptions in addressing individuals from various points on the medieval hierarchy. They rediscovered the classic structure of the letter in writers like Cicero and Pliny and provided models of the key classes of epistolary art: love, condolence, recommendation, legal, business, and so on. As these typologies suggest, the letter was being secularized, and it was also becoming more private, rather than public and political, in character. In the ensuing centuries, the formalization prescribed during this earlier period has been diluted, as any reader of undergraduate e-mail can attest. While the degree of structural formality in a message still helps to identify a letter, many examples of genuine correspondence dispense with all such structure.

6. Bernhard Siegert, *Relays: Literature as an Epoch of the Postal System*, trans. Kevin Repp (Stanford: Stanford University Press, 1999), 5.

7. Alain Boureau, "The Letter-Writing Norm, a Mediaeval Invention," in *Correspondence: Models of Letter-Writing from the Middle Ages to the Nineteenth Century*, Roger Chartier, Alain Boureau, and Cécile Dauphin, trans. Christopher Woodall (Princeton: Princeton University Press, 1997), 37.

For many analysts of correspondence, the qualities of privacy and confidentiality, rather than formal structure, are paramount. This position was first articulated by Christian Gellert in his 1751 treatise on letter writing. A true letter, according to Gellert, has a distinct number of known recipients or addressees. The sender must have a clear imaginative conception of the individuals on the receiving end of the message. Public letters, in his view, are just “writing.” Bismarck’s edited and published version of the Ems Dispatch, or Zola’s *J’accuse*, share certain formal structures and characteristics with letters, but they belong to a different genre from the personal letter. Diplomatic messages, memos, letters to newspaper editors, political statements taking the form of a letter like the Magna Carta—these are “works” created by the individual or organization in question, but only sealed letters intended for a discrete recipient or recipients fit the category of correspondence.

This strict definition of correspondence, if universally adopted, would carry significant ramifications in an electronic environment, where the line between public and private is more diffuse. It is a plausible claim. However, it should be noted that there have always been some situations when private letters have had unintended and unknown recipients. For centuries, as long as literacy rates remained uneven, many letters were meant to be read aloud before families, communities, businesses, or congregations. In the 17th century:

One can easily imagine the scene in the mother’s salon in which the daughter’s letter is read aloud and passed around with pride among friends and family. Learning to write the kind of letters that could be read aloud in company was necessary, since the daughter would be expected to write them for the rest of her life.⁸

In the 19th century, Ronald Zboray says correspondents tried “to perfect their orthography and grammar because they knew they would be on public display.”⁹ Additionally, collection of letters routinely contain items intended for many eyes to see, such as “To Whom It May Concern” letters; letters to newspapers, journals, and organizations; memos; and even suicide notes.¹⁰ Even when written by poets or novelists and then printed, these are generally treated as correspondence, not published works. On the other hand, missives that bear the formal structure of private correspondence can turn out to be solicitations or advertisements. The divide between public and private seems hopelessly compromised.

8. Dena Goodman, *Becoming a Woman in the Age of Letters* (Ithaca: Cornell University Press, 2009), 81.

9. Cited in Peter Wosh, “Going Postal,” *The American Archivist* (Spring 1998), 229.

10. The recent *Letters of Allen Ginsberg*, a selection of 165 of Ginsberg’s best specimens, includes items sent to the *New York Times*, the *New York Times Book Review*, *Time*, and *Poetry* magazine. Bill Morgan, ed., *The Letters of Allen Ginsberg* (Philadelphia: Da Capo Press, 2008).

In the preface to her collection of the letters of Abelard and Heloise, Constance Mews suggests that correspondence is a literary genre that deepens relationships through the exchange of words and ideas.¹¹ By emphasizing the dialogic element in letter-exchange, she captures Gellert's distinction between works and genuine letters without insisting on the problematic divide between the public and private. The enduring value of correspondence, both to the participants and to those who later read it, lies in the accumulation of the parts of another person—words, ideas, thoughts, and experiences—through a pattern of reciprocal gift-giving. This notion has a certain appeal. It does seem that true correspondence involves reciprocal exchange, a give-and-take between two individuals or entities separated in space-time.

The dialogic criteria would allow archives to dismiss messages that, while perhaps informational in content, are insufficient to classify as true correspondence. Obvious impostors like junk and mass mail could be summarily rejected, while legitimate and consistent grounds for excluding more equivocal subgenres, such as unsent or unopened letters, memos and other official messages, and open letters to the public, would be available. Items that lack words, like blank postcards, or that do not travel in a postal system, like hand-delivered letters, would still be acknowledged as correspondence, since they represent genuine efforts to communicate. Such a definition maintains the concept of correspondence as a unique and privileged classification, a literary genre that special collections help create and disseminate. In the process they model ideal patterns of communication that society at large would do well to imitate.

Unfortunately, once again, this definition of correspondence does not stand up under scrutiny. It may be entirely idealistic to think that correspondents seek the pure, transparent, intersubjective exchange of self that Mews envisions. Vincent Kaufmann notes that some of western literature's most important writers use letters not to deepen interpersonal relationships but to create distance between themselves and their correspondents. Writing letters can be a means of enhancing the author's own freedom and self-sufficiency. Kafka went to exceptional lengths to avoid meeting his fiancée, Felice Bauer, in person, even as he prized their correspondence above everything else in his life. While they exchanged thousands of letters over the years of their courtship, they only saw each other twice. He longed each day for the post to arrive, but their ultimate union—in Palestine as he told her (as if purposely placing it somewhere out of reach)—was impossible. Flaubert used his letters to Louise Colet to ensure his own rural isolation, the feeling of apartness that enabled him to work and “marinate” in solitude. Proust wrote thousands of letters canceling or declining social invitations, often using the occasion to detail at

11. Constant J. Mews, *The Lost Love Letters of Heloise and Abelard: Perceptions of Dialogue in Twelfth-Century France*, trans. Neville Chiavaroli and Constant J. Mews (New York: St. Martin's Press, 1999).

great length the medical complaints that prevented him from leaving his room. A host of other writers corresponded in similar ways. “At the heart of the epistolary mode, there is the need to destroy relationships—a mental attempt to dissolve what is solid, to replace a correspondent with his shadow, or to bury him beneath the image one has of him.”¹² It isn’t difficult to imagine nonliterary letter writers doing something similar in their missives. We write to explore our own uniqueness and to practice our ability to conceptualize a distant human being. Correspondence is quite often monologic.

Each of the factors discussed here—the postal system, privacy, dialogue, structure, and form—appears to have something to do with correspondence. However, it is hard to develop a litmus test that will enable the differentiation of correspondence from other document types in an absolute and unambiguous way. Instances of correspondence might display a tendency toward these characteristics; perhaps there is a Weberian ideal type of letter that is structured, private, dialogic, and posted, against which actual letters can be measured. There may be additional descriptive criteria to add as well. And, ultimately, questions of appraisal, rather than arrangement and description, intrude. The decision to classify an item as correspondence is preceded by the decision to collect it, which is usually dependent on appraisal of the collection as a whole. Blank postcards and form letters in the papers of Henry James are saved as correspondence. The same items in the estate of a person of local significance might be discarded, not arranged in some other category. They aren’t works, diaries, or drafts; they are either correspondence or trash. And of course the principle of original order can eliminate the need for decision making, because, if the recipient kept an item along with other correspondence, then that gives us license to call it correspondence, too. Thus, as long as most correspondence involved paper and ink, the problem was easily finessed. Repositories make decisions on an ad hoc basis that perhaps are not logically or universally consistent but work in practice. We can recognize correspondence, even if we cannot define it.

Electronic Correspondence

In the digital realm, however, things are liable to be much different. For many, the stock of paper correspondence stored in personal archives began to dwindle in the mid-1990s. While much personal correspondence has shifted to e-mail, large numbers of messages disappear through lax personal archiving, the unexpected closure or switching of accounts, and the rise and fall of service providers. Social networking sites increasingly provide the framework for much interpersonal communication, but they are flanked by other formats, whose number seems to be endlessly expanding: blog posts and comments, Web site comments, instant messages, chat

12. Vincent Kaufmann, *Post Scripts: The Writer’s Workshop*, trans. Deborah Treisman (Cambridge, Mass.: Harvard University Press, 1994), 107.

rooms, tweets, text messages, and so on. Increasingly most “letter”-writing takes place in a digital environment and is not confined to a single technological format. As a result, the inability to define correspondence suddenly is synonymous with the inability to collect and preserve it.

In addition to the usual concerns about preserving electronic records and formats, there is a classificatory dilemma: Which of these new formats should be classified as correspondence in a digital repository? Do the criteria discussed above provide any guidance? Is there even any sense in continuing to use the term “correspondence” at all?

In his *American Archivist* article on the postal system, Peter Wosh downplays the shift to a digital age. He argues that many of the debates prompted by electronic records echo issues raised in earlier communication revolutions, particularly that of the expansion and commodification of the postal system and the rise of instant messages in the form of telegraphy. The big questions—the digital divide, the power of capital and bureaucracy, the interplay between individual and society—were all hashed out then, though they remain vexing. On an arrangement level, no major rethinking is required: “[T]he current archival angst over controlling and keeping electronic information represents much ado about some relatively minor alterations in the nation’s informational infrastructure.”¹³ Yet Wosh seems to undercut his argument by conceding that new media have prompted some big changes in personality, society, and records:

Very different notions of authority, self and communication hold sway, and capturing the complexity of the new environment...means nothing less than redefining the very nature and reexamining the very purpose of a personal manuscript.¹⁴

He insists, however, that we understand the content of the emerging formats, rather than reflexively generating classificatory schemes for them without first achieving some kind of historical perspective, as this article is attempting to do in the epistolary genre.

In an electronic context, the criteria of transmission through a postal system and formal structure are no longer meaningful. All of the new communication modes pass through the Internet and have thus been “posted” in one way or another. They all are addressed to someone or to some location in an unambiguous way. For many communicants, formal features are no longer significant either. Superscrip-

13. Wosh, “Going Postal,” 238.

14. *Ibid.*, 239.

tions and signatures have disappeared along with the material foundation of paper and ink, replaced only with an evolving set of emoticons and abbreviations.

The issues of privacy and dialogue, however, remain relevant. E-mail is still primarily private, and it undoubtedly relies on dialogue (except in cases of spam and phishing). In that respect, private correspondence still exists; and, for most people, this is the form it takes. Some messages even use “Dear...” and “Sincerely...” However, it is not entirely clear that e-mail is really a correlate to the letter. James O’Toole notes, for example:

[E]-mail communications are less formal than their counterparts, the written or typed hard-copy letter...their content and style are more fluid, more immediate, and more conversational...they are, in fact, more like speaking than writing. Thus the better parallel for e-mail, some maintain, is the telephone conversation rather than the written letter.¹⁵

William Merrill Decker thinks e-mail is a hybrid of the letter and the telephone call. “[E]-mail involves an exchange of text, but not an exchange of materials associated by a metonymy of philosophical and theological resonance with the bodies and souls of the partners in the exchange. In the absence of a material exchange there generally exists an absence of record...”¹⁶ Without a record to ensure some sense of material reality and permanence, matters of form are ignored and the communication ceases to be a letter. There is not much to distinguish it from IMs, Tweets, and other brief dialogic electronic exchanges.

Under these circumstances, the line between correspondence and conversation is so blurred it ceases to be meaningful. Considering how much of the history of the 20th century was tied up in phone conversations, and how much of previous eras in unrecorded speech, this is a highly significant development. Individuals now conduct so much communication digitally that it is like having the equivalent of Nixon and Johnson tapes for millions of individuals. Imagine having a record of Saul Bellow’s phone calls with Philip Roth, or the Kennedy brothers’ youthful banter, or the Heisenberg-Bohr meeting in Copenhagen. In years to come, such interactions will likely be conducted, at least in part, through some form of captured electronic exchange. The historical record could be broadened considerably. Yet, at the same time, to do so risks engulfing collections in a sea of trivia and blather, raising questions of appraisal—not to mention privacy—that are more daunting than ever before.

15. James M. O’Toole, “‘Commendatory Letters’: An Archival Reading of the Venerable Bede,” *The American Archivist* 61:3 (1998): 283. Such arguments obviously owe a great deal to the influence of Marshall McLuhan. See *Understanding Media: The Extensions of Man* (New York: Signet, 1964).

16. William Merrill Decker, *Epistolary Practices: Letter Writing in America before Telecommunication* (Chapel Hill: University of North Carolina Press, 1998), 236.

Other new, emerging technologies straddle categories like public and private, correspondence and conversation, as well and potentially represent either a dearth or a surfeit of biographical information, depending on how well they are retained. Blogs are intentionally public and to some represent a public version of the diary. Catherine O'Sullivan says of the public-private divide, "many believe that this is an artificial and unnecessary dichotomy . . .," stating that "[blogs] are personal works with an undefined audience."¹⁷ But if the distinction between public and private is compromised—and assuming that the comment feature is an elemental part of the blog—then they may share more as a document type with correspondence. Dialogue and exchange are fundamental aspects of blogs and other epistolary writing, which is not the case with a diary. Bloggers generally are seeking communication in the form of comments. A person who might once have written a letter to a distant friend today will post a comment on a blog as an alternative. An uncommented-on blog is equivalent to an unanswered letter, not a private diary entry. Furthermore, a blog post that links to and comments on an external blog post is a form of epistolarity, too. Even if we follow Kaufmann and see letters as a way to keep interlocutors at a distance, their participation in the communicative act sets the blog apart from the diary: diaries can have audiences, but they do not have interlocutors.

Social networking sites may become the preferred location for interpersonal communication, if they haven't already. Here again it is difficult to use privacy as a privileged criterion in evaluating messages. While some messages continue to be sent to a single recipient, more common would appear to be the exchange that is freely observable to a wide network of acquaintances. Other exchanges are interpersonal in nature but can be publicly observed. In a sense, none of this is new: a dialogue carried out on a Facebook wall or a Twitter update is equivalent to a private letter that is passed around and read aloud among a family or congregation. The same exchange in a more public forum, like MySpace, is more like an open letter.

What is new is that budding young writers, politicians, artists, entrepreneurs, and activists are communicating in this manner, and there is no way at present to accumulate a corpus of relevant, meaningful communication from and to an individual that parallels the print-world series of correspondence. Anyone who has read biographies that were limited in their authors' use of the subject's letters (Ian Hamilton's notorious *In Search of J.D. Salinger*, for example, or any number of works on officially proscribed Soviet writers, many of whom in any case made a point of not writing letters) knows what a blow this potentially deals to the historical record. Despite the proliferation of electronically captured conversation, there is as likely to be a dearth of information as a surfeit. How is a scholar to gain access to decades of a person's

17. Catherine O'Sullivan, "Diaries, On-Line Diaries, and the Future Loss to Archives; or, Blogs, and the Blogging Bloggers Who Blog Them," *The American Archivist* 68 (2005), 63.

digital correspondence scattered over a host of Web sites and service providers, especially when many of those communications may have been made anonymously? How, once such collections are established, does one cut through the ephemera of daily life to find the specific materials appropriate to a researcher's interests? There is also the issue of long-term preservation. Since many of these epistolary statements are made on proprietary systems and stored at isolated server farms, neither creator (and copyright owner) nor a collecting repository has control over them. Those who post materials on commercial sites that later go bankrupt may find their reflections lost forever, unless they have maintained their own backup copies.

Relationships, Metadata, and the Semantic Web

The key to preserving an individual's correspondence may depend on the ability to establish relationships, ensconced in metadata, between far-flung digital objects. Emerging metadata, markup, and information modeling schemes may enable folksonomic or automated coding that allows for the collocation of all extant digital communications that fit the criteria of correspondence (criteria that would need to be rigorously defined). Advances in semantic digital libraries and the semantic Web in particular allow for the association of content from a single creator to be linked together. Ontologies will precisely delineate the category of correspondence, while drawing together all of an individual's addressed writings from a variety of resources, including blogs, social networking sites, news forums, and e-mail.

As the proposed basis of the semantic Web, RDF (Resource Description Framework) is the data model most likely to enable automated identification of this corpus of documents and messages. RDF envisions Web resources as a giant graph composed of object-attribute-value (or subject-predicate-object) triples that link distant resources together and enable the construction of logical statements about those resources that can be parsed and combined to generate further meaning. Ontologies, such as the Web Ontology Language (OWL) or the hundreds of others that have been developed, represent ways that multiple resources can be related in a single domain model. In these terms, it is easy to imagine the construction of a graph linking together every statement, exchange, or message attributed to a single individual or organization that would constitute the totality of that entity's meaningful interpersonal Web communication.


Unlike XML, the structure of RDF is flat rather than hierarchical. In the node-arc-node pattern that constitutes the data model, triples are never nested within one another. Because they are always separated, further instances of a related statement or exchange between two entities can be iterated together with other examples of the same to an infinite degree. Some have wondered if this structure makes RDF impractical for the compilation of catalogs that would quickly retrieve,

for instance, all existing works of a single author (or all instances of a person's correspondence).¹⁸ Yet RDF/OWL ontologies have been used successfully in a variety of digital humanities contexts; and, in theory, there are a number of ways in which the plethora of independent Web statements and exchanges made by an individual might be tied together in RDF/OWL as a collection featuring that person's correspondence, whether drawn from e-mails, social networking page dialogues, blog posts and comments, or other frameworks.¹⁹ At that point, a search query—in SPARQL or some other query language—would enable the collation of an individual's correspondence. Depending on the ontology used to define the category, this might broadly include all digital conversation, or it might be more narrowly focused on substantive exchanges. RDF/OWL thus acts as the basis of Web 3.0 where “social semantic spaces become deeply embedded in the day-to-day life of everyone on the [W]eb.”²⁰

18. Martha M. Yee, “Can Bibliographic Data Be Put Directly onto the Semantic Web?” *Information Technology and Libraries*, vol. 28, no.2 (June 2009): 57.

19. See, for example, Arianna Ciula, Paul Spence, and José Miguel Vieira, “Expressing Complex Associations in Medieval Historical Documents: The Henry III Fine Rolls Project,” in *Literary and Linguistic Computing*, vol. 23 (2008): 311–25.


20. Sebastian Ryszard Kruk and Bill McDaniel, eds., *Semantic Digital Libraries* (Berlin: Springer, 2009), 5.



TAVISTOCK BOOKS


**First Editions,
Rare & Collectible Books
in Most Subject Areas**

WWW.TAVBOOKS.COM



ANTHONY
BOOKSELLERS
ASSOCIATION of AMERICA

SECURE • SEARCHABLE



UNIVERSITY OF CALIFORNIA

1503 WEBSTER STREET • ALAMEDA, CALIFORNIA 94501
T: 510.814.0480 • F: 510.814.0486 • vjz@tavbooks.com

For the semantic Web to function in this way, bibliographic and metadata schemes would have to evolve to help provide meaningful connections among digital objects. Existing systems do not pay a lot of attention to categories like correspondence or letters. FRBR (Functional Requirements for Bibliographic Records), for instance, is based on the entity-relationship model that is one foundation of RDF, linking entities together through relationships that can be systematically recorded in a database.²¹ In FRBR's terms, based on Barbara Tillet's classification system, individual letters from the same author would carry a "shared-characteristic relationship" to each other based on the fact that they share a creator.²² Unfortunately, letters in manuscript form have never been treated as bibliographic objects—a "work," FRBR's basic and highest abstract category, generally refers to a physically or digitally published item, like a book or a Web site. By contrast, the TEI (Text Encoding Initiative), the W3C's standard for marking up literary and other texts in an electronic environment, makes substantial allowance for the letter as a text form, but only as traditionally defined. *TEI:P5*, the latest iteration of the standard, contains a whole vocabulary for the analysis and markup of letters, among other texts, capturing formal structure in the form of super- and subscriptions, dates, and postscripts.²³ This makes more sense in the context of digitized medieval letters than in today's formless electronic messages.

The metadata scheme that focuses most directly on relationships between digital objects is PREMIS: Preservation Metadata: Implementation Strategies. The five major entities of PREMIS—intellectual entities, objects, events, agents, and rights—all are linked by a whole series of semantic components, whose implementation on a broad scale would be hugely beneficial when it comes to keeping track of an individual's electronically published work.²⁴ Dublin Core (DC) metadata is also designed for Web resources. However, it does not exploit the entity-relationship structure like FRBR does; as a result, it is limited in its ability to collate linked data. While, in its unqualified form, DC is simple enough to be applied by creators, it is primarily designed for information discovery and retrieval rather than mapping the interrelationships between linked data. Furthermore, none of the DC fields enables an easy identification of content-oriented descriptors like letters. The closest that qualified DC comes to this is the "text" term within the DCMI controlled vocabulary for describing resource types in the "type" field, of which "examples

21. In fact, there is a RDF schema for FRBR: "Expression of Core FRBR Concepts in RDF," Davis and Newman (2005). Available online at <http://vocab.org/frbr/core.html>. [Accessed 30 August 2010].

22. Robert L. Maxwell, *FRBR: A Guide for the Perplexed* (Chicago: American Library Association, 2005), 100. See also Carol A. Bean and Rebecca Green, eds., *Relationships in the Organization of Knowledge* (Dordrecht and Boston: Kluwer, 2001).

23. *TEI:P5: Guidelines for Electronic Text Encoding and Interchange*, sec. 4.2. Available online at www.tei-c.org/release/doc/tei-p5-doc/en/Guidelines.pdf. [Accessed 11 December 2009].

24. *PREMIS Data Dictionary for Preservation Metadata*, v. 2.0. Available online at www.loc.gov/standards/premis/v2/premis-2-0.pdf. [Accessed 30 August 2010].

include books, letters, dissertations, poems, newspapers, articles, archives of mailing lists.”²⁵ In MODS (Metadata Object Description Schema), on the other hand, the combination of the “typeOfResource” and genre elements enables use of MARC21 Genre Terms, which do include “letter.”²⁶ Refinement of DC’s controlled vocabulary for “type” might enable better identification of instances of electronic “letters.”

Further challenges to mapping the universe of digital correspondence lie in privacy and copyright. As the W3C guidelines state, “[T]he capability to merge data from multiple sources, combined with the inferential power of OWL, does have potential for abuse. Users of OWL should be alert to the potential privacy implications.”²⁷ A person or organization may not wish to have information from far-flung sources be compiled in this manner and disseminated. Many people post messages anonymously for a reason. One of the chief objections to the use of semantic Web technologies to track and collate correspondence is the poten-

25. DCMI Type Vocabulary, Available online at <http://dublincore.org/documents/dcmi-type-vocabulary/>. [Accessed 30 August 2010].

26. Source Codes for Vocabularies, Rules, and Schemes. Available online at www.loc.gov/marc/sourcecode/genre/genelist.html. [Accessed 30 August 2010].

27. OWL Web Ontology Language Guide, sec. 2.3 “Data Aggregation and Privacy.” Available online at www.w3.org/TR/owl-guide/. [Accessed 16 December 2009].

CALADEX

LOGISTICS

PERSONALISED
TRANSPORTATION
MANAGEMENT

Caladex LLC is a leading provider of convenient and cost-effective transportation solutions ideally suited to meet the packing and shipping needs of fine book and map sellers, university libraries, auction houses and small galleries.

Caladex offers specific packing solutions for:


- Short term or long term archival storage*
- International and domestic shipping*
- Ocean (sea) freight including crating*

These unique services are available from any point in the USA and most international locations.

Our team of knowledgeable specialists brings years of experience to each assignment. This allows us to provide the service appropriate to your requirements with the utmost care and attention to detail.

www.caladex.com

Tel: (800) 643 4315
Fax: (718) 233 4159



tial intrusion on privacy from a variety of agencies. Clearly, for the semantic Web to operate, individuals would have to be in control of the nexus of links to their own work, and institutions would need to seek permission to preserve or disseminate the work or links to it. As the authors of a semantic Web textbook note, “The Web will only achieve its full potential when users have trust in its operations (security) and in the quality of information provided.”²⁸ The semantic Web will also require the ability to control who may view particular pieces of information.

Control over resources would include copyright as well. In a sense, copyright is the new privacy. In the context of the Web, it is the most powerful force governing issues of access and preservation. Individuals posting materials ought to be able to embed the copyright status of each message directly into the metadata, drawing the kinds of relationships between digital objects and rights that PREMIS envisions. They may choose to copyright more substantial communiqués that include personal thoughts, information, or expressions, while allowing more casual statements to float unimpeded through cyberspace. Using copyright in this way might provide particularly useful classificatory divisions for repositories seeking to preserve electronic documents. This would have the additional benefit of clearing up in advance significant copyright issues, which are bound otherwise to become increasingly murky, as biographers attempt to draw on twenty-year-old e-mails, blog posts, and Facebook status updates to do their work. According to current law, of course, the author should retain copyright over each of these formats even though the physical data is stored in proprietary databases. But in distinguishing levels of privacy or use rights by recipients or the general public, much future confusion could be avoided and problems of access and preservation would be mitigated.

Conclusion

We have been sending letters to each other for hundreds of years (and important people have been doing so for thousands). Everyone who has lived through the shift from “snail” to e-mail recognizes the changes it has wrought in communication style. More generally, the concept of privacy itself has been reconfigured—or perhaps simply assaulted and driven off into more and more confined areas. As a result, there is a potential treasure-trove of information out there waiting to be collected, offering new insights into the childhood, youth, social lives, work habits, and fundamental beliefs of individuals whose collected papers might one day be targeted by a repository.

28. Grigoris Antoniou and Frank van Harmelen, *Semantic Web Primer*, 2nd ed. (Cambridge, Mass.: The MIT Press, 2008), 19.

But the proliferation of digital correspondence comes with attendant dangers. So much interpersonal conversation could be preserved that none of it would be meaningful. Or so little might be effectively saved that the historical record would suffer enormously. With so much communication appearing in digital form, it appears inevitable that ontologies will be built enabling easier search and retrieval of authors' products on the Web as the semantic Web introduces a new era of cooperative information systems. The question is what role special collections departments will play in these developments. Although it is often suggested in other areas—from technology to legislation—that librarians and archivists need to be involved in a hands-on way, perhaps nowhere is that truer than in the evolution of the semantic Web. Unless issues of historical and technical preservation, privacy, and copyright are handled in a manner that meets the standards of the profession, the potential for using digital correspondence may be squandered. Any future ontology governing digital communication must take into account the history of correspondence as a genre and be able to define it in a way that makes sense in a digital environment.

**Antiquarian Booksellers' Association of America
2010/2011 Book Fair Calendar**

The Boston International Antiquarian Book Fair

November 12-14, 2010
at the Hynes Convention Center
More info at bostonbookfair.com

The California International Antiquarian Book Fair

February 11-13, 2011
at the Concourse Exhibition Center in San Francisco
More info at sfbookfair.com

The New York Antiquarian Book Fair

April 7-10, 2011
at the Park Avenue Armory
More info at nybookfair.com



Shop online 24/7, visit abaa.org!

The Antiquarian Booksellers' Association of America, Inc.
20 West 44th Street, Suite 507
New York, NY 10036
212.944.8291
hq@abaa.org

