

‘Nothing is the same as something else’: significant properties and notions of identity and originality

Introduction

This paper addresses some issues of identity (or ‘sameness’) in the context of records and archives. What does it mean to claim that one record, one archival object, is identical to another? Is it ever meaningful to make such a claim? What degree of trust can we place in claims of this kind? In our professional literature, discussion of the identity of records or other digital objects often arises in connection with digital preservation. If we have to migrate records to ensure their long-term accessibility, how far, if at all, can we trust that the migrated record is the same as its predecessor? What are we to make of the traditional concepts of *original* and *copy* in the digital era? This paper seeks to explore these questions and to relate them to understandings that records are of interest to diverse user communities.

The paper opens with a discussion of notions of identity, and introduces a connection with the commonly held view that some characteristics of a record may have particular significance. It then reviews various perspectives on significance and identity adopted in studies of digital preservation, imaging, transcription and editing, and suggests that attempts to determine the significant characteristics of records are problematical, not least because judgements about significance will vary from one user community to another. The final sections of the paper consider relations between the concepts of identity and originality, and offer some tentative conclusions about the implications of acts of ‘copying’ and the tensions that inevitably arise when decisions have to be made about the migration, conversion or transcription of records.

Identity and significance

Very broadly, there are two ways in which we conceive of identity or sameness. First, we can speak of persistence or *identity over time*: we may choose to believe that, despite the passage of time, an entity remains identical if it occupies the same space, or has the same components, or retains the same properties (or some combination of these). These notions allow us to speak of ‘the same bench’ that we sat on yesterday, or (a little more problematically) ‘the same watch’ that we had before the repairer dismantled it and reassembled its components. Alternatively, or additionally, we can attribute sameness to entities that we encounter discretely, but consider identical because they are *alike*. Norman Paskin (2003) wrote that ‘ultimately nothing is the same as something else’, and this appears to be a self-evident truth; but in practice we sometimes consider that the differences between two discrete objects are inessential, and treat the two as identical. This latter perception of identity will provide a primary focus for this paper.

Philosophical questions underlie this discussion, and concepts of identity have been debated by philosophers and scholarly thinkers from Heraclitus to Ferdinand de Saussure. Saussure’s (1973 p.151) example of discrete objects that are alike is a train that leaves Geneva for Paris every day at 8.45 p.m. Although the engine, the coaches and the train crew are different, we are very likely to consider these differences inconsequential; if timing and destination are the important factors, we may be inclined to say that the ‘same’ train leaves at 8.45 each day. There are innumerable other examples of objects that share so many properties that sameness can easily be attributed to them. Two cans of soup occupy different physical spaces, so in this sense (if no other) they are not identical; yet when we inspect the two cans, or two cars coming off a factory production line, or two instantiations of an edition of a printed book, we may be happy to say that each is the same as the other. Their sameness is a matter of interpretation and we can expect disagreement about

the appropriateness of describing any two or more entities as identical. But even in the case of so-called ‘identical’ twins, although we know that each of the pair is likely to have many distinguishing marks and individual characteristics, we also know that many people will see their differences as trivial or insignificant.

This is effectively the position reached by various commentators on digital preservation, who assert that although migration inevitably brings some changes it is an acceptable solution to the challenges of technological obsolescence provided it retains what are variously called the ‘significant properties’ (Cedars 2001; Hedstrom and Lee 2002; Knight and Pennock 2009), ‘salient features’ (Allison *et al.* 2005 p.370) or ‘essential characteristics’ (Hofman 2004 p.90) of the objects to be migrated. The structure, appearance and functionality of objects can all be at risk during migration, but even if some differences are introduced – or so the argument runs – this will be inconsequential if migration is carried out in such a way that no ‘significant’ loss occurs and the features necessary ‘for digital objects to remain accessible and meaningful’ (Hockx-Yu and Knight 2008 p.142) are preserved intact. Arthur Allison *et al.* (2005 p.371) argued for a cautious approach, suggesting that digital curators need to devise ‘auditable processes that ensure that the[se] features are not altered by the transformations of the bitstream’, but Kenneth Thibodeau (2002) wrote with apparent confidence that ‘it is possible to change the way a conceptual object is encoded ... without having any negative impact on its preservation’ and that ‘different digital encodings of the same conceptual object ... can preserve ... [its] essential characteristics’.

Similar issues, though not necessarily discussed under these labels, arise when records are converted from one medium to another (e.g. paper to microfilm; digital to paper; paper to digital) and when printed or electronic transcripts of manuscript originals are made. Like migration, these activities may be undertaken to support preservation – they can assist with preservation objectives by introducing more stable media, reducing handling of fragile originals, or providing ‘security copies’ for offsite storage as part of a vital records programme – but they are often motivated not so much by long-term preservation requirements as by desires to make archival resources more immediately available to wider audiences. Whatever the objective, guidelines for media conversion and transcription often affirm that certain essential aspects of the originals must be safeguarded and retained in the products that emerge from these activities.

In the literature on digital imaging, for example, Stephen Chapman and Anne Kenney (1996) contended that a successful imaging project requires the identification of ‘meaningful document attributes’ or ‘key features’ of the originals, to ensure the utility of the digitised products. In another publication, Kenney and Chapman (1996 p.26) referred to the need to determine ‘significant detail’ or ‘essential detail’ in materials to be digitised. More recently, Steven Puglia and Erin Rhodes (2007) argued that imaging projects should ‘replicate and carry forward ... the appropriate essential characteristics that tell us about the original resource’.

Writers on transcription and editing sometimes seek a kind of identity between a transcript and its original. As long ago as 1923, a team of authors of a report on editing historical documents asserted that ‘the transcriber should not omit anything that is, or insert anything that is not, in the manuscript’ (Anglo-American Historical Committee 1923 p.7); but the team (whose membership included Hilary Jenkinson) added the caveat ‘so far as the conditions of transcribing allow’. The caveat seems crucial since, as Claus Huitfeldt and C. M. Sperberg-McQueen (2008 p.297) noted, there are troublesome questions about what it means ‘to say that a manuscript symbol is “the same” as a typed or printed symbol in a transcription of the manuscript’.

Recognising that some degree of loss is inevitable when copies are made, scholars who have studied the transcription and editing of ‘literary’ works have often chosen to distinguish the

'substantives' of a work of literature from its 'accidentals', or 'signal' from 'noise'. The distinction between substantives and accidentals was first made in 1949 by the critic and bibliographer Walter Greg, who described the 'substantive readings' as those 'that affect the author's meaning or the essence of his expression', and the accidentals as 'spelling, punctuation, word-division, and the like, affecting mainly ... formal presentation' (Maxwell 1966 p.376). More recently, Greg's ideas have come under attack – not least because of the fashionable belief that meanings derive from readers' responses rather than authors' intentions – and some scholars have preferred an alternative approach, distinguishing the *signal* that works convey to their readers from the *noise* that can obscure it, using terminology derived from the 'communication theory' (or 'information theory') of Claude Shannon and Warren Weaver (Hammond and DelVecchio 1995 pp.46-8; Hunter 2007 pp.66, 76; Shannon and Weaver 1949).

Although Greg suggested that accidentals as well as substantives can be considered authoritative, many later commentators have seen accidentals or noise as unhelpful 'surface features' that can sometimes be ignored or eliminated in transcription (Driscoll 2006 p.254; Hammond and DelVecchio 1995 pp.47-8; Hunter 2007 pp.33, 76). In theories of literary editing, substantives, or signals, are commonly designated for protection on the grounds that they are 'what the reader needs' (Hammond and DelVecchio 1995 p.47) or the 'readings ... which are significant' (Hunter 2007 p.146). The terminology of substantives and accidentals, signals and noise, has not yet found its way into writings about the transcription and editing of archival records – a field that is under-theorised in comparison with 'literary' editing – but, like their literary cousins, writers on archival editing have affirmed that it is both possible and necessary to make informed decisions about the elements of a record that merit reproduction and those that do not (Hunnissett 1977; Harvey 2001).

Significant content

Although there is little recognition of common interests between editors and digital preservationists, in effect both groups believe that success in their mission depends on the ability to define what the preservation specialists call 'significant properties' or 'essential characteristics'.¹ This is the easy part, of course. The hard part is deciding what might be deemed essential or significant. The following sections of this paper examine some differing perspectives on where significance may lie.

In the digital arena, many writers have made distinctions between the *form* and the *content* of objects, and some have suggested that the significant properties needing protection are those that relate to content. In this view, the critical objective in any migration or conversion exercise is to preserve *information*, or *texts*. This contention is set out at length in two American papers that are closely linked, although composed thirteen years apart. The first, written in 1990 by Steven DeRose, David Durand, Elli Mylonas and Allen Renear, was entitled 'What is text, really?'. Its authors argued that 'the distinction is largely one of form versus content', and that the 'essential parts' of a digital document are 'content objects' such as 'paragraphs, quotations, emphatic phrases, and attributions', in contrast to aspects of 'appearance', which they described as 'superficial and transient' (DeRose *et al.* 1990 p.3). The second paper, by Allen Renear and David Dubin (2003), used these assumptions to suggest that, in libraries, the identity of 'documents' for

¹ The word 'significant' is ambiguous. It can connote 'important' or 'noteworthy'. It can also bear semiotic overtones; A is significant because it signifies B. Digital preservation literature rarely discusses what might be meant by 'significant', but in practice the former sense seems to prevail among preservation experts. Their parallel use of 'essential' appears to confirm this.

digital preservation purposes can be found at the level of the ordering of the 'content objects' discussed in the earlier article.

The tone of these two papers was largely theoretical, but the emphasis on content can be paralleled in Chapman and Kenney's work on digital imaging practice, which affirmed that images of 'research library materials' should convey 'all significant information contained in the original documents' and that the imaging process must be matched 'to the informational content of the original ... no more, no less' (Chapman and Kenney 1996; Kenney and Chapman 1996 p.8). To some degree it can also be paralleled in the work on digital preservation and migration undertaken by the Cedars project team in the United Kingdom, who emphasised *content* and *functionality* as the significant properties of digital objects (Cedars 2001 p.14).

While it is unsurprising to find such views among writers with disciplinary backgrounds in librarianship, similar opinions are also sometimes found in archival discourse. For example, the report of the *Effective Records Management Project* at Glasgow University used language reminiscent of the paper by DeRose *et al.* to recommend that digital preservation activities should focus on safeguarding the structured content of records. According to the report's authors:

'The record has a logical structure ... Logical structure ... may be ... sections ... subsections, paragraphs, lists, captions, footnotes, and so on. ... The presentation of a document is not the same as the structure ... It is this logical structure ... which must be preserved.' (Currall *et al.* 2002 Appendix 4 pp.8-9)

The report concluded that, in contrast to 'the practice and expectations associated with ... paper', the 'presentational form' of a digital record need not be stably maintained in a preservation system.

Recordness and authenticity

Other archival scholars have a rather different emphasis. In the 1990s, David Bearman and his colleagues at the University of Pittsburgh set out to examine the 'essential properties' of digital records and specify attributes of 'recordness', a quality that they equated with 'evidentiality' (Bearman and Sochats 1996). To Bearman, a 'complete' record is one that retains content, structure *and* context, and the crucial aspects of a record are not limited to its content but are to be found in its connection to appropriate metadata. In his view, 'the necessary and sufficient characteristics of data purporting to be records is a concrete set of metadata ... inextricably linked to, and retained with, the data associated with each business transaction', and the metadata required for 'recordness' are those that enable an audit to verify that records are 'comprehensive, identifiable ... complete ... and authorized' (Bearman 1996 pp.208-9).

Others suggest that the critical issue for records is their *authenticity*. In recent years, authenticity has become a contested notion, problematised by those who want to dispute the possibility of objective truths and binary oppositions between the authentic and the inauthentic (MacNeil and Mak 2007). Nevertheless, concepts of authenticity continue to occupy a central position in traditional understandings of archives and records management. Their core concern is with questions of identity over time: questions about whether and how we can trust that what we see or handle now is the same as what was seen, handled, transmitted, stored or used to effect an action at some previous time. The InterPARES project at the University of British Columbia (UBC) offered a traditional definition of an authentic record as 'a record that is what it purports to be and is free from tampering or corruption' (Duranti 2005 p.21). On the basis of this definition, one could argue that the authenticity of digital objects stands or falls by whether they remain

unchanged at bit level, and that any migration necessarily nullifies it, but the UBC team contended that authenticity can survive across migrations. Luciana Duranti (2002 p.13) affirmed that migration creates new *records*, on the grounds that it changes physical form; since 'the elements of a record's physical form are intended to convey meaning', any small change will 'generate a new and different record'. But the UBC team asserted that the new records can be 'authentic copies' of the records that have been migrated, provided they are made by the record creator in the ordinary course of business or are otherwise verified and authenticated (Duranti and MacNeil 1996 p.50). With this move, the grounds for authenticity are no longer limited to proving the identity of single objects over time; reliance is also placed on factors that allow identification of objects that are discrete but alike.

Duranti's statement that 'elements of ... physical form are intended to convey meaning' derives from her understanding of diplomatic theory. In the world of diplomatic scholarship, 'extrinsic' elements of physical form as well as 'intrinsic' elements of intellectual form are deemed critical for verification of authenticity (Duranti 1998 p.151). The elements of intellectual form (*elementi intrinseci* in Duranti's native Italian) 'depend on the content, or more precisely the formal structure in which the content ... is represented' (Iacovino 2001 p.126). Elements of physical form (*elementi estrinseci*) include script, format, inserts, colours, 'special signs', seals and, in the digital world, time stamps and digital signatures (Duranti 2002 p.13), and potentially these might be equally crucial. Duranti's concerns for the 'nature' and authenticity of records, as well as her emphasis on their custody and contextual relationships,² give her work a very different perspective from that of scholars mainly or exclusively interested in preserving content.

Following the diplomatic perspective that 'these formal elements are ... category specific' (Duranti 2002 p.13) and that, depending on the type of record, there might be some elements that could safely be lost without compromising the ability to verify authenticity, the InterPARES 'Authenticity Task Force' was charged with developing a typology of electronic records, identifying elements that would permit verification of the authenticity of each record type and establishing whether these elements could be moved 'to a place where they can more easily be preserved, without affecting validity' (Duranti 1999 p.159 and 2005 p.20). The task force recognised that 'authenticity ... is threatened', not just when migration is performed, but whenever electronic records are transmitted across space or time; but it proved unable to develop a useful typology of such records or correlate authenticity with 'the presence of specific documentary elements or annotations' (Duranti 2005 pp.21, 41). During its work, it shifted from the rigid objective of 'defining each element's specific role in ensuring ... authenticity' (Gilliland-Swetland and Eppard 2000) to the less foundationalist intention of 'ascertaining ... elements the creator *considers* essential for verifying authenticity' (MacNeil 2000 p.56; my italics). The task force's published report (Duranti 2005 pp.211, 216) recommended that creators should establish 'the documentary forms of records associated with each procedure' they perform, following either 'the requirements of the juridical system' or the creator's own needs. The deliverables of the second InterPARES project (Duranti and Preston 2008 p.726) placed similar reliance on the contribution of record creators, advising creators (among other precepts relating to 'fixity') that it 'would be wise' to identify the 'essential characteristics' of records, in terms of intrinsic elements 'that convey the action in which the record participates and its immediate context' and of extrinsic elements 'that constitute its external appearance'.

² See Duranti (1997). The InterPARES Authenticity Task Force recognised that, in traditional diplomatic theory, the 'elements most relevant to ... a record's authenticity' are 'typically found in the record's documentary form and in annotations', but that other elements are found 'outside' the record as part of its 'larger documentary and administrative framework' (MacNeil 2002 p.29).

Editing matters

A concern with ‘extrinsic elements’ is also increasingly characteristic of literary and textual scholarship. Researchers in these fields do not use the labels ‘intrinsic’ and ‘extrinsic’, and their notions of authenticity have developed on rather different lines from those of diplomatists, but many of them are turning to considerations of external or visual effects to supplement their traditional concentration on written text. Most studies remain focused on published ‘literary’ works, but there are signs of growing interest in unpublished works and archival records.

The scholarly exploration of the correspondence of the American poet Emily Dickinson provides an instructive example. Dickinson’s letters are records – persistent representations of the messages she sent to her sister-in-law or her friends – and they provide informational content (for example, they tell us that the name of the preacher in Amherst, Massachusetts, in 1854 was Aaron Warner),³ but these dimensions are rarely of much concern to literary scholars, who are more likely to be interested in Dickinson’s creativity, style or literary production. Her correspondence displays numerous ‘deviations from straightforward linear script’ (Mitchell 2005 p.271), such as unusual indentations and line breaks, and her use of white space is distinctive. Recent studies by textual scholars have often sought to emphasise and analyse these features on the grounds that they are, or may be, not trivial idiosyncrasies but ‘signs or elements in an apparatus of intentional structure’ (Mitchell 2005 p.274). Even Dickinson’s placing of individual letter characters, or the size and shape of her punctuation marks, have been interpreted as purposeful visual strategies.

In their analysis of Dickinson’s correspondence, Ellen Louise Hart and Martha Nell Smith argued that:

‘certain manuscript features – calligraphic orthography, punctuation, capitalization and line breaks – help to convey meaning. ... Dickinson used ... the placement of words in relation to embossments, attachments, and margins to convey meaning.’ (Hart and Smith 1998 p.xxiii)

The language employed by these writers is surprisingly similar to Duranti’s, but the ‘meanings’ that Hart and Smith invoke are not those of a diplomatist’s authenticity. They do not have a fixed view on the precise meaning of ‘manuscript features’ – this would be unacceptable in current modes of literary criticism – but wish to prompt readers to examine such features and find their own meanings in them. From this perspective, ‘Dickinson did not ... intend *a* meaning or *the* meaning by (for example) her line breaks, but she did intend them to be *meaningful*’ (Mitchell 2005 p.274).

Not all interpreters of Dickinson adhere to the view that she bequeathed such foresightful gifts to literary scholarship, deliberately determining a postmodern indeterminacy in every detail of her correspondence. However, even sceptics may want to examine these details to support their contributions to the debate. Dilemmas therefore arise for anyone planning a transcription or a published edition; it appears that *any* feature of Dickinson’s work might be deemed significant, and it becomes impossible to draw up supposedly definitive lists of features that could be disregarded with impunity. Textual scholars can dispute whether Dickinson’s use of different forms of the letter ‘d’ is meaningful (Mitchell 2005 p.46), but transcripts of her correspondence that do not differentiate them inhibit any judgement.

³ The preacher’s name is given in the letter printed by Hart and Smith (1998) p.49.

Editors of archival documents have often been historians, and have often believed that any aspect of a document ‘that could not conceivably be relevant in *historical* research ... can safely be ... ignored’ (Harvey 2001 p.10; my italics) and that alterations for ‘clarity’ or ‘the reader’s convenience’ can be made ‘without affecting ... the document in any important way’ (Tanselle 1990 p.266). In the United Kingdom, the guidelines for editing archival materials produced by R. F. Hunnisett in 1977 have been influential. Editors were advised that:

‘the original punctuation and ... use of capitals should not be followed slavishly. ... Additional punctuation should ... be introduced and the existing punctuation modified in order to allow easy comprehension of the text. ... Any unnecessary punctuation in the originals can be omitted. ... There is no merit in preserving each scribe’s eccentricities.’ (Hunnisett 1977 pp.23, 32-3)

Hunnisett assumed that his readers would be transcribing older organisational records, but similar advice (‘one may correct, without notice, obvious slips of the writer’s pen’) was given to editors of personal ‘manuscripts’ in the *Harvard Guide to American History* (Handlin *et al.* 1954 p.96). Such practices are obviously inimical to textual scholarship, and more recently there has been a reaction against them, much of it in response to criticisms made by an American textual scholar, G. Thomas Tanselle (1990 pp.7, 10), who argued that ‘false starts, excised words, slips of the pen, peculiar abbreviations, and unusual punctuation ... can often be important clues to the writer’s psychology and personality’ and that modernising spelling and punctuation is ‘inappropriate because it conceals some of the evidence that constitutes the reason for looking at the document in the first place’.⁴ In the United Kingdom, a distinguished editor of archival texts, P. D. A. Harvey (2001 p.43), has affirmed that ‘punctuation ... shows how writers thought about their writing and can thus be seen as a crucial part of the record’.

Emily Dickinson’s correspondence is of particular interest to much current scholarship,⁵ but it would be rash to assume that other archival materials will be of little concern to textual or literary scholars. Just as elements in Dickinson’s letters once thought trivial – such as words divided between two lines – have come to be viewed as important by such scholars, we may suppose that undervalued elements in other materials will be reassessed sooner or later. A Dickinson scholar has commented that ‘Jane Austen’s letters are full of instances where there is writing upside down on the page, and yet this has no significance’ (Mitchell 2005 p.15); but it is easy to imagine a researcher deciding to look for significance in Austen’s use of upside-down writing, and not much more difficult to conceive of the growth of a substantial scholarly industry investigating upside-down writing in 19th century documents. Interest need not be confined to the records of literary figures; a critic of Dickinson’s work has studied paragraph spacing and indentation in letters written by ‘non-literary’ individuals to provide comparators for his study of spacing and indentation in Dickinson’s own correspondence (Mitchell 2005 pp.288-90). Textual studies of archival materials have usually focused on personal letters and diaries – perceived as the most ‘literary’ of archival genres – but scholars have begun to recognise that there is no unbridgeable divide between literary and non-literary writing, and some now take an interest in the stylistic, rhetorical or visual devices used in medieval and early modern accounting documents, 19th and 20th century business correspondence or 21st century optometry patient records (Tebeaux 1999 and 2004; Beamer 2003; Varpio *et al.* 2007).⁶

⁴ Tanselle argued these points at length in a paper entitled ‘The editing of historical documents’, first published in 1978 and reprinted in Tanselle (1990) pp.218-73.

⁵ See the Dickinson Electronic Archives website, <http://www.emilydickinson.org/>.

⁶ Interest in the layout and form of records is not restricted to textual and rhetorical studies. In their sociological investigation of medical records, Heath and Luff (2000 p.37) noted that the position of each

Significance and identity cannot easily be circumscribed, even in the best known or most widely studied archival records. In the manuscript of the American Declaration of Independence, the words ‘free and independent states’ were written ‘in a distinctive Italic script’ (Armitage 2007 p.22). We might choose to argue that the italicisation of these words has no impact on the information content of the Declaration or on its authenticity. Presumably it also did not affect what, in the spirit of J. L. Austin (1962), we might call its ‘performativity’, its original ability to deliver, or to help deliver, the actions that its words represent; it seems unlikely that the United States would have become less independent if these four words had not been italicised.⁷ Nevertheless, we may be reluctant to consider a transcript or digitised copy of the Declaration which reproduces these words in italics identical to a version which does not italicise them, or to aver that the use of italics in the manuscript has no significance at all.

Custodial adventures

Despite occasional dissenting voices in recent years, most professionals have understood the key concepts in archival science, especially the central doctrines of provenance, original order and the *fonds d’archives*, as being ‘first and foremost ... linked to the creator’ (Cook 1993 p.31; Shilton and Srinivasan 2007 p.94). The ‘significant’ features discussed in the previous sections of this paper are also associated with the creators of records, the context of records’ creation and the activities and events that they represent. This emphasis on origination is particularly apparent in diplomatic studies, with their focus on ‘the elements of the document important to its juridical character, which exist from the time of creation of the document’ (Iacovino 2001 p.127). Most users of records also want to look back, in one way or another, to the time of the record’s creation and the circumstances that surrounded it. However, we should not forget that the creation of a record is only the beginning of its story. The record represents an event, but it also outlives the event and acquires a career of its own. Despite precepts in professional literature about the need for secure and continuous custody, or for meticulously designed systems that will control every aspect of recordkeeping from or even before the moment of creation, in practice we know that things rarely work this way. Records often undergo many vicissitudes during their lives. They are used and misused, ordered and disordered, lost and found, bought and sold. They can be added to one collection and then moved to another, or arranged in one way and then rearranged to meet changing needs or respond to changing whims.⁸ Inevitably we must consider the possibility of finding significance in features that result from later intervention as well as those bequeathed by the initial creator.

component of a record is critical both to the writing of the record and to making sense of it when it is read. Archaeologists Nissen, Damerow and Englund (1993 p.30) commented that interpretation of the clay tablets of the ancient Middle East depends in part on the layout of each tablet and the positioning of its marks; different types of transaction had different layouts that determined the organisation of the record’s contents. Nissen and his colleagues noted how the tablets resemble certain formalised records of our own era, such as cheques and balance sheets, which also to some degree rely on the geography of their components to fulfil the purposes for which they are designed.

⁷ As Armitage suggested (pp.80-1, citing Derrida (1986) p.13), the performativity of a declaration of independence may be open to question; who or what empowers a state to declare its independence? But my comment applies to *any* ‘dispositive’ record: the use of italics in the text will rarely affect the disposition that the record achieves.

⁸ For the ordering and reordering of records over time and space, see MacNeil (2008). For an exploration of the concept of *collections*, and of the tendency in archival practice to neglect collection histories, see Yeo (2009) and (forthcoming).

File covers, for example, may supply evidence about storage or arrangement systems employed either by creators or by subsequent custodians or collectors; they may also indicate levels of previous use or even identify named users. Paste marks, perhaps with fragments of former mountings still attached to them, can be indicative of collections in which documents were held in the past and the actions of the collectors who assembled them. Pin holes and folds can also indicate past methods of delivery or filing, or other interventions by custodians or users; traces of rust can suggest the length of time for which documents were pinned or clipped together; tears, stains and other marks of damage provide evidence of accidents; and any of these can be matched against similar features in other documents, to assist in making inferences relating to custodial or archival history. Such features are often lost when documents undergo conservation work, a process which may also sometimes lead to the disappearance of original file covers or bindings. Perhaps there is an even greater risk that they will fail to survive a media conversion initiative, as was noted by the historians who remarked that microfilming of records from the Nazi era in Germany made it impossible to match up the staple marks (Posner 1994 p.43). However, it is not impossible to take steps to preserve such features or at least ensure that they are documented in some way. Conservators and archivists should recognise that some users will be interested in the record's custodial history and its adventures over time, and should not assume that paste marks, folds, tears and stains are blemishes with no significance, merely because they are unsightly or of later date than a document's first inscription.

Digital dilemmas

Some of these concerns appear to belong to the paper world, but it is intriguing to consider whether similar issues might arise in digital environments. Digital preservation literature is full of exhortations about the possible need to preserve the 'look and feel' of digital objects, and we must certainly be cautious about attempts to identify aspects of a digital record's appearance that can be disregarded. Helen Heslop, Simon Davis and Andrew Wilson (2002 p.13) suggested that preserving 'colour in the user interface' may be inessential, but it is not hard to imagine that future researchers might be interested in the colour schemes available or provided by default to the creator or erstwhile recipient of a particular record. The InSPECT project framework (2008 p.22) suggested that it is 'preferential' to 'maintain some formatting and visual relationship between page elements' when migrating a 'structured text document', but 'not necessary to recreate the exact appearance of the text'; yet it seems rash to assume that the level of concern with 'exact appearance' that scholars bring to Emily Dickinson's letters will never be applied to archival records of the early 21st century.

Of course, there is the oft-cited difficulty that appearance is frequently unstable in the digital world. Fonts, line breaks and other aspects of screen display can vary, depending on the particular configuration of hardware and software used to view the record. This phenomenon is perhaps most obvious in – but is certainly not limited to – e-mail messages and other records created in plain text or hypertext markup language (HTML). As technology changes across space and time, the on-screen appearance of such records – the way they are experienced by users – is unlikely to remain constant.

Indeed, we can plausibly argue that instability of appearance and variability of user experience are themselves properties of potential interest to future researchers. Archival preservation strategies that rely on conversion to portable document format (PDF) can capture stable representations of the appearance of e-mail messages or word-processed documents as they were presented to a user on a given occasion, but in the wider world presentation is variable and users do not experience e-mail messages as PDF files. Moreover, e-mail normally retains the possibility of reusing content, especially when messages are forwarded or replies sent. PDF viewer

applications do not provide the same functionality as word-processors or e-mail clients, and e-mail metadata are necessarily handled differently when PDF is employed. However desirable it may be to protect digital archives from unauthorised amendment, there is an inescapable paradox in any attempt to preserve user experience of potentially alterable objects by converting them to a supposedly unalterable format.

Creators of digital records who use malleable formats may do so because they consider a fixed appearance unimportant, or merely because no other format is available to them. In other circumstances, however, creators may choose to use a file format that is designed to inhibit variability of output, together with a distinctive page layout or a non-standard font such as Papyrus or Comic Sans. In such cases, the creator is applying a level of deliberateness that a future user might find significant, and there is a much stronger argument for seeking to preserve a stability of appearance.⁹

In the digital world, such concerns may not be limited to aspects of appearance that are immediately visible on screen. Consider, for example, a word-processing application that offers a tick box labelled 'Keep with next', which allows creators of documents to instruct the application to move a given paragraph to the next page, if necessary, to ensure it remains on the same page as the paragraph that follows it. When the document is displayed, readers are not alerted to the fact that this box has been ticked; indeed, depending on how the application manages 'normal' page breaks, it is not unlikely that the two paragraphs will appear together by default, and the instruction to 'Keep with next' may then seem to have been an unnecessary precaution. Nevertheless, this instruction is represented in the binary file created by the word-processor. If the file is migrated, does the instruction need to survive in such a way that it remains recognisable in the post-migration environment? We might be inclined to say that it is redundant, especially if the two paragraphs appear on the same page by default; perhaps all that is needed is to ensure that the page breaks inserted by the original application are preserved in the migrated copy. But the precedent set by the textual study of Dickinson's letters may make us think again. When textual critics begin to take an interest in digital archives of the 21st century, some of them might argue that the ticking of this box by the document's creator constitutes significant evidence of intentionality or meaningfulness.

In contrast to paper documents, digital objects often carry large amounts of 'hidden' information. We may surmise that some of these *secreta*, such as hidden comments or formulae in spreadsheets, could be indicative of what Tanselle called 'the writer's psychology and personality', though as yet such aspects of digital composition have not been fully researched. Traces of later digital interventions are also different from those of the paper world; paste marks and pin holes are absent, but other traces may remain, perhaps in the form of detailed 'use metadata' gathered by transaction logs or recordkeeping systems programmed to capture evidence of a record's life history, or perhaps the scantier and less obvious traces that are retained by almost any digital object. Objects frequently carry metadata, hidden or apparent, indicating when they were last accessed, printed or modified. Content that has been changed, overwritten or

⁹ Care may be needed to achieve this where fonts are concerned, since computer systems do not usually store fonts in document files, and are programmed to substitute an alternative if a designated font is not available. (A timely demonstration of this potential pitfall occurred when I submitted this article to *Archival Science* via the PDF conversion tool on the journal's website; this tool had no difficulty in handling *Comic Sans*, but it proved unable to retain my selection of the 'Papyrus' font, and silently converted the word Papyrus to the somewhat different 'Arabic Typesetting' font.) Much of the computer industry remains pervaded by notions that only verbal content matters and that, ultimately, features such as fonts are not significant; archivists and users of archives may perhaps take a different view.

deleted, whether by the initial creator or by another party, may also survive concealed within the object itself or on the storage medium used when the changes were made. Forensic techniques, widely used to expose 'hidden' digital evidence in legal disputes, can also recover concealed content for cultural purposes. In Oxford, for example, the Bodleian Library has applied such techniques to the digital archives of the English playwright Alan Bennett, to reveal (with Bennett's consent) deleted drafts of his work (Thomas 2009). However, hidden metadata or concealed content may not survive a migration or conversion exercise unless care is taken to preserve them. To many users this may be of no consequence and a migrated copy that omits hidden information may appear identical to its unmigrated predecessor, but others may disagree and argue that identity has been lost.

Communities of practice and boundary objects

Migration, imaging and transcription are different practices requiring different skills, but they face the shared dilemma that identity conditions commanding general assent are rarely easy, and often impossible, to define. The best we are likely to be able to achieve – and even this may not be simple – is agreement that certain features are considered 'essential' for constituting identity in a particular context. As David Levy (2000 p.27) suggested, 'determinations of which properties matter are made in the context of purpose and use'. Instead of trying to assert that two or more discrete representations can be objectively or universally 'the same', it makes more sense to argue that X might be considered interchangeable with Y in specific circumstances.¹⁰

Consider Saussure's example of the Geneva-Paris train: in a context where the timings or the route of the train are significant, each daily train may be perceived as the same; in contexts where the rolling stock or the names of the train crew are significant, each is likely to be perceived as different. As Anthony Giddens (1979 p.16) noted, to a repair engineer or a train spotter it is the rolling stock that gives trains their identity, but to intending travellers the times of departure and arrival are the significant features that render trains identical or non-identical. Universal assent seems unachievable, but agreement about identity may be possible within a social domain or community of practice, a group of people who have common goals and have learnt to understand the world in a similar way. Train spotters, engineers, timetabling officials and travellers can be seen as instances of such communities. Community membership is fluid, of course, and differences in emphasis may co-exist within a single community, but it is in communities of practice that shared perceptions and priorities are most likely to be found.

Like trains, records are what the American sociologist Susan Leigh Star has called 'boundary objects', phenomena of interest to many communities (Yeo 2008). Star argued that such objects span the boundaries of diverse communities of practice (Star and Griesemer 1989; Bowker and Star 1999 pp.296-8); the objects are shared across these boundaries, but each community will probably seek to use them in different ways. Records represent activities, events or other occurrences (Yeo 2008 pp.135-6), but we cannot assume that every user community will interpret them as records or that the concerns of every user community will be centred on their representational aspects (Yeo 2010 p.105). Nor can we assume that all users will be interested in authenticity or provenance or content or original intellectual and physical form. Some may have multiple concerns, but the interests of others may be limited to a single aspect. When interacting with records, as with other boundary objects, each community will bring its own perspectives, its own ideas of what is significant and its own criteria of identity.

¹⁰ For a parallel argument, see Searle's maxim (1969 pp.35, 51-2) that 'institutional facts' are underlain by 'constitutive rules' in the form 'X counts as Y in context C'.

Paskin (2003) offered the example of two photocopies, one of which has his fingerprints on it while the other does not. Let us assume that the graphical layout and textual content are alike in both copies. The fingerprints probably have no significance to a community of business users, whose members are likely to recognise the two copies as identical, but considerable significance to the community of forensic scientists, whose members will perceive one copy to be crucially different from the other. As boundary objects, the photocopies are potentially of interest to both communities, but each community views them in its own way and forms its own opinions about their significant properties. Similarly, the embedding of records in structures reflecting creation contexts or custodial history may be highly significant to scholarly historians, but much less significant to genealogists searching for factual data about their ancestors or picture researchers seeking attractive images for a magazine. In the language of the Digital Preservation Testbed (2003 p.26), 'the "meaningful" and "relevant" aspects of a record' are 'in the eye of the beholder'.

For any given 'aspect' of records, or of railway trains, we may also want to ask how much divergence the members of a community are prepared to allow. How many carriages can be altered before the community ceases to consider today's 8.45 train identical to yesterday's? How many discrepancies in font size, or how much variation in colour, can be permitted before the community begins to deny the equivalence of the migrated record to its unmigrated predecessor? Questions like these lurk in the undergrowth of current approaches to digital preservation, but there can be no easy answers.

The philosopher Nelson Goodman famously asked whether a performance of a Beethoven symphony with a single wrong note still counts as a performance of the work in question. As Goodman (1976 p.187) pointed out, 'by a series of one-note errors we can go all the way from Beethoven's *Fifth Symphony* to *Three Blind Mice*', and it is impossible to be dogmatic about the numbers of errors that can occur before a performance ceases to be recognisable as the *Fifth Symphony*. Goodman believed that even a single wrong note disqualifies it, but others might display more latitude. We can surmise that a community of professional musicians might allow fewer wrong notes than communities of casual listeners or occasional concert-goers. Of course, some communities may introduce additional identity criteria, requiring, for example, that the performance employ contemporary instruments as well as follow the composer's score. Once again there may be a range of possibilities, from an insistence on antique instruments known to have been used at concerts in which Beethoven personally participated to a willingness to admit modern replicas of the instruments of his time. Some might say that much depends on *which* note is played wrongly or *which* instrument is non-contemporary. Regardless of whether our concern is with symphonies or trains or migrated records, opinions will vary, not only on properties considered significant, but also on the level of acceptable divergence.

Problematizing significance

I have rehearsed these arguments at some length because they must cast doubts on claims by some digital preservationists that records (or other 'classes of digital objects') have inherent 'essences' that can be determined or quantified as sets of 'significant properties ... defined in a rigorous and measurable manner' and preserved across episodes of migration (InSPECT 2006 pp.1-2; Brown 2007 p.7). As two recent papers in *Archives and Manuscripts* have suggested (Sloggett 2005 p.114; Collins, Collins and Garnaut 2007 p.94), concepts of significance appear to be 'culturally loaded':

‘We cannot speak of significance as an inherent attribute ... waiting only to be discerned. ... Significance, rather, is a quality that we assign ... based on the theoretical framework within which we happen to be thinking.’ (Tainter and Lucas 1983 p.714)

Notions of ‘significance’ in digital preservation, and in media conversion, transcription and editing, are no less problematical than the long-running debates about notions of ‘value’ in appraisal theory. The idea that records have ‘values’ can be found in the writings of archival pioneers such as G. Herbert Fowler (1939 pp.10-11) in the United Kingdom and Margaret Cross Norton (1939) and Philip C. Brooks (1940) in the United States, and has been widely associated with appraisal methodologies since its codification in T. R. Schellenberg’s influential *Modern Archives* (1956). Many newer ‘macro-appraisal’ methods also depend on assessing and prioritising value, even if this is sometimes disguised by using wording that emphasises assessment of the supposed ‘importance’ of record creators or functions rather than their ‘value’ (Cook 1992 pp.52-3). All these approaches assume that archivists can decide which records merit retention, or acquisition, on the basis of a determination of values ‘waiting only to be discerned’. But even if importance or value can have an objective existence – and by no means all students of axiology deny the possibility, or even the necessity, of absolute values¹¹ – there remains the problem, noted by Hans Booms (1987) and Terry Eastwood (1992 pp.81-3), that – theological or ideological dogma apart – there are no apprehensible standards or benchmarks that can be used to measure value absolutely. Despite the essentialist rhetoric of much professional literature, archivists have increasingly perceived that values can only be ascribed, not discovered.¹² All we can hope to know, and perhaps to measure, are the subjective values that communities and individuals assign.

Value and significance are kindred concepts, and similar notes of caution are sometimes found in writings about significance by digital preservation experts. At the beginning of the century, many reports and papers on digital preservation included statements that ‘a digital object’s significant properties are not absolute’ (Research Libraries Group 2002 p.23; cf. Cedars 2001 p.14), and that judgements would be needed to determine which properties are deemed significant. Recent initiatives have sometimes repeated these formulae, but in practice have often shown little recognition that significance can be subjective or fluid. Writers on ‘significant properties’ frequently prefer to see them as ‘inherent properties, which ... represent the qualities of the object’ (Brown 2007 p.7), or as properties that are significant to the long-term preservation of digital objects (Brown 2007 p.5) rather than to their various actual and potential users and uses.

Writing in a more philosophical vein, Allison *et al.* (2005 pp.370-1) also seem to imply that the ‘salient features’ of an object can be definitively identified. A variation on this approach, also tentatively suggested by the InterPARES project and by Allison *et al.*, is to attribute such features to *types* or *classes* of object. In the imaging field, Puglia and Rhodes (2007) have asserted that the essential characteristics to be safeguarded when digital images are made ‘will be unique to the collection/record/media type’. Undaunted by InterPARES’s failure to construct a usable typology, digital preservationists often lay similar emphasis on determining which properties ‘should be considered significant for different classes of record’ or ‘for a range of object types’ (InSPECT 2006 and 2008 p.10). For example, a project in Portugal (Ferreira, Baptista and Ramalho 2007) has attempted to compile lists of significant properties for preserving still images, text documents and relational databases, and a researcher in Austria has set out to define the significant properties of dynamic documents and digital art (Guttenbrunner 2008). The InSPECT project in the United

¹¹ See for example the arguments adduced by Attfield (1987) pp.25-31.

¹² In the axiological literature, useful introductions to the subjectivities of value include Smith (1988) and Goldman (2008).

Kingdom initially followed a similar route, seeking to identify the properties of raster images, e-mails and other formats, and suggesting that e-mails, for example, have '28 properties ... that must be preserved' (InSPECT 2008 p.25). Research of this kind is aimed at establishing systems to support the migration of particular file formats, by identifying 'the most appropriate target format ... which supports the designated significant properties ... at the lowest possible risk' (Brown 2007 p.10), and the inclusion of definitive descriptions of significant properties in file format registries and representation information registries such as PRONOM (<http://www.nationalarchives.gov.uk/pronom/>) and the Planets Characterisation Registry (Planets 2008). Such outcomes would undoubtedly be beneficial for migration technologists, providing them with useful practical tools constructed on a seemingly objective basis, but it is arguable that in most user communities notions of significance would be more likely to focus on the properties of record genres, such as reports, directives or contracts, than on the properties of particular versions of the .DOC format.

The InSPECT project adopted the idea of 'designated communities' from the Open Archival Information System (OAIS) model and noted that 'the importance attributed to a property may differ between communities', but its first published recommendations showed only minimal levels of interest in divergent notions of significance; instead, the initial emphasis of the project was on the development of 'a canonical list of significant properties' and the construction of a classification system and an assessment template, following the principle that 'if [a property] can be identified, it can be measured' (InSPECT 2008 pp.5-7, 16). The project report published in 2008 suggested two 'designated communities': 'institutions performing preservation' and a broad category of 'academics and other users' (InSPECT 2008 pp.7-8); but the 'significance levels' tabulated in the report appear to have been assigned more or less arbitrarily by members of the project team, and the report showed little real concern for the views of user communities or the diversity of their perceptions. However, a presentation by Angela Dappert and Adam Farquhar at the ECDL conference in 2009 reminded digital preservation specialists that significance is not universal and must always depend on the varying needs of 'stakeholders'; and the final outputs of the InSPECT project, which became publicly available while this article was being peer-reviewed for *Archival Science*, adopted a very different stance from the project's earlier publications, recommending a detailed evaluation of 'stakeholder functional requirements' as well as analysis of the properties of digital objects (Dappert and Farquhar 2009; InSPECT 2009).

Almost a century ago, in discussing documentary editing and transcription, the Anglo-American Historical Committee (1923 p.7) noted that an edited text 'will be used for several purposes. The historian will use it to establish ... facts; the student of diplomatic for its form and style; the philologist will expect an accurate representation of the spelling of the original; and in course of time other scholars may want to use it for purposes yet unknown.' Broadly similar observations could be made about the multiplicity of present and future users and uses of original paper records and of digital records in migrated or unmigrated form. As Marie-Anne Chabin (1999 p.112) remarked, 'anticiper toutes les utilisations ... est de nature à épuiser l'imagination la plus prolifique'. There are tensions between the highly diverse interests of potential users and the preservationists' desire for definitive solutions: tensions that, until very recently, most research in digital preservation has failed to acknowledge.

Archivists and user communities

In the preceding sections of this paper I have attempted to show that the putative sameness of discrete entities and the significance that may attach to particular properties are matters open to dispute and varying interpretations. As with most areas of dispute, there is also room for debate about whose interpretation might prevail, and who shall be the arbiter of differing views. The

United Kingdom Joint Information Systems Committee (JISC, undated) noted that significant properties may be those ‘considered significant by the creator and the users’, but most writers about significant properties have placed the task of identifying such properties firmly on the shoulders of archivists, librarians or other curators. Hannelore Dekeyser (2006 p.85), for example, asserted that ‘distinguishing between essential and incidental characteristics is the responsibility of the archivist’. In digital imaging, similarly, the consensus is that ‘curatorial judgments’ will be used to identify ‘key features’ or ‘meaningful document attributes’ for digitisation (Chapman and Kenney 1996; Puglia and Rhodes 2007). These approaches rest on assumptions about professional competence; it is the editor, or the archivist, who decides what user communities will find ‘significant’.

Here, too, parallels with other forms of appraisal are apparent. Cynics might agree with Booms’s aphorism (1987 p.82) that ‘documentary sources become valuable only when the archivist accords them value during the appraisal process’. Of course, archivists can study the activities of users and seek to uncover their views on significance or on identity criteria for migrated copies. User activity is measurable, at least to some degree, and in his review of appraisal methods Eastwood (1993 pp.117-18) contended that we can employ our knowledge of use as an empirical measure of value. If we need further, or more qualitative, data about users’ opinions, we can also undertake consultation exercises. Ross Harvey (2005 p.216) noted that, in a migration project at the Australian Department of Family and Community Services, the National Archives of Australia held a focus group, which ‘provided feedback about what was important to retain’. But user studies are far from infallible as a basis for assessing significance. Patterns of use – and users’ responses to consultation exercises – are often determined by other factors besides levels of value or significance to the users concerned. For example, issues of awareness and accessibility may come into play, and people may fail to use, or assign significance to, records or record properties that they do not know about or are unable to access easily.

Moreover, as the appraisal literature frequently emphasises, archivists can measure only past or current use, and can hold consultation exercises only with current users or current potential users. Methodologies such as those developed in Australia to assess significance of government functions (Cunningham and Oswald 2005) can only employ retrospective indicators. Even if past or present values may be measurable to some degree, such values may not remain static in future. Just as textual critics’ views of which properties are significant in Dickinson’s correspondence have changed over time, a feature or features that seem unimportant to one generation of researchers in any user community may be perceived as highly valuable by their successors. Eastwood (1993 p.118) suggested that knowledge of past use forms a basis for projecting future value, but there are limits to the feasibility of such projections and we cannot hope to foresee the uses of records for ‘purposes yet unknown’.

At the *What to Preserve?* conference in London in 2008,¹³ some speakers recognised that decisions about an object’s ‘significant properties’ need to take account of the perspectives of different groups of users; others assumed that a single user community can be identified for each resource, resource type or repository. The notion that preservation decisions need to serve only one community, or at most a limited number of communities, also underlies much of the advice offered in digital library literature, such as Priscilla Caplan’s (2009 pp.6-7) recommendation that ‘any institution creating or acquiring digital materials for a user community should think hard

¹³ What to preserve? The significant properties of digital objects. British Library Conference Centre, 7 April 2008.

about what features of those materials are important to that community'.¹⁴ Some archivists might argue that the primary user community for an archival service should comprise those users who employ records for probative purposes – those who seek evidence of the activities records represent – and that preservation criteria perceived as relevant to the needs of these users should take precedence over criteria relevant only to a record's general informational content, custodial adventures, cultural resonances or visual appeal. Broadly speaking, this is the position taken by Bearman and many members of the Australian recordkeeping school, and also by Duranti and the InterPARES project team.¹⁵ But to many commentators who seek a wider view of records and archives and their role in society, this position seems constrictive and exclusionary.¹⁶ Records represent particular activities or occurrences, but they have many affordances beyond their representational capacity. Some of these affordances are aesthetic; others are symbolic, or relate to personal or social identity and continuity (O'Toole 1993; Etherton 2006; Bastian and Alexander 2009). The memorial affordances of records are wide-ranging: records may support individual and collective memory, not only of the activities that gave rise to their creation, but also of innumerable other aspects of the world in which they were created, maintained or used. A record's affordances are affinitive (Yeo 2007 p.331), and perhaps few of them are wholly independent of evidentiality. As Cook (2000) suggested, 'without reliable evidence ... memory becomes counterfeit', and the degree to which users benefit from a record's other affordances may to a greater or lesser extent depend on their view of its evidentiality or authenticity. But archival records offer what Verne Harris (2001 p.41) called a 'cornucopia of meanings' and their affordances transcend any single aspect of recordkeeping or use. In the words of the British *Standard for Access to Archives*, they 'serve multiple ... different communities of researchers and ... types of users' (National Council on Archives 2003 p.12). For archivists, conceptual understandings of records as boundary objects combine with moral obligations to society, and with pragmatic requirements for archival services to meet the expectations of policy makers, employers and users, to militate against imposing an exclusive view of the ontology of records or attempting to limit their 'significant properties' to those deemed significant by a single user community.

Suggestions that professional endeavours should be focused on the interests of a single user community have always been open to dispute, but appear even less plausible in the digital era. In the world of analogue systems some dissemination projects were necessarily targeted at single audiences, but the online environment offers potential means of juxtaposing alternative images, alternative readings, alternative versions of finding aids, or combinations of any or all of these, to meet the needs of different communities in ways that would previously have been difficult or impossible. In the editing field, where there have been long-standing tensions between perceived needs for editions that are easy to read and demands that editors should attempt precise reproduction of every small detail of an original, extensible markup language (XML) can offer innovative solutions. XML allows transcribers to include both normalised and non-normalised readings, and users to choose which readings they prefer – an approach warmly welcomed by the users consulted by the LEADERS project team at University College London (Sexton *et al.* 2004

¹⁴ This line of thinking is often associated with the identification of a 'designated community', as recommended in the OAIS model for digital preservation. Although OAIS's formal definition of a 'designated community' ('an identified group of potential consumers ...') acknowledges that it 'may be composed of multiple user communities' (Consultative Committee for Space Data Systems 2002 p.1-10), this *proviso* is frequently overlooked. The mental exemplar appears to be the so-called 'special library', intended for a defined readership with distinctive interests.

¹⁵ Cf. Duranti (1998) p.171, where she cites Australian writers such as Frank Upward and Glenda Acland in support of her position.

¹⁶ See especially the critiques of the 'evidential' conception of records by Cook (2000), Harris (2001) and Greene (2002).

p.199).¹⁷ Moreover, in the digital world, as well as using predefined sets of resources chosen by an archivist or an editor, users from different communities can benefit from opportunities to build their own online collections of resources, supply their own transcriptions or annotations, or add their own descriptive contributions. Despite the preservation challenges that new technology brings, digital environments – especially those built around hypertext or the ‘social web’ – can open up a range of possible modes of engagement for user communities, in sharp contrast to the closed systems implicit in traditional paper-based representations of archives.

Intimations of originality

The last topics we need to consider are originality and the shifting relations between ‘originals’ and ‘copies’. Of course, no transcript – printed or electronic – can offer an exact reproduction of a manuscript original. The variable shape and spacing of individual letter forms is replaced by the consistency of print or digital fonts, and users wanting to examine handwriting to establish authorship, or to see whether careless formation of words and characters gives the impression that a writer was hurried, will necessarily reject transcripts and scholarly editions as inadequate for their needs. Two decades ago, it might have been possible to suggest that an appropriate response to such users would be to mimic Queen Marie Antoinette and dismiss them with a cry of ‘let them read *microfiches!*’ (Storey 1991 p.66). Today, perhaps, users dissatisfied with edited transcripts might be despatched to consult digital images rather than microform, but both microform and digital images may render watermarks invisible, colours distorted or the use of different writing instruments less distinct. Moreover, two-dimensional imaging techniques are inevitably deficient at reproducing three-dimensional materials such as seals, bindings, clips, threads and other filing devices. Facsimiles of documents (of the kind sometimes made for exhibition or educational use) may capture some of the features that conventional images lose, but – as malicious forgers occasionally discover to their cost – few facsimiles are able to reproduce such features as the age and chemical composition of original paper or ink. Ultimately, there is no detail so small that its absence can be considered wholly inconsequential; however trivial a feature may seem, we can never be certain that no community will consider it significant or deplore its omission when copies are made.

In Domhnall Mitchell’s (2005 p.48) phrase, ‘only the original is the original’, and many studies of users of older materials have found evidence of users’ preferences for originals rather than images or transcripts. Unsurprisingly, many of the historians in a Canadian survey (Duff, Craig and Cherry 2004 p.66) reported that they found photocopies and other reproductions useful for access purposes, but almost all (92%) stated that they ‘liked ... the original the most’. John Seely Brown and Paul Duguid (2000 pp.173-4) described Duguid’s encounter with a researcher smelling items of 18th century correspondence; the researcher hoped to detect faint scents of vinegar that would allow him to chart the progress of cholera (vinegar having been used to disinfect letters during cholera outbreaks). Like users seeking to study the mud from the trenches that clings to the diaries and letters of soldiers who fought in the First World War, such researchers will be thwarted if they are denied access to the originals.

Practical requirements apart, there may also be emotional satisfaction to be gained from consultation of originals. In an earlier user survey by Joan Cherry and Wendy Duff, several respondents reported the pleasure they felt in handling old documents or seeing ‘the real thing’ (Cherry and Duff 1999 pp.13-14; Duff and Cherry 2000), and similar comments were made by users consulted during the LEADERS project; one remarked that even if ‘you get ... dirty and

¹⁷ For an account of how this might be achieved using the TEI (Text Encoding Initiative) markup system, see Driscoll (2006) pp.258-61.

covered in all sorts' there is 'still something about seeing and touching the original', and another noted that 'you don't get the same thrill from a JPEG' (Sexton *et al.* 2004 pp.202-3). This is a further sense in which 'records ... mean more than what the words in them say' (O'Toole 2006 p.51). Some critics may seek to dismiss such responses as fetishism, but others might prefer to view them in the light of Walter Benjamin's (1970 p.223) coining of the term 'aura' to allude to some of the properties associated with original artefacts, or the rather more prosaic 'intrinsic value' mentioned in guidelines issued by the US National Archives (1982).

Notions of originality are often associated with the idea of uniqueness: the belief that, at least in certain fields of human endeavour, there can be only *one* original of a given work. The archetype of the singular original is the painting; only the original is the work of (let us say) Rembrandt, and it stands above and apart from any copies, whether forgeries or 'legitimate' reproductions. Archivists have traditionally understood the record in this way; there is one unique original and everything else is a copy.

Where records are concerned, however, some caution is needed. Many records, such as contracts or peace treaties, can be expected to have two or more original instantiations. Pierre Chaplais (2003 p.95) has described how English kings in the middle ages sometimes sent two 'identical' messages *propter viarum pericula*; each instantiation of the message was sent by a different route in case the perils of medieval road journeys prevented one of them from reaching its intended recipient. James O'Toole (1994 pp.643-4) and Hillel Schwartz (1996 p.225) have drawn attention to Thomas Jefferson's polygraph device that could be used to create two 'identical' documents in a single act of writing. We may assume that the two instantiations of a contract or a royal message are commonly drawn up or signed sequentially, and that their duplicate wording may be the work of a copyist, but the polygraph was an innovation that allowed both originals to be created simultaneously.

Uniqueness, like identity, is to some degree a matter of interpretation. Ultimately 'nothing is the same as something else', and strictly speaking even two letters polygraphed by Jefferson are not identical, since the physical spaces they occupy, the molecules of ink and perhaps the watermarks in the paper are all different; from this perspective each can be considered unique. However, if their distinct material properties are considered non-significant, the two instantiations may be treated as interchangeable, and in this sense a letter from Jefferson or a message from the king need not be seen as unique; but we cannot infer from this that it lacks originality.

In the digital world, multiple instantiations are the norm, and it is arguable that we encounter a multiplicity of originals on a scale unparalleled in pre-digital eras. For example, instantiations of an e-mail message may be created more or less automatically at a number of locations as the message travels from sender to recipient. Critics who assume that originals must be singular, or that temporal priority determines originality, may suggest that the only 'original' is the initial instantiation held in the temporary memory of the sender's computer before the message is despatched; but others are likely to question this perception. From the recipient's viewpoint, the original may be perceived as the instantiation stored in the recipient's computing space when the message arrives, or its manifestation on the recipient's computer screen when the message is viewed. A comparable claim to originality attaches to the sender's saved instantiation of the message. Further instantiations created at intermediate locations on its journey also have claims to originality, since all played their part in the message's transmission.

We may need to replace the binary differentiation of original and copy with a more nuanced understanding of originality. The transmission of e-mail entails a chain of replication, and from this perspective the instantiations of e-mail messages viewed by their recipients are *n*-th

generation copies; yet from the recipients' perspective these copies constitute the original messages received, in contrast to (for example) printouts of such messages made for a paper filing system. From a recordkeeping perspective, we might argue that all the copies created during an act of communication are original. In this way, what is a copy in one context becomes an original in another.¹⁸

Some commentators have suggested that the proliferation of seemingly identical instantiations means there are *no* originals, or at any rate no persistent originals, in the digital world. Levy (2000 pp.24-5) described the 'digital realm' as 'a realm in which, as far as I can tell, there are ... only copies'. Heather MacNeil and Bonnie Mak (2007 pp.45-6) have recently suggested that 'most digital resources that fall under the purview of ... archivists' might be perceived as allographic, following Goodman's (1976 p.113) distinction between autographic works (those having a single original or a finite number of originals) and allographic works that are open to infinite repetition. Goodman's analysis, however, remains controversial. His taxonomy requires identity criteria for allographic resources, and to resolve this requirement Goodman assumed that 'properties constitutive of the work' can be isolated from those merely 'contingent'. In the case of a performance of a symphony – his favoured example of an allographic mode – he argued that strict adherence to the score is the constitutive property that guarantees a 'genuine instance' of the work (Goodman 1976 pp.116-17, 186) and, as we have seen, that even a single wrong note disqualifies it. Many critics have questioned these assumptions or the corollary that other properties, such as contexts of production, can invariably be dismissed as irrelevant to identity (Wreen 1983; Levinson 1990 pp.89-106; Pillow 2002).

The assumptions that underlie Goodman's theory of allographic works bring us by another route to the 'signal' of the textual editors and the 'significant properties' of the digital preservationists, and they encounter similar difficulties. Not every user community will always consider wrong notes important or production histories unimportant. Goodman's concern was with works of creative art, but comparable issues arise with records. We cannot assume that everyone will deem any one instantiation identical to any other. Some may consider the record retained by the sender identical to the record received by the addressee, but others may insist on their difference, even though both represent the same occurrent and their textual content is indistinguishable. Besides its physical separateness, each record will have its own contexts of creation, maintenance and use. Each may carry different metadata. Moreover, if the records are digital, each may be rendered differently when viewed on different occasions, depending on the technical configurations used to view it. Some commentators are likely to insist that each viewing of *any* record – paper or digital – creates new meaning and a different 'performance'. But none of this need preclude users from discovering Benjaminian auras of originality, even in the digital realm. Asked about 'the equivalent of bibliophilia in relation to CD-ROMs or floppy disks', Jacques Derrida (2005 pp.28-9) replied: 'Some particular draft that was prepared or printed on some particular software, or some particular disk that stores a stage of a work in progress – these are the kinds of things that will be fetishized in the future.' Critics who write about fetishism often do so with disdain, but if Derrida is correct, original digital records may perhaps provide future generations of researchers

¹⁸ In paper records systems, the converse applies: it is the *outgoing* correspondence that is necessarily captured in the form of copies (carbon copies created simultaneously with the letters destined for mailing out, or press copies or photocopies created at a slightly later moment). But at a still later date, when archivists plan imaging or transcription projects for these paper records, the erstwhile copies become originals from which a further generation of copies will be derived.

with feelings of emotional satisfaction comparable to those reported by users of analogue originals in our own era.¹⁹

There are implications here for archivists who find themselves impelled to make copies for reasons of preservation or accessibility. Clearly, a copy of an Emily Dickinson letter made by a 21st century archivist is not the letter that Dickinson sent; a later copy of a peace treaty is not one of the instantiations of the treaty that effected a cessation of hostilities at a particular moment of history. In the digital world, an archivist's later copy of John Poindexter's famous 'Well done' e-mail message is not one of the instantiations of the message that played a part in communicating Poindexter's congratulations on a successful arms deal in 1986, even though it may be indistinguishable from them in almost every respect. 'Only the original is the original', and there may be potentially 'significant' properties of originals that are not reproducible or can only survive copying in an attenuated form. Originality itself may be significant and cannot be replicated.

Of course, copying can enable wider access, and is often seen as a key element in vital records programmes and other preservation initiatives. But if copying necessarily incurs losses that are likely to be deplored by one user community or another, we may acknowledge an obligation to keep such losses to a minimum. With born-digital records, it is easy to generate copies that are almost indistinguishable from their originals, but initiating changes that might have been avoidable would seem to be another matter. A level of variability is inherent in most records – paper records suffer physical decay, and digital records may be subject to differences in rendition or automatic updating of their embedded metadata – but such variability is always limited and offers no rationale for the introduction of wider changes that go beyond its bounds.²⁰

Conclusions

This paper has ranged widely over the terrain of identity and differentiation, significance and value, originality and uniqueness. Although these topics are contentious and there can be no simple answers to the questions that they pose, some concluding thoughts can be offered. First, the concept of originality may be elusive, but it must not be dismissed as meaningless in a digital era. Digitised image copies of paper records differ from their originals, not only because they lose affordances deriving from tangibility or weight or smell, but also because they add new features that their originals lack, such as alternative display modes, zooming, enhancement or ease of further reproduction. Digitised transcripts offer multiple encoding, linkage, search and retrieval options that undoubtedly benefit many users but supply an experience very different from use of the originals. Like microform imaging and paper transcription, digitisation projects often rely on an assumption that the experience of sensing no more and no less than the affordances of the originals can be sacrificed for the sake of the greater accessibility that the copies provide.

Where our concern is with imaging or transcribing paper records, we might argue that these differences are of less consequence, because users wanting to examine the originals can normally

¹⁹ Despite widespread assumptions by archivists and digital preservationists that digital storage media are inconsequential, Derrida's reply raises the possibility that not only the 'draft' but also the 'disk' will be significant to some user communities. Disks with labels or marks seem obvious candidates, but we may also note that the use of digital forensic techniques depends on access to relevant storage media; the forensic work at the Bodleian Library (Thomas 2009) has only been possible because of the survival of Bennett's original floppy disks. Perhaps it is time to rethink our notions that storage media rarely matter in the digital world.

²⁰ Cf. the comments on 'bounded variability' made by Duranti and Thibodeau (2006) pp.47-8.

still gain access to them. Migration of digital records brings similar concerns about differences between migrated objects and their unmigrated predecessors – we may doubt whether a message filtered through a migration exercise can ever be incontestably identical to its predecessor – but with digital records a further critical issue is that migration is driven by expectations that originals will *not* remain accessible over time. As a response to technological obsolescence, migration implies that users may only be able to access older digital records in their migrated form. With the passage of time a sequence of migrations may be required, and each migration is likely to take us further from the original. Arguably, migration is flawed as a preservation strategy and could more appropriately be perceived as an access strategy. But alternatives such as emulation are rarely straightforward; at present, for many archival institutions, migration may be the only practically affordable solution.²¹

The notion that systematic copying or migration can generate impartial representations of originals dies hard. In the editorial field, it has been partly laid to rest,²² but it remains forceful in digital preservation, where much effort has been directed at constructing automatable rule-based systems to minimise the need for human labour in managing collections. The search for methods of quantifying significant properties of digital objects forms part of this venture, which has usually allowed little space for recognising subjectivity or fuzziness. Classification systems for data about significant properties are predicated on beliefs that aspects of experience can be definitively fixed and formally codified, and attempts to construct canonical lists of such properties can be seen as motivated by a wish to find an objective basis for decision-making and selectivity, in line with modernistic desires to establish rational foundations for human behaviour. As Chabin (1999 p.133) remarked, ‘si les universaux archivistiques n’existaient pas, il faudrait les inventer’. From the perspective of critics sceptical of modernist rationality, the preservationists’ approach runs the risk of imposing rigidity where experience is fluid and simplification where it is multifarious. Giddens (1979 p.69) noted that ‘structure is both enabling and constraining’, and many of the structures that digital preservation specialists have proposed seem likely to confirm his dictum, facilitating the automation of preservation processes but constraining the potential for responding to the variability and complexity of users’ understandings of significance and identity.

In the absence of loss-free conversion techniques, an obvious precaution is to retain originals: advice that seems equally applicable to unmigrated digital records and to the original newspapers whose destruction was deplored in Nicholson Baker’s *Double Fold* (2001). In practice, we may find that both originals *and* copies are needed, perhaps because some users find it difficult to visit the archival institutions where originals are held, or because some unmigrated digital originals may be inaccessible until new solutions to technological obsolescence become available. But if digital records are to be migrated, or paper records digitised or transcribed, we must move beyond approaches that presuppose a static and universally specifiable significance. Where possible, we should endeavour to find flexible approaches that adapt to the needs and interests of different user communities and recognise that estimations of significance will vary across communities and

²¹ Frequent migrations may be less necessary than archivists once believed; superseded file formats ‘from consumer-oriented commercial software products’ now usually remain accessible at least in the short-to-medium term (Rusbridge 2006). The introduction of XML formats in these products seems likely to enhance this trend. But longer-term accessibility, particularly of rarer or more specialised formats, remains problematic, and reliance on backwards compatibility features provided by software vendors can be dangerous; the latest version of *Bloopersoft Millennium Plus* may be able to open digital records created using *Bloopersoft 1994* but display them in a distorted form. The challenges of technological change are not yet fully resolved.

²² See for example the discussion of recent textual and editorial scholarship in Shillingsburg (2006) pp.14-17. However, there are indications that claims to definitiveness may be resurfacing in the new editorial world of XML encoding and digital dissemination (Kichuk 2007 pp.296-7; Sutherland 2009 pp.17-18).

across time. Instead of assuming that a single method will serve every purpose, we should consider using different approaches to migration, transcription or media conversion to serve the needs of different communities. However, given the multiplicity of records and their range of potential uses, we must also acknowledge that it may be difficult to meet the needs of all user communities or to make provision for the exceptional community member with atypical requirements. It will be almost impossible to predict how the interests of existing communities may change or what new communities may be formed in the future.

Further difficulties may arise with modes of scholarship that might be described as minority interests. For every user who finds paste marks, letter forms or hidden metadata significant, there could be a hundred to whom such matters seem irrelevant. Large numbers of users may perhaps aver that their main or only concern is with textual content. In a world of limited resources, it may not be easy to justify the cost of taking minority interests into account when migration or conversion solutions are sought; there may be no practical alternative to reliance on bulk processing and meeting the perceived needs of the majority of users. Nevertheless, a difficulty remains. In discussing transcription of medieval documents, the Anglo-American Historical Committee (1923 p.7) affirmed that 'it is hardly possible to insist that an editor shall note every folly that a careless scribe may commit', and Hunnisett (1977 p.33) advised against seeking to preserve scribal eccentricities; yet editors may be reluctant to follow advice that leads them to neglect the small numbers of users who are interested in careless or eccentric scribes. More recently, in considering questions of digital identity, Dekeyser (2006 p.83) suggested that 'if the money owed in a contract is changed from 100 to 1,000 per item, the integrity of the contractual terms is clearly compromised', but if a photographic pixel is changed from one shade of grey to another, 'this is probably not at all relevant'. To most users, this will doubtless seem correct, but if any users are interested in the minutiae of pixel colouring they might have a different view of significance.²³

Decisions about migration, conversion or editing therefore present us with ethical dilemmas broadly comparable to the dilemmas implicit in the prior appraisal decisions that determine which records we collect or retain. Even when the digital world offers new techniques that promise responsiveness to the needs of diverse user communities, the costs of exploiting their full potential may be prohibitive.²⁴ If we cannot meet everyone's needs we may have no choice but to privilege certain features of records, and thus the needs of certain communities, in the actions we take. Our recognition that records are boundary objects, and our sense of obligation to multiple user communities that this recognition entails, cannot easily prevail over practical limitations imposed by restricted resources and finite budgets.

If we find ourselves obliged to make rulings about which communities are to be privileged and which features of records deemed significant, we can see this as a further form of appraisal. Since

²³ Cf. the contention by staff of the National Archives of Australia that 'what the researchers want' from Australian census records is 'reliable personal information'. This was claimed to demonstrate that microfilm copies suffice to 'capture the essence' of the records (Wilson 2005 p.24; Wilson and Platzer 2004), and these assertions were used to justify a decision that the originals need not be preserved after microfilming. Yet, as with many essentialist claims, what is represented as a brute fact (X is significant, Y is not; X is of the essence, Y is not) is merely a construction: built, in this case, on assumptions that few users will be interested in, for example, the types of paper or ink issued to census enumerators, or that such users do not matter very much. Whatever we think an 'essence' might be, it is not to be equated with 'predominant use'.

²⁴ Editing or transcription projects that use XML encoding, for example, are often laborious and highly resource-intensive. A project using XML markup to include both normalised and non-normalised readings 'will almost certainly require several layers of input' (Driscoll 2006 p.258).

these rulings involve the exercise of powers of inclusion and exclusion, they will necessarily be controversial, but – as with other appraisal decisions – we cannot abnegate our responsibilities. We can consult widely and apply our professional judgement, but we must acknowledge that we cannot be neutral; we must not lay claim to scientific certainties or present our decisions as mere discoveries of essential truths about significance. Transparency – publicising the decisions we have made and not seeking to pretend that they are anything other than constructed outcomes of fallible appraisal processes – is as crucial as an awareness of their consequences.

As Annemaree Lloyd (2007 p.53) noted, both the ‘concept of significance’ and the ‘process of assessing significance ... should be seen as areas of tension and contestation’. The identity of discrete entities is also a matter on which opinions will vary depending on the contexts in which judgement is made. The copies that emerge from acts of migration, conversion or transcription are neither incontrovertibly identical to their originals nor carriers of properties that are objectively significant, but they may have a future role in showing what archivists, editors or digital preservation specialists *thought* was significant. In this sense the copies will themselves become originals. In recordkeeping terms, they function as metarepresentations rather than original records (Yeo 2007 p.341), but with the passage of time they will nonetheless come to be seen as original *objects* affording evidence of what kinds of copies were once thought important and how those responsible for copying exercised their decision-making power. To scholars researching medieval society, a 19th century edition of a medieval account roll is merely a reproduction, but to those interested in the perceptions and practices of 19th century editors it has its own aura of originality. If they survive, a similar destiny awaits the products of the copying processes of the digital era.

Acknowledgements

I am grateful to Laura Millar and to Elizabeth Shepherd, Andrew Flinn and other members of the ICARUS research seminar group at University College London, who kindly commented on an earlier draft of this paper. I am also grateful to Elaine Penn for drawing my attention to the references to ‘value’ in the works of Margaret Cross Norton.

References

Web addresses were accessed in December 2009.

- Allison A, Currall J, Moss M, Stuart S (2005) ‘Digital identity matters’. *J Am Soc Inf Sci Tech* 56: 364-72
[Anglo-American Historical Committee] (1923) ‘Report on editing historical documents’. *Bull Inst Hist Res* 1: 6-25
Armitage D (2007) *The Declaration of Independence*. Harvard University Press, Cambridge, MA
Attfield R (1987) *A theory of value and obligation*. Croom Helm, London
Austin J L (1962) *How to do things with words*. Clarendon Press, Oxford
Baker N (2001) *Double fold: libraries and the assault on paper*. Random House, New York
Bastian J A, Alexander B (eds.) (2009) *Community archives: the shaping of memory*. Facet, London
Beamer L (2003) ‘Directness in Chinese business correspondence of the nineteenth century’. *J Bus Tech Commun* 17: 201-37
Bearman D (1996) ‘Item level control and electronic recordkeeping’. *Arch Mus Inform* 10: 195-245
Bearman D, Sochats K (1996) *Metadata requirements for evidence*. Available at <http://web.archive.org/web/19970707063459/http://www.lis.pitt.edu/~nhprc/BACartic.html>
Benjamin W (1970) ‘The work of art in the age of mechanical reproduction’. In: Arendt H (ed.) *Illuminations*. Jonathan Cape, London, pp.219-53
Booms H (1987) ‘Society and the formation of a documentary heritage: issues in the appraisal of archival sources’. *Archivaria* 24: 69-107

- Bowker G C, Star S L (1999) *Sorting things out: classification and its consequences*. MIT Press, Cambridge, MA
- Brooks P C (1940) 'The selection of records for preservation'. *Am Arch* 3: 221-34
- Brown A (2007) 'Developing practical approaches to active preservation'. *Int J Digit Curation* 2(1): 3-11
- Brown J S, Duguid P (2000) *The social life of information*. Harvard Business School Press, Boston
- Caplan P (2009) *Understanding PREMIS*. Available at <http://www.loc.gov/standards/premis/understanding-premis.pdf>
- Cedars (2001) *The Cedars Project report*. Available at <http://www.webarchive.org.uk/wayback/archive/20050410120000/http://www.leeds.ac.uk/cedars/pubconf/papers/projectReports/CedarsProjectReportToMar01.pdf>
- Chabin M-A (1999) *Je pense donc j'archive*. L'Harmattan, Paris
- Chaplais P (2003) *English diplomatic practice in the middle ages*. Hambledon and London, London
- Chapman S, Kenney A R (1996) 'Digital conversion of research library materials'. *D-Lib Mag* 2(10). Available at <http://www.dlib.org/dlib/october96/cornell/10chapman.html>
- Cherry J, Duff W (1999) *Improving access to early Canadiana*. Available at <http://web.archive.org/web/20041129085829/http://www.fis.utoronto.ca/research/programs/digital/final+report.pdf>
- Collins J, Collins S, Garnaut C (2007) 'Behind the image: assessing architectural drawings as cultural records'. *Arch Manuscr* 35(2): 86-107
- Consultative Committee for Space Data Systems (2002) *Reference model for an Open Archival Information System (OAIS)*. Available at <http://public.ccsds.org/publications/archive/650x0b1.pdf>
- Cook T (1992) 'Mind over matter: towards a new theory of archival appraisal'. In: Craig B L (ed.) *The archival imagination: essays in honour of Hugh A. Taylor*. Association of Canadian Archivists, Ottawa, pp.38-70
- Cook T (1993) 'The concept of the archival fonds in the post-custodial era'. *Archivaria* 35: 24-37
- Cook T (2000) *Beyond the screen: the records continuum and archival cultural heritage*. Available at http://www.archivists.org.au/files/Conference_Papers/2000/terrycook.pdf
- Cunningham A, Oswald R (2005) 'Some functions are more equal than others: the development of a macroappraisal strategy for the National Archives of Australia'. *Arch Sci* 5: 163-84
- Currall J, Johnson C E, Johnston P, Moss M S, Richmond L R (2002) *No going back? The final report of the Effective Records Management Project*. Available at <http://www.gla.ac.uk/InfoStrat/ERM/Docs/ERM-Final.pdf>
- Dappert A, Farquhar A (2009) *Significance is in the eye of the stakeholder*. Available at http://www.planets-project.eu/docs/papers/Dappert_SignificantCharacteristics_ECDL2009.pdf
- Dekeyser H (2006) 'Authenticity in bits and bytes'. In: Neef S, Van Dijck J, Ketelaar E (eds.) *Sign here: handwriting in the age of new media*. Amsterdam University Press, Amsterdam, pp.76-90
- DeRose S J, Durand D G, Mylonas E, Renear A H (1990) 'What is text, really?'. *J Comput High Educ* 1(2): 3-26
- Derrida J (1986) 'Declarations of independence'. *New Political Sci* 15: 7-15
- Derrida J (2005) *Paper machine*. Stanford University Press, Stanford, CA
- Digital Preservation Testbed (2003) *Emulation: context and current status*. Available at http://www.digitaleduurzaamheid.nl/bibliotheek/docs/White_paper_emulation_UK.pdf
- Driscoll M J (2006) 'Levels of transcription'. In: Burnard L, O'Brien O'Keeffe K, Unsworth J (eds.) *Electronic textual editing*. Modern Language Association of America, New York, pp.254-61
- Duff W, Cherry J (2000) 'Use of historical documents in a digital world: comparisons with original materials and microfiche'. *Inf Res* 6(1). Available at <http://informationr.net/ir/6-1/paper86.html>
- Duff W, Craig B, Cherry J (2004) 'Finding and using archival resources: a cross-Canada survey of historians studying Canadian history'. *Archivaria* 58: 51-80
- Duranti L (1997) 'The archival bond'. *Arch Mus Inform* 11: 213-18
- Duranti L (1998) *Diplomatics: new uses for an old science*. Scarecrow Press, Lanham, MD
- Duranti L (1999) 'Concepts and principles for the management of electronic records, or records management theory is archival diplomatics'. *Rec Manag J* 9: 153-75
- Duranti L (2002) 'The concept of electronic record'. In: Duranti L, Eastwood T, MacNeil H *Preservation of the integrity of electronic records*. Kluwer, Dordrecht, pp.9-22
- Duranti L (ed.) (2005) *The long-term preservation of authentic electronic records: findings of the InterPARES Project*. Archilab, San Miniato

- Duranti L, MacNeil H (1996) 'The protection of the integrity of electronic records: an overview of the UBC-MAS research project'. *Archivaria* 42: 46-67
- Duranti L, Thibodeau K (2006) 'The concept of record in interactive, experiential and dynamic environments: the view of InterPARES'. *Arch Sci* 6: 13-68
- Duranti L, Preston R (eds.) (2008) *International research on permanent authentic records in electronic systems (InterPARES 2): experimental, interactive and dynamic records*. Coop Libraria Editrice Università di Padova, Padua
- Eastwood T (1992) 'Towards a social theory of appraisal'. In: Craig B L (ed.) *The archival imagination: essays in honour of Hugh A. Taylor*. Association of Canadian Archivists, Ottawa, pp.71-89
- Eastwood T (1993) 'How goes it with appraisal?'. *Archivaria* 36: 111-21
- Etherton J (2006) 'The role of archives in the perception of self'. *J Soc Arch* 27: 227-46
- Ferreira M, Baptista A A, Ramalho J C (2007) 'An intelligent decision support system for digital preservation'. *Int J Digit Libr* 6: 295-304
- Fowler G H (1939) *The care of county muniments*, 3rd edn. County Councils Association, London
- Giddens A (1979) *Central problems in social theory*. Macmillan, Basingstoke
- Gilliland-Swetland A, Eppard P B (2000) 'Preserving the authenticity of contingent digital objects: the InterPares Project'. *D-Lib Mag* 6(7-8). Available at <http://www.dlib.org/dlib/july00/eppard/07eppard.html>
- Goldman A H (2008) 'The case against objective values'. *Ethical Theory Moral Pract* 11: 507-24
- Goodman N (1976) *Languages of art*, 2nd edn. Hackett Publishing Co., Indianapolis
- Greene M A (2002) 'The power of meaning: the archival mission in the postmodern age'. *Am Arch* 65: 42-55
- Guttenbrunner M (2008) *Preserving interactive content*. Available at http://publik.tuwien.ac.at/files/PubDat_172294.pdf
- Hammond A, DelVecchio D (1995) 'General textual introduction'. In: Gunby D, Carnegie D, Hammond A (eds.) *The works of John Webster*, vol.1. Cambridge University Press, Cambridge
- Handlin O et al. (1954) *Harvard guide to American history*. Belknap Press, Cambridge, MA
- Hart E L, Smith M N (eds.) (1998) *Open me carefully: Emily Dickinson's intimate letters to Susan Huntington Dickinson*. Paris Press, Ashfield, MA
- Harris V (2001) 'Law, evidence and electronic records: a strategic perspective from the global periphery'. *Comma* 2001-1/2: 29-44
- Harvey P D A (2001) *Editing historical records*. British Library, London
- Harvey R (2005) *Preserving digital materials*. K G Saur, Munich
- Heath C, Luff P (2000) *Technology in action*. Cambridge University Press, Cambridge
- Hedstrom M, Lee C A (2002) 'Significant properties of digital objects: definitions, applications, implications'. In: *Proceedings of the DLM-Forum 2002, Barcelona, 6-8 May 2002*. Office for Official Publications of the European Communities, Luxembourg, pp.218-23
- Heslop H, Davis S, Wilson A (2002) *An approach to the preservation of digital records*. Available at http://www.naa.gov.au/Images/An-approach-Green-Paper_tcm2-888.pdf
- Hockx-Yu H, Knight G (2008) 'What to preserve? Significant properties of digital objects'. *Int J Digit Curation* 3(1): 142-54
- Hofman H (2004) 'Can bits and bytes be authentic? Preserving the authenticity of digital objects'. *Arch Comput* 14(3): 85-98
- Huitfeldt C, Sperberg-McQueen C M (2008) 'What is transcription?'. *Lit Linguist Comput* 23: 295-310
- Hunnisett R F (1977) *Editing records for publication*. British Records Association, London
- Hunter M (2007) *Editing early modern texts*. Palgrave Macmillan, Basingstoke
- Iacovino L (2001) 'Common ground, different traditions: an Australian perspective on Italian diplomatics, archival science and business records'. *Arch Manuscr* 29(1): 118-38
- InSPECT (2006) *JISC project plan*. Available at <http://www.significantproperties.org.uk/documents/Inspect-ProjectPlan.pdf>
- InSPECT (2008) *Framework for the definition of significant properties*. Available at <http://www.significantproperties.org.uk/documents/wp33-propertiesreport-v1.pdf>
- InSPECT (2009) *Final report*. Available at <http://www.significantproperties.org.uk/inspect-finalreport.pdf>
- JISC (undated) *The significant properties of digital objects*. Available at <http://www.jisc.ac.uk/whatwedo/programmes/preservation/2008sigprops.aspx>

- Kenney A R, Chapman S (1996) *Digital imaging for libraries and archives*. Cornell University Library, Ithaca, NY
- Kichuk D (2007) 'Metamorphosis: remediation in *Early English Books Online*'. *Lit Linguist Comput* 22: 291-303
- Knight G, Pennock M (2009) 'Data without meaning: establishing the significant properties of digital research'. *Int J Digit Curation* 4(1): 159-74
- Levinson J (1990) *Music, art, and metaphysics*. Cornell University Press, Ithaca, NY
- Levy D M (2000) 'Where's Waldo? Reflections on copies and authenticity in a digital environment'. In: Council on Library and Information Resources, *Authenticity in a digital environment*. Council on Library and Information Resources, Washington, pp.24-31
- Lloyd A (2007) 'Guarding against collective amnesia? Making significance problematic: an exploration of issues'. *Libr Trends* 56: 53-65
- MacNeil H (2000) 'Providing grounds for trust: developing conceptual requirements for the long-term preservation of authentic electronic records'. *Archivaria* 50: 52-78
- MacNeil H (2002) 'Providing grounds for trust II: the findings of the authenticity taskforce of InterPARES'. *Archivaria* 54: 24-58
- MacNeil H (2008) 'Archivalterity: rethinking original order'. *Archivaria* 66: 1-24
- MacNeil H, Mak B (2007) 'Constructions of authenticity'. *Libr Trends* 56: 26-52
- Maxwell J C (ed.) (1966) *W. W. Greg: collected papers*. Clarendon Press, Oxford
- Mitchell D (2005) *Measures of possibility: Emily Dickinson's manuscripts*. University of Massachusetts Press, Amherst, MA
- National Archives [USA] (1982) *Intrinsic value in archival material*. Available at <http://www.archives.gov/research/alic/reference/archives-resources/archival-material-intrinsic-value.html>
- National Council on Archives [UK] (2003) *Standard for access to archives*. Available at www.nca.org.uk/materials/access_standard.pdf
- Nissen H J, Damerow P, Englund R K (1993) *Archaic bookkeeping*. University of Chicago Press, Chicago
- Norton M C (1939) 'Archives and libraries'. *Ill Libr* 21(March): 11-13. Reprinted in: Mitchell T W (ed.) (1975) *Norton on archives: the writings of Margaret Cross Norton on archival and records management*. Southern Illinois University Press, Carbondale, IL, pp.86-9
- O'Toole J M (1993) 'The symbolic significance of archives'. *Am Arch* 56: 234-55
- O'Toole J M (1994) 'On the idea of uniqueness'. *Am Arch* 57: 632-58
- O'Toole J M (2006) 'Between veneration and loathing: loving and hating documents'. In: Blouin F X, Rosenberg W G (eds.) *Archives, documentation and institutions of social memory: essays from the Sawyer Seminar*. University of Michigan Press, Ann Arbor, pp.43-53
- Paskin N (2003) 'On making and identifying a copy'. *D-Lib Mag* 9(1). Available at <http://www.dlib.org/dlib/january03/paskin/01paskin.html>
- Pillow K (2002) 'Versions and forgeries: a response to Kivy'. *J Aesthet Art Crit* 60: 177-9
- Planets (2008) *White paper: representation information registries*. Available at http://www.planets-project.eu/docs/reports/Planets_PC3-D7_RepInformationRegistries.pdf
- Posner G (1994) 'Letter from Berlin: secrets of the files'. *New Yorker* 14 March. Available at <http://www.posner.com/articles/bdc%20-%20the%20posner%20file.htm>
- Puglia S, Rhodes E (2007) 'Digital imaging: how far have we come and what still needs to be done?'. *RLG DigiNews* 11(1). Available at <http://www.worldcat.org/arcviewer/1/OCC/2007/08/08/0000070519/viewer/file2484.html>
- Renear A, Dubin D (2003) *Towards identity conditions for digital documents*. Available at <http://dc2003.ischool.washington.edu/Archive-03/03renear.pdf>
- Research Libraries Group (2002) *Trusted digital repositories: attributes and responsibilities*. Available at <http://www.oclc.org/research/activities/past/rlg/trustedrep/repositories.pdf>
- Rusbridge C (2006) 'Excuse me ... Some digital preservation fallacies?'. *Ariadne* 46. Available at www.ariadne.ac.uk/issue46/rusbridge
- Saussure F de (1973) *Cours de linguistique générale*. Payot, Paris
- Schellenberg T R (1956) *Modern archives: principles and techniques*. F W Cheshire, Melbourne
- Schwartz H (1996) *The culture of the copy*. Zone Books, New York
- Searle J R (1969) *Speech acts: an essay in the philosophy of language*. Cambridge University Press, Cambridge

- Sexton A, Yeo G, Turner C, Hockey S (2004) 'User feedback: testing the LEADERS demonstrator application'. *J Soc Arch* 25: 189-208
- Shannon C E, Weaver W (1949) *The mathematical theory of communication*. University of Illinois Press, Chicago
- Shillingsburg P L (2006) *From Gutenberg to Google*. Cambridge University Press, Cambridge
- Shilton K, Srinivasan R (2007) 'Participatory appraisal and arrangement for multicultural archival collections'. *Archivaria* 63: 87-101
- Sloggett R (2005) 'Valuing significance or signifying value? Culture in a global context'. *Arch Manuscr* 33(2): 110-29
- Smith B H (1988) *Contingencies of value: alternative perspectives for critical theory*. Harvard University Press, Cambridge, MA
- Star S L, Griesemer J R (1989) 'Institutional ecology, "translations", and boundary objects: amateurs and professionals in Berkeley's Museum of Vertebrate Zoology'. *Soc Stud Sci* 19: 387-420
- Storey M (1991) 'Creeping into print: editing the letters of John Clare'. In: Small I, Walsh M (eds.) *The theory and practice of text-editing*. Cambridge University Press, Cambridge, pp.62-89
- Sutherland K (2009) 'Being critical: paper-based editing and the digital environment'. In: Deegan M, Sutherland K (eds.) *Text editing, print and the digital world*. Ashgate, Farnham, pp.13-25
- Tainter J A, Lucas G J (1983) 'Epistemology of the significance concept'. *Am Antiq* 48: 707-19
- Tanselle G T (1990) *Textual criticism and scholarly editing*. University Press of Virginia, Charlottesville
- Tebeaux E (1999) 'Designing written business communication along the shifting cultural continuum'. *J Bus Tech Commun* 13: 49-85
- Tebeaux E (2004) 'Pillaging the tombs of non-canonical texts'. *J Bus Tech Commun* 18: 165-97
- Thibodeau K (2002) *Overview of technological approaches to digital preservation and challenges in coming years*. Available at <http://www.clir.org/pubs/reports/pub107/thibodeau.html>
- Thomas S (2009) 'Curating born-digital archives and manuscripts at the Bodleian Library'. Presentation at *Archives and Society* seminar, University of London, 3 November 2009
- Varpio L, Spafford M M, Schryer C F, Lingard L (2007) 'Seeing and listening: a visual and social analysis of optometric record-keeping practices'. *J Bus Tech Commun* 21: 343-75
- Wilson A (2005) 'A performance model and process for preserving digital records for long-term access'. In: *Archiving 2005: final program and proceedings*. Society for Imaging Science and Technology, Springfield, VA, pp.20-5
- Wilson A, Platzer C (2004) *The National Archives of Australia: approach to digital preservation*. Available at http://web.archive.org/web/20070830055422/http://naa.gov.au/recordkeeping/preservation/digital/Wilson_digital_preservation_Oct2004.pdf
- Wreen M (1983) 'Goodman on forgery'. *Philos Q* 33: 340-53
- Yeo G (2007) 'Concepts of record (1): evidence, information, and persistent representations'. *Am Arch* 70: 315-43
- Yeo G (2008) 'Concepts of record (2): prototypes and boundary objects'. *Am Arch* 71: 118-43
- Yeo G (2009) 'Custodial history, provenance, and the description of personal records'. *Libr Cult Rec* 44: 50-64
- Yeo G (2010) 'Debates about description'. In: Eastwood T, MacNeil H (eds.) *Currents of archival thinking*. Libraries Unlimited, Santa Barbara, CA, pp.89-114
- Yeo G (forthcoming) 'The conceptual fonds and the physical collection'. To be submitted to *Archivaria*