would be applied primarily in artistic frameworks where their use was clearly identified to the viewer of the photograph. In the 20th century, as photographs (and later moving images) became an increasingly central part of advertising, alterations to photographs without warning to the viewer became more commonplace.[74] In the past few years, however, photography has begun to migrate to digital technology; this shift, in conjunction with the ease with which digital images can be altered, combined, or even generated from scratch by computer graphics programs has created a situation where the implicit evidentiary value of images (without regard to their provenance) has been lost.

It is not entirely clear that this loss of evidentiary value has fully penetrated the public perception, despite recent developments such as advertising which interposes modern-day celebrities into historical films in commercials, very sophisticated special effects in movies, and even virtual reality technologies to which the public is now regularly exposed. But the fact remains that it is almost impossible to tell whether an image represents a record of a real event given current technology; moreover, while twenty years ago it required a great deal of sophistication to alter images or films, easy to use software that can be employed to perform such image manipulation (and, increasingly, image creation) is now widely available on inexpensive personal computers.

This change in the meaning of images has a number of implications for digital libraries. When viewing an image, one must always harbor a certain degree of skepticism about whether the image actually represents a real event or thing and to what extent this representation may have been altered. In essence, when using images, one must constantly be concerned with the source of the image, whether the image has been altered subsequent to its capture, and the purposes for which the creator of the image is making it available. Without facilities to track and verify the source of images, they have become meaningless as a part of the historical record. 75

12. Authenticating versions and sources

In the print environment a great deal is taken for granted about the integrity of documents. If an article appears in a journal it is extremely rare that the authorship that the journal lists for the article is called into question; when this happens it is typically framed either in the context of scientific and scholarly misconduct such as plagiarism and/or results in a lawsuit. Outside of scholarly circles the issue is probably more likely to revolve around the publisher's right to publish (that is, whether the rights holder indeed gave permission, or whether the claims to rights on the material are valid—for example, in the case of unpublished archival material) than whether the author

[74] While advertising was perhaps the greatest culprit in the undermining Of the Integrity of visual images as a record of events, there were many other contexts in which altered photographs were used: politically motivated changes and sensationalistic news reporting are two other common areas.

[75] Of course, just because an image cannot be verified as a representation of an actual thing or event does not mean that the image is not of interest. Computer generated or computer altered images are of vital importance as works of art, as hypothetical models of things (for example, a reconstruction of a dinosaur) and as records of culture. The problem is that the viewer of the image needs to know the context within which to understand the image.

attribution is false. With the exception perhaps of certain tabloids one would not normally assume that there was much reason to question authorship, Similarly, print publication naturally tends to produce specific editions of a work; if two people both have the same book or issue of a journal there would be little reason to question whether the two copies of the publication were indeed identical in content.[76] It is not that a publisher couldn't deliberately publish variant copies of what is labeled as the same edition, or deliberately misattribute authorship of material, but rather that it does not happen often and when it does the publisher is typically readily identifiable and can be sued by the aggrieved parties. Further, there is little motivation (other than general malice) to motivate most publishers to do this; a publisher would have to go to considerable trouble, expense and risk in order to do it.

Perceptions and concerns in the world of networked information are quite different. It is very easy for someone to distribute information over someone else's name, and hard to trace the person who does it in most cases. It is very easy to replace an electronic dataset with an updated copy, and, since there is no automatic system which distributes multiple copies of the original version to different sites (such as the distribution of an issue of a printed journal or a book) the replacement can have wide-reaching effects. The processes of authorship, which often involve a series of drafts that are circulated to various people, produce different versions which in an electronic environment can easily go into broad circulation; if each draft is not carefully labeled and dated it is difficult to tell which draft one is looking at, or whether one has the "final" version of a work. Because of the ease with which material can be taken from one document and inserted into another which can then be circulated to a large number of people quickly, there are concerns about quotation from obsolete or draft ("unpublished") versions of a work. Visionary authors such as the late Ithiel De Sola pool[77] have written that the world of networked information would lead to the demise of the "canonical" form of documents that print publication created, and that documents would become living texts that were continually adapted and annotated. Events thus far have suggested that De Sola Pool may have overstated the case. While it is common within small groups to have people annotating drafts of a document, they are typically ultimately brought to a final, "canonical" form. Further (and it is unclear whether this is due to current limitations in information technology such as groupware and collaborative authoring systems or whether it is a more basic problem having to do with the limits to the size of a group that can collaborate effectively) such continual annotation typically occurs among a relatively small community of collaborators or reviewers, and not among the full community of interested individuals on the net. In cases where a large scale, long term collaborative effort is taking place to develop and manage a "living document" such as On Mendelian Inhetitance *In Man* at Johns Hopkins, a fairly complex and formal editorial structure is set up to manage this collaboration, and considerable care is taken to validate and track updates and their

[76] While there would be little question in most people's minds about what they could expect from printed literature, this is not necessarily the case with other media, such as videotapes, audio recordings, or computer programs, although even here it would probably not occur to most people to doubt that identically labeled copies of a work were in fact identical.

[77] See his work on "The Culture of Electronic Print", reprinted/adapted in his book *Technologies* of freedom [de Sola Pool, 1983].

sources. In a real sense, these efforts are as much undertakings to create and manage databases as they are efforts to author documents.

Interestingly, aside from a few pranks and malicious acts (most commonly people sending electronic mail with false origin addresses) it is unclear whether the fears that people seem to harbor about the deceptive and mutable nature of the electronic environment are justified by real occurrences of problems. Also, those publishers who have risked the Internet environment have had less problems with piracy than one expect given the experience of software vendors, for example. The simple fact is that several people are successfully distributing publications on the Internet today for profit (though I don’t know how much real profit they are making, they are still in business after several years in some cases). Nonetheless, it seems clear that if network based information distribution is to become a widely accepted context for the sorts of archival materials that libraries currently acquire and provide access to in print these concerns must be addressed. Certainly, the development and wide implementation of technologies and practices to address these concerns will, at the least, lead to a far more robust and reliable environment, although from a strict cost-benefit analysis it may be difficult to fully justify the costs of addressing some of the fears one hears voiced.

To clarify the focus here, it is important to recognize that while the network will undoubtedly be used extensively for transacting commerce (including, as just one example, commerce in the form of acquisition of electronic information by individuals and organizations, which may involve activities such as the identification of the parties involved, the exchange of credit information for billing purposes, the assessment of charges against some type of account, and even the acceptance of the terms of a license agreement for copyrighted material) and there is, I believe, strong justification for ensuring that these commercial transactions are conducted in a technical environment that protects the security and confidentiality of all parties, the issues involved in protecting transactions are somewhat different from those involved in ensuring that a user of the network who finds a document somewhere can verify that the authorship attribution is true and that the copy which the user is looking has the same content as the version that the author “published.” The issues of protecting network commerce generally are outside the scope of this paper, other than simply to observe that for a market in network-based digital information to develop it will be necessary to develop and implement adequate measures to protect commerce on the network and also to conduct some form of electronic rights clearance. The remainder of this section will address issues of verifying authorship and integrity of contents, and the state of the art in technologies to accomplish these objectives.

Public key crptography and various higher level protocols layered above the basic cryptographic algorithms offer methods that can be used to effectively address both of these needs. A public key cryptosystem can be used to attach a signature to a digital object in such a way that the contents can be associated with a given individual or organization at a given time. There are well-established algorithms for computing “digests” of digital objects in such a way that it is extremely unlikely that any change to the object can be made without changing the value of the digest computed from it. Thus, by checking whether the digest for an object in a user’s possession is the same as the digest value that the author has signed and makes available as the signature of the current version of the work, it is straightforward to check whether one has the same object as the author or publisher distributed. These systems offer the additional feature of non-repudiation; it is possible to include capabilities so that one can prove that a given author actually distributed a document at a given time even if that author later denies it. Such capabilities can be seen today in the Internet in the privacy-enhanced mail system [Balenson, 1993; Kaliski, 1993; Kent, 1993; Linn, 1993].

While the basic technology exists to solve the problems in question (at least as long as one is satisfied with literal bit-by-bit equivalence of two digital objects as a definition of having the "same" document, which is often really overly restrictive, since it prevents any reformatting, character code conversion or other activities that might be needed to successfully move the document from one machine to another, even if these do not change the "content" of the document in any way) the operational problems of implementing these technologies on a large scale in environments such as the Internet are far from solved. There are least four barriers:

- Standards are needed. While the algorithms are well understood at a general level, in order to ensure interoperability among implementations agreements need to be reached at a much more specific level of detail and documented in standards. Parameters such as the precise types of signatures need to be defined, along with lengths of public/private key pairs, the exact computational algorithms to be used, and the supporting protocols and data interchange formats. The IETF specifications for message digest algorithms [Kaliski, 1992; Rivest, 1992a; Rivest, 1992b] are an important step in this direction, but more work is needed. It is also important to recognize that there is more to an effective system for authentication and verification than simply algorithms; there are application protocols which use these algorithms to be defined, along with an accompanying infrastructure of service providers (see below). An additional problem that must be resolved in the standards area is the seemingly continual conflict between the standards proposed or established by the Federal Government through the National Institute for Standards and Technology (NIST) and the standards that are favored by the commercial and research and education communities .[78]

- Patent issues need to be addressed. What is currently widely accepted as the best public key cryptosystem is called RSA (named after its inventors, Rivest, Shamir and Adelman); this was patented and commercial rights to this patent are licensed, as I understand it, to a company called RSA Data Systems incorporated. Similarly, Public Key Partners holds patents to a variety of public key and message digest algorithms; to make matters even more confusing the National Institute of Standards and Technology (NIST) has filed patent applications for some of the algorithms that it has developed and adopted as federal standards and proposed licensing these (on an exclusive basis) to Public Key Partners; again, there seems to be a provision for free use of the patents for at least some types of personal or non commercial use. The net effect of these patents on basic cryptographic technology is to promote considerable uncertainty about the status of the algorithms and to inhibit their incorporation in software of all types (but particularly public domain software; while large corporations can negotiate and pay for

[78] Recent examples of this problem include NIST's controversial adoption of the Digital Signature Standard Algorithm and the Secure Hash Standard (FIPS 180), as well as the widely-denounced proposal for the Clipper chip and its accompanying key escrow system. The continued unwillingness of the government to recognize the RSA public key algorithm as a standard despite its widespread use is another example.

license agreements with the rightsholders, individual software developers or university based software development groups that wish to distribute their work without charge typically are unable or unwilling to do so), despite the relatively liberal positions that the commercial rightsholders seem to be taking on personal use and use by the research and education communities. Further, the patent filings by NIST are regarded with considerable suspicion in some quarters; there is concern that in future these patents might be used as a means of controlling the use or implementation of the technology.

- Cryptographic technology is export restricted. This has caused two problems. The commercial information technology vendors have been somewhat reluctant to develop products which can only be marketed in the United States without major export complexities, particularly given that authentication and digital signature technology tend to be very basic building blocks for distributed systems. In addition, because the world of the Internet and of networked information is very clearly viewed as a global rather than a national enterprise, system developers and standards makers have been reluctant to use technologies that cannot be freely used internationally. The issue of the justification and implications of applying export controls to cryptographic technology is a very complex one that is well outside the scope of this paper; however, the impact of the current restrictions must be recognized. In addition, it should be observed that while the position of the United States on the export of cryptographic technology is crucial because of the nation's leadership in so many areas of information technology, other nations may also have laws related to the import, export and use of cryptographic technology that also create barriers to the free use of authentication and digital signatures on a global, Internet-wide basis.[79]

- Critical mass and infrastructure. Like so many things in the networked environment, these technologies will not come into wide use unless they are available on a large part of the installed base. Authors want to communicate; publishers and libraries want to make information available. If this information is not readily used without specialized cryptographic software that is difficult and/or costly to obtain, or that cannot be used outside the United States, it is unlikely that authors, publishers or libraries will use them. While in the specific applications under discussion here of verifying authorship and integrity of objects it should not be necessary to have specialized cryptographic software support simply to view the material but rather only to conduct verification,[80] it is really more of a question of whether the investment is worthwhile because enough

[79] There is a major public policy debate currently taking place about the appropriate balance between the rights of individuals and the private sector generally to privacy on the one hand and the desires of law enforcement and intelligence agencies to be able to monitor communications on the other. While the details of this debate are outside the scope of this paper, the interested reader might wish to review the history of export restrictions on the RSA algorithms, the recent proposal by the Clinton administration for the Clipper chip, and the deployment of the PGP ("pretty good privacy") computer software both inside and outside the US.

[80] In cases where cryptographic technology is being used in conjunction with rights clearing some proposals do call for the distribution of encrypted documents that cannot be read without both software to decrypt and the appropriate key. Somewhat similar approaches are being used today where vendors will distribute a variety of locked digital information on a CD-ROM (such as programs or font libraries) and then issue decryption keys on a file-by-file basis as the customer purchases these keys; one advantage to this approach is that one can phone order the information by providing a credit card and getting a key, without waiting for physical delivery of media containing the information being purchased.

people will make use of the services. In addition, it should be recognized that there is a substantial infrastructure needed to make public key cryptosystems and the applications constructed using them work on a large scale, including key providers, certification authorities, directories of public keys, and “notary public” servers (third parties that can witness signatures and contracts, or that can record the fact that a given entity had a given document at a given time and date as a registry function). The precise details of this infrastructure will vary depending on the standards, protocols and procedures that develop in support of an implementation of the technology; how these details change from one proposal to another are not important here, but recognizing that an investment in support infrastructure must be made is vital. Further, as indicated earlier under the discussion of standards, the conflicts between federally endorsed standards and standards favored by much of the user community are having the effect of fragmenting and confusing the user community, and greatly delaying the achievement of the necessary critical mass.

It is interesting to place the issues of authenticating authorship and document integrity in the broader context of the way in which migration to a networked information environment is beginning to suggest an “unbundling” of the services that publishers have traditionally provided in the print world. Print publishers serve as selectors of material; they prepare the material for distribution, distribute it, manage rights and sometimes royalty payments, and authenticate authorship and integrity, among other functions. In the network environment, it is clear that distribution (at least through a mechanism as crude as making a document available for anonymous FTP) can be done by anyone. It is clear that services that help people to identify material of interest such as abstracting and indexing services, reviewers, bibliographers, and ratings servces are likely to play an enlarged role in the electronic environment, and that these services can be quite separate from the persons or institutions that make material available. It may well be that authenticating and verifying the integrity of a document is at least optionally a separate, and separately priced (and perhaps costly!) service from simply obtaining a copy of the document.[81] If so, it will be interesting to see how much use is made of such services (outside of some specific environments, such as litigation, where cost is typically not much of a factor and such issues simply must be unambiguously established) and particularly the extent to which people are willing to pay to allay their fears about the electronic information environment.[82]

[81] The integrity of published works is not an entirely new problem. A number of publishers currently provide loose-leaf services for areas such as tax law; determining whether a user has the most current version of such a service is today an important question with a potentially high payoff.

[82] It is also necessary to consider other implications of establishing a chain of provenance for an electronic document. As discussed previously, technology for tracing the provenance of a document is well established, and depends on sophisticated cryptographic technology. It seems likely that US government, and perhaps other governments have agencies that are monitoring most international traffic, and that any encrypted traffic will attract their attention., since at least at present it is relatively rare In some nations use of cryptographic technology may be illegal, either across international boundaries or even within the national boundaries. Even if it’s not illegal, it may attract attention. Are scholars prepared to attract the attention of such communications security agencies as a consequence of maintaining verifiable ties to the scholarly record?