DRM: Doesn't Really Mean Digital Copyright Management

L Jean Camp
Associate Professor of Public Policy
Kennedy School Of Government
79 JFK St
Cmbridge, MA 02138
1-617-496-6331
Jean_Camp@harvard.edu

ABSTRACT

Copyright is a legal system embedded in a larger technological system. In order to examine the functions of copyright it is critical to examine the larger technological context of copyright: analog media and printed paper in particular. The copyright system includes both the explicit mechanisms implemented by law and the implicit mechanisms resulting from the technologically determinant features of paper and print. In order to prevent confusion between the legal, technical, and economic elements I refer to the whole as "copy accurate".

Digital rights management design should explicitly address legal issue in copyright and economics of paper, technology of mass produced analog media, and print culture. An examination of that entire system (copy accurate) yields a return to first principles for the design of digital rights management systems.

1. INTRODUCTION

In a world where paper is no longer the medium copyright no longer protects the message. Changes in that medium to other analog media – wax recordings, vinyl recordings, magnetic analog tape recordings - have been by gradual extensions of copyright. Yet digital media are different in a fundamental manner – a change in kind as opposed to a change in degree.

In designing systems to implement copyright in the digital age, that change in medium should be taken in account. Currently the change in media is taken into account when those things either potentially lost or inherently threatened are identified as revenue schemes for digital content owner [1].(Burk, 2001).

Digital Rights Management standards are being developed for digital content. Yet the question remains, what problems are the systems trying to solve? Most systems are explicitly modeled on copyright (e.g., [2] Piva, Bartolini, & Barni, 2002, [3] (Rowe, 2002) Those that use the metaphors of piracy and author's moral rights to define or defend the design goals (e.g., [4] Bolic, 2001) depend on the underlying philosophical support of copyright. Yet the debate needs to be expanded beyond the borders of copyright. Thus this examination is a utilitarian one, with a focus on the functions of copyright law when embedded in the economics of analog media and based on the technologically determined features of the press. To consider copyright for the digital right requires considering the nature of printed paper.

This paper is framed by the Western European experience of copyright; yet by focusing on function rather than motivation the results may be more widely useful.

Paper and analog mass reproduction technologies bind attribution to content, thus enabling referencing and attribution. Mass-produced analog content also has high levels of integrity not only because of the difficulty of altering analog content but also because of the widespread distribution. . This combination of technology and law enabled epistemological surety and literacy, as well as a functioning information market for the reward of authors. In order to prevent confusion between copyright (a legal regime) and the combination of law, economics and technology at the dawn of the print age I refer to the latter whole as the *copyright system*, and to the sets of functions the three create together *as copy accuracy.*

After illuminating the utilitarian aspects of the copyright system, I propose a set of functional requirements corresponding to these functions. I then use these function requirements to examine three digital rights management systems: Giovanni, the Content Scrambling System, and the Adobe E-book. The well-known tools that break both CSS and Adobe E-book are also included in this discussion.

This work was supported by the National Science Foundation under Grant No. 9985433. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

I conclude that while each of the DRM systems and the systems that defeat DRM has elements of the copyright system, none solves the set of problems the copyright system solved for the dawn of the print age. I further argue that were information all free, the information market might fail. "Copy accuracy" consists of the legal principles of the copyright system in 1710, the technologically deterministic elements of printing press technology, the economics of mass-produced analog content all of which are embedded in print culture. I close by proposing a return to copy accuracy as a basis for DRM design.

2. BACKGROUND AND MOTIVATION

Ownership systems for information are inherently social and political. [5] (Branscomb, 1995). Critics of digital rights management systems copyright have been concerned with the democratic implications of DRM Democratic concerns include fair use, [1] (Burk, 2001), conflict between property rights and speech rights [6], [7], [8] (Lessig, 2001, Litman, 1997, Netanel, 1996), the dangers of anti-circumvention laws [9], [10] (Samuelson, 1999; Lessig, 1999), and the threat to privacy. [11] (Cohen, 19960). Others have criticized the underlying economics of the extensions of copyright. [12], [13], [14] (Gordon, 1992; Gordon, 1997, Litman, 2001)

I also do not trace the history of the trade-off between ever more broad information protection and ever more specific definitions of fair use. [15] (Vaidhyanathan, 2001). Nor am I addressing the sources of intellectual property [16] (Fischer, 1999). The various ideological arguments over the theoretical economic models that underlie copyright are of tangential interest [8] (Netanel, 1996).

A critical issue and ignored issue is that copyright is and was a legal regime embedded in a technological system, just as deeply as motor vehicle controls embed the assumption of the automobile. This copyright system (as opposed to copyright per se) includes technical, economic, and legal elements that together serve more purposes than the law alone serves.

Given the fundamental technological blinkers on the current debate, this work begins with an extremely brief consideration of the nature of printing and paper. Readers are urged to examine the dominant primary sources for this work [17]., [18], [19], (Eisenstein, 1979, Castells, 1997, McLuhan, 1997;). Here I argue for the conception of copyright as part of a larger technological and legal system, rather than a legal system alone, for *copy accuracy*.

In the second section I argue that copy accuracy includes binding of reputation to document, content integrity, document persistence, intrinsically enabling availability and archiving. I then consider the new technological and economic system in which copyright is failing.. The fourth section is concerned with applications of the assertions in Section 2 and Section 3 on selected DRM systems. I conclude that while free as in beer may not be ideal for digital content markets, the current DRM designs are far from serving the functions of copy accuracy.

3. COPYRIGHT AND ITS CONTEXT

Copyright as a legal construct was created at a time of dramatic changes in the economics and technology of information. In particular copyright was created after the diffusion of the movable type press. In this section I focus on the three sets of changes technical, legal and economic. I do so in order extract from the technical and economic discontinuity the critical functions that have made copyright successful.

Once the fundamental contours of copyright had been sculpted the form proved so potent as to sweep the globe, along with the movable type press.

Copyright has been altered significantly in the past 300 years. In order to avoid the debates on the legitimacy, wisdom or efficacy of the incremental changes I focus on the moment on creation: 1710. At that time the technology of mass-produced analog content had passed the point of incubation and been widely diffused.

I argue implicitly (by method) and explicitly that a new approach is needed, one which performs the critically important functions of copy accuracy but in the fundamentally new information market.

3.1 Technical

Technological change has long been the driver for alteration of copyright. In fact, technological change is the basis of creation of copyright. As a class of policies, information property is the creation of the industrial revolution. Previously all information belonged to Crown or Church, with the rare author receiving some ownership. [20](Rose, 1993). The ability for individual tradable ownership of authorship and information rights were hotly contested and explicitly part of the debates and revolutions which raged across Europe in the eighteenth and nineteenth century. Individual ownership of ideas is in direct opposition to the absolute state.

Before the invention of the press, authoring was simple compared to copying. Copying enabled the survival of information from the Summarians, through the Roman and Byzantine empires, through Arab universities and finally to Europe. To copy a work was in no way theft but rather the only way to save a manuscript. Without laborious copying any authored document would be lost. Each book copied consisted of a set of articles selected by the copier. To copy was to edit, in that modern editing is the selection and ordering of material for inclusion. [17] , [21] (e.g., Eisenstien 1979; Febvre & Martin, 2000).

Copyright was created a moment of great discontinuity. Before the printing press, content was very expensive to produce and even more expensive to distribute. Effectively there was no distribution, excluding the occasion loan for copying. Before the printing press, to copy was to preserve. Every document was unique in the content between the covers. To fail to copy a document was to resign it to decay and loss.

Draft. Later works published as: L. Jean Camp, "First Principles of Copyright for DRM Design," *IEEE Internet Computing,* Vol.7, No. 3 pp. 59-65, May 2003. Previous version as "DRM Doesn't Really Mean Copyright" in the peer-reviewed *Proceedings of the 2003 ACM Conference on Computer and Communications Security*, ACM Press (NY NY).

The printing press enabled mass creation and necessitated distribution. Distribution and reproduction were intimately related in the nature of duplication via printing press.

Like ICTs today, the printing press changed the economics and politics of information in early modern Europe. Any literate middle-class person could author and publish a leaflet and have it distributed across London, Paris, or Berlin. The nobility decried the opinionated chatter of the merchant classes to no avail. After the defeat of the Royalists in England, those who controlled the equipment for data reproduction no longer had exclusive rights. Competition reigned the control of content was also lost. [22] (Holdsworth, 1938).

Before the press the media was relatively expensive, isolated copies meant ease of loss, and isolated handwritten copies could be altered more easily than printed widely archived copies. The expense of making multiple copies made archiving knowledge difficult, and the nature of hand selected copies resulted in difficulty in referencing and validation. Every copy was less than the original as scribal drift created a centuries-long decay of content not unlike the modern parlor game of telephone.

Even that most basic ordering – alphabetical order – was in no way standard in Europe before the printing press. "Amo comes before bibo because a is the first letter of the former and b is the first letter of the latter and a comes before b ….by the Grace of God working in me, I have devised this order." [17] (Eisenstien, 1979). To copy was to be the agent of the author, to serve the author's greatest interest. Not to copy was to resign the authors' words to certain destruction. Only after the printing press was invented was there a conflict between copier and author.

Printing creates multiple identical copies. Once distributed making changes in every copy is extremely difficult. In addition the physical nature of paper makes changing each individual copy difficult to change. In the parlance of computer security paper presents a high work factor to change in a manner that is difficult to detect. In order to make a convincing change it would necessary to locate multiple (perhaps most) copies and alter them in an identical manner.

Increased distribution increases surety. For example, for reliable timestamps [23] (Haber & Stornetta, 1991,) the outputs of hash trees are intermittently published in the NY Times. As the NY Times is widely archived, it would be inconceivably difficult to alter all versions of the published material. Similarly if a document exists in many archives not even the author can go back and alter the content without detection. (In contrast, when there are improved results or clarifications in papers on-line then the publisher or author can improve or correct in real time. In a paper version the improvements correspond to a new edition number or revised publication. Therefore after a claim becomes substantiated or an error corrected, referencing functions on paper can fail in the digital network.)

Mass production of printed material (or indeed any mass-produced analog material) and distribution of the material together create integrity in content. Furthermore, when the mass production includes attribution the linkage of attribution and content is strong. The linkage of attribution to content is used today [24] (Stallman, 1984) to enable an explicit reputation right.

Related to integrity is the issue of personal annotations. When alterations to the document are easy to detect, then annotations are clear. While it is possible to use version control and make annotations clear in a digital document such a result requires explicit design effort. In contrast, the ability to make personal annotations, including highlighting passages, marking pages, and adding comments in the margin, are integral to print technology.

Printed paper creates editions. That is, each edition is better and more polished than the last. Every edition offers confirmation of the first (in comparison with decayed individual content). Editions identify dates, including first editions where there is no second. Any edition can be referenced. Before the press, referencing required copying exact material with context, as no audience was likely to be able to check an original. Page numbers, editions, and the ability to reference a work such that a distant colleague can validate my assertions are a function of the press. Building upon the innate nature of the press, in that it produces identical copies identifiable at specific times and place, reference and review as used in modern scholarship were enabled.

Annotated quotations, peer-review based orderings, and claims of previous proven foundation via reference require some archival trusted information mechanism, integrity in documents including attribution, and some ordered manner to search the archive. In the case of the printed book the page itself and the nature of repeated printings provide the possibility of this certainty.

Now copyright is so deeply embedded into practice and thought that its success makes difficult the struggle for the appropriate response to the new global shift in the technology and economics of information

3.2 Legal

Copyright was created to address the issue of ownership and payment in a media that was relatively cheap, reasonably permanent, subject to references, straight-forward to archive and difficult to alterations without detectable trace.

At its creation, copyright was essentially an extension of freedom of communication, in that it replaced the Star Chamber, and fundamentally a mechanism for control, in that it required deposit in the Royal Library. Before the Copyright Act the dual goals of regulation of press had been (often brutal) control of the content and exercise of political power for enrichment of the powerful few. First the printing guilds, then under the Stationers Act, the booksellers (as merchants overtook tradesmen in power), and always the Crown controlled and profited.

The Parliamentary refusal to renew the Stationer's Act (at the first opportunity of Parliament to so refuse) resulted in a collapse of controls on licensing. Unique in the case of England a dual judicial system (the Star Chamber as well as the criminal and civil courts) had created a set of sometimes-incompatible rulings. There was no single body of judicial findings on which to build. Printers were printing and selling

as they wished. The information market was not functioning. Even the most ardent opponents of the Stationers' Act recognized the need for some form of commercial structure.

The Copyright Act was, "an Act for the encouragement of Learning, by vesting the copies of printed books in the authors or purchasers of such copies during the times therein mentioned." [22], Vol. X.

The Copyright Act was radical in the following aspects:

- The right went to the author, not to printer, bookseller or Crown.
- The right had a finite term (14 or 21 years). Previously patents or privileges did not expire.
- Violations of the Act were civil violations, never criminal violations.
- Multiple depository requirements were created.

Arguably the university depository requirements in the Copyright Act were not entirely radical. Those libraries (Oxford, Cambridge, Sion College, the Royal library, four unnamed Scottish Universities, and Edinburgh) were also notable in that the presses of the Colleges were members of the Stationers Guild previously. (S. Anne, Article 1, Section 8, Clause 8, Section V.)

Depository laws have an established history, dating from copy requirements of monarchs seeking to build personal/national libraries [25](Harris, 1995). (The division between a royal library and national one is difficult, as not only was the crown the state but also because some royal libraries were open to scholars while some national libraries strictly limited access.) Yet depository laws very rarely required deposit in more than one institution. Thus the depository requirement for the Royal library is in the traditional of depository requirements while the university requirements are something quantitatively and possibly qualitatively different.

The creation of the right to own an expression created a market right. It can be argued that the Copyright Act as signed by Queen Anne (thus becoming part of the Statute of Anne) did create the concept of fair use, requiring nine copies of each book so that each of the major universities could each have a high quality copy available in their libraries. Yet the counter argument, based on the role of libraries in the early eighteenth century argues against this interpretation, argues that fair use was a later construct. What is certain is that the requirements for deposit simplified the creation of scholarship. The ordering of content for scholarship and debate was made possible by the printing press.

(For the changes in copyright law with respect to changes in analog technology, see [6], [26] (Lessig, 2001, Sterling, 1998).)

3.3 Information Markets

The effects of printing are so ubiquitous as to be invisible. Again I refer the reader to the superior histories of the press and print noted at the beginning of this section.

"Information wants to be free" as a phrase may have originated in Stewart Brand's WELL (Whole earth 'Lectronic Link) but the concept was instituted (briefly, and as a notable failure) in revolutionary France. [27] (Darnton, 1982) The licensing of the pre-Revolutionary regime of the resulted in excessively controlled debate resulting in a degradation of dialogue provided by the underground. In the case of the French Revolutionary period the adoption of copyright followed a reign of completely free information. In the records of the French revolution the popular political discourse suggests that an excessive control of information or a complete lack of control of information both result in a dysfunctional information market leading to all discourse plummeting to the pornographic: truly the lowest common denominator. [27] (Darnton, 1982)

Britain similarly suffered a dysfunctional market after the refusal to renew the monopolistic Stationer's Act. In Britain, authors were in arms and publishers had difficulty securing copies of works from abroad. Printed paper, like other forms of information storage and transmission, created fundamental problems of economics and reliability of information. Copyright solved these problems for the printing press so well that the solution functioned with following analog mass-produced media.

In the networked digital age copying is costless. Alteration of documents and redistribution is trivial. [1]

The digital network means that copying and distribution are now extremely low cost [28](National Academy of Science, 2000). To distribute is now to copying, due to the inherent technology of digital reproduction.

From being inexpensive copying has become nearly free. Distribution costs have been reduced to almost nothing, a course of affairs which will continue at a rate surpassing even the decrease in cost of computing power [29](Odlyzko, 2000). Yet the functions of copyright are still necessary: production can be expensive either monetarily for a Hollywood studio or in sweat of the brow from a single creator. The simplification of misappropriation of goods does not imply that such a state of affairs creates a functional market.

A related issue now and in past implementations of information technology is the issue of epistemological surety. With widespread misappropriation, how does a reader know to trust the content, the attribution, or the consistency of a document? Paper addressed surety via distribution, archiving and technological determinism. There are no such mechanisms on the digital network.

[1] I recommend primary sources for further discussion. [40], [41], [28] (Shapiro & Varian, 1999; Whinston, Stahl & Choi,, 1997, National Academy of Science, 2000).

In terms of the function of copyright there are two further distinctions between paper and digital information. First, reputation is bound to paper. Second, distribution *is* effectively archiving in a paper medium. Third, filtering is integrated into publication.

When an author's name is added to a paper document removing it is difficult, and leaves traces of the action. The author's name stays on a book: on the cover, the title page, sometimes even the pages. It takes significant effort to convincingly alter authorship information on a book or a paper and redistribute. In contrast, maintaining author information is difficult in digital transmission, requiring the retransmission of an entire document and often even the difficult construction of the context of the document (e.g., URL, author, date visited).

Trusted third parties may be new in that cryptographers implement them using write-once optical media and digital cryptography but the concept of a trusted third party is an ancient as the library or record-keeping temple. While self-validating contracts were used for transactions that were critical only to the parties to the contract, third party validation was used for critical information. The earliest libraries were temples and then royal archives, where information for greater social importance was stored. The seat of government or the highest temple would hold the canonical document and copies could be provided to those institutions lower in the hierarchy. In this way there was a trusted third party where those concerned with the integrity of information could be validated [30] (Lerner, 1998).

Distribution is archiving in paper distribution. Books are distributed in a format that requires conscious decision to store or remove. Storage is simplified with books. Newspapers and books when distributed are distributed to individuals who might store them and to libraries that hold archives. In contrast, distribution via web sites requires distribution into caches that are automatically cleared. Saving requires conscious decision and deleting is effortless. Saving with contextual information for referencing is even more difficult. The critical problem of persistence is being addressed with great optimism but little result via the Web Consortium project the Semantic Web. [31] [(Berners-Lee, 2002)

Filtering is integrated with publishing, distribution, and copying in the world of the printed page. Editors choose the material to publish, and merchants choose which to provide. The fundamental physical nature of the printed page makes filtering a requirement for physical production and distribution just as the finite number of possible pages requires the selection processes of journals.

The issue of filtering is radically different. Herb Simon most famously coined the concept of an attention span economy. As the sheer volume of information increases those services, software, or practices that reduce information flow are becoming increasingly valuable. Corporations seek to decrease the information flow to consumers using affinity marketing and personalization. As an example, weather.com offers information for three selected zip codes versus offering every user the entire database of global weather conditions. The value of all the weather data on the globe is so much as to be of no value if I cannot sort it to determine if there will be rain on my picnic.

Now filtering requires more participants. Increasingly filtering is used to select distinct information for individuals rather than trusted selections per se. Compare the selection of texts for temple inclusion, to the library selection process, to the interactive selection processes of Amazon. In the previous two cases specialists make selections for a wide number of people based on professional training. The relative value of the information they selected enabled the pre-press library to survive. Depository laws, community support and democratic values enabled press-based libraries to survive. What is the value of filtering and what additional filtering models must be supported for the information economy to function in the digital networked age? And more to the point, how do current DRM technologies enable or prevent the filtering and rating function?

Currently libraries continue to receive public support but are overwhelmed at the cost and difficulty of archiving on-line material. Archives (such as the WayBack machine) are prohibited from providing archival access by copyright – an essential feature of the analog copyright system. In addition, individual users develop following with web pages, blogs, personal selections, and electronic subscription or newsletter services with few avenues for archival storage. Of these, arguably a blog is itself some sort of archive given the temporal ordering of items of interest to a selected community, although it is most compelling compared to the pre-museum curiosity collections [32], [33] (Rodzilla, 2002; Blood, 2002).

Thus in addition to the reputation market, monetary market, certainty of access/archiving, referencing systems, and binding of content to creators, the next copyright must function to allow the rewarding of filtering or to explicitly provide support for a filtering market. This is not a set of trivial design requirements. To examine the state thus fare, I examine four digital rights management systems.

4. COPYRIGHT TO 'COPY ACCURATE"

For three hundred years, dialogue thrived, the scientific revolution raged, and copyright established itself as the rational mechanism for content control across the globe. Vastly different cultures adopted copyright. Over the last century new technologies, such as the phonograph, radio, motion pictures, and even high-speed presses, have monotonically increased the expense of producing and distributing information, leading to fewer publishers demanding greater copyright protection. Each new technology has lead to new variations or clarifications in copyright. [14], [20].

Technological change resulting in changes in the economics of information has yielded changes in copyright. Phonographs were the first challenge to the functionality of copyright. When there was no medium for recording song or voice the author could be assured of payment when the music was reproduced via sheet music sales. Similarly the author of a play could obtain reimbursement for small productions from the value of the scripts sold. Yet the ability to record song or play changed that equation. When the movies came along there were new concerns about ownership. Both those who rejoice in and despair of digital networked technology identify it as radical and as ubiquitous a change as networked digital information. [18], [34], [35] (e.g., Wade, 1997; Castells, 1997; Beniger, 1989). In short, we are

again at a place not unlike the period of discontinuity where copyright was created. This discontinuity is so great as to require more than a clarification of copyright consistent with the technological (and legal) trajectory of the past century.

I consider copyright as created at the discontinuity. Essentially I seek a return to first principles to guide the design of digital rights management systems.

The problems of authentication and document integrity studied by scholars and legislators today study are not new, but ancient. Trustworthy information is not a new problem. Reliability of information, currently referred to as integrity in computer security, has an ancient history. In Babylon clay tablets were the mechanisms for recording information and contracts. Where there was significant concern that one holder of the contract might alter the clay on which the details were impressed, four "copies" of a contract were made. (While our concept of exact copies does not apply, the critical details were the same.) First two of the copies were baked. Then these two copies were encased in the clay of the other copies and baked inside, illustrating a need for self-validating contracts. [25] (Harris, 1995)

Because of the importance of persistence, integrity and reputation I propose that an implementation of the functionality and coherence of the analog copyright system for networked digital copyright be called copy accurate. Copy accurate ideally would enable the same functions of reference, filtering, persistence and surety as the copyright system when grounded in the mass copying made possible by the press. Copy accurate is technology grounded in the mass publication made possible by the digital network.

Practical application of the fundamental practice of peer review requires multiple copies of an identical document. Yet the printing press can only provide this function if there is an archive and a regulated way of identifying authorship, as well as enforcement functions to prevent widespread plagiarism. Copyright and the press *together* enable referencing and simplified the creation of communities of shared knowledge.[8], [17]. (Netanel, 1996; Eisenstein 1979).

Because of the integrity of a paper document, when the Copyright Act created a market right, it created two market rights. One right was the right to copy and resell the material; that is a physical or monetary right. The other right was the technologically-given right to have ones name associated with the material, derived from the right of authorship. The problem of illegal copying was distinct from the problem of usurpation even in the earliest days of the book [26](Johns, 1998). This continues today, as illegal copying is distinguished from plagiarism. (Should a student turn in a false paper it is no defense to say that a fair price was paid for the material.)

Recall that paper, print *and* copyright in 1710 made it possible to own rights to expressions of information. Ownership prohibited plagiarism from any work by extending the scope of works covered to all ASCII content. Access, persistence and integrity made information reliably available to royalty and scholars.

In summary then, here are the functions of the copyright system as embedded in the technology of the printed page. Here is copy accuracy:

1. Market
    a. Reputation
    b. Monetary value
2. Epistemological trust
    a. Persistence
    b. Content integrity
        c. Reputation and context integrity
3. Personalization
    a. Filtering
    b. Personal annotation

There are two elements of copyright law that must be acknowledge to prevent outrage by legal scholars. These are:

4. A human right of expression
5. A moral right

I have explicitly chosen not treat the third and fourth human rights aspects. It is my contention that the human rights elements do not provide guidance in addressing the development of digital rights systems, especially since these two rights can be in conflict. Kant created the concept of copyright as a human right in an essay on unlawful publishing in 1785 that used natural law to argue for moral ownership for the production of one's mind. [26] (Sterling, 1998) The argument for author's right supposes that the work bears "the mark of the author's personality". Ownership of the fruit of intellectual labor is now widely regarded as a human or cultural right. Freedom to access information and privacy rights are also human rights, confusing the matter further in the case of DRM. [37](United Nations, 1995). I include this to acknowledge the existence of these rights, and to identify that these do not make useful design guidelines, despite their vast fertility as ground for debate.

Given that the elicitation of author's rights followed the creation of copyright by some decades, and the American utilitarian insistence on rejecting human rights notions in the Berne Convention as late as 1956, I find this position defensible as well as practical.

Furthermore my focus is exclusively at the point of technological and economic discontinuity, where the economics of the press began to dominate the economics of information. Human rights based definitions of copyright, and the acceptance of those definitions, lagged the utilitarian functions.

I therefore propose copy accuracy.

5. DESCRIPTIONS OF THREE SYSTEMS

Digital rights management systems attempt to implement the copyright system in the digital realm. Changes in the technology from analog to digital offer users some new options, such as creating back-ups, while potentially removing other, such as ease of archive, personal mark-up and integrity. An implicit question answered by DRM design decisions is on what properties of paper are valuable enough to reproduce in DRM and which can be removed or made subject to payment. Those implications of design decisions are made explicit for three designs in this section.

In this section I describe three digital rights management systems: Giovanni,[2], Content Scrambling System[4], and the Adobe eBook[3]. Noticeably this requires examining DeCSS[4] and the Advanced eBook Processor[5]. In the use of the word secure, in no way do I intend to imply that all the systems here provide cryptographic security, only that security is the design goal of the system.

5.1 Giovanni

Blue Spike offers a suite of products, so my focus here will be on Giovanni. Giovanni is Blue Spike's digital watermarking technology. Giovanni can be used for identification, authentication and auditing of digital audio works. Giovanni is often classed as mechanism for protecting audio content, and indeed there is a company focus on making the watermarking inaudible. Yet Giovanni can be used for any content; and it is the explicit goal of Giovanni is to be available for all media.

In the Blue Spike model a producer will create content and then mark it as his or her own. This information will be presented for sale at the Blue Spike server.

The creator or owner of the content selects attribute data to be embedded in the content. The resulting secure content is then stored and made available over Blue Spike's servers.

Giovanni begins by generating a single random number from Giovanni's key and the seed for a hash function. This random number is divided into two segments, r_0 and r_1. r_0 is hashed with the attribution data to create a payload the correct size to data to embed, thereby creating the secure content. Then r_1 is used to determine the placement in the content of watermarked data generated with r_0. Thus the payload is encrypted and embedded into the content.

Blue Spike has a model that requires author registration. Blue Spike offers to manage the enforcement of the content as well, by searching for Giovanni signatures across the network.

Blue Spike includes the option of embedding purchaser sensitive information in addition to producer information in a good. It is not exactly clear from the documentation how this is done, presumably the same symmetric keys are used and the purchaser information is added to the hashed payload. In the case of the producers, the identifying information may be pseudonymous as long as the producer can prove him or herself the rightful owner. In the case of the customer the information is based on payment information meaning that the data embedded are personally identifiable and therefore sensitive.

5.2 Content Scrambling System

The content scrambling system (CSS) is the standard for the content protection system architecture (CSPA). As such, CSS is embedded into DVDs and DVD players. In the case of CSS the content is the DVD movie itself, and the metadata consists of a region code.

Region codes are a decimal digit that determines in which region a DVS can be sold. The region code prevents regional arbitrage by buyers and allows large-scale geographic price discrimination by merchants. Therefore a primary function of CSS is regional price discrimination[6].

CSS encrypts the contents of a DVD so that only approved readers can access the code. A key that decrypts each DVD is stored in any licensed DVD player. Every DVD player has a small set of player keys (in case one key should be compromised).

---

[2] Blue Spike, www.bluespike.com/giovanni/gdigmark.html

[3] Adobe eBook FAQ

www.adobe.com/support/ebookrdrfaq.html

[4] The DVD FAQ http://www.dvddemystified.com/dvdfaq.html

[5] Elmsoft www.elcomsoft.com/

[6] I use the phrase price discrimination in the economics sense, implying no moral wrong. In fact, there is a strong economic argument that for high fixed cost, low marginal cost goods price discrimination is necessary for the market to function. A most common form of price discrimination is the "stay over Saturday night" used by airlines to discriminate between corporate and vacation travelers. Notice that air travel providers are in a high fixed cost, low marginal cost business.

Every DVD is encrypted with a key, called the title key.

Each CSS-protected DVD begins with a hashed disk key (5 bytes). After the hashed value, the full disk key is then listed encrypted in every possible player key. There is a set of 409 player keys.

Every CSS licensee is given a player key. Thus if a CSS licensee implements an unacceptable player, the license can be revoked by removing the corresponding encrypted disk key.

Assuming the player has a valid key, the player confirms that it is using the correct key for the given disk by hashing the decryption of the disk key. The hashed, decrypted key should be equal to the 5 bytes at the beginning of the CSS block.

Once the disk key has been determined, the DVD player uses the disk key to decrypt a title key. The content is either encrypted in the title key or encrypted in a permutation of the title key and the list of encrypted disk keys.

DeCSS breaks the encryption provided by CSS so that a DVD can be decrypted and played on unlicensed players. This undermines the licensing strategies and the value of the region code. A primary function of DeCSS was to allow users to play DVDs on the Linux operating system using open code[7].

All players limit the number of times a region code can be altered. The Macintosh OS X DVD player limits the change to three times. Other players limit the number of changes five times. The number of times a player can allow the user to alter the region code is a function of the license associated with the DVD player key. Should any player manufacture a device that allowed arbitrary region code alterations then the license of that producer would be revoked.

5.3 eBook

The Adobe eBook software provides for digital encryption of content and is associated with a series of sellers who agree to provide copies of an eBook-protected book for resale. eBook merchants provide conversion to digital format, author rights management services via eBook and offer to provide distribution. Different providers offer different bundles of services associated with eBook. However eBook can be purchased with server software so there is no archival requirement with eBook services.

As with PDF, Adobe provides a free eBook reader compatible with the eBook digital rights management system.

In contrast with paper books eBooks expire.

In contrast with other digital books Adobe prevents cut and paste of significant sections. Adobe eBook also prohibits the use of text-to-audio readers.

The Advanced eBook Processor directly addresses the core question of which of the characteristics of paper books are worth preserving in the digital realm, with the eBook and the Advanced Processor having fundamentally different answers.

6. ANALYSIS

In this section I bring together the functions of copyright and the functions of the digital rights management systems. A hypothesis that must be addressed for this analysis to be useful is that all illegal copying substitutes for purchasing. While there is evidence to the contrary [38], [39]. (Osorio, 2002, Pahfl 2001) the concept of illegally copying as revenue lost will be assumed in the tabulated analysis. Then in each discussion the question will be briefly revisited.

If all illegal copying is a direct substitute for legal copying the obvious implication is that the value of increased access is negligible. There is simply loss of revenue for the author. If illegal copying functions as free advertisement, and encourages additional purchases then increased access yields increased revenue. If illegal copying yields widespread awareness of a work and thus increases distribution but not sales, then the author losses monetary value but gains reputation value.

There is a second conflict is evident between the functions of copy accuracy: how does availability alter the veracity of a document? When there is an increase in availability and a decrease in security in a document, what dominates? If the content are not available, or the availability is strictly controlled, then the useful functionally of surety is limited as fewer can learn from or build on the information. A timely example is the alteration of *E.T.: the extra terrestrial* for re-release. In the 2002 release potty humor was deleted, and police are seen with flashlights instead of guns. These digital changes would be misleading to any scholar using E.T. to study the culture of the seventies. These changes were made possible by the existence of few master copies. Similar controls of documents and copying technology allowed Stalin to alter official historical documents when previously favored officials were purged. In both cases, the alteration is detectable by virtue of copies in archives beyond the reach of the editor.

Having considered these common issues I now consider the systems described above according to copy accuracy. In each cell the function of copy accuracy is listed in the first column. Along each row the ability of a DRM system to fulfill that function is listed in the corresponding column. The increase or decrease in the ability of a system to support a particular function is noted by a minus or plus sign, respectively.

---

[7] Since a fundamental element of open code is the ability of users to alter the code, no provider of open source players can enforce a limit on the number of times the region code is altered.

The monetary incentive reflects the creation of a primary market. Recall that the author's incentive consists of both a direct right over the content as well as a reputation right. Attribution refers to the binding of the reputation attribute to the content. Archiving, attribution and access are inherent technologically deterministic elements of the paper-based copyright system.

Table 1. The eBook Compared to Paper and the Advanced eBook Processor with respect to the eBook

Functions of Copy Accuracy	**Adobe eBook**	**Advanced eBook Processor**
Tradable good	-	+
Tradable right	+	-
Attribution and Integrity	+	-
Persistence and Archiving	-	+
Access	-	+
Personal Annotation	-	+

Advanced eBook Processor enables the right of first sale (or secondary distribution) and illegal copying. The increase or decrease in market value is a function of the nature of illegal copying as substitution or complement for legal copies.

Note the Adobe eBook confirms the features of digital content that are a loss to consumers while negating the gains for consumers. Specifically the eBook removes first sale, archival storage, and the ability to mark passages. The eBook deletes the right of first sale and related rights (rentals). The eBook essentially takes all alterations of the copy accuracy resulting from technological change and negates those granted to the consumer.

Given that Advanced eBook Processor removes the tight binding between author and content, the implication is that eBook decreases reputation value. However this is subject to the issue of distribution as opposed to encryption as surety mechanisms.

In terms of access eBook arguably increases access in that it provides a portable format and support for creating digital books. However, in the long term, eBook decreases access not only because of explicit expiration of content access rights but also because of changes in format. Advanced eBook Processor increases access and availability because it prevents expiration and ensures availability when the Adobe-owned format is altered.

The eBook prohibits personal annotations. The Advanced eBook Processor does not prevent annotations but adds nothing to simplify highlighting, marking, or visible margin comments.

In contrast to other on-line book formats, eBook prevents cut and paste of significant sections. Of course, the centuries old tradition of transcription for plagiarism and illegal copying remains available. Thus in comparison with paper texts there is not effective change between an electronic book and a traditional book with respect to authorship and reputation value.

Table 2. CSS, DeCSS and Copy Accuracy

Functions of Copy Accuracy	**CSS**	**DeCSS**
Tradable good	-	+
Tradable right	+	+/-
Attribution and Integrity	+	-
Persistence and Archiving	-	+
Access	-	+
Personal Annotation	-	neutral

CSS does not prevent bulk reproduction and resale of content. When content are mass-produced the only issue would be that the region code remains. This means that a large-scale commercial production of illegal copies must purchase originals in the target market. In effect, this prevents consumers from purchasing illegally copied goods in another region and returning home to sell them. Also, by making it more likely that an illegal copy will not work in a CSS-compliant player, CSS may function to reduce demand for illegal copies of DVDs.

CSS increases the monetary value of the DVD because it allows regional price discrimination. CSS limits this value as the secondary market is decreased because first owners cannot resale across boundaries. If this constrains the right of first sale, there is a possible increase in monetary value to the owner.

DeCSS decrypts the content on the player, allowing excerpts and therefore misappropriation. It increases each of access and availability, In addition, it creates the ability to make derivative works. Unlike the case of books direct transcription is not an option in video content. Therefore DeCSS may increase the .value of the tradable right if it increases the use of a good for building other information goods. This reuse will also increase the reputation value.

CSS prevents annotation. DeCSS does not prevent annotation by design, but neither does it enable such annotation.

DeCSS enables archiving by creating an unprotected bit stream and allowing any player to be used. DeCSS increases access by allowing initial cross-border trades and an expanded secondary market.

Table 2. The eBook Compared to Paper and the Advanced eBook Processor with respect to the eBook

Functions of Copy Accuracy	**Giovanni**	**Free Information**
Tradable good	Neutral	-
Tradable right	+	Neutral
Attribution and Integrity	+	Neutral
Persistence and Archiving	+	+
Access	Neutral	+
Personal Annotation	+	+

Arguably the Giovanni system should be far preferable to those seeking to maintain fair use and access of traditional information markets. Giovanni offers critical functionality that is aligned with the value of copyright to society: archiving, maintenance of author information, and no innate restrictions on access. The basic concept, to allow people to view rather than to assume each viewer is about engage in a criminal act, is embedded in copyright.

Giovanni offers critical functions of copyright. Giovanni maintains copies as an archive as well as ensuring that those who have purchased services can alter the format of the material to keep up to date.

If the watermarking allows for a history of use, just as this document embeds references and quotation from other works, then Giovanni offers a method for building complex hyper-linked trustworthy documents. This depends on the existence of homomorphisms in the watermarking system used, and the size of the sample necessary to extract attribute data. (Currently Giovanni requires 7 seconds of CD quality sound or 104 pixels to extract the watermarked data.) The inclusion of watermarks could be used to identify annotations in an individual copy.

Giovanni does not function to provide a tree-like mechanism so that the interrelationship between documents can be traced. However, with market dominance it could provide something not unlike citeseer (http://citeseer.nj.nec.com/).

The practice of embedding consumer information assumes that only the consumer would release the data and does not address the possibility of data theft. In the case of data theft the consumer is subject to the inherent punishment of exposure of sensitive information. Given that identity theft costs the victim on the order of thousands and recovery takes an average of eighteen months this would be a punishment unlikely to be sustainable as a matter of policy. While the feature is optional, the choice is up to the producer. In the case of embedding information the producer is allowed to select the level of risk for the customer without customer input. In addition this option precludes the ability read anonymously.

Does Blue Spike does not solve the problem of persistent storage or search? The question with respect to Giovanni is if a single company can serve the role of the Royal Library – creation of effective copyright protection, archiving, and access. Of all the systems, Giovanni is the nearest to serving the functions of the copyright system for digital content.

7. CONCLUSIONS

Copyright is a set of narrow legal rights built on a technologically deterministic foundation and embedded in a particular economic system. Extracting the functions of the technology, the law, and the economics enables a return to first principles for the design of DRM systems. That set of functions, which I refer to as copy accuracy to prevent confusion with copyright per se, is not entirely provided by any DRM system. Systems which fall under the rubric of copyright today by virtue of protecting content sometimes do enhance the market element of copy accuracy, however, sometimes at the price of the other functions of the copyright system. Similarly code that breaks content control mechanisms enhances some functions of copyright and undermines others.

Draft. Later works published as: L. Jean Camp, ”First Principles of Copyright for DRM Design,” *IEEE Internet Computing,* Vol.7, No. 3 pp. 59-65, May 2003. Previous version as ”DRM Doesn't Really Mean Copyright” in the peer-reviewed *Proceedings of the 2003 ACM Conference on Computer and Communications Security*, ACM Press (NY NY).

It is a common assertion that authors want to be remunerated with some combination of reputation and wealth. If information can be said to have desire, then information wants to be used and trusted. The increasing value of the readers' attention span should also be considered in the information economy. Neither those who seek to manage authors' rights nor the defenders of access rights are entirely aligned with the purposes of copyright in that no system motivates and awards filtering, rating and distribution.

No DRM or DRM circumvention system implements personal annotation. No DRM or DRM circumvention system rewards filtering. It is worth noting that some peer-to-peer systems that are decried as being an assault on copyright are designed such that filtering is intrinsic to the system (for example, Napster).

Future work entails examination of more digital rights management systems.

Of course, the most interesting question is designing a system that fulfills completely the functions identified here as copy accuracy. These extensions of the work are currently being addressed in ongoing research.

8. ACKNOWLEDGMENTS

My thanks to William Fischer, III for hearing this as a spark of an idea, and to Ron Rivest for hosting me in a presentation to his research group. Thanks to the participants of Financial Cryptography 2002 and the Workshop on Public Values in Design in NYU School of Law.

9. REFERENCES

[1] D. Burk and J. Cohen, 2001, Fair Use Infrastructure for Rights Management Systems, *Harvard Journal of Law & Technology*, Vol. 15, No. 1, pp. 41- 83.

[2] Piva, Bartolini, Barni, 2002, Managing Copyright in an Open Network, *IEEE Internet Computing*, May/June, PP. 18-26.

[3] J. Rowe, 2002, "Tollbooth of the Mind", *the Christian Science Monitor*, June 27m 2002, pp.9 sec 2.

[4] R. Bolick, 2001, Publishers' Requirements for Digital Rights Management, *W3C DRM 2001 Workshop*. INRIA, Sophia Antipolis, France.

[5] Branscomb, 1995, *Who Owns Information*, Basic Books, New York, NY.

[6] Lessig, 2001 *The Future of Ideas*, Basic Books, New York, NY.

[7] J. Litman, 1997, Reforming Information Law in Copyright's Image, 22 *University of Dayton Law Review* 587: http://www.msen.com/~litman/dayton.htm.

[8] N. W. Netanel, 1996, Copyright and a Democratic Civil Society, 106 *Yale Law Journal* 283.

[9] P. Samuelson, 1999, "Why the anti-circumvention regulations need revision", *Communications of the ACM*; New York; Sep 1999.

[10] L. Lessig, 1991, *Code and Other Laws of Cyberspace*, Basic Books, NY, NY.

[11] Cohen, J. 1996, "A Right to Right to Read Anonymously: A Closer Look at Copyright Management in Cyberspace", 28, *Connecticut. Law review,* 981

[12] W. Gordon, 1992, "Asymmetric Market Failure and Prisoner's Dilemma in Intellectual Property," 17 *University of Dayton Law Review* 853.

[13] W. Gordon, 1997, "On the Economics of Copyright, Restitution and 'Fair Use': Systemic Versus Case-by-Case Responses to Market Failure," 8 *Journal of Law and Information Science* (Australia) 7.

[14] Litman, 2001, *Digital Copyright*, Prometheus Books, New York, NY

[15] Vaidhyanathan, 2001, *Copyrights and Copywrongs*, New York University Press, New York, NY.

[16] W.W. Fischer, III (1999) "The Growth of Intellectual Property: A History of the Ownership of Ideas in the United States," in *Eigentum im internationalen Vergleich* (Vandenhoeck & Ruprecht, 1999),pp. 265-91

[17] E. Eisenstein, 1979, "*The Printing Press as an Agent of Change*", Cambridge University Press, Cambridge, UK.

[18] M. Castells, 1997, *The Information Age: Economy, Society*, Culture Blackwell Publishers, Massachusetts.

[19] M. McLuhan ,1997, *The Guttenberg Galaxy*, University of Toronto Press, Toronto, Canada, pp 23-31

[20] M. Rose, 1993, *Authors and Owners,* Harvard University Press, Cambridge, MA.

[21] L. Febvre & H. Martin, 2000, *The Coming of the Book*, Verso, London, UK

[22] Holdsworth, 1938, *A History of English Law*, Methuen & Co Ltd, London, UK.

[23] S. Haber and W. S. Stornetta, 1991, "How to time-stamp a digital document", Journal of Cryptology: the journal of the International Association for Cryptologic Research, Vol. 3, No. 2, pp. 99-111.

[24] R. Stallman (1984) The GNU Manifesto, http://www.fsf.org/gnu/manifesto.html. Included in C. DeBona, S. Ockman and M. Stone (eds) *Open codes: Voices from the Open code Revolution*, O'Reilly, 1999

Draft. Later works published as: L. Jean Camp, "First Principles of Copyright for DRM Design," *IEEE Internet Computing,* Vol.7, No. 3 pp. 59-65, May 2003. Previous version as "DRM Doesn't Really Mean Copyright" in the peer-reviewed *Proceedings of the 2003 ACM Conference on Computer and Communications Security*, ACM Press (NY NY).

[25] M. H. Harris, 1995, *History of the Libraries in the Western World*, The Scarecrow press, London, UK.

[26] Sterling, LLB, 1998 *World Copyright Law*, Sweet & Maxwell, London.

[27] R. Darnton, 1982, *The Literary Underground of the Old Regime*, Harvard University Press, Cambridge MA

[28] National Academy of Science, 2000, *The Digital Dilemma: Intellectual Property in the Information Age* National Academy Press.

[29] A. Odlyzko, 2000, The History of Communications and its Implications for the Internet, *University of Minnesota Working Paper Series*, University of Minnesota, Minneapolis, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=235284; Posted: August 30, 2000; Last Accessed: May 6, 2002.

[30] Lerner, 1998, *The Story of Libraries from the Invention of Writing to the Computer Age*, Continuum Books, New York, NY.

[31] T. Berners-Lee, 2002, "The Semantic Web", 2002-04-18, Center for eBusiness@MIT, Cambridge, MA.

[32] R. Blood, 2002, *The Weblog Handbook*, Perseus Publishing, Cambridge, MA.

[33] J. Rodzilla, 2002, *We've Got Blog*, Perseus Publishing, Cambridge, MA.

[34] R. Wade, 1987, *The Spirit of the Web*, Sommerville House Books, CA.

[35] J. R. Beniger, 1989, *The Control Revolution : Technological and Economic Origins of the Information Society*, Harvard University Press.

[36] A. Johns, *The Nature of the Book: Print and Knowledge in the Making*, University of Chicago Press, Chicago, IL.

[37] United Nations, 1995, *The United Nations and Human Rights 1945-1995: The United Nations Blue Book Series. Vol. VII*, United Nations; New York, New York.

[38] A. Osorio, "Primary Income Loss and Secondary Network Effects in Illegal Copying of Software" *Information Technology Group, Center for International Development, Working Paper Series*. Harvard University, Cambridge, MA.

[39] M Pahfl, (2001) "Giving Music Away to Make Money, First Monday, Vol. 6 No 8. www.firstmonday.org/issues/issue6_8/pfahl/index.html

[40] A. Shapiro and H. Varian, 1999, *Information Rules* , Harvard Business School Press, Boston, MA, c1999.

[41] A. B. Whinston, Dale O. Stahl, Soon-Yong Choi,, 1997, *The Economics of Electronic Commerce*, Macmillan Technical Pub.. Indianapolis, IN.

Draft. Later works published as: L. Jean Camp, "First Principles of Copyright for DRM Design," *IEEE Internet Computing,* Vol.7, No. 3 pp. 59-65, May 2003. Previous version as "DRM Doesn't Really Mean Copyright" in the peer-reviewed *Proceedings of the 2003 ACM Conference on Computer and Communications Security*, ACM Press (NY NY).