

The Impact of Electronic Publishing on the Academic Community

Session 2: Legal and political issues

Managing copyright in a digital environment

Jon Bing

Norwegian Research Center for Computers and Law, Faculty of Law, University of Oslo,
PO Box 6702 St Olavs plass, N-0130 Oslo, Norway

©Portland Press Ltd., 1997.

[Copyright Information](#)

Expectations of electronic publishing

Working with information systems for providing authentic text (full text) since 1970, it is to some extent regrettable to admit to having failed to achieve what we set out to do. In the early 1970s online systems were introduced, giving access to central databases through terminals. This was at once perceived as a way for making material available to end users, perhaps emphasizing the academic communities and some professional groups of end users like health personnel and lawyers. Great advances were made in transforming the cryptic user interfaces of the systems of that time - typically involving the use of codes and strings of special characters unique to the single system - towards more user-friendly standards.

At the end of the 1970s, many of us felt we were on the brink of a new era of database publishing. The second generation of text retrieval systems were quite powerful, had interfaces based on menus with help functions, and within some domains, databases of considerable coverage had been established. Our expectations, however, failed in general to be realised.

An exception was, perhaps, bibliographical databases. Systems like Medline for medical literature or the more general systems like Orbit, had a considerable measure of success. They were not providing material in authentic text, but contained references to paper-based documents. They were mainly used by the library community, whose members were trained in using information systems and classification schemes, and who used the online systems for search only. If an interesting document was identified, it had to be acquired through conventional systems for inter-library lending, photocopying, etc.

In two other areas there were established significant online services. One major example was legal databases. Nearly all Western countries had at least one major text retrieval system for statutes, regulations or case law. The major U.S. systems - Lexis¹ and Westlaw² - had a major impact on the market. But even in these instances, they had difficulties with penetrating the legal community to the expected depth. Schemes were launched in the U.S.A. to provide these services to legal academics and students at favourable rates in order to train future customers to benefit from the powerful research facilities. But in most countries, the services were offered at a rate too high for academic institutions to take advantage of them.

The other major example was databases of editorial material from newspapers, news magazines, etc. Again, a major service was the twin of Lexis, known as Nexis. These were to a large extent used for background research by media, but did not penetrate to the general public or the academic community.

There are a number of lessons to be learned by these early developments. The major lesson is one in availability - although the systems were far more user friendly than the first generation of text retrieval systems, they still were rather cumbersome. This was demonstrated with a vengeance when systems for office automation based on microcomputers were introduced in the early 1980s. Compared to the user interface of a Macintosh and, later a Microsoft Windows application, the user interfaces of the second generation text retrieval systems were less than appealing. This in spite of the fact that the databases available through online systems were huge, the search strategies powerful and the speed of retrieval rather impressive.

During the 1980s, we failed to realise the perceived potential of database publishing, also within the scholarly sector. But unknown to most, in 1980 a development took place that was to change the future of computerized publishing. Tim Berners-Lee wrote a notebook program at CERN (European Organization for Nuclear Research) in June--December 1980, a program he named "Enquire-Within-Upon-Everything" (ENQUIRE), written for a Norsk Data running under the operation system SINTRAN III. This program allowed links to be made between arbitrary nodes, each node having a title, a type, and a list of bidirectional typed links.

Based on ENQUIRE, Berners-Lee wrote a proposal for information management in 1989 which was circulated internally at CERN, and the project was re-formulated in 1990 with Robert Cailliau as co-author, and was launched under the now well-known name World Wide Web (WWW). A browser was released in 1991, and WWW became first available on the Internet in August 1991. At the end of 1992, the number of servers had increased to 26, next year the alpha version of Mosaic was released by Marc Andreessen, and more than 200 servers were operational. Mosaic graduated to Netscape in 1994, and the servers had increased to 1\500. In 1995 the Web was the main theme for the G7 meeting hosted by the European Commission in Brussels.

This compact history is presented to bring home one point only: the WWW with associated browsers is very young. But the technology did for computerized access to material what IBM's 'personal computer' did for office systems. The technology has become immensely popular and widely available in a short time, it has actually graduated to become a mass medium with the public at large as a possible market, completely saturating academic institutions and other

educational institutions. The reason is, obviously, the user interface that is similar to the icon interfaces of office automation. It should be noted that retrieval facilities are rather modest, even the more powerful search engines have not the sophistication of the search language of the more conventional online systems, and performance³ is mediocre⁴.

Therefore, the current state is only sufficient to indicate the potential of the technology. Again we feel our expectations grow as the number of information services offered through the Internet grows and as they become more exciting. We observe, however, that there is a reluctance to offer certain types of material on the Internet, typically the material that is traded by traditional publishing houses. The reason is not the user interface, but rather the possibilities for managing the rights to the material.

Conventional database publishing managed rights in a rather cumbersome way: a contract was negotiated between the service provider and the end user, typically through an exchange of paper documents. The end user was assigned an account number and a password for identification and authentication. The service provider measured end-user exploitation of the service, typically using log-on time as a measure, and would periodically bill the end user with an invoice, which was paid by the end user using traditional payment systems⁵.

Within the Internet, two types of payment systems have been introduced. The first is an indirect way of payment: the information service provided may be sufficiently popular to attract advertisers. They pin their ads to the same pages that contain the information, and because these ads will be read by a large number of users, they are willing to pay the information provider for the privilege, and in this way indirectly finance the service. Newspapers - either based on the editorial material in paper editions, or designed directly for the web - are service providers typically relying on this indirect financing.

The second is a subscription scheme: the end user will subscribe on the basis of months or years to a service, and is given means for identification and authentication. This is a version of the conventional online database contract, but somewhat simplified as the service provider does not have the same means within the Internet technology to base the calculation of the fee on the user of the service: for instance, there is typically no notion of a 'session' on the Internet. Therefore the fee typically is a flat fee, and it is paid in conventional ways, again typically using existing credit card facilities. One service adopting this strategy is the Encyclopaedia Britannia.

We see also that this latter form of payment is somewhat cumbersome, and that it is not well designed to purchase small items of information, as the transaction costs for the payment itself are rather high. To facilitate this, one would have a more efficient payment mechanism, and there are emerging services that have such possibilities, like IBM's InfoMarket.

These emerging services are the seed to more well-developed electronic copyright management systems. There are several interesting services designed for the academic community, like the "electronic readings in management studies" (ERIMS) based on the Electronic Libraries Programme (eLib) in co-operation with Aston University, the University of Newcastle and the University of Sheffield. Another, older, initiative is Electronic Library Information Online

Retrieval (ELINOR) at De Montford University in co-operation with the British Library and IBM UK.

It is suggested that development at the moment is held back by a lack of appropriate copyright management systems designed for the environment of the Web. It is also suggested that there is a need for consensus within the industry on standards for such a system before it really will accelerate. Such a consensus need not necessarily be universal, possibly several systems could co-exist. But it needs to be sufficiently widely accepted to be more than a proprietary or closed system. As will be discussed below, we see tendencies towards such a development. We may be impatient with the speed of development, but our impatience may be somewhat soothed when we realise how young this technology really is.

Public sector as a provider of raw material for private sector

One of the slogans emerging for the networked environment is that "content is the king". By this is indicated the rather obvious, that to attract end users to a service, the service offered has to be useful, entertaining, educational or provide news coverage. It is not the technology as such that will create a market, but the content of the service offered to the user. A new terminology describing authors as 'content providers', and the sector of publishers and phonogram or videogram producers as 'content industry' has emerged, perhaps indicative of a change in basic values and policies more akin to the software industry than the traditional field, for instance, of scholarly publishing, where ideals relating to quality, education, the advancement of science and other issues of cultural policy have been central. Some of us are reluctant to adopt these terms as they seem to imply a change in traditional policy objectives.

As proposed earlier, the lack of a flexible, efficient and powerful solution for the management of rights currently holds back development. The 'content providers' are reluctant to launch services when no appropriate control of the material can be secured, therefore, more conventional solutions still prevail.

In the European Union, one has even before the advent of the Web looked for ways of stimulating the growth of the 'information market', which is also perceived as a key to development in other fields of European economy. It has therefore been tempting to look for 'content' not subject to the problems of the management of rights. In Europe generally, public material is not subject to copyright⁶. As mentioned briefly earlier, one of the more successful sectors for conventional online databases has been the legal systems, and a large fraction of the online services in Europe is based on such material⁷.

Pursuing this policy, the European Communities issued their 'synergy guidelines' which suggested for member countries that the material held in the public sector should be seen as an information resource to be developed, refined and offered to the market as value-added services by the private sector. At the same time, there was a tendency in many jurisdictions for public agencies themselves to develop information services based on the material held by such agencies, and offering such services commercially to the private sector. The 'synergy guidelines' have hardly had the effect originally intended. And the issue of payment for information services supplied by the public sector has become controversial in many countries. It has, for instance,

been suggested that it is hardly fair that a public agency, which has a right based on statutes to require information from the private sector, should be allowed to re-package information and sell it back to the private sector.

Without pursuing this issue in greater detail, it may be maintained that the balance between public and private sector still has to be struck. Although the material may not be protected by copyright, the material is still controlled by the public sector. In Europe, only a few jurisdictions have efficient freedom-of-information legislation that make it possible for private sector to require access to information to which value then may be added for a commercial service.

When developing the database directive, the original draft contained a provision that extended a compulsory licence to a private party in the case where a public agency had entered into an arrangement with one private party. In such a case, the another private party might require to have access to the same material, and exploit this commercially. This obviously was designed mainly to ensure competition in the market place, and to restrict the development of exclusive arrangements. The provision is deleted in the final directive. But the Magill decision may nevertheless imply that if a dominant market position is abused for an exclusive arrangement, a compulsory licence may be extended by the Commission. The Magill decision addressed a situation where copyright was the foundation of the dominant position, but one may speculate whether the argument would not also apply if a public agency contributed towards establishing a dominant position through exclusive arrangements with parties in the private sector.

As the 'synergy guidelines' were found to have less than the desired effect, the Commission has been preparing a green paper on Access to Governmental Information, which is still to be adopted. It is still to be seen what possible effect this paper may have on the issues discussed here.

Traditional rights management: control of physical objects

Something of the nature of conventional systems for the management of rights has already been implied, see above. We are well used to the notion that copyrighted works are of an intangible nature, and we are trained to make the distinction between the work as such, and a copy of that work, the former being an intangible object, while the latter is a physical object. In many practical cases this is also reflected in the contracts applying to the work, and consequently in the management of rights. For instance, when a novel is to be translated from one language, A, to another language, B, a contract will be signed transferring publishing rights for the new language. The resulting text of the novel in language B is recognised without effort as a representation of the original work, though fused with the creative effort of the translator.

There is, however, somewhat of a paradox that although the intangible nature of a copyrighted work is basic to our understanding of the law, and often reflected in the contractual relationships, the strategy for the management of rights is based on controlling the physical objects. In publishing, the control is focused on the printing of an edition, and the distribution and sale of the printed books through retailers, this commonly creating the basis for the remuneration of the author in the form of royalties. For films, conventionally control is exercised in the relatively few copies of the film available for distribution and rental to cinemas. The management of rights with

respect to videograms is similar to that of books, as is the case for phonograms. Also, the regimes for photocopying are oriented toward the physical objects, the paper copies, or the means for making such copies, the photocopying equipment.

There are, as already indicated, exceptions. The translation contract (or any other contract for adaptation) is based on different principles, and the broadcasting of works has led to the development of a rather specialized form of management of rights, where the collecting societies have found a special role for re-diffusion through cable networks.

It is, nevertheless, maintained that it is a characterization of the retail-oriented systems for rights management that they are based on control with physical objects, a characterization that may be interpreted as something of a paradox as it relates to intellectual property of intangible works.

Characteristic of the digital environment is the disappearance of physical copies: the case of "the invisible copies", as Michael Keplinger once put it. A machine-readable representation is a copy, and one may have a trade in such copies, typically illustrated by the trade in CD-ROMs, where the conventional systems for rights management may be successfully applied (and which may be part of the reason why these have grown popular, perhaps more popular than the technology merits).

But for a trade in nets, copies are not traded. Rather, a communication stream is initiated, signs being communicated through the net from the provider of the service to the end user. At the end user's, a new machine-readable copy may be downloaded, or the signs may be used to create displays on a screen, or a local paper copy may be printed out, or all of these. The trade in physical objects is replaced by a trade in signs, liberated from the media on which the signs more or less incidentally are stored, more or less temporarily.

In analysing such transactions, it is suggested that it is not sufficient to liken this to trade in physical objects, there is a qualitative difference. What the end user acquires is not so much a copy of the work, as a special competitive position with respect to the work. The position may be utilized for personal benefit, but equally easily for re-distribution through an intranet, as a basis for an adaptation, the production of an edition of paper-based copies, etc. It is suggested that if the transactions in this way are seen as a trade in competitive positions with respect to the work, rather as a trade in copies of the work, it is easier to understand the characteristics of the new market place for an electronic commerce in intellectual property.

Electronic copyright management systems

The policy of the European Communities seems to take as a basis the proposition that an adequate regime for the management of rights is necessary to unlock the potential of the market place of the networks. Several efforts have therefore been undertaken to establish a basis for the development of electronic copyright management systems (ECMSs)⁸. An important effort is CITED (copyright in transmitted electronic documents), a project producing a generic model of how to achieve a solution for digital environments, under which two demonstrators were implemented: ADONIS for a collection of journals in machine-readable form, and NARCISSE (network for art research computer systems in Europe) for fine art from several museums with

accompanying explanatory texts. Under the fourth framework program, CITED is a reference for further efforts and concrete prototypes, like the COPYCAT (copyright ownership protection in computer-assisted training) being tested in a 'hostile' university environment to have empirical evidence of its security.

One of the current projects attracting most attention is IMPRIMATUR (intellectual multimedia property rights model and terminology for universal reference), a two-part project where one part aims at implementing an ECMS according to the CITED concept, while the other part aims at arriving at a consensus among the relevant parties. As indicated above, such a consensus is necessary for a general market place to emerge, and it is not really a technical issue, but more an issue of balancing interests, and agreeing upon standards for interconnection and other elements, see later. IMPRIMATUR organises a number of 'Consensus Fora' for discussion between the relevant parties, not only within the European Communities, but also with representatives from the United States. The first Consensus Forum was organised in London in November 1996, the next was organised in Stockholm in May 1997.

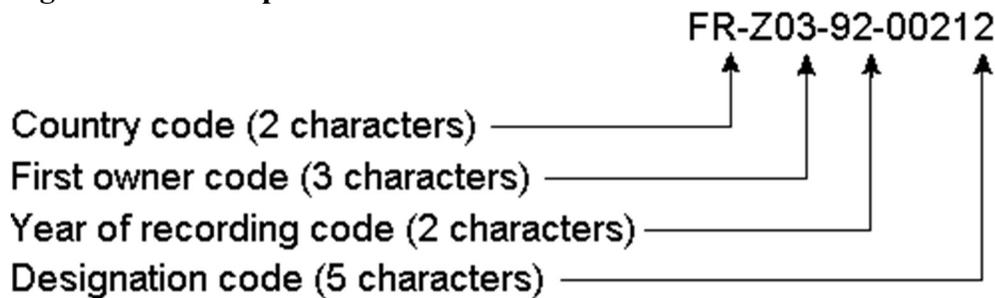
The elements necessary for the development of a viable electronic copyright management system may be listed briefly.

Identification

There is a need to identify the works subject to trade. Identifiers are well known in the conventional environment as the International Standard Book Number printed on the cover of each book, and made machine-readable by a bar code. However, this is not so much an identification of the literary work as the book as a type of goods (the bar code is part of the European Article Number scheme). The printing of the code openly on the book is an example of 'tattooing' in the jargon of the field.

A better example is the International Standard Recording Code (ISRC), which was developed by the International Organisation for Standardisation (ISO), the worldwide federation of standardization organizations, as an identifier for phonograms and audio-visual recordings, known as ISO 3901. The International Federation of the Phonographic Industries (IFPI) recommended as early as 1988 that ISRC should be adopted to identify the recordings on short-form music videos. ISO appointed IFPI as the International Registration Authority for ISRC in 1989, and IFPI is now working on a recommendation of introducing an ISRC into the sub-code of all digital sound recordings. An example of an ISRC is given in [Figure 1](#).

Figure 1. An example of an ISRC



There are currently efforts to develop such codes for many categories of works. They are often referred to as "copyright management information", and the WIPO (World Intellectual Property Organization) Copyright Treaty Art 12 includes provisions making it a criminal act to remove or tamper with this information.

The code will typically be a reference to a database that contains the details of the work, for instance the owner, or succession of owners, the licensing divided on geographical areas, media, etc. Using the identifier as key, one may access the detailed information and utilize this in the management system.

In addition to prohibiting the tampering with identifiers, one has also developed methods of 'watermarking' the code onto the work. The work in digital form is represented by a stream of binary groups, each group defining for musical works the frequency of sound, for pictorial works the qualities of the pixel in grey scale and colour code, etc. Such codes can be slightly altered without this having an effect on the quality of the reproduction of the work, but making it possible to derive the code using an algorithm for retrieving the altered groups from the work itself. Successful schemes have been launched for sound recordings, still and moving pictures. For sound recordings, one of the systems marketed embeds the code for every 4.7 seconds throughout the work: removing the watermark would obviously also render the work itself without value.

A related method is that of 'fingerprinting', which is a scheme for retaining information on the use of the work. An example is given in United States Digital Audio Recording Devices and Media Act, which requires devices for recording or playing digital sound recordings to include a device which keep track of (i) how many copies are made from an original, writing this information on to the original itself, and (ii) that a copy is made on the basis of an original. In this way, a serial copy management system is achieved, making it impossible to make more than three copies of each original. The WIPO Copyright Treaty Art 11 also included provisions prohibiting the sale of devices designed to circumvent such technical protection systems.

The result is that there is emerging a scheme of identifying each work by a code embedded in the work itself, a code which may be retrieved by a copyright management system, and used as a key for accessing databases containing detailed information on the rights associated with this work, for instance for distributing revenues.

Cryptography

To protect the work, it has for a long time been suggested that the work should be subject to encryption. This would make it possible only for somebody with access to the key to benefit from the work. The use of public key encryption systems, which are asymmetric encryption systems, has also for a long time been seen as a probable solution.

Very simply put, such a system works with two keys; a private key A, and a public key B, the latter being readily available from public databases. A work encrypted by A can only be decrypted by B, and vice versa. This would allow secure transactions, for instance if person B purchases a work from person A, person A would first encrypt the work with his or her own private key A, and then with the public key of B. On receipt, B would decrypt the work with his or her private key, and the public key of A. If the result is an understandable work, it proves that the work originated by A and was meant for B, and B only.

The scheme may seem rather complex, but as the encryption is enabled by end-user equipment and automatic retrieval of the public keys, the user will not really be bothered by what is going on. But this is only a crude sketch, in reality there will be further routines to make the communication as efficient as possible.

The technology allows many legal issues to be solved, obviously identification and authentication of the parties are secured. Regrettably, there have been legal obstacles for arriving at a consensus for using cryptography, mainly related to the reluctance of the law-enforcing community in allowing encrypted communication that cannot be intercepted by law enforcement officers. The United States has promoted a scheme of key escrow, which make the keys available for law enforcement, and France has strict limitations for the use of encryption. However, the recent OECD (Organisation for Economic Co-operation and Development) Cryptography Policy Guidelines⁹ may be an important step towards consensus. These guidelines reject key escrow and endorse voluntary, market-driven development of cryptography products.

Payment

It is important to find a scheme that allows on-the-spot purchases of items of rather low value. This implies that the transaction costs of payments have to be low, much lower than for the methods currently being the more common solutions, see earlier.

There are initiatives which at the moment seem promising, among them the British-based Mondex system that is based on a smart card that is 'loaded' by credit from the account of the end user. Again, it is too early to tell which of the initiatives will gather a sufficiently large following to become a consensus standard.

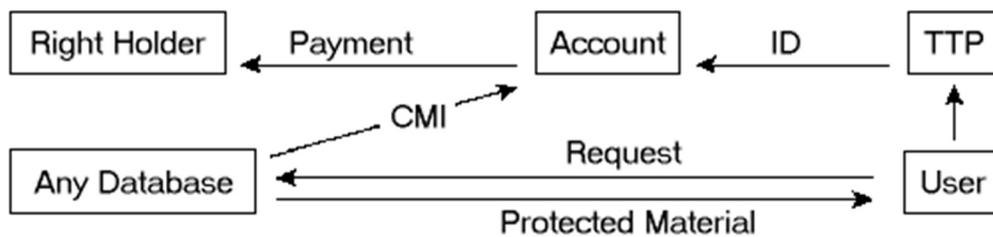
User identification

There must also be some way of ensuring the identity of the user. Today, one generally uses four-digit personal identification codes, but this system does not have a sufficient level of security, and also has severe drawbacks in user-friendliness. There are experiments with a smart

card with integrated devices for recognizing some biometrically unique feature of the user, for instance an optical finger-print scanner.

One may for instance imagine a PCMCIA card with these features, also containing the private encryption code of the end user, issued by a trusted third party. The card could be inserted in any device with an appropriate station for reading the card, making it possible for the end user to trade at any work station to which he or she would have access. As a summary, a rough model of the relations of the different elements in the system is indicated in [Figure 2](#).

Figure 2. A rough model showing the different elements of an electronic management system



The transaction is initiated by the user who identifies a work he or she wants to purchase. The material is not necessarily residing on a database under the control of the right holder, as it is in encrypted form and therefore protected in this way.

The system would then identify and authenticate the end user, checking that the necessary credit is available at the account of the end user. It would retrieve the copyright management information from the work, perhaps accessing a database to determine how the payment should be distributed. The work would then be transferred to the user and decrypted by his or her device, at the same time the payment would be deducted from his or her account and credited the appropriate accounts of the right holders.

This is but a crude sketch. There are several unsolved issues. One is that this sketch presumes that one purchases 'one copy'; this is, as argued above, a fallacy. One purchases some limited licence to the work. This can be implied or made explicit, but there should be possibilities to purchase more than one type of licence. For instance, a medical doctor may purchase an article from a journal for personal use, while a hospital may purchase the same article for internal dissemination within its intranet. The system should be able to cope with this, and would therefore have to be complemented by mechanisms for specifying licensing terms. This could be some sort of system for electronic data interchange (EDI) related to the Incoterms known from the trade in goods.

Another is the problem of disloyal or careless end users who do not respect the limitations of the licence. In this way, the control of the copyright management system may fail. There is,

however, in the embedded code of the work the potential for also controlling - even policing - secondary use.

But this latter aspect also indicates another legal policy issue; the system enables a detailed recording of consumer's use of works. With respect to data protection, one may find it less acceptable to keep a detailed track of purchases of music, videograms and other material from the home.

And finally, this is a system sketched for managing copyrighted material. But there is the possibility also to encrypt and protect material that is not protected. Some feel that the system offers the possibility for too much control, making the market pay also for material that today is freely available.

Therefore, there is a long way to go before we will see generally accepted and well-functioning copyright management systems for the digital environment.

Renewed expectations: unlocking the potential

There may be a long way ahead before a satisfactory copyright management system for the networked society has been developed. The possibility that one will develop such a system seems, however, rather high. The main reason is the consumer market, the steeply increasing number of people having access to Internet, the improved telecommunication capacity provided by ISDN (Integrated Services Digital Network) and, perhaps more importantly, a new earth-bound infrastructure based on broad-band fibre optics, as being introduced now in countries like the UK and Sweden.

The engine in the development will, I believe, be the consumer market. But with respect to public policy, the access gained by the academic community to the wealth of information resources will be as important. And this wealth of information is to a large extent provided by libraries. These two communities will be strongly influenced by the developments, and strategies to influence these developments by these communities are and should be sought.

The conventional market for publishing has many actors: authors, publishers, printing houses, distribution systems, libraries, mail-order companies (book clubs), retailers and readers. In this complex chain, only the authors and the readers will remain constant. The rest of this complex and interlocking organization making up the market will be heavily influenced by the on-going changes. One may only glance at what is going on in the market place with respect to positioning, mergers and failures to see the some of the first effects of these changes: they will be even stronger in the next few years. As the Chinese say, "we live in interesting times". It is not necessarily a word of comfort.

Discussion following presentation by Bing

Dallas discussed the question of rights in historical and artistic items and their images (for example, photographs). Bing explained that in such cases the copyright lies in the image not in

the original artefact. In principle, anyone could create his or her own image of a historical or artistic work.

¹ Lexis was at this time provided by Mead Data Central, it has been more recently acquired by Reed Elsevier.

² Westlaw was and is provided by West Publishing Co.; this company has been more recently acquired by the Thompson group.

³ For instance, measured in recall and precision.

⁴ Although no controlled experiments measuring the performance are available.

⁵ These are to a large extent based on computerized solutions, but using networks other than the one through which the end user accesses the information service, typically the closed networks of banks.

⁶ See Berne Convention Art 2(4). This is a matter of national legislation: the possibility is exploited rather differently in different jurisdictions, the UK still maintains the doctrine of Crown Copyright in statutes and other governmental material.

⁷ Around 1990 it was indicated that approximately 20% of online services were based on legal material.

⁸ Some prefer the abbreviation IPMS for 'intellectual property management systems'. This is probably more correct, as intellectual property rights other than copyright will have to be managed by the system. ECMS has, however, already become the standard term.

⁹ Adopted on 27 March 1997.

©Portland Press Ltd., 1997.

[Copyright Information](#)