Metadata and its relationship to copyright

Based on a presentation by Professor Dr. Christian Donle

Attorney at law of Preu Bohlig & Partner, Berlin, Germany.

1. Definition of Metadata

Metadata is structured data which advances the description and findability of sources of information.

Metadata provides basic information about an object, such as details of the author, title, date of release, information concerning the ownership and the rights for use.

2. Examples of Metadata Schemes

Dublin-Core-Metadata-Element-Set

A collection of simple and standardised conventions used to describe documents and other objects on the internet, to simplify searches with the help of metadata.

Exchangeable image file format (Exif)

Exif is a specification for the image file format used by digital cameras.

The metadata tags defined in the Exif standard cover a broad spectrum:

- Date and time information. Digital cameras record the current date and time and save this in the metadata
- Camera settings.
- Descriptions and copyright information.

IPTC

Metadata specified by the news and photo image industry to describe and manage media content which is widely in use for photos.

3. Metadata Types

There are three main types of metadata:

Descriptive metadata
- Describes a resource for purposes such as discovery and identification.

Structural metadata
- Indicates how compound objects are put together

Administrative metadata
- There are two types defined in a NISO document about metadata; rights management metadata and preservation metadata. It provides information to help manage a resource, such as when and how it was created, file type and other technical information, and who can access it.

(The NISO source document for metadata is at: http://www.niso.org/publications/press/UnderstandingMetadata.pdf)

4. Metadata and Digital-Rights Management

The media industry distinguishes between Rights Expressions and DRM.

Rights Expressions express rights in a human readable and machine readable way.

DRM (Digital Rights Management) are closed systems which impact directly on the usage of digital media objects: access to the object is denied if the user cannot prove the right to use it.

In the digital world, an unauthorised copy is indistinguishable from the original. For the cultural industry, DRM is becoming more crucial.

DRM systems allow technological protection measures beyond the ‘classic’ measures offered in copyright law.

Examples of DRM technological protection measures are:

- metadata
- digital watermarks
- copy control procedures

The advantages of embedding metadata in images

- Decrease in the number of Orphan Works
- Improved information retrieval
- Advancement of search quality for the author

5.Infringement may occur when …

a) Metadata is changed

If copyright data, creator or credit details are changed, this can amount to infringement.

b) Metadata is removed

XMP, EXIF and IPTC data is added to the 'Header 'of the image file. There can be interoperability problems when metadata is added by one software and read by another. Using image editing software can lead to deletion or conversion of metadata, both knowingly and unknowingly.

6. European Commission Directives relating to Metadata

i) DIRECTIVE 2001/29/EC

on the harmonisation of certain aspects of copyright and related rights in the information society

-Standardisation is a political goal:

(54) Important progress has been made in the international standardisation of technical systems of identification of works and protected subject matter in digital format. In an increasingly networked environment, differences between technological measures could lead to an incompatibility of systems within the Community. Compatibility and interoperability of the different systems should be encouraged. It would be highly desirable to encourage the development of global systems.

-Widespread distribution and use is a political goal:

(55) Technological development will facilitate the distribution of works, notably on networks, and this will entail the need for rights holders to identify better the work or other subject-matter, the author or any other rights holder, and to provide information about the terms and conditions of use of the work or other subject-matter in order to render easier the management of rights attached to them.

Enactment of Directive 2001/29/EC

- Digital information is an essential element in copyright.
- The spread of e-commerce demands a more specific definition of the terms of use of intellectual property rights

In Europe:

The Copyright Directive (Directive 2001/29/EC) was enacted to implement the WIPO Copyright Treaty.

For the first time, this Directive implemented the legal protection of DRM- Systems and "technological measures“.

In Germany:

The Copyright Directive has been incorporated into the German Copyright Act.
§§ 95a – 95d (UrhG) in order to achieve an effective protection against the circumvention of technological measures.
§ 95 Copyright Act protects against the conversion and removal of metadata.

ii) Legal Protection Directive 2001/29/EC

(56) There is, however, the danger that illegal activities might be carried out in order to remove or alter the electronic copyright-management information attached to it, or otherwise to distribute, import for distribution, broadcast, communicate to the public or make available to the public works or other protected subject-matter from which such information has been removed without authority. In order to avoid fragmented legal approaches that could potentially hinder the functioning of the internal market, there is a need to provide for harmonised legal protection against any of these activities.

iii) Legal Protection Directive 2001/29/EC

Article 7 Obligations concerning rights-management information

1. Adequate legal protection against any person knowingly performing without authority any of the following acts:
   - the removal or alteration of any electronic rights-management information;
   - the distribution, importation for distribution, broadcasting, communication or making available to the public of works or other subject-matter protected … from which electronic rights-management information has been removed or altered without authority,

if such person knows, or has reasonable grounds to know, that by so doing he is inducing, enabling, facilitating or concealing an infringement of any copyright or any rights related to copyright as provided by law, or of the sui generis right ….

2. For the purposes of this Directive, the expression "rights-management information" means any information provided by rights holders which identifies the work … or covered by the sui generis right …, the author or any other rights holder, or information about the terms and conditions of use of the work or other subject-matter, and any numbers or codes that represent such information.

iv) Legal Protection II

Germany: § 95 c German Copyright Act - Legal protection for essential information

Information defining the terms under which a work can be used as defined by the rights holder may not be removed or altered, if essential information is embedded in a the copy of a work or the information is shown together with the work when the work is transmitted. If the information is removed or altered without authorisation the infringer knows or expects that this will cause the infringement of copyright and related rights or will conceal the fact that copyright has been infringed.

Rights information is electronic information that identifies the author of a work or other rights holders. It also contains data and codes which describe the terms under which a work can be used.

If the rights information of a work has been removed or altered without authorisation it may not to be published, traded, sent, shared or made available to the public, if the person infringing knows or expects that this will cause the infringement of copyright and related rights or will conceal the fact that copyright has been infringed.

v) Legal Protection Directive 2001/29/EC

- Removal of Metadata is prohibited in the EU
- Removal is subject to injunctions
- Removal is subject to criminal sanctions
- Works with removed Metadata may not be published and distributed
- Persons or entities delivering tools to remove Metadata are subject to civil actions (injunctions, damages) under Copyright Laws and criminal sanctions.

vi) Legal Protection Directive 2001/29/EC

Any circumvention of the protection of Metadata is prohibited in the EU

E.g.:
Services that remove Metadata
Distribution of Hardware or Software to remove Metadata
Advertising of such devices
Providing of storage systems to distribute such data (P2P)
Links to such data systems

7. Obstacles to metadata protection

- The widespread scanning and publishing of printed works by third parties (e.g. Google books) that produce digital data without appropriate metadata can create problems.

- Software that deletes metadata is still available.

- There have been very few cases in Europe where the removal of Metadata has been subject to court proceedings.
    - Associations of rightholders or agencies should enforce Metadata in online or other digital media.
    - Prosecution is permitted in many EU-countries only through class actions – not by the individual author.

- Authors and agencies need to improve systems so that images are not distributed without the appropriate metadata.
    - There are voices in the market that argue that deleting metadata has become common practice
    - The human or moral right to be recognised as the author is endangered by the deletion of metadata

- Online images should *always* have metadata embedded in them. This is not always the case at the moment.

- Authors and agencies should modify their standard terms and conditions to include:
    - “Digital publication only with metadata.”

Legal Protection of Metadata and Databases

Nancy E. Wolff of Cowan, DeBaets Abrahams & Sheppard, LLP

http://www.cdas.com

This short article outlines the legal protection, or lack of legal protection, afforded to metadata and electronic databases in the United States. It addresses U.S. copyright law, the Digital Millennium Copyright Act, recent failed database protection legislation, and the doctrine of "hot news" misappropriation. The article concludes with a broad outline of the state of database protection in the European Union.

Is metadata protected under US law?

The first question is whether metadata and/or databases are protected under U.S. copyright law? In answering this question, it is important to understand that U.S. copyright law, as set forth in the U.S. Copyright Act of 1976, protects only "original works of authorship." In *Feist Publications, Inc.* v. *Rural Telephone Co., Inc.*, the Supreme Court determined that an "original" work is one that is "independently created by the author (as opposed to copied from other works)" and which "possesses at least some minimal degree of creativity." This means that copyright protection never extends to raw facts, including facts which comprise the information contained in a photograph's or other work's metadata (e.g. the author's name, title of the work, time/date of creation, image resolution, etc.).

Nonetheless, copyright protection may be available for the selection, coordination, or arrangement of certain factual compilation (though protection never extends to the facts themselves), but only to the extent that the compilation's selection, coordination, or arrangement is sufficiently creative – i.e. "original."

Copyright and creativity - *Feist*

The creativity needed for copyright protection to attach is relatively low; however, as set forth in *Feist*, there are certain "categor[ies] of works in which the creative spark is utterly lacking or so trivial as to be virtually nonexistent…[S]election and arrangement of facts cannot be so mechanical or routine as to require no creativity whatsoever."

Therefore, in *Feist* the court found that a telephone directory which merely arranged entries alphabetically by last name and provided basic and obvious factual information such as a person's "name, town and phone number" was not sufficiently "original" to merit copyright protection because it lacked creativity. *Feist* also rejected the idea that copyright protection should be extended to

reward the labor, skill and expense involved in compiling the telephone directory at issue, [a principal which is also enshrined in the TRIPS and WIPO copyright treaties.] As a result, *Feist* sounded a conclusive and final blow to what is commonly called the "sweat of the brow" theory of copyright protection.

In addition to telephone directories, courts have found blank forms, including collections of legal forms, arrangements of real estate data, and databases of factual information to be insufficiently original and not protected by copyright law. Further, it is likely that most compilations of metadata, as stored or managed in a database or registry, are particularly unlikely to be protected by copyright because such databases derive their value from efficient organization and comprehensiveness.

Consequently, the coordination, selection, and arrangement of such database information will likely be so obvious, basic and lacking in creativity that it cannot be deemed "original."

As stated in *Feist*, "common sense tells us that 100 uncopyrightable facts do not magically change their status when gathered together in one place."

Nonetheless, theoretically at least, copyright law provides a thin layer of protection for electronic databases that are sufficiently creative, though it is difficult to imagine exactly what such databases would have to look like.

Other attempts to expand protection

There have been two recent congressional attempts to expand the legal protection afforded to databases and other factual compilations beyond the protection afforded by the Copyright Act: the Consumer Access to Information Act of 2004 (CAIA) and the Database and Collections of Information Misappropriation Act of 2003 (DCIMA).

However, in both cases the proposed legislation failed to make it out of the House of Representatives. CAIA would have classified database misappropriation as unfair competition and would have provided the Federal Trade Commission with authority to enforce its provisions.

Similarly, the DCIMA would have imposed civil liability for making available a "substantial" part of the information contained in a database created, maintained or generated by another person without authorization.

Both pieces of legislation would have provided greater legal protection for compilations of metadata if they had passed into law.

The DMCA

The Digital Millennium Copyright Act (DMCA) is one piece of legislation which actually does provide an avenue by which an author may seek to protect the removal of metadata from their copyrighted work, though it does not protect the metadata itself.

Specifically, the DMCA, which was signed into law by President Clinton in 1998, prohibits the unauthorized, intentional removal or alteration of "copyright management information" (CMI) for the purpose of inducing, enabling or facilitating copyright infringement.

CMI includes the following: the title or identifying information of a work, the name of the author or copyright owner, the terms and conditions of use of a work, and identifying numbers or symbols, including hyperlinks, referring to such information, and any information required by the Register of Copyrights.

To be deemed CMI under the DMCA, some courts (notably the courts of New Jersey) require that the information be affixed to the author's work through the use of digital rights management technology. Most courts, however, simply require that the information be affixed to the author's work by digital means.

In any event, CMI must be provided "in connection with" the work. Broadly speaking, this means embedded in a work's metadata or in close proximity to the copyrighted work.

Additionally, although damages can be substantial under the DMCA ($2,500-$25,000) there have been few successful cases involving the removal of CMI because the DMCA requires a plaintiff to establish that the defendant removed CMI *with an intent to infringe*.

This can be difficult to establish given that many versions of popular design software, notably Photoshop CS3 or earlier, automatically strip metadata from images by default when the software's "print to web" function is used.

Hot News

The doctrine of "hot news" misappropriation could conceivably provide another legal avenue for protecting databases of factual information, including metadata.

In 1918, the U.S. Supreme Court first articulated the basis for a "hot news" claim and classified it as a part of unfair competition law in *International News Service v. Associated Press*.

The doctrine was originally envisioned as protecting the costly efforts of newspaper publishers and others in gathering time-sensitive and commercially valuable information, including factual information that, at least since *Feist*, is not protectable under copyright law.

Essentially, the "hot news" doctrine protects information gatherers from the free-riding efforts of competitors for a short period of time. Over the years, however, "hot news" misappropriation has received a fair amount of criticism for running counter to First Amendment freedom of the press values, and the doctrine has remained largely dormant.

Nonetheless, the "hot news" doctrine has made somewhat of a revival in recent years. In *Barclays Capital, Inc. v. TheFlyOnTheWall.com*, the Southern District of New York issued an injunction against TheFlyOnTheWall.com on the basis of "hot news" misappropriation.

The injunction stopped the website from publishing summaries of financial research (i.e. factual information) compiled by Wall Street firms until two hours after the research had first been published.

Additionally, and most interestingly, Iprevo Holdings filed a complaint against Goldman Sachs on May 5, 2010 in New York under a "hot news" theory, claiming that its database of contact information of people in the financial industry was dynamic and, thus, constituted "hot news," and that Goldman downloaded and used the information without authorization.

It remains to be seen, however, whether Iprevo Holdings will actually be able to show that its database of information was "time-sensitive" and that Goldman's free-riding efforts reduced Iprevo's incentive to compile its database such that Iprevo's existence was substantially threatened.

If the court follows *Barclays Capital*, Iprevo must prove both of these elements to succeed with its "hot news" claim. Future claims that call for the protection of metadata and/or databases as "hot news" are also likely to face these substantial hurtles.

European Directives

Lastly, and strictly by way of drawing a comparison between U.S. and European law, it should be noted that the European Union provides two levels of legal protection to databases under Directive 96/9/EC. First, the Directive provides that "original" databases, defined as databases which are the "intellectual creation" of the author, are entitled to copyright protection. Secondly, the Directive gives "non-original" databases, defined as databases where there has been "qualitatively or quantitatively a substantial investment in either the obtaining, verification or presentation of the contents," a right, regardless of copyright, to prevent the extraction or reutilization of the information contained therein for fifteen years. According to the European Union's evaluation of Directive 96/9/EC, the Directive's passage has had no discernable effect on the production of databases generally. Additionally, the evaluation found that the European Union's share of the global database market has decreased relative to that of the United States since the Directive's implementation.

Nancy E. Wolff, Partner, Cowan, De . aets, Abrahams & Sheppard, LLP, New York, NY

www.cdas.com

Protecting copyright – a summary

Metadata fields required to claim copyright protection

- the title or identifying information of a work
- the name of the author or copyright owner
- the terms and conditions of use of a work, identifying numbers or symbols, including hyperlinks, referring to such information,
- any information required by the Register of Copyrights
- The values of the fields above must be provided “in connection with” the work.

What must not be done

- Intentional removal of values from fields listed above
- Unintentional removal of field values by software