Document Watermarking

Introduction

Everyone has heard about watermarks, but the Oxford English Dictionary lists seven different common meanings for the word, quite apart from those used in the IT industry. So to help identify what is meant by watermarking please read the following sections that describe the commonest meanings in more detail, and then go on to describe which watermark methods are supported by Locklizard and how they may be implemented.

Watermarks in printed documents

Most people are familiar with two types of document watermarks which can be found in banknotes or on checks. In banknotes, these are recognizable designs that are **put into the paper** on which the documents are printed, whilst in checks they tend to be specific patterns. These watermarks are normally used to prevent people from being able to make fake copies, and, therefore, to be confident that the banknote or document is authentic.

There are also are a number of specialized printing techniques that make it possible to have printed watermarks that will vanish or appear if a protected document is photocopied or scanned. However, these rely upon high quality printing processes if they are to be successfully created. That is why documents using this technique are normally professionally printed, and distributed as paper documents rather than in electronic form.

There are other printing techniques that produce 'raised' printing, use magnetic inks or inks that will change colour if they get wet (either with water or other liquids), but they are very specialized and are used together with watermarks to provide higher levels of document protection and copy prevention.

Digital watermarks

With the advent of digital works (pictures, music, film and so on) digital watermarks have also been developed, however they are not always used in quite the same way as watermarks in documents. This is because, at least for the moment, it seems impractical to try and prevent people from making copies of computerised files, as computers will readily make perfect copies of the files they hold – something essential for backup and recovery.

The first type of digital watermarks you will see are those which are visible or obvious, and are intended to be so. These are usually images that are superimposed upon a still picture or a moving picture. The intention is either that in the event the images are copied then the ownership is not in dispute, or to prevent any realistic commercial use of the images if they are copied because their quality would not be acceptable.

The second type of watermarks are invisible. These digital watermarks are created by embedding extra information, commonly in the form of digital patterns, into the computer files containing the images or sounds to be protected. For this to be successful, the addition of the image to the information in the file must have no noticeable effect as far as the person seeing the subsequent picture or hearing the sound. There are two reasons for this. The first is not to reduce user satisfaction. The second is to make it much more difficult for someone to remove the pattern because they do not know what they are looking for.

In this case, the owner is using the watermark to help identify the source of the copy, since they cannot prevent the copying. The principle is to 'scan' the images or sounds and recognize the pattern, thus 'prove' the ownership of the original so that the owner can prosecute those found making or storing illegal copies. This type of watermark is often used in the film and music industries to identify pirated copies.

However, it is not a very useful method if you are trying to protect information such as text or similar. This is because to hide these kinds of watermarks in a file you need to be able to alter lots of bits without anyone being able to notice. But in text files there is nowhere to 'hide' the watermark. Also, whilst these watermarks act to identify the owner of the original they don't tell you anything about who was actually authorized to use them.

Several patents appear to have been granted over the use of digital watermarks as a method of demonstrating that a document that is encoded as a picture is genuine (but not necessarily original). These rely upon mathematical calculations performed on the picture, which has embedded in it, a unique numeric identification. This has a similar characteristic to the watermark manufactured into paper documents.

Watermarks in electronic documents

In Digital Rights Management (DRM) document protection solutions, it is possible to add 'printed' watermarks to documents being shown on a screen or printed out. These document watermarks may be static or dynamic, and they may also act in support of copy preventing or copy revealing. Typically, a combination of these is used.

For these watermarks to be effective, they need to be arranged in such a way that any attempt to remove or alter them is itself evident.

Static watermarks are those which do not change regardless of who opens and processes the watermarked document. They are used in the same as those on the banknote. It doesn't matter who gets the banknote, the watermark is the same. Static watermarks may be used to prove the authenticity of the document, or to make it difficult to produce another document and then pretend it is authentic. They are also used to identify the owner (or copyright controller/or the document). Static watermarks may also be used so that when a document is copied, a previously 'hidden' watermark is revealed which shows that the copied document is not an original.

Dynamic watermarks in the physical world were created by affixing a seal to a document, or stamping it (as a bank or post office does with documents, even today). These kinds of watermarks are used to identify the institution or individual associating themselves with the authenticity of the document. In PDF document watermarking terms these are watermarks that are added at the time of viewing or printing which identify the individual/enterprise that is authorized to use the protected PDF document.

Dynamic watermarks may be used as a form of copy resisting, because the individual allowing the watermarked document to be copied has their own identity linked to it, and they most likely do not want to be identified as the source of piracy.

Static watermarks may also be used for copy preventing. In this approach, a diffraction pattern (sometimes referred to as a Moiré fringe pattern) can be used. Whilst the human eye is subtle enough to be able to ignore the pattern, mechanical devices such as scanners and photocopiers become confused by the presence of the pattern and produce substandard copies or cannot convert the graphic scanned image back into text accurately. Obviously there is a balance between having a pattern that gives a document that can be used, and having a pattern strong enough to resist removal.

Using watermarks in Locklizard Safeguard document copy protection

When applying watermarks to electronic documents it is important to make sure that the documents themselves are adequately protected so that a user cannot easily remove the watermark from the text and pictures. And that is often the problem with watermarks applied to PDF documents, because as long as you know the permissions password (or you use a $10 piece of software to break it) then you can easily remove any watermarks that have been applied.

The Safeguard document protection system does not allow any modification of the protected content by the licensed user, so if the user is able to print out the document (only possible if they are allowed to do so), they must scan it back in, and then try to edit the result so that they could falsify the information that the document contained or pass it on to others without fear of identification. Since Locklizard do not use passwords, there are no automatic hacking tools that grant access to a document so that it can be edited or presented with different branding. The Safeguard document protection system also prevents screen grabbing so that content cannot be easily captured and manipulated with photo editing software.

Locklizard Safeguard allows you to add a number of different types of watermark to visible and printed documents protected by Locklizard. Watermarks are applied as a separate process from controlling and displaying text content, so watermarks shown when a document is viewed on screen may be quite different from those applied when a document is printed. You, as the publisher, decide whether watermarks will be displayed at view and/or print time and what those watermarks will look like.

We also provide with our document protection system a number of graphic watermarks that may help provide printed copies that resist document scanning. The objective of these watermarks is to make an original printed copy of an acceptable quality to the eye, but to cause OCR scanning packages to fail to convert text correctly when operating in recognition mode, and to cause poor quality documents to be reproduced. Publishers need to experiment with these to see which image type and opacity gives a result that on the one hand preserves the image quality appropriate to your clients, whilst on the other hand produces a result that would be suitably degraded if a printed copy is scanned.

You can also add dynamic text watermarks to identity the name and email address of the person/organization that is authorized to use/print the protected document so that their identity is linked to the document more than once. This, for example, would resist people trying to transfer a certification to another organization. Publishers only ever need to protect the document once for all users, rather than protecting documents for individual users, as our document protection software will automatically apply the correct watermark details at view and/or print time (user name, email address, date/time stamp).

Using watermarks to determine if a protected document has been altered

If a Locklizard protected document contains a visible watermark, then the document has to be scanned as a graphic to be reproduced properly on the forgery. The forger then has the problem of inserting the fake text in such a way as not to be evident when superimposed on both the underlying existing text and watermark. Whilst anything is theoretically possible, given enough time and money, it might be cheaper for the forgers to obtain the real document (i.e. a proper certification document).

Here, of course, we are assuming that you would be using a graphic watermark that is of sufficient complexity that it will prove inconvenient to forge. The document cannot be copied or altered in the protected electronic PDF format (unlike normal PDF files where the watermark can be easily removed), so in order to modify the document the user would have to print it (assuming they are authorized to do so), scan it back in, and then work out a way how they could alter it without damaging the display of the logo watermark, which would be very difficult to do.

Advantages of using document watermarks

So, using watermarks with protected documents offers many advantages to a publisher as a:

- copying deterrent;
- means of identifying the source of a printed document;
- means of determining whether a document has been altered.