International Standard

ISO/IEC 23078-1

First edition
2024-06

Information technology — Specification of digital rights management (DRM) technology for digital publications —

Part 1: Overview of copyright protection technologies in use in the publishing industry

Reference number
ISO/IEC 23078-1:2024(en)

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives or www.iec.ch/members_experts/refdocs).

ISO and IEC draw attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO and IEC take no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO and IEC had not received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at www.iso.org/patents and https://patents.iec.ch. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see www.iso.org/iso/foreword.html. In the IEC, see www.iec.ch/understanding-standards.

This document was prepared by Joint Technical Committee ISO/IEC JTC1, *Information technology,* Subcommittee SC 34*, Document description and processing languages*.

This document cancels and replaces ISO/IEC TS 23078-1:2020, which has been technically revised.

The main changes are as follows:

— in 3.3, the definition has been improved.

A list of all parts in the ISO/IEC 23078 series can be found on the ISO and IEC websites.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html and www.iec.ch/national-committees.

Introduction

Ever since digital publications have grown in popularity, copyright protection has been an important issue for authors and publishers.

While the distribution of digital publications around the world is mostly based on the open EPUB standard, most retailers are using proprietary technologies to enforce usage constraints on digital publications in order to impede oversharing of copyrighted content. The high level of interoperability and accessibility gained by the use of a standard publishing format is therefore cancelled by the use of proprietary and closed technologies: digital publications are only readable on specific devices or reading applications (a retailer "locked-in" syndrome); digital publications may not be accessed anymore if the distributor which protected the publication goes out of business or if the DRM technology evolves drastically. As a result, users are deprived of any control over their digital publications.

In reaction to these hindrances, watermarking and fingerprinting technologies have also been developed for digital publications. These are sometimes called "social DRM" which is a good way to describe the effect of the visible marks embedded into the content. Thanks to their presence and the personal information they contain, the “licensee” cares about the use of the content he/she has acquired: one would not like to see content associated with one's personal information freely shared on the web. But the term “social DRM” is misleading also, as watermarking and fingerprinting techniques do not enforce technical control on the use of digital media.

Requirements related to security levels differ depending on which part of the digital publishing market is addressed. Many trade publishers, in different countries, are satisfied with a protection based on watermarking; but in many other situations, publishers require a solution which technically enforces the digital rights they provide to their users. This is where DRM technologies come into play.

In most use cases, publishers are happy to adopt a DRM solution which guarantees an easy transfer of publications between devices and a certain level of fair-use, and provides permanent access to the publications acquired by their customers. However, in certain use cases, publishers require a stronger protection measure, which limits the capability for users to transfer publications from one device to another.

Information technology — Specification of digital rights management (DRM) technology for digital publications —

Part 1: Overview of copyright protection technologies in use in the publishing industry

1 Scope

This document describes three types of copyright protection technologies in use in the publishing industry:

— digital rights management (DRM) free protection, i.e. technologies which do not rely on content encryption but rather use content fingerprinting or watermarking, adequate for use cases where user convenience is the top priority;

— user key-based DRM protection, adequate where user constraints are limited;

— device key-based DRM protection, adequate where the transfer of publications from one device to another is severely constrained.

2 Normative references

There are no normative references in this document.

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

ISO and IEC maintain terminology databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at https://www.electropedia.org/

3.1
digital publication
set of constituent resources and associated metadata, organized together in a uniquely identifiable grouping

3.2
digital rights management
DRM
systematic approach to copyright protection to prevent unauthorized redistribution of digital media and restrict the ways consumers can use the content they've acquired

3.3
distributor
digital publication (3.1) retailer, public library, academic library or specialized intermediary that facilitates the distribution of electronic content between publishers and retailers

3.4
protected publication
digital publication (3.1) on which a *DRM* (3.2) solution has been applied

4 DRM free protection

4.1 General

Many users and librarians prefer plain digital publications to encrypted ones because of their undeniable advantages in terms of usability, portability or long-term preservation. On the other hand, many publishers are opposed to releasing their valuable contents in plaintext due to concerns about copyright infringement. Under this circumstance, some service providers adopt a protection measure which does not rely on encryption, such as fingerprinting or watermarking.

4.2 Fingerprinting

Fingerprinting means analysing content and extracting a unique set of inherent properties resilient to content transformation. Fingerprinted content is identified in a non-ambiguous way and therefore some use the term “content DNA” to describe a fingerprint.

Content fingerprint does not involve modifying the publication: the fingerprint is kept in a database and used to check if some random content is identical to the fingerprinted content. A user never sees any visible evidence that a digital fingerprint exists for the content he/she has acquired.

Digital fingerprinting can be used to track overshared content across the web with the help of monitoring services using specific web crawlers. But this technology is most actively used when a user wants to upload a digital publication onto a content sharing platform; before the upload is accepted, this platform may use a fingerprinting service and database to know if the upload is legal, which mitigates the risk of a lawsuit.

The difficulty is that a proper fingerprinting solution should be able to identify content independently of potential minor transformations and independently of a packaging format (e.g. EPUB vs PDF), which may be in practice more difficult for the digital publications than it is for audio-visual content. This document does not intend to find an open solution to this tricky problem.

4.3 Watermarking

Watermarking, on the other hand, is based on the creation of a personalized copy of a digital publication, which contains visible and/or invisible information tying the publication to its "licensee". The visible information, which is often personal information, acts as a reminder of the fact that free sharing of the digital publication is not allowed by the publisher, but it may also contain usage rules. The invisible information is usually based on multiple redundant techniques; it is meant to be difficult to suppress from the publication and a rogue user never knows if all watermarks have been deleted or not.

Because digital publication file formats such as EPUB are based on text characters, it is not possible to apply an algorithm to alter the text itself without being immediately perceptible; the text of the digital publication should remain as the author and publisher intended. Therefore, invisible watermarking techniques are applied to the space between characters on a line or the space between lines, in the form of non-printed character added to the textual content, or embedded in the illustrations of the digital publication (image steganography), etc.

Watermarking is based on obfuscation: a provider of a watermarking solution has to find imaginative ways to hide a payload of personal information into different parts of the publication; therefore, it is not possible to define open watermarking techniques, and this document focuses on DRM techniques instead.

Watermarking generally applies to retail sales, as well as certain special situations such as pre-release distribution of review copies; it is more rarely used with models such as library digital publication lending because it lacks technical enforcement for the end of a loan.

Watermarking is process intensive: the whole content should be processed for each "licensee" before distribution. Also, a watermarked EPUB cannot be cached efficiently on the web because of its customization. This is not the case with DRM technologies as described in Clause 5.

NOTE The European General Data Protection Regulation (GDPR) guards citizens against unapproved processing of personal information, imposes security measures to the storage of such information, requires proper response to breaches, and gives users the right to have their personal information deleted if no longer used.

All these issues should be properly handled before a watermarking solution is put in place by a distributor of digital publications.

5 DRM protection

5.1 General

In many situations, publishers require a solution which technically enforces the digital rights they provide to their users.

As of today, DRM technologies mainly rely on content encryption. Only the holder of the content decryption key can decipher the content of a digital publication and read it. The difficulty lies in transferring the content decryption key to the licensee in a safe way and making sure that a rogue user does not save the deciphered content in a clear format and overshare it.

The content decryption key is provided to the user in a "license", which also contains usage rights provided to the licensee by the copyright holder. Such rights can be a date before which (or after which) the content cannot be deciphered, the number of pages the licensee may print, or the number of characters the licensee may copy. A software application compliant with the DRM solution constrains the access to the digital publication to the scope of these rights.

DRM technologies are especially useful in the domain of electronic library lending (i.e. e-lending), where a digital publication is provided to a patron for a limited period of time. By definition of a loan, after some time, the patron should not be able to read the digital publication anymore; only a DRM can make such a rule enforceable.

The main criticisms against currently deployed DRM technologies are that they tend to lock users into proprietary solutions, be over-complex to use, lack accessibility for people confronted with specific disabilities (blindness, dyslexia …), leak private data, and block users from moving digital publications from a device to another when needed.

This document introduces two kinds of encryption-based protection, adapted to different levels of protection required by publishers and distributors. These vendor-neutral solutions aim at striking a proper balance between the requirements of authors and publishers in terms of copyright protection, and the needs of users in terms of ease of use and accessibility.

5.2 User key-based protection

5.2.1 General

A user-key based protection is based on a passphrase shared between the provider and the user of a digital publication. On compliant reading systems, the sole use of the passphrase unlocks the access to the digital publication only if the license associated with the publication allows such operation. Readium LCP is an example of user key-based protection for digital publications.

Such a solution targets the requirements specified in 5.2.2 and 5.2.3.

5.2.2 Requirements from publishers and distributors

— The solution must be built on reliable encryption technologies. It must be very difficult to break the encryption using brute-force attack.

— It must be possible to encrypt each publication with a different encryption key. It must not be possible to decrypt a large set of protected publications using a single decryption key.

— It must not be possible to derive the decryption key by only using information found in the license.

— It must not be possible to modify the rights embedded in the license.

— It must be possible to specify a time range outside of which reading the publication is not allowed.

— It must be possible to specify the number of pages the licensee can print, for the whole lifetime of the license.

— It must be possible to specify the number of characters the licensee can copy, for the whole lifetime of the license.

— It must be possible to extend the duration of a loan. Patrons don't always finish reading in time and request a loan extension for printed books; the same applies to digital publications.

— It must be possible to return a loan before the initial expiry time. Patrons may finish reading early; as the number of simultaneous loans is restricted in public libraries, as the number of simultaneous loans for a patron is also limited and as there may be people waiting for a digital publication, this adds fluidity to the library workflow.

— The solution must provide a solid safeguard against oversharing. The main issue for a trade publisher is to find protected content freely shared on the web, along with the key which opens it. If this happens, the distributor of the license must have a way to revoke the license and therefore make the key unusable.

— It must be possible to implement the solution on a wide range of distribution platforms and vendor lock-in must be avoided.

— Support related to DRM issues must be limited. The cost of support of DRM solutions is related to the number of issues users are facing when manipulating DRM protected digital publications.

— The solution must be suitable for different formats of digital publications. While EPUB is currently the preferred format, it must be possible to use the DRM technology on other formats, like W3C Web Publications or PDF.

5.2.3 Requirements from users

— The presence of a DRM must not be a hindrance for the use of digital publications. A user must not need to create a specific "DRM" account before being able to read DRM protected digital publications. The user must not experience exotic DRM errors when opening a digital publication for which he/she has acquired rights. As much as possible, the DRM must be invisible.

— It must be possible to get access to DRM protected digital publications from a wide range of devices and applications. Some users have specific needs which require specialized reading devices or applications: this is especially the case of people with visual disabilities who need text-to-speech functionalities, but also for dyslexic people who need reading applications with enhanced helpers.

— It must be possible to import (side load) a protected publication into a device or reading application.

— Importing a protected publication and reading it must be possible when offline, as users are not always connected and in certain countries, web access is still very expensive.

— A DRM solution must not leak personal information without explicit consent. Some personal information is given to a digital publication retailer by its client or public library patron when opening an account but must not leak to any third party (e.g. a DRM provider).

— In most situations, it must be possible to freely transfer a digital publication from a reading device or application to another. It must be possible to read the same digital publication on different devices: people read on a smartphone when commuting, on an e-ink device at night, change their smartphone

every 18 months. Limiting drastically the number of devices which is able to open a digital publication in its whole lifetime is a hindrance for users and a large source of support for distributors.

5.3 Device key-based protection

5.3.1 General

In certain situations, publishers have additional requirements.

A device-key based protection is adequate for use cases where transfer of publications from one device to another should be severely constrained. It is a variant of the user-key based protection, in which an additional asymmetric key is used to lock a license to a specific device, corresponding to a use case where an a posteriori control of the number of devices attached to a license is not an option; it aims at publishing high value content, where "fair use" is not accepted by its publisher. This is for instance the case in academic publishing for costly textbooks or high-priced market research reports. This can be also the case for textbooks licensed for use on a precise number of devices.

Such a solution targets the requirements introduced in 5.2, plus the additional requirements specified in 5.3.2 and 5.3.3.

5.3.2 Requirements from publishers and distributors

— The use of a protected digital publication must be locked to a specific device at a given time.

— Copying a digital publication from a reading device to another must be possible only within a limited number of devices, specified by the publisher of the digital publication.

— All specifications related to DRM interoperability must be fully open to the public in order to avoid the risk of a leak of closed but sensitive information with which the DRM protection mechanism can be disarmed in a short time. Nonetheless, the open specification must be secure enough.

— The potential impact of a compromised DRM must not be of global scale. Even if a one-click hack tool is released on the black market, the impact of such hack must be limited to a single distributor.

5.3.3 Requirements from users

— As long as the device is online, users must not experience difficulties for copying a digital publication from a reading device to another.