

<b>Digital Preservation Team</b>	<b>Preservation Assessment:</b> Mobipocket Format Preservation Assessment	<b>Date:</b> 27/06/2019
		<b>Version:</b> 1.1

## ***Mobipocket Format Preservation Assessment***

### *Document History*

Date	Version	Author(s)	Circulation
09/02/2018	1.0	Michael Day, David Russo, Simon Whibley	External
27/06/2019	1.1	Simon Whibley, Michael Day	External

British Library Digital Preservation Team  
[digitalpreservation@bl.uk](mailto:digitalpreservation@bl.uk)



This work is licensed under the [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

<b>Digital Preservation Team</b>	<b>Preservation Assessment:</b> Mobipocket Format Preservation Assessment	<b>Date:</b> 27/06/2019
		<b>Version:</b> 1.1

## 1. Introduction

This document provides a high-level, non-collection-specific assessment of the Mobipocket file format with regard to long-term preservation risks and the practicalities of preserving data in the format.

This format assessment is one of a series of assessments carried out by the British Library's Digital Preservation Team. An explanation of the criteria used in this assessment is provided in italics below most headings.

### 1.1 Scope

This document is specifically focused on the Mobipocket file format – sometimes also known as the MOBI format – although it will also note relationships with other formats and standards as necessary, including the Open eBook Publication Structure (OEBPS) and XHTML, as well as the derived AZW and KF8 formats used by Amazon. A general overview of eBook formats is provided in a separate file format assessment (1).

Please note that this assessment considers format issues only, and does not explore other factors essential to a preservation planning exercise, e.g. the collection specific characteristics that need to be considered before implementing preservation actions.

### 1.2 Mobipocket Summary

Mobipocket is an eBook file format originally created by Mobipocket.com in the early 2000s, but perhaps more commonly associated with Amazon, Inc., who sold eBook content in its associated formats from 2005, and later for use with its Kindle eBook reading devices. The Digital Preservation Coalition's Technical Watch report on *Preserving eBooks* provides the following general overview of Mobipocket's features (2 pp. 13-14):

MOBI allows DRM protection to be used, and supports reflowable content (fixed format text is supported only by inserting an image of a page [...]). Reflowable content in a MOBI package is included as XHTML. MOBI also includes CSS for rendition of the XHTML content. MOBI files are delivered as compressed binary files. The format may include advanced navigation controls and supports indexing. MOBI also allows JavaScript and frames. A MOBI eBook may have either a .mobi or .prc file extension. Amazon's KF8 and AZW are two variants of MOBI specific to Amazon's Kindle eBook reader device; they both use a different DRM scheme from the standard MOBI format. AZW does not allow JavaScript (KF8, the more recent Kindle format, does allow it).

The format itself comprises two distinct elements: an authored source format consisting of the marked-up text and structured resources which make up a book's content and metadata, and a compiled binary format which more efficiently packages (and optionally encrypts) that content into a single file for distribution and access (3). While the Mobipocket source format was documented in Mobipocket.com's Developer Center until its closure in 2016, no official specification for the compiled format has ever been published.

Mobipocket's source format was based on the Open eBook Publication Structure, also known as the Open eBook format (OEB), but with its own proprietary extensions and mark-up support (4). Like OEB, the source format for Mobipocket consists of an XML file that combines a list of source files (HTML files containing text, with formatting and images) with descriptive metadata about the eBook (e.g. title, author, publisher, cover). Some of Mobipocket's proprietary extensions included support for JavaScript, indexes, databases, SQL queries, HTML forms and framesets, and other formatting elements (5). While these files could be exchanged between actors in the publishing and retail supply chains, they would always need to be compiled into a binary format for delivery to end users.

Once compiled with Mobipocket.com's Creator or Mobigen software, the eBook would be viewable on reading devices, but no longer modifiable. Mobipocket's compiled format can be identified by the file extensions .mobi or .prc and ultimately derives from the PalmDOC format, a popular document format for portable devices running Palm OS circa the year 2000 (the .prc extension was used because the PalmOS did not support any file extensions except .prc or .pdb) (6). It may be compressed or uncompressed, as well as encrypted or unencrypted. What information is publically available on the compiled format has largely been gathered through reverse engineering (6).

<b>Digital Preservation Team</b>	<b>Preservation Assessment:</b> Mobipocket Format Preservation Assessment	<b>Date:</b> 27/06/2019
		<b>Version:</b> 1.1

## 2. Assessment

### 2.1 Development Status

*A summary of the development history of the format and an indication of its current status*

The Mobipocket format was originally developed by a French company called Mobipocket SA (sometimes known as Mobipocket.com), which was founded in March 2000. Mobipocket's original aim was to develop an eBook format that could render content on the handheld devices then known as PDAs (personal digital assistants). Dedicated eBook readers were in short supply in the early 2000s, so companies like Mobipocket and Peanut Press (later Palm Digital Media) were interested in developing formats that would enable eBooks to be read efficiently on PDAs and, eventually, on smartphones. In the early 2000s, Mobipocket developed a proprietary rendering technology and a hyphenation system for supporting the display of text on small screens (7). The company also invested heavily in digital rights management (DRM), which was used to secure eBooks sold by Mobipocket and its partners by linking downloaded eBooks with the serial numbers of their reading software.

Mobipocket was built upon the eBook and web standards of the early 2000s. It was essentially an implementation of the Open eBook format, but with the inclusion of proprietary extensions and DRM support, it was positioned as a format that could help protect a publisher's investment in eBook content. By April 2007, Mobipocket.com offered over 40,000 eBook titles for download in the Mobipocket format as well as the latest versions of the Mobipocket Reader and Creator software (8).

Amazon acquired Mobipocket in 2005. From that point, it distributed copyright-protected eBooks in AZW format, a proprietary variant of Mobipocket that supported a new DRM scheme able to restrict the use of files to particular devices (4). Within a few years of the acquisition, development on Mobipocket's Reader and Creator software ceased, both software packages receiving their final updates in mid-2008 with versions 6.2.608 and 4.2.41 respectively.

Amazon continued the format's development in private, and, in 2012, launched a new format known as KF8, which was based on AZW. The new format supported HTML5 and CSS3, thus enabling publishers and authors to include audio, video, and some level of interactivity for compatible readers. For backwards compatibility, KF8 files contain an additional version of the content in the older Mobipocket format (4). A newer format Kindle Format 10 (KFX) was released in August 2015 (9), and is able to support the "Enhanced Typesetting" available in more recent generations of Kindle viewers and apps (10).

The Mobipocket.com website remained available until it was finally closed by Amazon in November 2016.

### 2.2 Adoption and Usage

*An impression of how widely used the file format is, with reference to use in other memory organisations and their practical experiences of working with the format*

#### Mobipocket adoption

The Mobipocket file format was designed as a means to support the publication of eBooks that could be displayed on small, handheld devices. With reader software available for a wide range of devices, it became a popular delivery mechanism for the delivery of eBooks. Mobipocket.com was acquired by Amazon in 2005, who adapted the format to build up its own eBook operations.

While the format was reasonably supported by smaller reading devices and mobile phones, by the late 2010s, Mobipocket files were not usable on all eBook readers. In a 2011 newspaper comparison of eBook formats, Jack Schofield commented that Mobipocket had been a reasonable choice of format in the past, but noted that its main benefit by 2011 was the ease of creating Amazon's AZW files (11):

It is based on an earlier XHTML-based Open eBook standard, which was superseded by ePub, so it could be regarded as a "legacy" format. It is not supported by many leading e-readers including the Sony, Barnes & Noble and Kobo models. However, MobiPocket has two things in its favour: it is supported by software for many mobile phones, and Amazon's AZW is a version of MobiPocket with a new file extension and different identification numbers. This makes it relatively simple to convert a MobiPocket file into an AZW file. There is even a fix that enables MobiPocket vendors to enable the DRM-protected books they have sold to be redownloaded in Kindle format.

Many things changed when Amazon acquired Mobipocket in 2005 (12). According to one technology blog, the company was "stripped for parts and allowed to wither for years before being forced to shut down its retail and distribution operation in 2011" (13). From that point on, Mobipocket.com was no longer in the

<b>Digital Preservation Team</b>	<b>Preservation Assessment:</b> Mobipocket Format Preservation Assessment	<b>Date:</b> 27/06/2019
		<b>Version:</b> 1.1

eBook sales market and official support for the Mobipocket file format effectively ceased (although the reader and conversion software remained available for download for another five years).

A business assessment at the time recorded some of the reasons why Amazon purchased Mobipocket (12):

Amazon already offers eBooks and more than 1 million e-documents on its site, using downloadable software from Microsoft Corp. and Adobe Systems Inc. The purchase of Mobipocket will allow Amazon to use its own software to diversify product distribution methods, rather than relying on third-party providers.

After the takeover, Amazon promoted a variant of the Mobipocket format for use with its Kindle device. As has already been noted, early Kindles used a format called AZW, which was essentially the Mobipocket format using its high compression option (14). There is, therefore, a significant amount of overlap between Mobipocket files and AZW, to the extent that both files without DRM are essentially identical (15). For example, one way of generating AZW files would be to create Mobipocket files and then rename the extensions to AZW (14).

### **Mobipocket in Memory Institutions**

Few memory institutions appear to be collecting eBooks in Mobipocket format. Most are primarily focused on other formats, such as EPUB or PDF.

On behalf of the UK Legal Deposit Libraries, the British Library has been collecting eBooks as part of Non-Print Legal Deposit (NPLD) since 2013. To date, this has mainly included content in EPUB and PDF. Decisions on how to process the small amount of content received in Mobipocket format have yet to be finalised, and this assessment is one contribution to that process.

As of 2019, the Library of Congress has approximately 23,000 Mobipocket files created for the Internet Archive as part of a digitisation project. The Library does not provide direct access to the files but they are downloadable via the Internet Archive (16). Mobipocket is not included in the Library of Congress's Recommended Formats Statement for Textual Works - Digital (17).

## **2.3 Software Support**

### **2.3.1 Rendering Software Support**

*An overall impression of software support for rendering the format with reference to: typical desktop software; and current support on British Library reading room PCs*

While it was in operation, Mobipocket.com offered several software products, which broadly included both reader software and a suite of creator/conversion tools focused on publishers and other content creators. While functionality and target markets evolved over the years, major Mobipocket.com products included:

- Mobipocket Reader: a freely-downloadable eBook reader program that was initially designed to run on the popular PDA environments of the early 2000s, namely Palm OS, Windows CE, Pocket PC, Psion Epos 32, and Franklin eBookMan (18). In due course, versions of the reader became available for other environments, including Windows PCs, as well as mobile operating systems including Blackberry, Windows Mobile, and Symbian. The final official version of Mobipocket Reader appears to have been version 6.2, which was released in 2008 (19). The reader software was able to display eBook content in Mobipocket format, but also supported advanced features like annotation.
- Mobipocket Creator<sup>1</sup> (formerly Mobipocket Publisher): a toolkit that was designed to be used by eBook publishers to generate content in the Mobipocket format, initially from HTML and TXT files (7). Later on, Mobipocket.com produced a free "Personal Edition" of Publisher that was intended to help authors create their own eBooks, e.g. from Word or PDF files (20).
- Mobipocket Emulator: a PDA simulator allowing users to preview how their compiled Mobipocket eBook might display on devices with differing screen and font sizes (21).

These software products have not been available for download from the Mobipocket.com website since late 2016, when the website was officially closed down.

<sup>1</sup> The self-publishing service, Kindle Direct Publishing, has a list of supported formats (58) and Mobi (.prc) files created with Mobipocket Creator are no longer supported.

<b>Digital Preservation Team</b>	<b>Preservation Assessment:</b> Mobipocket Format Preservation Assessment	<b>Date:</b> 27/06/2019
		<b>Version:</b> 1.1

Unencrypted Mobipocket files can also be read unaltered on Amazon's Kindle devices and applications. Amazon currently markets its free Kindle App for four main technical ecosystems: iOS (Apple), Android, Mac and PC (22). The PC version of this application was previously known as Kindle for PC (23), and sometimes more generically as Kindle Reader for Desktop (24).

Other reading software that claims to be able to render Mobipocket files include: Calibre, Sumatra PDF, Stanza Desktop, Mobi File Reader, FBReader, Okular, and IceCream Ebook Reader (25) (24). Of these software products, Calibre positions itself as an eBook manager, able to deal with multiple formats (including Mobipocket and EPUB) – assuming that they are not protected by DRM (26) (27).

### 2.3.2 Preservation Software Support

*An impression of the availability and effectiveness of software for managing and preserving instances of the file format*

#### **Format identification**

Mobipocket has an outline record listed as “PocketMobi (Palm Resource) File” (PUID: fmt/396) by the PRONOM technical registry (28). It is therefore identified correctly by Siegfried (29) and DROID (30) which both use the PRONOM database. The format is also identified by Apache Tika (31). However, none of the identification tools tested were able to determine specific versions of the Mobipocket files though version information in the file is usually dependent on the generation tool (6) (32).

#### **Validation, Conformance Checking and Detecting Preservation Risks**

JHOVE does not have a module that identifies or validates Mobipocket content (33), nor does any other software appear to exist for validating compiled Mobipocket files. Mobigen and KindleGen software may provide a certain degree of validation for Mobipocket source files, but are both limited to the format versions and features which they support generating.

#### **Metadata Extraction**

The Mobipocket format itself contains some descriptive metadata, which is stored either in the MOBI header, or the EXTH header that follows in some files (6). However, citing Goyal (34), the Digital Preservation Coalition Technical Watch Report on Preserving eBooks notes that there is anecdotal evidence that “the metadata (such as title and author) within formats such as MOBI and EPUB is neither robust nor accurate” (2 p. 17).

Notwithstanding this, there is a Python-based tool called KindleUnpack, which allows the extraction of MOBI source fields and should therefore be able to extract metadata from unencrypted Mobipocket files, the source HTML and image files (35), and in some cases even the log created during compilation but doesn't work with DRM'd files.

In practice, some descriptive metadata about eBooks might be available directly from publishers, e.g. in the form of ONIX (2 p. 17).

#### **Migration**

The need for authors and publishers to provide eBooks in different formats has created a market for eBook migration tools and services, including to and from Mobipocket. A number of web-based file conversion services are able to convert Mobipocket files to and from formats like EPUB or PDF, such as DocsPal (36), Zamzar (37) and EPUB to MOBI Converter (38), as well as open source software like Calibre, which supports the conversion to and from many other input and output formats, including Mobipocket (26).

The official Mobipocket Creator software (which is no longer available for download) allowed users to create Mobipocket eBooks from EPUB, Microsoft Office, PDF, and several other file formats, as did the Windows version of Mobipocket Reader (6). Alongside Creator and Reader, Mobipocket.com also offered Mobigen, a configurable command-line tool for compiling Mobipocket eBooks from HTML, OEB 1.0 and 2.0, or Mobipocket source files (39). Mobigen was formerly known as PRCGen, and would later develop into Amazon's KindleGen (40).

Mobipocket Office Companion was a commercial utility for exporting Microsoft Office files for viewing on Mobipocket Reader-compatible systems. Version 1 supported Word 97, PowerPoint 97, and Excel 97 formats (41), while version 2 supported Word 2000, PowerPoint 2000, Excel 2000, Outlook 2000, Access 2000, FrontPage 2000, and Visio 2002 formats. Both versions of the software offered a two-week trial, and

<b>Digital Preservation Team</b>	<b>Preservation Assessment:</b> Mobipocket Format Preservation Assessment	<b>Date:</b> 27/06/2019
		<b>Version:</b> 1.1

were available from 2001 to 2006 from Mobipocket.com (42). Some of their functionality may have been merged into later versions of Mobipocket Reader for Windows.

KindleGen is Amazon's evolution of the Mobigen command-line tool, which, like its predecessor, enables a number of source formats to be converted into the appropriate version of Mobipocket, including KF8 (43). As of 2019, the KindleGen information page is still available (43) but is not referenced in the Kindle Direct Publishing site launched in 2017 (44). Mobipocket is also one of the formats supported by Amazon's Kindle Previewer 3 tool (45) which enables authors, publishers and service providers to see how ebooks might render when delivered to Kindle customers (and which also supports Amazon's Enhanced Typesetting engine (10)).

Discoverysoft's PDF ePUB to Kindle and PDF to ePub MOBI tools allow the conversion of EPUB or PDF into Mobipocket format (46). Similarly, Secursoft's EPUB to Kindle converter supports Mobipocket as an output format (47). Likewise, the book creation tool for Mac, Vellum (48), allows you to import and export in a variety of formats include Mobipocket (49).

Finally, KindleUnpack is a Python-based tool that can re-create Mobipocket (or AZW) source files from their compiled forms, though only if unprotected by DRM. Not all information may be recoverable, however, e.g. the filenames used by the source files may not necessarily be the same as those used originally, but "an unpacked set of files should be able to [...] recreate the same [compiled form] using standard MOBI or Kindle generating tools". It can also extract the original source files if they were separately embedded during compilation (35).

## 2.4 Documentation and Guidance

*An indication of the availability of practical documentation or guidance with specific reference to the facilitation of any recommended actions*

There seems to be no formal documentation currently available about the Mobipocket format on the web. Up until 2016, the Mobipocket.com website contained detailed information regarding the source format, and some background information on the compiled format, as part of its developer pages, which were primarily intended to support publishers and authors wanting to create content in the format. These pages are now only available through the Internet Archive (3).

Some third party, reverse-engineered, documentation about the compiled Mobipocket format, including information on headers and metadata, is available from the MobileRead wiki (6).

## 2.5 Complexity

*An impression of the complexity of the format with respect to the impact this is likely to have on the British Library managing or working with content in this format. What level of expertise in the format is required to have confidence in management and preservation?*

While the Mobipocket source format was built upon the Open eBook standard, the eBooks actually distributed to consumers use a proprietary binary format that is often compressed, and may be encrypted with one of three different DRM schemes (3; 50).

The compression algorithms used in Mobipocket, whilst not common, are fairly well documented and should require no special handling as long as rendering software remains available. Knowledge of the types of DRM encryption that can be found would be important when considering access and migration requirements. Successful decryption of two of the three Mobipocket encryption schemes, for example, are currently contingent on information either only available from the original reading environments or from the eBook's creator, and so should be considered as close to item acquisition as possible.

## 2.6 Embedded or Attached Content

*The potential for embedding or attaching files of similar or different formats, and the likely implications of this*

Mobipocket files may include some embedded files, but these would chiefly be those incorporated within the XHTML component of an eBook. For Mobipocket files, these images would be stored as raster graphics, as there was no scope within the original reader ecosystem to support vector graphics (or audio files and other multimedia). In the original Mobipocket documentation, it was recommended that images in the (non-consumer-facing) source files should be in GIF format for graphics that include text (e.g. mathematical equations, tables) and JPEG for photographs, but with a 63 KB limit with anything above that resized or recompressed (51).

<b>Digital Preservation Team</b>	<b>Preservation Assessment:</b> Mobipocket Format Preservation Assessment	<b>Date:</b> 27/06/2019
		<b>Version:</b> 1.1

In general, embedded content is not handled very well in Amazon's Kindle ecosystem. The eBooks Architects blog (52) reports:

Amazon's support for embedded media is very limited. It only works when embedded in the old Mobipocket 7 format,<sup>2</sup> not in the new Kindle Format 8, and it is only supported in the Kindle iOS App and in 2nd Generation or later Kindle Fire devices. It does not currently work in the First Generation Kindle Fire or the E Ink Kindle devices.

Besides content intended for rendering, Mobipocket files compiled with later versions of KindleGen may also embed a compressed form of the original source material, as well as compilation metadata such as the KindleGen parameters used (6).

## 2.7 External Dependencies

*An indication of the possibility of content external to an instance of the file format that is complimentary or even essential to the intellectual content of the instance*

Mobipocket files are designed to package all necessary content, so the main external dependencies relate to the software applications required to render them. Even so, certain applications may capture and store personalised information, or user-generated content, associated with the individual eBooks that could provide a valuable context around how they were used.

Official Mobipocket and Amazon reading software, for instance, often allowed users to create personal annotations, bookmarks, and other content that would then be stored in a separate .mbp file (Mobipocket Notes File) alongside the eBook. Like Mobipocket, the format of this file is proprietary, may be encrypted, and has only been partially reverse engineered, allowing only certain information to be extracted under certain circumstances (53).

## 2.8 Legal Issues

*Legal impediments to the use, management or preservation of instances of the file format*

The main legal impediment to the use, management and preservation of Mobipocket files would be the potential presence of DRM. Mobipocket was designed to be used in a tightly controlled environment whereby the use of files could be restricted to particular devices or applications. In practice, however, many Mobipocket eBooks are provided in forms free of DRM.

## 2.9 Technical Protection Mechanisms

*Encryption, Digital Rights Management and any other technical mechanisms that might restrict usage, management or preservation of instances of the file format*

Mobipocket was designed to be used with DRM schemes that would restrict reading to particular devices or applications. The inclusion of DRM is a more general issue for eBooks and eBook formats. The Digital Preservation Coalition Technical Watch Report on *Preserving eBooks* provides a useful summary (2 p. 10):

DRM constitutes a challenge in the preservation of eBooks. It is challenging from the business model perspective – if a book is 'sold' with limited use, what preservation rights come with that sale? And it is also challenging from a technical perspective, where DRM can impede the preservability of an eBook. Further, a change in digital rights technologies can be seen as a specialized case of format obsolescence.

[...]

If however DRM – and, more particularly, its encryption – cannot be removed by preservation institutions from the preservation copy of an object, that object will be opaque to future viewers. Preservation in such cases would be byte-level preservation of an unintelligible digital object.

More specifically, traditional Mobipocket files may be encrypted in one of three ways: content encryption, content encryption with a password, or content encryption with DRM. The first option simply limits a Mobipocket eBook to being viewed through an official Mobipocket reading application, and is not tied to any specific accounts, devices, or installations. Encryption with a password works similarly, but with the addition of requiring a password upon access. Encryption with DRM ties a Mobipocket eBook to one or more of a user's Mobipocket reader installations, using those installation's serial numbers, or personal identifiers (50).

---

<sup>2</sup> Mobipocket 7 is an alternative name for the Mobipocket format.

<b>Digital Preservation Team</b>	<b>Preservation Assessment:</b> Mobipocket Format Preservation Assessment	<b>Date:</b> 27/06/2019
		<b>Version:</b> 1.1

There are a number of tools and plugins available online, along with various guides which explain how to remove simple content encryption or DRM from Mobipocket files (54; 55; 56), although the legal status of doing so is not always clear. It should also be noted that the removal of Mobipocket DRM from an eBook currently requires knowing the personal identifiers of the readers for which it was originally encrypted – information which may no longer be available.

## 2.10 Other Preservation Risks

*Other evidence based preservation risks, noting that many known preservation risks are format specific and do not easily fit under any of the sustainability factors above*

While the format is supported by a wide-range of software applications, Mobipocket is essentially a proprietary format. While this in itself may not represent a significant preservation risk, there is a perception that it, especially when combined with DRM, may make content stored in the format ultimately difficult to preserve. The Digital Preservation Coalition's Technical Watch Report on Preserving eBooks concluded that, "although Amazon's Kindle formats (MOBI, AZW, and KF8) dominate in the market of books for use on dedicated readers or applications, the proprietary nature of those formats, especially in combination with their DRM regimes, comprise a long-term preservation risk" (2 p. 16).

## 2.11 Preservation Risk Summary

*A summary of preservation risks and recommended actions (where possible).*

As a key part of the Kindle eBook ecosystem, the Mobipocket format is reasonably widely used and its use is supported by a range of software applications. However, while Mobipocket source files are built upon open web standards (e.g. HTML and CSS) and the Open eBook standard, the eBooks actually distributed to consumers are in a binary format that uses additional compression. While there is software available (e.g., KindleUnpack) that claims to be able to re-create the component source files of a Mobipocket file, this will only work if the item is unprotected by DRM.

In addition, Amazon has focused since 2012 on the distribution of eBook content in more advanced formats than Mobipocket, initially KF8 – which is based on an updated set of web standards (e.g. HTML5, CSS3). The relationships between the main Amazon formats (Mobipocket, AZW, KF8 and now, KFX) are complex, and DRM-based content is often additionally 'locked-down' to particular devices or instances of reader software. There are tools available that can help circumvent this, but these may not be legal or robust.

Despite the advent of KF8, Mobipocket has retained an important role within the eBook ecosystem based around the Amazon Kindle, not least because it acts a kind of *de facto* baseline standard. Some publishers continue to distribute Mobipocket editions of works because they assume that there will be a potential market for them from Kindle users (and the cost of offering an additional format option may not be that prohibitive). For the same reason, some organisations that distribute digitised books – e.g. the Internet Archive – include Mobipocket as one of their main download options (the other formats include EPUB, PDF, DAISY, plain text, source image files, etc.).

Mobipocket has a number of associated preservation risks, which can be summarised under the following headings:

- **Proprietary format**
  - While built upon open standards, Mobipocket remains essentially a proprietary format. This represents a preservation risk, albeit one that may be partly mitigated by the existence of tools that are able to recreate source files (e.g. KindleUnpack) or can convert Mobipocket into other eBook formats (e.g. Calibre).
- **Official documentation and support**
  - There seems to be no openly available official documentation on the Mobipocket format. In partial mitigation, some inferred technical information about the format has been collated on the MobileRead wiki (6). The (now defunct) Mobipocket.com website did contain some basic information on the format and how it was constructed, but this is only available now from the Internet Archive (3). The closure of the Mobipocket.com web site in 2016 suggests that Amazon already views Mobipocket as a legacy format.
- **Availability of reader software**
  - The Mobipocket Reader software is no longer being updated and has not been available for download from the Mobipocket.com web site since 2016. Mitigating this, the Mobipocket format itself can be read on a wide range of other eBook devices and applications. Amazon's own Kindle App is a popular consumer choice, but may be too



<b>Digital Preservation Team</b>	<b>Preservation Assessment:</b> Mobipocket Format Preservation Assessment	<b>Date:</b> 27/06/2019
		<b>Version:</b> 1.1

strongly linked to the Kindle ecosystem to be useful as a preservation and access platform. Other eBook reader or management programs, e.g. Calibre, may be able to provide a more neutral choice of platform. From a preservation planning perspective, it would be useful to consider which particular reader or management software would need to be preserved in order to provide on-going access.

- **Encryption**
  - As with many eBooks, a major preservation risk with content in Mobipocket format would be the presence of DRM. In the UK, the Legal Deposit Libraries (Non-Print Works) Regulations (57) currently oblige publishers to provide "any other data or information necessary to access" works – which would include DRM – but the presence of DRM or other encryption would still need to be checked as part of ingest processing and, if possible, removed prior to ingest. Items with DRM may not be able to be preserved or made accessible in any meaningful manner, though items with the simplest form of Mobipocket encryption could probably still be accessible, as long as access to official Mobipocket reading software is also preserved.
- **Invalid or badly-formed files**
  - There are no third-party tools available to validate Mobipocket files. As the format is not open and there is no public specification, the development of such a tool outside of Amazon (or previously Mobipocket) make this extremely difficult. It is possible that Amazon's KindleGen (43) tool may include some validation features for Mobipocket source material, but this would need to be investigated in more detail (and they may not work outside the context of that tool's purpose). There is also the Kindle Previewer tool, which is able to emulate how eBooks display across Kindle devices and apps (45), although it is not primarily a format validation tool.
- **Format stability**
  - Mobipocket is part of a family of related formats used in Amazon's Kindle ecosystem. Amazon's focus now seems to be mainly on KF8 and the even newer KFX format, which is more capable in a number of ways, but may mean that continued support for older format types (e.g., Mobipocket, AZW) will reduce over time. This will need to be monitored.
- **Embedded content**
  - Embedded content is less likely to be an issue than with other eBook formats, as these are likely to primarily be images, and mostly in "web-friendly" formats like JPEG or GIF. Embedded content, however, will need to be rendered correctly when providing access.

### 3. Recommendations for Action

*Recommended actions in usage and handling of the format. Recommend actions in the support or development of software applications that provide, or have the potential to provide, significant risk mitigation for the format. Note that these recommendations do not take into account other requirements such as those driven by specific British Library collections, or non-preservation issues such as resourcing.*

#### **Handling Recommendations**

- All Mobipocket files deposited / received should be tested for the presence and type of DRM. Any auxiliary files associated with the eBooks should also be considered for preservation.

#### **Knowledge Recommendations**

- Research is required into validation tools that could be used for Mobipocket files though the lack of a formal specification makes development unlikely.

#### **Software Recommendations**

- The collection of reading software capable of rendering or otherwise handling the Mobipocket format and its auxiliary files, with particular attention paid to official Mobipocket and Amazon software, such as Mobipocket Reader and the Kindle applications.

<b>Digital Preservation Team</b>	<b>Preservation Assessment:</b> Mobipocket Format Preservation Assessment	<b>Date:</b> 27/06/2019
		<b>Version:</b> 1.1

### Monitoring Recommendations

- Review the status of the Mobipocket format annually, in conjunction with Kindle eBook formats more generally.

## 4. References

- Whibley, Simon, May, Peter and Wheatley, Paul.** eBook Summary Assessment. *DPC Wiki*. [Online] 3 February 2017. [Cited: 27 June 2019.] [http://wiki.dpconline.org/images/0/04/EBook\\_Summary\\_Assessment\\_v1.0.pdf](http://wiki.dpconline.org/images/0/04/EBook_Summary_Assessment_v1.0.pdf).
- Kirchhof, Amy, Morrisey, Sheila M.** Preserving eBooks. *Digital Preservation Coalition*. [Online] June 2014. [Cited: 19 June 2017.] DPC Technology Watch Report 14-01. <http://dx.doi.org/10.7207/twr14-01>.
- Mobipocket.com.** What is the Mobipocket file format. *Internet Archive*. [Online] 24 April 2008. [Cited: 13 June 2017.] <https://web.archive.org/web/20160414103204/http://www.mobipocket.com/dev/article.asp?BaseFolder=prcgen&File=mobiformat.htm>.
- Bläsi, Christoph and Rothlauf, Franz.** *On the interoperability of eBook formats*. Brussels : European and International Booksellers Federation, 2013.
- Mobipocket.com.** Creating content. *Internet Archive*. [Online] 08 March 2016. [Cited: 12 April 2017.] <https://web.archive.org/web/20160308093338/http://www.mobipocket.com/dev/article.asp?BaseFolder=prcgen>.
- MobileRead.** MOBI. *MobileRead wiki*. [Online] 28 September 2016. [Cited: 8 August 2017.] <https://wiki.mobileread.com/wiki/Mobi>.
- Felici, Jim.** Mobipocket making headway. *Seybold Report: Analyzing Publishing Technologies*. 4 February 2001, Vol. 1, no. 1, p. 38. Accessed via EBSCOhost Business Source Complete.
- Mobipocket.com.** About Mobipocket. [Online] 19 April 2007. [Cited: 1 August 2017.] Available from the Internet Archive. <https://web.archive.org/web/20070419211111/http://www.mobipocket.com:80/EN/Corporate/Profil.asp?Language=EN>.
- Kindle File Format. *Wikipedia*. [Online] 15 May 2019. [Cited: 21 June 2019.] [https://en.wikipedia.org/wiki/Kindle\\_File\\_Format](https://en.wikipedia.org/wiki/Kindle_File_Format).
- Amazon.** Enhanced typesetting and reading customization. [Online] [Cited: 27 June 2019.] <https://www.amazon.com/b/?node=11516960011>.
- Schofield, Jack.** Which is the best format for ebooks? *The Guardian*. [Online] 15 September 2011. [Cited: 13 June 2017.] <https://www.theguardian.com/technology/askjack/2011/sep/15/ebook-format-drm-kindle>.
- Bolt, Kristen Millares.** Retail notebook: Amazon buys e-book software. *Seattle Post-Intelligencer Reporter*. [Online] 15 April 2005. [Cited: 30 March 2017.] <http://www.seattlepi.com/business/article/Retail-Notebook-Amazon-buys-e-book-software-1171055.php>.
- Hoffelder, Nate.** Mobipocket to turn off its website and servers on 31 October. *The Digital Reader*. [Online] 23 October 2016. [Cited: 30 March 2017.] <http://the-digital-reader.com/2016/10/23/mobipocket-turn-off-website-servers-31-october/>.
- MobileRead.** AZW. *MobileRead wiki*. [Online] 31 August 2016. [Cited: 14 June 2017.] <https://wiki.mobileread.com/wiki/AZW>.
- Basu, Saikat.** GT Explains: What is the difference between EPUB, MOBI, AZW and PDF eBook formats? *Guiding Tech*. [Online] 8 April 2015. [Cited: 15 June 2017.] <http://www.guidingtech.com/9661/difference-between-epub-mobi-azw-pdf-ebook-formats/>.
- Sustainability of Digital Formats: Mobipocket File Format. [Online] 20 February 2019. [Cited: 20 June 2019.] <https://www.loc.gov/preservation/digital/formats/fdd/fdd000472.shtml>.
- Recommended Formats Statement - Textual Works - Digital. *Library of Congress*. [Online] 2019. [Cited: 20 June 2019.] <https://www.loc.gov/preservation/resources/rfs/textmus.html#digital>.
- Mobipocket.com.** Mobipocket Reader 4.1 for your PDA. [Online] 8 November 2001. [Cited: 30 March 2017.] Available from the Internet Archive. <https://web.archive.org/web/20011108185547/http://www.mobipocket.com/en/DownloadSoft/ProductDetailsReader.asp>.
- . Mobipocket Reader Desktop 6.2. [Online] 17 March 2016. [Cited: 30 March 2017.] Available from the Internet Archive.

<b>Digital Preservation Team</b>	<b>Preservation Assessment:</b> Mobipocket Format Preservation Assessment	<b>Date:</b> 27/06/2019
		<b>Version:</b> 1.1

- <https://web.archive.org/web/20160317103248/http://www.mobipocket.com/en/DownloadSoft/ProductDetailsReader.asp>.
20. —. MobiPocket Publisher 3.0. [Online] 6 February 2002. [Cited: 30 March 2017.] Available from the Internet Archive.  
<https://web.archive.org/web/20020206044331/http://www.mobipocket.com/en/DownloadSoft/DownloadPublisherStep1.asp>.
21. —. Mobipocket Emulator. *Internet Archive*. [Online] 26 October 2016. [Cited: 25 July 2017.]  
<https://web.archive.org/web/20161026215204/http://www.mobipocket.com/dev/article.asp?BaseFolder=emul>.
22. **Amazon**. Get the top rated free Kindle App. *Amazon*. [Online] [Cited: 21 June 2017.]  
<https://www.amazon.co.uk/kindle-dbs/fd/kcp>.
23. —. Kindle for PC. *Amazon*. [Online] [Cited: 21 June 2017.]  
<https://www.amazon.co.uk/gp/help/customer/display.html?nodeId=201245960>.
24. Top 7 Free Mobi Readers for PC – Windows 7/8.1/10 | Read Mobi Ebooks on Computer & Laptop. *TechnoRange*. [Online] 5 February 2016. [Cited: 15 June 2017.]  
<http://www.technorange.com/2016/02/top-7-free-mobi-readers-for-pc-windows-7-8-1-10-read-mobi-ebooks-on-computer-laptop/>.
25. **Fisher, Tim**. What is a MOBI File? *Lifewire*. [Online] 9 March 2017. [Cited: 15 June 2017.] <https://www.lifewire.com/what-is-a-mobi-file-2621997>.
26. Wikipedia. Calibre (software). [Online] [Cited: 15 June 2017.]  
[https://en.wikipedia.org/wiki/Calibre\\_\(software\)](https://en.wikipedia.org/wiki/Calibre_(software)).
27. **Roy, Lachlan**. A user's guide to Calibre eBook Manager . *MUO*. [Online] 9 June 2011. [Cited: 15 June 2017.] <http://www.makeuseof.com/tag/download-open-book-managing-your-ebooks-with-calibre/>.
28. Details for: PocketMobi (Palm Resource) File. *PRONOM the technical registry*. [Online] 11 June 2012. [Cited: 16 June 2017.]  
<http://www.nationalarchives.gov.uk/PRONOM/fmt/396>.
29. Siegfried. *ITforArchivists*. [Online] 2017. [Cited: 10 April 2017.]  
<http://www.itforarchivists.com/siegfried>.
30. Download DROID: file format identification tool. *National Archives*. [Online] 2015. [Cited: 22 May 2015.] <http://www.nationalarchives.gov.uk/information-management/manage-information/preserving-digital-records/droid/>.
31. **Apache Software Foundation**. Apache Tika. [Online] 2015. [Cited: 5 May 2015.]  
<https://tika.apache.org/0.5/formats.html>.
32. **Goyal, Kovid**. mobi.txt. *GitHub*. [Online] 9 January 2017. [Cited: 24 June 2019.]  
[https://github.com/kovidgoyal/calibre/blob/master/format\\_docs/pdb/mobi.txt](https://github.com/kovidgoyal/calibre/blob/master/format_docs/pdb/mobi.txt).
33. Standard Modules. *JHOVE*. [Online] October 2015. [Cited: 13 April 2017.]  
<http://jhove.openpreservation.org/modules/>.
34. **Goyal, K**. Editing Ebook metadata. *Calibre Manual*. 2013.
35. **MobileRead**. KindleUnpack. *MobileRead wiki*. [Online] 1 October 2016. [Cited: 19 June 2017.] <https://wiki.mobileread.com/wiki/KindleUnpack>.
36. DocsPal. [Online] [Cited: 15 June 2017.] <http://www.docspal.com/>.
37. Zamzar. [Online] [Cited: 15 June 2017.] <http://www.zamzar.com/>.
38. EPUB to MOBI Converter. *Reedsy*. [Online] 2019. [Cited: 24 June 2019.]  
<https://blog.reedsy.com/epub-to-mobi-converter/>.
39. **Mobipocket.com**. Mobipocket Developer Center. *Internet Archive*. [Online] 05 July 2009. [Cited: 11 July 2017.]  
<https://web.archive.org/web/20090705194622/http://www.mobipocket.com/dev/default.asp>.
40. **Amazon**. Release Notes. *KindleGen Manual*. [Online] 2014. [Cited: 11 July 2017.]  
<https://kindlegen.s3.amazonaws.com/docs/english/Release%20Notes.html>.
41. **Mobipocket.com**. Mobipocket Office Companion Product Details. *Internet Archive*. [Online] 21 November 2001. [Cited: 10 June 2017.]  
<https://web.archive.org/web/20011121082401/http://www.mobipocket.com:80/en/DownloadSoft/ProductDetailsOfficeCompanion.asp>.
42. —. Mobipocket Office Companion 2.1. *Internet Archive*. [Online] 16 July 2006. [Cited: 11 July 2017.]  
<https://web.archive.org/web/20060716084416/http://www.mobipocket.com:80/en/DownloadSoft/ProductDetailsOfficeComp.asp>.
43. **Amazon**. KindleGen. *Amazon.com*. [Online] [Cited: 22 June 2017.]  
<https://www.amazon.com/gp/feature.html?docId=1000765211>.

<b>Digital Preservation Team</b>	<b>Preservation Assessment:</b> Mobipocket Format Preservation Assessment	<b>Date:</b> 27/06/2019
		<b>Version:</b> 1.1

44. —. Kindle Direct Publishing. *Amazon*. [Online] 2019. [Cited: 20 June 2019.] [https://kdp.amazon.com/en\\_US/](https://kdp.amazon.com/en_US/).
45. —. Kindle Previewer. *Amazon.com*. [Online] [Cited: 22 June 2017.] <https://www.amazon.com/gp/feature.html?docId=1000765261>.
46. PDF ePUB to Kindle Tool. *Discoverysoft*. [Online] 2016. [Cited: 16 June 2017.] <http://www.ebook-tool.com/pdf-epub-to-kindle-tool.htm>.
47. EPUB to Kindle Converter. *Secursoft.net*. [Online] 2017. [Cited: 16 June 2017.] <https://epub-to-kindle-converter.secursoft.net/>.
48. Vellum. *Vellum*. [Online] 29 May 2019. [Cited: 24 June 2019.] <https://vellum.pub/>.
49. Review: Vellum Software for Book Formatting. *Reedsy*. [Online] 30 April 2019. [Cited: 24 June 2019.] <https://blog.reedsy.com/vellum-software/>.
50. **Mobipocket.com**. Security/Encryption Options. *Internet Archive*. [Online] 7 September 2015. [Cited: 3 August 2017.] <https://web.archive.org/web/20150907235354/http://www.mobipocket.com/dev/article.asp?BaseFolder=creatorpublisher&File=security.htm>.
51. —. Image support and display. *Internet Archive*. [Online] 14 March 2016. [Cited: 19 June 2017.] <https://web.archive.org/web/20160314211408/http://mobipocket.com/dev/article.asp?BaseFolder=prcgen&File=images.htm>.
52. Enhanced eBooks. *eBook Architects*. [Online] 28 April 2016. [Cited: 19 June 2017.] <http://ebookarchitects.com/learn-about-ebooks/enhanced-ebooks/>.
53. idleloop. Mobipocket notes file extractor. *Idleloop's site*. [Online] [Cited: 8 August 2017.] [http://www.angelfire.com/ego2/idleloop/mbp\\_reader.html](http://www.angelfire.com/ego2/idleloop/mbp_reader.html).
54. **Slangen, Simon**. How to remove DRM from MOBI and PRC eBooks. *MakeUseOf*. [Online] 16 June 2009. [Cited: 19 June 2017.] <http://www.makeuseof.com/tag/how-to-strip-mobi-and-prc-ebooks-of-encryption/>.
55. **Apprentice Alf**. Ebook formats, DRM and you -- a guide for the perplexed. *Apprentices Against DRM*. [Online] 13 January 2011. [Cited: 19 June 2017.] <https://apprenticealf.wordpress.com/2011/01/13/ebooks-formats-drm-and-you-%E2%80%94-a-guide-for-the-perplexed/>.
56. **Sorrel, Charlie**. How to strip DRM from Kindle E-Books and others. *Wired*. [Online] 17 January 2011. [Cited: 16 June 2017.] <https://www.wired.com/2011/01/how-to-strip-drm-from-kindle-e-books-and-others/>.
57. S.I. 2013 No. 777: The Legal Deposit Libraries (Non-Print Works) Regulations 2013. *Gov.uk*. [Online] 5 April 2013. [Cited: 21 June 2017.] <http://www.legislation.gov.uk/uksi/2013/777/contents/made>.
58. Supported eBook Formats. *Kindle Direct Publishing*. [Online] 2019. [Cited: 20 June 2019.] [https://kdp.amazon.com/en\\_US/help/topic/G200634390](https://kdp.amazon.com/en_US/help/topic/G200634390).

< - end - >