

[Sustainability of Digital Formats: Planning for Library of Congress Collections](#)

Search this site

Go

[Introduction](#) | [Sustainability Factors](#) | [Content Categories](#) | [Format Descriptions](#) | [Contact](#)
[Format Description Categories](#) >> [Browse Alphabetical List](#)

7z File Format

[>> Back](#)
Table of Contents

- [Identification and description](#)
- [Local use](#)
- [Sustainability factors](#)
- [Quality and functionality factors](#)
- [File type signifiers](#)
- [Notes](#)
- [Format specifications](#)
- [Useful references](#)

Format Description Properties

- ID: fdd000539
- Short name: 7z
- Content categories: aggregate
- Format Category: file-format
- Other facets: container-bundle, binary, structured
- Last significant FDD update: 2023-09-06
- Draft status: Full

Identification and description

Full name	7z File Format
Description	<p>7z is an archive file format that supports a variety of compression methods and encryption. 7z was initially the default format for 7-Zip, developed by Igor Pavlov in 1999 that was used to compress groups of files. There is no formal format specification for 7z but the LZMA compression specification serves as an "unofficial" specification with information on the file and byte structure. There are other specification options from third-party sources including but not limited to py7zr, a 7z python library that also documents the file's key characteristics and structure. According to 7-zip.org, the official homepage for the 7-Zip file archiving program, the 7z file format has the following main features:</p> <ul style="list-style-type: none">• An open architecture, which allows the support for a variety of compression methods. (See below)• High compression ratio (dependent on the compression method used)• AES-256 encryption• Large file size support, up to 1600 terabytes• Unicode file names• Archive header compression• Support for solid compression, where multiple files of a similar type are compressed within a single stream.

7z files are compatible with several compression methods because of its open architecture. More compression options may follow in time. The following are a list of compression methods, as stated by [7-zip.org](https://7zip.org) that are integrated with the 7z file format:

- [LZMA](#), an optimized version of the LZ77 compression algorithm. LZMA is the default compression method for 7z files.
- LZMA2, an improved version of the LZMA compression method
- PPMD
- BCJ, converter for 32 bit x86 executables
- BCJ2, converter for 32 bit x86 executables
- Bzip2, BWT algorithm
- Deflate, another LZ77 based compression algorithm

Per [7-zip.org's 7z history](https://7zip.org/7z-history), beginning in 2008, 7z supported UTF-8 encoding for filenames. 7z archives now have 3 distinct modes:

- If there are no required symbols, then the default mode uses UTF-8.
- mcu-switch: if there are non-ASCII symbols then 7-Zip uses UTF-8 encoding.
- mcl-switch: 7-Zip uses local code page.

The file structure for 7z files begins with a “signature header”. The header can contain information about the version number where version numbers consist of two bytes. A major version consists of the bytes "0x00" and a minor version has the bytes "0x04". As explained in the [LZMA 7z file format specification](#), the data block is placed after the signature header known as "Packed Streams." The signature header should consist of 32 bytes with the first six bytes of any given 7z file consisting of the same bytes as follows, {'7', 'z', 0xBC, 0xAF, 0x27, 0x1C}. "Packed Streams" is optional as the data block may be empty. From [LZMA's GitHub](#), 7z files use little endian encoding where the least significant bytes are stored first. The majority of the file structure is a sequence of "Property IDs" which dictate the type of information contained within the 7z file. The py7zr documentation lists all of the associated IDs for each type of property including:

- Header
- Fileinfo
- Substreams info
- Folder

The "Files Information" or "Fileinfo" Property ID can contain information about the files within the archive including: number of files, names, sizes, timestamps, and select attributes such as permissions.

The file structure also contains Codec IDs which correspond to specific compression methods (such as the ones listed above). Some of those Codec IDs are defined as follows:

- LZMA - 0x030101
- BCJ - 0x04
- LZMA2 - 0x21
- Bzip2 - 0x040202
- Deflate - 0x040108

Additional compression Codec IDs can be found in the py7zr documentation as well as [LZMA's GitHub](#), which provides information about many of the aforementioned compression methods. The 7z file format also supports [solid compression](#), where several files of a similar type are compressed within a single stream to avoid redundancies.

7z files utilize [AES \(Advanced Encryption Standard\)](#) encryption. As stated in its definitive [standard](#), the encryption uses a cypher that converts the data to an unintelligible form called ciphertext while the decryption converts the data back to its original format, or plaintext. The AES encryption method is able to use cryptographic keys of 128, 192, and 256 bits. The 7z format also supports file naming encryption.

Peazip (another compression program similar to 7zip, Winzip, among others) [GitHub](#) released a compression benchmark that compared a variety of compression software programs including Peazip, Winzip, WinRar, and 7-Zip. Most of the focus of this benchmark pinpoints the strengths and weaknesses of the particular compression programs, but it should be noted that the results summary highlighted that 7z files provided better compression but slower speeds than [RAR](#) files while decompression speeds were comparable to [ZIP](#) and [RAR](#) files making it a good candidate for content distribution.

For more information on the history of 7z and the 7-Zip application, see [Notes](#).

A note on terminology. In computing and in many standards specifications, 7z, ZIP, RAR, tar and the like are classified as archive files. This site is using the term aggregate instead of archive when defining quality and functional factors because the latter term archive has broader community use beyond the definitions of these formats. The term aggregate is used here instead to convey the basic function of bringing disparate parts together into a single collective object but also with the added features of compression, potential for encryption, error detection and more. See [Quality and Functionality Factors](#) for more information.

Production phase

May be used at any life-cycle phase for bundling files. When compressed with an external software program, maybe used at any life-cycle phase for packaging files for exchange and portability.

Local use

LC experience or existing holdings

The Library has over 2,000 7z files inventoried on long-term storage. 7z files have frequently been used as the delivery format for large collections of e-resource content.

LC preference

The Library of Congress Recommended Formats Statement (RFS) lists 7z as an Acceptable format for [Datasets](#).

Sustainability factors

Disclosure

The file format specification is partially documented in the LZMA's [GitHub](#), openly available since 2015. The specification is found in a plain text document in [GitHub](#). It should be noted that the GitHub space is described as the "Unofficial" Git mirror of LZMA releases. An additional unofficial specification for the 7z python library is also publicly available at [Py7zr](#).

Documentation

The 2015 file format specification is publicly available on [GitHub](#) and was jointly developed by Igor Pavlov and Jordan Justen. There is also a file format specification from a third party source from [py7zr](#), a 7z python library. Py7zr is written and maintained by Hiroshi Miura and also maintains its own [GitHub](#) as well.

Adoption

Below is a non-comprehensive list of applications that can open 7z files. This list can be found at [7-zip.org](#)

	<p>Windows:</p> <ul style="list-style-type: none"> • WinRAR • PowerArchiver • Peazip <p>MAC OS X:</p> <ul style="list-style-type: none"> • Unarchiver • Commander One • Keka <p>Linux:</p> <ul style="list-style-type: none"> • p7zip <p>7z files can be created using the 7-Zip graphic interface as well as via command line operations using both UNIX and LINUX systems. Examples of documentation for command line operations to create 7z files can be found at 7ziphelp, which provides tutorials on using different 7-Zip application functions.</p> <p>There is also considerable support from a variety of software programs to convert 7z files to other archive file formats. Lifewire's entry for 7z files highlights software programs such as Zamzar that can convert 7z files to other archive formats such as ZIP, RAR, ISO_image, and tar files.</p>
Licensing and patents	No licensing and patent restrictions for 7z files. 7-Zip uses the GNU LGPL, GNU Lesser Genera - Public License , which is a free-software license. License text within the context of 7-Zip are explained on 7-Zip's website .
Transparency	Transparency of compressed 7z files are dependent upon tools, software programs, and necessary algorithms used to read the file.
Self-documentation	<p>The 7z file format provides no additional metadata support beyond what is needed to support unpacking the archive and extracting the component items into a file system.</p> <p>Accessibility Features</p> <p>No specific features in the file format. Features to support accessibility would be found in the bundled and compressed files (such as embedded captions and subtitles in audiovisual content, tagged and structured text in textual documents, and alt text for images). Aggregate files can also contain separate files for transcripts, timed text or captions as part of the bundled package. See Relationships to other formats for details.</p>
External dependencies	7z files can be created using the 7-Zip graphic interface as well as via command line operations using both UNIX and LINUX systems. Examples of documentation for command line operations to create 7z files can be found at 7ziphelp , which provides tutorials on using different 7-Zip application functions. No external dependencies beyond available software to extract and decompress a compressed 7z files.
Technical protection considerations	<p>7z files can also utilize AES (Advanced Encryption Standard) encryption. The bit key is generated from a passphrase supplied by users that is based on the SHA-256 hash function. As stated in its definitive standard, the encryption uses a cypher that converts the data to an unintelligible form called ciphertext while the decryption converts the data back to its original format, or plaintext. The AES encryption method is able to use cryptographic keys of 128, 192, and 256 bits. The 7z format also supports file naming encryption.</p> <p>There are a variety of documentation sources such as the University of Victory University Systems provide information for users on how to use</p>

the 7-Zip program to encrypt 7z archive files.

Quality and functionality factors

Aggregate	
Compression	The 7z file format supports solid compression , where several files of a similar type are compressed within a single stream to avoid redundancies. Solid compression is also supported with RAR and tar archive file formats. The 7z file format uses a variety of compression methods including: LZMA, LZMA2, Bzip2, PPMd and DEFLATE
Support for Error Detection	7z files store CRC-32 checksums in the "Next Header CRC" a component of the "Signature Header" of the file. 7z files consist of data blocks known as packed streams where each packed stream has a CRC-32 checksum, which may be checked upon extraction. Uncompressed data in 7z files known as the "Coders Information" also contains CRC-32 checksums in the header section. 7-zip.org provides step-by-step documentation on how to recover corrupted 7z archive files.
Beyond normal functionality	The 7-Zip archive program allows the 7z file format to support multi-part archives such as 7z.0001, 7z.0002, etc. These multi-part files can also be reassembled using the "Combine Files" menu option in 7-Zip. Multi-part 7z files can also be unpacked in the command line using the 7z UNIX and LINUX systems. All segments of the multi-part archive must be in the same folder or server space to run command line operations to unpack the files.

File type signifiers and format identifiers

Tag	Value	Note
Filename extension	7z s7z	.s7z files are compressed 7-Zip files specifically from Mac OS. Read more at Reviversoft's s7z overview .
Internet Media Type	application/x-7z-compressed	There is no registration at IANA for an Internet Media Type for the 7z file format. Particularly of note is the presence of the "x" in the mediatype. Per RFC's Media Type Specifications and Registration Procedures , "names with "x" as the first facet may be used for types that are intended exclusively for use in private, local environments. These mediatypes cannot be registered and are strictly intended for use with the parties exchanging them." See RFC Section 3.4 for more information. See developer.mozilla.org or the Wikipedia entry for 7z .
Magic numbers	'7', 'z', 0xBC, 0xAF, 0x27, 0x1C	See Wikipedia entry for 7z .
Magic numbers	37 7A BC AF 27 1C	See GCK's File Signature Table at garykessler.net
Uniform Type Identifier (Mac OS)	org.7-zip.7zip-archive	See the Wikipedia entry for 7z .
Pronom PUID	fmt/484	For 7zip file format, see https://www.nationalarchives.gov.uk/PRONOM/fmt/484 . Outline record only .
Wikidata Title ID	Q270131	For 7z, family of archive file formats used by 7-Zip. See https://www.wikidata.org/wiki/Q270131 .
Wikidata Title ID	Q27492089	For 7z format description, technical specification of the 7z file format. See https://www.wikidata.org/wiki/Q27492089 .

General	
History	<p>7-Zip.org provides an extensive history of the archive program and file format on 7-Zip's site. The original beta version was first released on January 2, 1999. The 7z file format was designated the default format for the 7z.exe and 7za.exe applications. Further support for additional compression programs such as bzip2, gzip, tar, and zip was developed in late 2001. 7-Zip's history log chronicles ongoing bug fixes, software improvements, and additional functionality with the latest entry on December 26, 2021.</p>

Format specifications

- [LZMA compressed data and LZMA file format specification \(draft\)](#). (<https://github.com/jljusten/LZMA-SDK/blob/master/DOC/lzma-specification.txt>).
- [7z File Format Specification](#). (https://py7zr.readthedocs.io/en/latest/archive_format.html#).

Useful references

URLs

- Third Party File Format Specifications.
 - [Py7zr – 7z library on python 7z file format specification](#). (<https://py7zr.readthedocs.io/en/latest/index.html>). Unofficial file format specification.
- Wikipedia References and Sources
 - [Wikipedia entry for 7z](#). (<https://en.wikipedia.org/wiki/7z>).
 - [Wikipedia entry for List of archive formats](#). (https://en.wikipedia.org/wiki/List_of_archive_formats).
 - [Wikipedia entry for 7-Zip](#). (<https://en.wikipedia.org/wiki/7-Zip>).
 - [Wikipedia for LMZA](#). (https://en.wikipedia.org/wiki/Lempel%E2%80%93Ziv%E2%80%93Markov_chain_algorithm).
 - [Wikipedia entry for bzip2](#). (<https://en.wikipedia.org/wiki/Bzip2>).
 - [Wikipedia entry for solid compression](#). (https://en.wikipedia.org/wiki/Solid_compression).
 - [Wikipedia entry for partial matching](#). (https://en.wikipedia.org/wiki/Prediction_by_partial_matching).
- General References and Sources
 - [7z file format](#). (<https://docs.fileformat.com/compression/7z/>).
 - [7z File – Opening and Understanding](#). (<https://file.org/extension/7z>).
 - [7Z File Extension](#). (<https://fileinfo.com/extension/7z>).
 - [7-Zip. License for use and distribution](#). (<https://www.7-zip.org/license.txt>).
 - [7-Zip. How to recover corrupted 7z archive](#). (<https://www.7-zip.org/recover.html>).
 - [Compression benchmark: 7-Zip, PeaZip, WinRar, Winzip comparison](#). (<https://peazip.github.io/peazip-compression-benchmark.html>).
 - [s7z File Extension](#). (<https://www.reviversoft.com/en/file-extensions/s7z>).
 - [Announcing the ADVANCED ENCRYPTION STANDARD \(AES\)](#). (<https://nvlpubs.nist.gov/nistpubs/FIPS/NIST.FIPS.197.pdf>).
 - [Advanced Encryption Standard \(AES\)](#). (<https://www.nist.gov/publications/advanced-encryption-standard-aes>).
 - [7-Zip Command Line Examples — All Syntaxes Shared Here](#). (<https://7ziphelp.com/7zip-command-line>).
 - [Command-line Guide for Linux, Mac & Windows](#). (<https://info.nrao.edu/computing/guide/file-access-and-archiving/7zip/7z-7za-command-line-guide>).
- PRONOM and Wikidata IDs
 - [PRONOM entry for fmt/484](#) (<https://www.nationalarchives.gov.uk/PRONOM/fmt/484>). Information in PRONOM from UK National Archives about 7z: fmt/484. Outline record only.
 - [Wikidata entry for Q270131](#) (<https://www.wikidata.org/wiki/Q270131>). Information in Wikidata entry for 7z file format family. Wikidata Title ID: Q270131.
 - [Wikidata entry for Q27492089](#) (<https://www.wikidata.org/wiki/Q27492089>). Information in Wikidata entry for 7Z format description. Wikidata Title ID: Q27492089.

[Digital Preservation Home](#) | [Digital Formats Home](#)