

NAME

`make_bagit` - Create a BagIt archive of a folder, according to RFC 8493

SYNOPSIS

```
make_bagit -b input_folder
make_bagit -i input_folder -o output_folder
make_bagit (-h|-x)
```

DESCRIPTION

Bash AVpres is a collection of Bash script for audiovisual preservation. One of these small programs is **make_bagit**. It creates an archive of a folder, according to the "BagIt File Packaging Format", as designed by the Library of Congress in the U.S.A. and standardised under RFC 8493.

OPTIONS**BASIC OPTIONS**

-b *input_folder*, **--bagit=***input_folder*
input folder

This replaces the *input_folder* by its BagIt archive, containing the original folder, and is faster.

-i *input_folder*, **--input=***input_folder*
input folder

-o *output_folder*, **--output=***output_folder*
output BagIt folder

This copies the content of the *input_folder* into the *output_folder* while creating a BagIt archive and requires more time.

ADVANCED OPTIONS

The arguments of the advanced options can be overwritten by the user. Note that any string containing spaces must be quoted.

--algorithm=(md5|sha1|sha256|sha512)

RFC 8493 considers MD5, SHA-1, SHA-256 and SHA-512, which is the default:

```
--algorithm=' sha512'
```

make_bagit and **verify_bagit** support all four hash algorithms, but the use of different and/or multiple algorithms in the same BagIt archive is currently not allowed.

The script needs an external command to compute recursively the hash checksums of all elements inside the folder.

--md5=*md5_tool*

The default for MD5 is:

```
--md5=' /usr/local/bin/md5deep -r'
```

--sha1=*sha1_tool*

The default for SHA-1 is:

```
--sha1=' /usr/local/bin/sha1deep -r'
```

--sha256=*sha256_tool*

The default for SHA-256 is:

```
--sha256=' /usr/local/bin/sha256deep -r'
```

--sha512=*sha512_tool*

The default for SHA-512 is more complex, because there is no recursive function to call:

```
--sha512=' -find . -type f -exec sha512sum "{}" +'
```

--cp=copy_tool

The script needs a copy command when the **-i** and **-o** options are used.

By default, **gcp** (the GNU flavour of **cp**) is used rather than the classic **cp**, because it offers more possibilities, such as preserving the files' original date and time metadata. The default configuration is:

```
--cp='/usr/local/bin/gcp --preserve=mode,timestamps -r'
```

--bagit_version=bagit_version

RFC 8493 defines BagIt version 1.0 and **make_bagit** uses this by default, but, for compatibility reasons, it supports also older version tags (see Notes):

```
--bagit_version='1.0'
```

INFORMATIVE OPTIONS**-h, --help**

display a help message

-x, --options

display the advanced options with their default arguments

NOTES

The structure of a BagIt archive is:

```
bagit.txt
data/<archive_content>
manifest-<algorithm>.txt
[bag-info.txt]
[tagmanifest-<algorithm>.txt]
```

Only the 'bagit.txt' and 'manifest-<algorithm>.txt' files are mandatory, in addition to the 'data' folder, which contains the actual content to archive.

RFC 8493 defines BagIt version 1.0 and **make_bagit** uses this by default, but, for compatibility reasons, the values from 0.93 to 0.97 are also considered valid and supported.

According to RFC 8493, the 'bag-info.txt' and 'tagmanifest-<algorithm>.txt' files are optional and **make_bagit** creates these always. The 'bag-info.txt' file contains metadata elements describing the bag and its payload and is intended primarily for human use. Yet the 'Payload-Oxum' metadata is used by **verify_bagit** to quickly detect incomplete bags, before performing any checksum validation. The 'tagmanifest-<algorithm>.txt' file contains the checksum manifest of all the other files generated.

The optional 'fetch.txt' file is currently not supported neither by **make_bagit** nor **verify_bagit**.

In addition, for sustainability and transparency reasons, **make_bagit** embeds into the BagIt folder its own source code as well as **verify_bagit**'s one: 'make_bagit.txt' is the script used to actually generate the BagIt archive and 'verify_bagit.txt' can be used to check its integrity.

Note that the '.DS_Store' files generated by macOS are deleted and the spaces in folder and file names are replaced by underscore signs before the BagIt archive is created.

Temporary log files are stored at:

```
/tmp/AVpres/make_bagit.XXXXXXXXXX
```

The log files can be used for debugging, for example by running **cat** on the address prompted with fatal error messages:

```
cat /tmp/AVpres/make_bagit.XXXXXXXXXX
```

SEE ALSO

RFC 8493, "The BagIt File Packaging Format (V1.0)", October 2018:
<https://www.rfc-editor.org/info/rfc8493>

RFC 1321, "The MD5 Message-Digest Algorithm", April 1992:
<https://www.rfc-editor.org/info/rfc1321>

RFC 3174, "US Secure Hash Algorithm (SHA1)", September 2001:
<https://www.rfc-editor.org/info/rfc3174>

"Descriptions of SHA-256, SHA-384, and SHA-512":
<https://web.archive.org/web/20130526224224/http://csrc.nist.gov/groups/STM/cavp/documents/shs/sha256-384-512.pdf>

verify_bagit(1).

md5deep(1), **sha1deep**(1), **sha256deep**(1) and **sha512sum**(1).

COPYRIGHT

Copyright (c) 2014-2020 by Reto Kromer

LICENSE

The **make_bagit** Bash script is released under a 3-Clause BSD License.

DISCLAIMER

The **make_bagit** Bash script is provided "as is" without warranty or support of any kind.