



## Preserving Digital Records

Guidance for Donors and Depositors

Using Bagger

---

### Introduction

This guidance is for anyone considering donating or depositing digital records with Gloucestershire Archives. It aims to introduce a piece of software called *Bagger* (see glossary) that will package up your digital records to ensure safe transfer to the archives.

Gloucestershire Archives collects, preserves and provides access to records which are of historical significance and record everyday life for the counties of Gloucestershire and South Gloucestershire. Digital records have the same potential to become valuable historical resources as paper-based records. Many of the new collections received by Gloucestershire Archives include both digital and paper records.

---

### Bagger – a transfer tool

#### What is it?

To ensure the safe transfer of digital material from the depositor to Gloucestershire Archives we prefer that you use software called Bagger. The Library of Congress developed this tool specifically to support the reliable transfer of digital files. Bagger collects your files, packs them in a folder called a bag, (see Glossary), allows you to add delivery information and includes a list of the contents of the package – just as you would if you were transferring paper files to us.

Sometimes the delivery process can damage a physical package, for example, it might get wet in the rain. The delivery process can also damage an electronic package. Bagger, using a specification called Bag-it (see Glossary), allows you to check the package is complete and whole (or ‘valid’) before it leaves you and allows us to check it is complete and whole, ‘valid’ when it arrives with us. As Bagger is also a component of our in-house digital preservation tool we are easily able to process a package once it arrives here.

#### What does a package look like?

The bag is a single folder or directory – it acts like the container in which you might pack your paper files to transfer them to us. The name of this folder is the name of the bag, in this example it is *Gloucestershire Archives bag*.

Within the file there is one folder called data and 4 additional text files; bagit.txt, manifest-sha1.txt, bag-info.txt and tagmanifest-sha1.txt.

# Gloucestershire Archives



The screenshot shows a Windows File Explorer window with the following details:

- Title Bar:** Gloucestershire Archives bag
- Menu Bar:** File, Edit, View, Favorites, Tools, Help
- Toolbar:** Back, Forward, Search, Folders, a folder icon with a blue oval around it, and a refresh icon.
- Address Bar:** Address: C:\Documents and Settings\cecollin\Desktop\Gloucestershire Archives bag
- File and Folder Tasks:** Make a new folder, Publish this folder to the Web
- Other Places:** Desktop, My Documents, User on 100852L, My Network Places
- Details:** Gloucestershire Archives bag, File Folder, Date Modified: 02 July 2013, 13:43
- Content Area:** A table showing the contents of the 'data' folder:

Name	Size	Type	Date Modified
data		File Folder	02/07/2013 13:43
bag-info.txt	1 KB	Text Document	02/07/2013 10:01
bagit.txt	1 KB	Text Document	02/07/2013 10:01
manifest-sha1.txt	1 KB	Text Document	02/07/2013 10:01
tagmanifest-sha1.txt	1 KB	Text Document	02/07/2013 10:01

The data folder contains the content itself. This holds your files, which can be single files, whole directories or multiple levels of sub-directories as well as any type of file such as documents, pictures, sound or movie files. This is also known as the 'payload'.

The file called bagit.txt is the bag declaration, which contains details of the version of Bag-It used to create the bag.

The manifest-sha1.txt file lists the content in the data folder and lists each file's checksum (see Glossary). Bagger calculates this automatically.

The bag-info.txt contains the data recorded in the *Bag-Info* tab (number of files, date of creation and size).

The tagmanifest-sha1.txt lists the previous three text files and their checksums to show that these files have not been changed or damaged.

# Gloucestershire Archives



## Where is it?

To use Bagger, you will need to install it on your computer. You can download the software from <https://sourceforge.net/projects/loc-xferutils/files/loc-bagger/>

The document presumes that users are operating in a Windows environment. If this is not the case, please get in touch with us at Gloucestershire Archives for further help.

## Resolve Java Error

Bagger requires Java Runtime Environment (JRE) be installed and accessible by Bagger. JRE is widely used, and most computers will already have a copy. If the computer you are working with (a) does not have a copy or (b) the computer cannot correctly locate its copy, you will receive an error when you attempt to open Bagger.

**If you do not receive an error, you may skip this section and proceed directly to 'How do I use it?' However, if you do receive a Java error, you will need to follow these steps with administrator rights to your computer. If you do not have administrator rights, your computer will refuse to allow you to make the following changes. If you are not sure if you have administrator rights, try the following steps and stop if you receive a message indicating that you need administrator rights to proceed.**

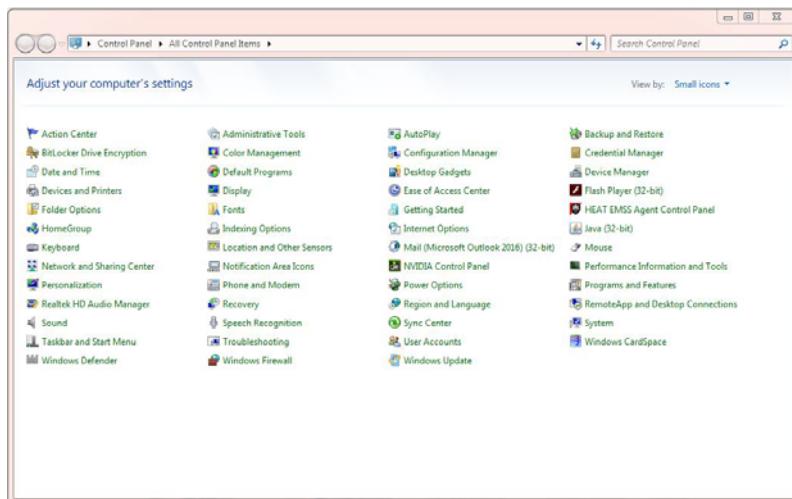
### Confirm that JRE is installed

You can check your computer to see if you have a JRE installed by navigating to the Control Panel.

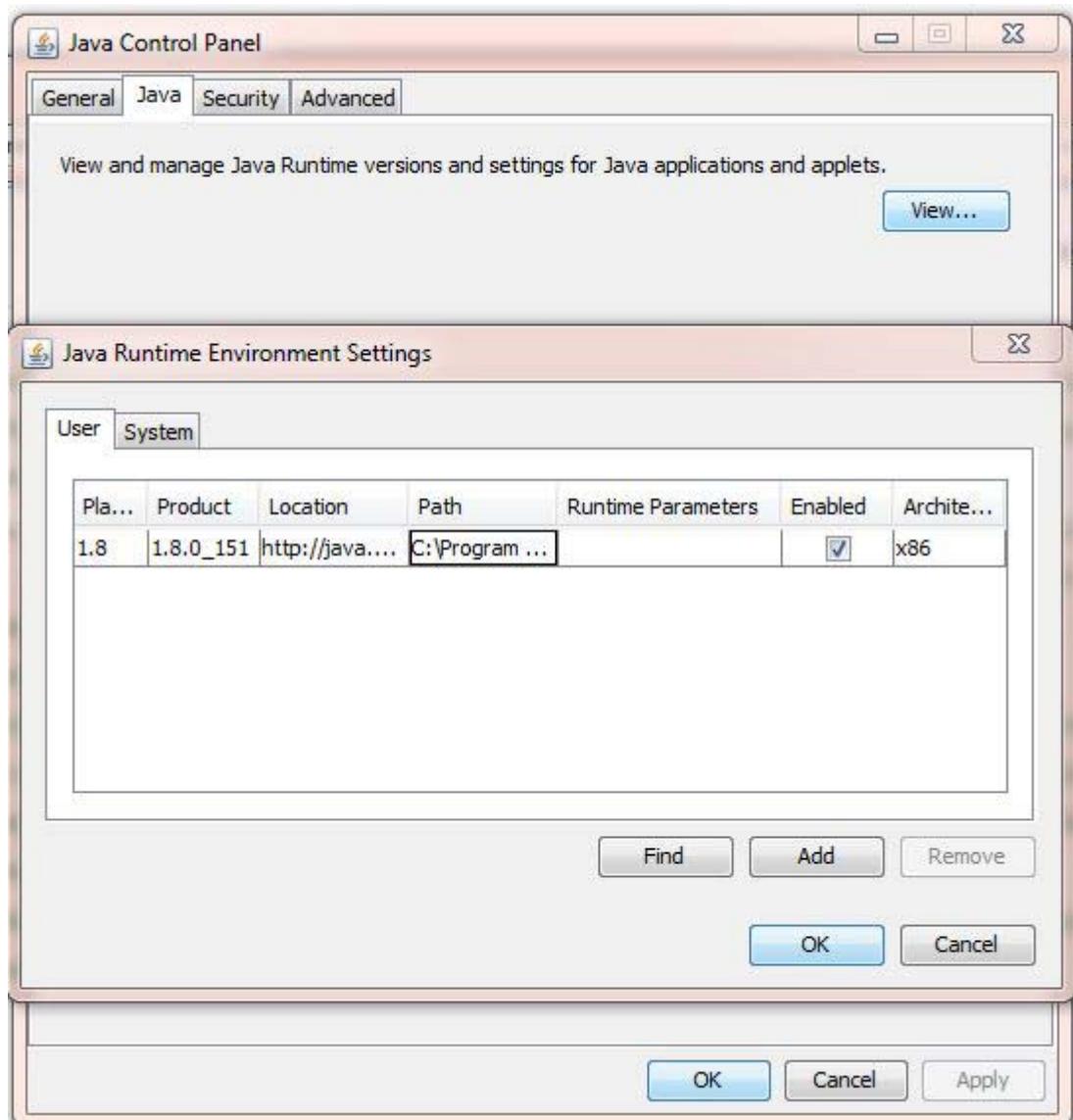
For Windows 7, select **Start -> Control Panel**

For Windows 10, search for "Control Panel" in the search bar

If your Control Panel is in Category View (see below), click on **Category** and change it to **Small Icons**. Double-click on **Java** to open the Java Control Panel



Select the **Java** tab to view the screen that provides access to the Java Runtime Environment Settings. Select the **View...** button, and this will display the Java Runtime Versions you have installed.



This displays the version and the installation location.

If it appears that you have a JRE installed, skip down to step 1.2 to confirm that the JRE environmental variable is defined.

#### **Download and Install JRE if it is not already installed.**

1. Oracle: <http://www.oracle.com/technetwork/java/javase/downloads/index.html>

2. CNET: [http://download.cnet.com/Java-Runtime-Environment-JRE/3000-2356\\_4-10009607.html](http://download.cnet.com/Java-Runtime-Environment-JRE/3000-2356_4-10009607.html)

Note the JRE is generally installed to the C:\Program Files or C:\Program Files (x86) folder.

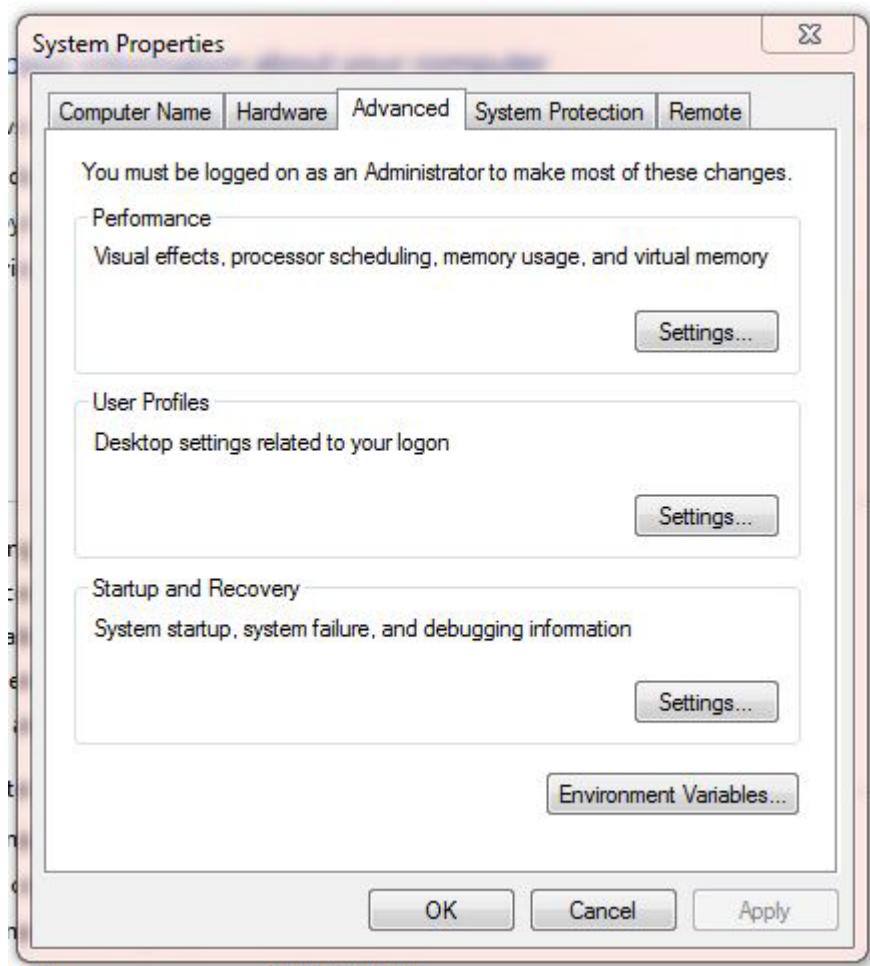
#### **Configure JRE environmental variable**

Navigate to the **Advanced system settings**.

- For Windows 7: From the **Start** menu, right click on the **Computer** button. From right-click menu, select the **Properties** option. Select the **Advanced system settings** link on the left.
- For Windows 10: navigate to the Control Panel by searching “control panel” in the search bar. Click on System, then Advanced system settings.

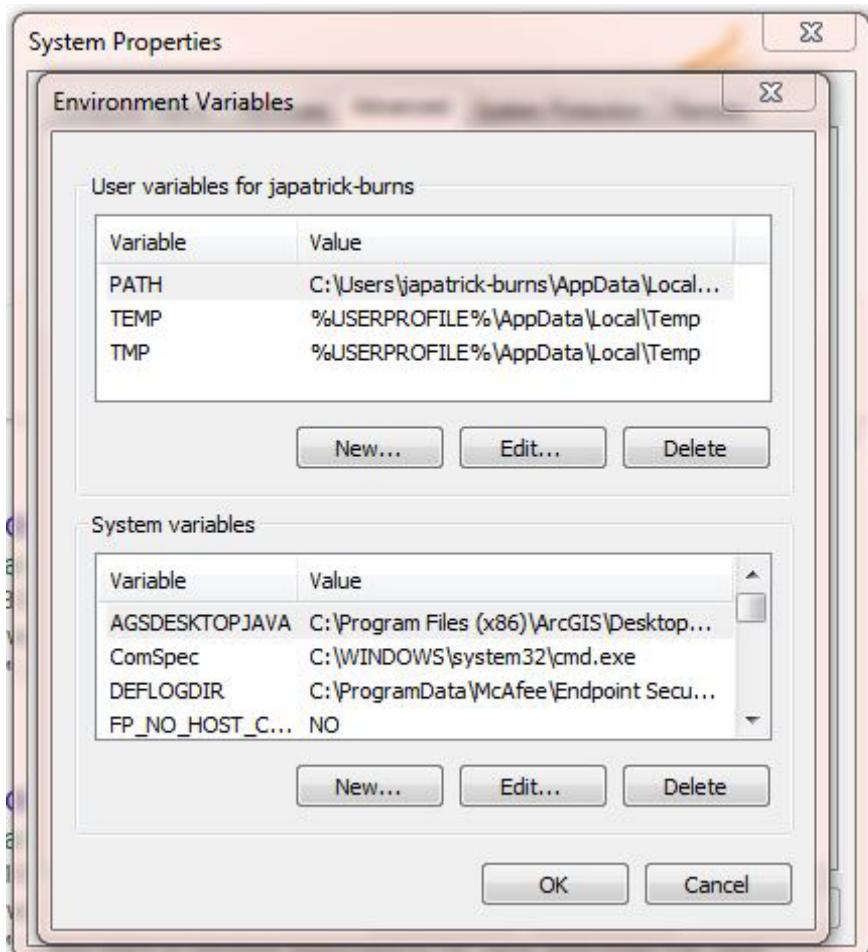
Select the **Environment Variables** button near the bottom of the dialog box.

# Gloucestershire Archives



An Environment Variables window will appear. First verify that the JAVA\_HOME environment variable is **not** defined.

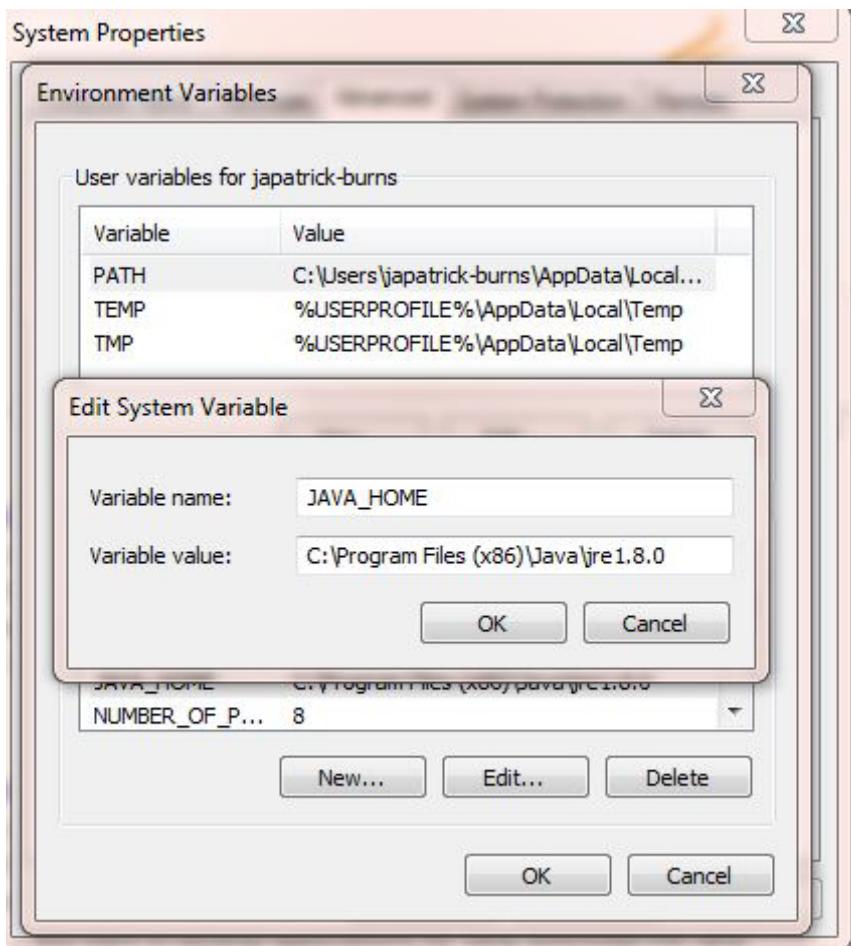
The environment variables are listed alphabetically. Scroll down the list, until you see where the environment variables would be listed that start with J. In the example below, there is no JAVA\_HOME variable listed in this display, so it is not yet defined. We will create it in the following steps.



To create a new Environmental variable, select the **New** button in the lower half of the window.  
A new **Edit System Variable** window will appear.

In the *Variable name* field, type JAVA\_HOME. In the *Variable value* field, type the complete path to the folder that contains the JAVA bin folder, the lib folder, and the COPYRIGHT file. This will be **above** the bin folder that contains java.exe (Note: this will **not** be the bin folder, but the folder up one level from the bin folder).

# Gloucestershire Archives



Click **OK** to close the **Edit System Variable** window.

Review the list of system variables and verify the JAVA\_HOME variable is there.

The system variables are listed alphabetically in the lower window. Scroll down through your list of system variables to confirm the JAVA\_HOME variable was properly assigned.

Click **OK** to close the **Environment Variables** window.

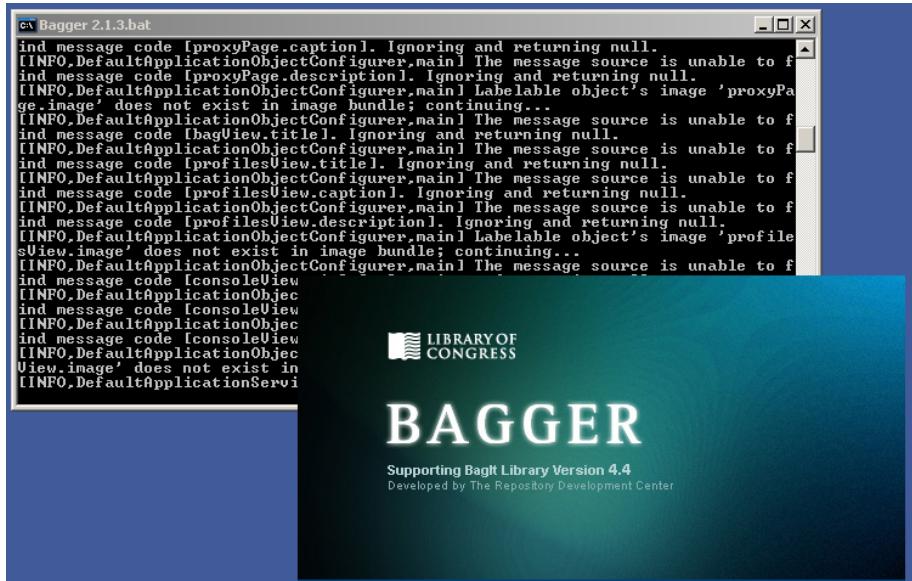
Click **OK** to close the **System Properties** window.

# Gloucestershire Archives



## How do I use it?

To start Bagger you need to double click the file bagger.bat. You will see a terminal, and then the Bagger start up screen.

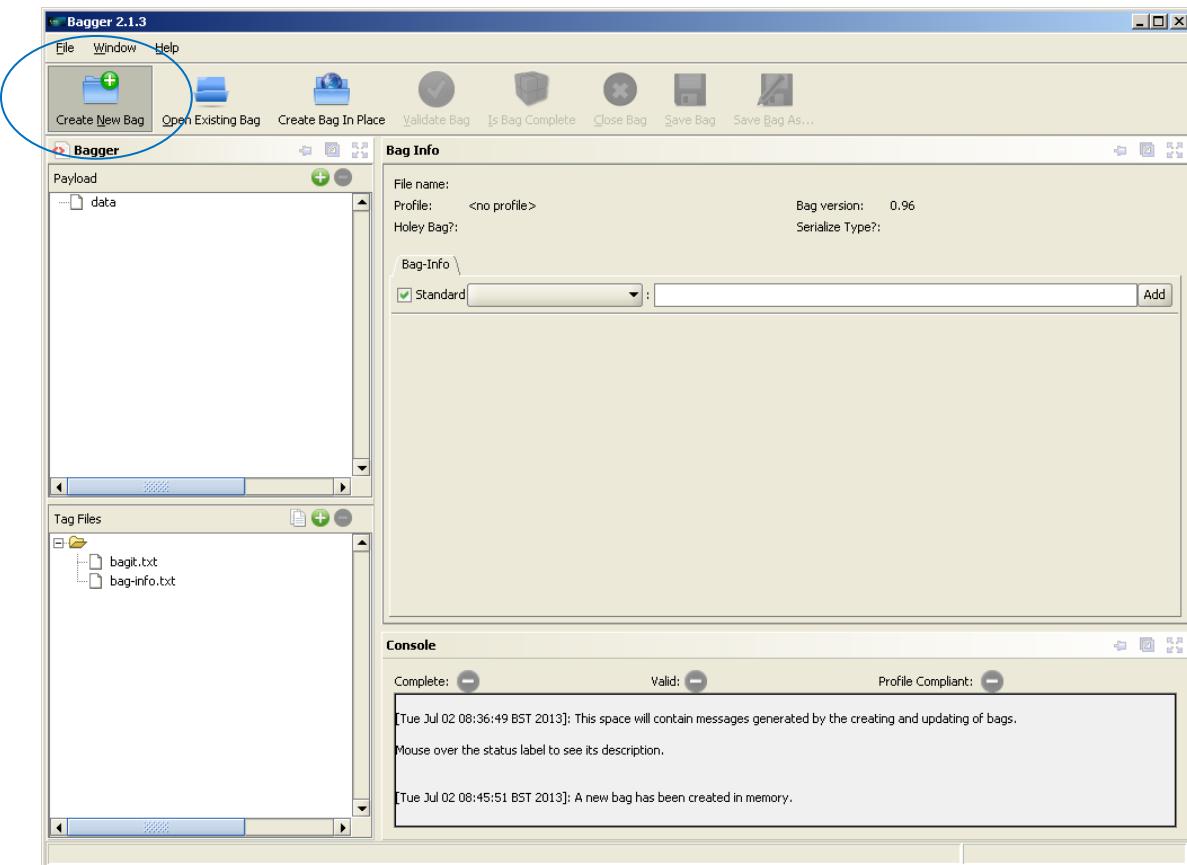


# Gloucestershire Archives

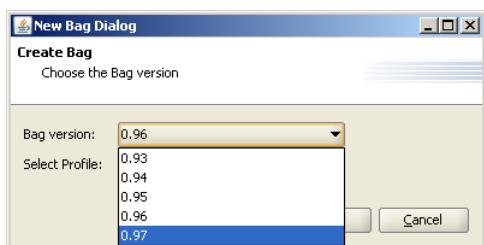


## To create a bag?

Click on the *Create New Bag* button.



You will see the *New Bag* dialog box.

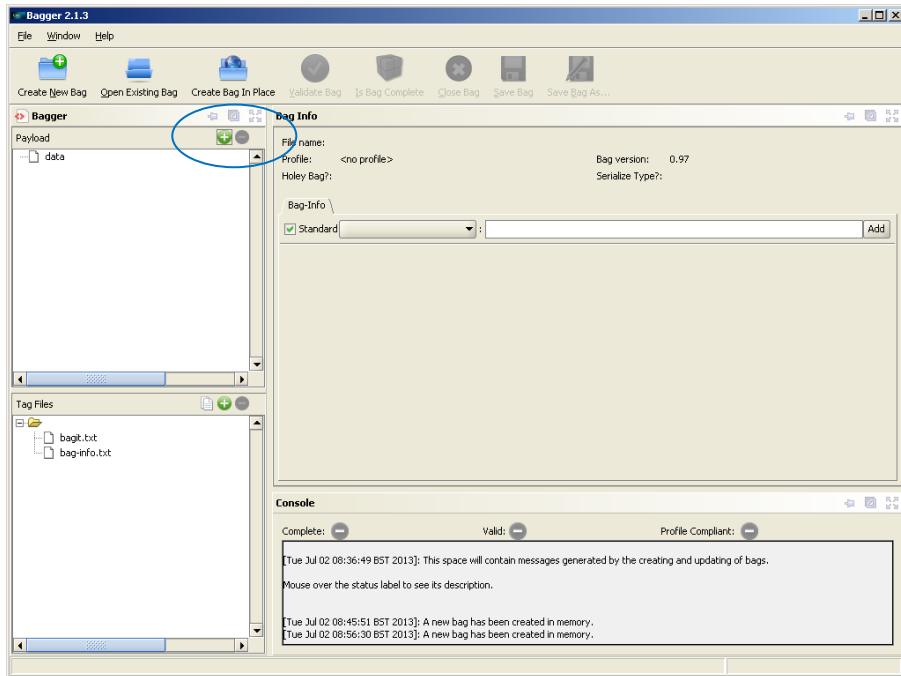


If *Bag version* doesn't default to 0.97, select it.

Select the <no profile> option for the *Select Profile* field. Click *ok*.

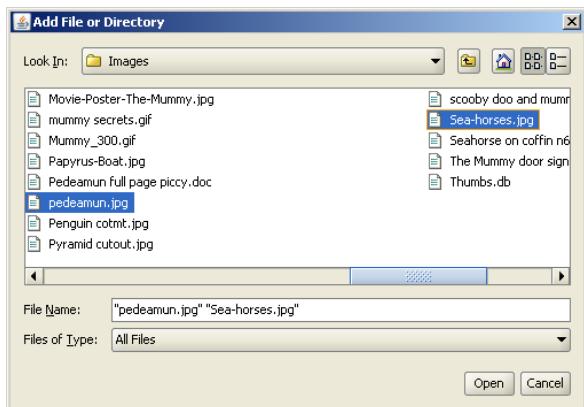
In the 'Payload' area of the tool, click the green button with a cross.

# Gloucestershire Archives



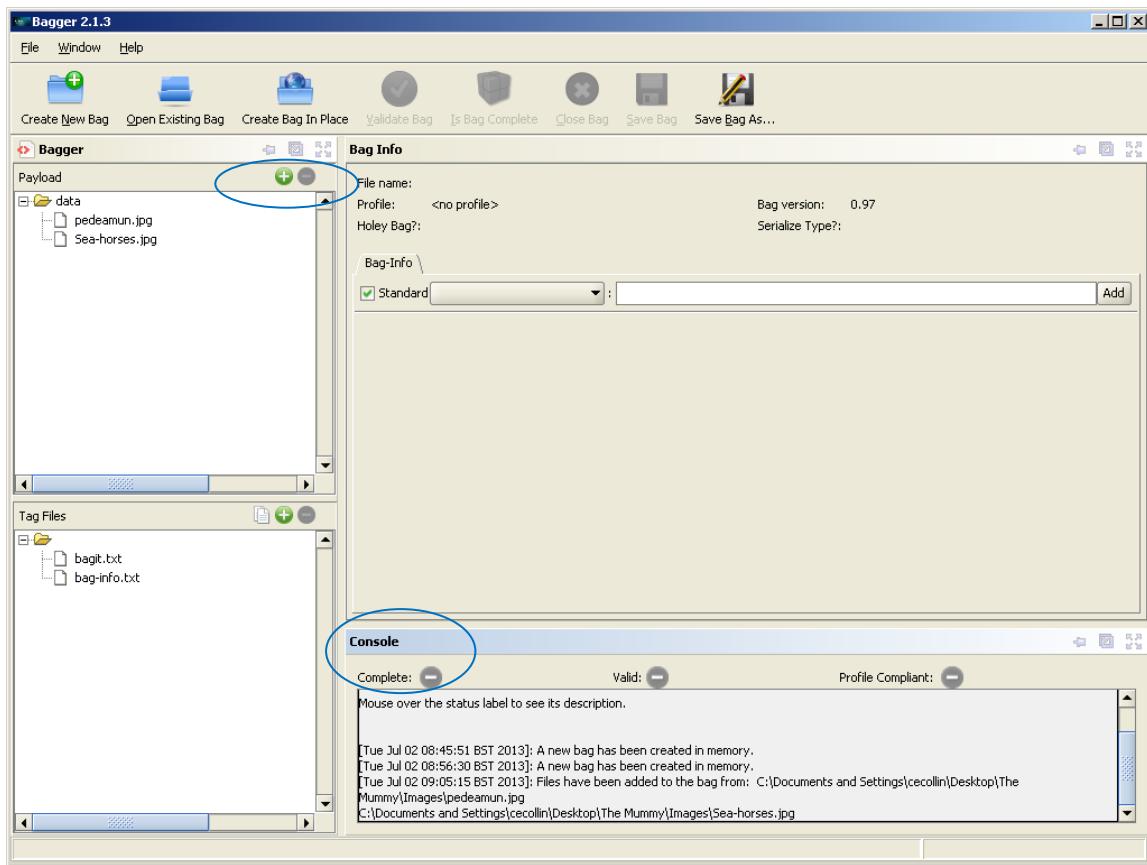
This allows you to browse your computer and select the files or folders that you want to add to your bag.

You can select more than one item at a time by using the *Ctrl* key.



Once you have selected your files or folders, click open. You can add any type of file or folder into the bag.

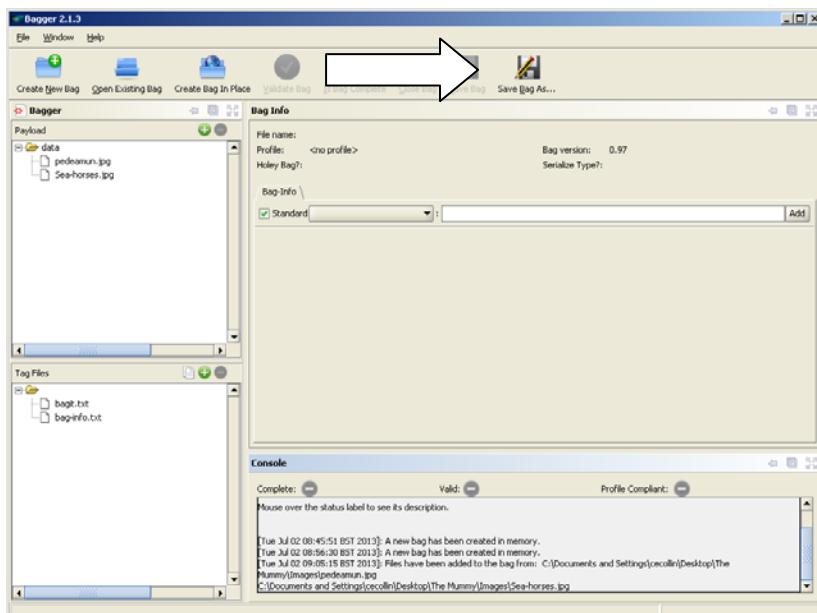
# Gloucestershire Archives



You can use the green and red buttons to add or subtract files. If you look at the console area of the tool you will see a list of actions that have been carried out.

Once you are happy that all of your files are in the payload, you can save your bag.

Click Save Bag As



# Gloucestershire Archives



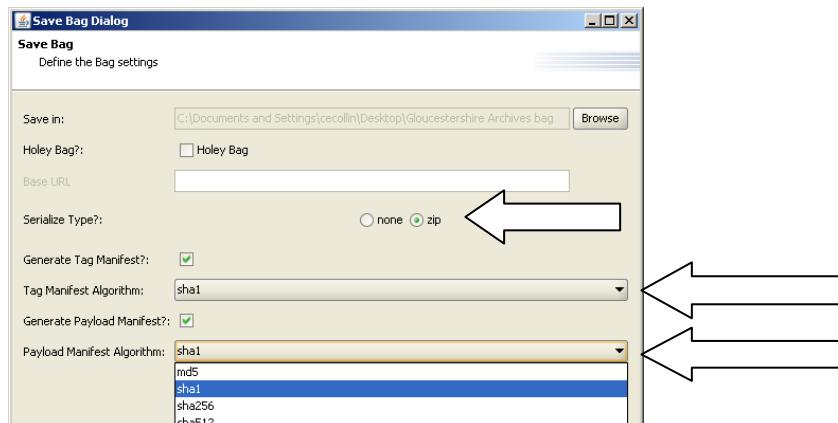
In the *Save In* field browse your filing system to find a place to store the bag. You can keep it on your machine, or you may want to put it on removable media, such as a USB stick.



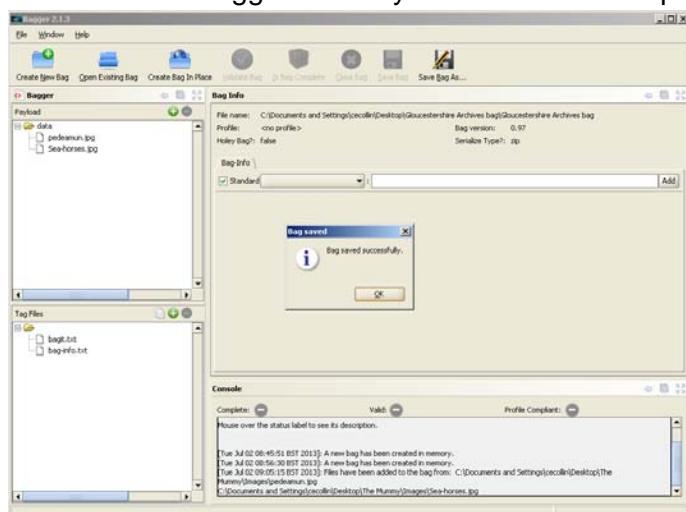
Give your bag a name – this can be anything that you like. Leave the *Files of Type* field blank.

Once you have named your bag you need to set the bag properties. Make sure that you select the zip option.

Make sure that the *Tag Manifest Algorithm* and the *Payload Manifest Algorithm* are set to sha-1.



Then click ok. Bagger will tell you when it has completed the save.



# Gloucestershire Archives



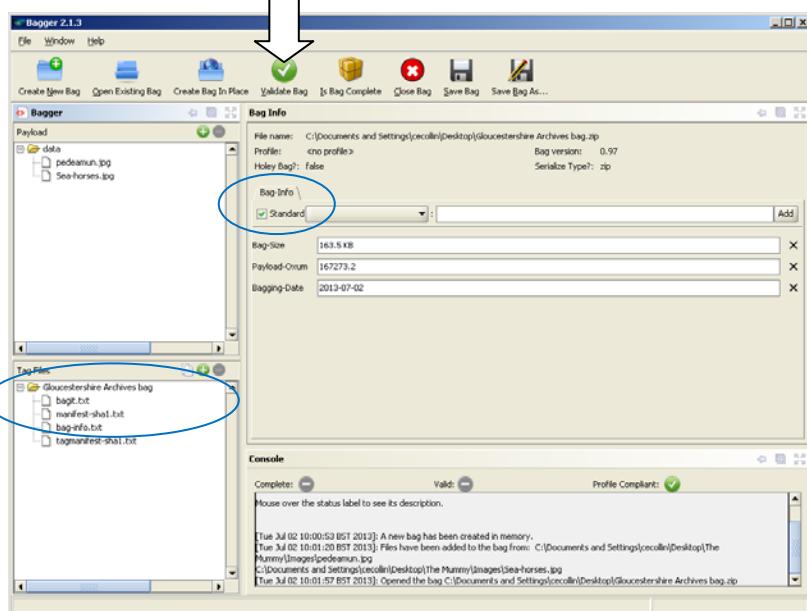
If your bag is very large, the saving process may take some time. You should see an egg timer for the duration of the process.

Once your bag is successfully saved, it will appear in the tag files part of the tool.

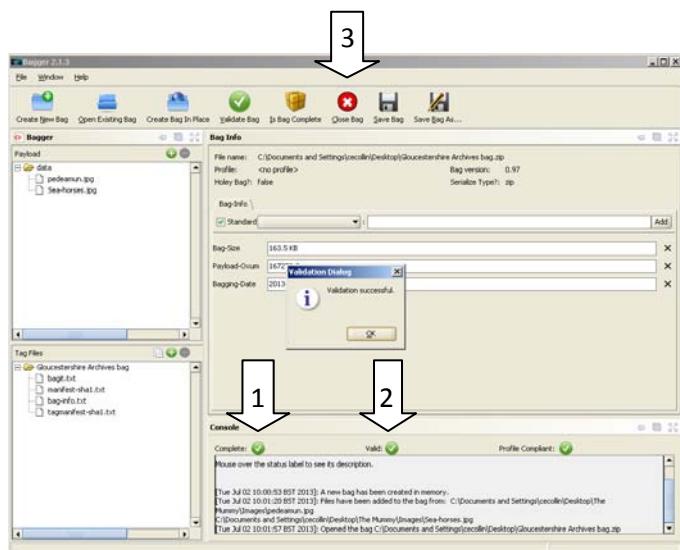
You will also see *Bag-Info*, which gives the size of the bag, the number of files and the date it was created etc.

The final action to carry out before transferring it to Gloucestershire Archives is to *validate* the bag.

Click the big green tick at the top of the screen.



Bagger checks that the bag is *complete* (arrow 1 below) and *valid* (arrow 2 below). Now you can close the bag (arrow 3 below). You don't need to save it as you already have done.





Now you can transfer this package to Gloucestershire Archives. You can email it to us, as long as it is smaller than 10MB or you can use a CD or DVD or a USB stick or external hard drive.

There is lots of information about Bagger. You may find the following links helpful,

The Library of Congress gives a very short introduction to the tool here

<http://www.digitalpreservation.gov//multimedia/videos/bagit0609.html>)

The Government Records Branch of North Carolina give some excellent instructions for downloading Bagger (as above).

[http://project.wdl.org/arab\\_peninsula/workshop2012/en/doha\\_workshop\\_2012\\_bagger\\_en.pdf](http://project.wdl.org/arab_peninsula/workshop2012/en/doha_workshop_2012_bagger_en.pdf)

The State Archives of North Carolina have produced a video to help with installing Bagger, which you may find useful. <http://shelf3d.com/NkhqdetV99M> .

## Glossary of Key Terms

**Bag:** a package that combines original files and checksums which can be validated and transferred safely.

**Bag-It:** specification developed by Library of Congress to package files for transfer.

**Bagger:** GUI or graphical user interface to enable easy creation of a bag

**Checksum or hash:** digital fingerprint of a file. Unique string of numbers and letters created by an algorithm running through every byte of your file. If a single byte in your file changes the checksum string will also change completely.

**Digital record:** 'A record created, and/or maintained, by means of digital computer technology. It includes records that are born-digital or have undergone conversion from a non-digital format.'<sup>1</sup> The term is often used interchangeably with 'electronic record'.

**Folder:** A computer folder is a graphical representation to support user management of computer files.

**File:** A computer file is a piece of digital information, which is recognised by a computer and presented and interpreted as one unit. It can usefully be regarded as the digital version of a paper document.

---

<sup>1</sup> SAA Glossary: <http://www.archivists.org/glossary/index.asp> <accessed 23/07/2011>.