

# Automatic Archives Download Script Administrator Guide

## Article Number

000075453

## Document Summary

In order to comply with data retention policies and reduce utilized space, administrators often want to move course archives to another location for longer-term storage. In SaaS, course archives are created automatically, and daily back-ups of course archives are stored in the cloud.

We've developed scripts written in Python to allow admins to copy archives locally or to an AWS S3 bucket prior to deleting them from Learn.

## Document Type

Administrator Documentation

## Information

# How to Download the Script

The latest version of the script is included with your Blackboard Learn application. To download the script go to **System Admin Course Settings Automatic Archive Settings**. The link "Download the Automatic Archive Bulk Downloader script" will always download the latest version.

## DOWNLOAD AUTOMATIC ARCHIVES

*Automatic course archives can be downloaded via the Content Collection or by using the bulk download script. See*

### Download specific automatic archives

To see all available automatic course archives, go to the `/internal/autoArchive` directory in your Content Collection and use the YYYYMMDD format (e.g. SampleCourseId01\_20210125). Select the date on which the archive you want view.  
[See available automatic course archives](#)

### Download automatic archives in bulk

*In order to run the bulk download script, you will need Python 3. If you don't have it already, download and install*

Select the link below to download the Automatic Archive Downloader script. Make a note of where the download is saved below.

[Download the Automatic Archive Bulk Downloader script](#)

### Install pip

Python installers for Windows include pip by default. If using Linux or MacOS you will likely need to install pip if you

# How to Use the Script

The utility script is operated from the command line by calling the Python runtime executable followed by the script and several parameters. In some modes a feed file list is used.

You need to be in the directory into which you have saved the script. And you should've followed the instructions in the *Related Links and Resources* section to install the required packages and the virtual environment and pre-configure Learn to allow WebDAV to be used.

## Invocation

The basic invocation follows. The characters "<", ">", "[" and "]" are CLI meta-syntax and should not be typed. See the synopsis for details.

```
python3 auto_archive_download.py [-h] -s <SITE> -u <USER> [ -f <FILE> | -b <YYYY>/<MM>/<DD>
```

Synopsis of Arguments

Option type	Short Form	Long Form	Description
<b>Authorization</b>	-s <i>site</i>	--site <i>site</i>	Front-end hostname of the site with protocol. EG: https://example.blackboard.com
	-u <i>user</i>	--user <i>user</i>	User ID of the user to authorize with. The user must be a System Administrator.  <i>Only "challenge/response" authentication such as LDAP and Default authentication are possible, CAS and SAML are not. The IETF RFC that governs WebDAV does not make any provision for "portal" based authentication.</i>  <i>If you have the default provider disabled you may need to create or use an alternative hostname. Contact Support for assistance.</i>
<b>Feed-File Mode</b>	-f <i>file</i>	--file <i>file</i>	A feed-file containing a list of course ID.  <i>This argument is required if the date mode is not used.</i>
<b>Date Mode</b>	-b <i>date</i>	--begin <i>date</i>	Date in big-endian format YYYY/MM/DD, for example 2021/01/01
	-e <i>date</i>	--end <i>date</i>	<i>When date mode is used all courses created from the specified begin-date to the specified end-date date are downloaded. Both options are mandatory if date mode is used. Date mode uses the creation date of the course, not the creation date of the archive files.</i>  <i>If this is used in conjunction with the feed-file mode, then courses must meet both criteria.</i>
			The output argument supports local paths and S3 paths. S3 and Local Mode options  Input any path for example the name of a folder. This can be an absolute path or a relative one. Follow the shell quoting rules of

<b>Output</b>	-o <i>output</i>	# <i>output</i>	<i>output</i>	your operating system
				<b>Local</b> <b>Examples:</b> d:\archive /mnt/ad0slc/archive archive  <i>The folder should already exist.</i>
				<b>S3</b> Input a standard S3 bucket path like <code>s3://mybucket/mypath/</code>  In S3 mode the <a href="#">standard AWS environment variable</a> are read automatically or a profile can be used.
<b>Other</b>	-h			Print a help message
	-d			Print debug messages while operating

## Examples

When you run the script you'll need to change the directory to the path where the script is located. The below examples assume that you follow the setup instructions at the end of this document and that this path is the same path where the virtual environment folder is located and the virtual environment is named "**env**." If you configured it differently you'll need to update the paths accordingly.

### Downloading a Range of Course Archives by Date

The script has two basic modes of operation. They can be used together or apart. The first is **Date Mode**: you specify a begin and end date and all the archives are then downloaded automatically for courses created between those dates. This can be useful for archiving terms.

This example is for Windows. The command interpreter prompt is shown in green light type, this is to indicate when the virtual environment becomes active. Don't type the prompt. Only type the portions in bold black type. In this example, we'll download courses created between January 1, 2021 and March 1, 2021

```
c:\> cd %userprofile%
c:\users\jsmith> .\env\scripts\activate
(env) c:\users\jsmith> py auto_archive_downloader.py -site example.blackboard.com -u admini:
```

The script will prompt for a password and begin downloading. Here is what each line does:

1. Change to the directory that the script, and the virtual environment folder are in. For the purposes of this example, that's the user's personal home folder.
2. Activate the virtual environment.
3. The virtual environment is now active and the script is executed. On Windows, the Python runtime executable is called 'py'

### Example for Macintosh or UNIX

The shell prompt is shown in light green. This is to illustrate when the virtual environment becomes active.

*Note that line 2 may be* `source ./env/scripts/activate` in some configurations.

```
jsmith@localhost:/$ cd $HOME
jsmith@localhost:~$ source ./env/bin/activate
(env) jsmith@localhost:~$ python3 auto_archive_downloader.py -site example.blackboard.com -i
```

If "python3" does not work see the note under "Installing and Configuring Python" below for Macintosh and Unix regarding the command "python" or "python3"

## Downloading a specific list of courses

The other major mode of the utility is **feed-file** mode. In this mode, the script takes an input list feed file and operate only on the specified courses.

### Feed File Format

You need to create a plaintext file consisting of the course IDs. There should be one course per line, and no other information can be included. The new line is the delimiter, don't include a comma or tab.

For example:

```
eng101
com304
phy202
```

Use the "Course ID", don't use the PK1, course name or batch\_uid ("External course key"). Use the full ID: a sub-string will not work.

### Running the Script with a Feed-File

Now run the script. As with the previous example, don't type the Windows command interpreter prompt. Pass the feed-file path to the script with the `-f` or `--file` arguments.

```
c:\> cd %userprofile%
c:\users\jsmith> .\env\scripts\activate
(env) c:\users\jsmith> py auto_archive_downloader.py -site example.blackboard.com -u admini:
```

## Using S3 Buckets

The above examples show downloading the archives to a folder on a drive "d" attached to the local machine called "**bulkarchiveoutput**." You can also use the script to **upload the archives to an S3 bucket**. You'll need to create an S3 bucket and an access key with enough rights to upload files. These steps are not detailed in this manual. Please consult the Official AWS manuals or AWS support for details.

To use the S3 bucket, export the AWS Access Key ID and secret key as environment variables with the AWS standard names. Provide the bucket path in the standard S3 URI-notation as the output option instead of a local folder path. In the following examples the S3 path is `s3://mybucket/myfolder/`

### S3 Configuration on Windows

Use the `set` command to set the variables

```
c:\> cd %userprofile%
```



1. Download the appropriate Windows installer package from <https://www.python.org>. Python for Windows includes `pip` but it doesn't include "requests" or "boto" modules, so those will be set up below.
2. Open a powershell or "cmd.exe" window.
3. Configure a virtual environment by typing the following commands. This command will create a virtual environment called "**env**" in your personal directory (usually: `c:\users\jsmith` or similar). The third line creates the `env` virtual environment. Depending on your settings, your command prompt will usually change to show the name of the environment.

```
c:\> cd %userprofile%
c:\users\jsmith> py -m pip install --user virtualenv
c:\users\jsmith> py -m venv env
```

4. Activate the virtual environment, and install the "requests" and "boto" packages into it.

```
c:\users\jsmith> .\env\scripts\activate
(env) c:\users\jsmith> pip install boto3 requests
```

## Macintosh and Unix/Linux

*Note about 'python' vs 'python3': depending on your OS configuration: you might need to type **python3** or you might need to type **python**. It is required to use Python version 3, Python version 2 will not work. The question is simply under which name the interpreter is installed. If Python versions 2 and 3 are co-installed, or on some older versions of Linux and Unix you might need to explicitly type `python3` to use the version 3 interpreter. However on some OS configurations, especially newer ones, Python version 2 is absent or depreciated and the name of the Python runtime interpreter is simply 'python'.*

1. Install Python 3 from [python.org](http://python.org), or from your operating system package repositories. If PIP is included then skip to step 3
2. If PIP is not included, install PIP:

```
jsmith@localhost:/$ python3 -m pip install --user --upgrade pip
```

3. For Macintosh, or desktop Unix open a terminal emulator.
4. Configure a virtual environment by typing the following commands. This command will create a virtual environment called "**env**" in your personal directory (usually: `/home/jsmith`, `/usr/local/home/jsmith` or similar). The third line creates the "`env`" virtual environment. Depending on your settings, your shell prompt will usually change to show the name of the environment. The following example is for the BASH shell:

```
jsmith@localhost:/$ cd $HOME
jsmith@localhost:~$ python3 -m pip install --user virtualenv
jsmith@localhost:~$ python3 -m venv env
```

5. Activate the virtual environment, and install the "requests" and "boto" packages into it. *Note that line 1 may be `source ./env/scripts/activate` in some configurations.*

```
jsmith@localhost:~$ source ./env/bin/activate
(env) jsmith@localhost:~$ pip install boto3 requests
```

## Other Operating Systems

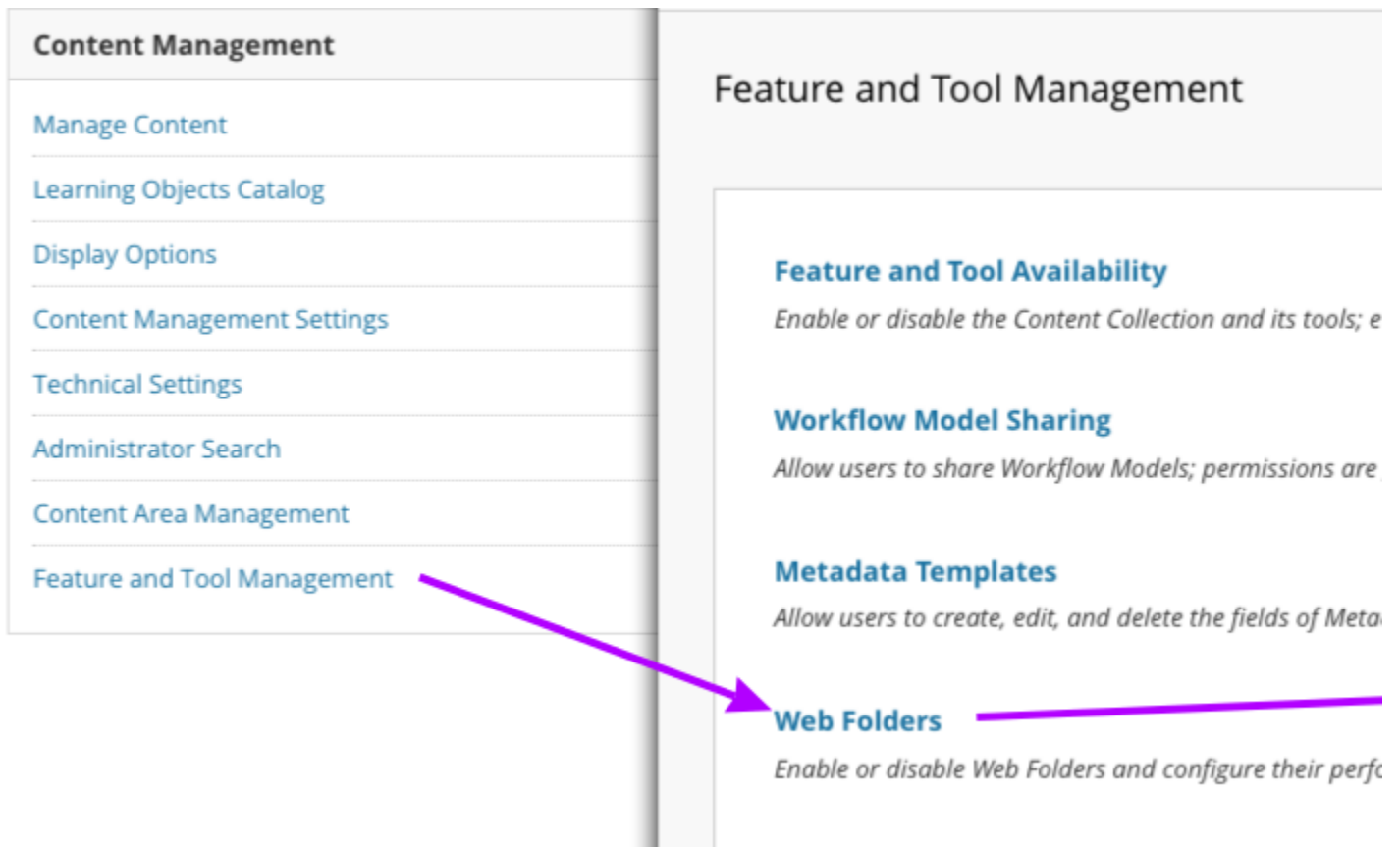
This script has only been officially tested with Windows, Macintosh and Linux. It may be possible to use other Operating Systems. Most other OSes will be configured relatively similarly to Macintosh and Unix. Consult the manuals of the other OS for information on adapting the instructions.

## Pre-configuration of Learn

In order to use the script WebDAV must be enabled. WebDAV availability is controlled by the "Web Folder" setting.

To check or enable WebDAV availability

1. Go to **System Admin Content Management module Feature & Tool Management Web Folders**
2. For "**Web Folders Available**" select the **Yes** radio button



## Frequently Asked Questions

*There's several archive files for the courses. Which archive will the script select?*

The script will select the latest available archive.

*Can I run more than one instance of the script at a time?*

It is not recommended to run multiple instances of the script concurrently. Archives can be large and performance issues are possible.

## ***Where can I get help installing Python and/or with PIP and Python virtual environments?***

Python lists help resources on its website: <https://www.python.org/about/help/>

An example is listed in *Related Links and Resources* but Enterprise client support can't assist with the installation or configuration of Python.

## ***Can you do this process for me?***

Blackboard Enterprise support is unable to operate this process for clients. Management of this data is a client process. For a charge, Consulting services may be available. To get started speak with your account manager.

## ***Can I use this script with Self or Managed Hosting?***

Currently: this script is only for SaaS hosting.

## ***Can I download to Google Drive, OneDrive, or some other cloud service provider?***

This is not currently an option. However you can download the files locally then re-upload them via the third-party vendor tools.

## ***Does this script pipeline from Xythos to S3, or does it download and re-upload?***

It downloads and re-uploads. It's not an S3 to S3 copy.

## ***There was an error, and the script terminated. What should I do?***

The script currently doesn't support retry or progress. If the script fails part way through, inspect the console output to find out where it errored out and address the root cause. Then, you'll need to update the feed file to reflect the new 'starting position,' and retry.

## ***Access Denied errors occur with the S3 bucket, but the permissions appear correct***

Check that the trailing slash is used after the folder. for instance `s3://mybucket/myfolder/` instead of `s3://mybucket/myfolder` . The latter may not work.

---

### **Summary**

This guide will get you started with the 'auto\_archive\_download.py' script available from Learn 3900.10 and higher to download automatic archives for long-term storage.

---

### **Title**

Automatic Archives Download Script Administrator Guide

---