

HTRC Feature Reader Installation

July 10, 2016

0.1 Installing Anaconda

For this workshop, we'll need to install the HTRC Feature Reader library for Python, alongside the data science libraries that it depends on.

We'll install it using Anaconda, an easy-to-install Python distribution that already includes most of the dependencies for the HTRC Feature Reader.

To install Anaconda, download the installer for your system from the Anaconda download page (<https://www.continuum.io/downloads>) and follow their instructions for installation of either the Windows 64-bit Graphical Installer or the Mac OS X 64-bit Graphical Installer. You can choose either version of Python for this lesson, but we recommend Python 3.

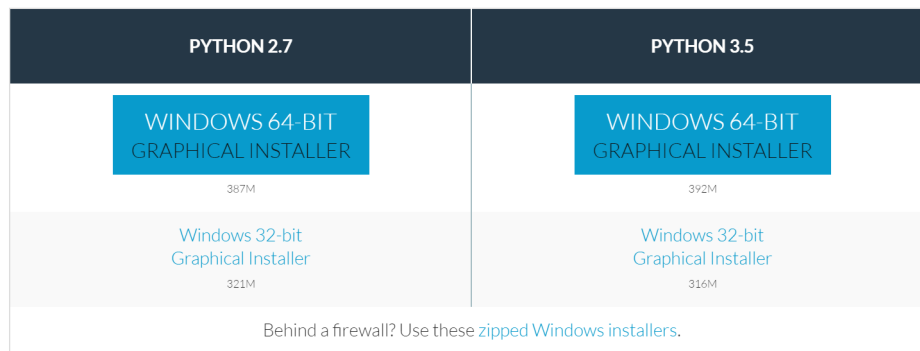


Figure 1: Conda Install

Advanced users: Why Anaconda? The HTRC Feature Reader depends on some fast low-level libraries that - because of optimizations specific to your system - may be slow if improperly installed. Anaconda make it easier to get the installation right.

0.2 Installing the HTRC Feature Reader

Once you have installed Anaconda, you need to open a command line to add the HTRC Feature Reader. This (should) be the only time you need a command line in this workshop!

First open a terminal application:

- Windows: Open 'Command Prompt' from the Start Menu and type: **activate**.
- Mac OS/Linux: Open 'Terminal' from Applications and type **source activate**.

Now, you need to type one command:

```
pip install htrc-feature-reader gensim
```

```
C:\Users\organism2>activate
Deactivating environment "C:\Users\organism2\AppData\Local\Continuum\Anaconda3"...
Activating environment "C:\Users\organism2\AppData\Local\Continuum\Anaconda3"...

[Anaconda3] C:\Users\organism2>
```

Figure 2: If Anaconda was properly installed, your command line should look like similar to this

This command installs the HTRC Feature Reader and its necessary dependencies. It also installs a library that we'll use for an optional module at the end of the lesson (if it causes issues, just remove the **gensim** part for now).

That's it! At this point you have everything necessary to start reading HTRC Feature Reader files.

psst, advanced users: If you installed the HTRC Feature Reader without Anaconda, you'll need to install two additional libraries for this lesson with `pip install matplotlib jupyter`. If you think your code is going slow, you should check that Numpy has access to [BLAS and LAPACK libraries](#) and install [Pandas recommended packages](#). The rest is up to you, advanced user!

Addendum: If you get an error about **ujson** with the above pip command, try to first install it separately with `conda install ujson`, then try the pip command again.