

Getting NLTK Up and Running on Mac OS X

Please note that there is a more direct version of these instructions that walks you through setting up everything around a Python 2.7 installation: read it [here](#).

So somehow, somewhere, you got interested in natural language processing, and the Natural Language Toolkit available for Python strikes you as one reasonable place to start. First of all, congratulations for wanting to go the building-block route as opposed to the already assembled route. The first few steps are a bit more complex, but I think you will be gratified pretty quickly with how much you can do and how quickly you are doing it. More importantly, you are doing it for yourself.

Here are the steps that we are going to take to get NLTK up and running on Mac OS X:

1. Install Xcode (and then the Xcode development tools).
2. Install [MacPorts](#).
3. Use MacPorts to install all the Python libraries you need, including NLTK.
4. There is no step four.

First, Xcode

Before you do anything else, you will need to open the [App Store](#) and download Xcode, Apple's Developer's Suite for creating OS X applications.

I'm not entirely clear what there is in Xcode or what Xcode installs that is needed for package managers like MacPorts, but MacPorts requires it, so go do it ... by the way, I think the new way this happens is that the App Store will install an application that you will find in the Applications folder and that you have to click on to install Xcode. (Again, I don't know why that is.)

As I note in an updated version of this [list of instructions](#), there is now supposed to be a shortcut from the command line to install what you need, but I have had various results and I have received reports of others having the same, various, results. Until this gets cleared up, my best recommendation is to install Xcode the usual way, and then to proceed as MacPorts directs you to.

Install MacPorts

Download the [Mac OS X Package pkg Installer](#) and step through the GUI install.

MacPorts should, as part of the install process, run `sudo port selfupdate -v` but you can always run it again. You know, just to make yourself feel better.

Use MacPorts to Install Python and the Libraries You Need for NLTK

Now you'll need to install a version of Python. In my case, I am building a setup around Python 2.7, and so I entered `sudo port -v install python27`. The `-v` option gives you a verbose description of what's happening. Be prepared to watch a lot of stuff scroll by. (If you'd rather not see all that and having the machine quietly do its thing, you can leave the `-v` off. Good for you for having quiet confidence in your Mac.)

How did I know to type in `python27` and not just `python`? Good question. MacPorts gives you some nice functionality with its `search` feature, which you can use to find MacPort portfiles. If you type in:

```
port search python
```

A whole lot of stuff is going to fly by, but you can scroll to the middle of the list to see all the versions of Python that are available for you to install. As of this writing, you can install everything from Python 2.4 to Python 3.1.

By the way, perhaps you are savvy enough to know that your Mac already comes with Python (and Perl and Ruby and PHP and goodness knows what else) already installed. Yes, it does, but the consensus seems to be that leave those system installations well enough alone. If you screw up Python, you want it to be an instance the system doesn't need. Don't worry. I'm terrible at this coding business and I've yet to screw up Python. Python is very good at telling you that you've screwed and, thank you very much, it won't follow your orders down the path to electronic perdition.

We are going to install the most recent version of Python 2, which is Python 2.7. (As of this writing, Python 3 is still considered a draft version of Python — it's complicated beyond my ability to explain. It's just as well, as we will see in a minute, the NLTK support only runs up to Python 2.7.

```
port install python27
```

Once that's done, and I should remind you that depending upon your connection speed and the size of any particular portfile, this could take a while, you will want to make it so that your computer turns to your nice custom install of Python and not the one that came with the system. I usually accomplish this by editing my `.bash_profile`, but this did not work for me. Luckily, MacPorts has the solution:

```
sudo port install python_select
```

Once you've done this, enter `sudo port select --set python python27` and you're done with your base installation of Python. Now it's time to process some natural languages, or process languages naturally, or ... you know, whatever it is we are going to do with the NLTK.

Install NLTK, But First Some Dependencies

In all honesty, here is where the magic really happens. I know that sounds weird, especially when we are talking

about something that happens at the command line, but, honestly, MacPorts, makes the rest of this so easy that you might just wonder, I tell you, why everyone prefers clanging around with GUI installer packages that you have to go find, download, open, and click on.

Anytime you want to install anything using MacPorts, the best place to start is to see if a *portfile* is available. (Else you can't.) Using the new-found power of the `port` command, this is quite simple:

```
port search nltk
```

Right? We want to see if the NLTK is available as a *portfile* and we want to see what, if any, versions are available. Here is what the search turns up:

```
py-nltk @2.0.1rc1 (python, textproc)
    Natural Language Toolkit

py24-nltk @2.0.1rc1 (python, textproc)
    Natural Language Toolkit

py25-nltk @2.0.1rc1 (python, textproc)
    Natural Language Toolkit

py26-nltk @2.0.1rc1 (python, textproc)
    Natural Language Toolkit

py27-nltk @2.0.1rc1 (python, textproc)
    Natural Language Toolkit

Found 5 ports.
```

Great! There's a version of the NLTK available that matches the version of Python we just installed. But before you install the NLTK, you will want to know what other Python modules it requires. (Okay, sidebar here, this is really something MacPorts does for you, but I can't help being a bit of a control freak and wanting to take care of this myself.) Again, MacPorts has your back. Enter:

```
port deps py27-nltk
```

And MacPorts reports back:

```
Full Name: py27-nltk @2.0.1rc1_0
Library Dependencies: py27-numpy, py27-yaml, py27-matplotlib
```

Now all you have to do is `port install` and then add each of those. While you're at it, you might as well add `py27-scipy` to your list. Don't ask; just do it.

Conclusion

That's it. You're done. Fire up IDLE, `import nltk`, and you can do an amazing assortment of things.

If you like, you can show your love for the makers of the NLTK and buy their book. It's available both at [O'Reilly](#) in a variety of formats and at [Amazon](#). Personally, I find the [NLTK Cookbook](#) a little less helpful, but this could be purely a matter of a cognitive style mismatch. It happens. It's not the author's fault. It's me. Really.

Updates

The first version of this post detailed how to “get up and running” using the HomeBrew package manager, but after running into a number of difficulties, I discovered that something about the HomeBrew setup just doesn't work. I updated this post to point to the now preferred use of MacPorts, but this post continues to be the most popular. Since I believe it's turning up in search engines because of the obviousness of the post title, and because I don't want to send people down the path of pain, I have removed the HomeBrew directions and replaced them with a much more detailed version of the MacPorts installation path. If you would like to try HomeBrew, please drop me a note — my contact info is on the [About](#) page — and I'll be glad to send you directions.

(Visited 605 times, 1 visits today)

Related

[The Complete Python for Text Analysis](#)
2013 December 29
In "note"

[Getting up and running with NLTK](#)
2011 April 8
In "note"

[Python in XCode](#)
2013 July 23
In "note"

*Posted on 2012 January 31 by johnlaudun. This entry was posted in **note** and tagged **python**, **terminal**, **unix**. Bookmark the **permalink**. Edit*

One thought on “Getting NLTK Up and Running on Mac OS X”

ping! [logbook « Text Analytics 101 \(Edit\)](#)

Comments are closed.