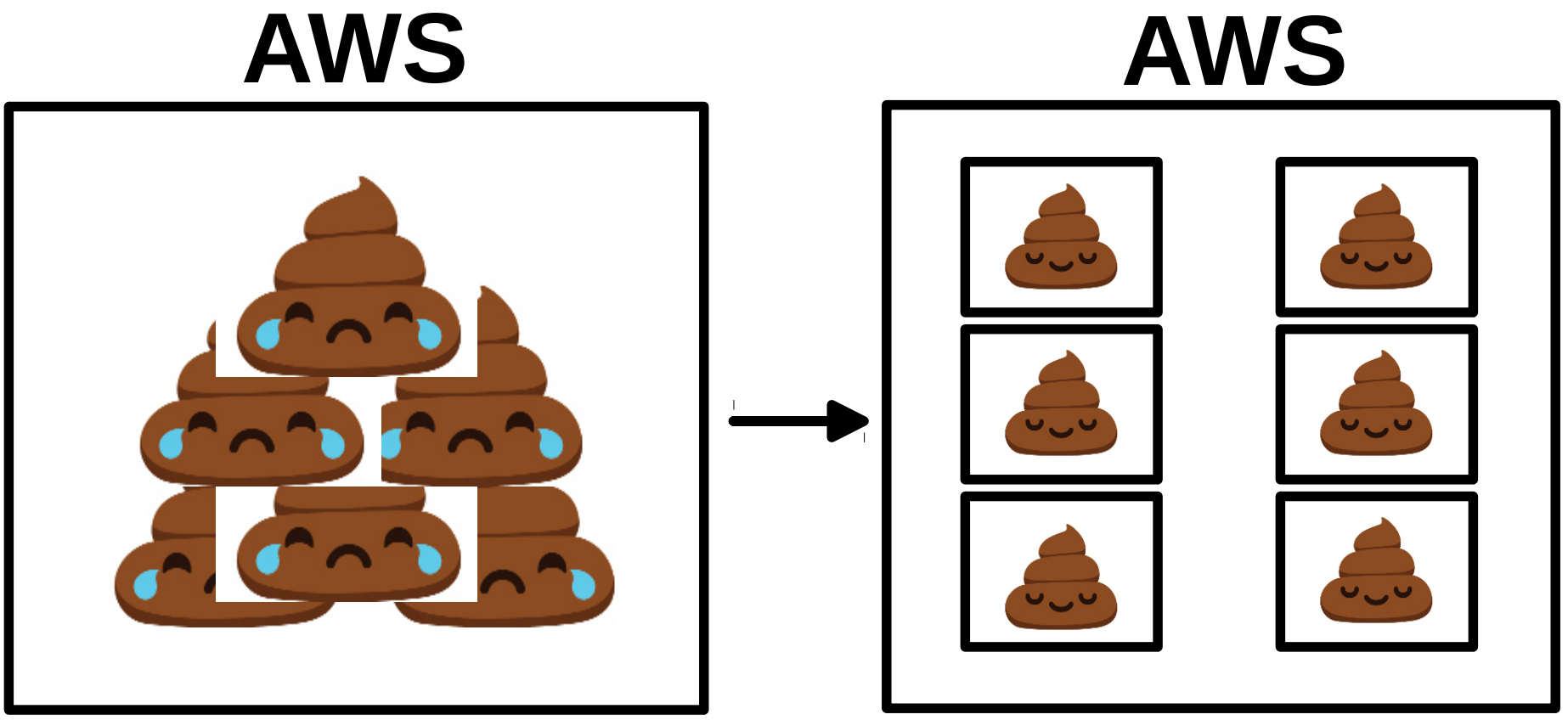
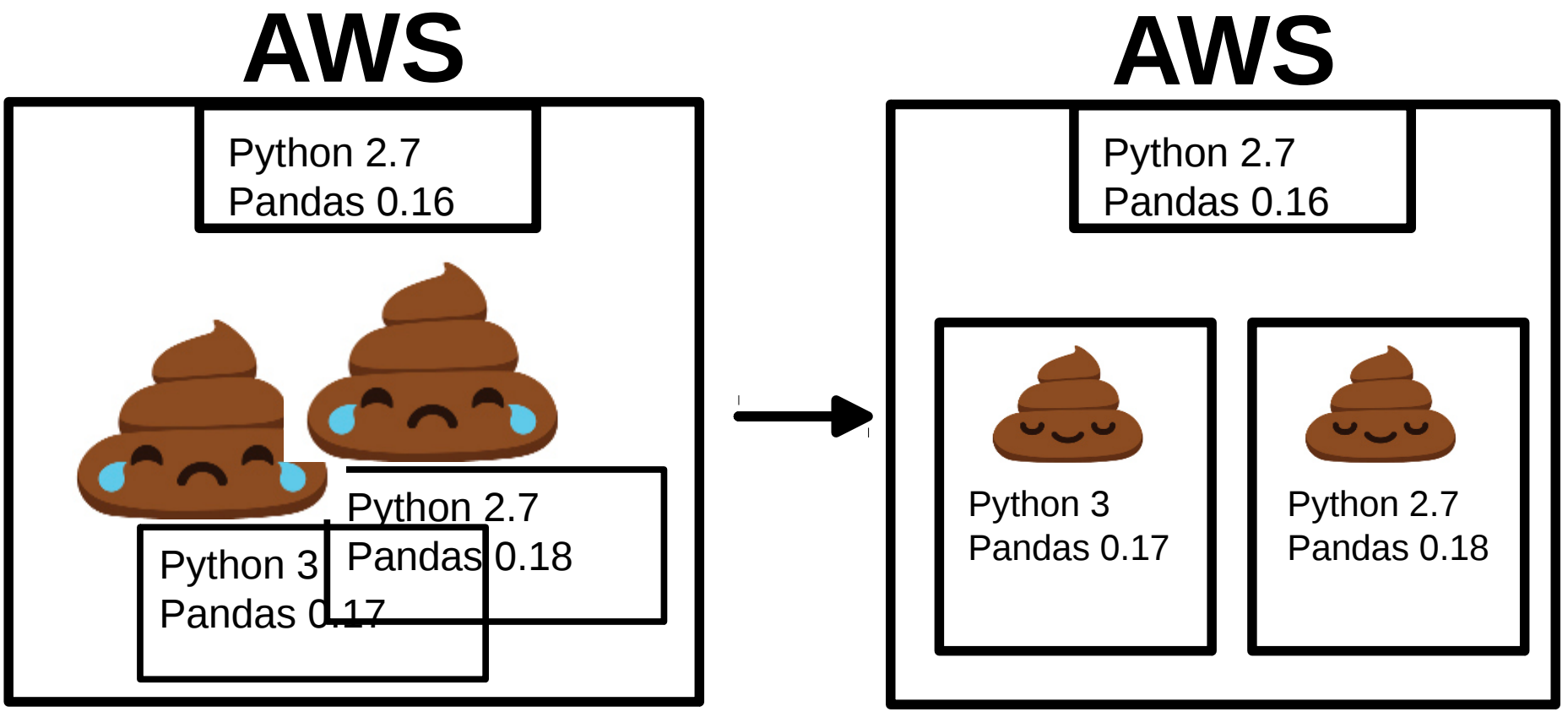
Virtual Environments Workshop (11/2/2016)

A quick intro and exercise, followed by a summary and links to resources on the last two slides.

Conda Virtual Environments



Virtual Environments are lil' portable sandboxes where you can run your code, the way you intended.



Let's make environments!

Log into AWS (or use conda on your own cpu)
Figure out which environments already exist

>> *conda info --envs*

Make a new environment

>> *conda create --name $ENV_NAME python=$X.X $PACKAGES*

Ex: >> *conda create --name isaac_poop python=3.5 pandas*

Activate the environment

>> *source activate isaac_poop*

How to add packages

Conda recommends you install them when you create the environment to avoid dependency conflicts. That might not be practical. To add a new package, do:

>> *conda install $PACKAGE_NAME*

If there is a dependency conflict, conda's still got your back:

Try to install an old version of pandas:

>> *conda install pandas=0.16*

```
The following packages will be DOWNGRADED due to dependency conflicts:

    numpy:  1.11.2-py35_0      --> 1.10.4-py35_2
    pandas: 0.19.0-np111py35_0 --> 0.16.2-np110py35_0

Proceed ([y]/n)? n
```

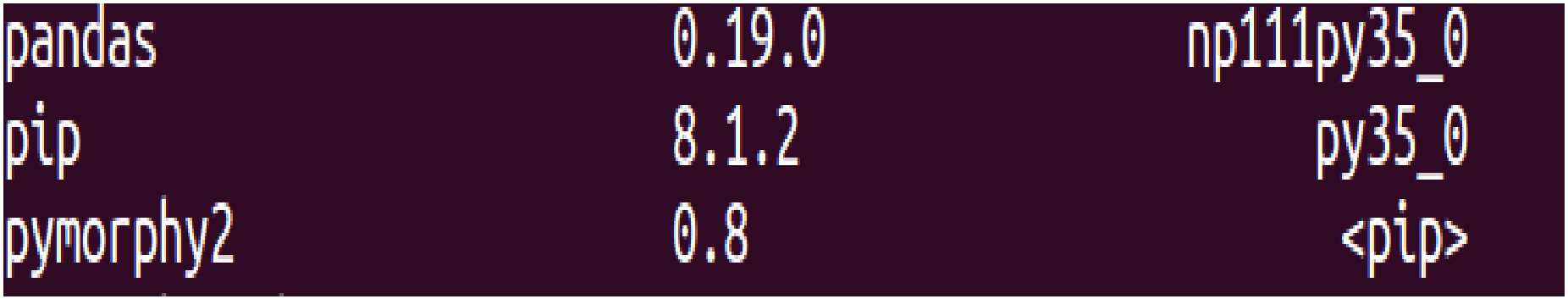
Cool story, bro. But what about non-Anaconda packages?

Pip still works!

>> *pip install pymorphy2*

Check which packages are installed

>> *conda list*



Cool story, bro. But what about non-Python packages?

Try Bioconda – bioinformatics packages for conda (i.e. bowtie, bioconductor pacakges, etc), or google it!

https://bioconda.github.io/index.html

Installation

With an activated Bioconda channel (see 2. Set up channels), install with:

```
conda install bowtie2
```

and update with:

```
conda update bowtie2
```

container ready

A Docker container is available at https://quay.io/repository/biocontainers/bowtie2.

Mmkay, how do I share an environment?

Simple. Export your environment to the current folder:

>> conda env export > environment.yml

>> nano environment.yml

```
name: isaac_poop
channels: !!python/tuple
- !!python/unicode 'defaults'
dependencies:
- mkl=11.3.3=0
- numpy=1.11.2=py35_0
- pandas=0.19.0=np111py35_0
- pip=8.1.2=py35_0
```

Conda dependencies

[...]

```
- pip:
  - dawg-python==0.7.2
  - docopt==0.6.2
  - pymorphy2==0.8
  - pymorphy2-dicts==2.4.393442.3710985
prefix: /home/ubuntu/anaconda2/envs/isaac_poop
```

pip dependencies

Sharing is easy.

Exit and delete the current environment

Exit

>> *source deactivate*

Delete

>> *conda remove -n $ENV_NAME --all*

Now re-create the environment using environment.yml

Note: You must be in the folder containing environment.yml

>> *conda env create -f environment.yml*

conda is a better choice than other virtual environment options

Mindshare

Conda is already installed on AWS and we know how to work it

Extensibility

Can be used for non-python packages:

Using R with conda

Decreased Filesize

Conda hardlinks to installed packages in different environments (we don't want the bloat of 20 people each downloading 5 versions of scipy)

Support

It's supported by a company with some $, so less likely to become vaporware (but is still open-source)

Recap

Make a new environment

>> conda create --name $ENV_NAME python=$X.X $PACKAGES

Activate the environment

>> source activate $ENV_NAME

Deactivate the environment

>> source deactivate

Install a package

>> conda install $PACKAGE_NAME

Export your environment:

>> conda env export > environment.yml

Load an environment (in folder with environment.yml)

>> conda env create -f environment.yml

Delete an environment

>> conda remove --name $ENV_NAME --all

Delete a package

>> conda remove --name $ENV_NAME $PACKAGE_NAME

Resources

Cheat sheet
(but don't use their way of making a requirements file)

Sample Workflow

Docs

Poop emojis