

# **Lab1 Getting started with NP**

**Introduction to Human language Technology**

# Preparations

- Make sure you have administrator rights on your laptop to install whatever you want
- Install software in the same environment (see later), check if you are in the same environment
- Understand the basics of the command line (windows) or terminal (mac, linux)
- Install Anaconda with Python 3.8 or higher
- Familiarise yourself with Jupyter notebooks or labs
- Use a “plain” (not Word) text editor for inspecting data:
  - Windows: Notepad++
  - Mac/Linux: Atom
- Useful (gotta learn that anyway later) install Git to interact with GitHub:
  - Install Git: <https://git-scm.com/book/en/v2/Getting-Started-Installing-Git>

# Getting the Lab notebooks

*Check regularly for updates*

Get the download link

- <https://github.com/cltl/ma-hlt-labs>

- Git installed:

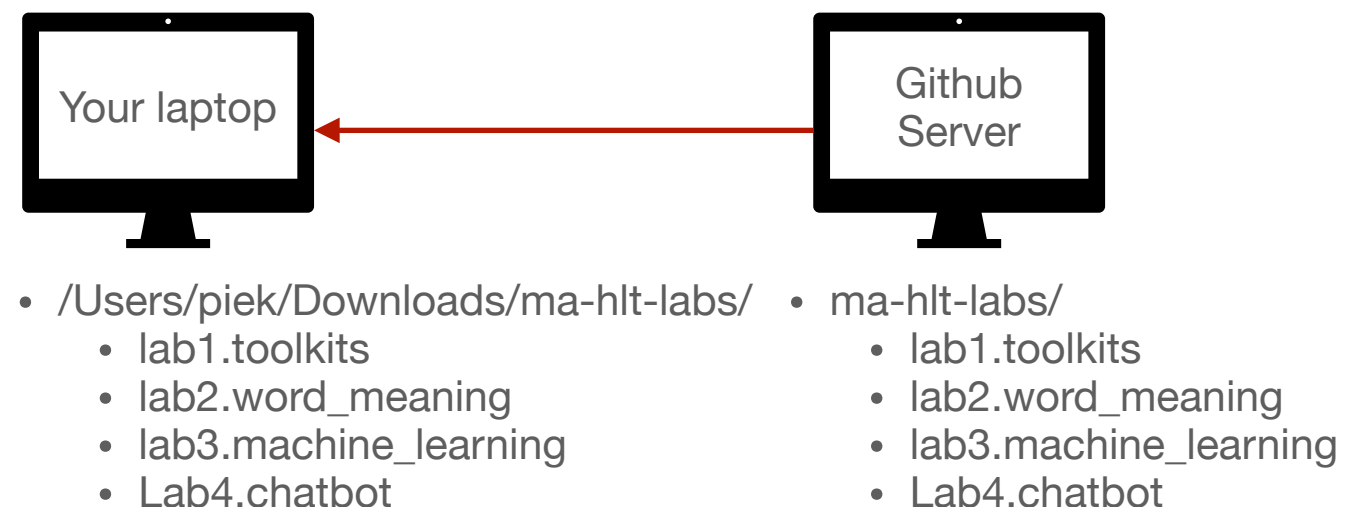
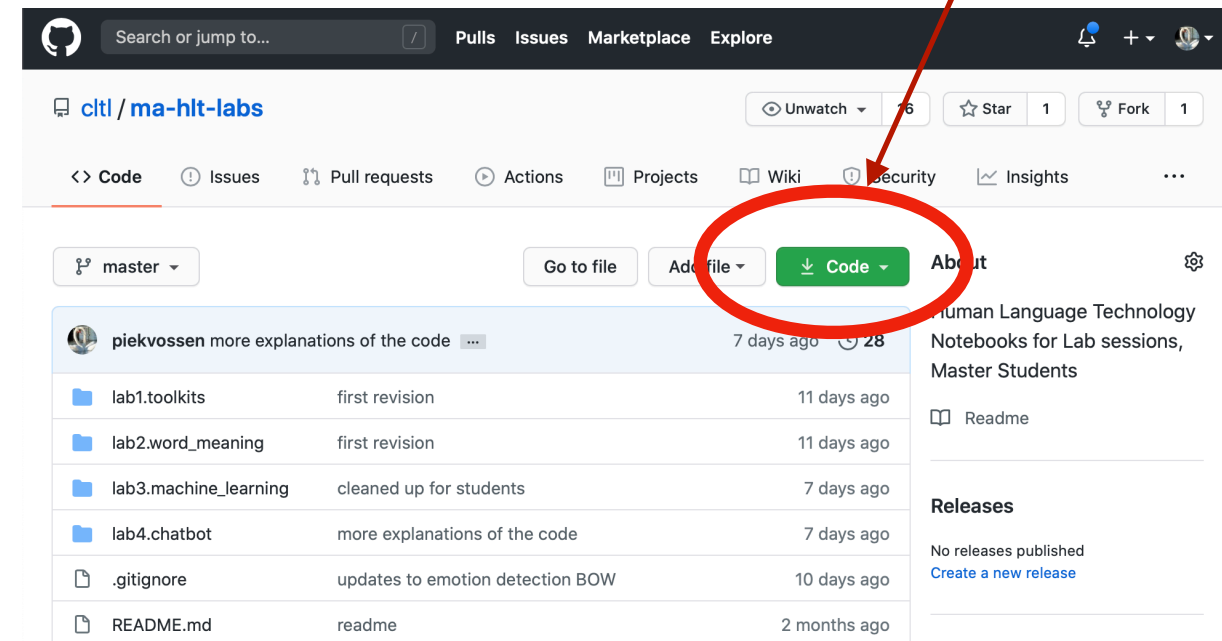
- Clone with ssh:

- `> git clone git@github.com:cltl/ma-hlt-labs.git`

- No local Git installed:

- Download ZIP file

- Unpack anywhere



# Some Git

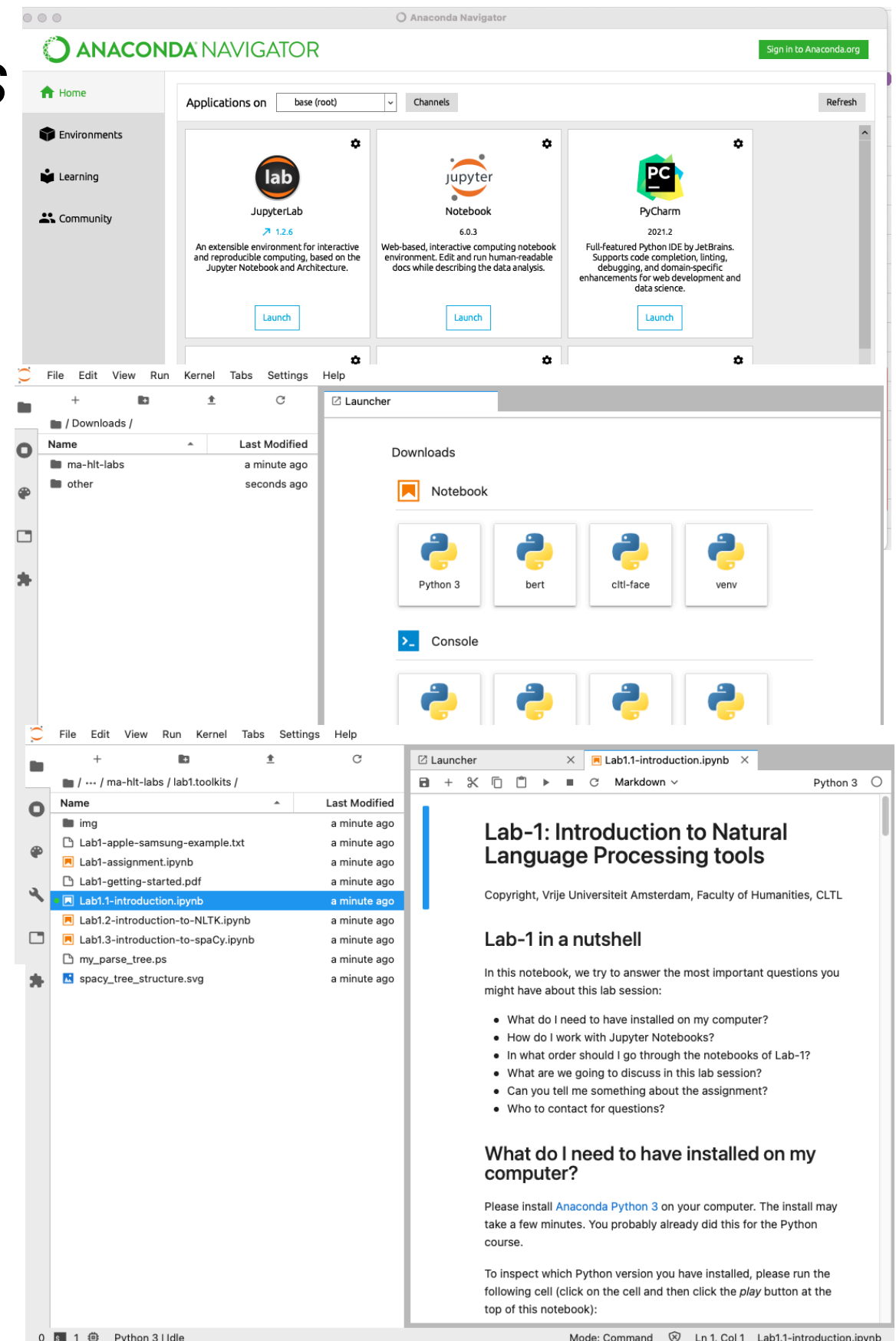
*Only for those who use Git from command line*

- At any moment in time you can check for updates on the server
  - Use “cd” to navigate to the directory with your local copy
  - This is a git directory (there is an invisible git flag with each file downloaded from a GitHub repository)
  - Type the command: “git pull”
  - Any updates will be downloaded from the server BUT if you changed a file Git will say you are out of sync.
- WARNING:
  - You can change your local copy of the Git repository as much as you like but your local copy is no longer in sync
  - Git will see that there is a difference and refuse to perform “pull” to avoid that your local changes are destroyed:
    - Type the command: “git stash” to ignore your local changes (these are lost) after which you can do “git pull”
    - Rename the directory of your local copy, e.g. “ma-hlt-labs-mine” and do a new “git clone” to get two copies: your local version and the latest version from the server.

# Anaconda Navigator

## *First steps to using notebooks*

- After launching the Anaconda GUI you get a window as shown here with various tools
- When click on the launch button in the JupyterLab or JupyterNotebook panel, it will launch a browser with the Jupyter application opening in the Download folder in my case, where I cloned ma-htl-labs from Github
- Entering ma-htl-labs and lab1-toolkits, I find notebook files with the extension .ipynb
- Double clicking on Lab1.1-introduction.ipynb loads the first notebook with all instruction in your browser
- You are ready to go!



# Virtual environments

## Installing software within a save silo

- When installing many packages with even more dependencies, conflicts may arise between versions
- It is wise to create a new “virtual environment” (venv) first and install and run your code within (a kind of silo to install and run safely)
- Basic instructions can be found here:
  - <https://docs.python.org/3/library/venv.html>
- Two options:
  - Within the Anaconda navigator
  - From command line

The screenshot displays the Anaconda Navigator application window. The top bar shows the Anaconda logo and the name "ANACONDA NAVIGATOR". On the right side of the top bar, it indicates the user is signed in as "Piek.Vossen" with a "Sign out" button.

The left sidebar contains navigation options: Home, Environments (highlighted), Learning, and Community. At the bottom of the sidebar are links to Documentation and Developer Blog, along with social media icons for Twitter, YouTube, and GitHub.




The main area is divided into two panels. The left panel, titled "Environments", lists several environments: "base (root)" (selected), "anaconda3", "evaluating-conversations-a...", and "thought-selection". Below this list are buttons for "Create", "Clone", "Import", and "Remove".

The right panel shows a search results table for the query "jupyter". The table has columns for Name, Description, and Version. It lists 19 packages matching the search criteria. A red circle highlights the "Create" button in the bottom-left corner of the interface.

Name	Description	Version
_ipyw_jlab_nb_ex...	A configuration metapackage for enabling anaconda-bundled jupyter extensions	0.1.0
ipykernel	Ipython kernel for jupyter	5.1.4
ipywidgets	Jupyter interactive widgets	7.5.1
jupyter	Jupyter metapackage. install all the jupyter components in one go.	1.0.0
jupyter-client		5.3.4
jupyter-console		6.1.0
jupyter-core		4.6.1
jupyter_client	Jupyter protocol implementation and client libraries	5.3.4
jupyter_console	Jupyter terminal console	6.1.0
jupyter_core	Core common functionality of jupyter projects.	4.6.1
jupyterlab	Jupyterlab pre-alpha	1.2.6
jupyterlab-server		1.0.6
jupyterlab_server	A set of server components for jupyterlab and jupyterlab like applications.	1.0.6

19 packages available matching "jupyter"

19 packages available matching "jupyter"

 Home Environments Learning Community[Documentation](#)[Developer Blog](#)

Search Environments



base (root)



anaconda3

evaluating-conversations-a...

thought-selection

Installed

Channels

Update index...

jupyter



Name



T

Description

Version



\_ipyw\_jlab\_nb\_ex...



A configuration metapackage for enabling anaconda-bundled jupyter extensions

0.1.0



ipykernel



Ipython kernel for jupyter

[5.1.4](#)

ipywidgets



Jupyter interactive widgets

7.5.1



jupyter



Jupyter metapackage. install all the jupyter components in one go.

1.0.0

Create new environment



Name: hlt\_venv

Location: /opt/anaconda3/envs/hlt\_venv

Packages: ☒ Python

3.8

☐ R

r

Cancel

Create



Create



Clone



Import



Remove



jupyterlab-server



1.0.6



jupyterlab\_server







A set of server components for jupyterlab and jupyterlab like applications.

[1.0.6](#)

19 packages available matching "jupyter"



-  Home
-  Environments
-  Learning
-  Community

[Documentation](#)[Developer Blog](#)

base (root)

anaconda3

evaluating-coversations-a...

hlt\_venv 

thought-selection



Create



Clone



Import

































Remove

Installed

Channels

Update index...

Name	T	Description	Version
 ca-certificates		Certificates for use with other packages.	2022.4.26
 certifi		Python package for providing mozilla's ca bundle.	2022.6.15
 libcxx		C++ standard library	12.0.0
 libffi		A portable foreign function interface library.	3.3
 ncurses		Library for text-based user interfaces	6.3
 openssl		Openssl is an open-source implementation of the ssl and tls protocols	1.1.1q
 pip		Pypa recommended tool for installing python packages	22.1.2
 python		General purpose programming language	3.8.13
 readline		Library for editing command lines as they are typed in	8.1.2
 setuptools		Download, build, install, upgrade, and uninstall python packages	61.2.0
 sqlite		Implements a self-contained, zero-configuration, sql database engine.	3.38.5
 tk		A dynamic programming language with gui support. bundles tcl and tk.	8.6.12
 wheel		A built-package format for python.	0.37.1
 xz		Data compression software with high compression ratio	5.2.5
 zlib		Massively spiffy yet delicately unobtrusive compression library	1.2.12

15 packages available

Home

Environments

Learning

Community

Search Environments

base (root)

anaconda3

evaluating-conversations-a...

hlt\_venv

thought-selection

Not installed

Channels

Update index...

jupyter

Name

Description

Version



\_ipyw\_jlab\_nb\_ex...



A configuration metapackage for enabling anaconda-bundled jupyter extensions

0.1.0



hdijupyterutils



Project with useful classes/methods for all projects created by the hdinsight team at microsoft around jupyter

0.12.9



ipykernel



Ipython kernel for jupyter

5.3.0



jupyterlab\_widgets



Jupyter interactive widgets

7.5.1



jupyter



Jupyter metapackage. install all the jupyter components in one go.

1.0.0



jupyter\_client



Jupyter protocol implementation and client libraries

6.1.3



jupyter\_console



Jupyter terminal console

6.1.0



jupyter\_core



Core common functionality of jupyter projects.

4.6.3



jupyter\_dashboar...



An add-on for jupyter notebook

0.9.1



jupyter\_kernel\_ga...



Jupyter kernel gateway

2.4.0



jupyter\_telemetry



0.0.5



jupyterhub



Multi-user server for jupyter notebooks

1.0.0



jupyterlab



Jupyterlab pre-alpha

2.1.5



jupyterlab\_launcher



A launcher for jupyterlab based applications.

0.6.0



jupyterlab\_server



A set of server components for jupyterlab and jupyterlab like applications.

1.1.5

- Open Terminal
- Open with Python
- Open with IPython
- Open with Jupyter Notebook

Install Packages

81 packages will be installed

	Name	Unlink	Link	Channel
1	jupyterlab	-	3.3.2	pkgs/main
2	*anyio	-	3.5.0	pkgs/main
3	*appnope	-	0.1.2	pkgs/main
4	*argon2-cffi	-	21.3.0	pkgs/main
5	*argon2-cffi-bindings	-	21.2.0	pkgs/main
6	*asttokens	-	2.0.5	pkgs/main

\* indicates the package is a dependency of a selected package

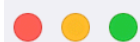
Cancel

Apply

30 packages available matching "jupyter" 1 package selected

Apply

Clear

[Home](#)[Environments](#)[Learning](#)[Community](#)[Documentation](#)[Developer Blog](#)

Not installed



Channels

Update index...

jupyter



base (root)

anaconda3

evaluating-conversations-a...

hlt\_venv



thought-selection

Open Terminal

Open with Python

Open with IPython

Open with Jupyter Notebook

Name



T

Description

Version



\_ipyw\_jlab\_nb\_ex...



A configuration metapackage for enabling anaconda-bundled jupyter extensions

0.1.0



hdijupyterutils



Project with useful classes/methods for all projects created by the hdinsight team at microsoft around jupyter

0.12.9



ipywidgets



Jupyter interactive widgets

7.5.1



jupyter



Jupyter metapackage. install all the jupyter components in one go.

1.0.0



jupyter\_console



Jupyter terminal console

6.1.0



jupyter\_dashboar...



An add-on for jupyter notebook

0.9.1



jupyter\_kernel\_ga...



Jupyter kernel gateway

2.4.0



jupyter\_telemetry



0.0.5



jupyterhub



Multi-user server for jupyter notebooks

1.0.0



jupyterlab\_launcher



A launcher for jupyterlab based applications.

0.6.0



metakernel



Metakernel for jupyter.

0.24.3



nb\_conda



Conda environment and package access extension from within jupyter

2.2.1



nb\_conda\_kernels



Launch jupyter kernels for any installed conda environment

2.2.3



nbpresent



Next generation slides for jupyter notebooks

3.0.2



nbserverproxy



Jupyter server extension to proxy web services

0.8.8



pivottablejs



Pivottable.js integration for jupyter/ipython

0.0.0



Create



Clone



Import

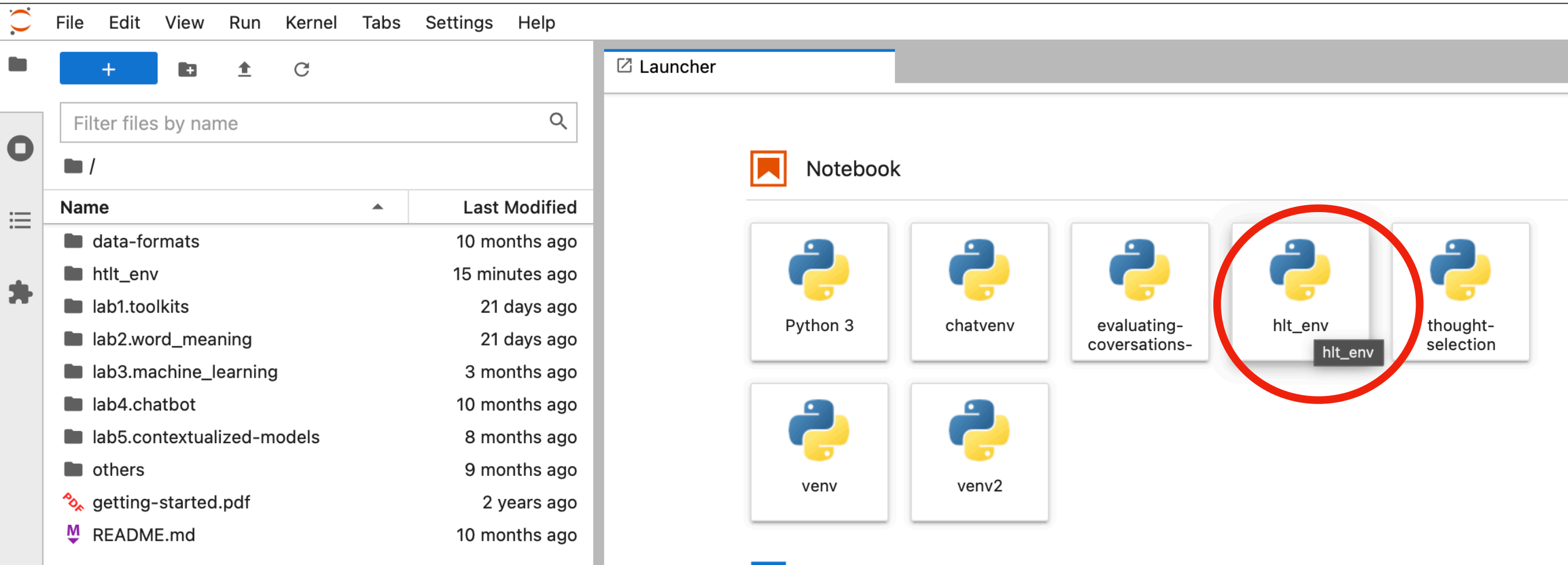


Remove

22 packages available matching "jupyter"

# Virtual environments

## Starting from Anaconda



The screenshot displays the Anaconda Launcher interface. On the left, a file explorer shows a directory structure with folders like 'data-formats', 'hlt\_env', 'lab1.toolkits', 'lab2.word\_meaning', 'lab3.machine\_learning', 'lab4.chatbot', 'lab5.contextualized-models', and 'others'. It also lists files 'getting-started.pdf' and 'README.md'. The right pane, titled 'Launcher', shows a grid of virtual environments, each represented by a Python logo icon and a name. The environments are: 'Python 3', 'chatenv', 'evaluating-conversations-', 'hlt\_env', 'thought-selection', 'venv', and 'venv2'. The 'hlt\_env' environment is circled in red, indicating it is the selected or active environment.

Name	Last Modified
data-formats	10 months ago
hlt_env	15 minutes ago
lab1.toolkits	21 days ago
lab2.word_meaning	21 days ago
lab3.machine_learning	3 months ago
lab4.chatbot	10 months ago
lab5.contextualized-models	8 months ago
others	9 months ago
getting-started.pdf	2 years ago
README.md	10 months ago

Launcher

Notebook

Python 3

chatenv

evaluating-conversations-

hlt\_env

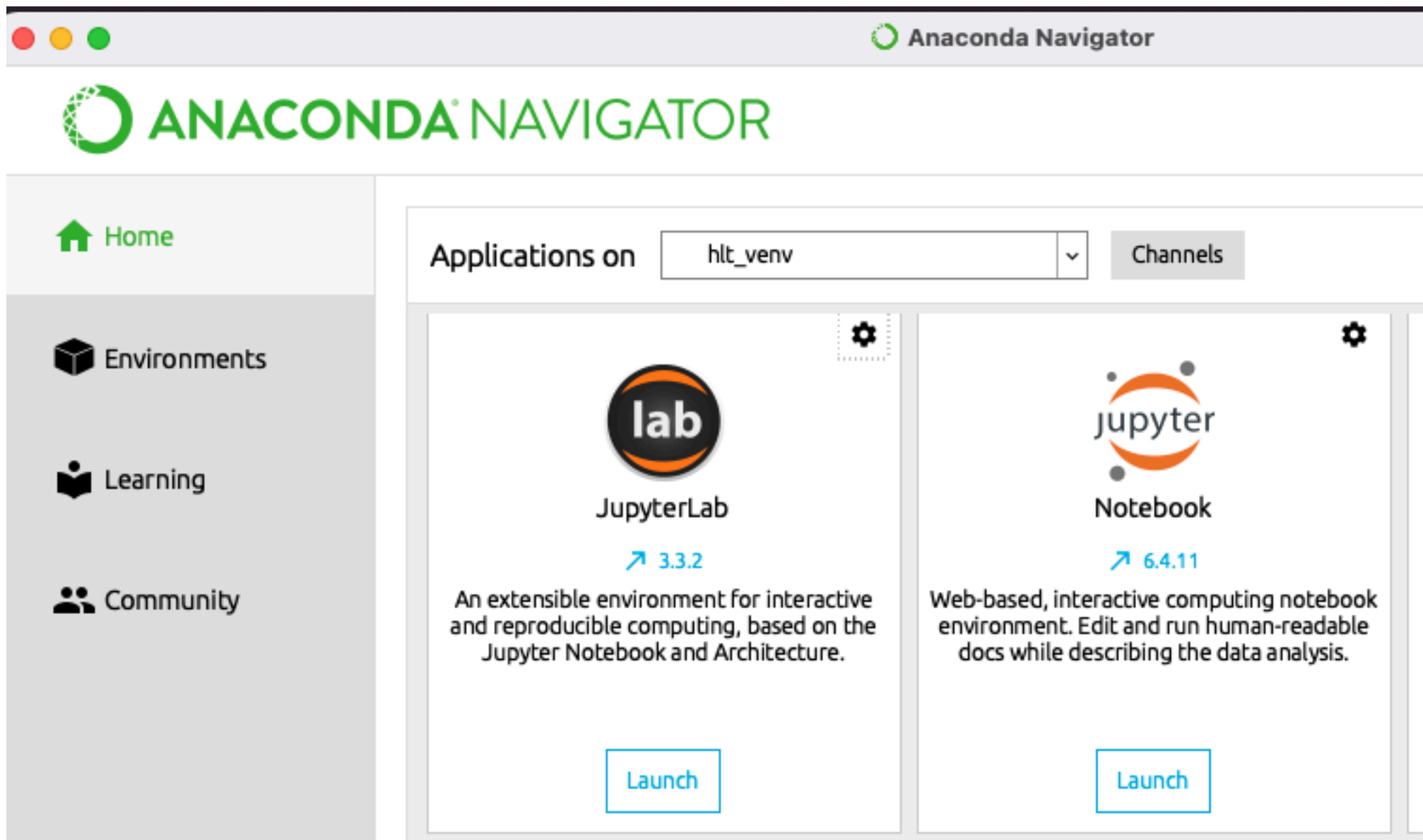
thought-selection

venv

venv2

# Virtual environments

## Installing software within a save silo



# Virtual environments

## Command line

- Open a terminal and navigate to the folder “ma-hlt-labs”
  - Do: `python -m venv hltl_env`
- Check if the directory “hltl\_env” is created locally
  - Do: `source hltl_env/bin/activate`
- Check if your prompt changed from (base):
  - (base) `piek@PTJMs-MacBook-Pro ma-hlt-labs %`
  - TO:
  - (hltl\_env) (base) `piek@PTJMs-MacBook-Pro ma-hlt-labs %`

# Virtual environments

## Command line

- If the prompt changed, the virtual environment is active.
  - Do: pip list
  - This shows the packages activated in this environment

```
(htlt_env) (base) piek@PTJMs-MacBook-Pro ma-hlt-labs % pip list
```

Package	Version
-----	-----
pip	21.2.3
setuptools	57.4.0

- Do: pip install
- Do: pip list

```
(htlt_env) (base) piek@PTJMs-MacBook-Pro ma-hlt-labs % pip list
```

Package	Version
-----	-----
click	8.1.3
joblib	1.1.0
nlTK	3.7
pip	21.2.3
regex	2022.7.9
setuptools	57.4.0
tqdm	4.64.0

- Make jupyter lab/notebook use the new virtual environment “htlt\_env”
  - Do: pip install ipykernel
  - Do: python -m ipykernel install --name=htlt\_env
  - Do: ipython kernel install --user --name=htlt\_env
  - Do: jupyter lab within the virtual environment

# Virtual environments

## Command line

The screenshot displays the JupyterLab application interface. On the left, a file browser sidebar shows a directory structure with folders like 'data-formats', 'hlt\_env', 'lab1.toolkits', 'lab2.word\_meaning', 'lab3.machine\_learning', 'lab4.chatbot', 'lab5.contextualized-models', and 'others', along with files 'getting-started.pdf' and 'README.md'. The 'hlt\_env' folder is highlighted. On the right, the 'Launcher' tab is active, showing a grid of virtual environments. The 'hlt\_env' environment is circled in red, indicating it is the selected environment. The environments shown are Python 3, chatvenv, evaluating-conversations-, hlt\_env, thought-selection, venv, and venv2.

File Edit View Run Kernel Tabs Settings Help

Filter files by name

Name Last Modified

data-formats	10 months ago
hlt_env	15 minutes ago
lab1.toolkits	21 days ago
lab2.word_meaning	21 days ago
lab3.machine_learning	3 months ago
lab4.chatbot	10 months ago
lab5.contextualized-models	8 months ago
others	9 months ago
getting-started.pdf	2 years ago
README.md	10 months ago

Launcher

Notebook

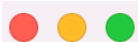
Python 3 chatvenv evaluating-conversations- hlt\_env thought-selection venv venv2



# Installing software

*Software packages and toolkits are installed from command line*

- Further instructions can be found on Canvas in the Basic Computer Skills modules
- Linux, Mac (Unix) users can launch a terminal (search for this application and keep it in your dock)
- Windows users can launch a command line window
  - <https://docs.microsoft.com/en-us/windows/terminal/get-started>
  - or
  - Simulate a unix/linux terminal on windows using Git-bash: <https://www.atlassian.com/git/tutorials/git-bash>
  - or
  - Open a terminal from the Anaconda navigator (see next slide for a screen dump)
- In Jupyter notebooks you can also carry out command line instructions in a cell by prefixing it with “%”
  - %ls -l
  - %pip install nltk
  - %mkdir test
  - %cd test
  - Etc.....



Home

Environments

Learning

Community

Search Environments

Installed

Channels

Update index...

Search Packages

	Name	T	Description	Version
base (root)				
anaconda3	anyio			3.5.0
evaluating-conversations-a...	appnope		Disable app nap on os x 10.9	0.1.2
hlt_venv				21.3.0
thought-selection				21.2.0
				2.0.5
	attrs		Attrs is the python package that will bring back the joy of writing classes by relieving you from the drudgery of implementing object protocols (aka dunder methods).	21.4.0
	babel		Utilities to internationalize and localize python applications	2.9.1
	backcall		Specifications for callback functions passed in to an api	0.2.0
	beautifulsoup4		Python library designed for screen-scraping	4.11.1
	bleach		Easy, whitelist-based html-sanitizing tool	4.1.0
	brotlipy			0.7.0
	ca-certificates		Certificates for use with other packages.	2022.4.26
	certifi		Python package for providing mozilla's ca bundle.	2022.6.15
	cffi		Foreign function interface for python calling c code.	1.15.0
	charset-normalizer			2.0.4

- Open Terminal
- Open with Python
- Open with IPython
- Open with Jupyter Notebook

Documentation

Developer Blog



Create

Clone

Import

Remove

107 packages available