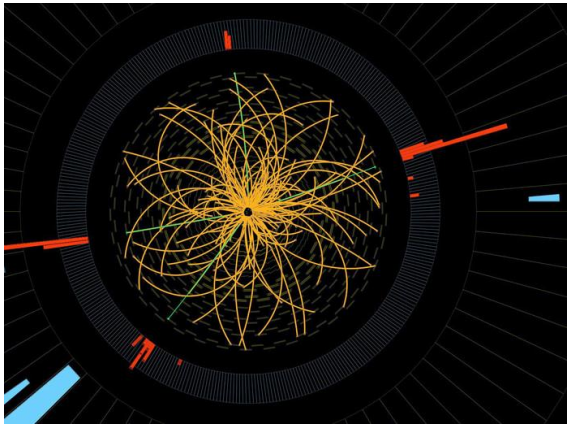
Jupyter Notebook for Scientific Software

Umar Farooq

Ilgwon Park

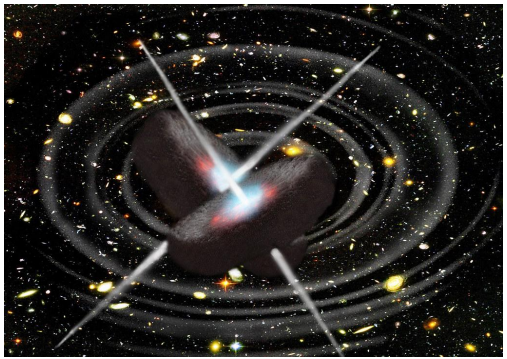
Gangmin Park

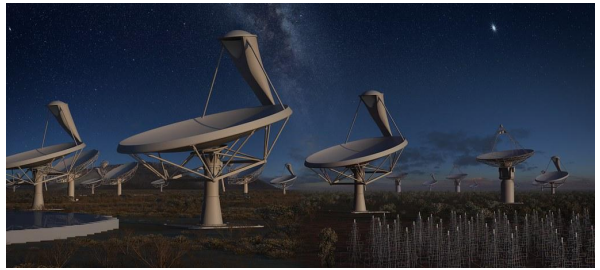
Cyril Briand



Contents

- Learning goals
- Progress made
- Gravitational waves
- Interferometry
- Jupyter notebooks
- Signal processing
- Analysis of data from LIGO
- Future work
- Personal reflections
- Advice to next group
- Questions



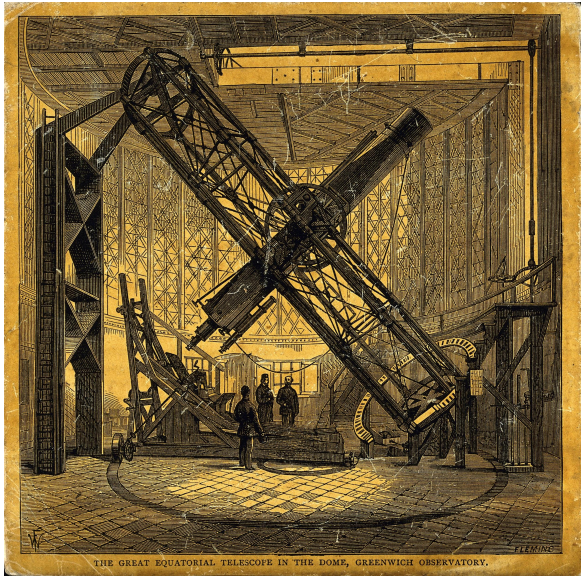


Learning Goals

- Learn about Scientific Software
- Introduction to Jupyter notebooks, git, github, latex
- FAIR principles for Scientific Software
- Practice working on specific software
- Improve team work, communication skills

Progress

- Gravitational Wave detection from LIGO data
- Produced Documentation for our work
- Individual reflections
- Advise for next summer school students
- Future work



Gravitational waves

What are gravitational waves?

Gravitational waves are disturbances in the curvature of spacetime.

They are created by accelerated mass.

They distort spacetime by squeezing and stretching it. For example two black holes orbiting each other would produce gravitational waves.

Gravitational waves can penetrate medium that are opaque to electromagnetic waves.

Gravitational waves observatories

How do we observe them?

Ligo and Virgo

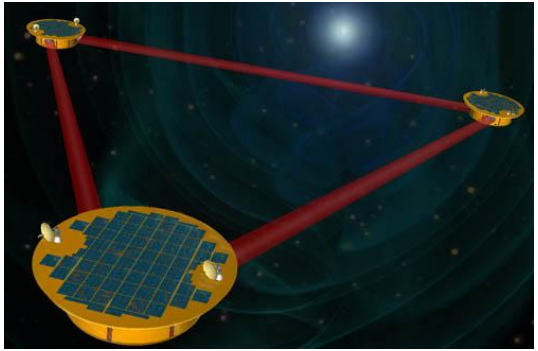
- Laser Interferometer Gravitational-Wave Observatory

KAGRA

- Large Scale Cryogenic Gravitational Wave Telescope

LISA

- Laser Interferometer Space Antenna

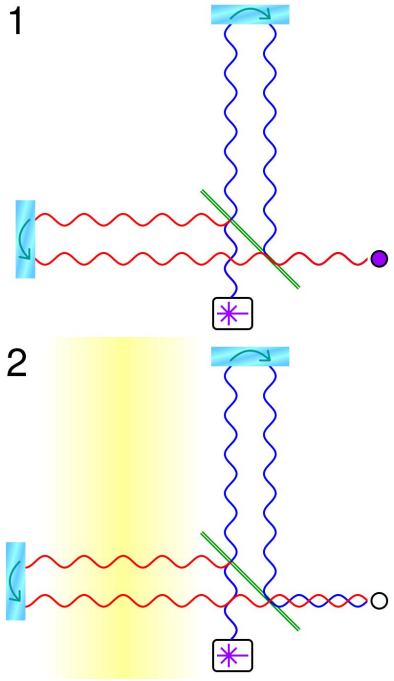


Interferometry

What is interferometry

Interferometry is a techniques that uses superposition of electromagnetic waves.

Interferometry for gravitational observatories:
Laser are calibrated so that they perfectly merge. As the arms of the observatories stretch and contract, tiny modifications appears on the lasers phases.



jupyter notebook

What is jupyter notebook?

An *open-source web application* that allows you to *create and share documents* that contain live code, equations, visualizations and narrative text.

- data cleaning and transformation
- numerical simulation
- statistical modeling
- data visualization
- machine learning

jupyter notebook

Characteristics

Language of choices

The Notebook has support for over 40 programming languages, including Python, R, Julia, and Scala.

Share Notebooks

Notebooks can be shared with others using email, Dropbox, GitHub and the Jupyter Notebook Viewer.

Interactive outputs

Your code can produce rich, interactive output: HTML, images, videos, LaTeX, and custom MIME types.

Big data integration

Leverage big data tools, such as Apache Spark, from Python, R and Scala. Explore that same data with pandas, scikit-learn, ggplot2, TensorFlow.

jupyter notebook

Components

iPython

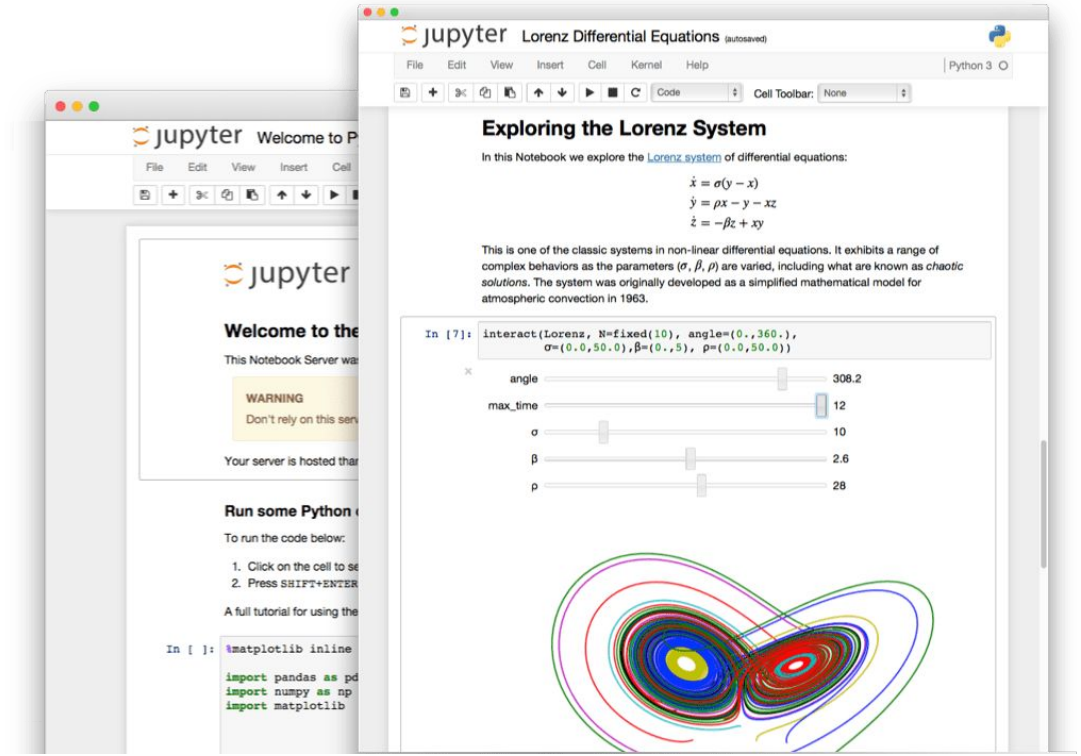
Interactive computational environment, in which you can combine code execution, rich text, mathematics, plots and rich media.

- A powerful interactive shell.
- A kernel for Jupyter.
- Support for interactive data visualization and use of GUI toolkits.
- Flexible, embeddable interpreters to load into your own projects.
- Easy to use, high performance tools for parallel computing.

jupyter notebook

Usage

Scientists are able to work on the Lorenz System which is based on the physics to explain a simplified mathematical model for atmospheric convection.

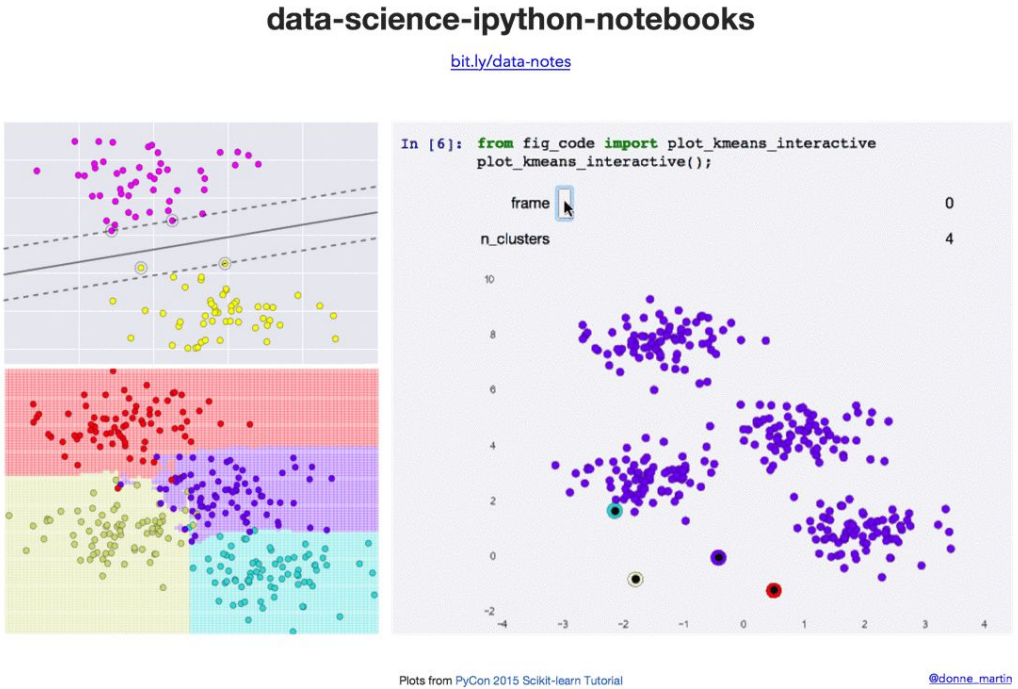


jupyter notebook

Usage

Data Scientist handles the data to make clusters based on the raw data.

Easy to test code and share the result with others.

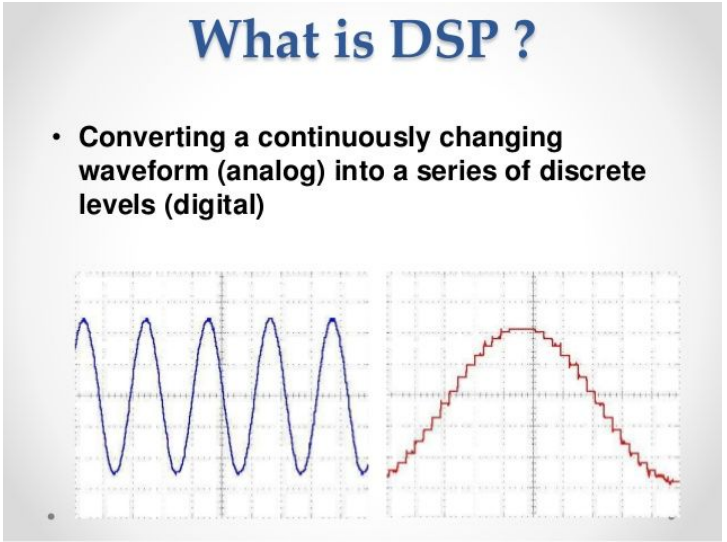


Digital Signal Processing

What is it?

The process of *analyzing and modifying a signal* to *optimize* or *improve* its *efficiency or performance*. It involves applying various mathematical and computational algorithms to analog and digital signals to produce a signal that's of *higher quality* than the original signal.

Signals need to be processed so that the information that they contain can be *displayed*, *analyzed*, or *converted* to another type of signal that may be of use.

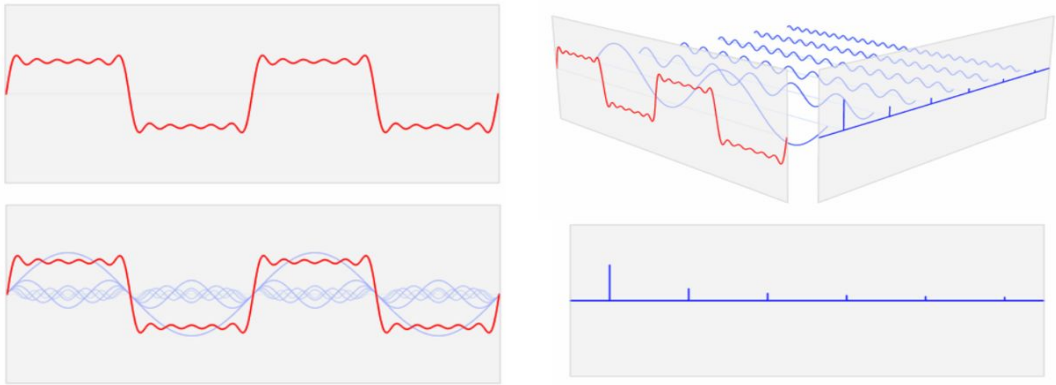


Fourier Transform

What Fourier Transform is?

A mathematical construct connecting *time* and *frequency domains*. The basic idea is that complicated but periodic functions (signals) can be written as the ***sum of simple waves***, mathematically represented by ***sines and cosines*** of ***different frequencies and amplitudes.***

The Fourier transform tells us which wave (frequency) contributes the most in defining the signal.

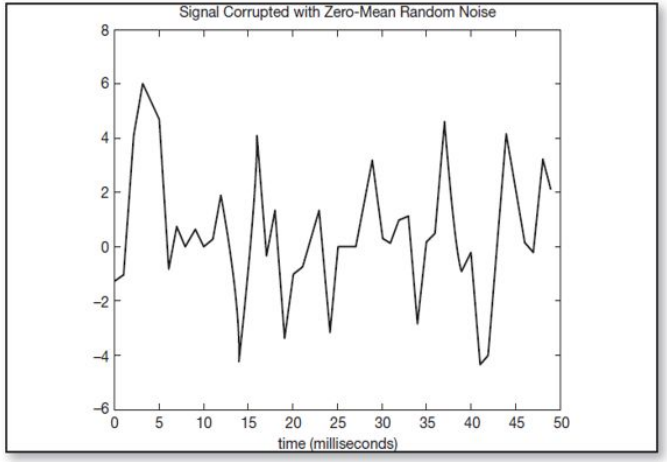


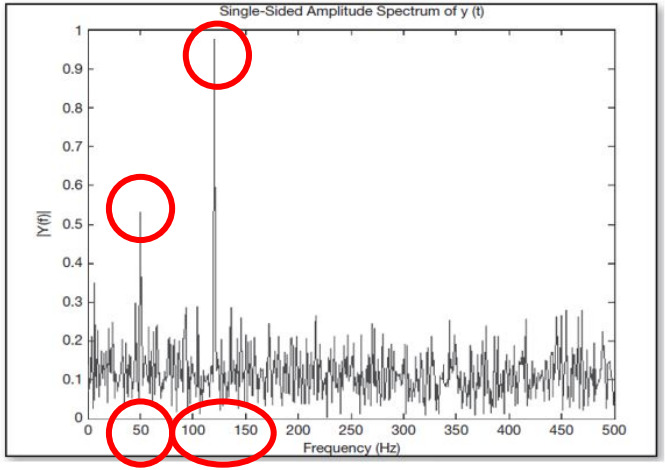
Fourier Transform

What Fourier Transform is?

The Fourier transform calculates **the correlation** between the **signal** and the **sine/cosine waves**.

Two clear spikes can be seen in the frequency domain in correspondence of **50** and **120 Hz**.





Limit

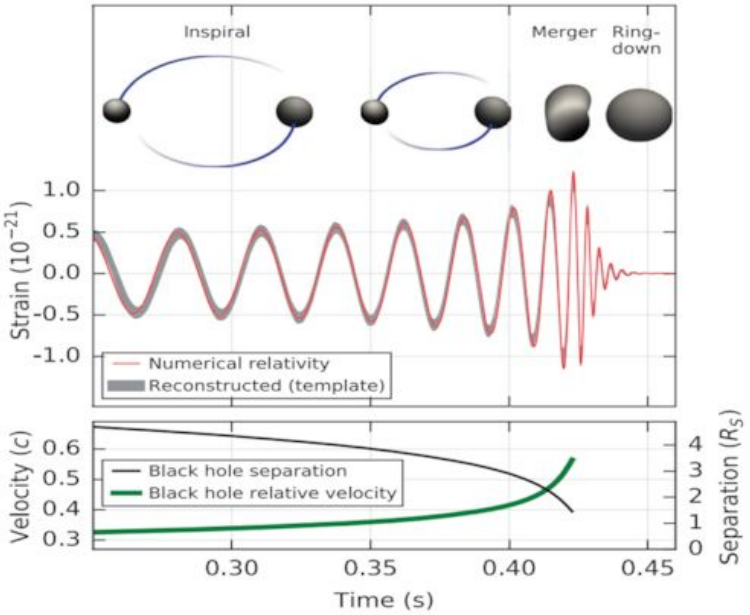
It cannot be used with market data, which are naturally discrete-time.

Solution

- DFT(Discrete Fourier Transform) => Slow
- FFT(Fast Fourier Transform) => Super fast

BINARY BLACK HOLE SIGNALS IN LIGO OPEN DATA

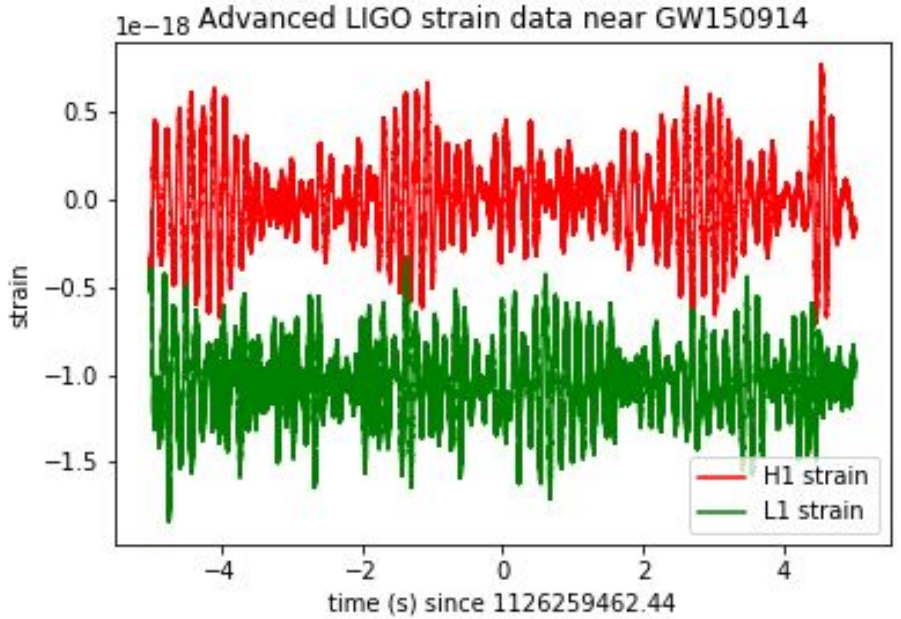
Causes of GW = merge of Binary Black Hole, Binary Neutron Star and something..



<PROCESS of tutorial>

Frequency

Raw data →



Specific range of frequencies -> information from BBH events is in advance.

Dominated by low frequency noise; there is no way to see a signal here, without some signal processing.

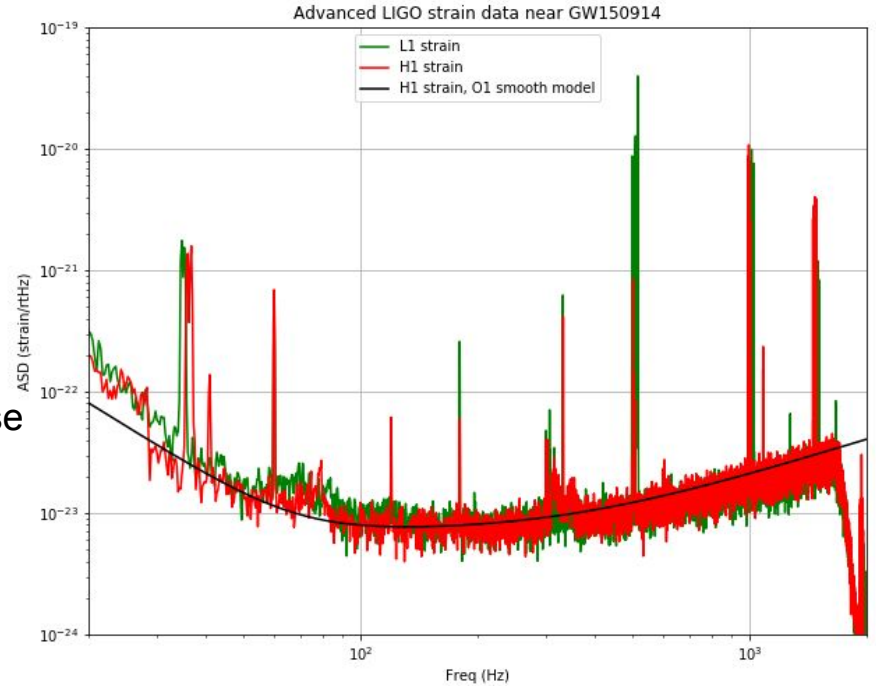
ASD (Amplitude Spectral Density)

Idea of the frequency content of the data.

The ASD = power spectral densities(PSDs)^2
The PSD = averages of (FFTs)^2.

Frequency range :
f_min = 20 Hz and f_max = 2000 Hz.
Below f_min -> not properly calibrated, high noise

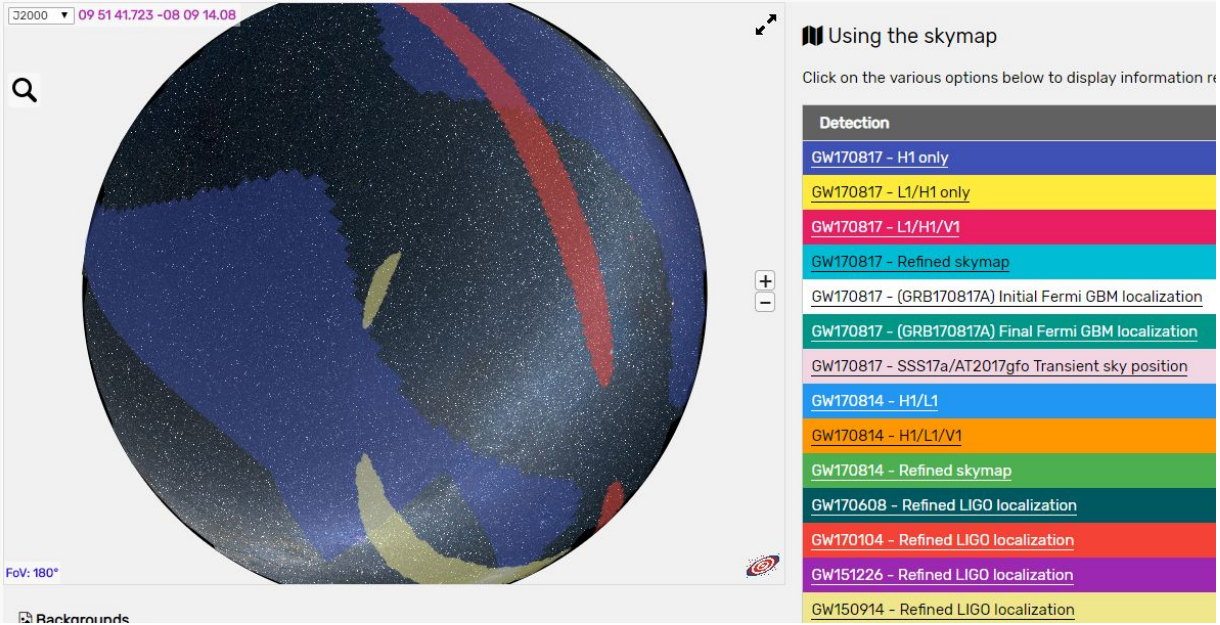
We can’t see the signal in this plot
because this plot is entirely dominated
by instrumental noise.



BNS and BBH range

BNS used to quantify the performance of a detector, corresponding roughly to the average distance at which we can detect a binary neutron star merger

BBH range >> BNS range, Systems with higher masses are much ‘louder’



Whitening

‘Normalization’

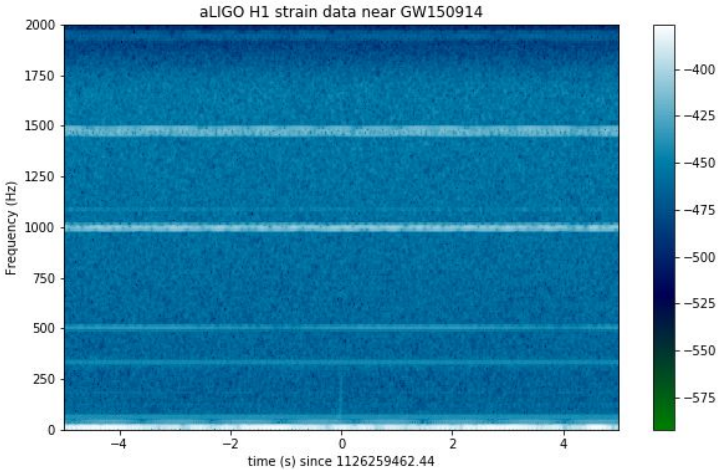
Noise fluctuations are much larger at low and high frequencies and near spectral lines.

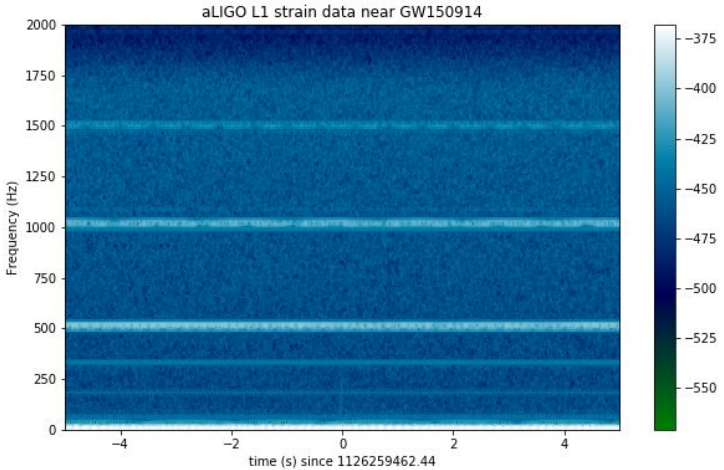
Whitening do not require any prior knowledge of spectral lines; only the data are needed.

Whitening of a signal means, smoothing down all the frequency contents of the signal to same intensity.

→ **Next process will be done by whitened data**

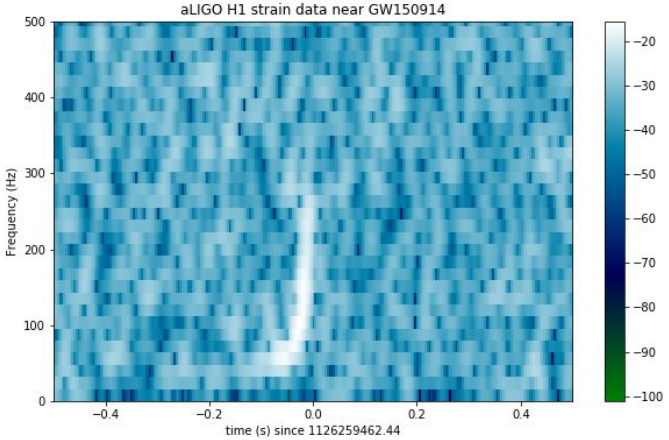
Spectrogram

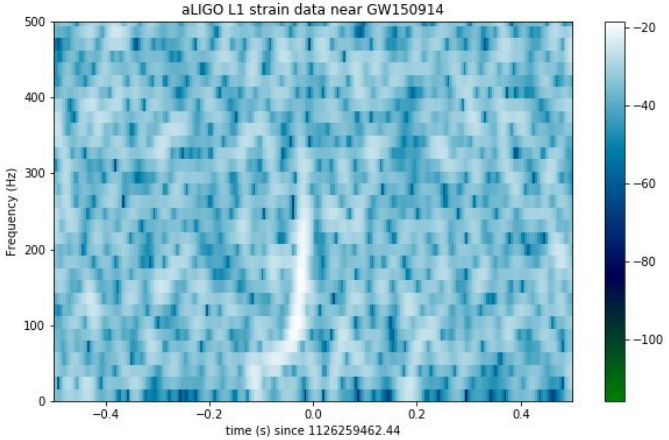




The lines at multiples of 500 Hz are the harmonics of the "violin modes" of the fibers holding up the mirrors of the Advanced LIGO interferometers.

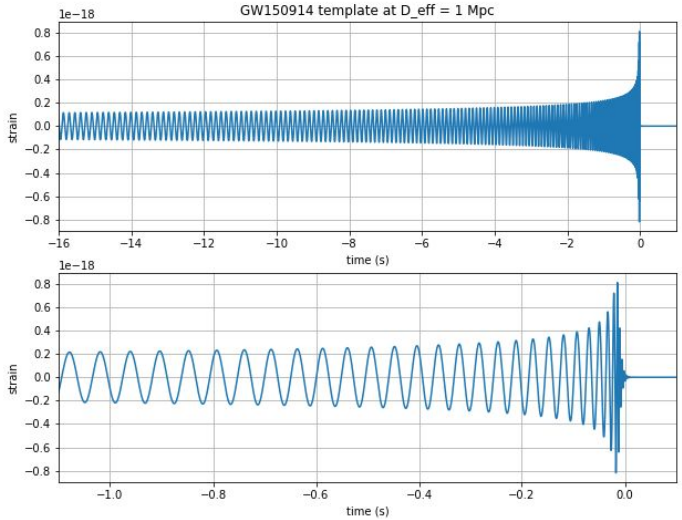
Spectrogram - Zoom in

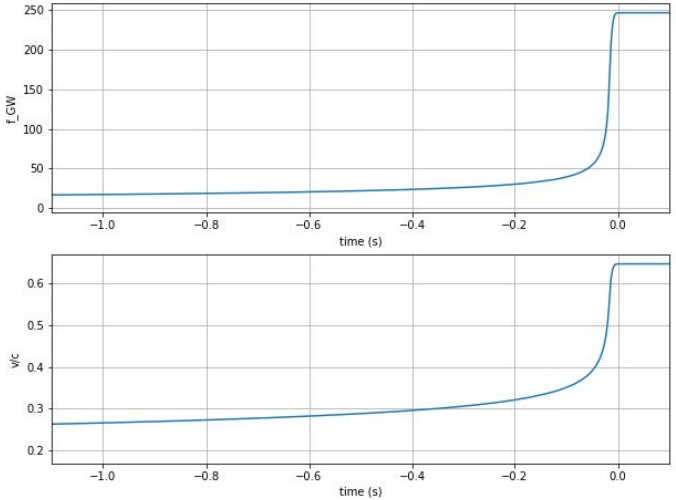




Loud (high SNR) signals
Compact object mergers show a characteristic "chirp" as the signal rises in frequency.

The result of LIGO analysis of BBH using parameters with Waveform Template



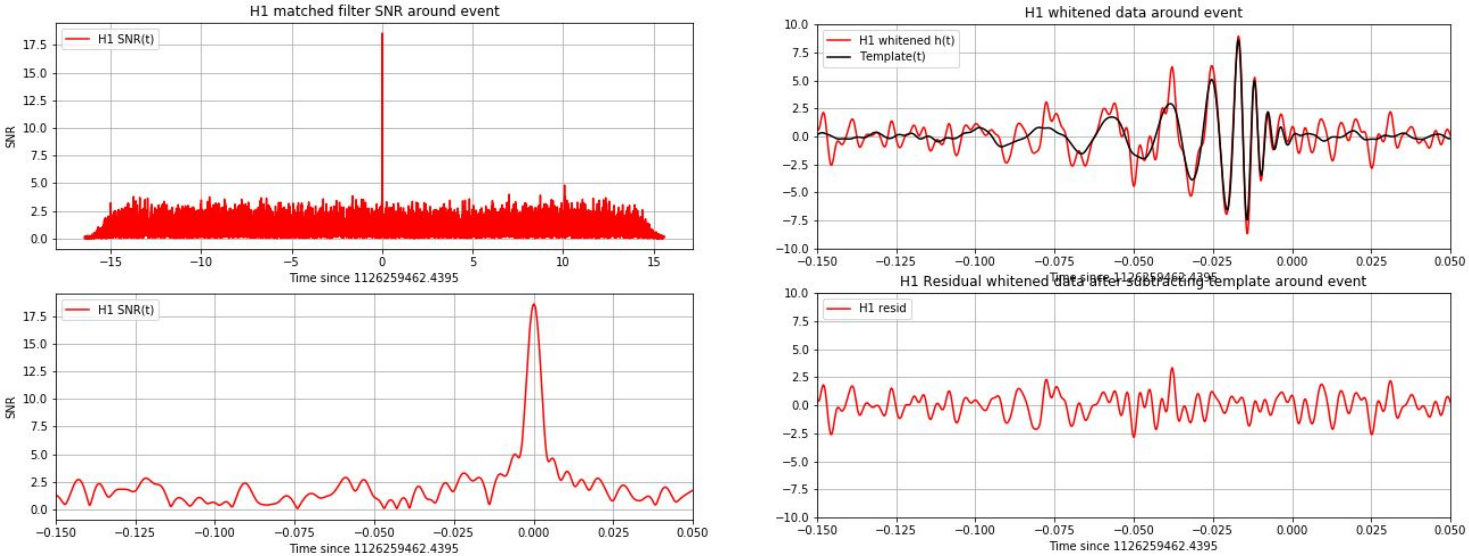


‘Prediction’

Analysis require a lot of computation by supercomputer

In the tutorial, use only one specific sample template

Matched filtering to find the signal (avoid the noise and match with the template)



High SNR means high power of signal

Using the template, do the matched filtering (extract the noise to clear signal)

Make sound files (visualization)

Frequency-shifted template for easy to hear

Frequency-shifted H1 for easy to hear

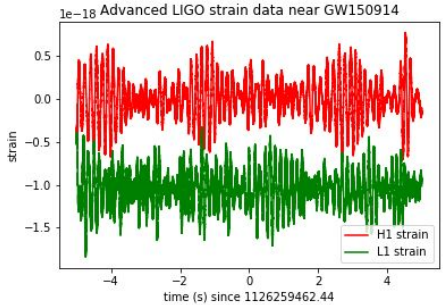
Frequency-shifted L1 for easy to hear

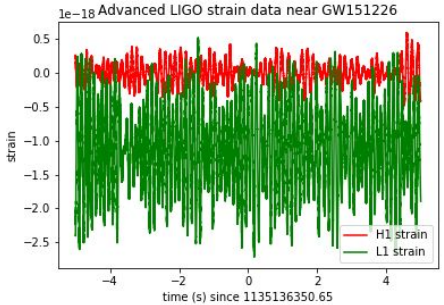
Finally we got the signal!

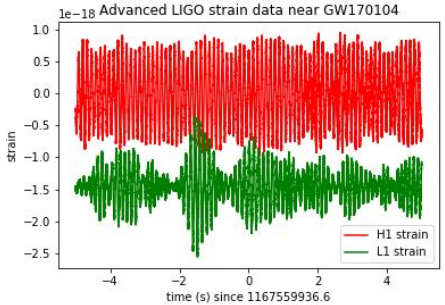
Sample from the GW150914

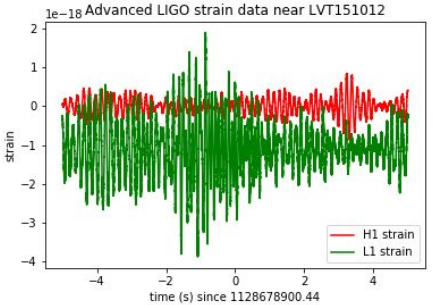
Comparing with 4 samples

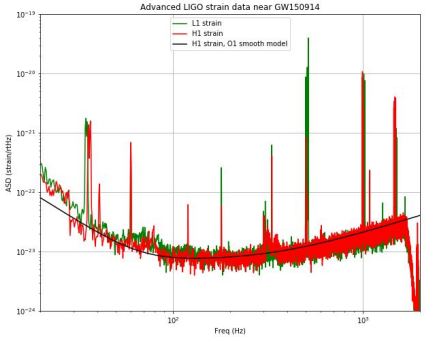
LVT(LIGO - Virgo Trigger) means the signal with insufficient confidence.

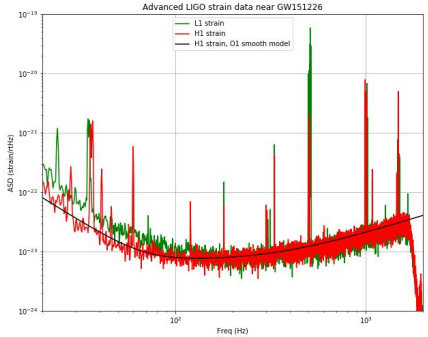


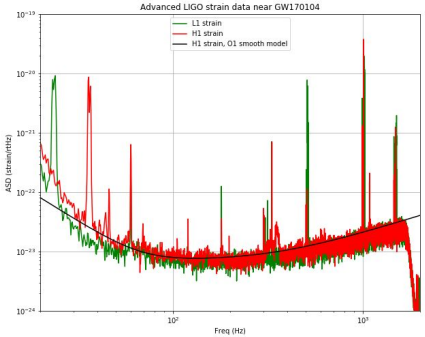


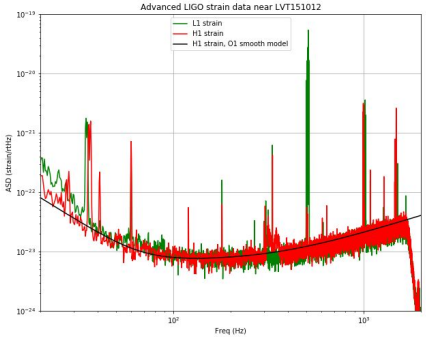




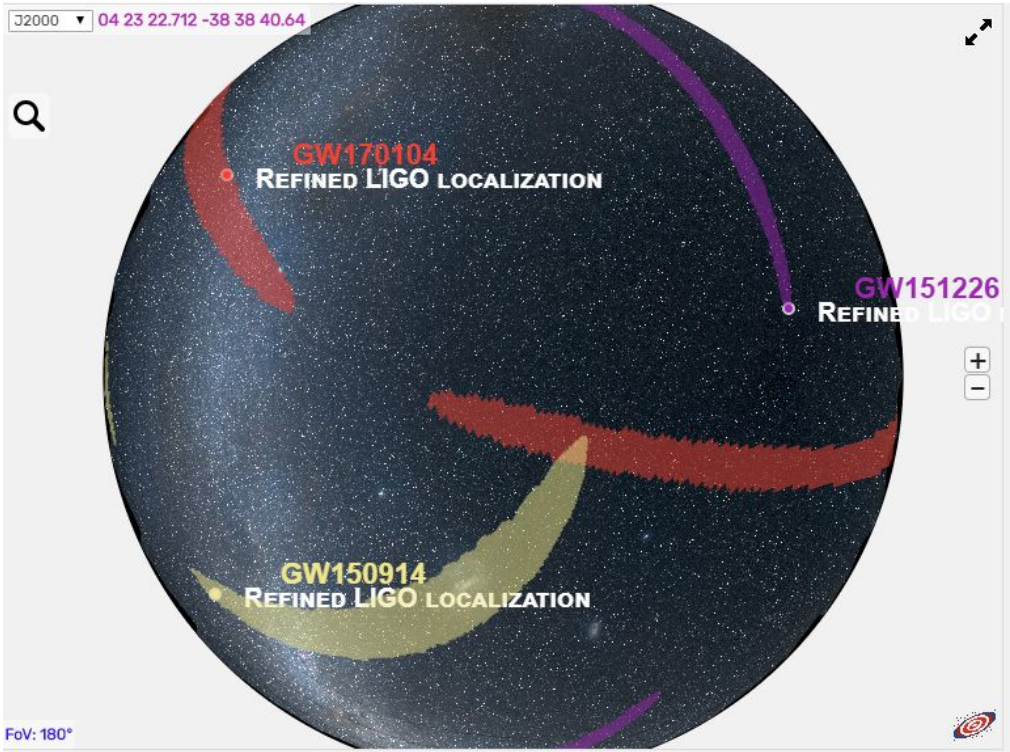


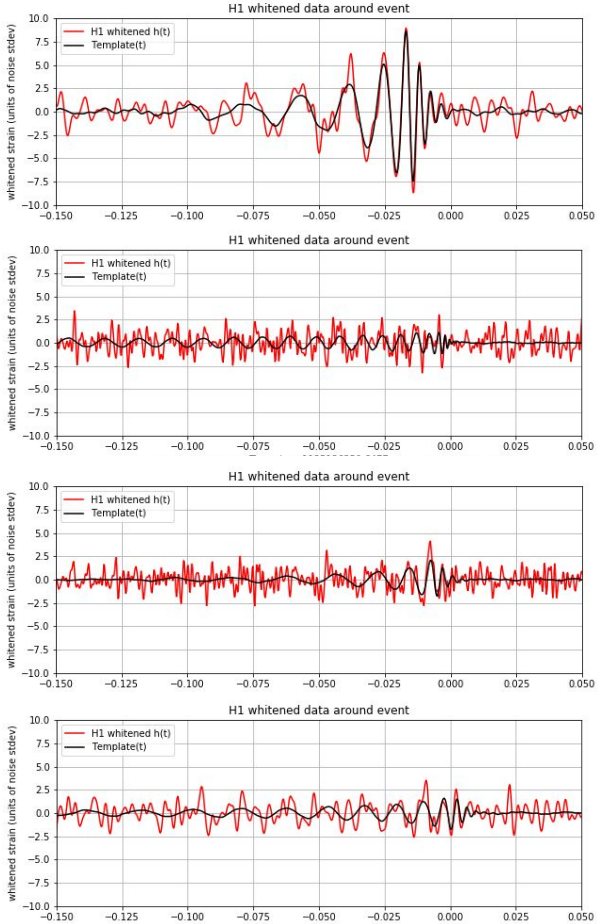




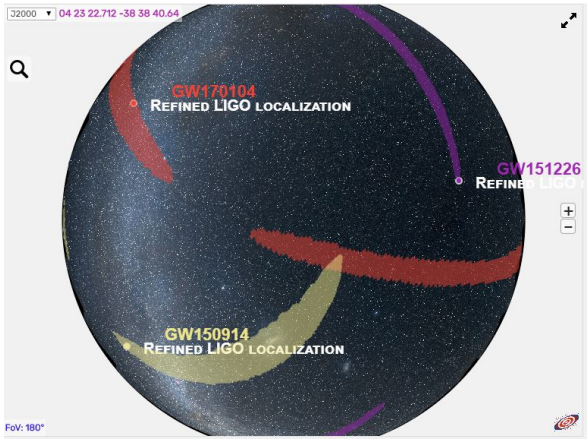


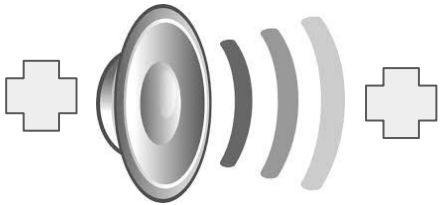
Comparing with 4 samples

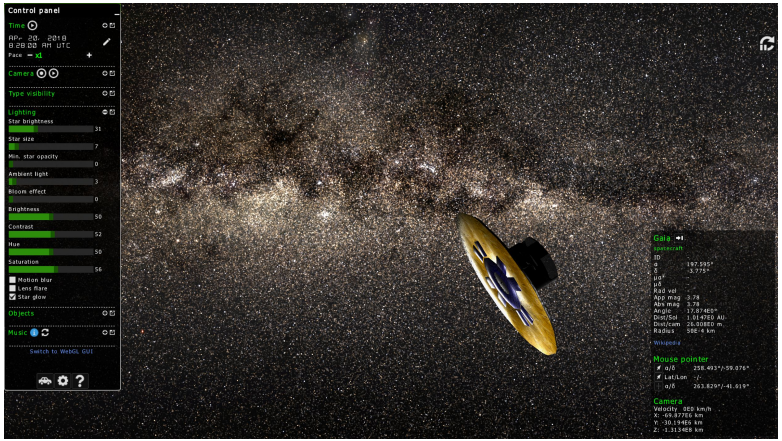


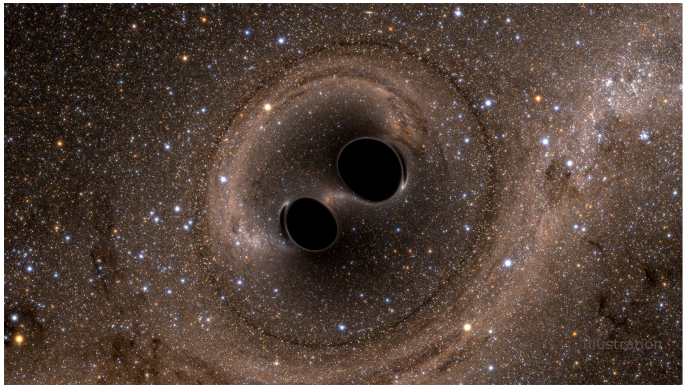


Future plan









make simulation of BBH and BNS!

Personal Reflection
Gangmin

- Fourier transforms can be used to analyze the frequency components of a signal.
- Through matched filtering, hidden signals can be found in the frequency.
- The importance of visualizing data.
- Process of LIGO and astronomy

Personal Reflection
ilgwon

- Apply Physics to the software
- Learning basic knowledge about Fourier transform
- Great opportunity to participate astronomy project

Personal Reflection
Cyril

- Use programing language to visualise physical phenomenon with Jupyter Notebook
- Signal whitening techniques
- Knowledge on gravitational waves

Personal Reflection
Umar

- Software and computer systems are crucial for large scientific projects
- Data analysis is also very important
- Domain knowledge helps immensely
- Learned about team work
- The importance of communication skills

Our advice to the next students

- ❏ Settle on a project and objectives
- ❏ Start early with the project
- ❏ Document everything (installation too)
- ❏ First focus on domain knowledge
- ❏ Use FAIR principles in your code

Thank you

Questions?