FAIR Public Data: The Last Mile

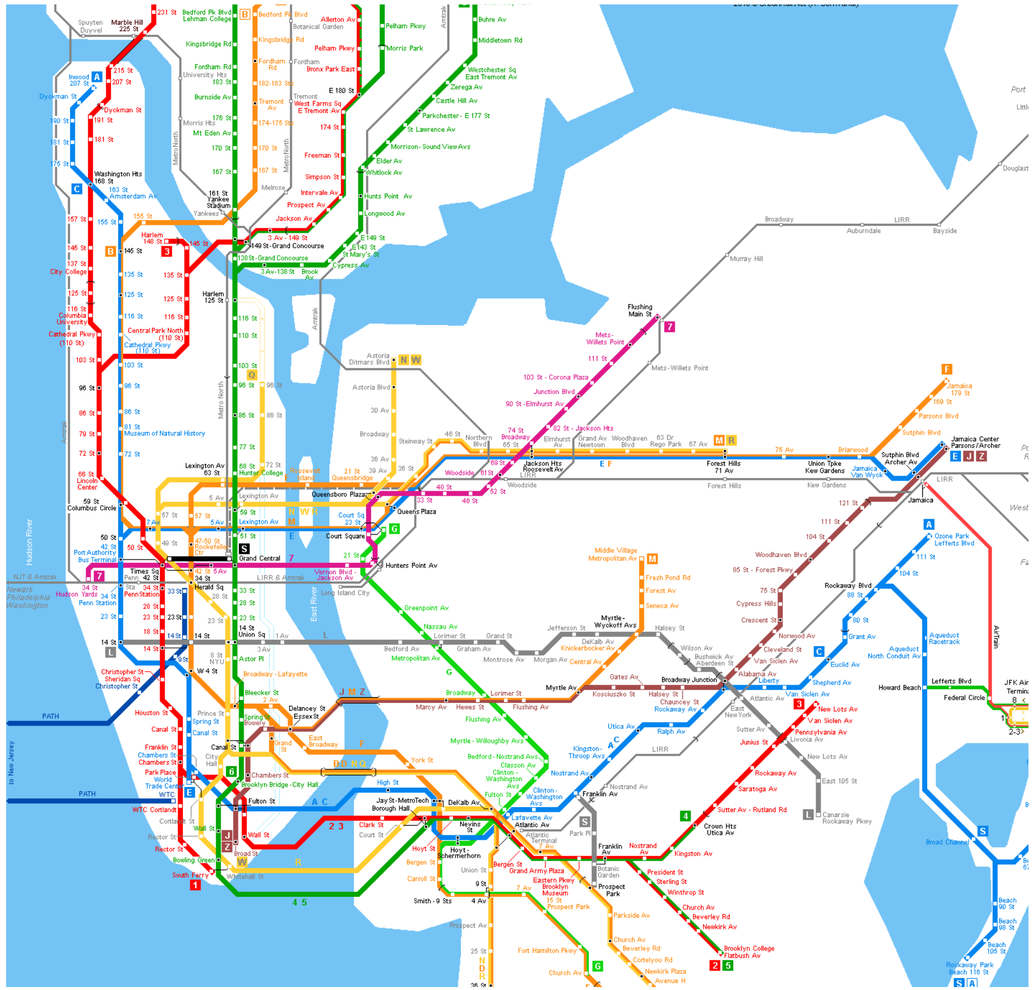
Jonah Kanner
LIGO Lab, Caltech

January 17, 2024 - G2302406-v2

The Last Mile Problem

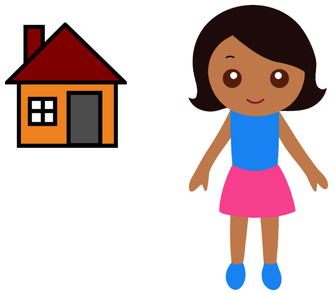
Transportation

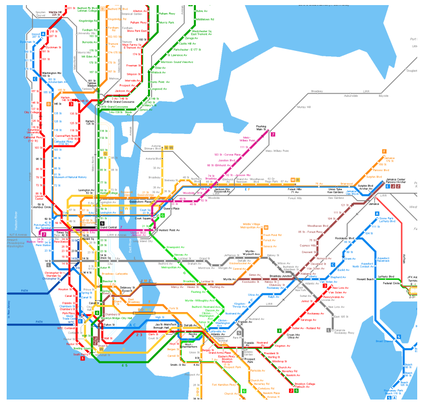
Last Mile



The Last Mile Problem

Public Data





Expert Networks

"Last Mile"

High School Student
Undergraduates
Grad Students
Experts in other field
Amateurs
Artists

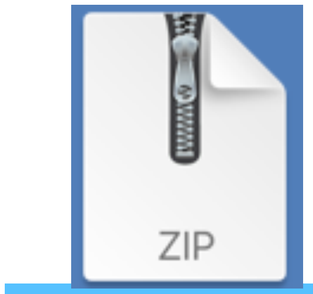
Gaps in:
Access
Knowledge
Resources

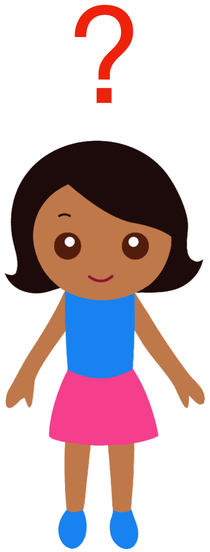
Data
Software
Journal Articles
Conferences
Colleagues


Arroyo Pkwy
730
8KZN054

Sure, our data are public … but:

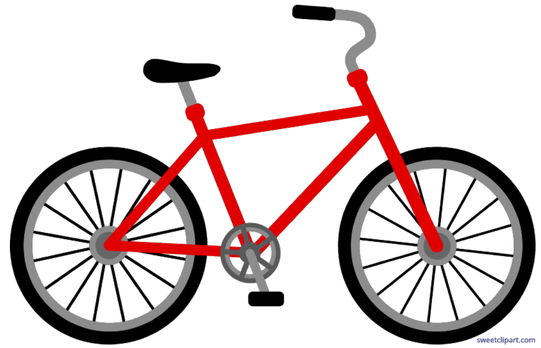
- Are the data easy to find and download?
- Do I recognize the file format? Can I figure out how to open it?
- Can I load the data in a spreadsheet or text file?
- Are there “secret steps” to processing the data?
- Can I find the software? Can I get it installed on my computer?
- Once the software is installed, can I figure out how to use it?
- Do I know where to ask for help when I get stuck?





Solutions for LIGO data

- Data in multiple formats (GWF and HDF5 and “streaming”)
- Software examples to show people exactly how to get started
  - Focus on basic tasks: loading, pre-processing, and plotting
- Use online tools and services, so no software installation is needed
  - (Google co-lab , mybinder , streamlit)
- Link to resources: software libraries, related data, papers, tools, web services
- Workshops and online courses
- Help Desk and Discussion Forum
- Integrated platform: gwosc.org



Everything at https://gwosc.org

Software Examples In Your Browser

Jupyter Notebooks

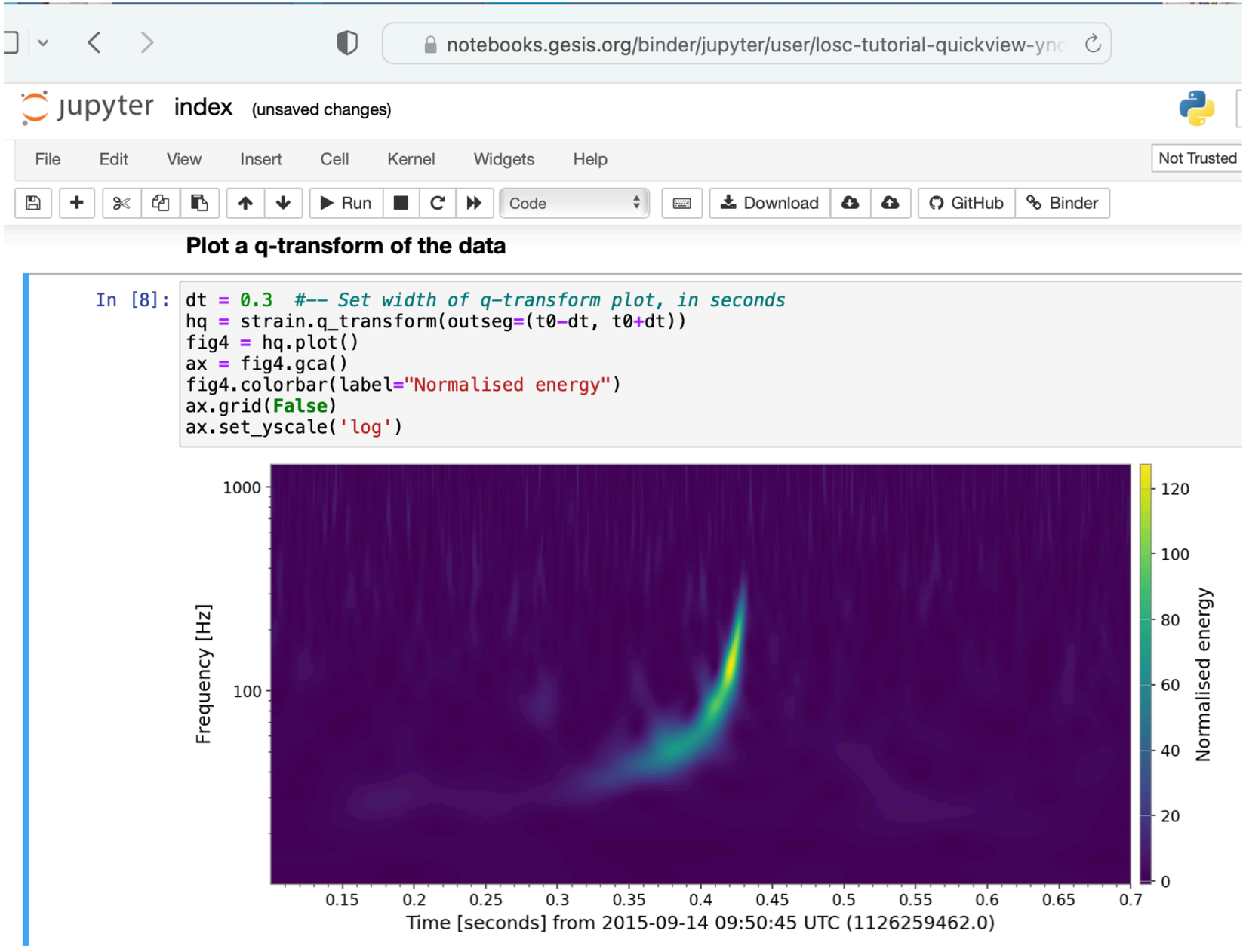
google co-lab
mybinder

Specialized libraries:

`GWpy, pyCBC, bilby`

No installation

gwosc.org/tutorials

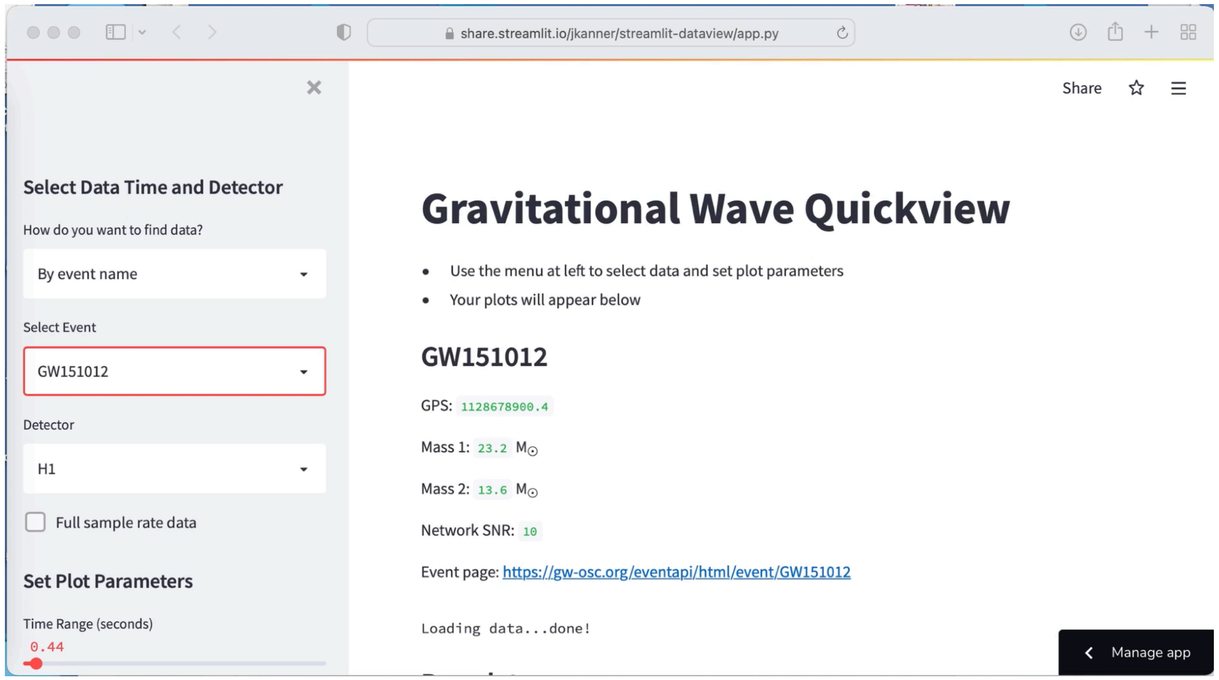


Web Apps or GUIs

Remove the need to program!

https://gwosc.org/path

- Plot data with no programming
- “Pre-process” data (whiten, filter, etc.)
- Export common file types (e.g. CSV)
- Introduction to signal processing



- **Common Request:**
  “I’d like to download processed data to in a CSV or text file”

Open Data Workshops

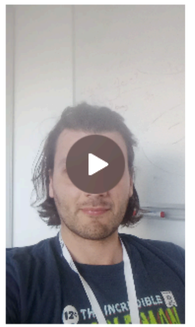
- Annual Event
- Junior scientists prepare material, lecture, and mentor
- Includes “hands on” software examples + challenge problems
- Recently: Hybrid and Scalable
- Online course
  - Enroll at any time

2023 Open Data Workshop
2000+ Participants
15 Locations + Virtual

Shreejit Jadhav
PhD Student
Inter-University Centre for Astronomy and Astrophysics (IUCAA), Pune, India

Leïla Haegel
Researcher
Astroparticles and Cosmology Laboratory, France

Simone Mastrogiovanni
Postdoc
ARTEMIS, Nice Observatory, France



Summary

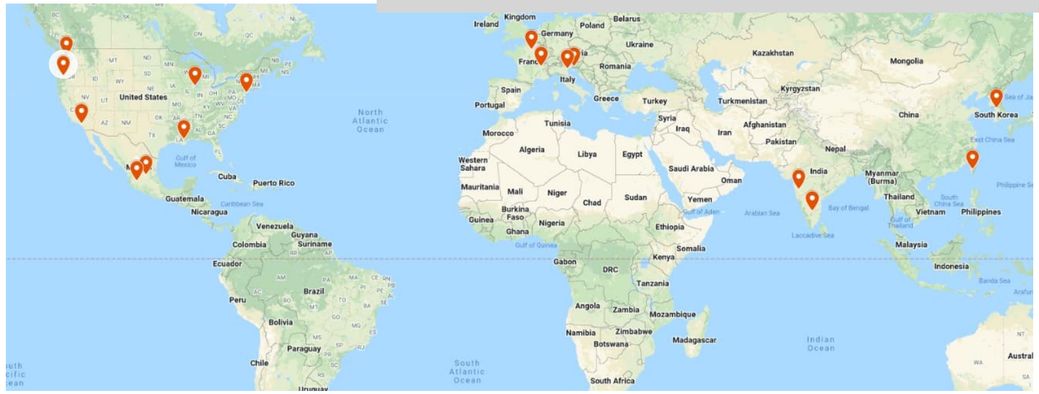
- Focus on the user experience
- The most basic tasks are always the most important
- Create a pathway for new researchers to follow
  - Provide support, and use the feedback to make improvements

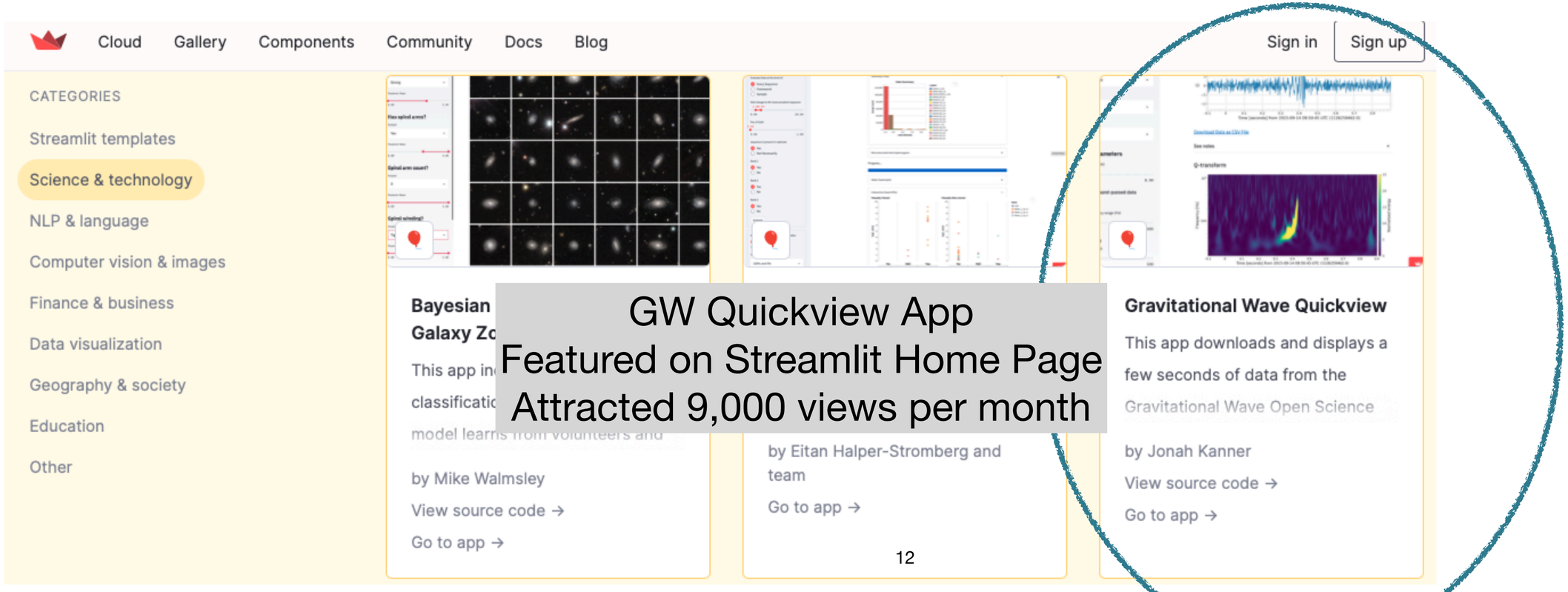
Thank you!

Supporting the Community

2022 Open Data Workshop
1000+ Participants
15 Locations + Virtual

Discussion forum: https://ask.igwn.org
E-mail help desk: gwosc@igwn.org
Online Course: https://gw-odw.thinkific.com
Web apps: https://gwosc.org/path
Tutorials & Workshops: https://gwosc.org/tutorials





GW Quickview App
Featured on Streamlit Home Page
Attracted 9,000 views per month


Pasadena, CA


GW Open Data Workshop
Study Hub
LIGO VIRGO KAGRA
San Luis Potosí, Mexico

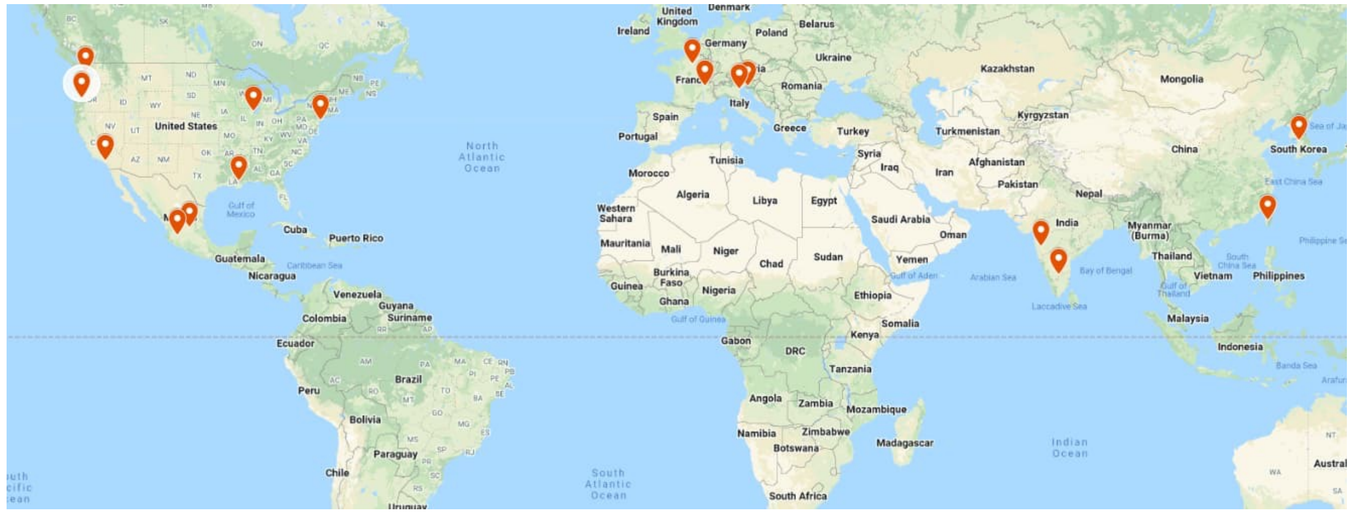

Trieste, Italy


Maria Paola Vaccaro
Kumar
Ankur C
Arnab Lahiry
Simon Prunet
European Virtual Hub
Simone Mastrogiovanni [Mentor]


Guadalajara, Mexico


LIGO Livingston


New York


United Kingdom
Ireland
Denmark
Belarus
Poland
Germany
Ukraine
Kazakhstan
Mongolia
Romania
Italy
Spain
Portugal
Greece
Turkey
Kyrgyzstan
Turkmenistan
China
South Korea
Tunisia
Syria
Iraq
Iran
Afghanistan
Pakistan
Nepal
Morocco
Algeria
Libya
Egypt
Western Sahara
Saudi Arabia
India
Myanmar (Burma)
Thailand
Vietnam
Philippines
Mauritania
Mali
Niger
Chad
Sudan
Oman
Yemen
Burkina Faso
Guinea
Nigeria
Ethiopia
Ghana
Somalia
Kenya
Gabon
DRC
Tanzania
Malaysia
Indonesia
Angola
Zambia
Mozambique
Namibia
Zimbabwe
Botswana
Madagascar
South Africa
United States
Cuba
Puerto Rico
Guatemala
Nicaragua
Venezuela
Guyana
Suriname
Colombia
Ecuador
Brazil
Peru
Bolivia
Paraguay
Chile
North Atlantic Ocean
South Atlantic Ocean
Indian Ocean


Paris


Evanston, IL

Lyon, France


Seoul, Korea


Padova, Italy


ICTS, Bangalore, India

Getting Help

Need to hear from people using data

- GWOSC Help Desk, via e-mail: gwosc@igwn.org
- New: LIGO/Virgo/KAGRA discussion forum: https://ask.igwn.org
  - Vera Rubin Telescope has an active discussion forum, with thousands of posts
- Discussion Board / Help Desk monitored both by GWOSC staff and volunteers in LIGO/Virgo/KAGRA collaboration

Provide direct support AND Collect user feedback

ask.igwn.org

Collapse

Welcome to the gravitational wave community forum

A community for discussion of gravitational wave science with LIGO, Virgo, and KAGRA.

Gravitational Wave Science

Post questions and announcements related to gravitational wave science, education, and careers.

Help with Data Analysis

Post questions and tips for finding, downloading, and analyzing gravitational wave data in the Data Analysis category.

Learn about this forum

These links describe usage guidelines for this forum and overviews of key observatories.

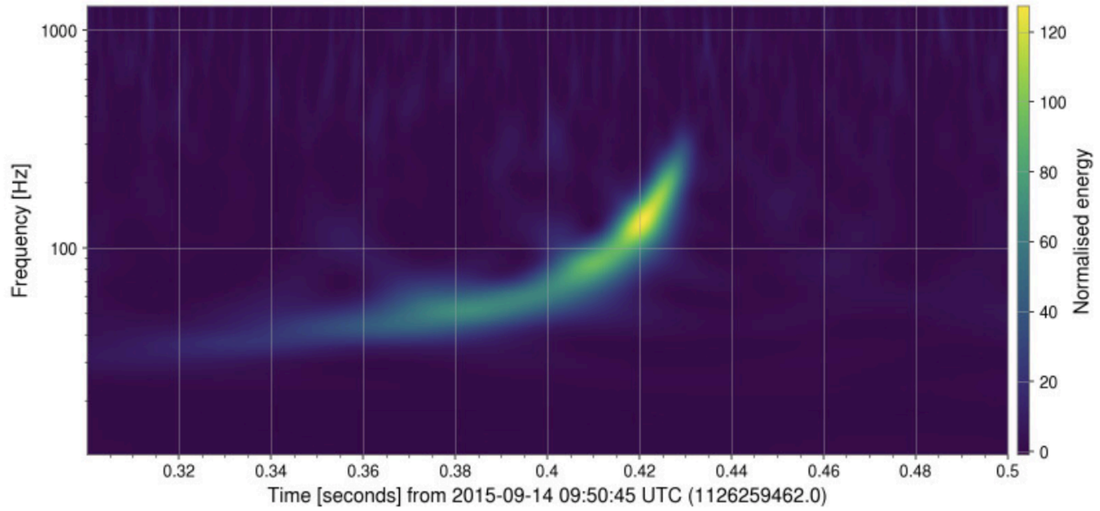
- Community guidelines
- LIGO Laboratory and the LSC
- Virgo Observatory
- KAGRA Observatory

A

anshul21 7d

Hi

I am trying to get the frequency information from a q-transform plot shown below.



I am using the following code snippet to get that information, but I see that the following code prints the frequency values for the complete q-transform plot.

Jul 17

1 / 3
Jul 17

6d ago

Impacts of Open Data

Around 6,000 visitors (12,000 sessions)
to GWOSC each month

Over a million strain file downloads over 6 months

250 Papers in 2 years (2020 + 2021)

Open Data Workshops with hundreds of participants