

Introduction to open and reproducible workflows in the public service

Workshop for EW2, February 12, 2021

LIVE POLL TIME

About us

Monica - Dr. Monica Granados is a Policy Advisor in the Communication and Open Science Section in the Science and Technology Branch. She has a PhD in Ecology from McGill University where she discovered her passion to make science more accessible and transparent. She is also a recent graduate of the Frictionless Fellows Programme.

Noushin - Dr. Noushin Nabavi is an Economist at the BC Ministry of Health in the Pharmaceutical Analytics Unit. She earned her PhD in Cell and Systems Biology from the University of Toronto and has worked on research and development projects related to human health and disease prior to she joined the public service.

Learning goals

— — —

1. Define reproducible research and open workflows
2. Discuss current issues surrounding reproducibility
3. Discuss solutions and important components of reproducibility
4. Tools used for reproducible and open research
5. Reproducible workflow demo

The problem

- Government is closed
- The public doesn't know how decisions are made
- Duplication of work within the government
- Types of reproducibility

LIVE POLL TIME

Open and reproducible workflows as a solution

Claerbout's Principle (1980s):

The scholarship does not only consist of theorems and proofs but also (and perhaps even more important) of data, computer code and a runtime environment which provides readers with the possibility to reproduce all tables and figures in an article.



Image here and next two slides were created by Scriberia for The Turing Way community and is used under a CC-BY licence.

Open and reproducible workflows as a solution in government

- Openness and transparency in government
- Evidence based decision making
- Efficiency
- Enable others to use our methods or help improve ours

Barriers to open and reproducible workflows

- Directives
- Approval of management
- Data privacy

Tools

Data

Methods

Results

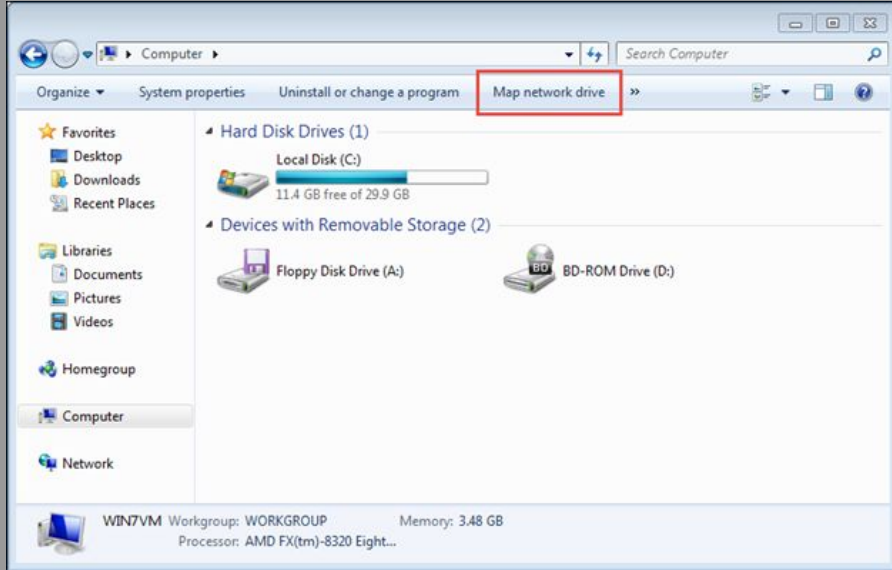
Products

Data

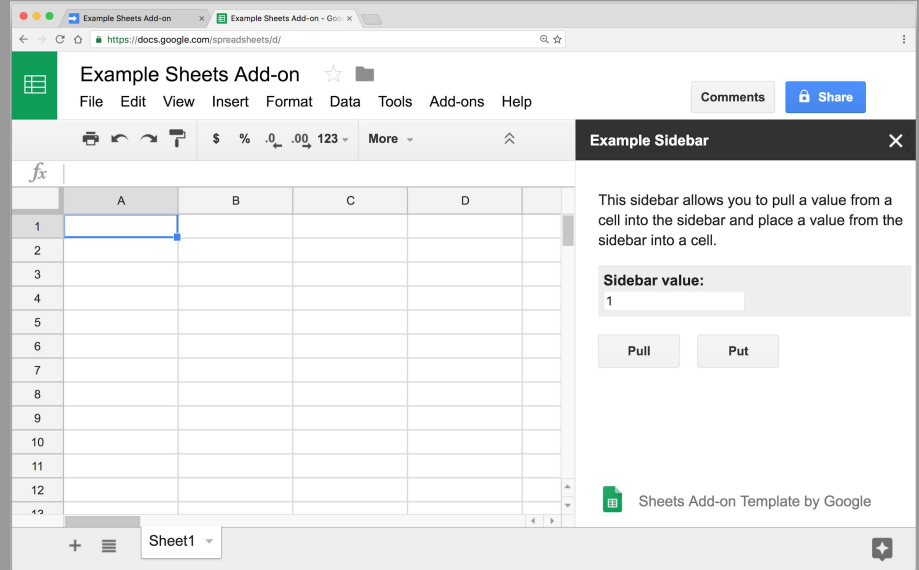
How can you make the data findable, accessible, interoperable, and reusable?

- Shared drive/LAN/Sharepoint
- Google sheets / OpenRefine
- Open government portal
- GitHub

Examples



Local area network



Google sheets

[Pull requests](#)[Issues](#)[Marketplace](#)[Explore](#)[CSSEGISandData / COVID-19](#)[Watch](#)

930

[★ Unstar](#)

25.6k

[Fork](#)

16.4k

[Code](#)[Issues](#) 1.4k[Pull requests](#) 281[Actions](#)[Projects](#)[Wiki](#)[Security](#)[Insights](#)[master](#)[COVID-19 / csse_covid19_data / csse_covid19_time_series /](#)[Go to file](#)[Add file](#)[...](#)

CSSEGISandData Automated update for delayed data for US

434fcb8 12 hours ago

[History](#)

..

[.gitignore](#)

update

12 months ago

[Errata.csv](#)

Patch Nova Scotia stale data from 20210124-20210202

4 days ago

[README.md](#)

Update README

9 months ago

[time_series_covid19_confirmed_US.csv](#)

Automated update for delayed data for US

12 hours ago

[time_series_covid19_confirmed_global.csv](#)

Automated update for delayed data for US

12 hours ago

[time_series_covid19_deaths_US.csv](#)

Automated update for delayed data for US

12 hours ago

[time_series_covid19_deaths_global.csv](#)

Automated update for delayed data for US

12 hours ago

[time_series_covid19_recovered_global.csv](#)

Automated update

15 hours ago

COVID-19 daily data updates on JH Github

Methods

- GCconnex
- Protocols.io
- GitHub

How can you make the process by which you do your work more open and reproducible? How could others access this work?

Examples

GCconnex

GCpediaGCintranetGCdirectoryGCcollab

Français

Search GCconnex

News FeedGroupsCareer▼The WireBlogsCommunities▼More▼

Policy

Community Newsfeed

This news feed shows content tagged with policypolicy developmentpolitiqueélaboration des politiques and was identified as useful for the following communities: Policy Specialists

New OECD report on "How do laws and regulations affect competitiveness"

French will follow Good afternoon, Last year the Government of Canada partnered with the OECD to help identify ways to incorporate economic and competitiveness considerations when designing and implementing...

Started by Simon Beaulieu 4 days ago

GCConnex

cols.io

Search

FeaturesPlansBlogSign

ProtocolStepsGuidelinesMaterialsForksMetadataMetrics

plates are **Lysis Buffer** according to the number of plates to be filled.

| | A | B | C |
|---|----------------------------|---------------|----------------|
| 1 | Reagent | 96-well plate | 384-well plate |
| 2 | NEB HF Phusion buffer (5x) | 1.1 µL | 4.4 µL |
| 3 | Proteinase K (20 mg/mL) | 27.5 µL | 110 µL |
| 4 | UltraPure Water | 411.4 µL | 1645.6 µL |
| 5 | Total | 440 µL | 1760 µL |

< >

Prepare 96/384 well plate(s) containing 4 µL **Lysis Buffer** per well.

Add 1 µL barcoded oligo-dT primer [2 µM] (E3V6NEXT adapter) to each well (12-/64-channel pipette).

Lysis plates with barcode primers can be prepared ahead of time and stored at -20 °C

Protocols.io

Results

1. Shared drive/LAN/Sharepoint
2. Google docs
3. Github

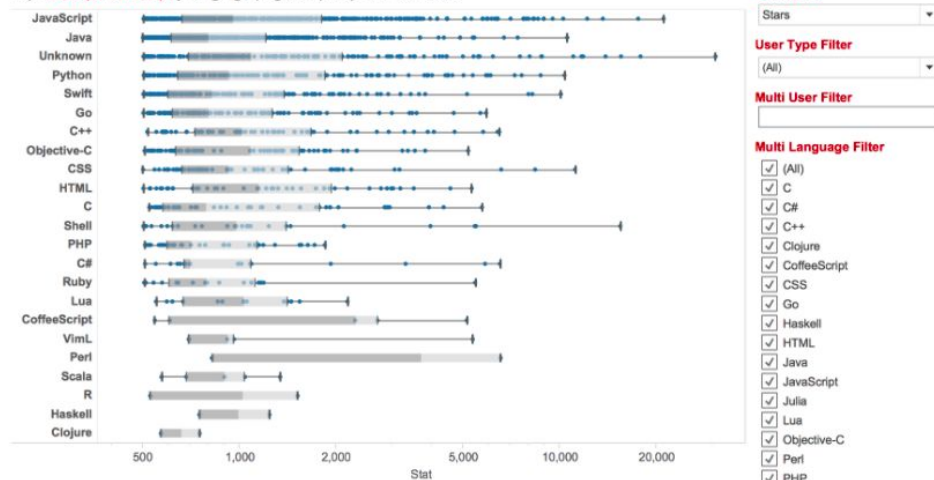
How can we adequately share the results with team members and tweak the methods to answer new questions/test new hypotheses?

Examples



Examples

Repo Stats (Select Stat) By Language (Log Scale): Repos >= 500 Stars



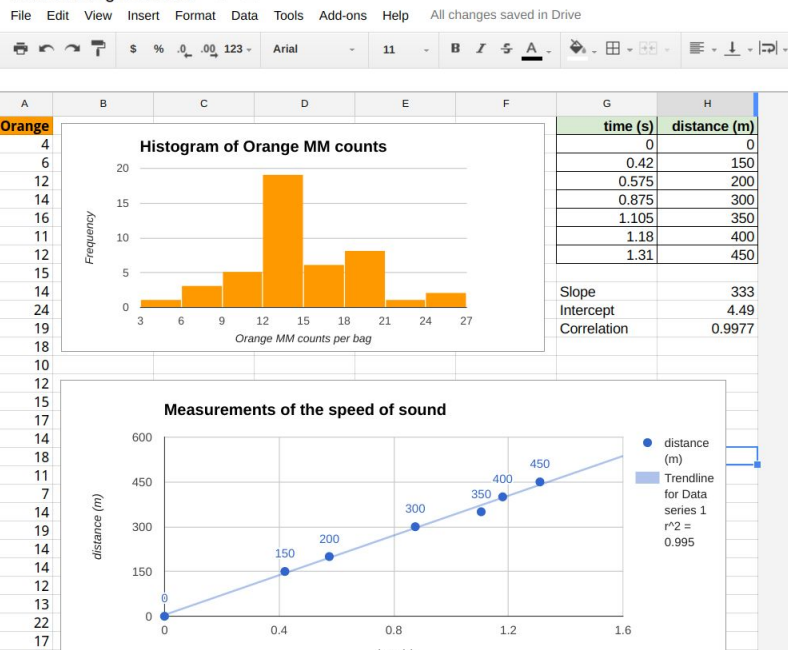
Viz Interactive Website

<http://www.donnamartin.com/viz>

Github

<https://github.com/donnamartin/viz>

Text 6: Google Sheets



Google sheets

Communications products

Once your work is done - how can you make it most accessible?

- Cochrane reports
- Open access publications
- Public talks
- Media
- Github

Giving and receiving credit

How do you make sure that others credit your work and you credit others?

- Code licences (MIT/Open Canada)
- Information licences (Creative Commons)
- Github



Creative Commons Licenses

[version française](#)

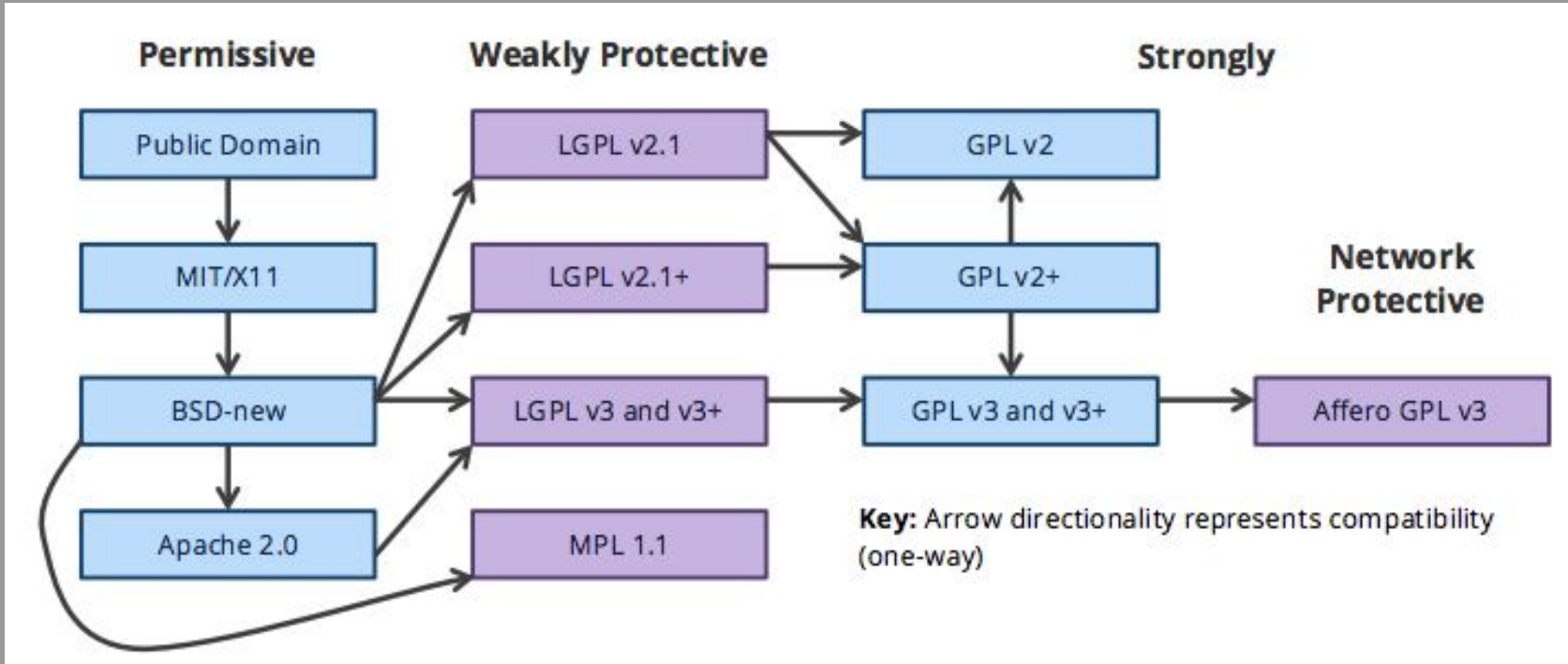
The possible combinations

\\ Choose your own license here: creativecommons.org/choose

| CC License | Button | Explanations | Sharing allowed ? | Exploitation allowed ? | Remix allowed ? |
|---------------------------------------|--------|---|-------------------|------------------------|-----------------|
| Attribution | | <ul style="list-style-type: none"> You must give CREDIT to the original work's creator. Commercial usage of the work is allowed. You can remix the work. Sharing is (always) allowed for commercial and non-commercial purpose. | | | |
| Share Alike | | <ul style="list-style-type: none"> You must give CREDIT to the original work's creator. Commercial usage of the work is allowed. You can remix the work. If you remix this work, you're forced to use exactly the same license for your derivative version of this work. Sharing is (always) allowed for commercial and non-commercial purpose. | | | |
| No Derivative | | <ul style="list-style-type: none"> You must give CREDIT to the original work's creator. Commercial usage of the work is allowed. You CANNOT remix the work. Sharing is (always) allowed for commercial and non-commercial purpose. | | | |
| Non-Commercial | | <ul style="list-style-type: none"> You must give CREDIT to the original work's creator. Commercial usage of the work is NOT allowed. You can remix the work. If you remix this work, you're forced to use exactly the same license for your derivative version of this work. Sharing is allowed for non-commercial purpose only. | | | |
| Non-Commercial + Share Alike | | <ul style="list-style-type: none"> You must give CREDIT to the original work's creator. Commercial usage of the work is NOT allowed. You can remix the work. If you remix this work, you're forced to use exactly the same license for your derivative version of this work. Sharing is allowed for non-commercial purpose only. | | | |
| Non-Commercial + No Derivative | | <ul style="list-style-type: none"> You must give CREDIT to the original work's creator. Commercial usage of the work is NOT allowed. You CANNOT remix the work. Sharing is allowed for non-commercial purpose only. | | | |

Chart embedded in the article [Creative Commons 101](#) by Guillaume Déziel is made available under the terms of the [Creative Commons \(CC BY-NC-SA\) license](#). The permissions not covered by this licence may be obtained at guilaudedeziel.com/contact.

Open source licences



LIVE POLL TIME

Questions?

Break Time!

**What does a end to end reproducible project
look like?**

Final quiz