Galaxy

Open-source, transparent, and reproducible bioinformatics

PUBS2015

Franklin Bristow and Eric Marinier
National Microbiology Laboratory
Public Health Agency of Canada

Overview

- Who are we?
- Data
- Bioinformatics
- Open-source software
- Reproducibility
- Galaxy

Who Are We?

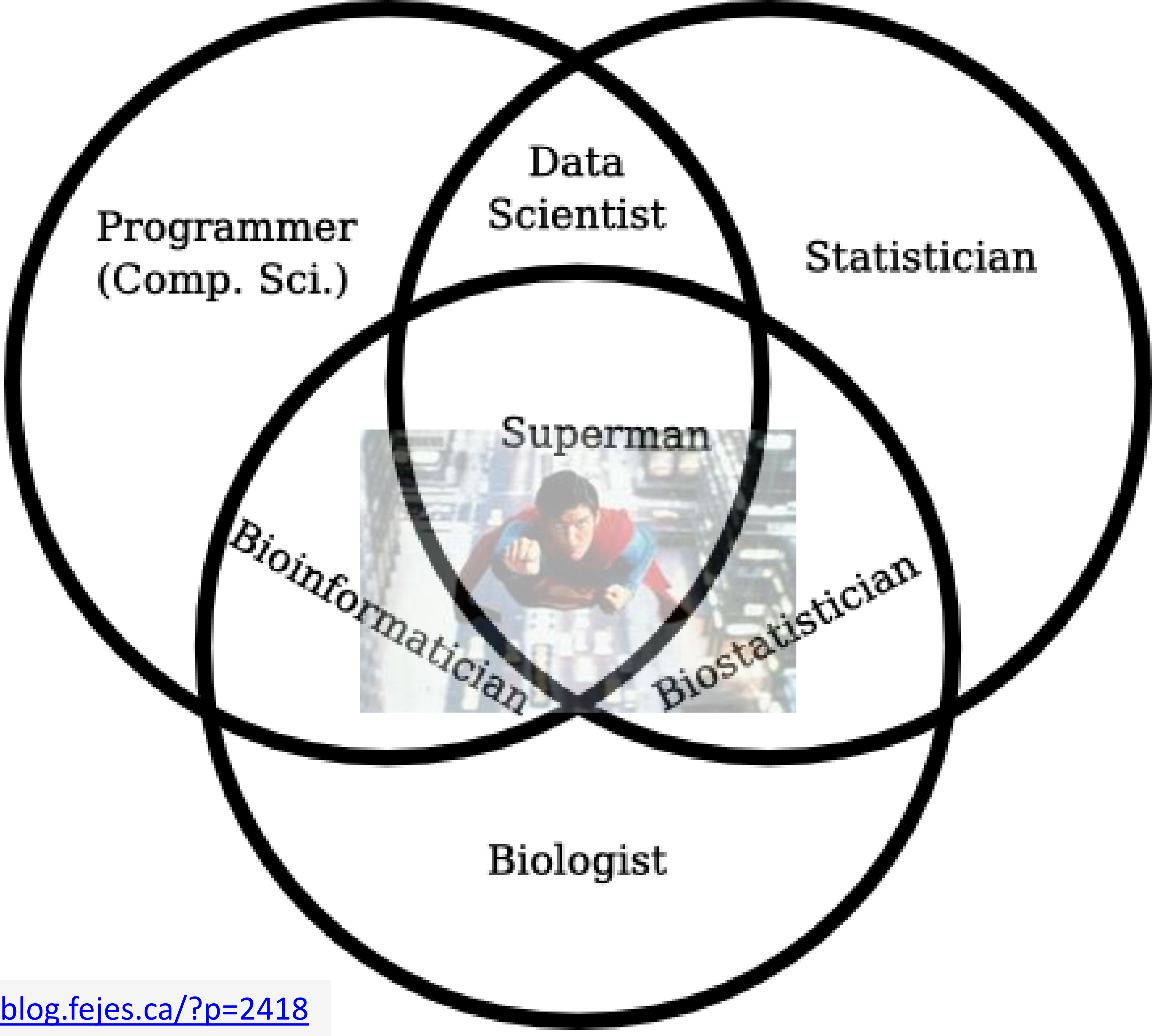
- Franklin Bristow
  - franklin.bristow@phac-aspc.gc.ca
  - https://github.com/fbristow
- Eric Marinier
  - eric.marinier@phac-aspc.gc.ca
  - https://github.com/emarinier

Data

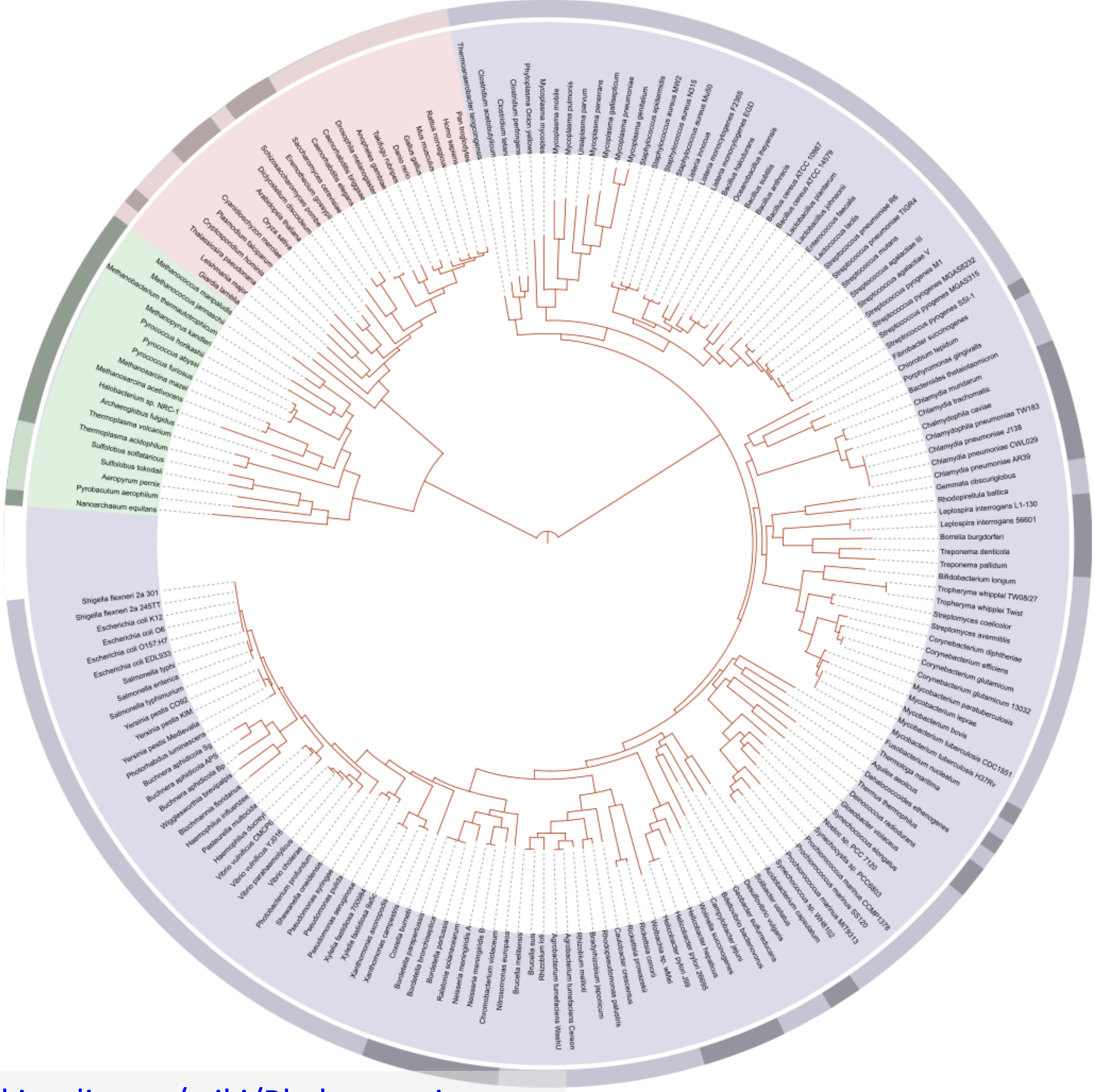
http://discardstudies.files.wordpress.com/2013/04/450_messy1_080516.jpg

Data

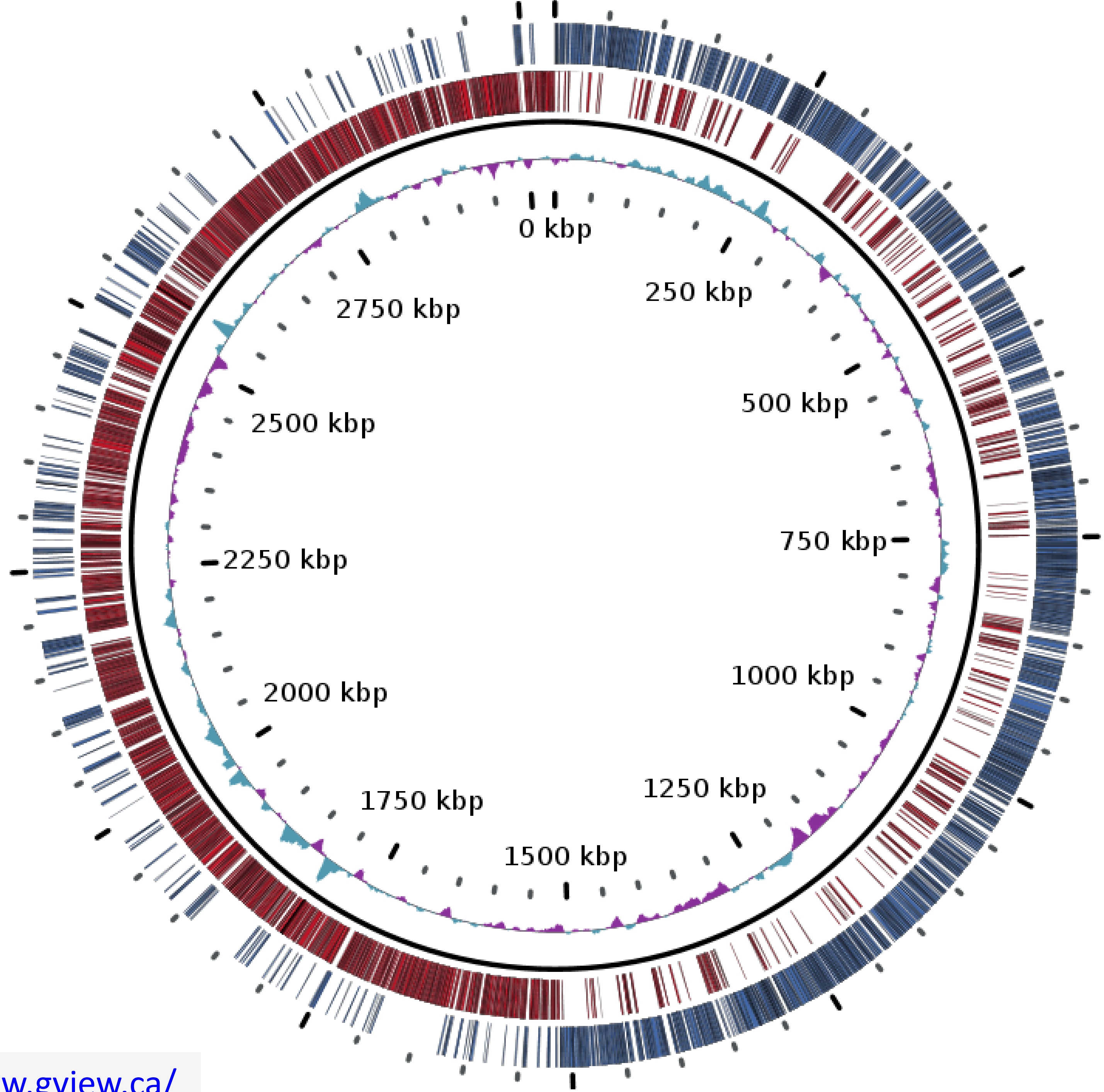
http://prodstatics3cdn1.purewow.com/images/articles/2014_08/Drought400x290.jpg



http://blog.fejes.ca/?p=2418



http://en.wikipedia.org/wiki/Phylogenetic_tree

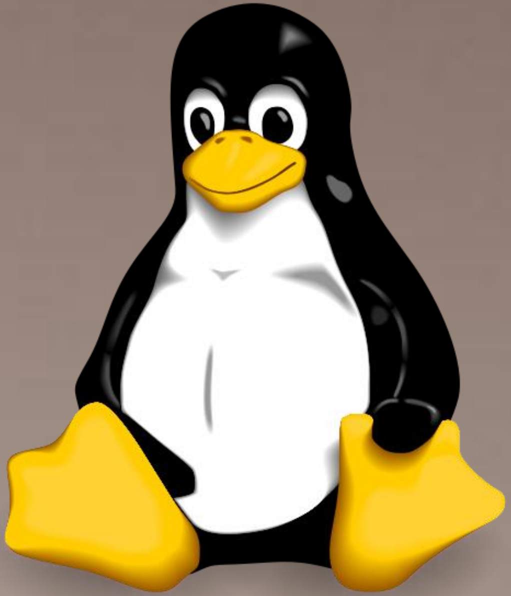


https://www.gview.ca/

Open-Source Software

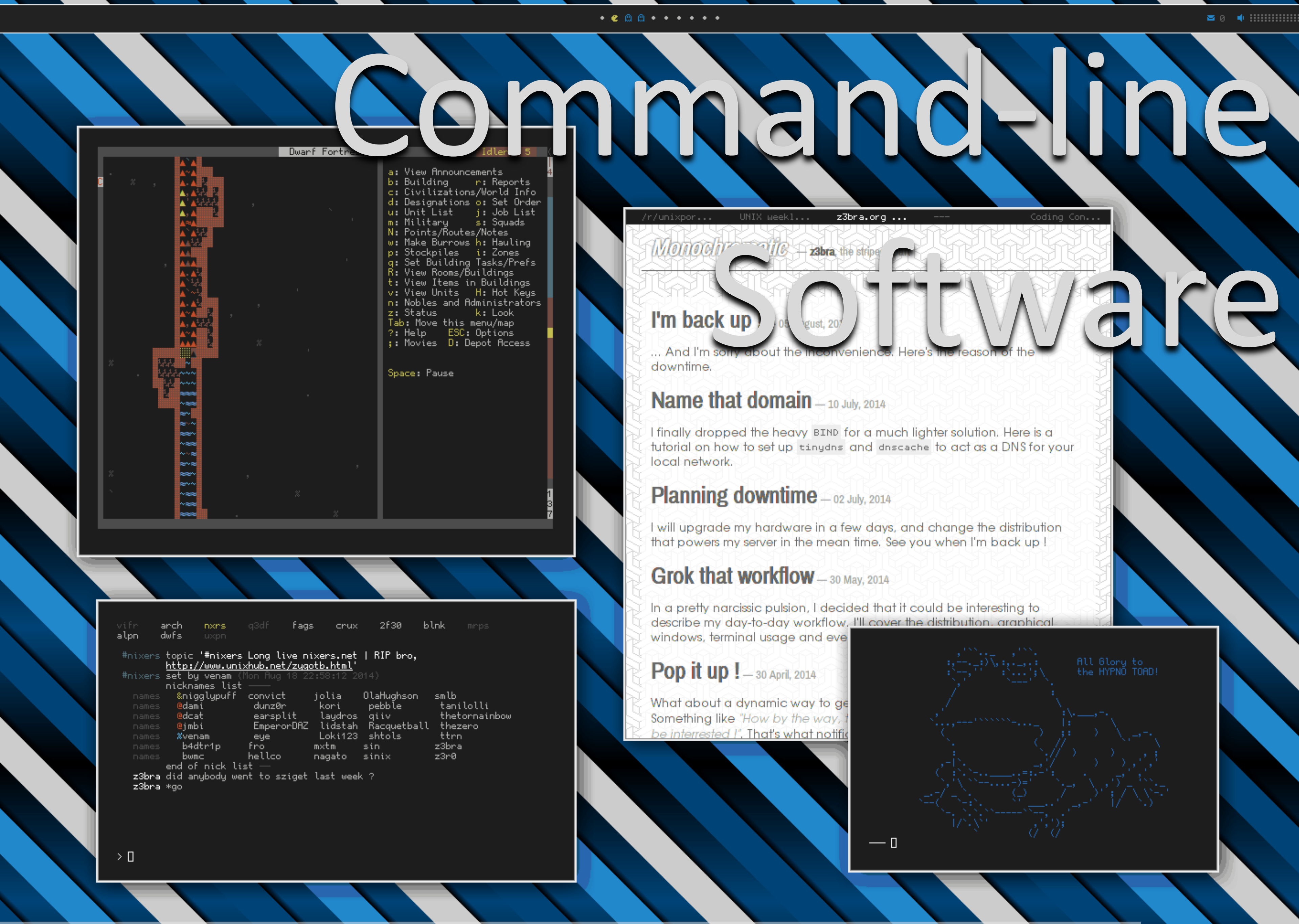
http://en.wikipedia.org/wiki/Richard_Stallman

Open-Source Software



http://en.wikipedia.org/wiki/Tux
http://en.wikipedia.org/wiki/Linus_Torvalds

Command-line Software



```
Dwarf Forti                         Idler  5

a: View Announcements
b: Building      r: Reports
c: Civilizations/World Info
d: Designations  o: Set Order
u: Unit List     j: Job List
m: Military      s: Squads
N: Points/Routes/Notes
w: Make Burrows  h: Hauling
p: Stockpiles    i: Zones
q: Set Building Tasks/Prefs
R: View Rooms/Buildings
t: View Items in Buildings
v: View Units    H: Hot Keys
n: Nobles and Administrators
z: Status        k: Look
Tab: Move this menu/map
?: Help     ESC: Options
;: Movies   D: Depot Access


Space: Pause
```

/r/unixpor... UNIX weekl... z3bra.org ... --- Coding Con...

Monochromatic — **z3bra**, the stripe

I'm back up — 05 August, 20

... And I'm sorry about the inconvenience. Here's the reason of the downtime.

Name that domain — 10 July, 2014

I finally dropped the heavy `BIND` for a much lighter solution. Here is a tutorial on how to set up `tinydns` and `dnscache` to act as a DNS for your local network.

Planning downtime — 02 July, 2014

I will upgrade my hardware in a few days, and change the distribution that powers my server in the mean time. See you when I'm back up !

Grok that workflow — 30 May, 2014

In a pretty narcissic pulsion, I decided that it could be interesting to describe my day-to-day workflow. I'll cover the distribution, graphical windows, terminal usage and eve

Pop it up ! — 30 April, 2014

What about a dynamic way to ge Something like *"How by the way, t be interested !"*. That's what notific

```
vifr     arch     nxrs     q3df     fags     crux     2f30     blnk     mrps
alpn     dwfs     uxpn

 #nixers topic '#nixers Long live nixers.net | RIP bro,
         http://www.unixhub.net/zygotb.html'
 #nixers set by venam (Mon Aug 18 22:58:12 2014)
         nicknames list ----
   names   &nigglypuff  convict     jolia    OlaHughson   smlb
   names   @dami         dunz0r      kori     pebble       tanilolli
   names   @dcat         earsplit    laydros  qiiv         thetornainbow
   names   @jmbi         EmperorDAZ  lidstah  Racquetball  thezero
   names   %venam        eye         Loki123  shtols       ttrn
   names    b4dtr1p     fro         mxtm     sin          z3bra
   names    bwmc        hellco      nagato   sinix        z3r0
         end of nick list --
   z3bra did anybody went to sziget last week ?
   z3bra *go


 > ▯
```

```
 — ▯
```

http://www.reddit.com/r/unixporn/comments/2dzs13/2bwm_lolcat_is_for_sissies/

Reproducibility

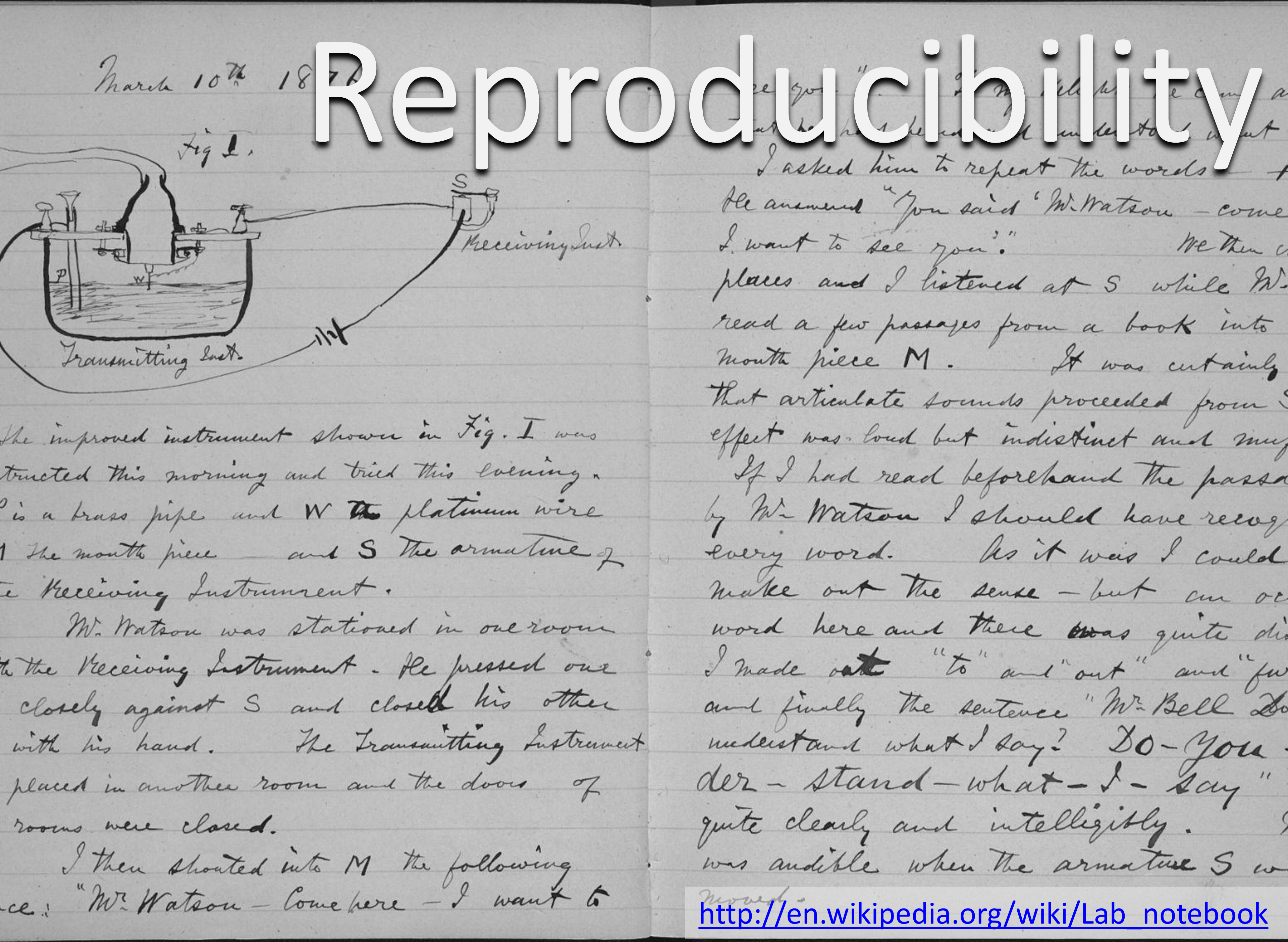


Fig. I.
S
Receiving Inst.
P
W
Transmitting Inst.

March 10th 18

he improved instrument shown in Fig. I was
tructed this morning and tried this evening.
is a brass pipe and W ~~the~~ platinum wire
the mouth piece — and S the armature of
e Receiving Instrument.

Mr. Watson was stationed in one room
h the Receiving Instrument. He pressed one
closely against S and closed his other
with his hand. The Transmitting Instrument
placed in another room and the doors of
rooms were closed.

I then shouted into M the following
ce: "Mr. Watson — Come here — I want to

see you." To my delight he came a
that he had heard and understood what

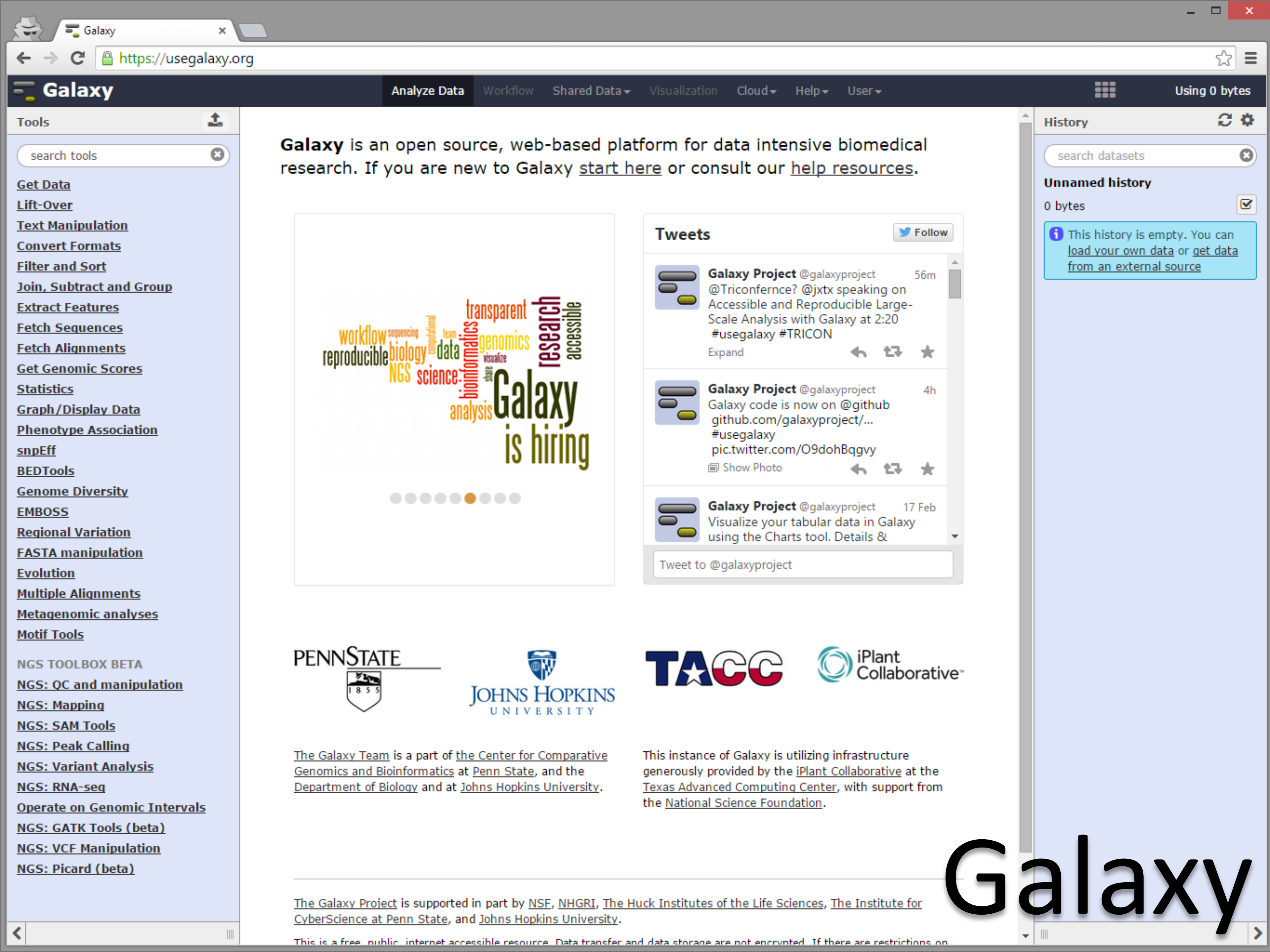
I asked him to repeat the words —
He answered "You said 'Mr. Watson — come
I want to see you.'" We then c
places and I listened at S while Mr.
read a few passages from a book into
mouth piece M. It was certainly
that articulate sounds proceeded from
effect was loud but indistinct and mu

If I had read beforehand the passa
by Mr. Watson I should have recog
every word. As it was I could
make out the sense — but an oc
word here and there ~~was~~ was quite di
I made ~~out~~ "to" and "out" and "fu
and finally the sentence "Mr. Bell Do
understand what I say? Do — you —
der — stand — what — I — say"
quite clearly and intelligibly.
was audible when the armature S w
moved.

http://en.wikipedia.org/wiki/Lab_notebook

Reproducibility

http://gizmodo.com/why-3d-printing-is-overhyped-i-should-know-i-do-it-fo-508176750



This history is empty. You can load your own data or get data from an external source

Galaxy is an open source, web-based platform for data intensive biomedical research. If you are new to Galaxy start here or consult our help resources.

Galaxy Project @galaxyproject 56m
@Triconfernce? @jxtx speaking on Accessible and Reproducible Large-Scale Analysis with Galaxy at 2:20 #usegalaxy #TRICON

Galaxy Project @galaxyproject 4h
Galaxy code is now on @github github.com/galaxyproject/... #usegalaxy pic.twitter.com/O9dohBqgvy

Galaxy Project @galaxyproject 17 Feb
Visualize your tabular data in Galaxy using the Charts tool. Details &

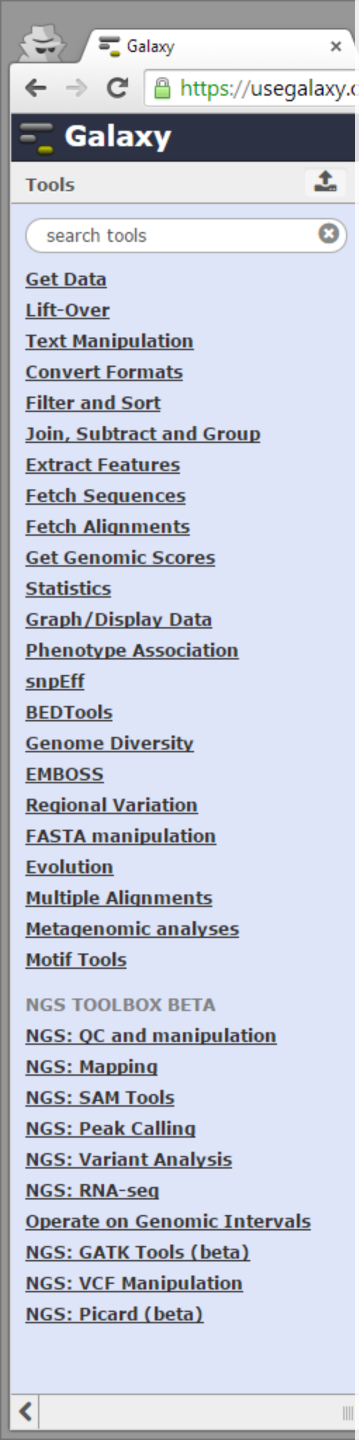
The Galaxy Team is a part of the Center for Comparative Genomics and Bioinformatics at Penn State, and the Department of Biology and at Johns Hopkins University.

This instance of Galaxy is utilizing infrastructure generously provided by the iPlant Collaborative at the Texas Advanced Computing Center, with support from the National Science Foundation.

The Galaxy Project is supported in part by NSF, NHGRI, The Huck Institutes of the Life Sciences, The Institute for CyberScience at Penn State, and Johns Hopkins University.

This is a free, public, internet accessible resource. Data transfer and data storage are not encrypted. If there are restrictions on

Galaxy



Galaxy

https://usegalaxy.org

Galaxy

Analyze Data | Workflow | Shared Data | Visualization | Cloud | Help | User

Using 0 bytes

Tools

search tools

- Get Data
- Lift-Over
- Text Manipulation
- Convert Formats
- Filter and Sort
- Join, Subtract and Group
- Extract Features
- Fetch Sequences
- Fetch Alignments
- Get Genomic Scores
- Statistics
- Graph/Display Data
- Phenotype Association
- snpEff
- BEDTools
- Genome Diversity
- EMBOSS
- Regional Variation
- FASTA manipulation
- Evolution
- Multiple Alignments
- Metagenomic analyses
- Motif Tools

NGS TOOLBOX BETA

- NGS: QC and manipulation
- NGS: Mapping
- NGS: SAM Tools
- NGS: Peak Calling
- NGS: Variant Analysis
- NGS: RNA-seq
- Operate on Genomic Intervals
- NGS: GATK Tools (beta)
- NGS: VCF Manipulation
- NGS: Picard (beta)

History

search datasets

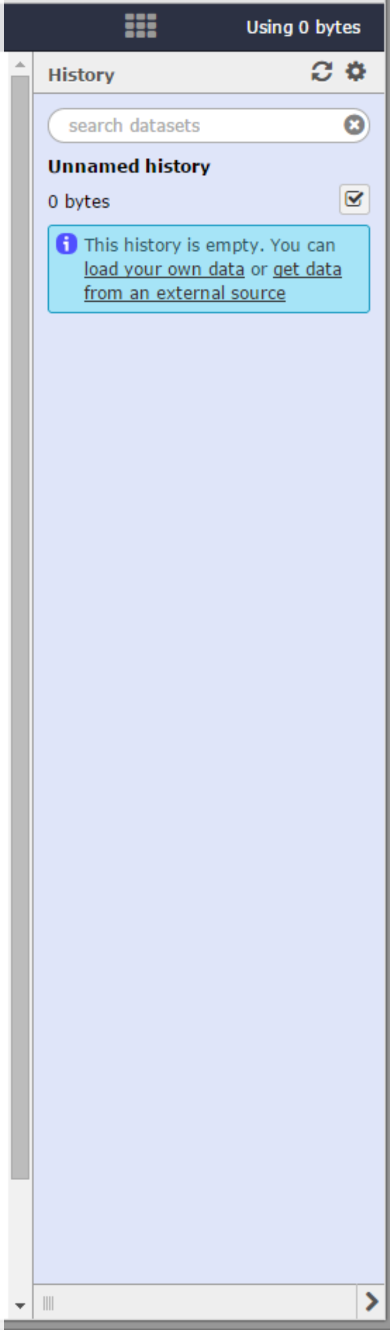
Unnamed history

The Galaxy Project is supported in part by NSF, NHGRI, The Huck Institutes of the Life Sciences, The Institute for CyberScience at Penn State, and Johns Hopkins University.

Toolbox

http://cdn.toysrus.com.au/www/732/files/581070-home-depot-18-piece-tool-box.jpg

History



Using 0 bytes

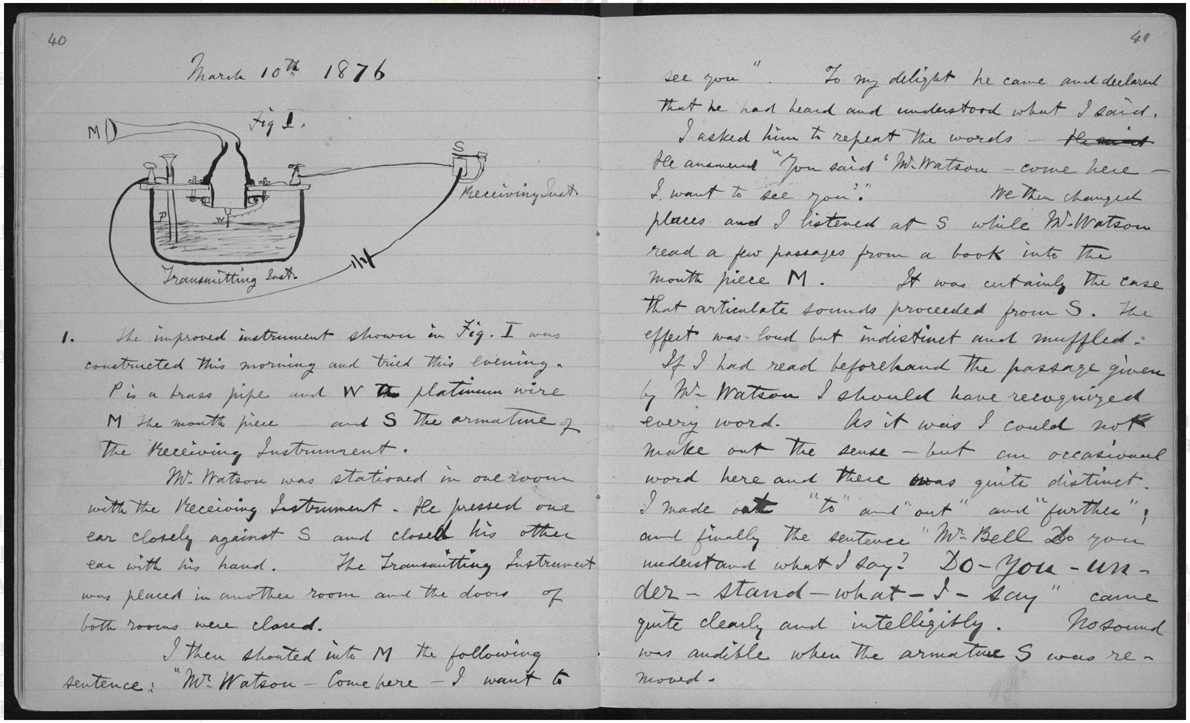
History

search datasets

Unnamed history

0 bytes

This history is empty. You can load your own data or get data from an external source



M
Fig 1.
S
Receiving Inst.
P
W
Transmitting Inst.

40

March 10th 1876

1. The improved instrument shown in Fig. I was constructed this morning and tried this evening.

P is a brass pipe and W ~~the~~ the platinum wire M the mouth piece — and S the armature of the Receiving Instrument.

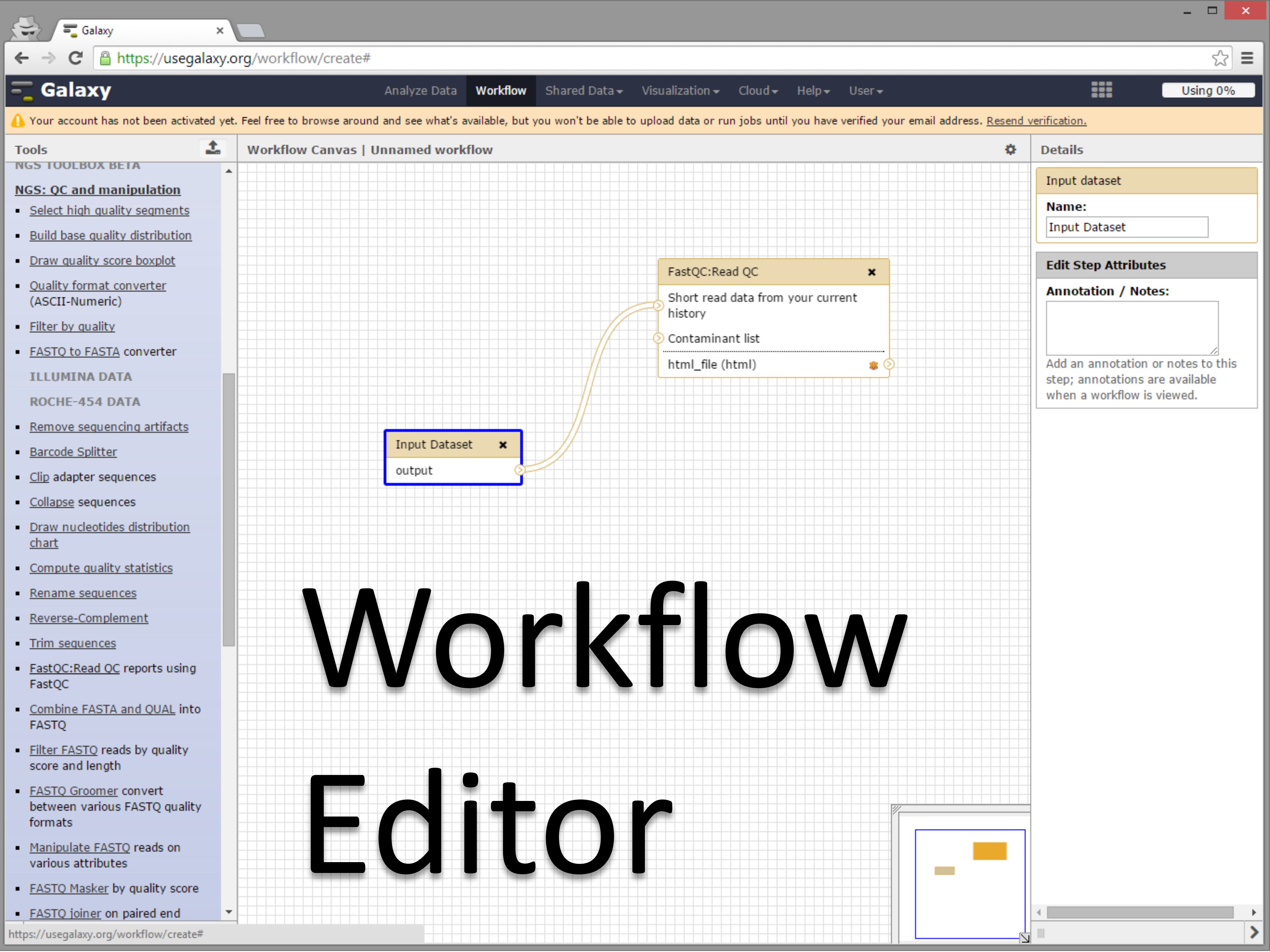
Mr. Watson was stationed in one room with the Receiving Instrument. He pressed one ear closely against S and closed his other ear with his hand. The Transmitting Instrument was placed in another room and the doors of both rooms were closed.

I then shouted into M the following sentence: "Mr. Watson — Come here — I want to see you". To my delight he came and declared that he had heard and understood what I said.

41

I asked him to repeat the words — ~~He said~~ He answered "You said 'Mr. Watson — come here — I want to see you'." We then changed places and I listened at S while Mr. Watson read a few passages from a book into the mouth piece M. It was certainly the case that articulate sounds proceeded from S. The effect was loud but indistinct and muffled.

If I had read beforehand the passage given by Mr. Watson I should have recognized every word. As it was I could not make out the sense — but an occasional word here and there was quite distinct. I made out "to" and "out" and "further"; and finally the sentence "Mr. Bell Do you understand what I say? Do - you - understand - what - I - say" came quite clearly and intelligibly. No sound was audible when the armature S was removed.



Workflow Editor

Galaxy
https://usegalaxy.org/workflow/create#

Galaxy | Analyze Data | Workflow | Shared Data | Visualization | Cloud | Help | User | Using 0%

Your account has not been activated yet. Feel free to browse around and see what's available, but you won't be able to upload data or run jobs until you have verified your email address. Resend verification.

Tools

NGS TOOLBOX BETA

NGS: QC and manipulation

- Select high quality segments
- Build base quality distribution
- Draw quality score boxplot
- Quality format converter (ASCII-Numeric)
- Filter by quality
- FASTQ to FASTA converter

ILLUMINA DATA

ROCHE-454 DATA

- Remove sequencing artifacts
- Barcode Splitter
- Clip adapter sequences
- Collapse sequences
- Draw nucleotides distribution chart
- Compute quality statistics
- Rename sequences
- Reverse-Complement
- Trim sequences
- FastQC:Read QC reports using FastQC
- Combine FASTA and QUAL into FASTQ
- Filter FASTQ reads by quality score and length
- FASTQ Groomer convert between various FASTQ quality formats
- Manipulate FASTQ reads on various attributes
- FASTQ Masker by quality score
- FASTQ joiner on paired end

Workflow Canvas | Unnamed workflow

FastQC:Read QC
Short read data from your current history
Contaminant list
html_file (html)

Input Dataset
output

Details

Input dataset
Name: Input Dataset

Edit Step Attributes
Annotation / Notes:
Add an annotation or notes to this step; annotations are available when a workflow is viewed.

https://usegalaxy.org/workflow/create#

Links

- Lecture notes and exercises:
  - https://github.com/emarinier/
- Public Galaxy instances:
  - https://usegalaxy.org
  - https://orione.crs4.it