

making the web work for science

kaitlin thaney

@kaythaney ; @mozillascience

colorado state / 07 may 2015

mozilla

doing good is part of our code



knight-mozilla

OpenNews





mozilla
Science Lab

help researchers use the
power of the open web to
change science's future.

learning around
open source, data sharing
needed to further open practice;
empowering others to lead in
their communities.

learning around

open source, data sharing:

sprints, collaborate, study groups

empowering others:

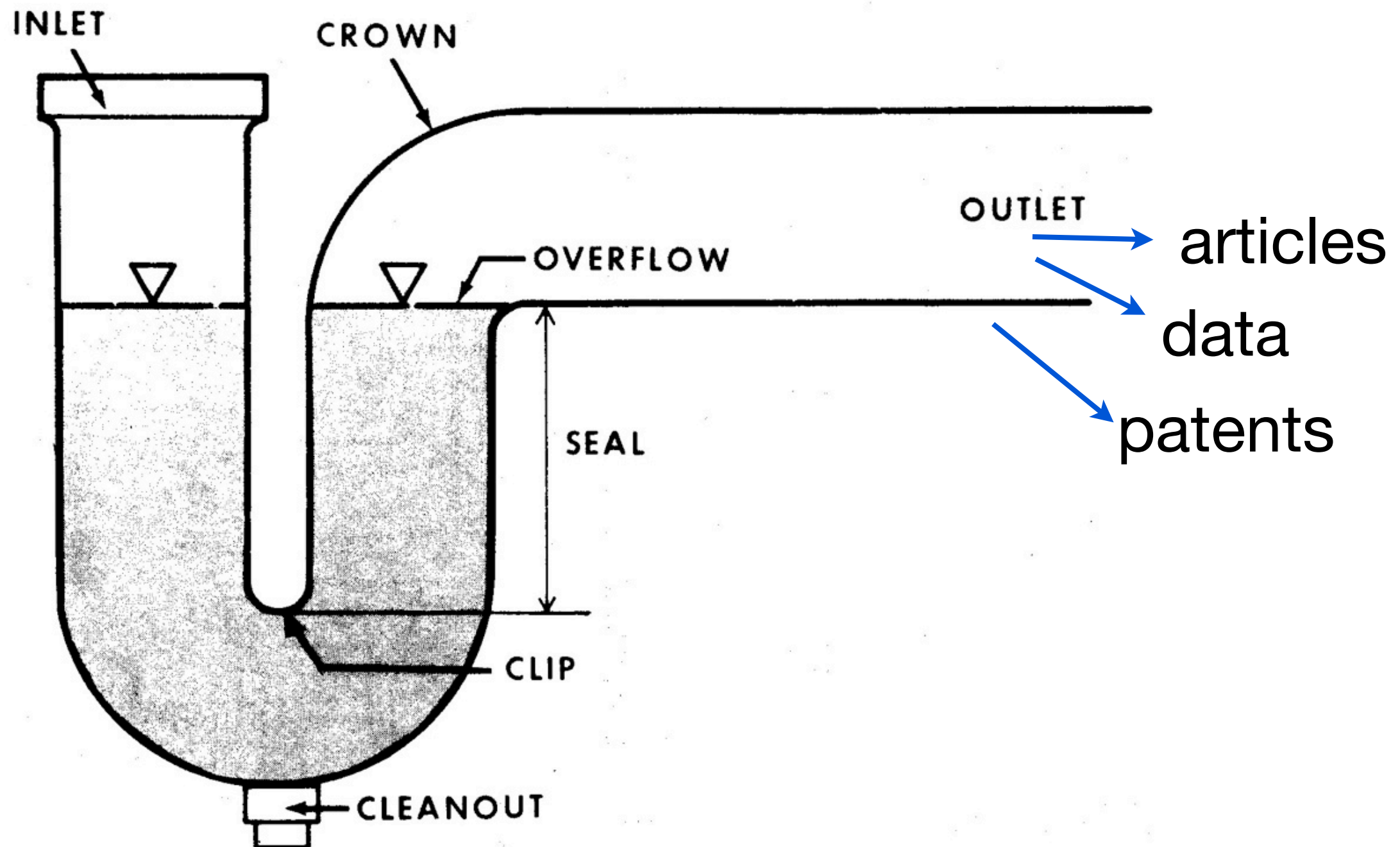
fellowships, mentorship, resbaz

(0)

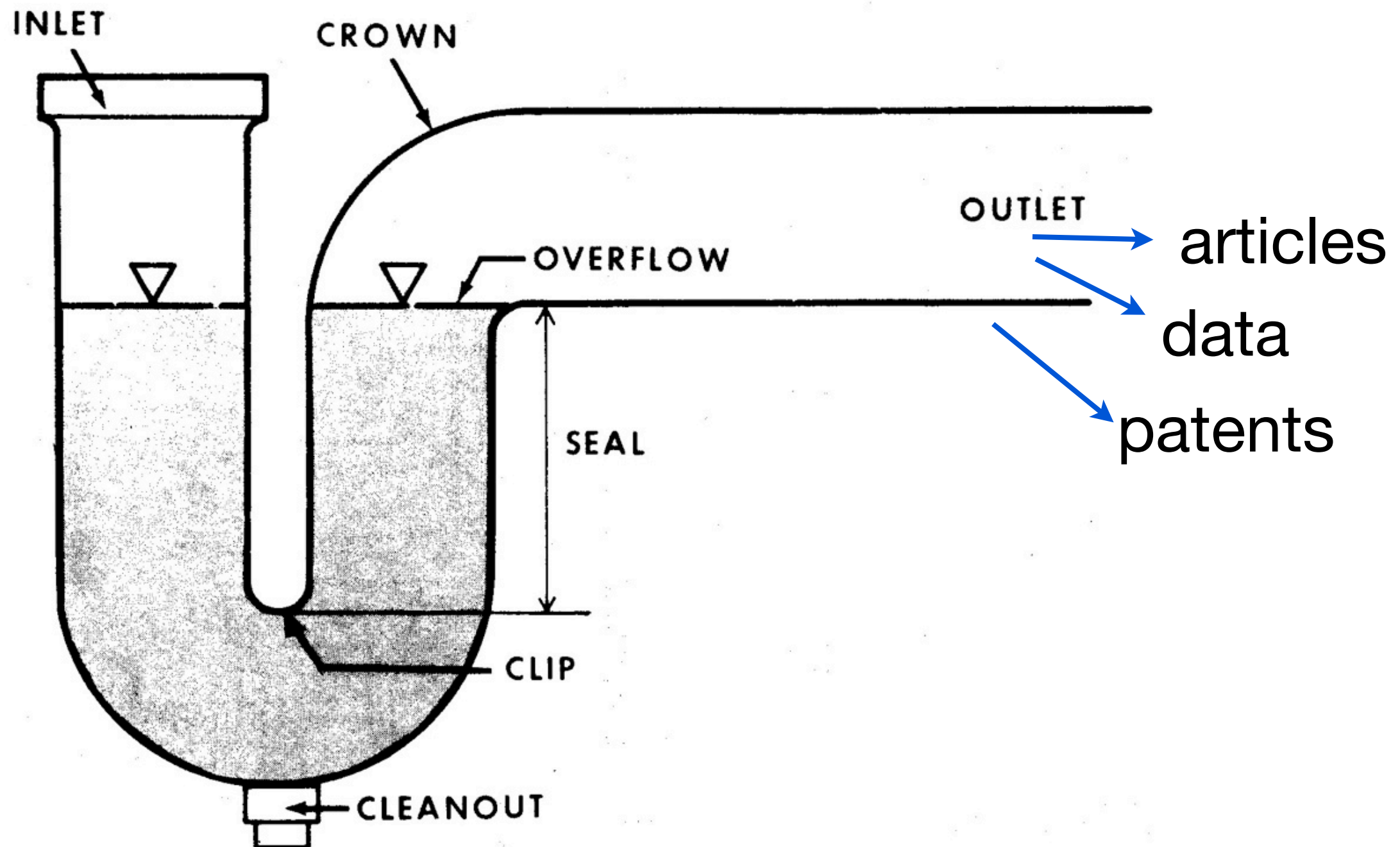
our current systems are
designed to create
friction.

despite original intentions.

current state of science



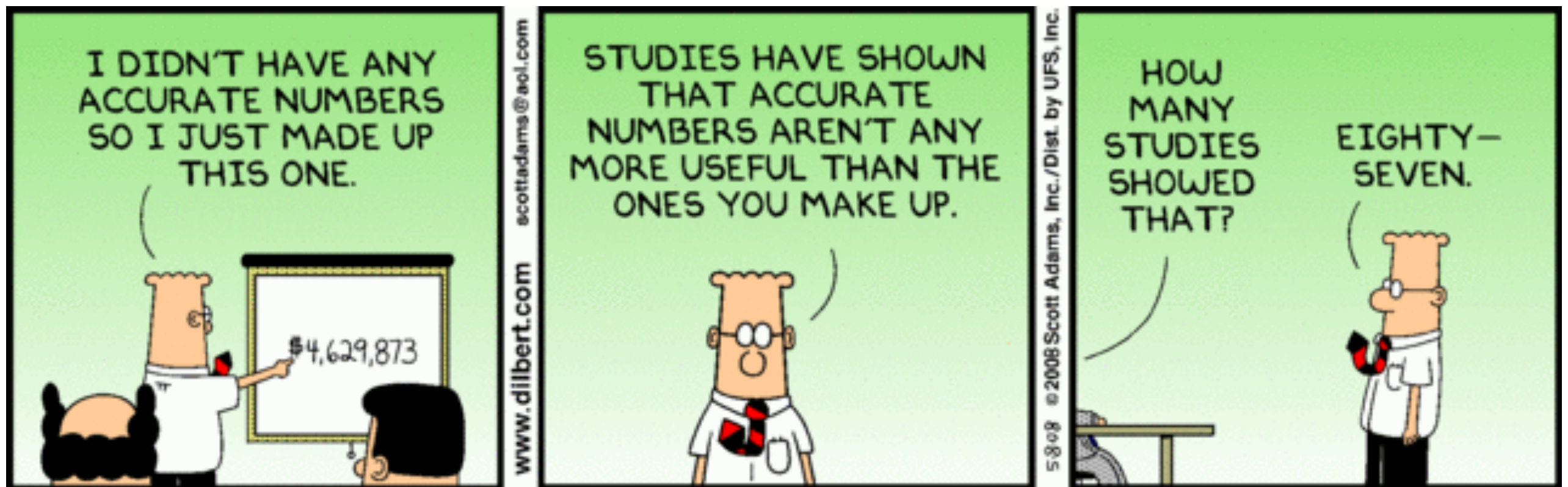
some have a firehose



“traditions last not because they are excellent, but because influential people are **averse** to change and because of the sheer burdens of **transition to a better state ...**”

Cass Sunstein

downside of output-driven recognition systems



“There’s greater reward,
and more temptation to
bend the rules.”

- David Resnik, bioethicist

(1)

leveraging the power of
the web for scholarship

leveraging the power of
the **web** for scholarship





(if we're lucky...)





power, performance, scale

“web-enabled research”

- access to content, data, code, materials.
- distributed work environments, participatory.
- emergence of “web-native” tools, efficiency.
- rewards for openness, interop, collaboration, sharing.
- reuse, recomputability, transparency.



how this maps to research

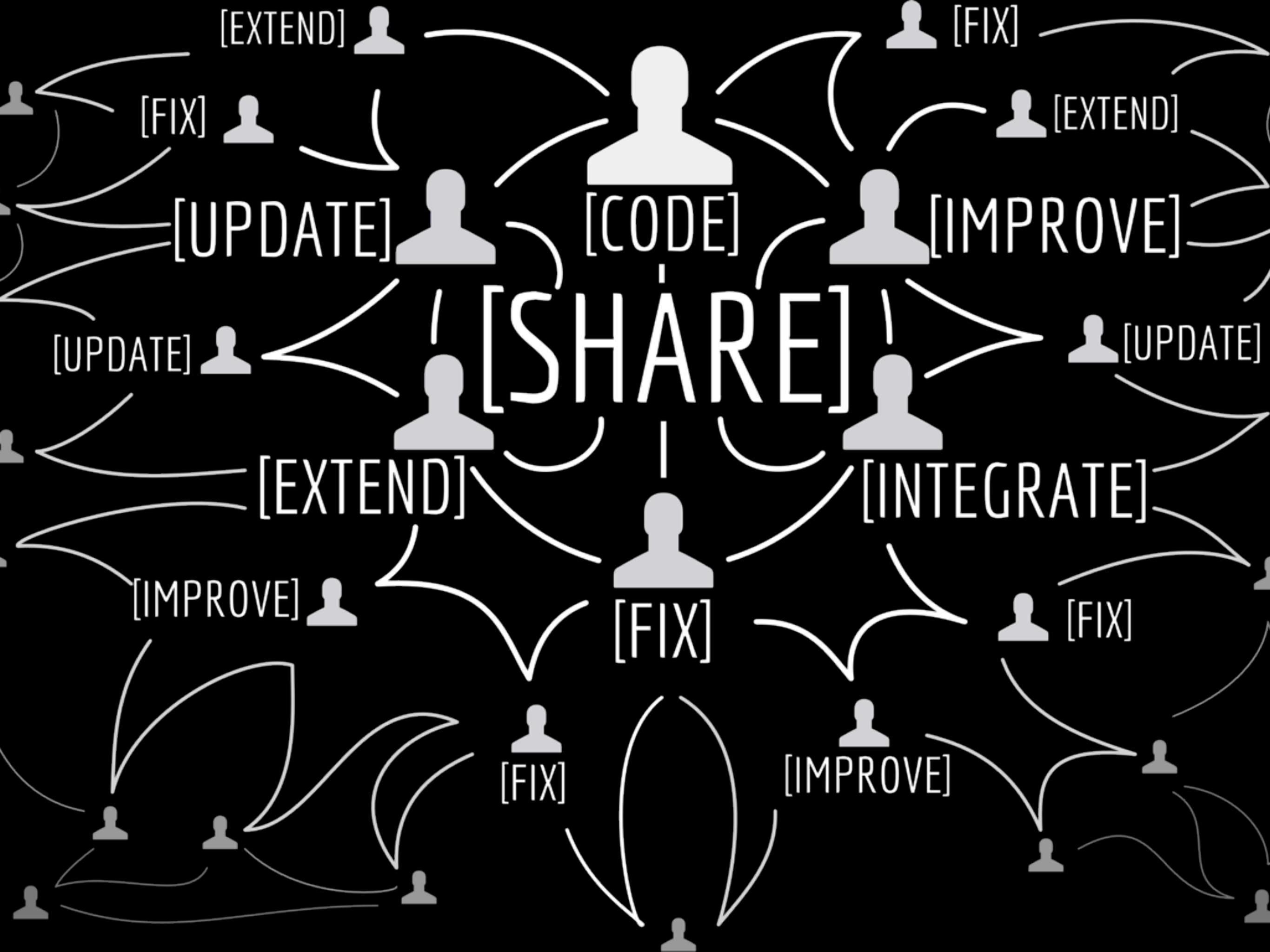
community	technology	practices
collaborative	interoperable	open review
participatory	discoverable	data management
recognition	open tools	sharing / reuse
mentorship	designed for reuse	documentation / versioning

(2)

learning from (+ through)
open source

**applying lessons from open source
development to science**







Mozilla plan seeks to debug scientific code

Software experiment raises prospect of extra peer review.

```

37 GG.add_edge(correspondence[ed[0]],correspondence[ed[1]]);
38
39 winsize = 1 + int(math.floor(GG.number_of_nodes()*0.1));
40 ofile.write(str(winsize) + ',')
41 runs = int(timesteps/winsize);
42 for s in range(runs-1):
43     chance = np.random.random_sample();
44     if chance < 0.1:#5*GG.number_of_nodes()/nodes:
45         ofile.write(str(s+1) + ',');
46         if i in scheduled:
47             temp = scheduled[i];
48             temp.append(s);
49             scheduled[i] = temp;
50         else:
51             temp = list();
52             temp.append(s);
53             scheduled[i] = temp;
54     centrality = nx.betweenness_centrality(GG);
55     for node in GG:
56         #print centrality[node];
57     wins[node,(s)*winsize + int(centrality[node]*200)-1] = i+1;
58     for ad in nx.edges_iter(GG,node):
59         offset = np.random.randint(1,3,[1,1]);
60         wins[ad,(s)*winsize + int(centrality[node]*200) + offset] = i+1;
61 ofile.seek(-1, os.SEEK_CUR);

```

Other reviews

blake

[View](#)

[Reply](#)

Just calling "int(whatever)" will take the floor, so you don't need the call to "math.floor".

An issue was opened.

[Fixed](#)

[Drop](#)



mozillascience / OAButton

forked from OAButton/OAButton_old

Watch

2

★ Star

2

Open Access button – webs

768 commits



branch: master

This branch is 12 commits behind O

force HTTPS for mapbox



crankycoder authored on Jun

deploy

merged changes from prod

11 months ago

oabutton

force HTTPS for mapbox

9 months ago

requirements

Slim down requirements.txt and split up into separate files.

a year ago

scripts

added an email dump script until the admin is extended to have a link

10 months ago

.gitignore

convert back to postgresql

a year ago

.travis.yml

Add phantomjs qunit runner (and make travis run it).

a year ago

INSTRUCTIONS.md

Updated README and added a note about Postgres with a local user into...

a year ago

open, iterative
development
the “work in progress” effect

<> Code

! Issues

Pull re

Pulse

Graph

HTTPS clone

https://g

You can clone

Subversion. C

Clone

Do

code as a research object

what's needed to reuse ?

Data inspection

The codes for variable site are PC=Potter Cove AB= Admiral Bay

```
```{r dataInspection, echo=FALSE, message=FALSE, warning=FALSE}
```

```
chla <- read.table("Meteo_stat.txt", header=T)
pairs(na.omit(chla[,2:6]), panel=function(x,y) { points(x,y); lines(lowess(x,y))})
```

```
#source("panelutils.R")
#pairs(chlaR, panel=panel.smooth, diag.panel=panel.hist, main="")
#hist(chlaR$ClorMAX, col="bisque", right=FALSE)
#hist(sqrt(chlaR$ClorMAX), col="bisque", right=FALSE)
```

```
#hist(chlaR$DegDay_NM, col="bisque", right=FALSE)
#hist(chlaR$ENSO_NM, col="bisque", right=FALSE)
#hist(chlaR$SAM_NM, col="bisque", right=FALSE)
```

<http://bit.ly/mozfiggit>





PUBLIC **mozilla/science / fidgit** **Unwatch** 3 **Star** 4 **Fork** 4  
forked from arfon/fidgit

An ungodly union of GitHub and Figshare <http://fidgit.arfon.org> — Edit

10 commits 1 branch 3 releases 1 contributor

branch: master **fidgit** / +

This branch is 0 commits ahead and 0 commits behind master Pull Request Compare

Update README.md		
arfon	authored on Oct 22, 2013	latest commit a1bc75c349
screens	Screenshots	9 months ago
tmp	Initial commit	9 months ago
views	Initial commit	9 months ago
.gitignore	Initial commit	9 months ago
Gemfile	Initial commit	9 months ago
Gemfile.lock	Initial commit	9 months ago
Procfile	Initial commit	9 months ago
README.md	Update README.md	8 months ago

Code  
Pull Requests 0  
Wiki  
Pulse  
Graphs  
Network  
Settings

HTTPS clone URL  
<https://github.com/m>  
You can clone with [HTTPS](#), [SSH](#), or [Subversion](#).  
Clone in Desktop

To help people find your project, it's useful to provide as much of this information as possible:

[mozillascience/fidgit](https://www.mozillascience.org/fidgit)

http://

First Last <first.last@example.com>

The current maintainer of the project

First Last, First Last, First Last

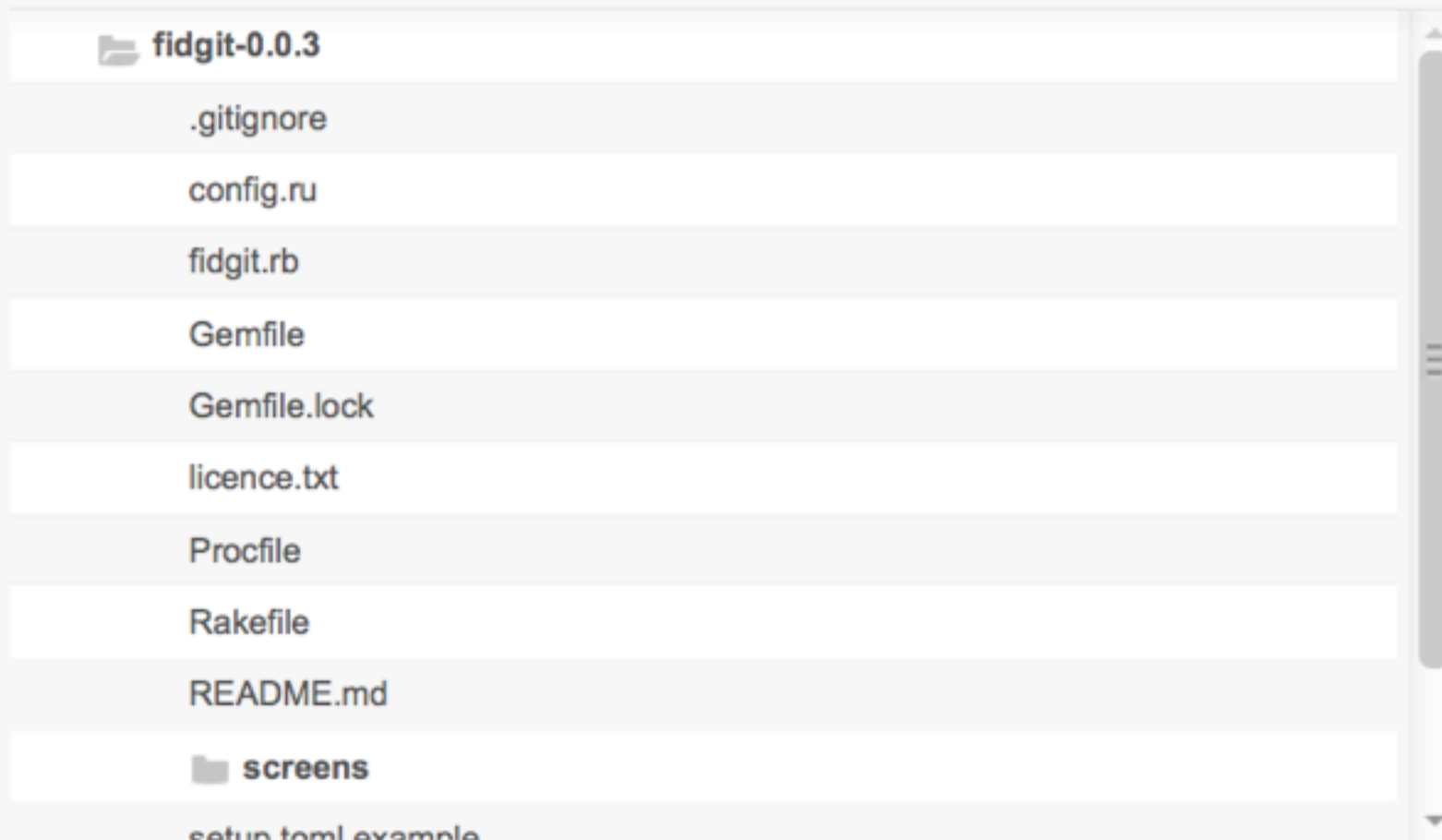
### People who contributed to the project

## An ungodly union of GitHub and Figshare

Summarise this project in a single sentence

Select a license...

# Fidgit - DOIs for code



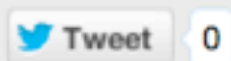
[Enlarge](#)

[Download](#)

Share this:



0



0



0

[Embed\\*](#)

Cite this:

Smith, Arfon (2013): Fidgit - DOIs for code. figshare.  
<http://dx.doi.org/10.6084/m9.figshare.828487>

\*The embed functionality can only be used for non commercial purposes... [more](#)

## Description

391  
views

0  
shares

cites  
coming  
soon

Published on 19 Oct 2013 - 22:27 (GMT)

Filesize is 683.36 KB

## Categories

- [Software Engineering](#)

## Authors

[Arfon Smith](#)

## Tags

- [software](#)

## License (what's this?)

CC-BY

## Export

- [Export to RefWorks](#)
- [Export to BibTeX](#)
- [Export to Ref. manager](#)

# (community driven) metadata for software discovery: JSON-LD

---

```
"@context": "http://schema.org",
"@type": "Code",
"name": "Fidgit",
"codeRepository": "https://github.com/arfon/fidgit",
"citation": "http://dx.doi.org/10.6084/m9.figshare.828487",
"description": "An ungodly union of GitHub and Figshare http://fidgit.arfon.org",
"dateCreated": "2013-10-19",
"license": "MIT",
"author": {
 "@type": "Person",
 "name": "Arfon Smith",
 "@id": "http://orcid.org/0000-0002-3957-2474",
 "email": "arfon@github.com"
}
}
```

<http://bit.ly/mozfiggit>





This objectives of this “Software Discovery Index” would be:

1. to assign standard and unambiguous identifiers to reference all software,
2. to track specific metadata features that describe that software, and
3. to enable robust querying of all relevant information for users.

If broadly used, this Software Discovery Index will form a cornerstone in a software ecosystem that benefits software developers, software users, journal publishers, and funding agencies.

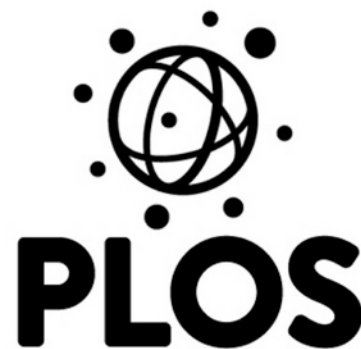
<http://softwarediscoveryindex.org/report/>



**BioMed Central**  
The Open Access Publisher



mozilla  
**Science Lab**



**ORCID**

Connecting Research  
and Researchers



mozilla

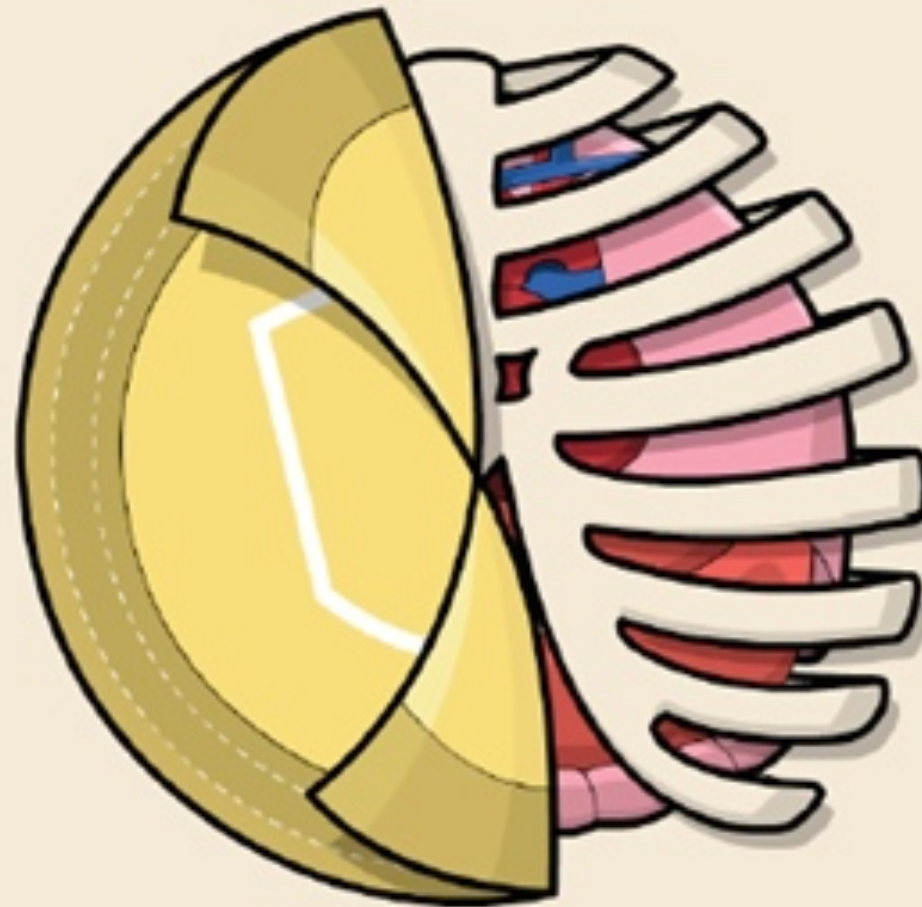
~~OpenBadges~~

Digital  
Credentials



Image CC BY Kyle Bowen

Badge image



OPEN BADGES  
ANATOMY

Badge name

Description

Criteria

Issuer

Evidence

Date issued

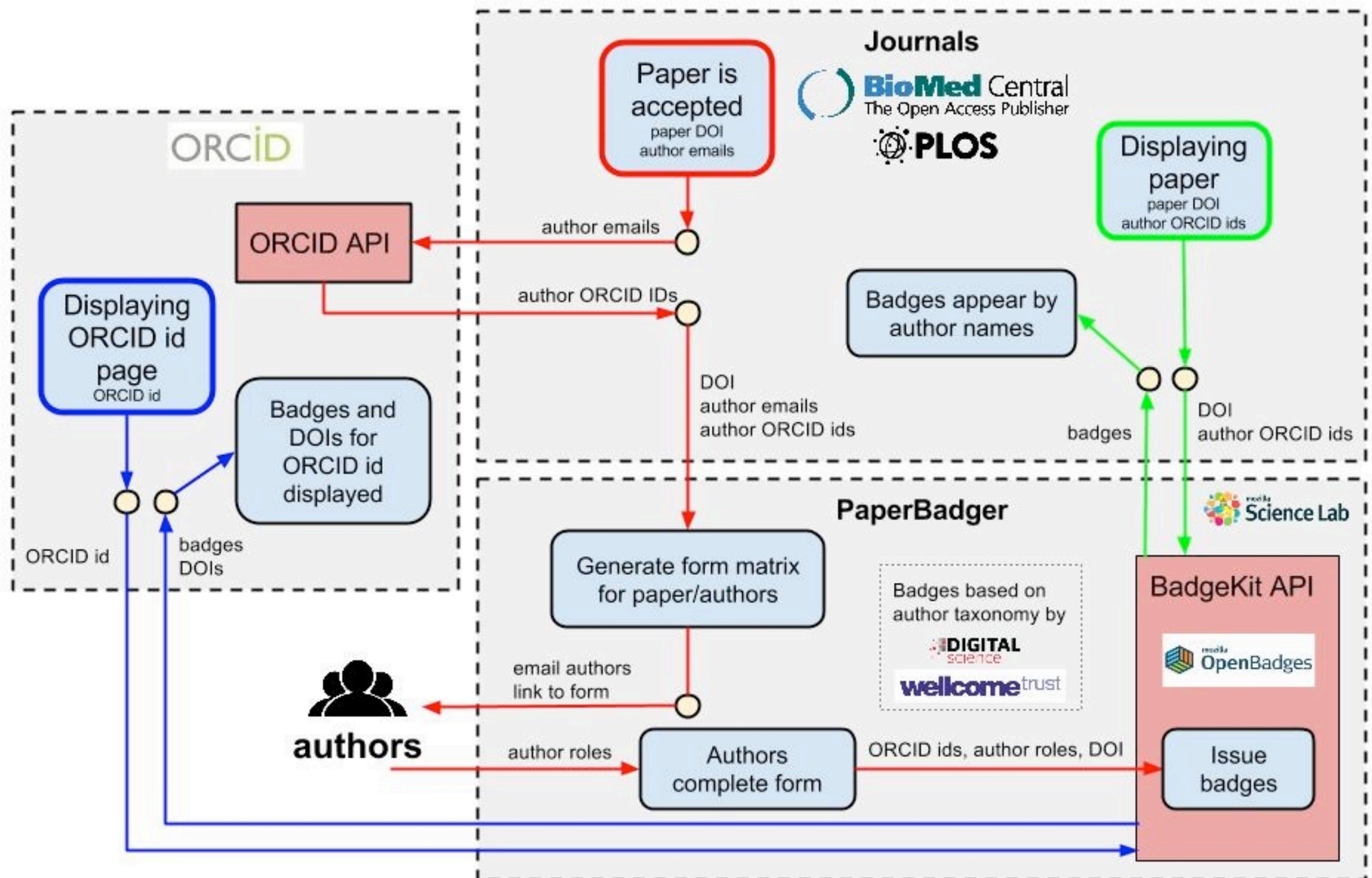
Standards

Tags



# Contributorship Badges - Proposed Implementation / Workflow

Option 1: ORCID implements ID lookup/creation via email address



<http://mozillascience.org/contributorship-badges-a-new-project/>

# Collaborate

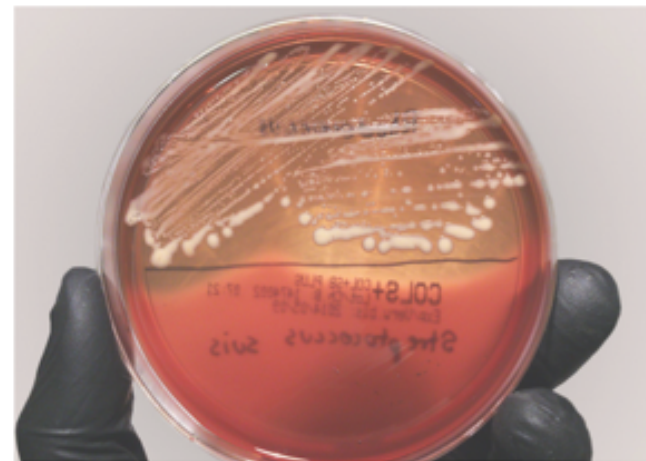
Featured software projects for science in Collaborate



Contributorship Badges



Chemix



Pathogens & Disease Immunity

Code as a  
Research  
Object

Code as a Research Object:  
Standardizing Metadata

Now open for new project submissions!

Submit your project

We're testing out new ways to help the Science Lab community build tools and solve problems together. [Read more >](#)

**<http://www.mozillascience.org/collaborate>**

(3)

our practices are  
limiting us.

**how to further adoption of  
open, web-enabled science?**

# infrastructure layers for efficient, reproducible research

## research

open tools  
standards  
best practices  
research objects  
scientific software  
repositories

## social capital

incentives  
recognition / P&T  
interdisciplinarity  
collaboration  
community dialogue

## capacity

training  
mentorship  
professional dev  
new policies  
recognition

**stakeholders:** universities, researchers,  
tool dev, funders, publishers ...



# **fostering a (sustainable) community of practitioners**



# supports needed for “professional development”





# Mozilla Study Group Handbook

Mozilla Study Groups are a fun and informal way to meet up with friends and colleagues to share code with programming problems and work together. This handbook describes everything you'll need to start your own study group.

Made by the [Mozilla Science Lab](#)

## Introduction

Writing and working with code is becoming a necessary part of daily life for scientists and researchers in almost every field - and yet, the skills to do so are still not included in most formal curricula, few places exist to support and encourage the discussion of scientific code and coding, and opportunities to collaborate on this crucial element of research are rare and hard to find. As a result, students and new researchers struggle to build the skills they need to pursue the science they aspire to, leading to an [epidemic of imposter syndrome](#) in the lab, and risking the waste of a tremendous amount of time and money as we fail to identify places to share and reuse research code, and instead reinvent the wheel, over and over.

### Chapters

1. [Introduction](#)
2. [Organizing a Mozilla Study Group](#)
3. [Event Ideas](#)
4. [Lesson Ideas](#)

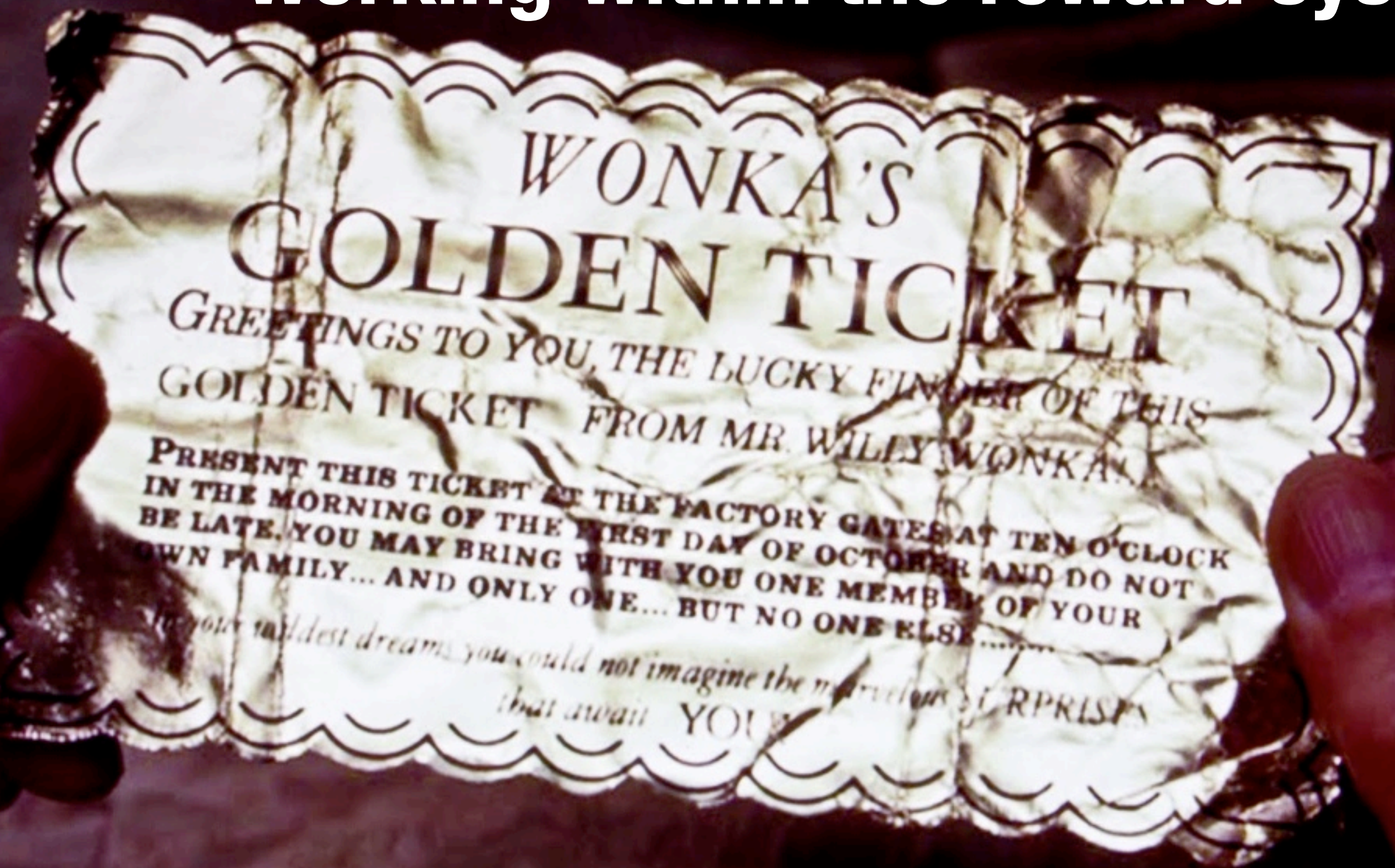
### Subsections in the Introduction

1. [Introduction](#)
2. [Hack This Book](#)
3. [Study Groups & Mozilla Science Lab](#)

<https://mozillascience.github.io/studyGroupHandbook/>



working within the reward system.







**<http://openresearchbadges.org/>**

Course materials from the Open Science Training Initiative

28 commits

1 branch

0 releases

1 contributor

branch: master OpenScienceTraining

Update README.md

StilettoFiend authored 3 months ago

latest commit 8c821f1f3b

LegoSessions	Slide decks from #solo13lego	3 months ago
LICENCE.txt	Rename cc-by-licence to LICENCE.txt	10 months ago
Lecture1A-OpenAndRepro-B...	Lecture 1 - beta version of slides added in ODP and PPTX formats	4 months ago
Lecture1A-OpenAndRepro-B...	Files added for Lego-based sessions	3 months ago
Lecture1B-ExampleRotation...	Lecture 1B: Example handout to accompany RBL assessment	5 months ago
Lecture1B-ExampleRotation...	Lecture 1B: Example handout to accompany RBL assessment	5 months ago
Lecture1B-RotationOutline.odp	Lecture 1B: ODP, PPT slides for rotation-based learning	5 months ago
Lecture1B-RotationOutline.p...	Lecture 1B: ODP, PPT slides for rotation-based learning	5 months ago
Lecture2-Licensing.odp	Lecture 2 - slides added in ODP and PPTX formats	5 months ago
Lecture2-Licensing.pptx	Lecture 2 - slides added in ODP and PPTX formats	5 months ago

Pull Requests 0

Wiki

Pulse

Graphs

Network

HTTPS clone URL

https://github.com/

You can clone with HTTPS, SSH, or Subversion.

Clone in Desktop

Download ZIP



OpenSci



DATA CARPENTRY

MAKING DATA SCIENCE MORE EFFICIENT





**ResBaz**

RESEARCH BAZAAR

[resbaz.edu.au](http://resbaz.edu.au)



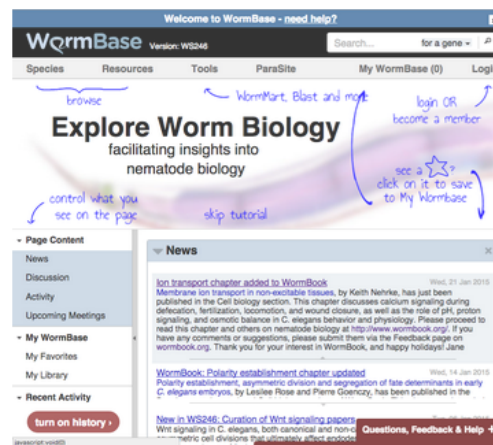




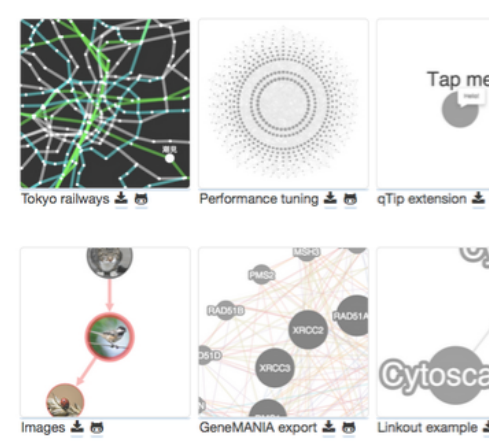
Pathogens & Disease  
Immunity



Contributorship Badges



WormBase Website



Cytoscape.js



Matplotlibdash

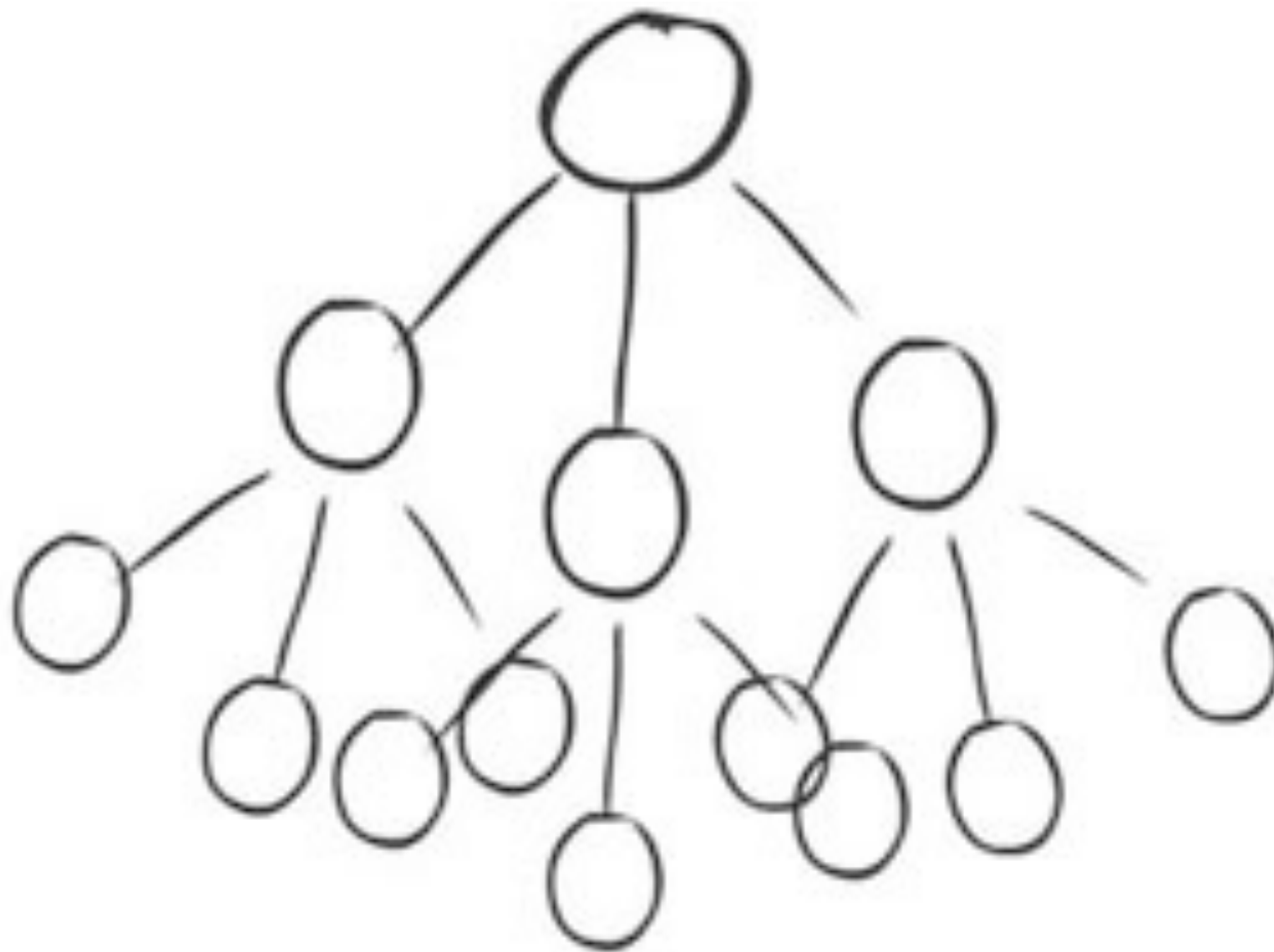
**next global sprint: june 4-5, 2015**  
**[mozillascience.org/collaborate](http://mozillascience.org/collaborate)**



*call opens soon!*

# The Mozilla Open Science Fellowship

The Mozilla Open Science Fellowship enables early-career researchers to spend 10 months working on open, web-enabled research and further open science within the community. Mozilla Open Science Fellows will learn open research skills, write code and help build community during their fellowship year.



lowering barriers to entry  
(not expectations)



# focus on building capacity, not just more nodes.

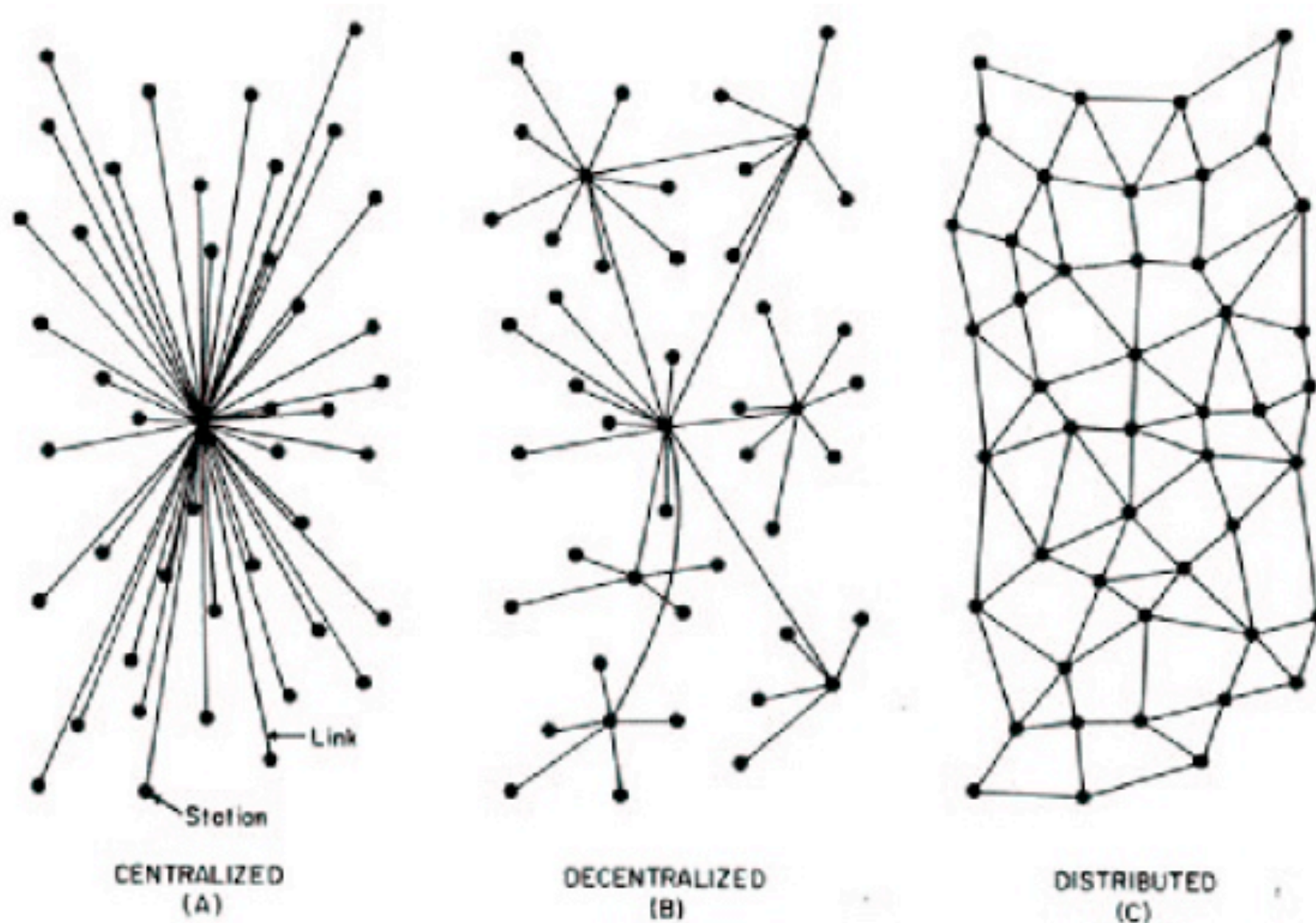


FIG. 1 — Centralized, Decentralized and Distributed Networks

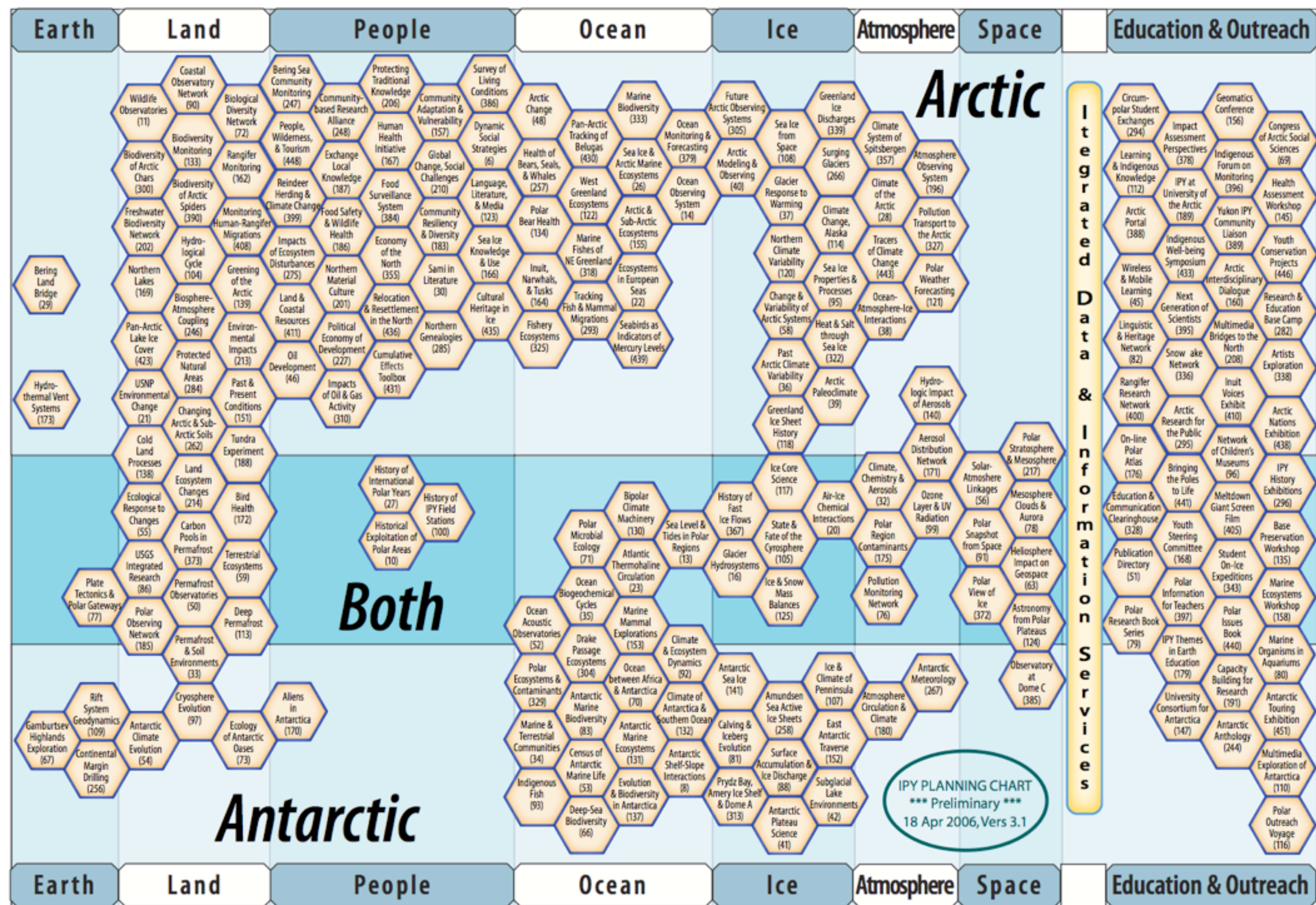


(4)

shifting practice  
(and getting it to stick)  
is challenging.

**open science's collective action problem.**







# **can we do the same for research on the web?**

63 nations

10,000 scientists

50,000 participants



1. bake reproducible  
practices into the fabric of  
research.

2. design to unlock latent  
potential of our systems.  
**(most of the technology is already there.)**

3. rethink how we reward  
researchers and support  
roles.

**(and don't be afraid to hit refresh.)**

4. use cases as means to  
show potential, sell trust  
and drive adoption.



we're here to help.

**<http://mozillascience.org>**  
**[sciencelab@mozillafoundation.org](mailto:sciencelab@mozillafoundation.org)**

kaitlin@mozillafoundation.org  
@kaythaney ; @mozillascience

special thanks:



ALFRED P. SLOAN  
FOUNDATION

THE LEONA M. AND HARRY B.  
**HELMSLEY**  
CHARITABLE TRUST

