

# Institutional Analytics Dashboards with SHARE: The UC San Diego Experience

Declan Fleming  
UCSD Library  
@declan

Jeffrey Spies  
COS | SHARE | UVA  
@jeffspies

# Acknowledgements: People & Funders

## UCSD Library

Arwen Hutt

David Minor

Brian Schottlaender

Aislinn Sotelo

Huawei Wang

Ho Jung Yoo

## COS Labs (R&D)

Josh Bird

Samuel Eagan

## COS

Laura and John Arnold Foundation

## SHARE

IMLS

Alfred P. Sloan Foundation

# DIGITAL COLLECTIONS

Currently totaling 110,943 research data sets, images, documents, video and audio recordings.

Or browse by [Collection](#), [Creator](#), [Format](#), or [Topic](#)

# Browse by Collection: UCSD Research Data Collections

View current collections below. Click an image or link for a full description.

Search Research Data Collections



< 1 2 > Go

Sort: title 20 per page Advanced Search



## [An annotated checklist of the bees \(Hymenoptera: Anthophila\) of San Diego County, California](#)

This database represents an ongoing survey to document all the bee species that occur within the bounds of San Diego County. As this project is still ongoing, the database will be updated periodically with additions and/or corrections.

Creation Date: July 2015



## [Bee Research Methods: Video Demonstrations](#)

This collection shows bee behaviors and methods used to study bee behaviors.

Creation Date: 2005-2014



## [California Cooperative Oceanic Fisheries Investigations \(CalCOFI\): Acoustic and Trawl Data](#)

This is a voluminous data set containing acoustic and trawl data collected on CalCOFI (California Cooperative Oceanic Fisheries Investigations) cruises.

Creation Date: 2009 to 2015,



## [California Smokers Cohort \(CSC\)](#)

The objective of this survey was to collect statewide data on tobacco-related behaviors, knowledge of and attitudes towards smoking, and monitoring the effect of State-initiated programs on individual target populations.

Creation Date: 2011-,



## [California Tobacco Survey \(CTS\)](#)

The objective of these surveys is to collect representative statewide data on cigarette smoking behavior, attitudes towards smoking, media exposure to smoking, and use of tobacco products other than cigarettes, from populations living in California.

Creation Date: 1990-



## [Carbonaceous Aerosol Particle Measurements from Cal-Mex 2010 in Tijuana, Mexico](#)

Single Particle Soot Photometer (SP2) measurements of aerosol refractory black carbon (BC) content for individual particles and Fourier Transform Infrared (FTIR) spectroscopic analysis of organic functional groups of submicron particles at Parque Morelos during Cal-Mex 2010 from 15 May to 30 June.

Creation Date: 2010-05-15 to 2010-06-30

Collections »

## Data from: A Living Vector Field Reveals Constraints on Galactose Network Induction in Yeast

### About this collection

#### Authors

- Rifkin, Scott A.
- Stockwell, Sarah R.

#### Date Collected

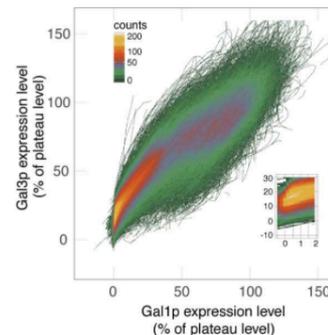
2015-2016

#### Date Issued

2017

#### Cite This Work

Stockwell, Sarah R; Rifkin, Scott A (2017): Data from: A Living Vector Field Reveals Constraints on Galactose Network Induction in Yeast. UC San Diego Library Digital Collections. <https://doi.org/10.6075/J0C24TCX>



View Collection Items

Search this collection...



PubMed



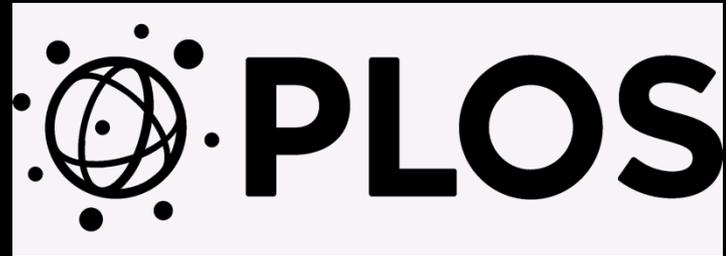
Open Science Framework

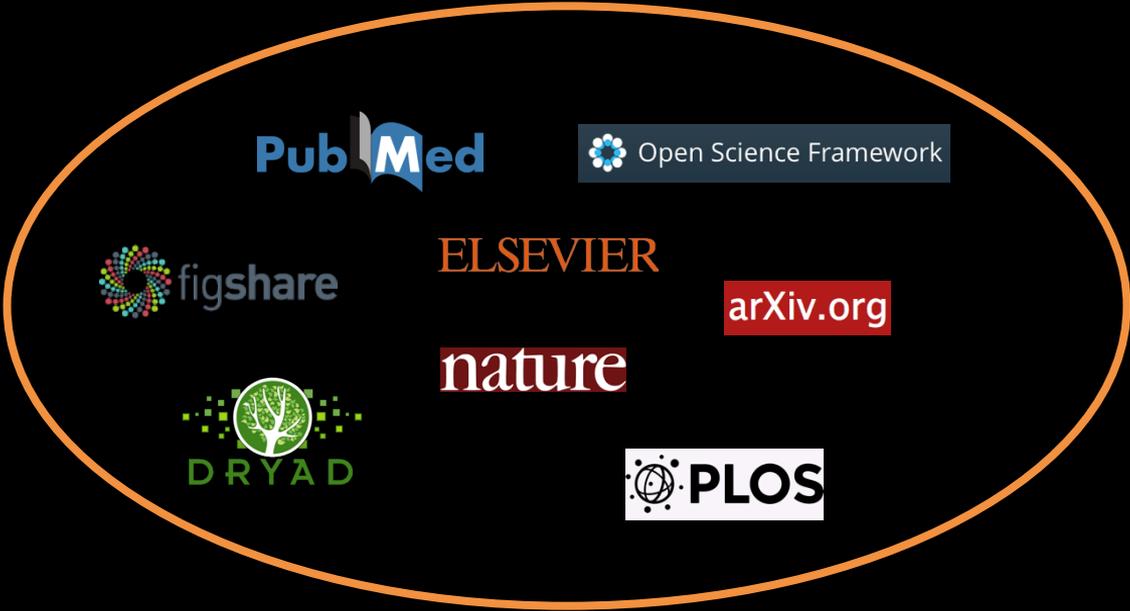


ELSEVIER

arXiv.org

nature





PubMed

 Open Science Framework

 figshare

ELSEVIER

arXiv.org

nature

 DRYAD

 PLOS

# Options

- Spend \$\$\$ with a commercial entity
  - Control issues
  - Trust issues
- Create something from scratch
  - No staff for this
  - Not enough money for this
  - End up with a special snowflake



# Ideally we wanted to ...

- Work with a community to create something that works locally but is reproducible elsewhere
- Leverage work/knowledge that already exists
- Access information/data/metadata that we don't even know about





SHARE

<https://share.osf.io>



Search scholarly works



[Advanced Search](#)

**PUBLICATIONS**



**PREPRINTS**



**SOURCES**



The [SHARE infrastructure](#) was developed by the [Association of Research Libraries](#) in partnership with the [Center for Open Science](#). SHARE is funded by the [Institute of Museum and Library Services](#) and the [Alfred P. Sloan Foundation](#). The SHARE initiative was founded in 2013 by [ARL](#), the [Association of American Universities \(AAU\)](#), and the [Association of Public and Land-grant Universities \(APLU\)](#).

For more information about the organization, go to [share-research.org](https://share-research.org). For more information about the project, including [presentations](#) and [working group details](#) check out [SHARE on the OSF](#).

22,274,169 events as of Mar 23, 2017 Found 15,818,301 events in 0.147 seconds Aggregating 147 sources



View query body

Clear filters

< 1 2 3 ... 1,000 >

1,580,831 additional pages not shown

Sort by: Relevance

Source

Add Source filter



Date

All time

Type

- Data Set
- Patent
- Poster
- Presentation
- Publication** ✖
- Repository
- Retraction
- Software
- Not Categorized

Tag

### THE CROWNS OF CLEOPATRA VII: AN ICONOGRAPHICAL ANALYTICAL STUDY

Al Shafei, Hanan Khamis

Article

DataCite MDS, Zenodo

Abstract The aim of this research is to study the types of crowns worn by Cleopatra VII. The paper explores their shapes and significance, besides explaining the purpose of wearing them and identifying how common were these crowns among Cleopatra VII's antecedents among the Ptolemaic queens. It also explores the common depiction of Cleopatra VII's ...

hellenistic egyptian crown hathoric cleopatra vii +9

Last updated Mar 29, 2016

Published by Zenodo

### IMOTRIS : Intermodales Transport Routing Informations-System ; Teilvorhaben: IGD: Schlussbericht

Fraunhofer-Institut für Graphische Datenverarbeitung - Flach, Guntram

Report

DataCite MDS

computer science engineering 2 mb online-ressource (pdf-date): 31 s report +4

Last updated Mar 16, 2016

Published by Technische Informationsbibliothek u. Universitätsbibliothek

Discours veritable des horribles meurtres et massacres commis et perpetrez de sang froid, par les troupes du Duc de Savoye, conduictes par Dom Amedée, bastard de la dicte Savoye, sur les pauvres paysans du Bailliage de Gex, et mandement de Gaillart et Terry, pres de Geneve, sans aucune exception de aage ou sexe, tant hommes, femmes, qu'enfans masles et femelles.

Discours

Book

DataCite MDS

Other GLN-3420 Other Dubois, Stoer, n° 1590/8 Other CDM, p. 127

in-8° book text e-rara ethz +2

Last updated May 13, 2016

Published by s. n. « Jean Le Court »

22,274,169 events as of Mar 23, 2017

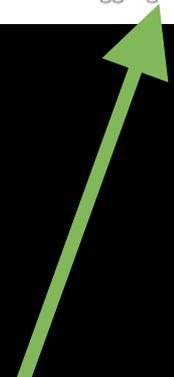
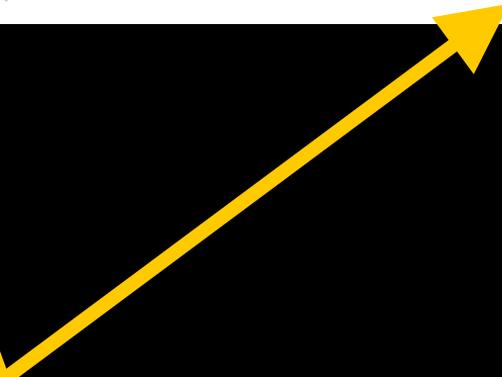
Found 15,818,301 events in 0.147 seconds

Aggregating 147 sources

LOTS  
of data

Performant

Comprehensive



What if ... *what if* we could use  
the SHARE corpus to tell a story  
about our campus research?

It has the metadata about outputs.

It has information about publications.

It has data about funders.

*How do we tie all of these things together and present them?*

# Let's create a dashboard!

- Simple in concept
- Somewhat difficult to implement
  - Institutional naming
  - Inconsistent entries
  - No clear ties between funders and data



# So ... LET'S DO IT

UCSD and CoS had extensive discussions on:

- How to identify our campus within the corpus
- How to present a wide variety of outputs
- How to brand the dashboard

# Timeline

## December

- UCSD visits CoS
  - UCSD RDCP Director and Curator, Metadata Director and Digital Metadata Manager, Technology Director and Developer
  - CoS CoFounder, Many Devs, Many UI folks, Database Dev
  - Rick Johnson from SHARE
- Set up a Slack channel

## January

- Rough prototype

## February

- “Good enough” to show UL

## March

- Branding and UI tweaks
- Present and gather feedback

<http://tritonshare.ucsd.edu>

Data and research output shared by the UC San Diego research community and indexed by SHARE.

[Read more](#)

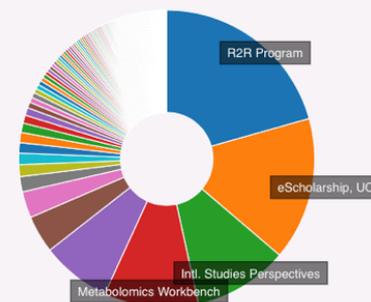
Total Results

**11,255**

## Highlighted Collections

- 1 CAVEcam Virtual Reality Photography Collection
- 2 Data from: Carbonic Anhydrases, EPF2 and a Novel Protease Mediate CO2 Control of Stomatal Development
- 3 Heavy Metals in the Ocean Insect, Halobates
- 4 Keith Rayner Eye Movements in Reading Data Collection
- 5 Stack Gas and Plume Aerosol Measurements from Renewable Diesel and Ultra Low Sulfur Diesel in At-Sea Operation of Research Vessel Robert Gordon Sproul

## Data Providers

[View all](#)

## Types

project: 31.18%

creative work: 25.43%

article: 19.58%

data set: 11.12%

dissertation: 7.75%

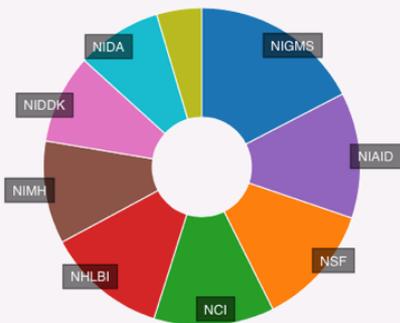
## Top Contributors

**1674** Scripps Institution of Oceanography**564** SIO Geological Data Center (GDC)

252	DOC/NOAA/NESDIS/NODC > National Oceanographic Data Center, NESDIS
59	Ucsd Language And Development Lab
35	Argonne
28	University of California
19	Univ. of California
19	Scripps Institution of Oceanography (SIO) ; University of California, San Diego (UC San Diego)
18	Marco Pereira

[View more](#)

#### Awards



#### Top Tags

1578	disease
1466	modeling
1456	cells
1453	human
1377	time
1314	health
1188	molecular
1165	field_expedition
1165	r2r
1157	response

[View more](#)

Data and research output shared by the UC San Diego research community and indexed by SHARE. [Read more](#) 

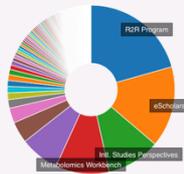
Total Results

11,255

## Highlighted Collections

- 1 CAVEcam Virtual Reality Photography Collection
- 2 Data from: Carbonic Anhydrases, EPF2 and a Novel Protease Mediate CO2 Control of Stomatal Development
- 3 Heavy Metals in the Ocean Insect, Halobates
- 4 Keith Rayner Eye Movements in Reading Data Collection
- 5 Stack Gas and Plume Aerosol Measurements from Renewable Diesel and Ultra Low Sulfur Diesel in At-Sea Operation of Research Vessel Robert S. Sproul

## Data Providers

[View all](#)

## Types



## Top Contributors

- |      |                                     |
|------|-------------------------------------|
| 1674 | Scripps Institution of Oceanography |
| 564  | SIO Geological Data Center (GDC)    |

Manually curated

Created on the fly via query

252	DOC/NOAA/NESDIS/NOI/ National Oceanographic Data Center, NESDIS
59	Language And Development Lab
28	Argonne
28	University of California
19	Univ. of California
19	Scripps Institution of Oceanography (SIO) University of California, San Diego (UC San Diego)
18	Marco Pereira

[View more](#)

## Awards



## Top Tags

- |      |                  |
|------|------------------|
| 1578 | disease          |
| 1466 | modeling         |
| 1456 | cells            |
| 1453 | human            |
| 1377 | time             |
| 1314 | health           |
| 1188 | molecular        |
| 1165 | field_expedition |
| 1165 | r2r              |
| 1157 | response         |
- [View more](#)



All UC San Diego Library ✕

Contributors 64025-884-465 ✕

data set

Friday, February 5, 2016, GMT

Hydrochemistry measured on water bottle samples at station SPac\_1968\_HH\_170W/8-1

Provided by: PANGAEA - Data Publisher for Earth &amp; Environmental Science

[Details >](#)

data set

Friday, February 5, 2016, GMT

Hydrochemistry measured on water bottle samples at station SPac\_1968\_HH\_170W/10

Provided by: PANGAEA - Data Publisher for Earth &amp; Environmental Science

[Details >](#)

data set

Friday, February 5, 2016, GMT

Hydrochemistry measured on water bottle samples at station SPac\_1968\_HH\_170W/9

Provided by: PANGAEA - Data Publisher for Earth &amp; Environmental Science

[Details >](#)

data set

Friday, February 5, 2016, GMT

Hydrochemistry measured on water bottle samples at station SPac\_1970\_HH\_170W/5

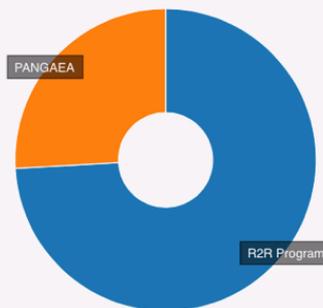
Provided by: PANGAEA - Data Publisher for Earth &amp; Environmental Science

[Details >](#)

Total Results

1,674

Data Providers

[View all](#)

Contributors

1674 Scripps Institution of Oceanography

564 SIO Geological Data Center (GDC)

22 University of California

# Hydrochemistry measured on water bottle samples at station SPac\_1968\_HH\_170W/8-1

[data set](#)

## Contributors

- [Scripps Institution of Oceanography](#)
- [James Swift](#)

## Tags

[hakuho maru](#) [kh-68-4](#) [salinity](#) [temperature](#) [pressure](#) [dataset](#) [pangaea](#) [t1b](#) [text/tab-separated-values](#) [90 data points](#) [water](#) [depth](#)

## Data Providers

- [PANGAEA - Data Publisher for Earth & Environmental Science](#)

Date: Friday, February 5, 2016, GMT

Date Created: Wednesday, November 30, 2016, GMT

Date Modified: Thursday, December 1, 2016, GMT

Date Published:

Date Updated: Friday, February 5, 2016, GMT

Language: eng

## Identifiers

- [oai://oai.datacite.org/370680](https://oai.datacite.org/370680)
- <http://dx.doi.org/10.1594/PANGAEA.477838>

## Sources

- [DataCite MDS](#)

<http://tritonshare.ucsd.edu>

## Up next ...

- General refinement of everything
- Improvement of diverse metadata in catalog
- Connecting improved metadata to sources

And now to Jeff....

A quick dive into architecture



html

+



js

+



css

+



SHARE API

hosted statically on Github sites.

Note: we did put a caching proxy in between the client and SHARE to be nice.

# Approach

# UCSD came to COS with plans to....

Design a schema

Build a database

Run a database server

Create a web app

Run a web app server

Design a front-end

Collect and curate data from around campus

# They left COS with the primary responsibility to....

Collect and curate data from around campus

Design a schema

Build a database

Run a database server

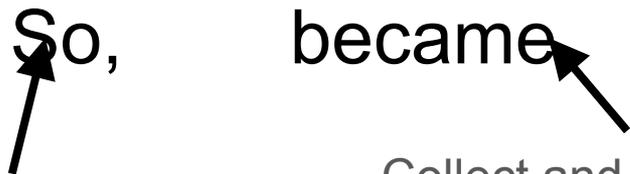
Create a web app

Run a web app server

Design a front-end

Collect and curate data from around  
campus

So,



became

Collect and curate data from around  
campus

# This was a good deal for both groups

The data from UCSD was essential to build a UCSD dashboard (obviously)

That data didn't currently exist in a curated form

They are curation expert

They know their community

Plus COS/SHARE received

A ton of feedback on dashboard design

Insight into curation workflows

A dedicated partner

A scalable approach via  
partnerships

COS & SHARE can't reach every community. That's not scalable.

But we can support members within those communities--that know them and speak their language.

# http://osf.io/preprints

The image shows a browser window displaying the OSF Preprints website. The browser's address bar shows the URL <https://osf.io/preprints/>. The website header includes the OSF logo and the text "Open Science Framework". Navigation links include "Browse", "Support", "Sign Up", and "Sign In". A secondary navigation bar features "OSF Preprints", "Add a preprint", and "Search". The main content area has a dark blue background with the text "OSF Preprints" and "The open preprint repository network". A search bar contains the placeholder text "Search preprints..." and a "Search" button. Below the search bar, it states "350,056 searchable preprints as of September 27, 2016". A green "Add a preprint" button is centered, with a link "See an example" below it. The bottom section, titled "Browse by subject", displays a grid of blue buttons for various academic fields: Architecture, Arts and Humanities, Biology and life sciences, Business, Computer and information sciences, and Earth sciences.

OSF Preprints

The open preprint repository network

Search preprints... Search

350,056 searchable preprints as of September 27, 2016

or

Add a preprint

[See an example](#)

Browse by subject

- Architecture
- Arts and Humanities
- Biology and life sciences
- Business
- Computer and information sciences
- Earth sciences

SocArXiv Preprints

[Add a preprint](#) [Search](#) [Sign Up](#) [Sign in](#)

# SOC ARXIV

Open archive of the social sciences  
[Powered by OSF Preprints](#)

Search preprints...  
605 searchable preprints as of December 12, 2016

or

[Add a preprint](#)  
[See an example](#)

Browse by subject

- Arts and Humanities
- Law

PsyArXiv Preprints

[Add a preprint](#) [Search](#)

# PsyArXiv

A free preprint service for the psychological sciences.  
[Powered by OSF Preprints](#)

Search preprints...  
145 searchable preprints as of December 12, 2016

or

[Add a preprint](#)  
[See an example](#)

Browse by subject

- Engineering
- Life Sciences
- Social and Behavioral Sciences

engrXiv Preprints

[Add a preprint](#) [Search](#) [Sign Up](#) [Sign in](#)

# engrXiv ARCHIVE

The open archive of engineering.  
[Powered by OSF Preprints](#)

Search preprints... [Search](#)  
31 searchable preprints as of December 12, 2016

or

[Add a preprint](#)  
[See an example](#)

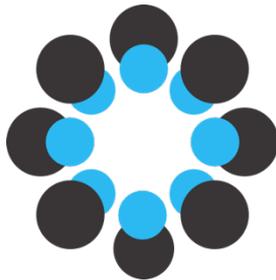
Browse by subject

- Architecture
- Engineering
- Physical Sciences and Mathematics

The power of public goods



SHARE is a **free, open dataset** of research activity across the research workflow supported by **free, open source tools**.



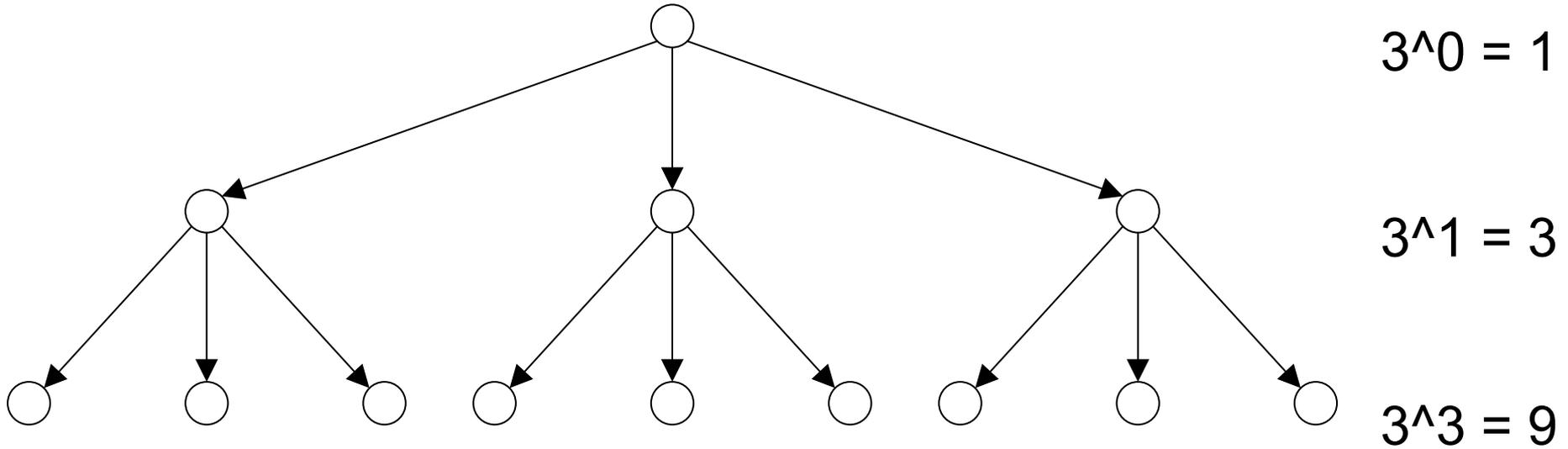
# Public good

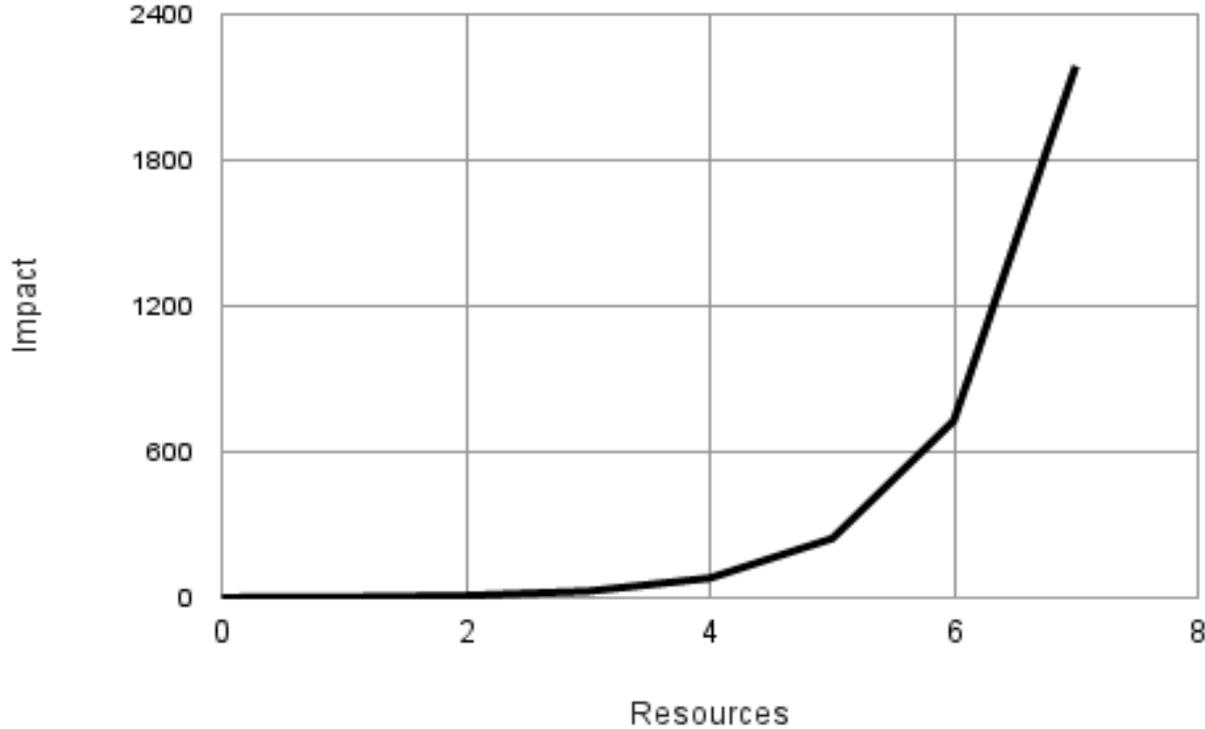
A commodity or service that is provided **without profit** to **all members of a society**, either by the government or a private individual or organization. (Oxford)

A good that is both **non-excludable** and **non-rivalrous** in that individuals cannot be effectively excluded from use and where use by one individual does not reduce availability to others. (Wikipedia)

Openness fosters inclusivity, collaboration,  
and innovation.

# Scaling





$$3^7 = 2187$$

$$3^6 = 729$$

$$3^5 = 243$$

$$3^4 = 81$$

$$3^3 = 9$$

$$3^1 = 3$$

$$3^0 = 1$$

# How to achieve efficient scaling

Engage broadly (and allowing others to engage broadly)

Facilitate experts being experts

Especially as force multipliers--make others more efficient

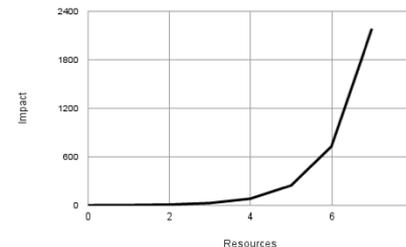
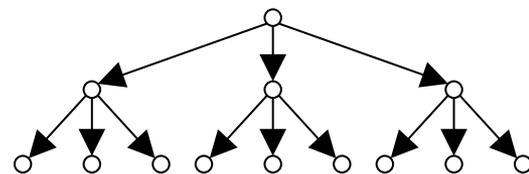
Facilitate (unknown) innovations

By

Respecting current incentives and current workflows

Allowing people to be selfish

Creating virtuous cycles



# Why is this important?

We need to collect and make use of campus analytics before big, for-profit publishers sell them to us and lock us in like we are with text.

The use of institutional identifiers is the most important task that the metadata community can take up.

# How you can get involved

Contribute to SHARE

Use SHARE

Use institutional identifiers (also use person & content identifiers!)

Create a dashboard or talk to us about building one

Contribute to other public goods

Start with a comprehensive focus e.g. datasets or a department

People like “comprehensive”

Be selfish

Find this presentation at

<https://osf.io/views/cni2017spring>

20170403\_cni\_fleming-spies.pptx

Delete

Check out

Share

Download

View

Revisions

Filter

Presentations

OSF Storage

2015.10.GHC.general.s...

20160107\_cendi\_spies...

20160128\_uva\_dev\_ps...

20160205\_rpi\_rcos\_sp...

20160308\_sparc\_spies...

Page: 1 of 82

Automatic Zoom

# Institutional Analytics Dashboards with SHARE: The UC San Diego Experience

Declan Fleming

Jeffrey Spies