

ACCELERATING SYNTHESIS SCIENCE THROUGH REPRODUCIBLE SCIENCE PRACTICES

Matthew B. Jones

*National Center for Ecological Analysis and Synthesis
University of California Santa Barbara*



@metamattj

jones@nceas.ucsb.edu

<https://orcid.org/0000-0003-0077-4738>

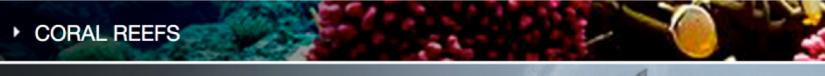


Ecological Synthesis

Marine Systems

- ESTUARINE AND MARINE NURSERIES 
- RECRUITMENT PATTERNS 
- DEEP SEA BIODIVERSITY 
- ECOSYSTEM-BASED MANAGEMENT 
- MARINE PROTECTED AREAS 

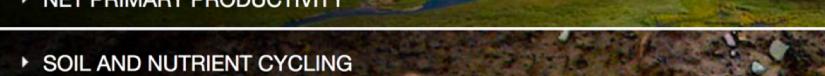
Threats and Population Declines

- SEAGRASS ECOSYSTEMS 
- CORAL REEFS 
- MARINE MAMMALS 
- SEA TURTLES 
- FISHING 
- CLIMATE CHANGE 

Understanding Ocean Health

- MEASURING BIODIVERSITY 
- ECOSYSTEM SERVICES 
- MAPPING HUMAN IMPACTS 
- OCEAN HEALTH INDEX 
- OCEAN TIPPING POINTS 

Climate and Ecosystems

- ARCTIC ECOSYSTEMS 
- FIRE REGIMES 
- FORESTS 
- FRESHWATER AND WETLAND ECOSYSTEMS 
- NET PRIMARY PRODUCTIVITY 
- SOIL AND NUTRIENT CYCLING 
- PERMAFROST 



Reproducible
Science



Provenance



Citation

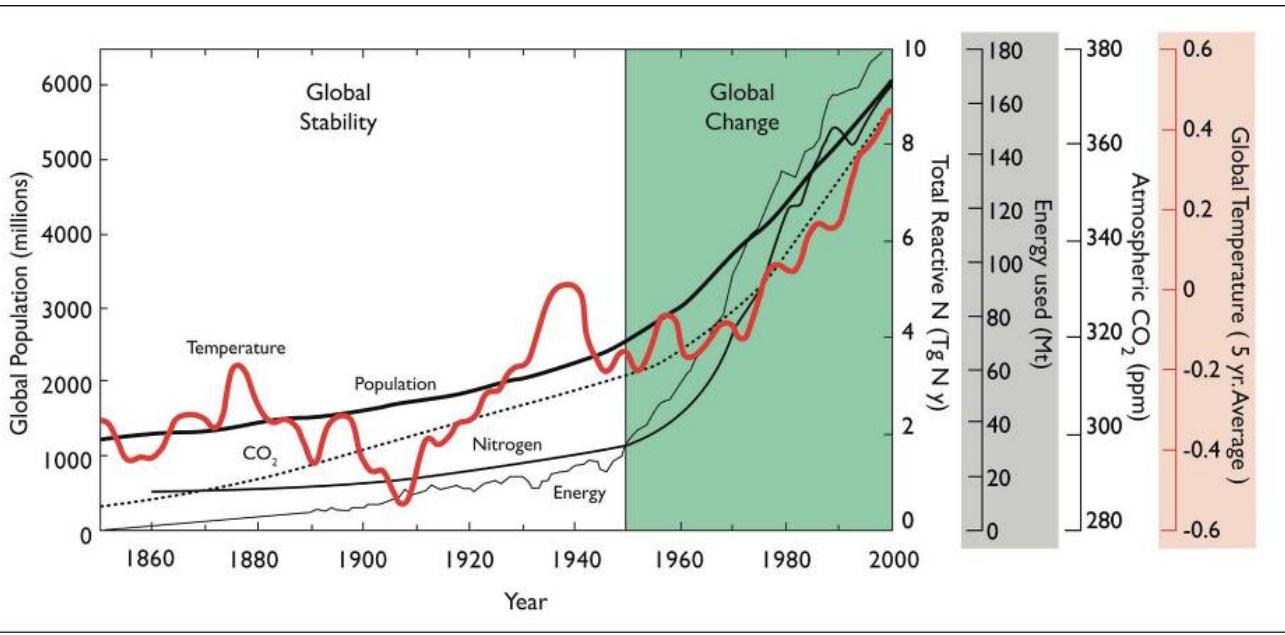


Synthesis

Reproducible Science



Trust in Science



What **data**?
What **methods**?
What **parameter settings**?

Can we **trust** these data and methods?

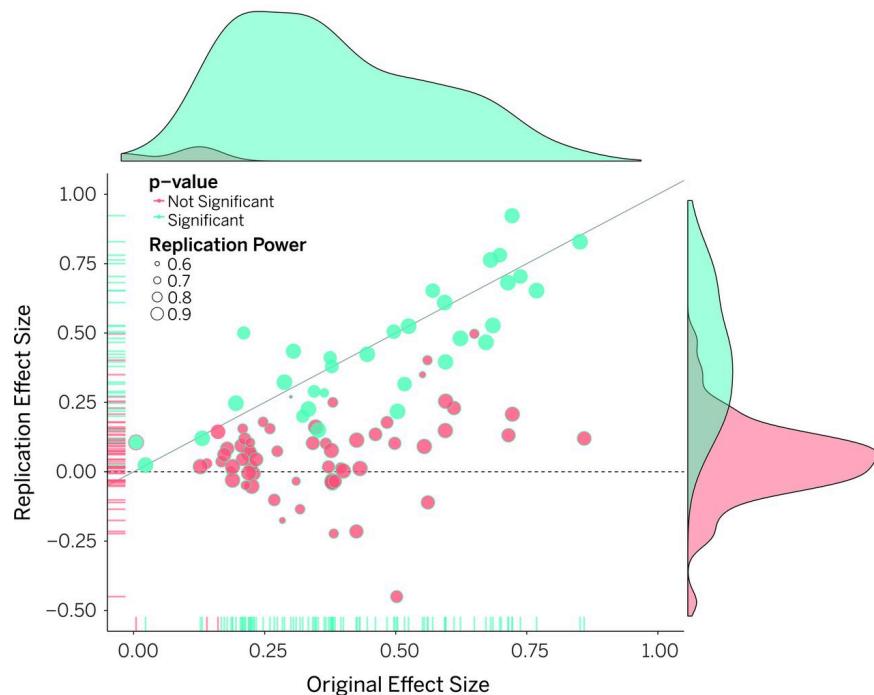
Reproducibility Crisis

“Most research findings are false for most research designs and for most fields”

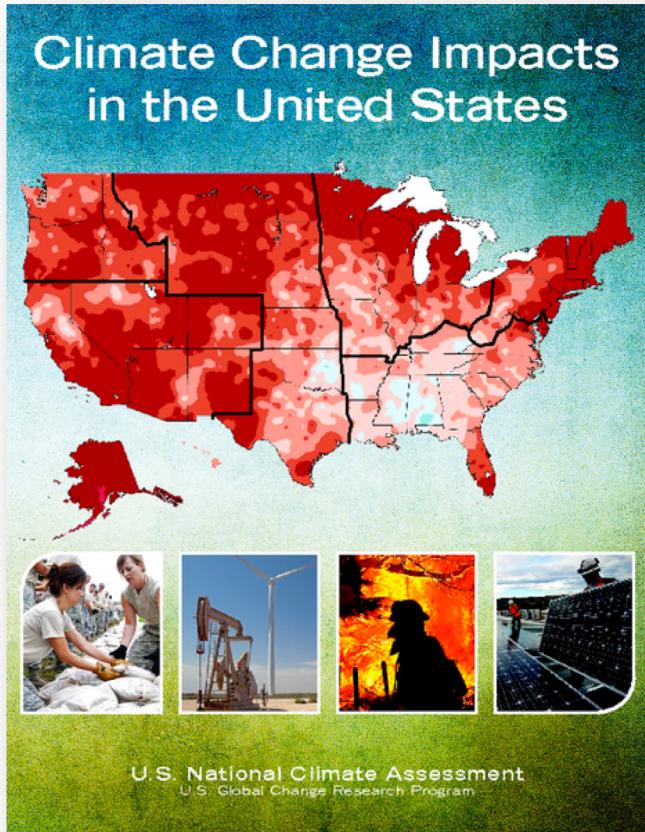
Ioannidis, 2005

“Most replication effects were smaller than original results”

Open Science Collaboration, 2015



National Climate Assessment

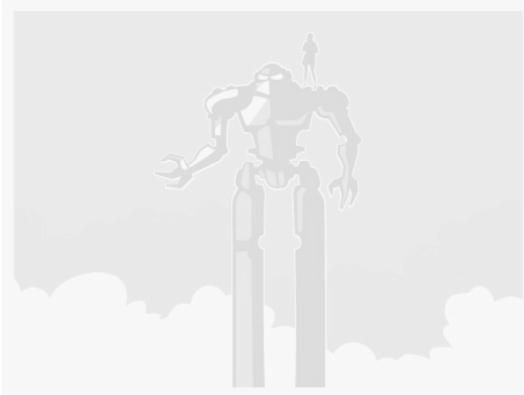


"This report is the result of a **three-year** analytical effort by a team of **over 300 experts**, overseen by a broadly constituted Federal Advisory Committee of **60 members**. It was developed from information and analyses gathered in over 70 workshops and listening sessions held across the country."

Computational Reproducibility

Facilitate transparency by
capturing and **communicating**
scientific workflows

Increase **trust in science**



Stand on the shoulders of giants
(build on work that came before)

Give credit for that **secondary**
usage enabling **easy attribution**

Practical Reproducibility



Preserve the data



Preserve the software workflow



Document what you did



Describe how to interpret it all



[Clear all filters](#)

Search ?

Search phrase



DATASETS 1 TO 25 OF 44

1 2 Next

Sort by Most recent

My Search

sasap



Jeanette Clark and Rich Brenner. 2017. [Sockeye salmon brood tables, northeastern Pacific, 1922-2016](#). Knowledge Network for Biocomplexity. urn:uuid:c11dff42-b988-437a-afee-58fc62dcd1dc.



Commercial Fisheries Entry Commission. 2018. [Commercial Fisheries Entry Commission Basic Information Table, 1975-2016](#). Knowledge Network for Biocomplexity. urn:uuid:8f351735-baf9-451a-b821-c1117ebf5a5e



Andrew Munro and Eric Volk. 2018. [Summary of Pacific Salmon Escapement Goals in Alaska with a Review of Escapements from 2001 to 2009](#). Knowledge Network for Biocomplexity. urn:uuid:d62539fd-3025-48d0-a1c3-5a903de1f269.



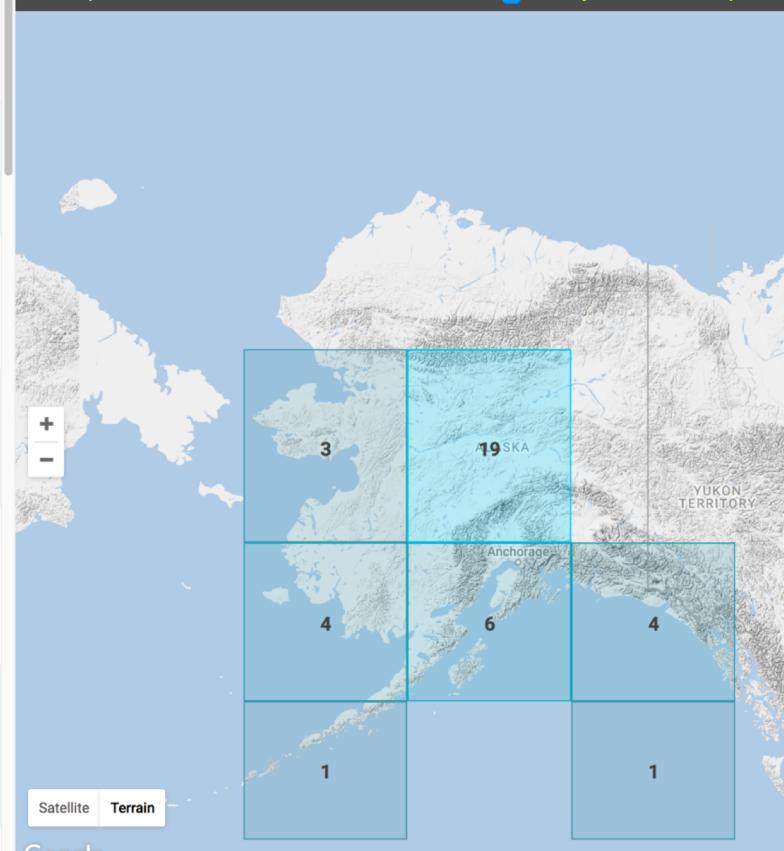
Alaska Department of Labor and Workforce Development, Research and Analysis Section. 2018. [Alaskan fishing industry employee counts by month, grouped by region and fish species from 2000-2016](#). Knowledge Network for Biocomplexity. urn:uuid:32958097-0ad3-428a-aba9-c37e804be0ef.



Alaska Department of Labor and Workforce Development Research & Analysis Section. 2018. [Alaskan fishing industry employee counts by month, subsetted by region and fish species](#). Knowledge Network for Biocomplexity. urn:uuid:4bbc9577-e81f-40f4-b4ca-9c740092bab



Commercial Fisheries Entry Commission. 2018. [Commercial Fisheries Entry Commission Permit Earnings, 1975-2016](#). Knowledge Network for Biocomplexity.

[Hide Map »](#) Limit my search to the map area

Google

Map data ©2018 Google, INEGI, SK telecom, ZENRIN | 500 km | Terms of Use



Global
Data Coverage



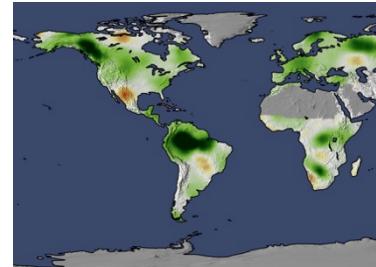
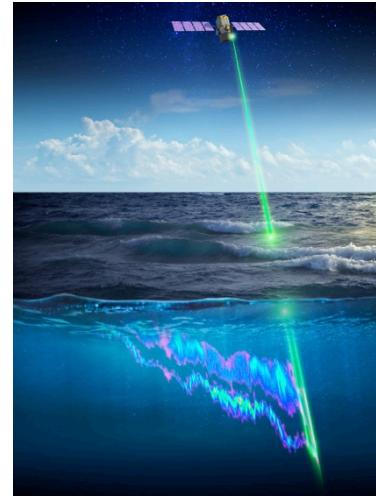
800K
Data Packages



40
Repositories



143K
Contributors





Reproducible
Science



Provenance



Citation

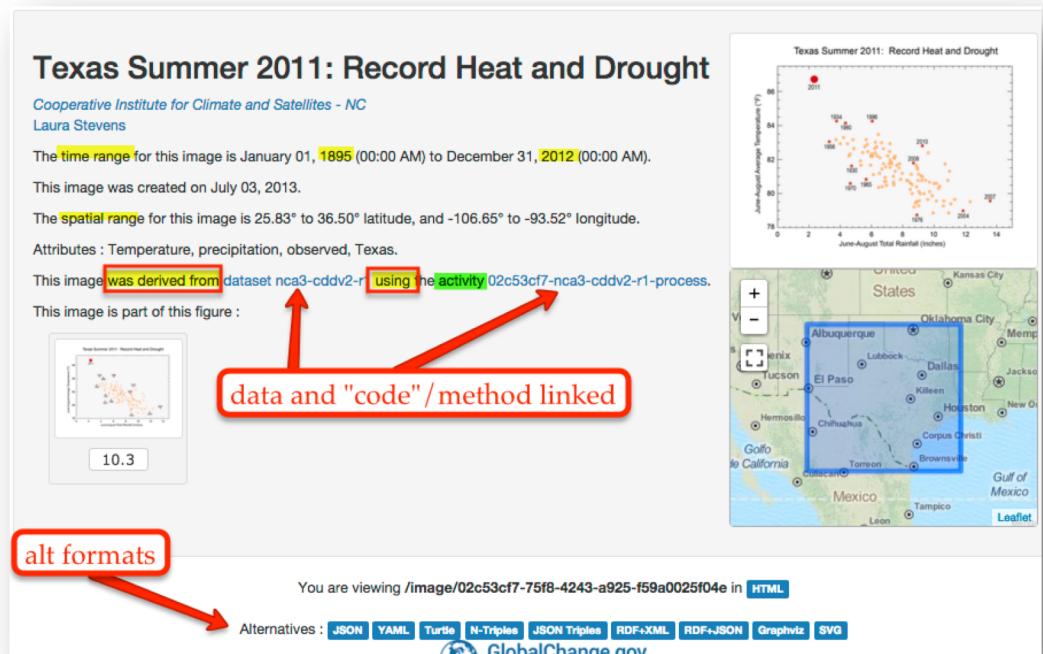


Synthesis

Computational Provenance

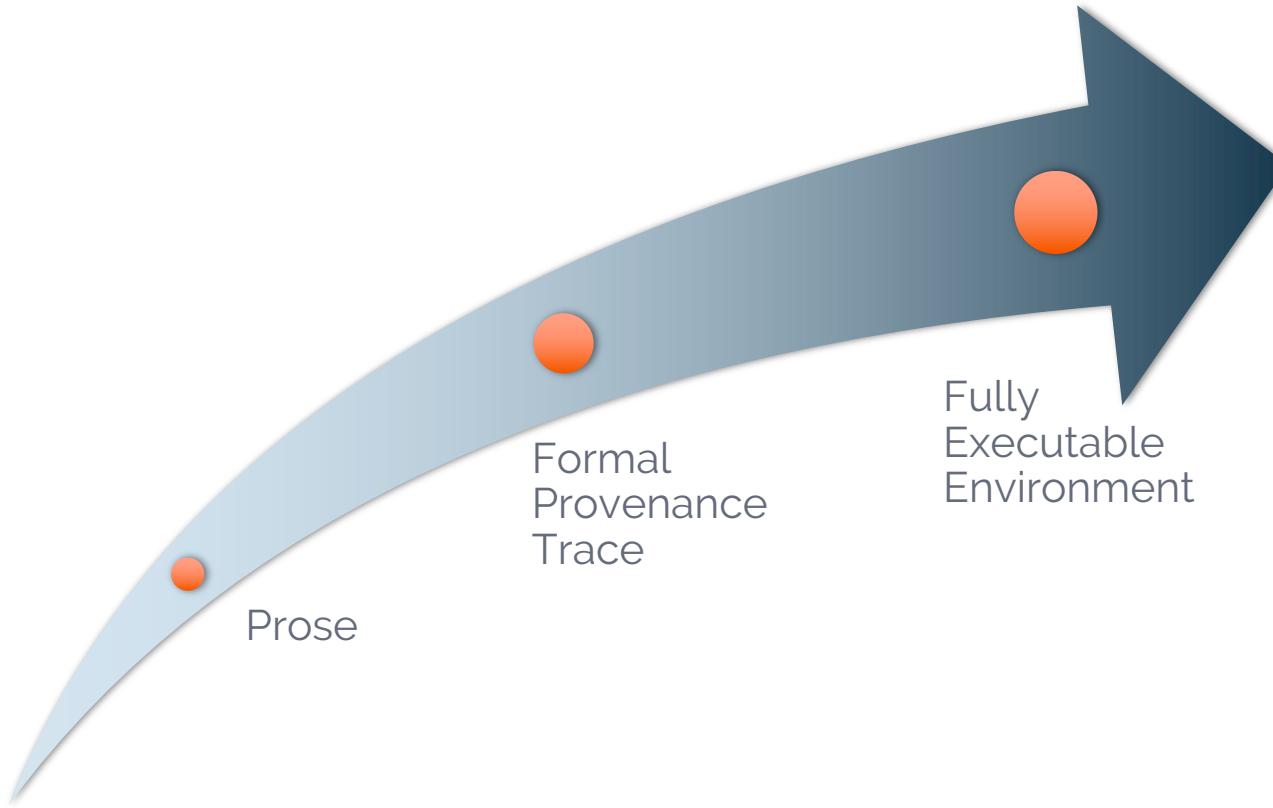
Origin, processing history of data

- Input data
- Workflow/scripts
- Output data
- Figures
- Understand methods, dataflow, and dependencies



Provenance

Origin and processing history of artifacts



Provenance in DataONE

Phase II Goal: Facilitate reproducible science

- Track **data derivation** history
- Track data **inputs** and **outputs** of analyses
- Track analysis and model **executions**
- Preserve and document software **workflows**
- Link all of these to **publications**

Provenance for Science Workflows



ProvONE – an extension of W3C PROV

See purl.dataone.org/provone-v1-dev

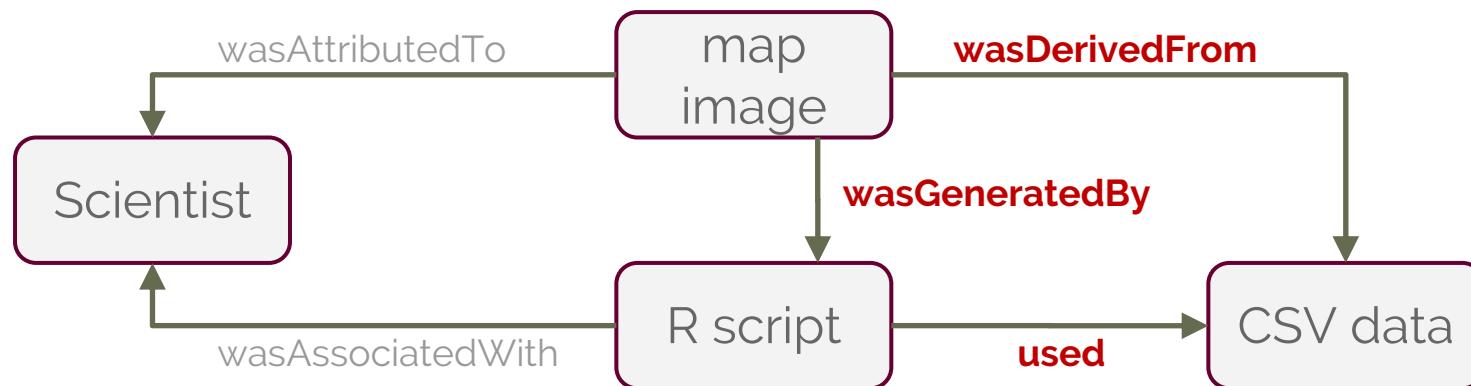


Provenance for Science Workflows

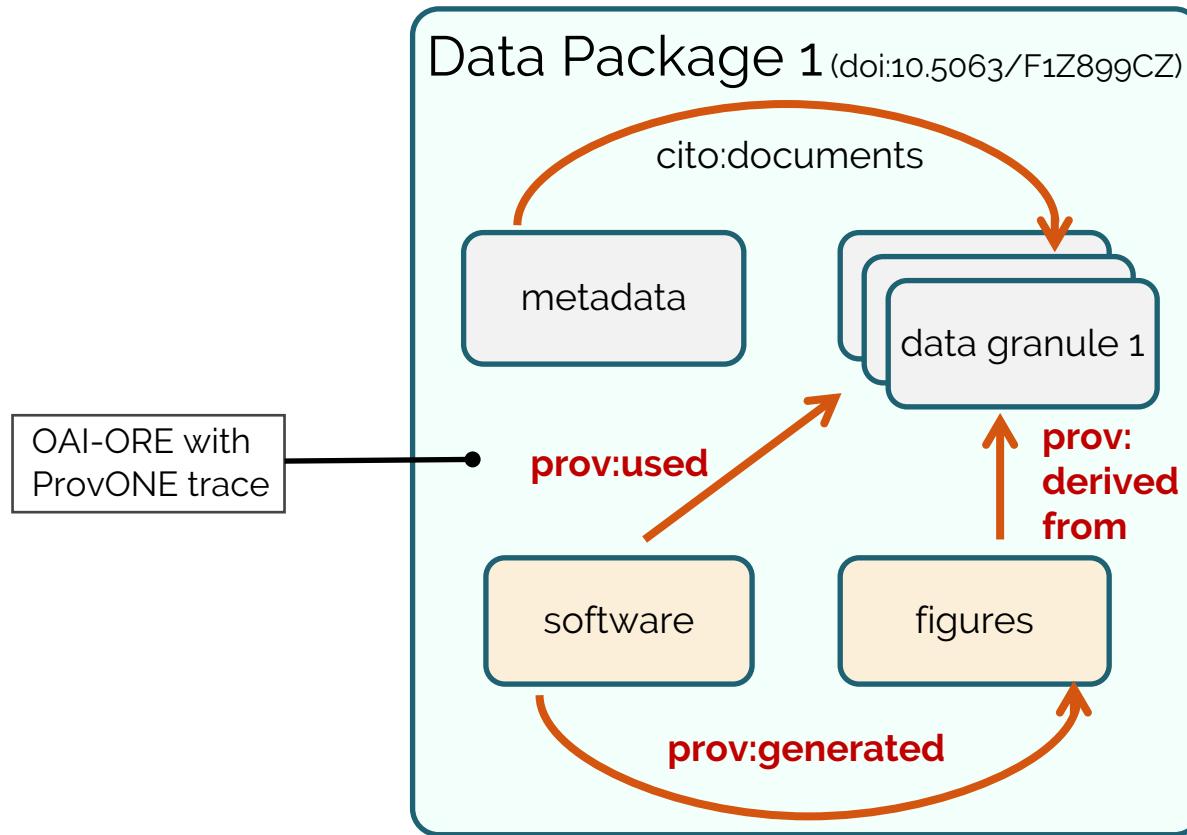


ProvONE – an extension of W3C PROV

See purl.dataone.org/provone-v1-dev

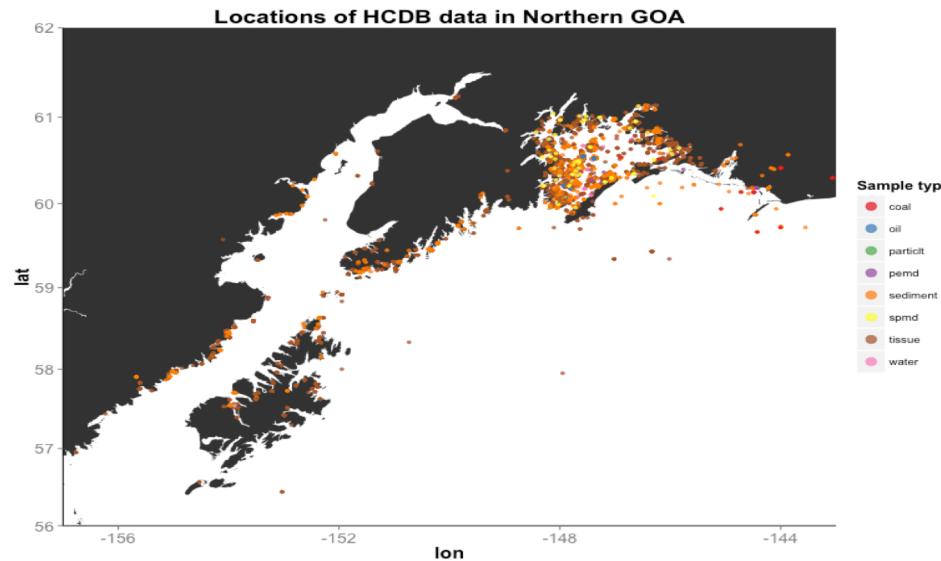


Data Package with Provenance



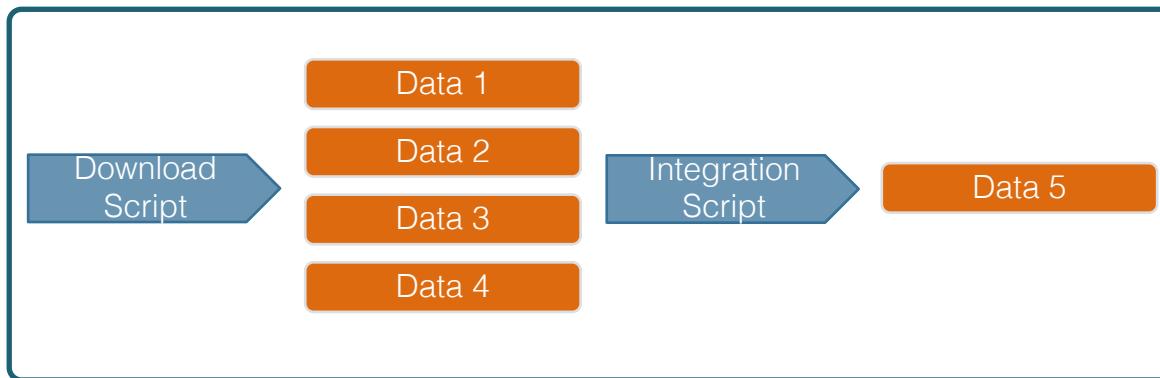
Hydrocarbon Data Example

Mark Carls. 2017. Analysis of hydrocarbons following the Exxon Valdez oil spill, Gulf of Alaska, 1989 - 2014. Arctic Data Center.

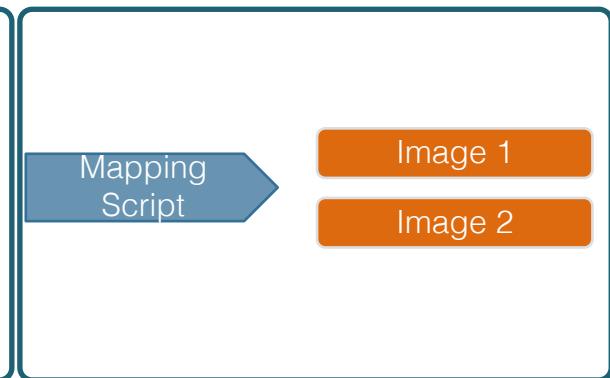


Publishing Data Workflows

Dataset C



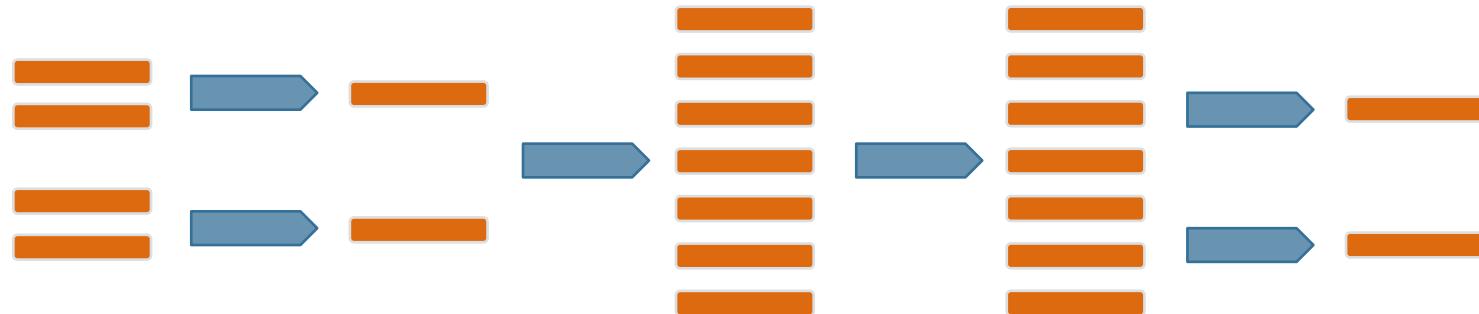
Dataset D



Hydrocarbon Data Example

Complex Workflows

Simplified view of complex workflows



Provenance Display

DataONE Search

About News Participate Resources Education Data

DATAONE SEARCH: [Search](#) [Summary](#)

Jump to: [DOI or ID](#) [Go](#)

[Sign in](#) or [Sign up](#)

[Back to search](#) | Search / Metadata

Mark Carls. 2017. Analysis of hydrocarbons following the Exxon Valdez oil spill, Gulf of Alaska, 1989 - 2014. Gulf of Alaska Data Portal. urn:uuid:3249ada0-afe3-4dd6-875e-0f7928a4c171.



[Copy Citation](#)

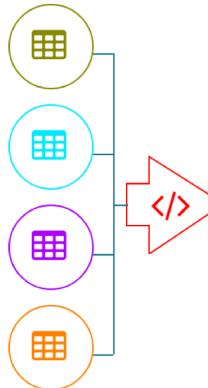
Files in this dataset Package: urn:uuid:1d23e155-3ef5-47c6-9612-027c80855e8d

Name	File type	Size	Download all
Metadata: metadata.xml	EML v2.1.1	140 KB	Download
Total_Aromatic_Alkanes_PWS.csv	text/csv	3 MB	Download
CollectionMethods.csv	text/csv	793 B	Download
Non-EVOS_SINs.csv	text/csv	3 KB	Download

[Show 8 more items in this data set](#)

Data Table, Image, and Other Data Details

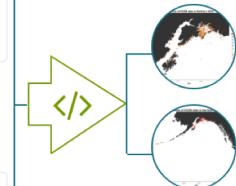
4 sources



Data Table

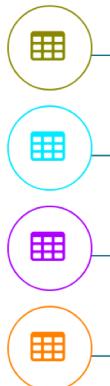
Entity Name	Total_Aromatic_Alkanes_PWS.csv										
	Download										
Description	Combined dataset from PAH, Alkane and Sample tables documenting samples collected after the Exxon Valdez oil spill in Prince William Sound, AK										
Object Name	Total_Aromatic_Alkanes_PWS.csv										
Online Distribution Info	https://cn.dataone.org/cn/v2/resolve/urn:uuid:44108e76-405d-4d58-b1b3-fb4b55e3fff9										
Size	2801033 byte										
Text Format	<table><tr><td>Number of Header Lines</td><td>1</td></tr><tr><td>Record Delimiter</td><td>#x0A</td></tr><tr><td>Attribute Orientation</td><td>column</td></tr><tr><td>Simple Text</td><td></td></tr><tr><td>Field Delimiter</td><td>,</td></tr></table>	Number of Header Lines	1	Record Delimiter	#x0A	Attribute Orientation	column	Simple Text		Field Delimiter	,
Number of Header Lines	1										
Record Delimiter	#x0A										
Attribute Orientation	column										
Simple Text											
Field Delimiter	,										
Number Of Records	12142										

2 derivations



Data Table, Image, and Other Data Details

4 sources



Source Program

Total_PAH_and_Alkanes_GoA_Hydrocarbons_Clean.R

Citation

[View »](#)

This program generated the data you are currently viewing, [Total_Aromatic_Alkanes_PWS.csv](#).

This program used [PAH.csv](#), [Sample.csv](#), [Non-EVOS_SINs.csv](#) and (and 1 more .

Alkanes_PWS.csv

from PAH, Alkane and Sample tables documenting samples collected after the oil spill in Prince William Sound, AK

Alkanes_PWS.csv

<https://doi.org/cn/v2/resolve/urn:uuid:44108e76-405d-4d58-b1b3-fb4b55e3fff9>

Text Format

Number of Header Lines

1

Record Delimiter

#x0A

Attribute Orientation

column

Simple Text

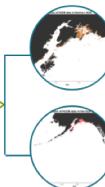
Field Delimiter

,

Number Of Records

12142

2 derivations



Web Provenance Editor

Deployed by Arctic Data Center

The screenshot shows the NSF Arctic Data Center's Web Provenance Editor interface. At the top, there is a navigation bar with links for Data, Support, About, and a green 'Submit Data' button. A user profile for 'Christopher Jones' is also visible. Below the navigation bar, the main content area is titled 'Data Table, Image, and Other Data Details'. It displays a single data table entry for 'Total_Aromatic_Alkanes_PWS.csv'. The table includes fields for Entity Name, Description, Object Name, Online Distribution Info, Size, and Text Format. There are also sections for '0 sources' and '0 derivations' with 'Add' buttons. The background features a blue banner with a snowy mountain scene.

Data Table					
Entity Name	Total_Aromatic_Alkanes_PWS.csv				
Description	Combined dataset from PAH, Alkane and Sample tables documenting samples collected after the Exxon Valdez oil spill in Prince William Sound, AK				
Object Name	Total_Aromatic_Alkanes_PWS.csv				
Online Distribution Info	https://cn-stage.test.dataone.org/cn/v2/resolve/urn:uuid:df984766-dd89-4e57-b97e-350506d7007e				
Size	2801033 byte				
Text Format	<table><tr><td>Number of Header Lines</td><td>1</td></tr><tr><td>Record Delimiter</td><td>#x0A</td></tr></table>	Number of Header Lines	1	Record Delimiter	#x0A
Number of Header Lines	1				
Record Delimiter	#x0A				



Reproducible
Science



Provenance



Citation



Synthesis

Credit where credit is due

Indexing and exposing data citations in international data repository networks



ALFRED P. SLOAN
FOUNDATION



University of California
CDL
California Digital Library



Force11 Data Citation Principles

1. Importance of data citation
2. **Credit and Attribution**
3. **Evidence**
4. Unique Identification
5. Access
6. **Persistence**
7. **Specificity** and Verifiability
8. Interoperability and Flexibility

Transitive Credit

When a user cites a pub, we know:

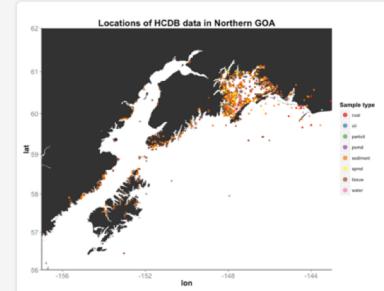
- **Which data** produced it
- **What software** produced it
- What was **derived** from it
- **Who to credit** down the attribution stack

See: Katz & Smith. 2014. **Implementing Transitive Credit with JSON-LD.**
arXiv:1407.51

Derived image

Map of sampling locations in the Northern Gulf of Alaska

Citation
Mark Carls. 2015. **Hydrocarbon database, Gulf of Alaska.** MN
Demo 2. urn:uuid:bf71c38b-22b2-469e-8983-734ec0ab19cb.

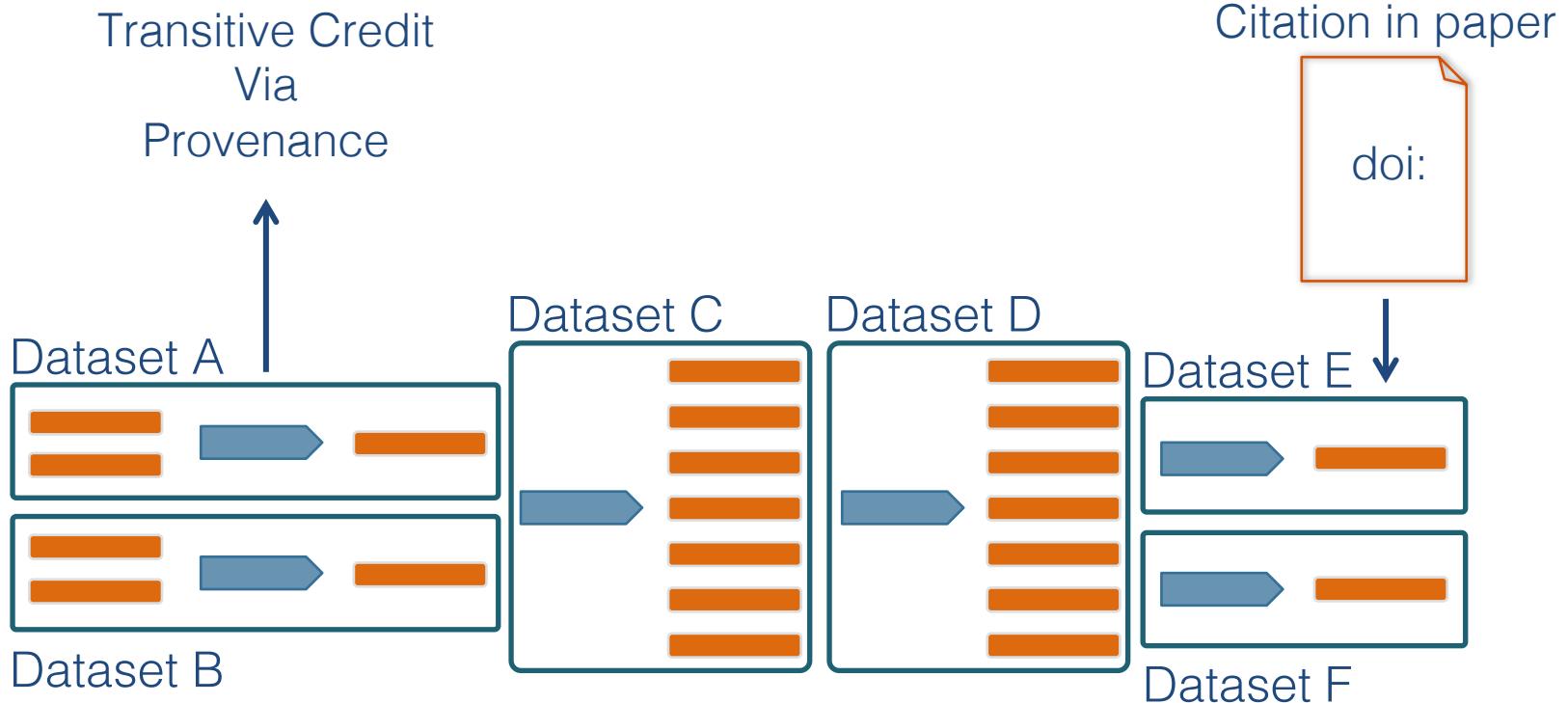


View »

This image was generated by the program you are currently viewing, [Locations map R script](#).

This image was derived from [Total_Aromatic_Alkanes_PWS.csv](#).

Citing multi-generational workflows



Evolution of the Living Paper



Scholarly Publications

1st Gen

Prose

2nd Gen

Prose

+ Data

3rd Gen

Prose

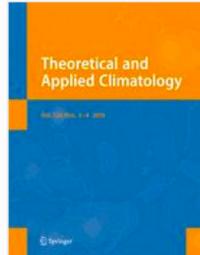
+ Data

+ Code

Prose + Data + Code + Provenance

Prose + Data + Code + Provenance + Execution Environment





[Theoretical and Applied Climatology](#)

... November 2016, Volume 126, [Issue 3–4](#), pp 699–703 | [Cite as](#)

Learning from mistakes in climate research

Authors

Authors and affiliations

Rasmus E. Benestad , Dana Nuccitelli, Stephan Lewandowsky, Katharine Hayhoe, Hans Olav Hygen, Rob van Dorland,

John Cook

Open Access | Original Paper

First Online: 20 August 2015

3.2k

103k

18

Shares

Downloads

Citations



replicationDemos

help

Meta

demo

html

R

replicationDemos.rdb

replicationDemos.rdx

replicationDemos

data

Rdata.rdx

Rdata.rdb

Rdata.rds

INDEX

NAMESPACE

DESCRIPTION

Ships with an R package



Edzer Pebesma

@edzerpebesma

Follow

Replying to @jhollist @metamattj

It is on CRAN, but in Archived; I could install it after installing a bunch of other Archived packages from source, and could run a number of examples. Another number depended on web resources no longer available.

5:04 AM - 14 Jul 2019



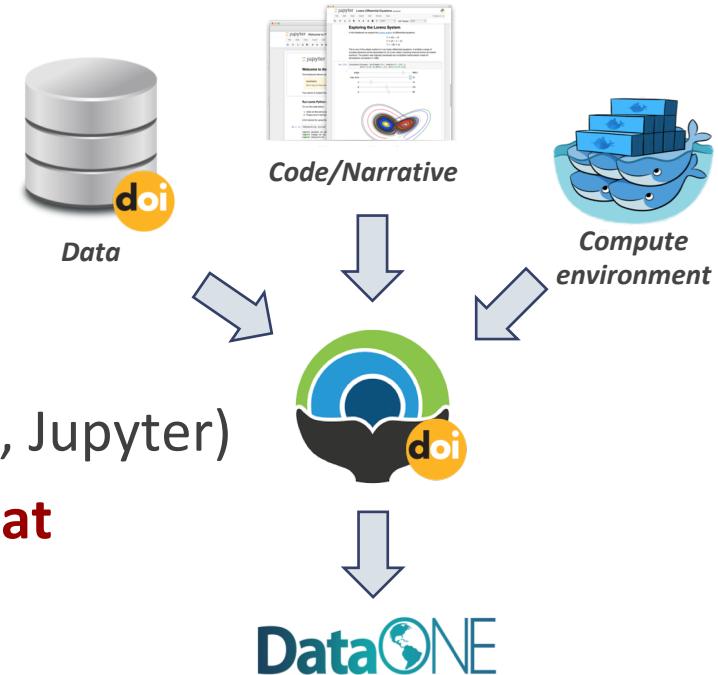
Parsing Reproducibility

- **Empirical Reproducibility:**
 - traditional empirical experiments, e.g. at the bench/lab
- **Statistical Reproducibility:**
 - statistical methodology used permits generalizability of data inferences
- **Computational Reproducibility:**
 - transparency of computational steps that produce scientific findings



What exactly is (in) a Tale?

- Tale = executable **research object**, i.e.
 - **data** (references)
 - + **code** (computational methods)
 - + **narrative** (traditional science story)
 - + **compute environment** (e.g. RStudio, Jupyter)
- Captured in a **standards-based tale format** complete with metadata





Browse Tales Launch to add to Launched Tales list

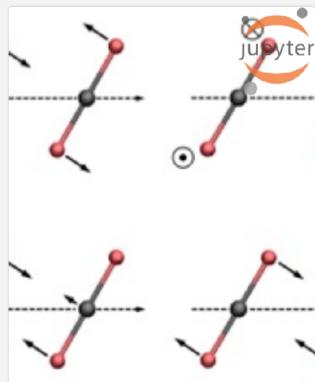


Search tales...



All

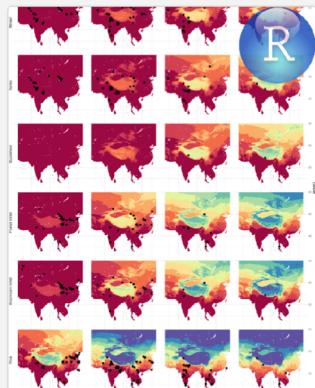
Switch to list view



COMPUTATIONAL CHEMISTRY

Anharmonic vibrational structure of...

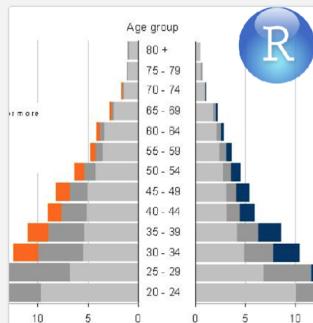
This project produces all of the data from the Anharmonic vibrational structure of the carbon dioxide dimer with a many-body potential energy surface journal article. The project solves the vibrational Schrodinger equation for the CO₂ monomer and dimer.



ARCHAEOLOGY

Climate change stimulated agricultu...

Ancient farmers experienced climate change at the local level through variations in the yields of their staple crops. However, archaeologists have had difficulty in determining where, when, and how changes in climate affected ancient farmers. We



ECONOMICS

L2-Boosting for Economic Applicatio...

Replication package for: L₂-Boosting for Economic Applications

The authors present the L₂-Boosting algorithm and two variants, namely post-Boosting and orthogonal Boosting. Building on work by M. and G. (2010), they

Launched Tales



L2-Boosting for Economic Application...

Browse Existing Tales ...





Compose

Create a new Tale by pairing a compute environment with a dataset



Tale name:

L2-Boosting for Economic Applications

Compute environment:



RStudio (rocker/geospatial)

Input data:

[Launch New Tale](#)

... Compose New Tales ...

Environments

Search compute environments...



RStudio (rocker/geospatial)



Jupyter Classic



RStudio



Jupyter Lab



L2-Boosting for Economic Application...
Ye Luo and Martin Spindler

Interact Files Metadata



File Edit Code View Plots Session Build Debug Profile Tools Help

Install.R x Sim_AER_V3.R x Source on Save Run Source

```

1: #####
2: # L2-Boosting for Economic Applications
3: #####
4: # Parameter for simulation study
5: rm(list=ls())
6: source("DGP.R")
7: source("helper.R")
8: R <- 500 # number of repitions
9: set.seed(12345)
10: library(MASS)
11: library(mvtnorm)
12: library(hdm)
13: library(newboost) # can be downloaded from R-Forge or requested by the a
14: #####
15: # IV Estimation

1:1: (Top Level) : R Script

Console Terminal
/WholeTale/workspace/
```

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

```

> load("./WholeTale/workspace/Sim_AER.RData")
> |
```

rstudio Project: (None)

Environment History Connections Jobs

Global Environment -

Data

data	List of 3
ds	num [1:90, 1] -1.24 -0.974 1.33 -0.154 -0...
ED	List of 6
ED1	List of 6
EDB	List of 6

Files Plots Packages Help Viewer

New Folder Upload Delete Rename More

WholeTale : workspace

Name	Size	Modified
apt.txt	5 B	Mar 6, 2019, 1:43 PM
DGP.R	1.5 KB	Mar 5, 2019, 3:36 PM
helper.R	9.2 KB	Mar 5, 2019, 3:36 PM
install.R	148 B	Mar 5, 2019, 3:36 PM
Readme.pdf	60.7 KB	Mar 5, 2019, 3:36 PM
runtime.txt	13 B	Mar 5, 2019, 3:36 PM
Sim_AER.RData	6.6 MB	Mar 5, 2019, 4:14 PM
Sim_AER_V3.R	5.3 KB	Mar 5, 2019, 3:46 PM

Launched Tales

L2-Boosting for Economic Application...

Daniel White and Lilian Alessa. Humans and Hydrology at High Latitudes: Water Use Information. Arctic Data Center. doi:10.5065/D6862DM8.

15 Citations 0 **Downloads** 183 **Views** 72 **Copy Citation** **Analyze**

Choose an analysis environment to interactively explore this dataset online using Whole Tale.

RSStudio Jupyter Notebook

Files in this dataset Package: resource_map_doi:10.5065/D6862DM8

Name	File type	Size	Action
Metadata: science_metadata.xml	EML v2.1.1	8 KB	65 views Download
estimated_use_of_water_in_US_2000.pdf	PDF	6 MB	6 downloads Download
estimated_use_of_water_in_US_2005.pdf	PDF	5 MB	5 downloads Download
first_nations_canada_water_and_wastewater_systems.pdf	PDF	365 KB	4 downloads Download

Show 13 more items in this data set

General

Identifier doi:10.5065/D6862DM8

Search About User Guide Support Sign Up Log In

AJPS AMERICAN JOURNAL of POLITICAL SCIENCE

political Science (AJPS) Dataverse (Midwest Political Science Association) aips.org

Journal of Political Science (AJPS) Dataverse > Replication Data for: Greater Expectations: A Field Experiment to Improve Accountability in Mali

downloads Contact Share

Data for: Greater Expectations: A Field Experiment to Improve Accountability in Mali [Version 3.0]

Replication Data for: Greater Expectations: A Field Experiment to Improve Accountability in Mali [DOI:10.5065/D6862DM8, Harvard Dataverse, V3, UNF:6:CCPqASJQMfVBr0+Hn/PMIQ== [fileUNF]]

Analyze in WT Explore Cite Dataset Learn about Data Citation Standards

I argue that if citizens systematically underestimate what their government can and should do for them, then they will hold politicians to a lower standard and sanction poor performers less often. A field experiment across 85 localities in Mali in which randomly assigned localities receive a civics course identifies the effect of raising voter expectations of government on their willingness to hold leaders accountable. The course provides information about local government capacity and responsibility as well as how local politicians perform relative to others, effectively raising voter expectations of what local governments can and should do. Survey experiments among individuals in treated and control communities (N=5,560) suggest that people in treated villages are indeed more likely to sanction poor performers and vote based on performance more often. A behavioral outcome – the likelihood that villagers challenge local leaders at a town hall meeting – adds external validity to survey findings.

Social Sciences
Government accountability, Voting behavior, Field experiments

Gottlieb, Jessica. 2016. "Greater Expectations: A Field Experiment to Improve Accountability in Mali." *American Journal of Political Science* 60 (1): 143-157. doi: 10.1111/ajps.12186

Search this dataset Find

https://nla.nla.gov.au/nla/nlaobj/348592595/dataset/harvard.edu%2Fdataset.xhtml%3Epersistentid%3DdmvSA1079109%2FJVNP%2E200408&name=Replication+Data+for+3A+Greater+Expectations%3A+A+Field+Experiment+to+Improve+Accountability+in+Mali



*... Integrate Data
Repos with
Whole Tale!*

- Enables **turnkey exploratory data analysis** on existing published datasets
- **DataONE and Dataverse networks cover > 90 major research repositories!**



Publish Tale

Publishing will create an immutable copy of your Tale with a DOI. [i](#)

This process will allow another user to easily rerun your published analysis using the WholeTale platform.

Please choose a target repository: *

DataONE-The Knowledge Network for Biocomplexity

[More Details ▾](#)

Your published Tale will include everything that has been uploaded to its associated workspace.

The following required files will be generated and published along with the Tale itself:

- Quantifying FAIR: metadata improvement and guidance in the DataONE repository network
 - manifest.json [i](#)
 - environment.json [i](#)
 - LICENSE [i](#)
 - README.md [i](#)
 - metadata.xml [i](#)

This process will allow another user to easily rerun your published analysis using the WholeTale platform.

For more information about publishing, please consult the [Publishing Guide](#).

[Cancel](#)

[Publish](#) ✓

*... Publish Data,
Code, and
Environment*

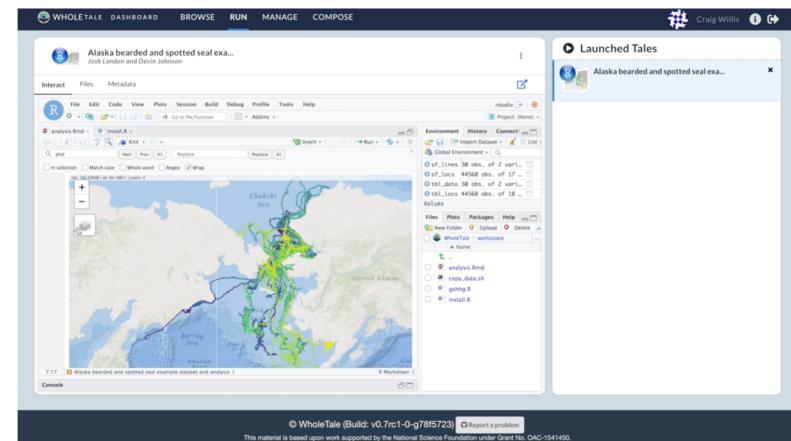
- Enables **full circle** reproducibility to **DataONE** repositories that accept API deposits



Whole Tale Forecast Demo

Demonstration of a model to predict the movement paths of seals using satellite telemetry data.

Based on analysis and models by:
Josh London and Devin Johnson
NOAA Marine Mammal Laboratory



<https://youtu.be/MI5d7r5OtCk>



Reproducible
Science



Provenance



Citation



Synthesis



NCEAS

National Center for Ecological Analysis and Synthesis

State of Alaska's Salmon and People

8 SASAP working groups

1: Bio-physical State of Knowledge of Salmon Distribution & Habitat

Leads: Peter Westley and Dan Rinella

2: Sociocultural and Economic Dimensions of Salmon Systems

Leads: Courtney Carothers, Jessica Black, Tobias Schworer

3: Governance and Subsistence

Leads: Steve Langdon, Taylor Brelsford, James Fall

4: Consistency, Causes, and Consequences of Declining Size and Age of Alaskan Salmon

Leads: Eric P. Palkovacs, Peter Westley, Bert Lewis

5: Well-Being and Alaska Salmon Systems

Leads: Rachel Donkersloot, Jessica C. Black, Courtney Carothers

6: Interacting Effects of Ocean Climate and At-Sea Competition on Alaskan Salmon

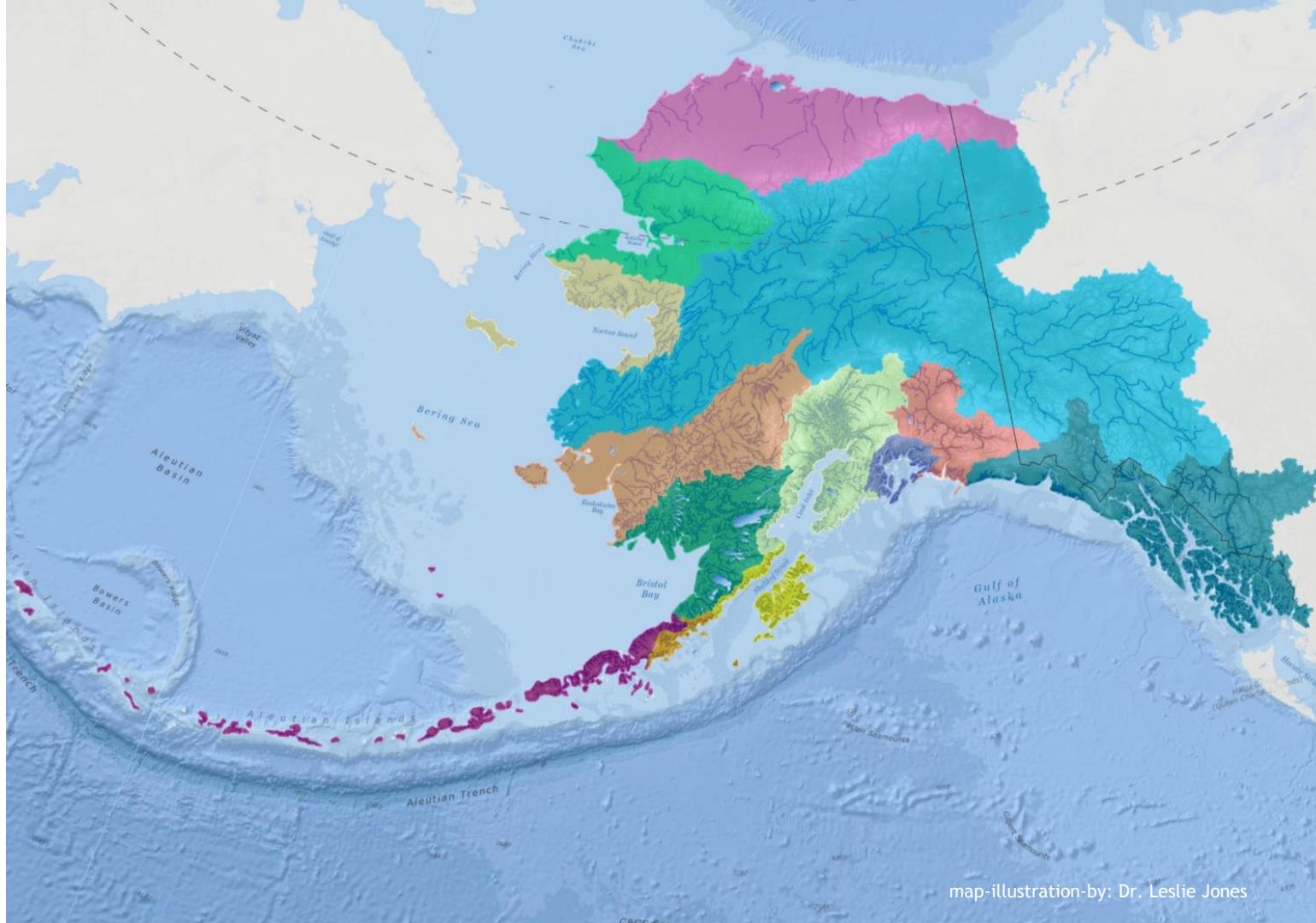
Leads: Peter S. Rand, Robert W. Campbell, Kristen B. Gorman

7: Using Participatory Modeling to Empower Community Engagement in Salmon Science

Leads: Michael L. Jones

8: Kenai Lowlands Salmon Research Synthesis and Design Tools for Integrated Watershed Management

Leads: Coowe Walker, Mark Rains, Ryan King, Charles Simenstad, Dennis Whigham



map-illustration-by: Dr. Leslie Jones

h

| Home / Search / Metadata

Jeanette Clark and Rich Brenner. 2017. Sockeye salmon brood tables, northeastern Pacific, 1922-2016. Knowledge Network for Biocomplexity. urn:uuid:c11dff42-b988-437a-afee-58fc62dcd1dc.

[Copy Citation](#)[Quality report](#)

Files in this dataset Package: resource_map_urn:uuid:c11dff42-b988-437a-afee-58fc62dcd1dc

Name	File type	Size	Downloads	Download All
Metadata: broodTable_metadata.xml	EML v2.1.1	37 KB	5 views	Download
BroodTables.csv	More info	text/csv	449 KB	61 downloads
StockInfo.csv	More info	text/csv	19 KB	2 downloads
SourceInfo.csv	More info	text/csv	723 B	2 downloads
broodTableProcessing.Rmd	More info	application/R	19 KB	3 downloads
broodTableProcessing.html	More info	HTML	1 MB	9 downloads

[▲ Show less](#)

30 inputs

Other Entity

1 outputs



view more ▾

Entity Name **broodTableProcessing.Rmd**

Download

Data Object Type:

Other

Physical Structure Description:

Object Name **broodTableProcessing.Rmd**

Source Data

urn:uuid:514f65fa-7f6b-4276-b502-4f46834d309b

Citation

[View ▾](#)

This data prov_hasDerivations
BroodTables.csv.

This data was used by the program you
are currently viewing, </>
broodTableProcessing.Rmd.

This data was used as an input to
create **BroodTables.csv.**



287e7d4799c089a59fb180125e1
d By SHA1

ne

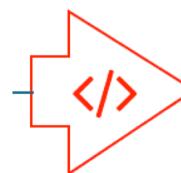
R

<https://dataone.org/cn/v2/resolve>
cd46e4-095b-4f25-918f-de

Python, Jupyter and Rmarkdown as Provenance

```
01-brood-table-integration.Rmd x
31
32 + ## Datasets
33
34 As part of the SASAP project, brood tables for 48 Sockeye salmon stocks were collected.
35 Table 2.1 shows a list of these stocks, along with other regional and location
36 information.
37
38 ````{r, echo = FALSE}
39 stocks <- read.csv('data/original/StockInfo.csv', stringsAsFactors = F)
40 ````{r, echo = FALSE}
41 datatable(stocks[, c('Stock.ID','Stock' , 'Region', 'Sub.Region')], rownames = FALSE,
42 caption = "Stock information")
43 ````{r, echo = FALSE}
44 These stocks range geographically from Washington to Alaska. Although temporal coverage
45 varies by stock, many of the brood tables were updated in 2016, and some have
46 reconstructions dating back to 1922.
47 Figure 2.1 indicates the approximate location of the salmon stocks in Table 2.1.
48 ````{r, echo = FALSE}
49 salmon = makeIcon('images/salmon_tiny.png',
50   'images/salmon_big.png',
51   26, 14)
52
53 m <- leaflet(stocks) %>%
54 setView(-median(stocks$Lon), median(stocks$Lat), zoom = 4) %>%
55 addTiles() %>%
56 addMarkers(~Lon, ~Lat, icon = salmon)
57
58 m
59
60 ````{r, echo = FALSE}
61 Figure 2.1: Location of stocks used in this data integration. Salmonid icon by Servien
62 (vectorized by T. Michael Keesey)
63 [CC-BY-SA](https://creativecommons.org/licenses/by-sa/3.0/), available at
64 [Phylopic](http://phylopic.org/)

37:72 R Markdown
```



2.2 Datasets

As part of the SASAP project, brood tables for 48 Sockeye salmon stocks were collected. Table 2.1 shows a list of these stocks, along with other regional and location information.

Stock.ID	Stock	Region	Sub.Region
101	Washington	WA	WA
102	E.Stuart	Fraser River	Fraser Early Stuart
103	Bowron	Fraser River	Fraser Early Summer
104	Fennell	Fraser River	Fraser Early Summer
105	Gates	Fraser River	Fraser Early Summer
106	Nadina	Fraser River	Fraser Early Summer
107	Pitt	Fraser River	Fraser Early Summer
108	Raft	Fraser River	Fraser Early Summer
109	Scotch	Fraser River	Fraser Early Summer
110	Seymour	Fraser River	Fraser Early Summer

Showing 1 to 10 of 54 entries Previous 1 2 3 4 5 6 Next

These stocks range geographically from Washington to Alaska. Although temporal coverage varies by stock, many of the brood tables were updated in 2016, and some have reconstructions dating back to 1922.

Figure 2.1 indicates the approximate location of the salmon stocks in Table 2.1.

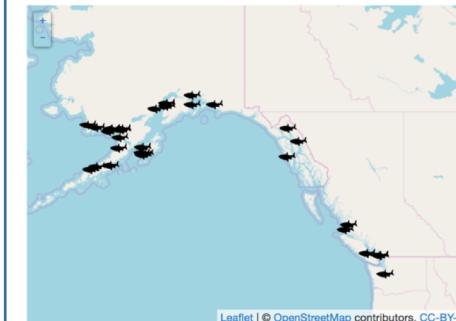


Figure 2.1: Location of stocks used in this data integration. Salmonid icon by Servien (vectorized by T. Michael Keesey)

Foundational Infrastructure

Providing ***findable, accessible*** data with ***interoperable*** infrastructure
enabling long term data ***reuse*** for synthesis

