

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>



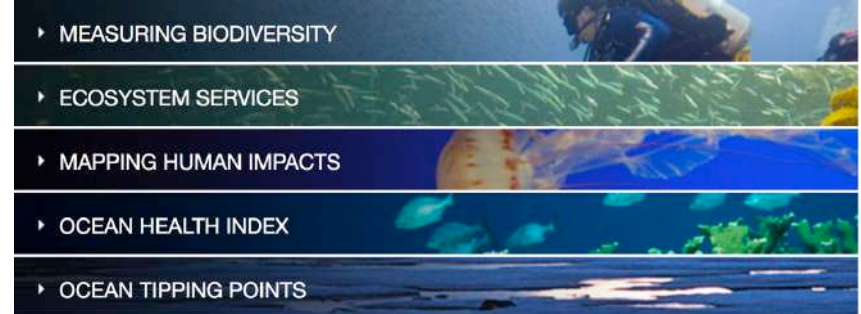
Marine Systems



Threats and Population Declines



Understanding Ocean Health



Climate and Ecosystems





Provenance



Citation



Synthesis

Reproducible Science



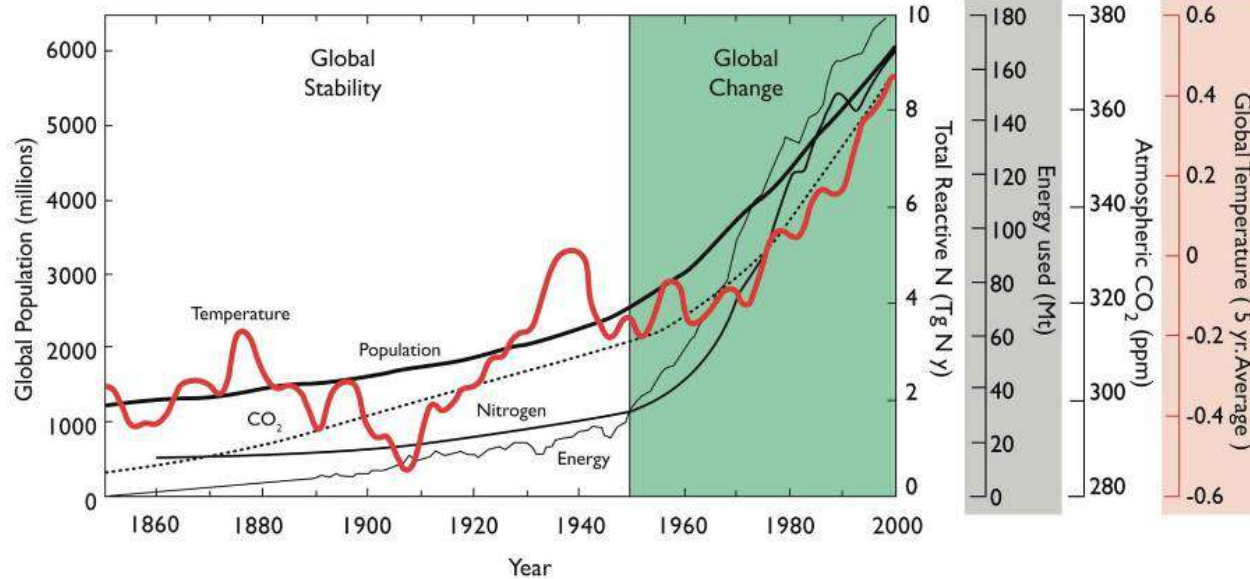
Climate Change
Fisheries
Sustainability
Subsistence



Science
Governance
Regulation
Policy



Trust in Science



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

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

Smith et al. (2009) Ecology doi:10.1890/08-1815.1

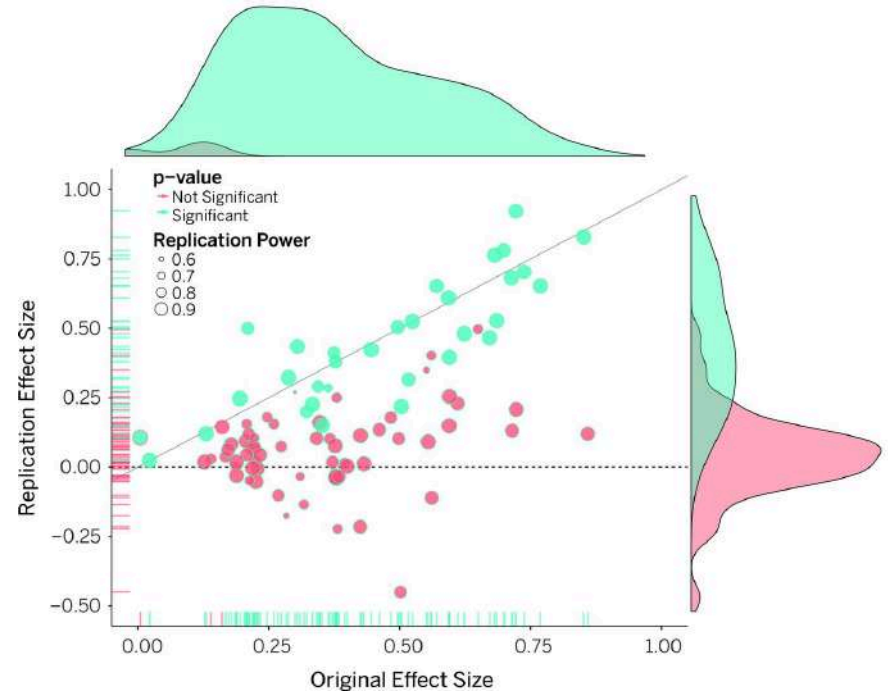
Reproducibility Crisis

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

Ioannidis, 2005

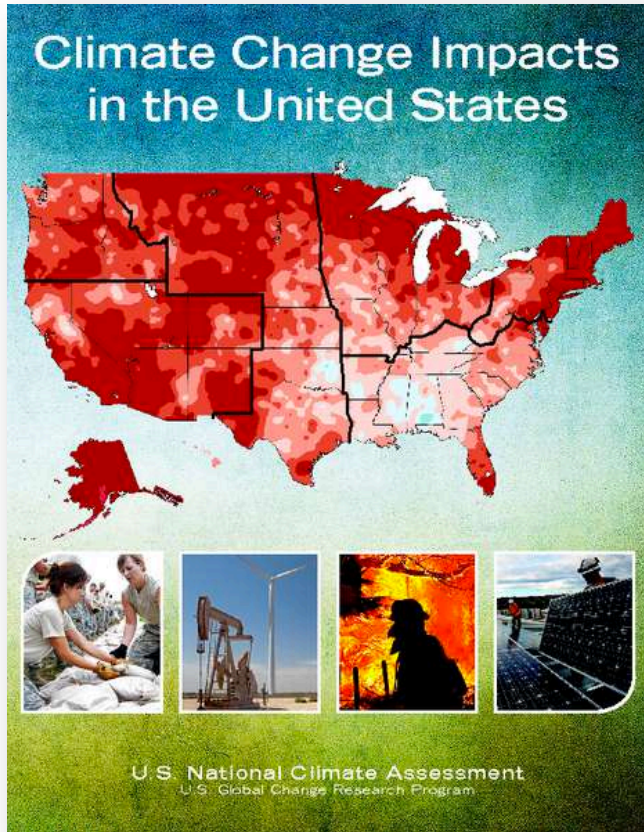
“Most replication effects were smaller than original results”

Open Science Collaboration, 2015



doi:10.1126/science.aac4716

National Climate Assessment

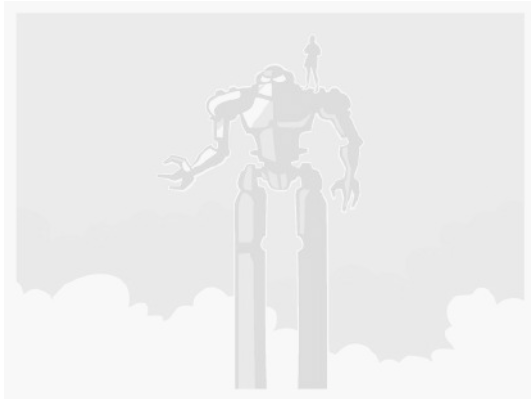


“This report is the result of a **three-year** analytical effort by a team of **over 300 experts**, overseen by a broadly constituted Federal Advisory Committee of **60 members**. It was developed from information and analyses gathered in over 70 workshops and listening sessions held across the country.”

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



My Search

sasap ✕

Filter by:

▶ Data attribute

▶ Data files

▶ Creator

▶ Year

▶ Identifier

▶ Taxon

▶ Location

DATASETS 1 TO 25 OF 44

1 2 Next

Sort by Most recent ▾



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.



5



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.



12



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.



10



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.



9



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-e811-40f4-b4ca-9c740092baba.

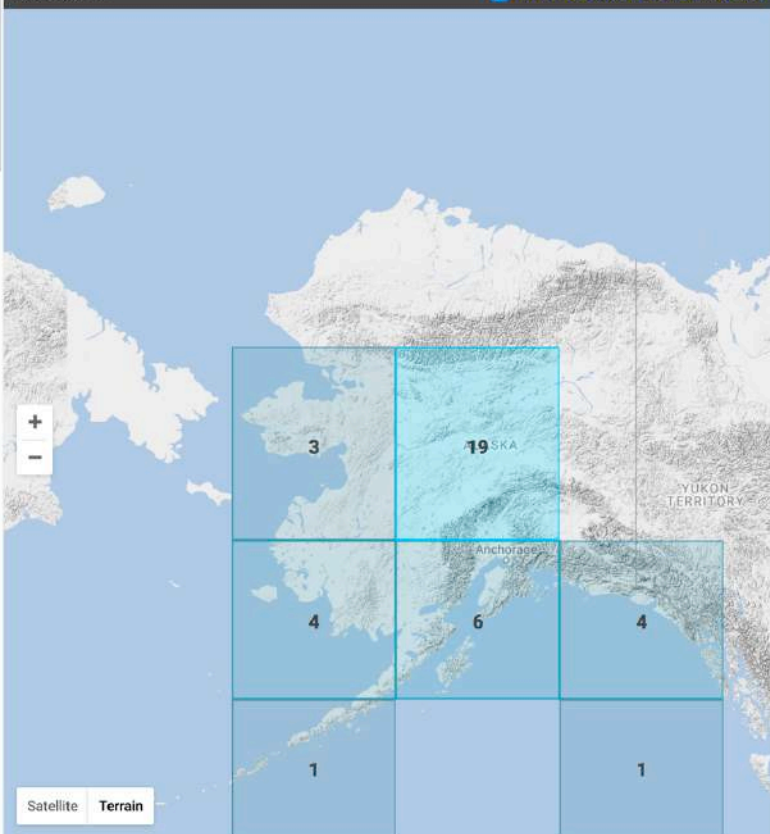


5



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

+

-

Satellite

Terrain

Google

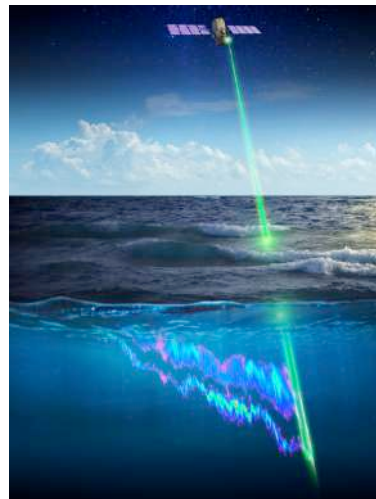
Map data ©2018 Google, INEGI, SK telecom, ZENRIN 500 km Terms of Use



Global
Data Coverage



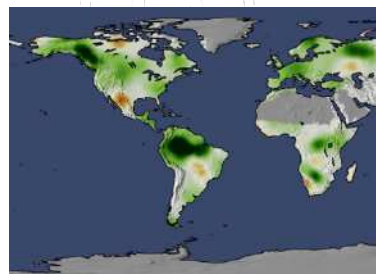
800K
Data Packages



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

Texas Summer 2011: Record Heat and Drought
Cooperative Institute for Climate and Satellites - NC
Laura Stevens

The **time range** for this image is January 01, **1895** (00:00 AM) to December 31, **2012** (00:00 AM).
This image was created on July 03, 2013.

The **spatial range** for this image is 25.83° to 35.50° latitude, and -105.85° to -83.52° longitude.

Attributes : Temperature, precipitation, observed, Texas.

This image **was derived from** dataset **nca3-cddv2-r1** **using** the activity **02c53cf7-nca3-cddv2-r1-process**.

This image is part of this figure :

alt formats

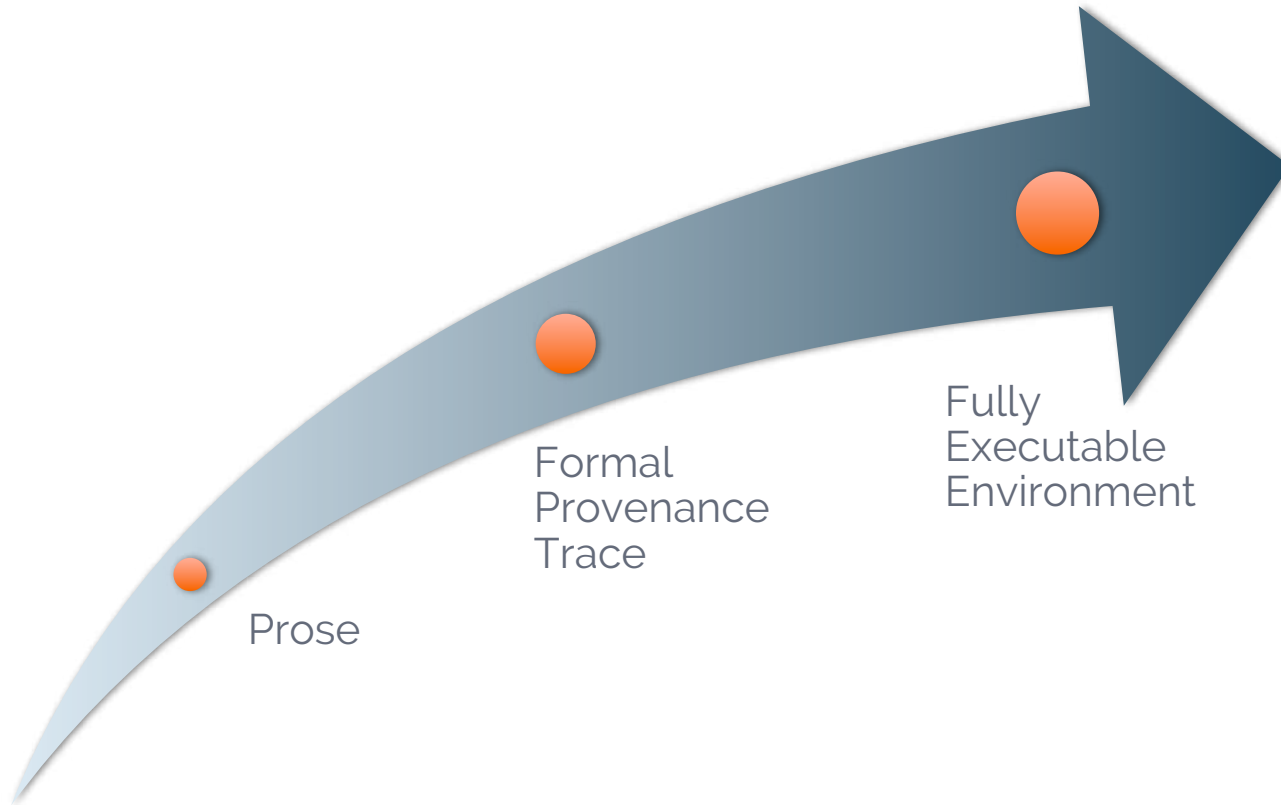
You are viewing /image/02c53cf7-75f8-4243-a925-f59a0025f04e in **HTML**.

Alternatives : [JSON](#) [YAML](#) [Turtle](#) [N-Triples](#) [JSON-Triples](#) [RDF+XML](#) [RDF+JSON](#) [Graphviz](#) [SVG](#)

GlobalChange.gov

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

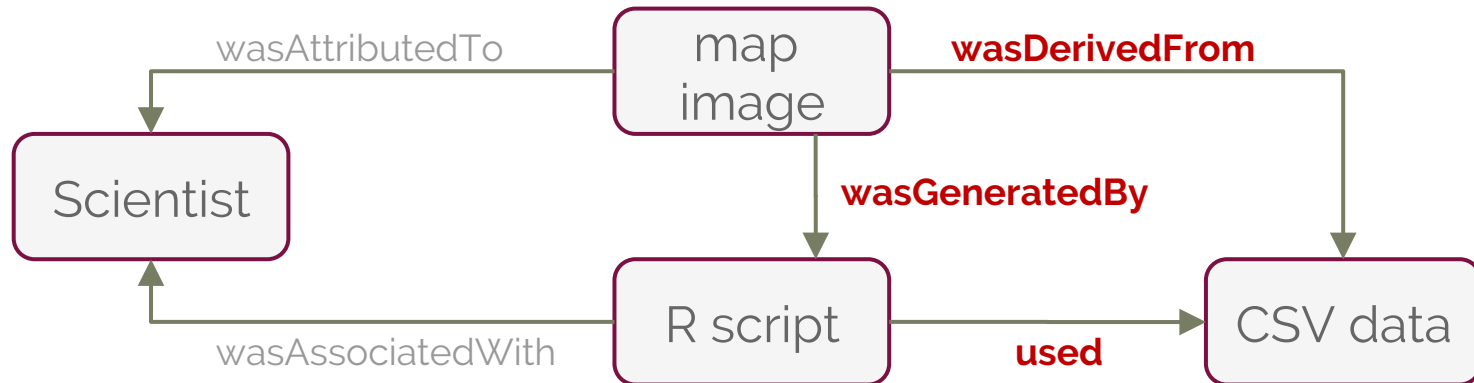


Provenance for Science Workflows

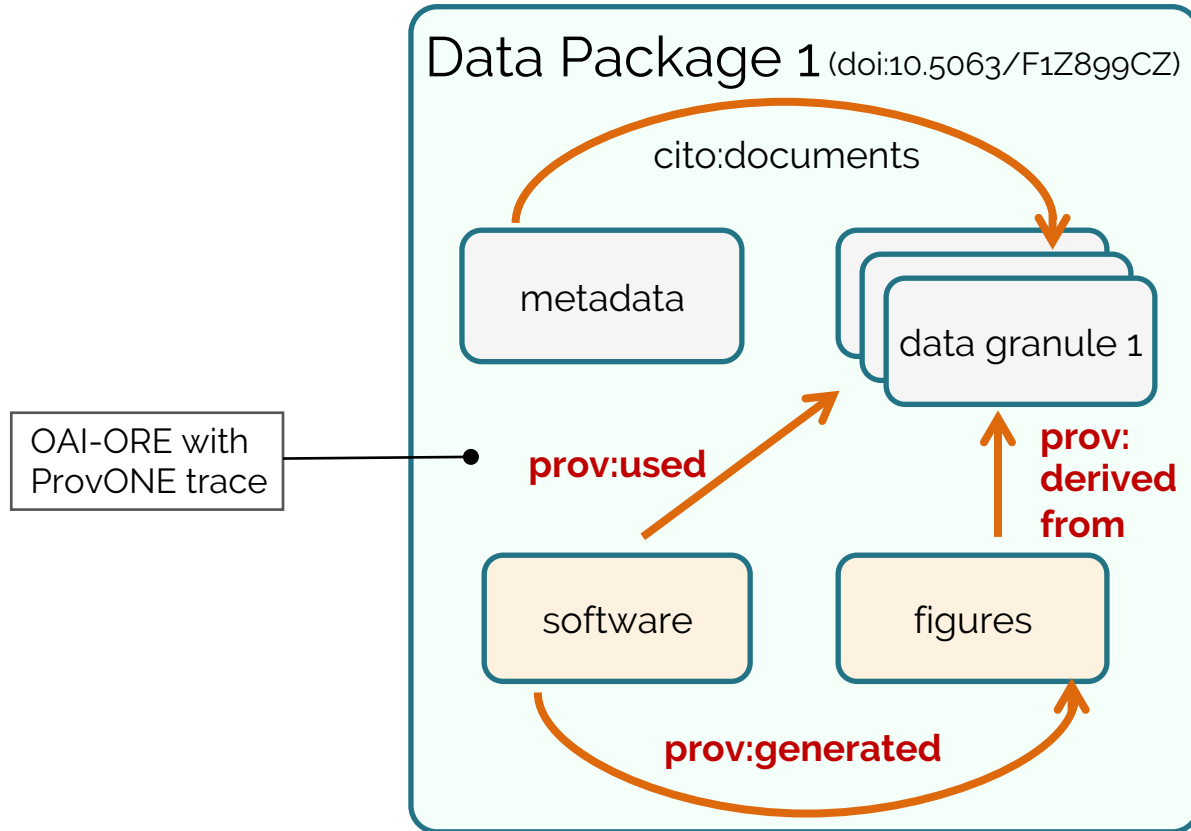


ProvONE – an extension of W3C PROV

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

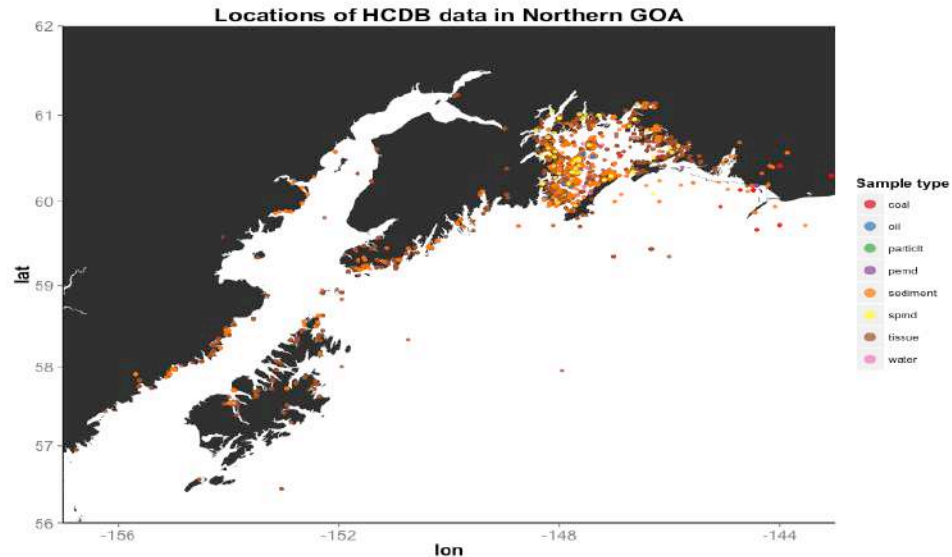


Data Package with Provenance



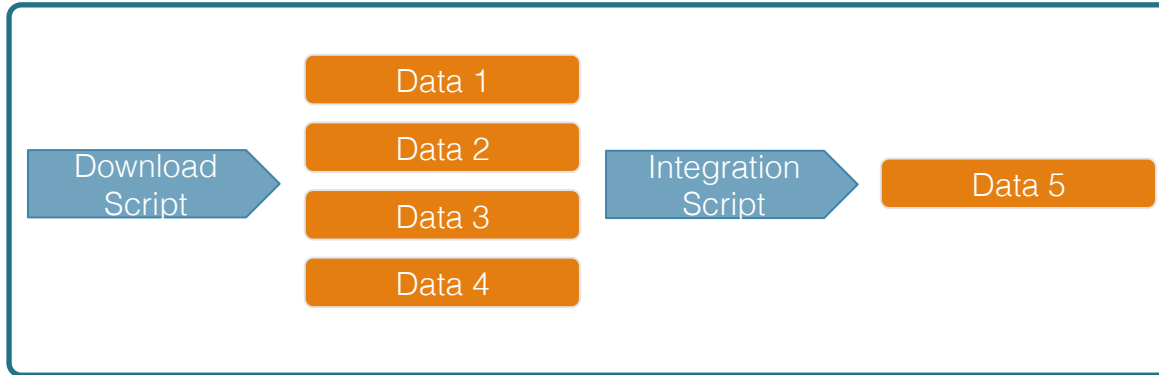
Hydrocarbon Data Example

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

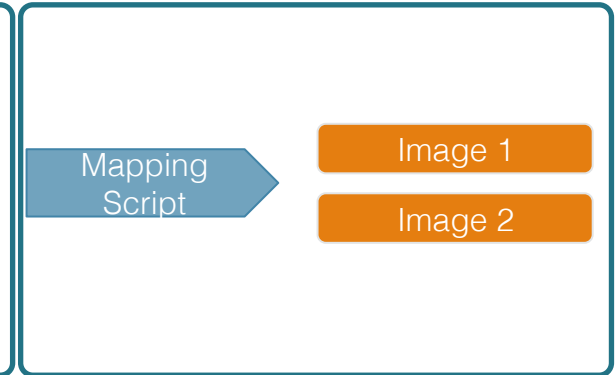


Publishing Data Workflows

Dataset C



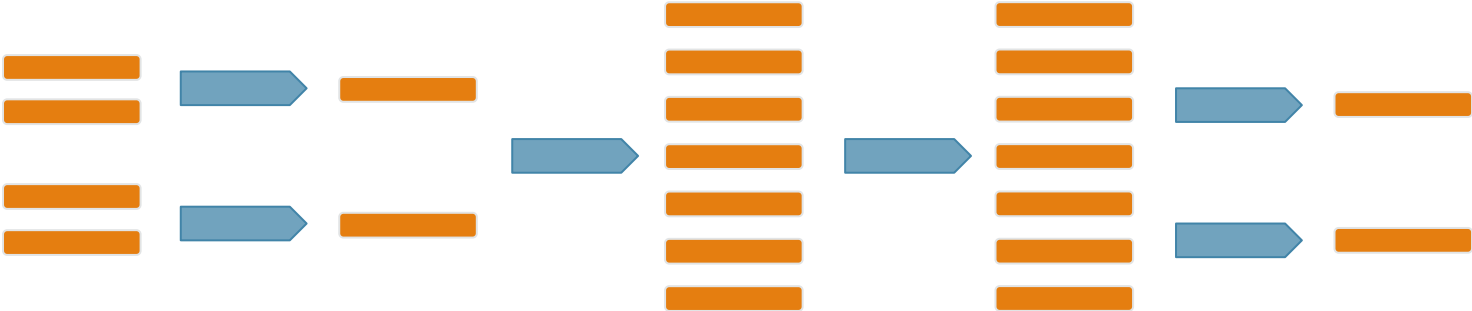
Dataset D



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

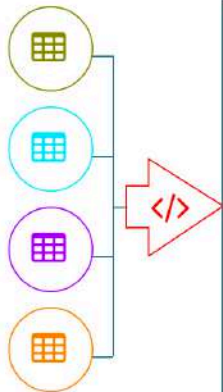
Files in this dataset Package: urn:uuid:1d23e155-3ef5-47c6-9612-027c80855e8d

Name	File type	Size	Download all
Metadata: metadata.xml	EML v2.1.1	140 KB	112 views
Total_Aromatic_Alkanes_PWS.csv	More info text/csv	3 MB	3 downloads
CollectionMethods.csv	More info text/csv	793 B	2 downloads
Non-EVOS_SINs.csv	More info text/csv	3 KB	

Show 8 more items in this data set

Data Table, Image, and Other Data Details

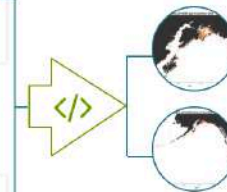
4 sources



Data Table

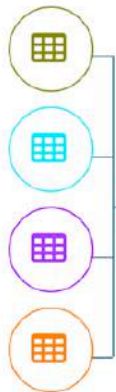
Entity Name	Total_Aromatic_Alkanes_PWS.csv										
	Download										
Description	Combined dataset from PAH, Alkane and Sample tables documenting samples collected after the Exxon Valdez oil spill in Prince William Sound, AK										
Object Name	Total_Aromatic_Alkanes_PWS.csv										
Online Distribution 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 colspan="2">Simple Text</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

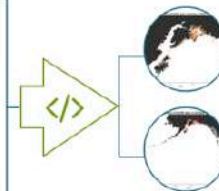
Citation

[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)**.

2 derivations



Text Format

Number of Header Lines	1
Record Delimiter	#x0A
Attribute Orientation	column
Simple Text	
Field Delimiter	,

Number Of Records

12142

Web Provenance Editor

Deployed by Arctic Data Center

The screenshot displays the NSF Arctic Data Center Web Provenance Editor interface. At the top, the NSF Arctic Data Center logo is on the left, and navigation links for Data, Support, and About are in the center. A green 'Submit Data' button and a user profile for Christopher Jones are on the right. The main content area is titled 'Data Table, Image, and Other Data Details'. Below this, there are sections for '0 sources' and '0 derivations', each with an 'Add' button. The central 'Data Table' section shows details for an entity named 'Total_Aromatic_Alkanes_PWS.csv'. It includes a 'Download' button, a description of the dataset, the object name, online distribution info, size, and text format details.

NSF Arctic Data Center

Data Support About Submit Data Christopher Jones

Data Table, Image, and Other Data Details

0 sources

0 derivations

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-stage.test.dataone.org/cn/v2/resolve/urn:uuid:df984766-dd89-4e57-b97e-350506d7007e>

Size: 2801033 byte

Text Format

Number of Header Lines	1
Record Delimiter	#x0A



Add source data to Total_Aromatic_Alkanes_PWS.csv

Choose files in this dataset:

CollectionMethods.csv
 hcdbSamplesGOA.png
 hcdbSampleLocs.png
 PAH.csv
 Alkane.csv
 Non-EVOS_SINs.csv
 Sample.csv

Done

Online Distribution Info

<https://cn-stage.test.dataone.org/cn/v2/resolve/urn:uuid:d1984766-dd89-4e57-b97e-350506d7007e>

Size 2601033 byte

Text Format	Number of Header Lines	1
	Record Delimiter	#x0A
	Attribute Orientation	column
	Simple Text	
	Field Delimiter	,

Number Of Records 12142



Data Table, Image, and Other Data Details

4 sources



Data Table

Entity Name **Total_Aromatic_Alkanes_PWS.csv**
[Download](#)

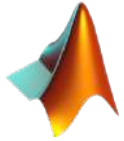
 Description
 Combined dataset from PAH, Alkane and Sample tables documenting samples collected after the Exxon Valdez oil spill in Prince William Sound, AK
Object Name **Total_Aromatic_Alkanes_PWS.csv**
 Online Distribution Info
<https://cn-stage.test.dataone.org/cn/v2/resolve/urn:uuid:df984766-dd89-4e57-b97e-350506d7007e>
Size **2801033 byte**
 Text Format

Number of Header Lines	1
Record Delimiter	#x0A
Attribute Orientation	column
Simple Text	
Field Delimiter	,

0 derivations


[Save](#)

Provenance Editing



Matlab DataONE Toolbox



Recordr R Library



YesWorkflow Tool

MetacatUI
Web Provenance Editor

Data Table, Image, and Other Data Details

0 sources 0 derivations

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

The screenshot shows a web interface for editing provenance. It features a central 'Data Table' section with a table containing one row. The table has columns for 'Entity Name' and 'Description'. The 'Entity Name' is 'Total_Aromatic_Alkanes_PWS.csv' and the 'Description' is 'Combined dataset from PAH, Alkane and Sample tables documenting samples collected after the Exxon Valdez oil spill in Prince William Sound, AK'. There are 'Add' buttons and arrows on the left and right sides of the table, and a 'Download' button below the 'Entity Name'.



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



DataONE

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:

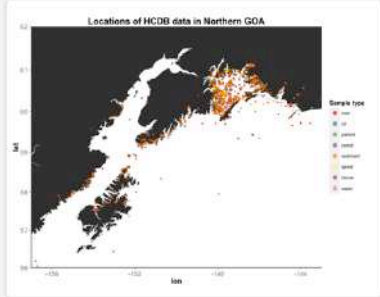
- **Which data** produced it
- **What software** produced it
- What was **derived** from it
- **Who to credit** down the attribution stack

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.



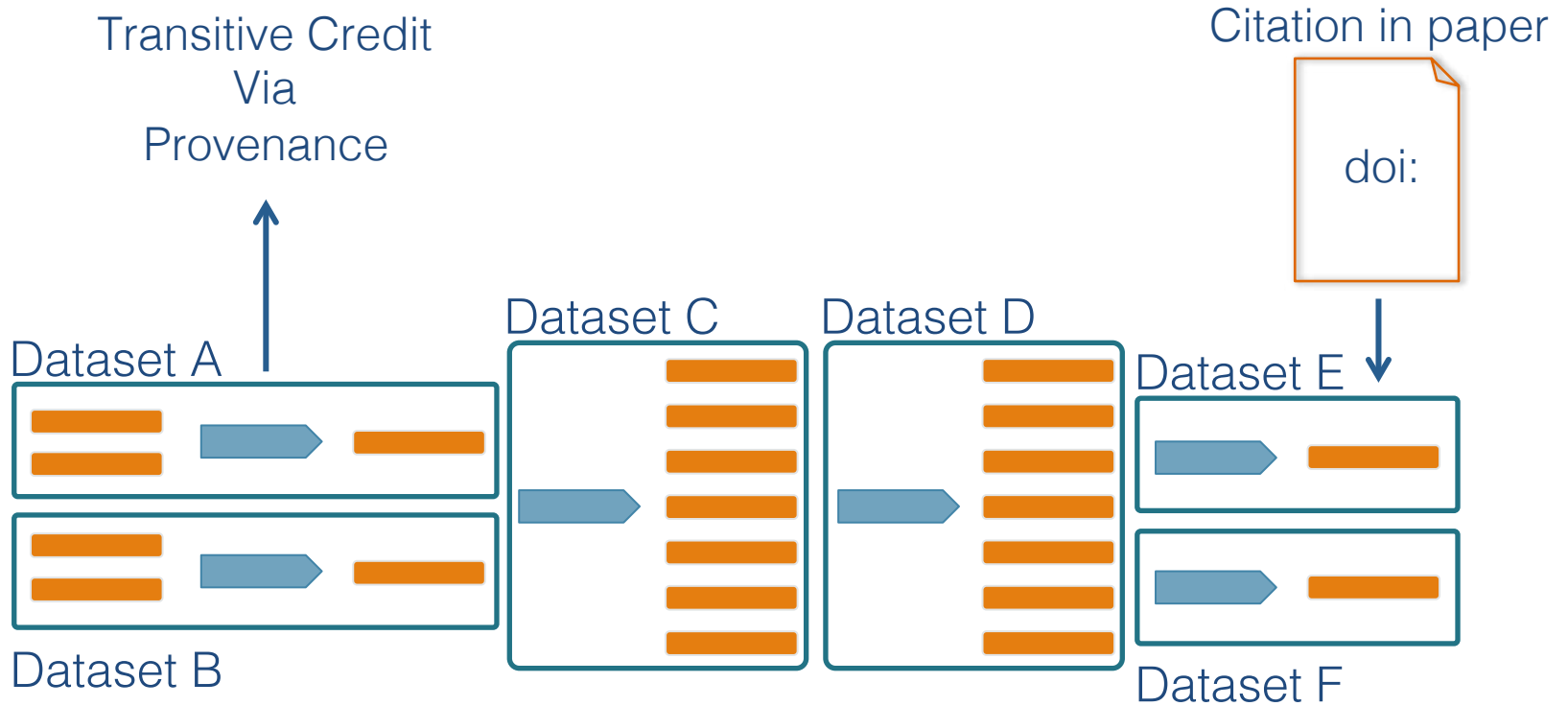
Locations of HCDB data in Northern GOA

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

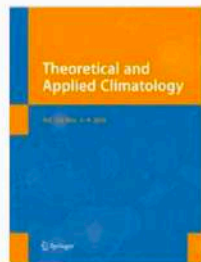
1 st Gen	Prose		
2 nd Gen	Prose	+ Data	
3 rd Gen	Prose	+ Data	+ Code

Prose + Data + Code + **Provenance**

Prose + Data + Code + **Provenance** +
Execution Environment

DataONE

WHOLETALE




[Theoretical and Applied Climatology](#)

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

Learning from mistakes in climate research

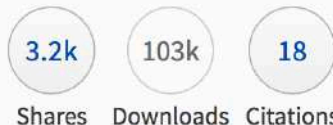
Authors

[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



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

Simplifying **Computational Reproducibility** in Whole Tale



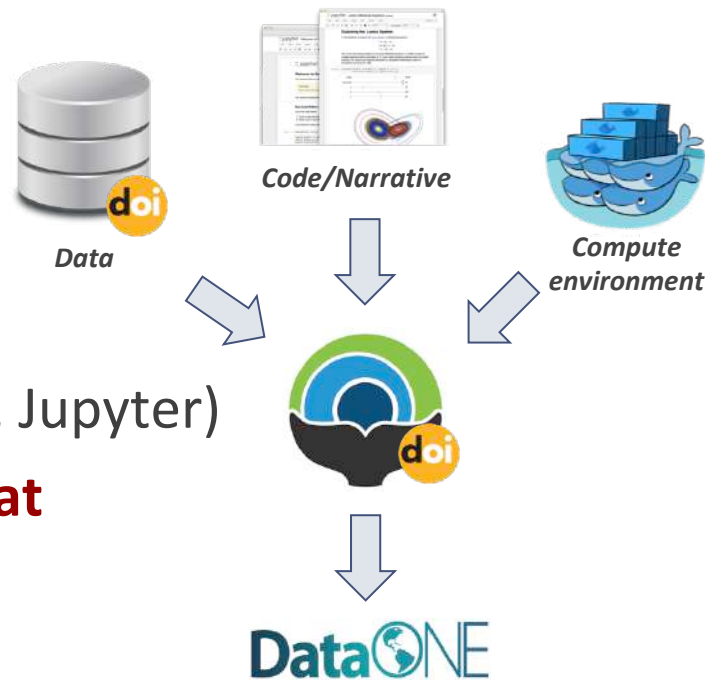
- Researchers can **easily package and share *tales***:
 - **Data, Code, and Compute Environment**
 - to **re-create the computational results** from a scientific study
 - achieving **computational reproducibility**
 - thus **“setting the default to reproducible.”**
- Also empowers users to verify and extend results with **different data, methods, and environments.**

V. Stodden, D. H. Bailey, J. Borwein, R. J. LeVeque, W. Rider, and W. Stein. (2013). *Setting the Default to Reproducible: Reproducibility in Computational and Experimental Mathematics*, ICERM workshop (2013)



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



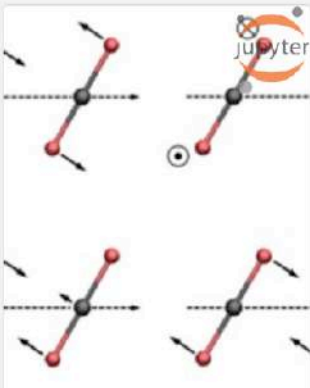


Browse Tales *Launch to add to Launched Tales list*



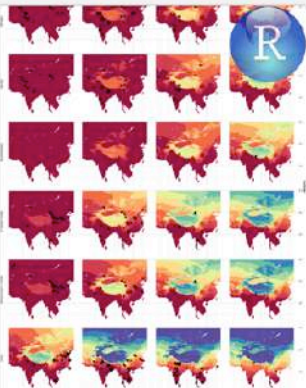
Search tales

All Switch to list view



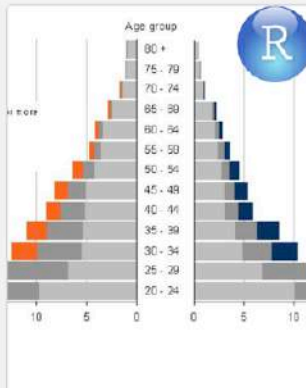
COMPUTATIONAL CHEMISTRY
Anharmonic vibrational structure of...

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 results in the Boosting for Economic Applications

Launched Tales


L2-Boosting for Economic Applicatio... x

Browse Existing Tales ...



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

Tale name:

Compute environment:  RStudio (rocker/geospatial)





Input data:

[Launch New Tale](#)

... Compose New Tales ...

Environments

Search compute environments...

-  RStudio (rocker/geospatial) i
-  Jupyter Classic i
-  RStudio i
-  Jupyter Lab i



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

Interact Files Metadata

File Edit Code View Plots Session Build Debug Profile Tools Help

```

1 # #####
2 # L2-Boosting for Economic Applications
3 # #####
4 # Parameter for simulation study
5 nn(list=ls())
6 source("DGP.R")
7 source("helper.R")
8 R <- 500 # number of repetitions
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
    
```

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

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

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")
> |
    
```

Launched Tales

L2-Boosting for Economic Applicatio...

...

Run & Interact with Tales

...

Search / Metadata

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

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

Citations 0

Downloads 183

Views 72

Copy Citation

Analyze

RStudio

Jupyter Notebook

Files in this dataset Package: resource_msp_id10.5065/D6862DM8

Name	File type	Size	Views	Download All
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

Downloads

Contact Share

Data for: Greater Expectations: A Field Experiment to Improve Accountability in Mali Version 2.0

Replication Data for: Greater Expectations: A Field Experiment to Improve Accountability in Mali (CVN/29010, Harvard Dataverse, V3, UNF:6:CCpg4SjQMVBnD-HvPM1Q== [fileURL])

Explore Cite Dataset

Analyze in WT

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 55 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

... 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

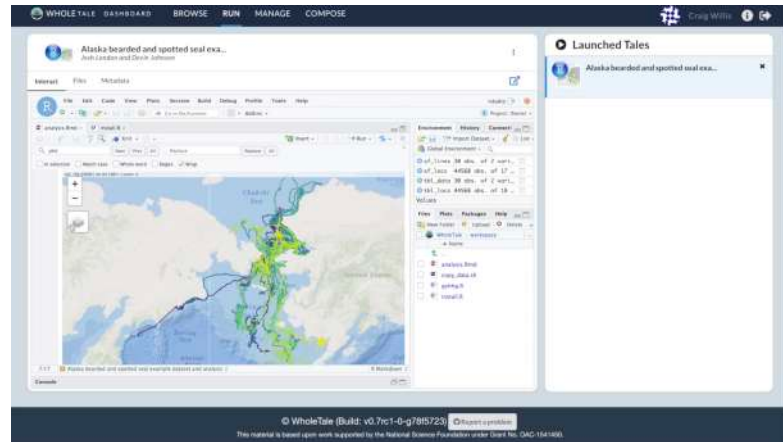


Whole Tale Forecast Demo

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

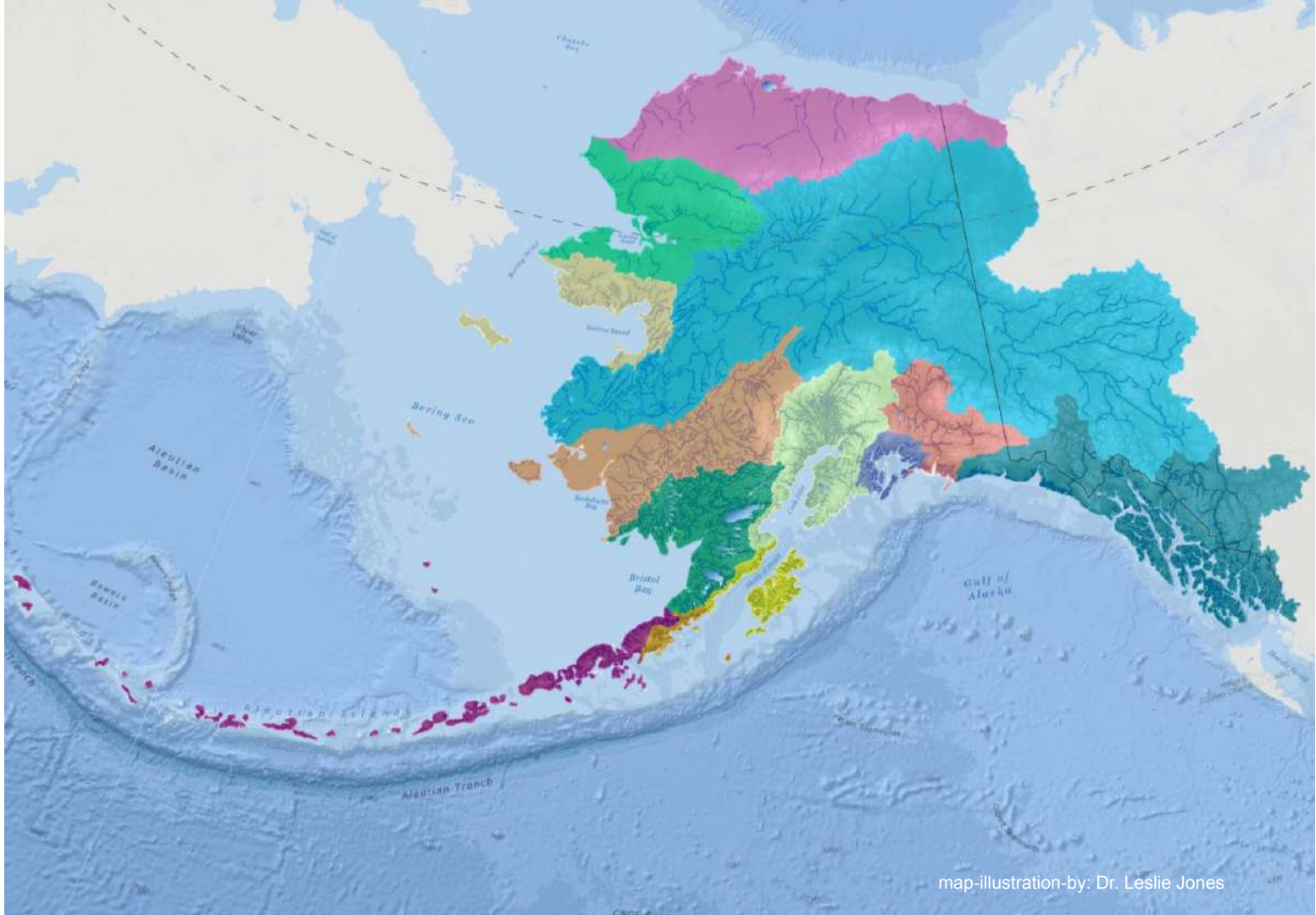
Leads: Peter S. Rand, Robert W. Campbell, Kristen B. Gorman

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

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	Download
StockInfo.csv	More info text/csv	19 KB	2 downloads	Download
SourceInfo.csv	More info text/csv	723 B	2 downloads	Download
broodTableProcessing.Rmd	More info application/R	19 KB	3 downloads	Download
broodTableProcessing.html	More info HTML	1 MB	9 downloads	Download

[▲ Show less](#)

30 inputs



view more

Other Entity

Entity Name

Download

Data Object Type:
Other

Physical Structure Description:

Object Name

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.

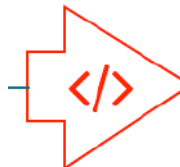
This data was used as an input to
create [BroodTables.csv](#).

1 outputs



Rmarkdown as Provenance

```
01-brood-table-integration.Rmd
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 ```
41
42 ```{r, echo = FALSE}
43 datatable(stocks[, c('Stock.ID', 'Stock', 'Region', 'Sub.Region')], rownames = FALSE,
44 caption = "Stock information")
45 ```
46
47 These stocks range geographically from Washington to Alaska. Although temporal coverage
48 varies by stock, many of the brood tables were updated in 2016, and some have
49 reconstructions dating back to 1922.
50
51 Figure 2.1 indicates the approximate location of the salmon stocks in Table 2.1.
52
53 ```{r, echo = FALSE}
54 salmon = makeIcon('images/salmon_tiny.png',
55 'images/salmon_big.png',
56 26, 14)
57
58 m <- leaflet(stocks) %>%
59   setView(-median(stocks$Lon), median(stocks$Lat), zoom = 4) %>%
60   addTiles() %>%
61   addMarkers(~Lon, ~Lat, icon = salmon)
62
63 m
64 ```
65
66 Figure 2.1: Location of stocks used in this data integration. Salmonid icon by Servien
67 (vectorized by T. Michael Keesey)
68 [CC-BY-SA](https://creativecommons.org/licenses/by-sa/3.0/), available at
69 Phylonic (https://www.nhvlonic.org/)
37:72 Chunk 2 = 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.

Show 10 entries Search:

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)

SASAP

Group

Group Id: SASAP

4 years, 6 months

Contributor since August 4, 2013

2 contributions

4,862 downloads

24 members

Krista B Oke

<http://orcid.org/0000-0002-5415-3534>

Josh Baron

<http://orcid.org/0000-0002-4286-6959>

Rich Brenner

<http://orcid.org/0000-0001-7209-9757>

★ Jeanette Clark

<http://orcid.org/0000-0003-4703-1974>

First 1 2 3 4 5 6 Last

DATASETS 1 TO 5 OF 60

1 2 3 ... 12 Next

Sort by Most recent



Alaska Department of Fish and Game, Division of Commercial Fisheries, Central Region. 2018. **Chinook age, sex, and length data from East Side Cook Inlet, Alaska, 1970-2012**. Knowledge Network for Biocomplexity. urn:uuid:16763faf-9ad6-4a95-bfc9-97d60957e499.



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.



Alaska Department of Fish and Game. 2018. **Salmon age, sex, and length data from Lower Cook Inlet, Alaska, 1961-2014**. Knowledge Network for Biocomplexity. urn:uuid:99e94ab7-b822-458e-88b3-df0ed1964378.



Jared Kibebe and Leslie Jones. 2018. **Glaciers in Alaska with subsetting by watershed and SASAP region**. Knowledge Network for Biocomplexity. urn:uuid:874e1ba2-48d2-4d31-b3fb-682aaf7e984b.

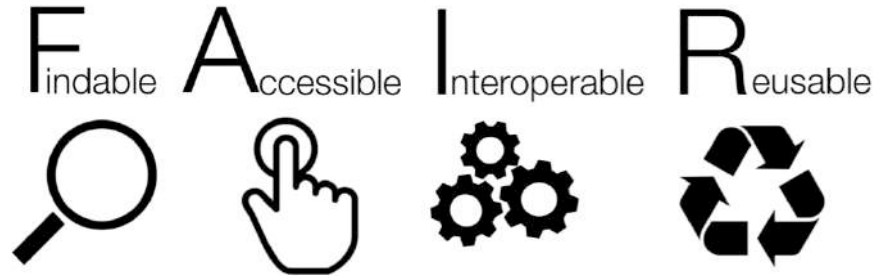


Jeanette Clark, Rich Brenner, and Bert Lewis. 2018. **Compiled age, sex, and length data for Alaskan salmon**. Knowledge Network for Biocomplexity. urn:uuid:63a9c8df-3543-44fe-a5d0-746469318f18.



Foundational Infrastructure

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



<https://www.force11.org/fairprinciples>