Geospatial Data Preservation and Access
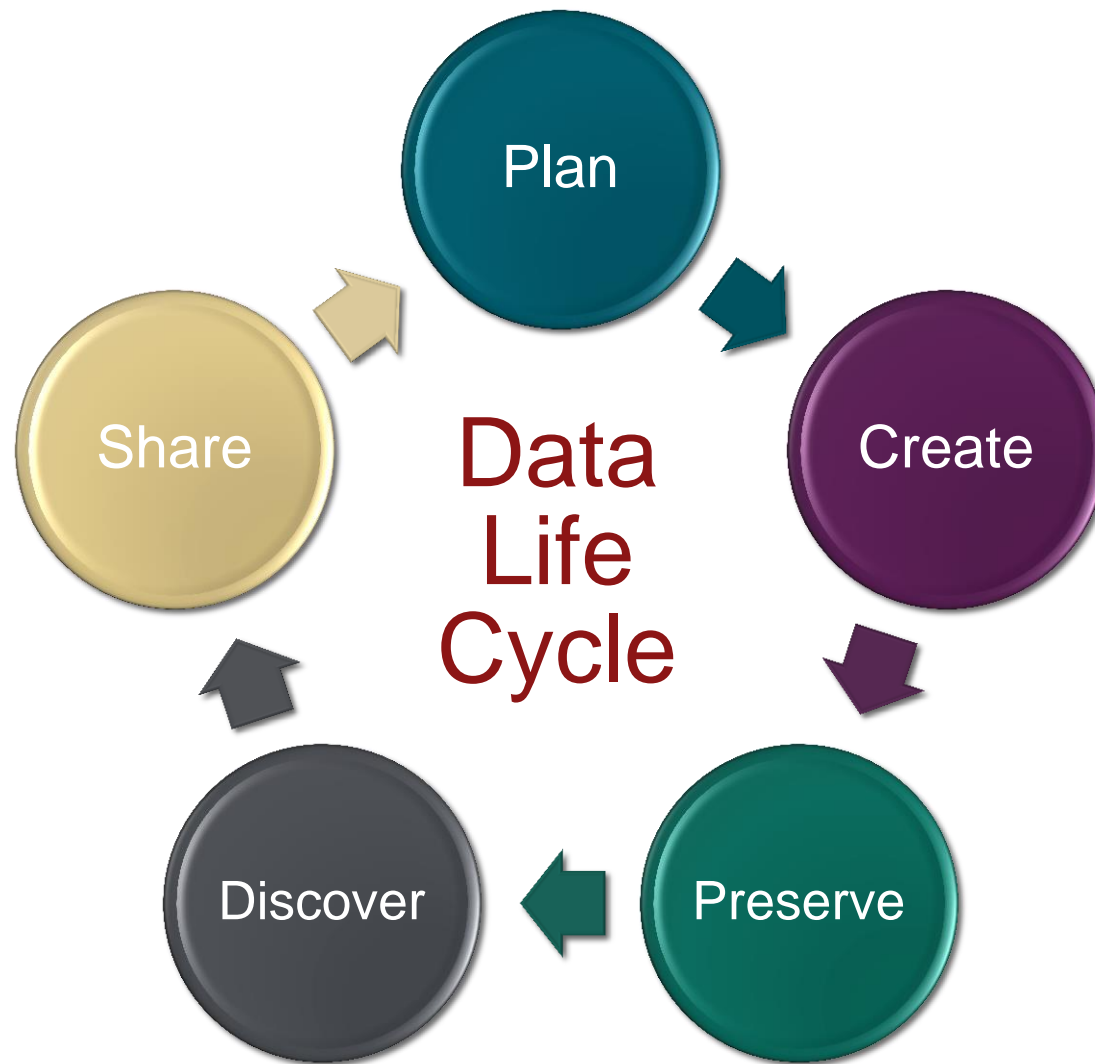
EASE OF DEPOSIT, EASE OF ACCESS

Julie Sweetkind-Singer

June 25, 2014

NGAC meeting

Plan

Create

Preserve

Discover

Share

Data Life Cycle

STANFORD UNIVERSITY LIBRARIES

Do-It-Yourself Data Sharing

- Put it up on the web
  - stored on local server
  - delivery of files via web site
  - no preservation
  - no discovery
  - no guarantee URL will be stable

Image by Flickr user kurtxio CC BY 2.0

- Provide it upon request
  - stored anywhere
  - delivered via email or physically
  - no preservation
  - no discovery
  - no guarantee you'll remember where it is

Image by Aaron Concannon CC BY-NC-ND 2.0

Stanford Digital Repository

○ Management of scholarly information resources of enduring value to the university

○ Content is preserved in a robust, reliable, and secure environment

○ Content is discoverable via catalog and Google crawling

○ Content is accessible via PURL downloads

○ Scholars around the world access our digital content for their research

SDR Facts

o **Storage**: redundant in geo-diverse locations, audited systematically to ensure bit preservation

o **Format**: spinning disks & tape

o **Volume**: 130+ TB of unique content (~500,000 objects)

o **Content**: books, manuscripts, maps, photographs, audio, video, data, web sites, theses, dissertations, articles, etc.

o **Scope**: no restrictions on file formats

o **Users**: faculty, students, researchers, library and other staff

SDR Online Deposits

o **Publications**: articles, working papers, pre-prints

o **Technical reports**

o **Conference materials**: slides, posters, etc.

o **Research data**: spreadsheets, images, video, etc.

o **Student work**: honors theses, class research projects

o **Archival records**: documents, event recordings, oral histories

SDR Features

o License options

o Embargo options

o Public access levels for finding, viewing, and downloading content

   o World access

   o Stanford-only access

o Indexed in SearchWorks (library catalog)

o Assigned PURL

o Version management

Preservation

Salmon Conservation

Sharing

Image by Andrew Russell CC BY 2.0

Sharing

Stanford Digital Repository

Data Supplement for "Range-wide selection of catchments for Pacific salmon conservation."

Author/Creator:	Pinsky, Malin L (Author)
	Springmeyer, Dane B. (Author)
	Goslin, Matthew N. (Author)
	Augerot, Xanthippe (Author)
Language:	English
Date created:	2009
Type of resource:	Software, multimedia
Genre:	Dataset
Abstract:	Freshwater ecosystems are declining in quality globally, but a lack of data inhibits identification of areas valuable for conservation across national borders. We developed a biological measure of conservation value for six species of Pacific salmon (Oncorhynchus spp.) in catchments of the northern Pacific across Canada, China, Japan, Russia, and the United States. We based the measure on abundance and life-history richness and a model-based method that filled data gaps. Catchments with high conservation value ranged from Califor ... more »
Preferred Citation:	Pinsky, ML, Springmeyer, DB, Goslin, MN, Augerot, X (2009). Data Supplement for "Range-wide selection of catchments for Pacific salmon conservation." Stanford Digital Repository. Available at http://purl.stanford.edu/zc193vn8689.
Related Publication:	Pinsky, M. L., Springmeyer, D. B., Goslin, M. N. and Augerot, X. (2009), Range-Wide Selection of Catchments for Pacific Salmon Conservation. Conservation Biology, 23: 680–691. doi: 10.1111/j.1523-1739.2008.01156.x. Available at http://onlinelibrary.wiley.com/doi/10.1111/j.1523-1739.2008.01156.x/full.
Contact:	malin.pinsky@gmail.com
Collection:	Hopkins Marine Station Collection
Related item:	Wild Salmon Center
	State of the Salmon
	Abundance Estimates of the Pacific Salmon Conservation Assessment Database, 1978–2008
Subject:	conservation planning
	freshwater protected areas
	landscape variables
	northern Pacific
	Oncorhynchus
	salmon áreas dulceacuícolas protegidas
	Oncorhynchus
	Pacífico del norte
	planificación de la conservación
	salmón
	variables de paisaje

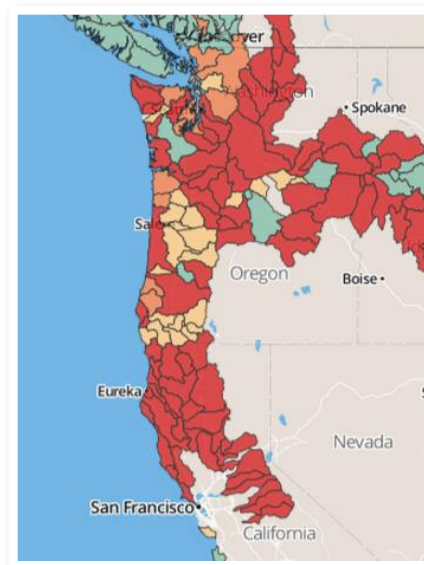| File (Download)	Description
Pinsky_etal_AbundanceDatabase2008.csv (2 MB)	Abundance Database as published
Hydro1kshapefile.zip (6 MB)	Catchment ArcGIS shapefile to be used for georeferencing the Abundance and Run-timing databases (link CatchmentID in the database to NLEVELS in the shapefile). Shapefile includes river names.
Abundance_Database_Metadata.pdf (70 KB)	Describes the content of each field in the Abundance Database (Pinsky_etal_AbundanceDatabase2008.csv and Pinsky_etal_2009_AbundanceData-2010-09-29.csv)
RunTiming_Database_Metadata.pdf (70 KB)	Describes the content of each field in the Run Timing Database (Pinsky_etal_2009_RunTimingData-2010-09-29.csv)
Pinsky_etal_Sources.pdf (90 KB)	List of data sources for Abundance Database
Pinsky et al. 2009 Conservation Biology 23(3)-680–691 Supplemental Data_v2.pdf (40 KB)	Overview of this set of supplemental data files
Pinsky_etal_2009_AbundanceData-2010-09-29.csv (3 MB)	Revised Abundance Database; Removed duplicate sockeye abundance record from AK Peninsula (CatchmentID 15917)
Pinsky_etal_2009_RunTimingData-2010-09-29.csv (200 KB)	Run Timing Database

Use and Reproduction:	User agrees that, where applicable, content will not be used to identify or to otherwise infringe the privacy or confidentiality rights of individuals. Content distributed via the Stanford Digital Repository may be subject to additional license and use restrictions applied by the depositor.
License:	This work is licensed under a Open Data Commons Public Domain Dedication and License (PDDL)

STANFORD UNIVERSITY LIBRARIES

Stanford Geoportal

PUTTING DATA ON THE MAP

http://purl.stanford.edu/vv853br8653

Dedicated geospatial portal

Data preview

Attribute data preview

Metadata information and data download

Future Directions

- Better capabilities for deposit of large data sets
  - Files that are > 5GB
  - Many files
  - Both
- Dedicated discovery interfaces: GeoBlacklight
- Cost model for larger projects
- Custom licenses/use/reproduction statements