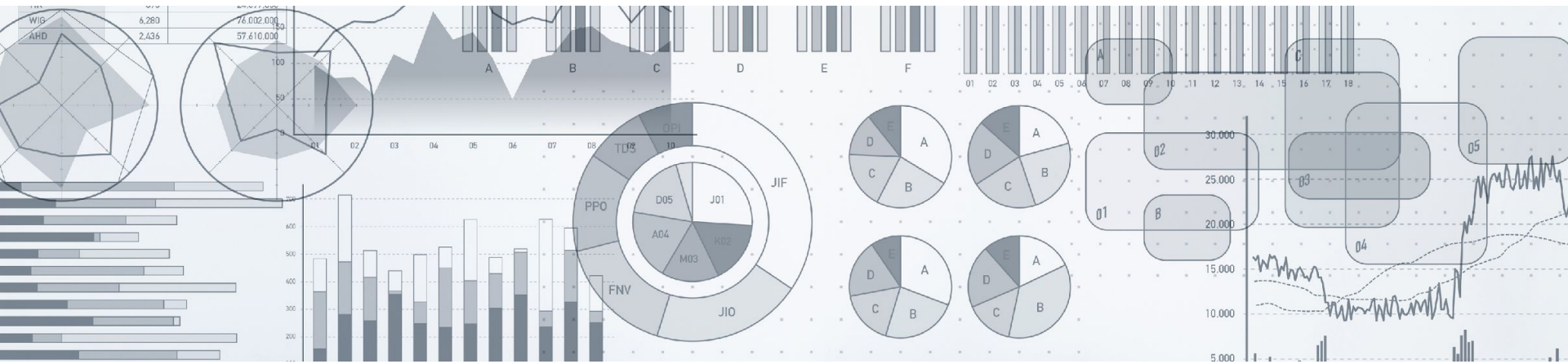


Managing and Sharing your PEER Data and Publications Meeting and exceeding USAID's Public Access Plan



Public Access to Research

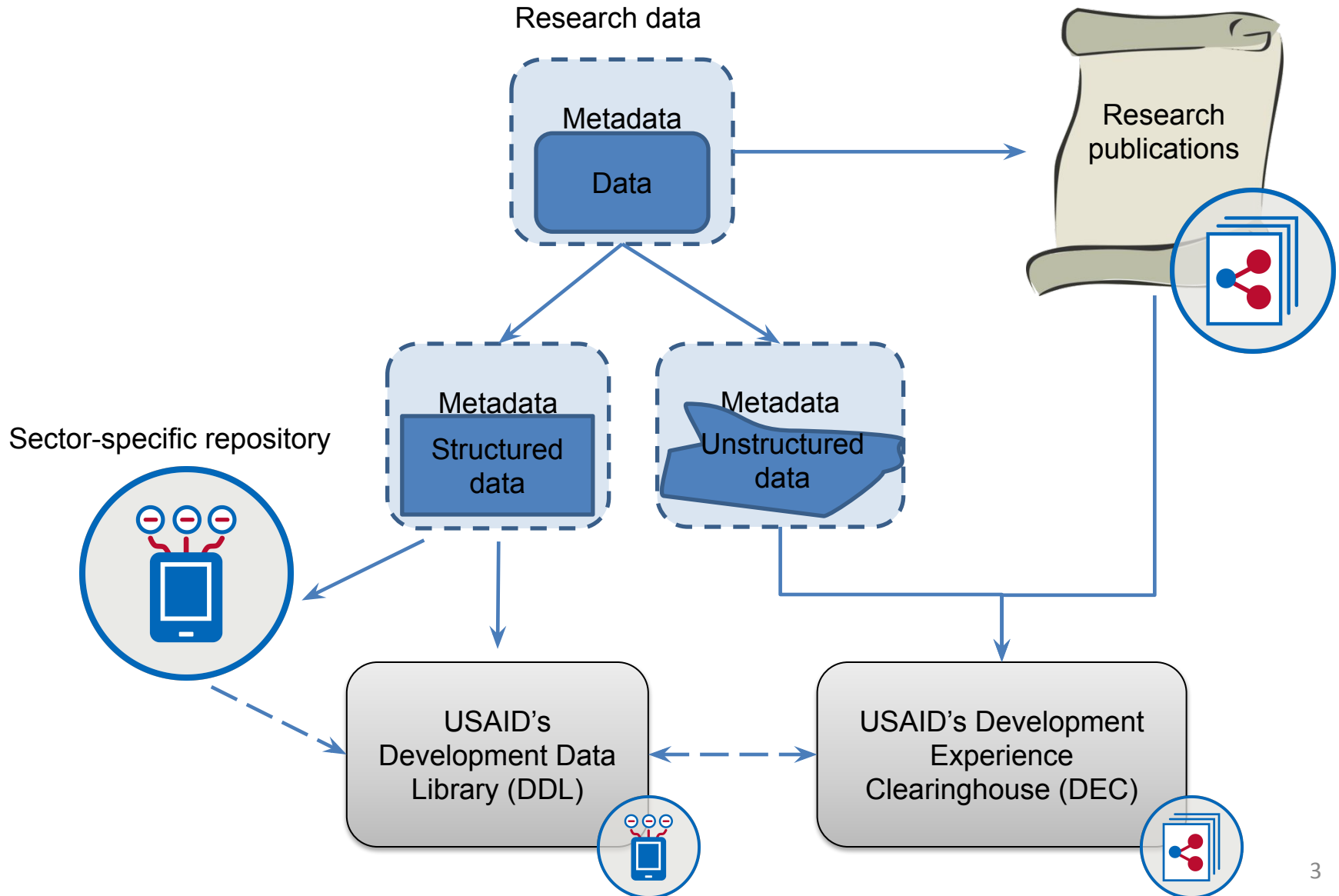
USAID research programs **strive to:**

- Support best practices for managing publications and scientific research data
- Increase public access to results of scientific research
- Facilitate the use of research

Why? To enhance scientific discovery and improve development outcomes

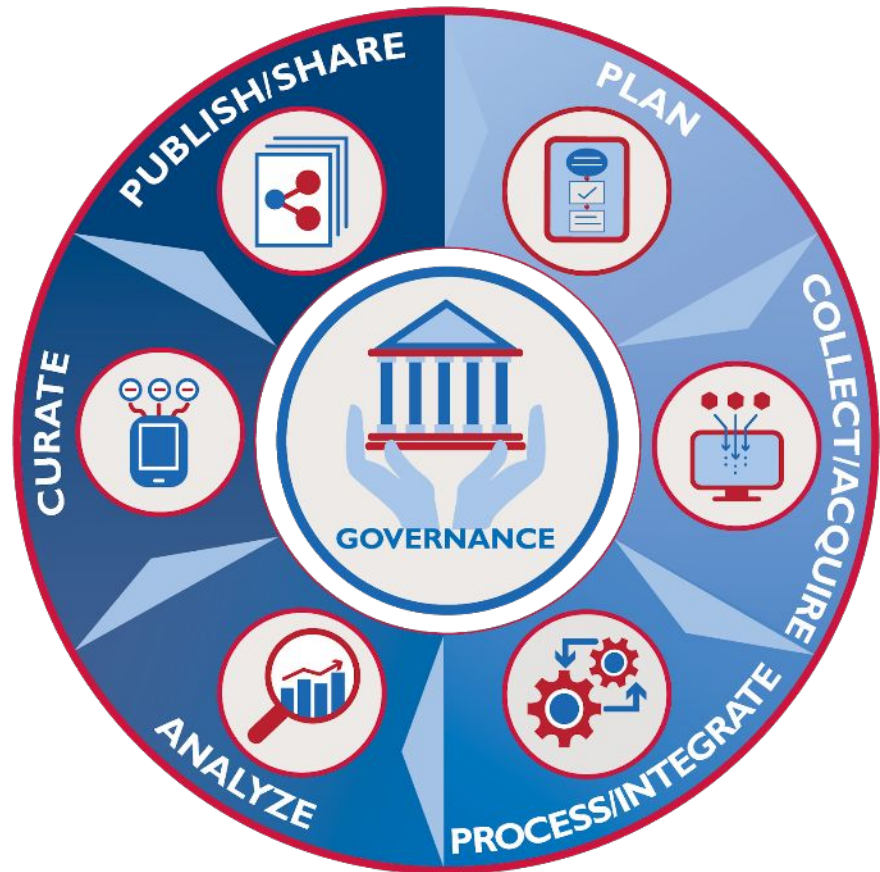


Pathway for Data and Publications

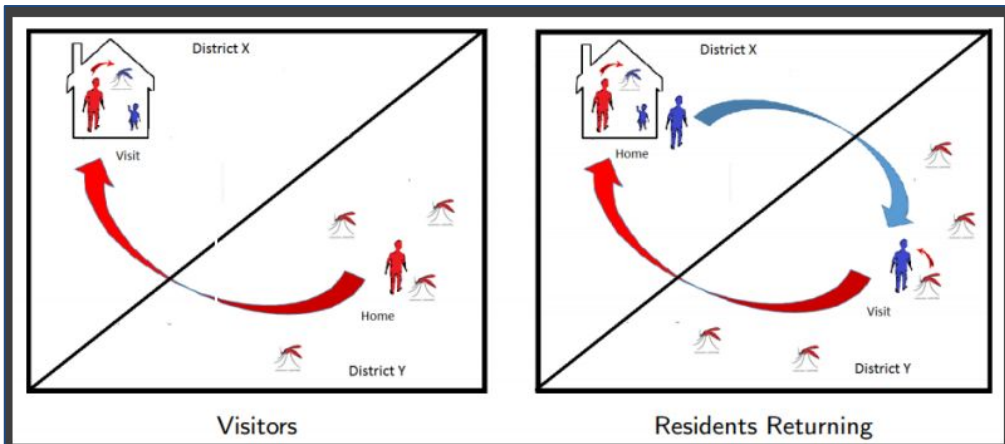


Data Management

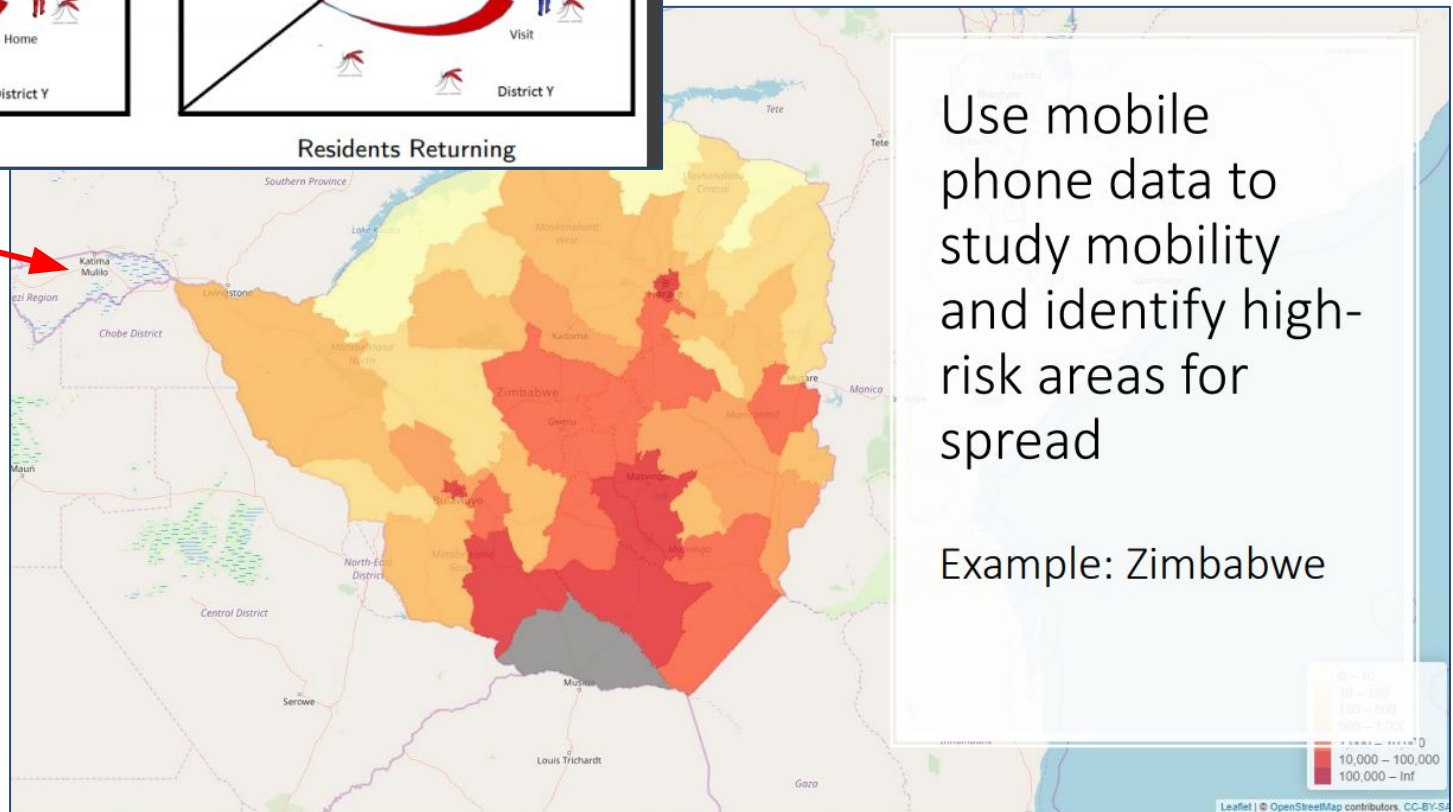
- Manage data responsibly throughout lifecycle
- Produce high-quality, usable data
- Include necessary information for reuse
- Protect data and make sure you get credit



Example: covid-mobile-data

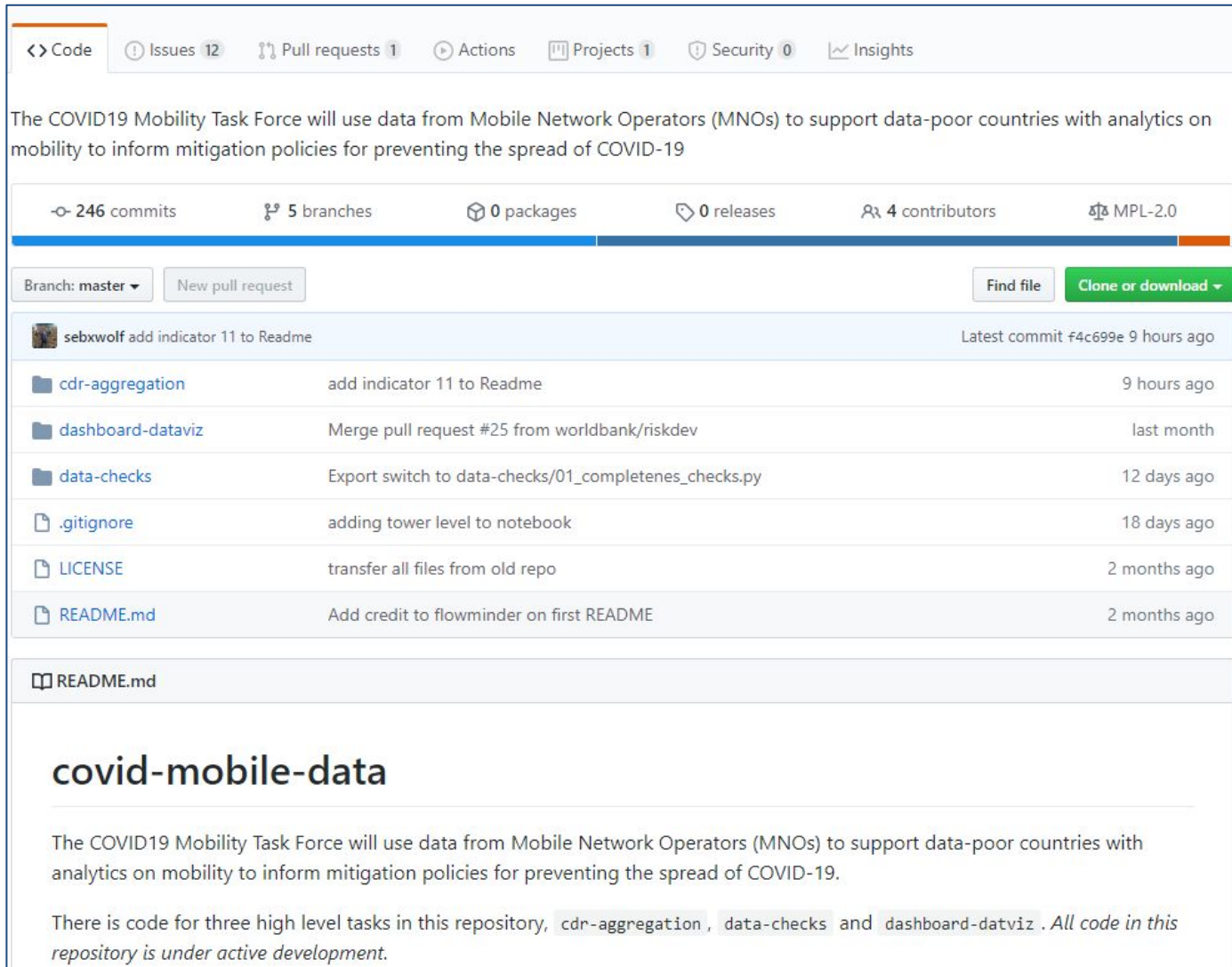


Data originally used to understand the spread of Malaria helps identify high risk areas for COVID-19.



Source: Sveta Milusheva, Development Impact Evaluation (DIME), World Bank, accessed 6/16/2020 from http://cega.berkeley.edu/wp-content/uploads/2020/04/Sveta_MeasureDev2020_Slides.pdf

Example: [covid-mobile-data](#)



The screenshot shows the GitHub repository page for `covid-mobile-data`. At the top, there are navigation tabs for Code, Issues (12), Pull requests (1), Actions, Projects (1), Security (0), and Insights. Below this is a description: "The COVID19 Mobility Task Force will use data from Mobile Network Operators (MNOs) to support data-poor countries with analytics on mobility to inform mitigation policies for preventing the spread of COVID-19".

Repository statistics are shown: 246 commits, 5 branches, 0 packages, 0 releases, 4 contributors, and MPL-2.0 license. There are buttons for "Branch: master", "New pull request", "Find file", and "Clone or download".

A commit history table is visible:

Commit	Message	Time
sebxwolf	add indicator 11 to Readme	Latest commit f4c699e 9 hours ago
cdm-aggregation	add indicator 11 to Readme	9 hours ago
dashboard-dataviz	Merge pull request #25 from worldbank/riskdev	last month
data-checks	Export switch to data-checks/01_completeness_checks.py	12 days ago
.gitignore	adding tower level to notebook	18 days ago
LICENSE	transfer all files from old repo	2 months ago
README.md	Add credit to flowminder on first README	2 months ago

Below the commit history is the `README.md` file content:

covid-mobile-data

The COVID19 Mobility Task Force will use data from Mobile Network Operators (MNOs) to support data-poor countries with analytics on mobility to inform mitigation policies for preventing the spread of COVID-19.

There is code for three high level tasks in this repository, `cdm-aggregation`, `data-checks` and `dashboard-dataviz`. *All code in this repository is under active development.*

Source: COVID19 Mobility Task Force [worldbank/covid-mobile-data](#)

Sharing Research Data

Goal

- Open Data is the idea that data should be freely available for everyone to use whenever possible
- Makes science more efficient and more adaptive - faster learning from project results across wider audiences
- Preserve access to your data with a persistent identifier

Research context

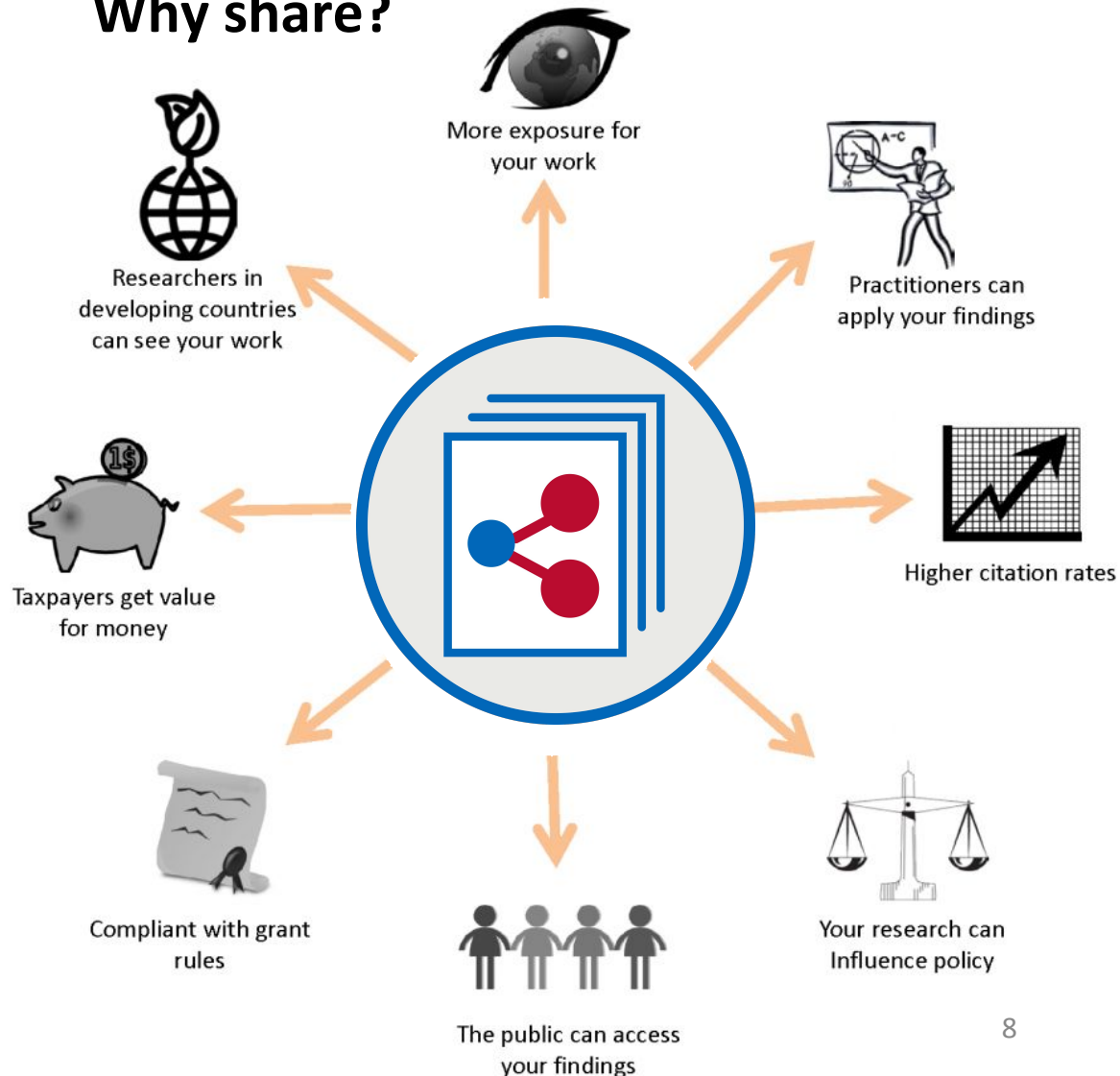
- Public Access mandates mean that research funded by the US government should be accessible
- The *data* is the research product, not the journal article
- Many repositories exist for housing data in perpetuity

Sharing Research Data

How to share:

- Release with journal article
- Peer-reviewed data paper
- Citizen science

Why share?



Sharing Research Data Responsibly

- Identify sensitive data
 - Personally identifiable information (e.g. human subjects)
 - Data restricted by country laws or regulations
- Protect privacy
 - Aggregate data, remove personal identifiers, redact sensitive information, restrict access, etc.
- Respect the agency of human subjects
 - Insist on informed consent, follow human subject protocols, allow for correction of data

Example: Snapshot Serengeti



**Motion
activated
cameras**

Source: dx.doi.org/10.1038/sdata.2015.26

Example: Snapshot Serengeti

English

SNAPSHOT SERENGETI

Home About Classify Profile Discuss Blog Authors



Fire No animals present

Looks like ▾

Pattern	Color	Horns	Tail	Build
Aardvark	Genet			Porcupine
Aardwolf	Giraffe			Reedbuck
Baboon	Guinea fowl			Reptiles
Bat	Hare			Rhinoceros
Bat-eared fox	Hartebeest			Rodents
Bird (other)	Hippopotamus			Secretary bird
Buffalo	Honey-badger			Serval
Bushbuck	Hyena (spotted)			Steenbok
Cattle	Hyena (striped)			Topi
Caracal	Impala			Vervet monkey
Cheetah	Insect/Spider			Vulture
Civet	Jackal			Warthog
Dik dik	Kori bustard			Waterbuck
Duiker	Leopard			Wildcat
Eland	Lion (female or cub)			Wildebeest
Elephant	Lion (male)			Zebra
Gazelle (Grant's)	Mongoose			Zorilla
Gazelle (Thomson's)	Ostrich			Human

Source: www.snapshotserengeti.org

Example: Snapshot Serengeti

nature.com > scientific data > data descriptors > article > article metrics

a natureresearch journal

SCIENTIFIC DATA



Search



E-alert



Submit



Login

Article metrics for:

Snapshot Serengeti, high-frequency annotated camera trap images of 40 mammalian species in an African savanna

Online attention



Altmetric score (what's this?)

- Tweeted by 111
- Blogged by 15
- On 13 Facebook pages
- Mentioned in 6 Google+ posts
- Picked up by 45 news outlets
- 109 readers on Mendeley
- 1 readers on Citeulike

This Altmetric score means that the article is:

- in the 99th percentile (ranked 348th) of the 222,467 tracked articles of a similar age in all journals

One of the most popular articles from 2015 on Altmetric

Importance for PEER

- Prepare PEER awardees for latest academic research trends
- Increase the visibility of research of PEER Principal Investigators
- Expand networks and enable new opportunities for funding



PEER Award Agreement

Article VIII - Data Rights and Publications

(a) All data, written materials, photographs, drawings, or other information created or generated under this Subaward (the "Subject Data"), and the copyrights therein in all media and languages throughout the world will be irrevocably assigned to and owned by Subawardee. NAS and the U.S. Government shall have an irrevocable, transferable, royalty-free, non-exclusive worldwide license in all media and languages now or hereafter known to reproduce, disseminate, publish, prepare derivative works in, or otherwise utilize all Subject Data, including without limitation, the right to authorize others to use and disseminate such Subject Data.

(b) **Attribution:** Subawardee shall cite in any published article that such article is derived from the Subject Data funded in whole or part by NAS and USAID under the USAID Prime Award Number AID-OAA-A-11-00012, and that any opinions, findings, conclusions, or recommendations expressed in such article are those of the authors alone, and do not necessarily reflect the views of USAID or NAS.

(c) **Electronic Program Deliverables:** Subawardee shall provide the Senior Program Officer with an electronic copy of all deliverables and any publications produced with project funds.

(d) **Submission of Datasets to the Development Data Library:** Subawardee shall submit data under this Subaward to the Development Data Library in accordance to Section M23 of the Standard Provisions for Non-U.S. Nongovernmental Organizations, Attachment C.

Data Management for the PEER Program

1. Choose a repository for your data
2. Create a short data management plan

**1st
year
of your
PEER
project**

Data and Publication Sharing for the PEER Program

3. Submit data to a repository and link to USAID's Development Data Library (DDL)

4. Submit any publication author's manuscripts to USAID's document platform - the Development Experience Clearinghouse (DEC)

**Year
2-3
of your
PEER
project**

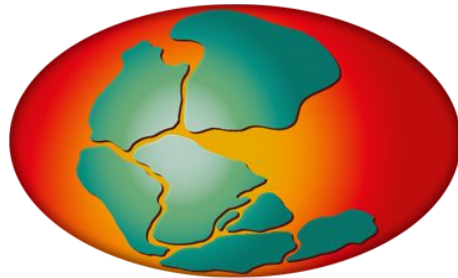
Step 1: Choose a Repository

Choose a data repository where your data will be seen



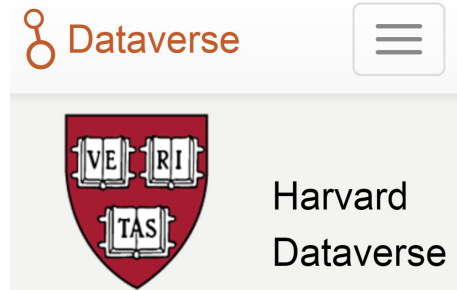
DEVELOPMENT DATA LIBRARY (DDL)

An Effort of USAID AidScape™



PANGAEA.

**Earth Science:
www.pangaea.de**



**General Data:
dataverse.harvard
.edu**



DRYAD

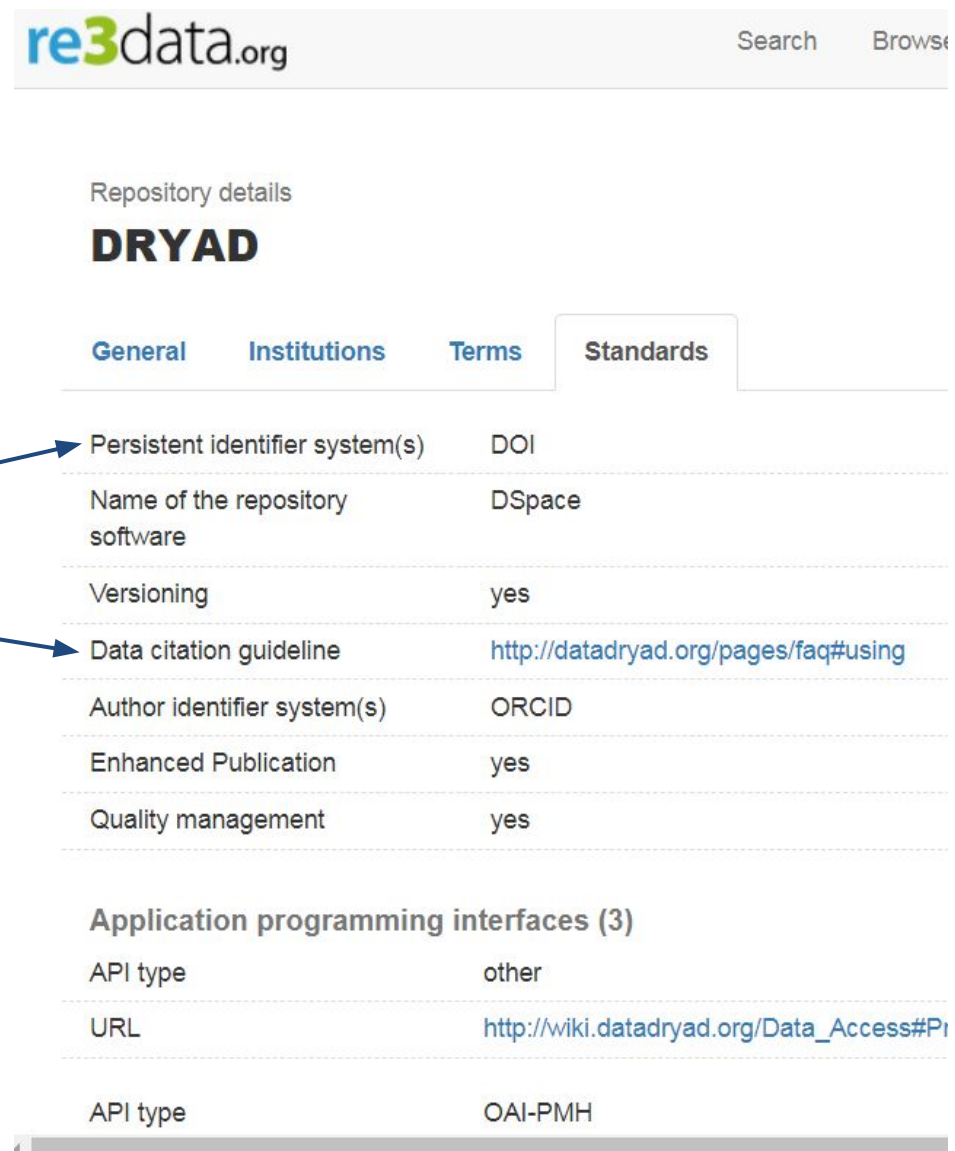
**General Data:
datadryad.org**

Step 1: Choose a Repository

How do I know if a repository is good?

1. Has a long term data management plan
2. Has a commitment to preserving your data or preservation plan (in mission statement or terms of service)
3. Provides a persistent identifier (ideally DOI) and a URL to the data
4. Has data citation guidelines to ensure you get credit for your work.
5. Makes data widely available and free for public use with clear licensing and use terms, [CC-0](#) or [CC-BY](#) licenses recommended
6. Allows wide sharing of metadata

For further guidance, see the [Development Data Overview](#)



re3data.org Search Browse

Repository details

DRYAD

General Institutions Terms Standards

Persistent identifier system(s)	DOI
Name of the repository software	DSpace
Versioning	yes
Data citation guideline	http://datadryad.org/pages/faq#using
Author identifier system(s)	ORCID
Enhanced Publication	yes
Quality management	yes
Application programming interfaces (3)	
API type	other
URL	http://wiki.datadryad.org/Data_Access#Pr
API type	OAI-PMH

Step 1: Choose a Repository

A good opportunity to work with your US partner colleagues



Step 2: Data Management Plan

Draft a data management plan

1 or 2 page document:

- how you will handle your data as it is collected
- how you handle data after research project ends

Submitted with PEER
year 1 annual report

[Data management plan instructions](#)

The screenshot shows the USAID PEER Cycle 6 Data Management Plan form. At the top, there is the USAID logo and navigation links for Search, Tools, Communications, and Reporting. Below the header, a breadcrumb trail indicates the current page: Process Manager / PEER Cycle 6 / Data Management Plan / Preview. A blue-bordered box contains a message: "Fields with an asterisk (*) are required." The form fields include: "Project name*" (text input), "Last name of PI" (text input), "Institution" (text input), and "Country" (dropdown menu). Below the form fields, there are four expandable sections: "Section A: Types of Data Produced", "Section B: Metadata", "Section C: Data Storage and Preservation", and "Section D: Policies for Access and Sharing".

Step 2: Data Management Plan

The DMP in Foundant covers:

1. Types of data produced
2. Metadata
3. Data storage and preservation
4. Policies for access and sharing

It is a living document that you can update over life of project!

Step 2: Data Management Plan

Example Question: Metadata Standard

Which metadata standards will you use and why have you chosen them?*

For many types of data, there are standards / guidelines for what metadata to use. To choose a metadata standard, it is helpful to consider the repositories where you would like to upload your data. Many repositories recommend metadata standards. Some recommended tools are:

- [DDI](#) - for Social, Behavioral, and Economic Sciences - You can use [Nesstar Publisher](#) to manage the data
- [EML](#) - for Earth, Environmental, and Ecological Sciences - You can use [Morpho](#) to manage the data

To answer this question, do one of the following:

- Choose a metadata standard that is popular in your field and say why you chose it; or
- Explain that you were unable to find an appropriate standard.

Samples:

- *The biological and ecological data will be structured in Ecological Metadata Language (EML).*
- *All physical and chemical time series data will be formatted to follow the standard operating procedures for ocean acidification research as described by Riebesell et al. (2010)*

Step 2: Data Management Plan

- Ideally all data can be shared immediately, but this may not be possible in all cases
- You will be asked to identify privacy concerns
 - May need to remove personally identifiable information (e.g. human subjects)
 - May need permission to share certain country data (e.g. water quality or river flow)
- Data can be private for up to 12 months under embargo if data is sensitive or there are publishing concerns



General Resources

- [PEER Data Management Webpage](#)
- [Development Data Overview](#)
 - E.g.: Metadata Standards, Persistent Identifiers, Data Anonymization, etc.
- [USAID's Development Data Library \(DDL\)](#)
- USAID's publication library, [the Development Experience Clearinghouse](#) (DEC)

Repository Resources

- Social science and public health research
 - USAID Development Data Library (DDL):
<https://www.usaid.gov/data>
 - Inter-university Consortium for Political and Social Research (ICPSR)
<https://www.icpsr.umich.edu/web/pages/deposit/index.html>
- Natural Sciences:
 - <http://www.sciencemag.org/authors/science-editorial-policies#data-deposition>
 - <https://www.nature.com/sdata/policies/repositories>
- Repository registries:
 - <https://www.coretrustseal.org/why-certification/certified-repositories/>
 - Registry of research data repositories <https://www.re3data.org/>
 - Fairsharing.org
 - [Tips](#) for selecting a data repository from PLOS ONE

Extra Slides

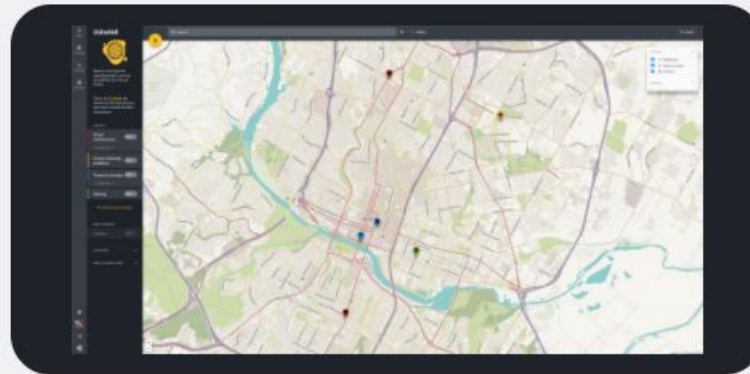
Example: Quakemap Nepal

QUAKEMAP

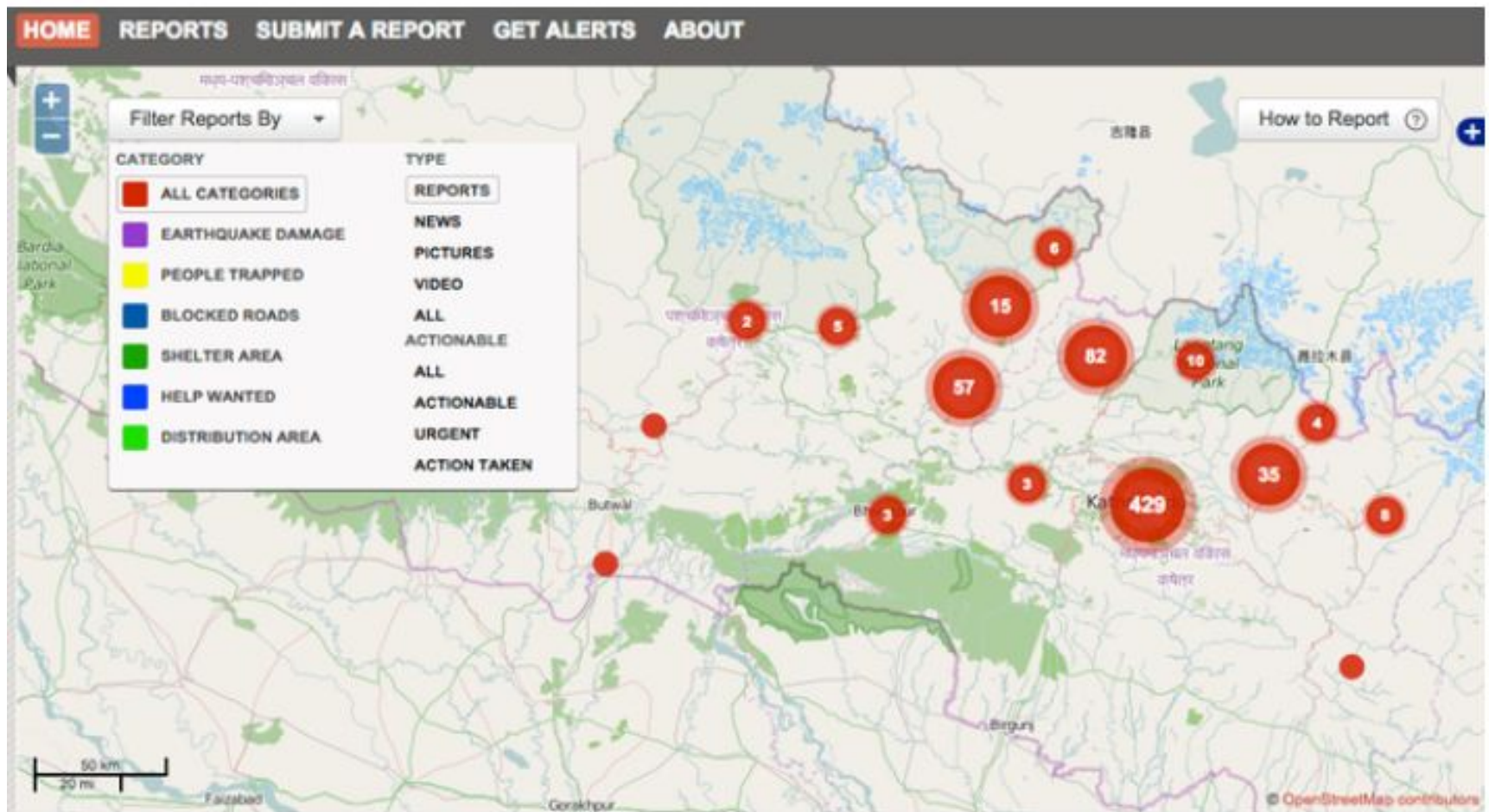
A deployment to help match those affected by Nepal's earthquake with ongoing relief efforts being conducted by various government, non-government and volunteer groups.



Kathmandu Living Labs



Example: Quakemap Nepal



Trustworthy, usable data unlocked rapid response to earthquake

Example: Quakemap Nepal

VIEW

Map

List

Bar Chart

Timeline

