
Supplementary information to: Credit data generators for data reuse

To accompany a Comment published in Nature **570**, 30–32 (2019)
<https://www.nature.com/articles/d41586-019-01715-4>

by Heather H. Pierce, Anurupa Dev, Emily Statham and Barbara E. Bierer

	Page
1. Participant list	2
2. Data tracking process	5
3. Action items	6

Implementing a System to Enable Credit for Data Sharing April 9, 2018

*Association of American Medical Colleges
655 K Street, N.W., Washington, D.C. 20001*

ATTENDEE LIST

George Alter

Research Professor
ICPSR

Gil Alterovitz

Presidential Innovation Fellow
White House Presidential Innovation
Fellow Program

Joyce Backus

Associate Director
National Library of Medicine

Jacquelyn Bendall

Director, Research Compliance
Council on Governmental Relations

Gordon Bernard

Executive Vice President for Research
Vanderbilt University

Barbara Bierer

Faculty Director
MRCT Center

Richard Bookman

Director and Senior Advisor
University of Miami

Brian Bot

Principal Scientist
Sage Bionetworks

Philip Bourne

Director & Professor
University of Virginia

Timothy Clark

Associate Professor
University of Virginia

Merce Crosas

Chief Data Science and Technology
Officer, IQSS
Harvard University

Jacques Demotes-Mainard

Director General
ECRIN

Anurupa Dev

Lead Specialist, Science Policy
Association of American Medical
Colleges

Jeffrey Drazen

Editor-in-Chief
New England Journal of Medicine

Rebecca English

Program Officer
National Academies

David Fearon

Sr. Data Management Specialist
Johns Hopkins University

Megan Force

Editor, Data Citation Index
Clarivate Analytics

Maryrose Franko

Executive Director
Health Research Alliance

Jason Gerson

Senior Program Officer
PCORI

Cynthia Grossman

Director
FasterCures

Janet Hall

Director, Precision Medicine
Platform
American Heart Association

Robert Hanisch

Director, Office of Data and
Informatics
NIST

Anna Hatch

DORA Community Manager
American Society for Cell Biology

Denise Hills

Director, Energy Investigations
Geological Survey of Alabama

Kristi Holmes

Director, Galter Library & Assoc
Prof, Prev Med
Northwestern University

India Hook-Barnard

Director of Research Strategy
University of California, San Francisco

Mike Huerta

Assoc Dir & Coordinator, Data &
Open Science
National Library of Medicine, NIH

Shanda Hunt

Public Health Librarian & Data
Curation Specialist
University of Minnesota

David Kennedy

Professor
University of Massachusetts Medical
School

Steven Kern

Deputy Director, Quantitative
Sciences
Bill & Melinda Gates Foundation

Rebecca Li

Executive Director
Vivli

Jennifer Lin

Director of Product Management
Crossref

Bernard Lo

President & CEO
The Greenwall Foundation

Daniella Lowenberg

Product Manager/ Research Data
Specialist
California Digital Library

Maryann Martone

Professor Emeritus
UCSD School of Medicine

Leslie McIntosh

Executive Director
Research Data Alliance

Jose Merino

US Research Editor
The BMJ

Pam Miller

Assistant to the Editor, Special Projects
New England Journal of Medicine

Elizabeth Moss

Librarian
ICPSR

Holly Murray

Associate Publisher
F1000

Mark Musen

Professor
Stanford University

Elizabeth Myers

Program Director for Medical
Research
Doris Duke Charitable Foundation

Amy Nurnberger

Program Head, Data Management
Services
MIT Libraries

Eric Olson

Membership Specialist
ORCID

Mark Paalman

Senior Editor
Wiley

Laura Paglione

Director, Strategic Initiatives
ORCID

Dina Paltoo

Director, Div of Scientific Data Sharing
Policy
Office of Science Policy, NIH

Laurence Peiperl

Chief Editor, PLOS Medicine
Public Library of Science

Heather Pierce

Senior Director for Science Policy
Association of American Medical
Colleges

Eleonora Presani

Product Manager, Scopus
Elsevier

Kacy Redd

Assistant Vice President, STEM
Education Policy
APLU

Jim Reecy

Associate Vice President for
Research
Iowa State University

Alyssa Reimer

Associate Director, Research
Partnerships
The Michael J. Fox Foundation

Frank Rockhold

Professor
Duke University

Judy Ruttenberg

Program Director for Strategic
Initiatives
Association of Research Libraries

Joshua Schnell

Director, Research Analytics
Clarivate Analytics

Margaret Shupnik

Sr. Associate Dean for Research
University of Virginia School of
Medicine

Tobin Smith

Vice President for Policy
Association of American
Universities

Courtney Soderberg

Statistician and Methodologist
Center for Open Science

William Southerland

Professor & Director
Howard University College of
Medicine

Emily Statham

Project Coordinator
MRCT Center

Katie Steen

Policy Associate
Association of American
Universities

Deborah Sweet

Vice President of Editorial
Cell Press

Darren Taichman

Executive Editor
Annals of Internal Medicine

Neil Thakur

Special Assistant
NIH

Dan Valen

Director, Business Development &
Strategy
Figshare

Valda Vinson

Deputy Editor, Research
Science Magazine

Joanne Waldstreicher

Chief Medical Officer
Johnson & Johnson

John Wilbanks

Chief Commons Officer
Sage Bionetworks

Kathleen Zackowski

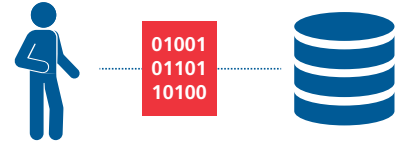
Director, Patient Mgmt Care &
Rehabilitation Res
National MS Society

CREDIT FOR DATA SHARING

PROCESS AND ROLES

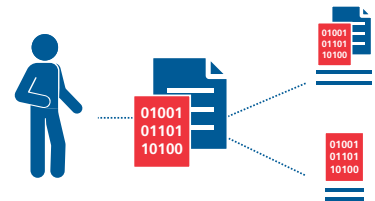
1

Researchers deposit datasets into repositories, associating each dataset with one or more ORCID iDs, which identify individuals. Repositories assign unique persistent identifiers (PIDs) to each dataset.



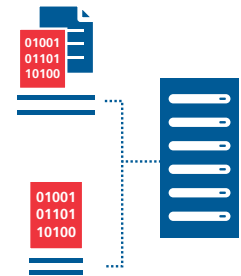
2

Authors submit the dataset PID(s) with each manuscript to enable future tracking.



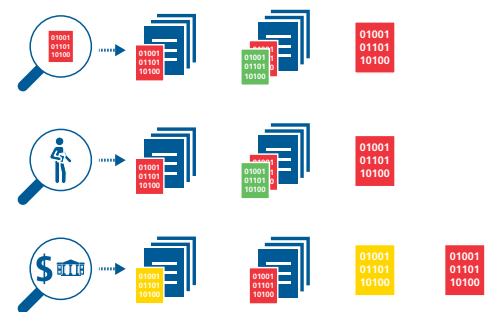
3

Publishers make links between papers and datasets available in the citation metadata.



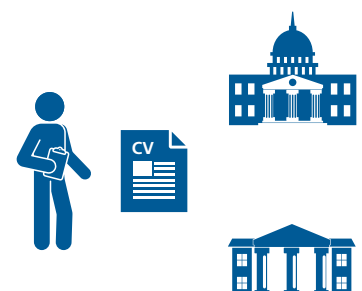
4

Searchable systems allow tracking of datasets through dataset PID, ORCID iD, or grant or organization identifiers.



5

Individuals and organizations use citation information and metrics to assess impact of shared data.



Credit for Data Sharing: Recommended Stakeholder Actions

Process Step	Stakeholder	Action	Guidelines and Current Initiatives
Researchers deposit datasets into repositories, associating each dataset with one or more ORCID iDs, which identify individuals. Repositories assign unique persistent identifiers (PIDs) to each dataset.	Researcher	Deposit dataset and associated metadata in an appropriate general purpose or discipline-specific repository.	Springer Nature recommended repositories ; FAIRsharing.org FAIR (Findable, Accessible, Interoperable, Usable) Guiding Principles for Scientific Data Management and Stewardship (2016)
	Repository	Meet the minimum requirements to support the FORCE11 Joint Declaration of Data Citation Principles (JDDCP), including the assignment of a globally unique persistent identifier (PID) to the dataset, which resolves to a landing page; and provision of the appropriate metadata for dataset discoverability on dataset landing pages.	A Data Citation Roadmap for Scholarly Data Repositories (2019) DataCite
	Repository	Collect one or more ORCID iDs through authenticated pathways and link them to the dataset and share dataset metadata with the researcher via their ORCID record.	ORCID Workflow: Repository Systems
	Repository	Facilitate the consistent and credible reporting of research data usage.	Code of Practice for Research Data Usage Metrics (2018)
Authors submit the dataset PID(s) with each manuscript to enable future tracking.	Researcher	Formally cite the dataset underlying a specific claim in the manuscript in the reference section, including author, year, title, PID, version, and repository and following publisher guidelines.	FORCE11 Citation Principles
	Publisher	During pre-submission, submission, production, and publication, implement data citation and related policies in accordance with the JDDCP. Publishers can enable researchers to share dataset metadata directly from the ORCID	A Data Citation Roadmap for Scientific Publishers (2018) ORCID Workflow: Publishing Systems

Process Step	Stakeholder	Action	Guidelines and Current Initiatives
		record. Include a Data Availability Statement in the publication.	
	Publisher	Assert that the deposition of data in a repository does not constitute “publication” that would prevent the submission of research manuscripts based on those datasets.	See, e.g., F1000 Data Policies
Publishers make links between papers and datasets available in the citation metadata.	Publisher	Deposit data citations by including them in the Crossref metadata when registering content. Use of the relationship section to assert the data link is highly recommended, as this allows for identification of any dataset PIDs and is not limited to DataCite DOIs.	Crossref Data & Software Citation Deposit Guide for Publishers
Searchable systems allow tracking of datasets through dataset PID, ORCID iD, or grant or organization identifiers.	Repository, Funder, Institution	Funder platforms, institutional research information management systems, and other technology initiatives can collect citations and usage metrics from DataCite/Crossref EventData and provide them to end-users in a searchable interface which would include datasets whether or not they are linked to publications.	Make Data Count Research Data Alliance
Individuals and organizations use citation information and metrics to assess impact of shared data.	Researcher	Add information/metrics on shared datasets to CVs and funding reports.	
	Funder	Track datasets generated from grants and recognize their value as a stand-alone output of research.	
	Institution	Incorporate metrics from shared datasets into the promotion and tenure process and other evaluations of research impact along with traditional publications.	