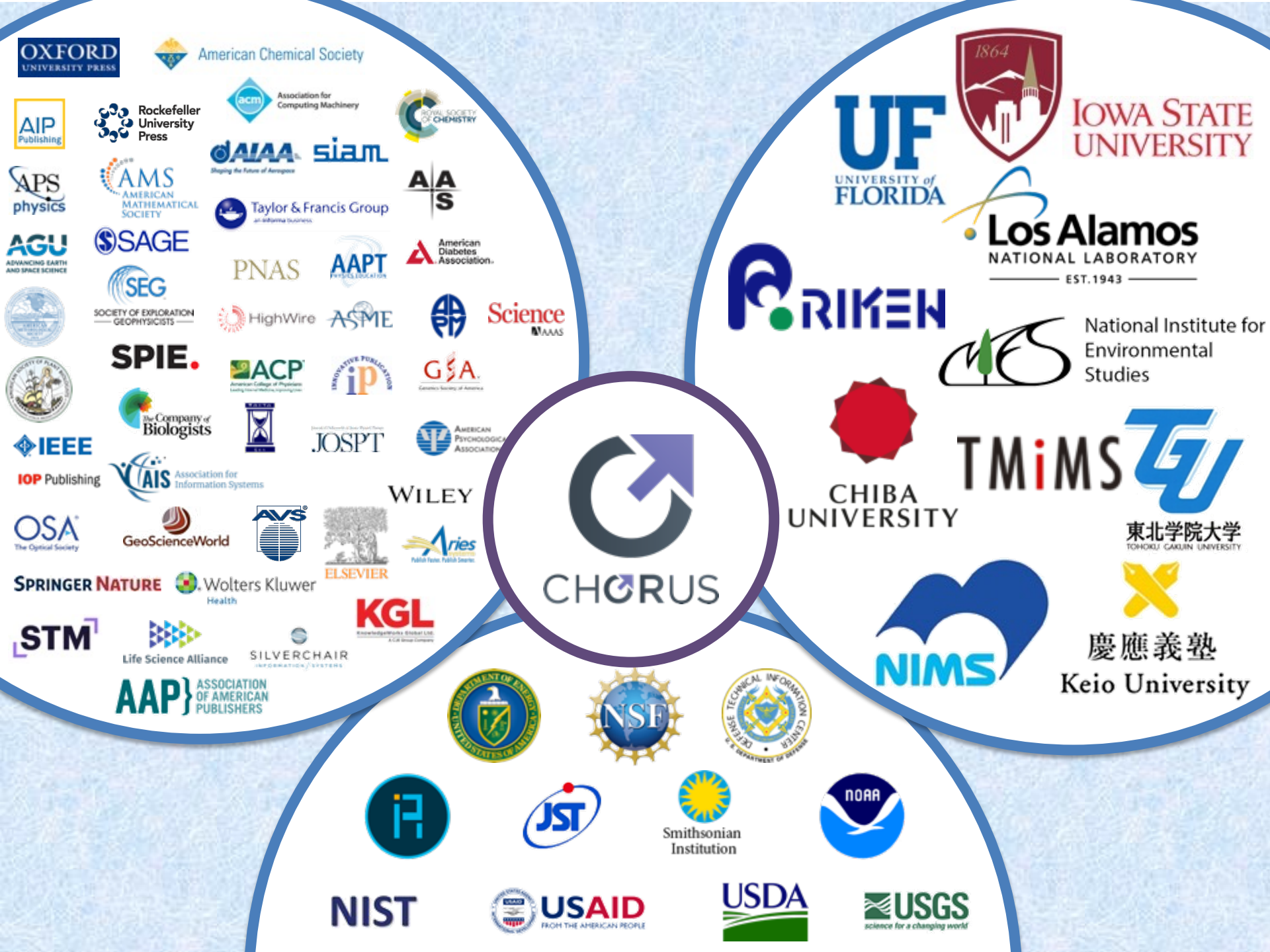
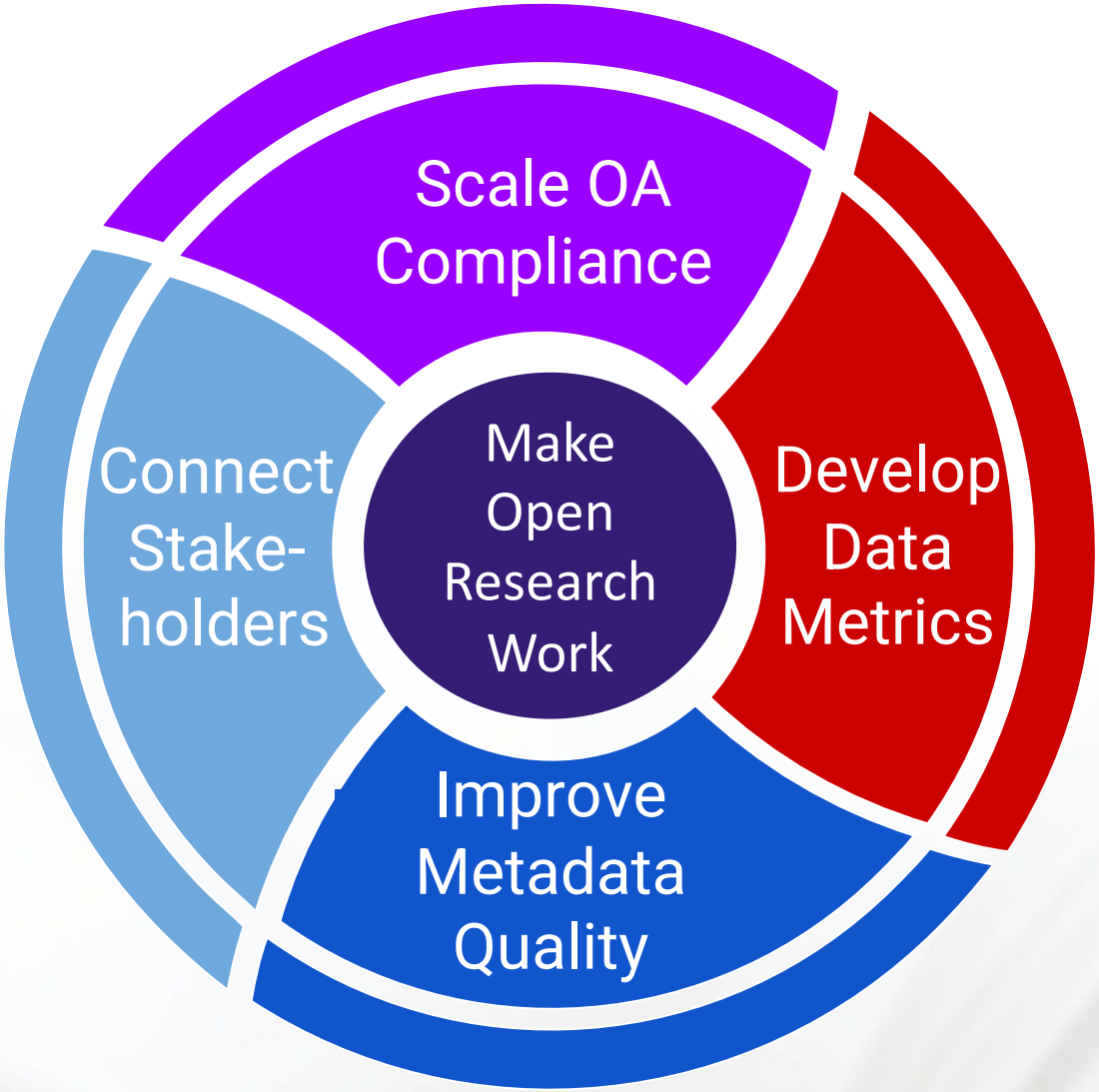
CHORUS and FAIR Data

JOSS / CHORUS Forum: FAIR Data from Funding to Publication
17/18 June 2021

Howard Ratner, Executive Director
hratner@chorusaccess.org
www.chorusaccess.org | @chorusaccess.org


OXFORD UNIVERSITY PRESS
American Chemical Society
AIP Publishing
Rockefeller University Press
Association for Computing Machinery
ROYAL SOCIETY OF CHEMISTRY
AIAA
siam
APS physics
AMS AMERICAN MATHEMATICAL SOCIETY
AAS
Taylor & Francis Group
AGU ADVANCING EARTH AND SPACE SCIENCE
SAGE
American Diabetes Association.
PNAS
AAPT
SEG SOCIETY OF EXPLORATION GEOPHYSICISTS
HighWire
ASME
Science AAAS
SPIE.
ACP
iP
GSA
The Company of Biologists
JOSPT
IEEE
AMERICAN PSYCHOLOGICAL ASSOCIATION
IOP Publishing
AIS Association for Information Systems
WILEY
OSA The Optical Society
GeoScienceWorld
AVS
Aries
ELSEVIER
SPRINGER NATURE
Wolters Kluwer Health
KGL
STM
Life Science Alliance
SILVERCHAIR
AAP ASSOCIATION OF AMERICAN PUBLISHERS
1864
UF UNIVERSITY of FLORIDA
IOWA STATE UNIVERSITY
Los Alamos NATIONAL LABORATORY EST. 1943
RIKEN
National Institute for Environmental Studies
CHIBA UNIVERSITY
TMiMS
東北学院大学 TOHOKU GAKUIN UNIVERSITY
NIMS
慶應義塾 Keio University
CHORUS
NSF
Smithsonian Institution
NOAA
JST
NIST
USAID FROM THE AMERICAN PEOPLE
USDA
USGS science for a changing world

CHORUS 2021 Goals

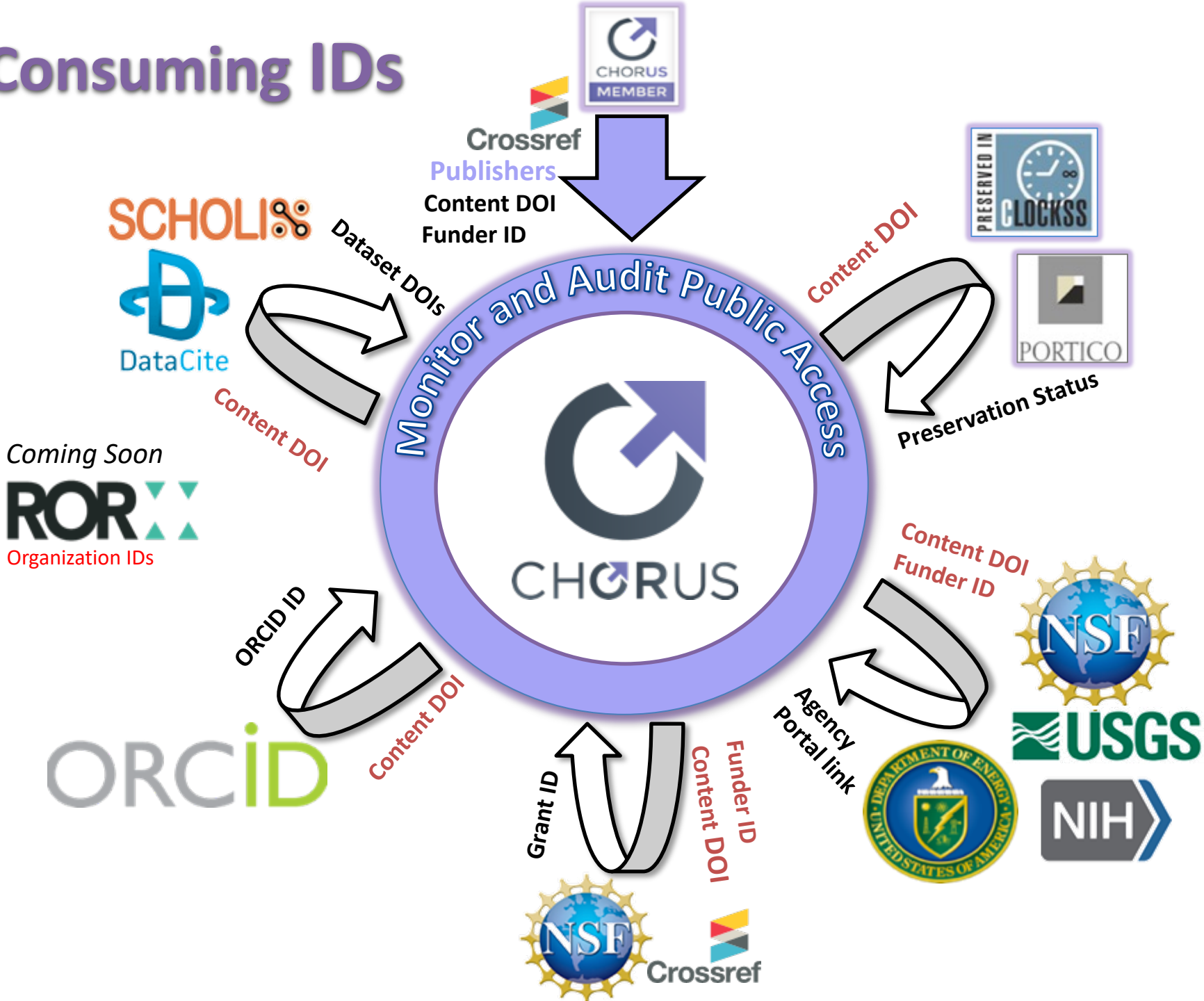


PUBLISHERS

INSTITUTIONS

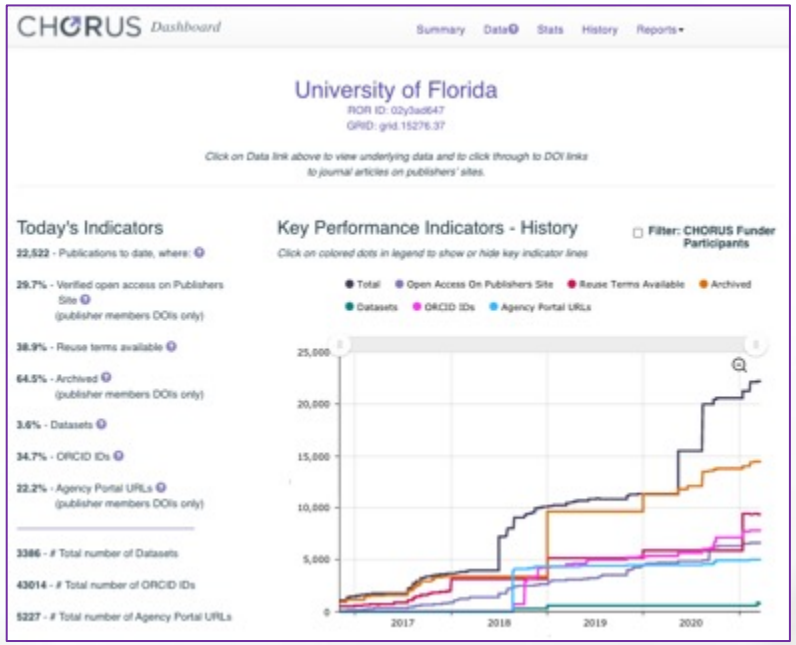
FUNDERS

Consuming IDs



Publication and Compliance Data Services

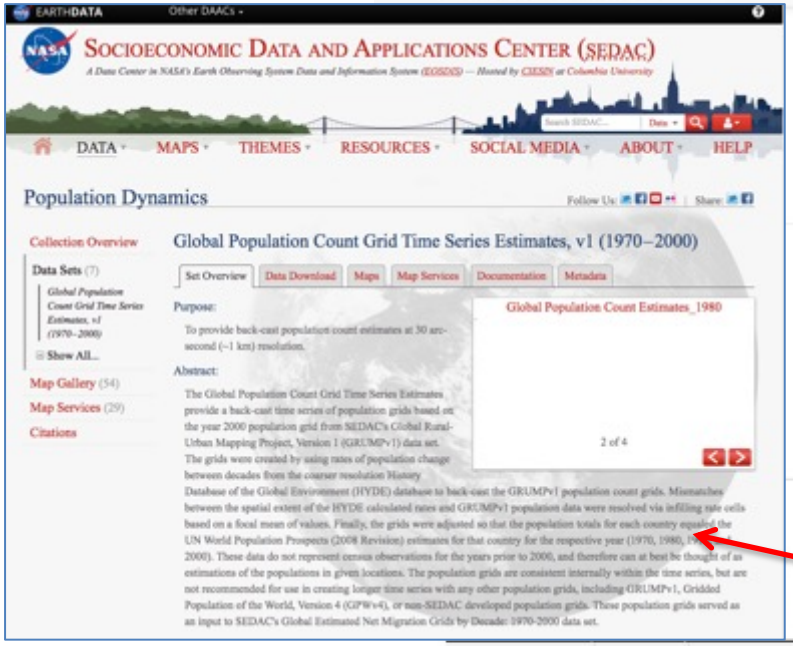
API, metadata feeds and dashboards for monitoring and tracking open research information



- Funder Reports
- Institution Reports
- Publisher Reports
- Dataset Reports

CHORUS Dashboard

Summary Data Stats History Reports



University of Florida

Record level information about Publisher Member content.
Click on DOIs to be taken to articles on publisher site.

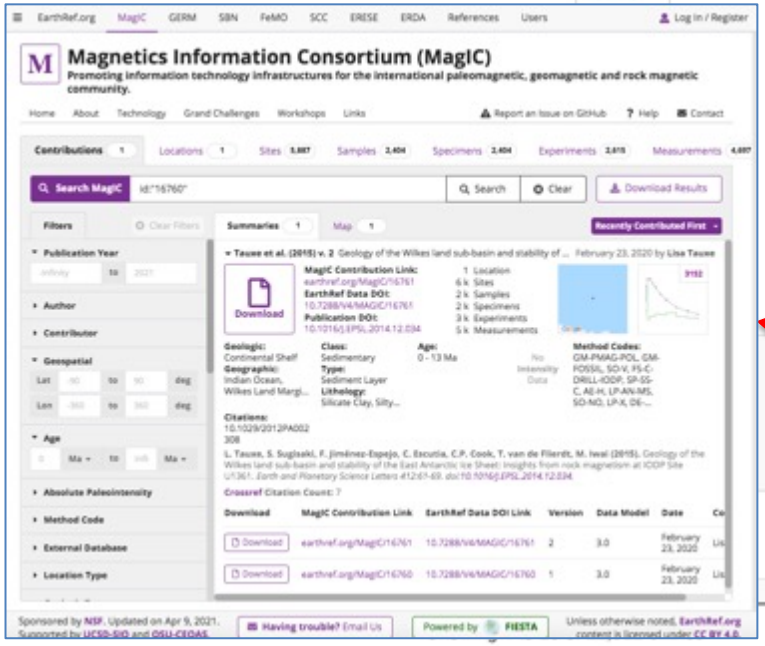
Indicator History Details

3/25/2021

VERIFIED OPEN ACCESS ON PUBLISHERS SITE - YES

Search:

Updated	Grant ID	Archived	Datasets (beta)
3/7/2021	NIH:U54GM111274 NIH:R01 AI102939-05	Yes	http://dx.doi.org/10.7927/h4cc0xnv
	NSF:EAR1141840 NSF:OCE1058858 NSF:1141840 NSF:1058858	Yes	http://dx.doi.org/10.7288/v4/magic/16749 http://dx.doi.org/10.7288/v4/magic/12152 http://dx.doi.org/10.7288/v4/magic/16760 http://dx.doi.org/10.7288/v4/magic/16761
	NSF:(EAR0608998 and EAR0609339)	Yes	http://dx.doi.org/10.7288/v4/magic/16345 http://dx.doi.org/10.7288/v4/magic/16347 http://dx.doi.org/10.7288/v4/magic/16346



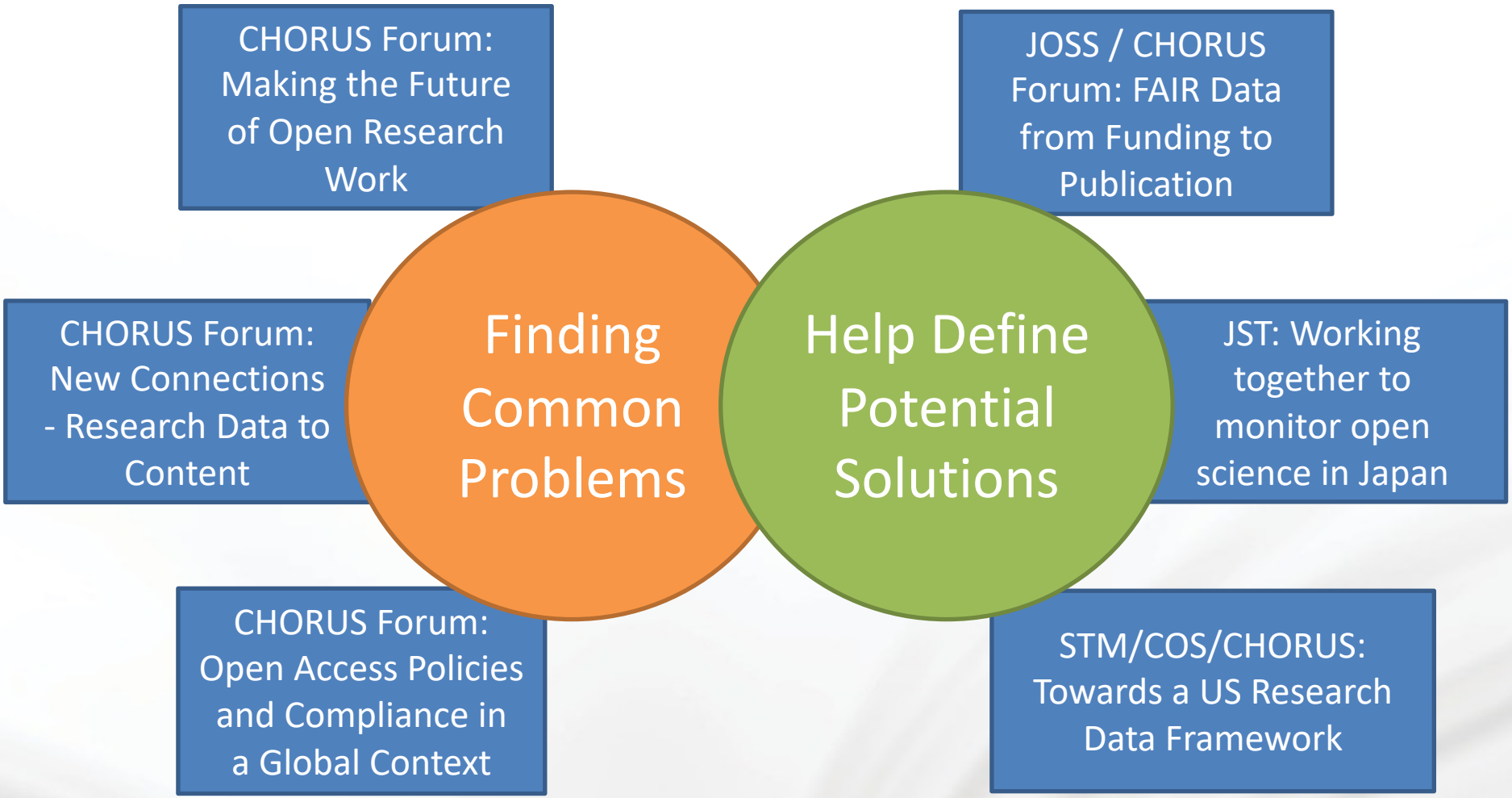
Previous 1 2 3 4 5 … 901 Next

Enhanced Dataset Reporting

NEW

Article DOI	Journal Title	Agency	Dataset Creators	Dataset DOI	Dataset Type	Dataset Repository Name	Dataset Title	Date Collected at Repository	Funder Name	Award #	Award Title	Rights
10.1073/pnas.1604386113	Proceedings of the National Academy of Sciences	National Oceanic and Atmospheric Administration,National Science Foundation	Ganguli, Poulomi; Paprotny, Dominik; Hasan, Mehedi; Güntner, Andreas; Merz, Bruno	10.5880/gfz.4.4.2019.003		GFZ Data Services	Compound flood drivers for northwestern Europe in high-resolution EURO-CORDEX Simulations	2019	Alexander von Humboldt Foundation, Germany	D05417008	Humboldt Research Fellowship for early career researchers	CC BY 4.0
10.1186/s12862-019-1465-5	BMC Evolutionary Biology	Directorate for Biological Sciences,National Oceanic and Atmospheric Administration	Schott, Eric; Lella, Santiago; Tsvetan Bachvaroff; L. Amzel; Vasta, Gerardo	10.6084/m9.figshare.c.4583777.v1	Collection	Figshare	Lacking catalase, a protistan parasite draws on its photosynthetic ancestry to complete an antioxidant repertoire with ascorbate peroxidase	2019				CC BY 4.0
10.1242/jeb.189233	The Journal of Experimental Biology	Hampden-Sydney College; Harris Award; National Oceanic and Atmospheric Administration; North Slope Borough, Alaska,	Werth, Alexander J.; Rita, Diego; Rosario, Michael V.; Moore, Michael J.; Sformo, Todd L.	10.5061/dryad.73rm81p	dataset	Dryad		2018	National Science Foundation	1656691		CC0 1.0 Universal (CC0 1.0) Public Domain Dedication

Convening Stakeholder Discussions

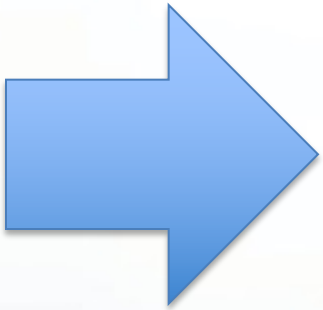


Helping to improve metadata

Crossref

DataCite
FIND, ACCESS, AND REUSE DATA

Gap Analysis
- ORCID IDs
- Reuse License Metadata
- Publication Dates



Gap Analysis (in planning)
- ORCID IDs?
- Creators?
- Institutions?
- Funders?

CHORUS Publisher Data Availability Policies Index

NEW!

PUBLISHER DATA AVAILABILITY POLICIES INDEX

Over the last few years, publishers have been making their Data Availability Policies known either at the publisher level or at the journal level. These policies range in their mandate, but most require authors to make all data necessary to replicate their study's findings publicly available without restriction at the time of publication. When specific legal or ethical restrictions prohibit public sharing of a data set, authors must indicate how others may obtain access to the data. Often, when submitting a manuscript to the publisher, authors must provide a Data Availability Statement describing compliance with the data availability policy. If the article is accepted for publication, the Data Availability Statement will be published as part of the article.

CHORUS has created a centralized index of our member publishers' policies with links to the publisher's site. This chart will be updated at least annually.

Please let us know about any updates or requested enhancements by contacting us at info@chorusaccess.org.

CHORUS Publisher Member	Data Availability Policy URL	Policy Type
AAAS/Science	https://www.sciencemag.org/authors/science-journals-editorial-policies	Publisher
American Association of Physicists in Medicine	https://www.aapm.org/org/policies/details.asp?id=329&type=PP	Publisher
American Association of Physics Teachers	TBD	
American Astronomical Society	https://journals.aas.org/policies/	Publisher
Association for Computing Machinery	https://www.acm.org/acm-policies-and-procedures	Publisher
American College of Physicians	https://annals.org/aim/pages/authors	Publisher
American Chemical Society	https://publish.acs.org/publish/publishing_policies https://publish.acs.org/publish/author_guidelines?coden=acscii#data_requirements	Publisher
American Diabetes Association	https://professional.diabetes.org/sites/professional.diabetes.org/files/media/2018_data_resource_sharing_policy.pdf	Publisher
American Geophysical Union	https://www.agu.org/Publish-with-AGU/Publish/Author-Resources/Data-for-Authors	Publisher
American Institute of Aeronautics and Astronautics	https://www.aiaa.org/publications/books/Publication-Policies	Publisher
AIP Publishing	https://publishing.aip.org/resources/researchers/open-science/research-data-policy/	Publisher
American Meteorological Society	https://www.ametsoc.org/index.cfm/ams/publications/ethical-guidelines-and-ams-policies/data-archiving-and-citation/	Publisher
American Mathematical Society	TBD	
American Psychological Association	https://www.apa.org/research/responsible/data	Publisher
American Physical Society	TBD	
American Society of Mechanical Engineers	https://www.asme.org/publications-submissions/journals/information-for-authors/data-policy	Publisher

https://www.chorusaccess.org/resources/chorus-for-publishers/publisher-data-availability-policies-index/

Accelerating Open and FAIR Data Practices Across the Earth, Space, and Environmental Sciences: A Pilot with the NSF to Support Public Access to Research Data

Grant 2025364

Two-year project aimed at implementing FAIR data practices across the Earth, space, and environmental sciences such that, by the end of the project:

- Scope: AGU data citations for research funded by NSF grants are captured in the NSF Public Access Repository (PAR 2.0)
- Knowledge of leading practices and workflows around data citation are well known across the AGU community
- Ensure NSF earth science data are publicly accessible, linked to publications and grants, and credited
- Present dataset reports:
  - linked articles,
  - name and type of dataset,
  - repository and creator names,
  - subject classification, and
  - reuse license information
- Help AGU and others improve data citation deposit into the NSF Public Access Repository (NSF PAR)
- Develop and further open data citation best practices

Participants:

AGU | NSF | CHORUS | Dryad | Earth Science Information Partners (ESIP) | Wiley

All learnings and practices from the work will be openly shared

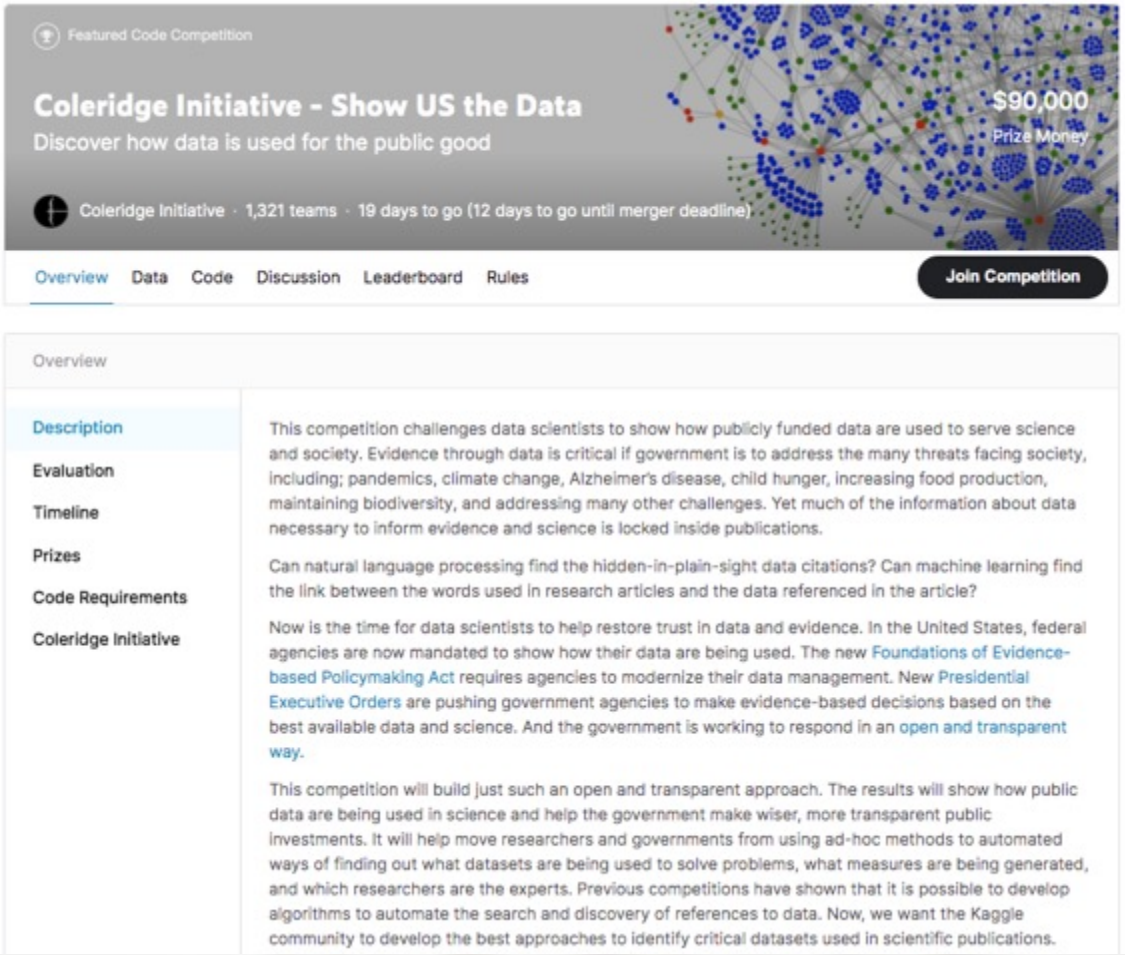
https://fromtheprow.agu.org/nsf-grant-takes-agus-data-and-publishing-practices-to-the-next-level/

COLERIDGE INITIATIVE

Rich Context Project

Apply machine-learning and natural language processing techniques that searches publications to

- Find what datasets are in the publications
- Show how they have been used
- Find other experts who have used the data
- Identify other related datasets
- Show impact of funded datasets



https://www.kaggle.com/c/coleridgeinitiative-show-us-the-data/

How might we?

Project Goal

How might we enable government agencies to show how, why and by whom their data are being used?

The objective + ask

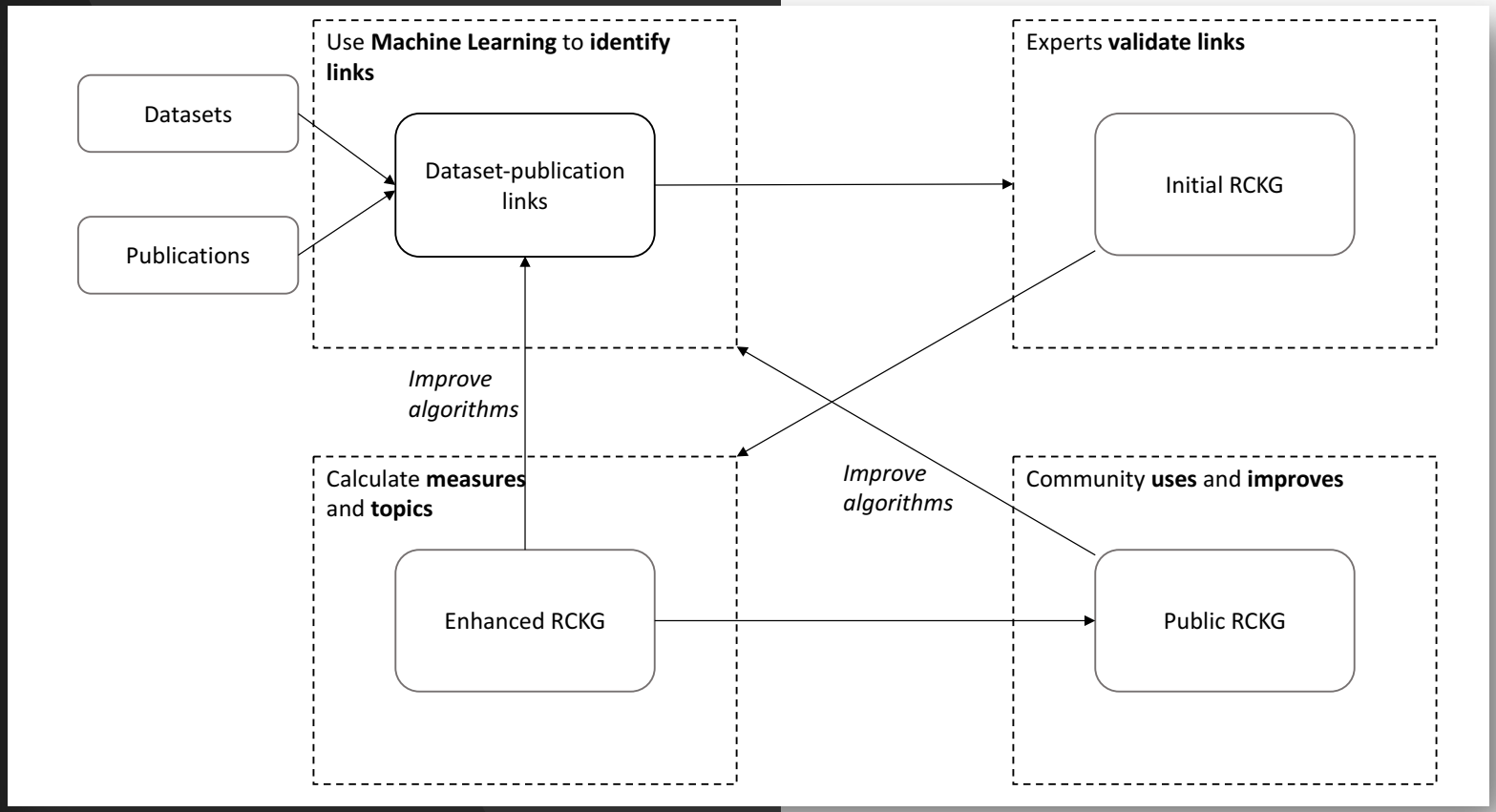
Demonstrate the value of data as a strategic asset to the public and legislators.

Design a data scorecard that can serve as a source of information about users, data sources and applications of data.

Credit: Julia Lane, Coleridge Initiative

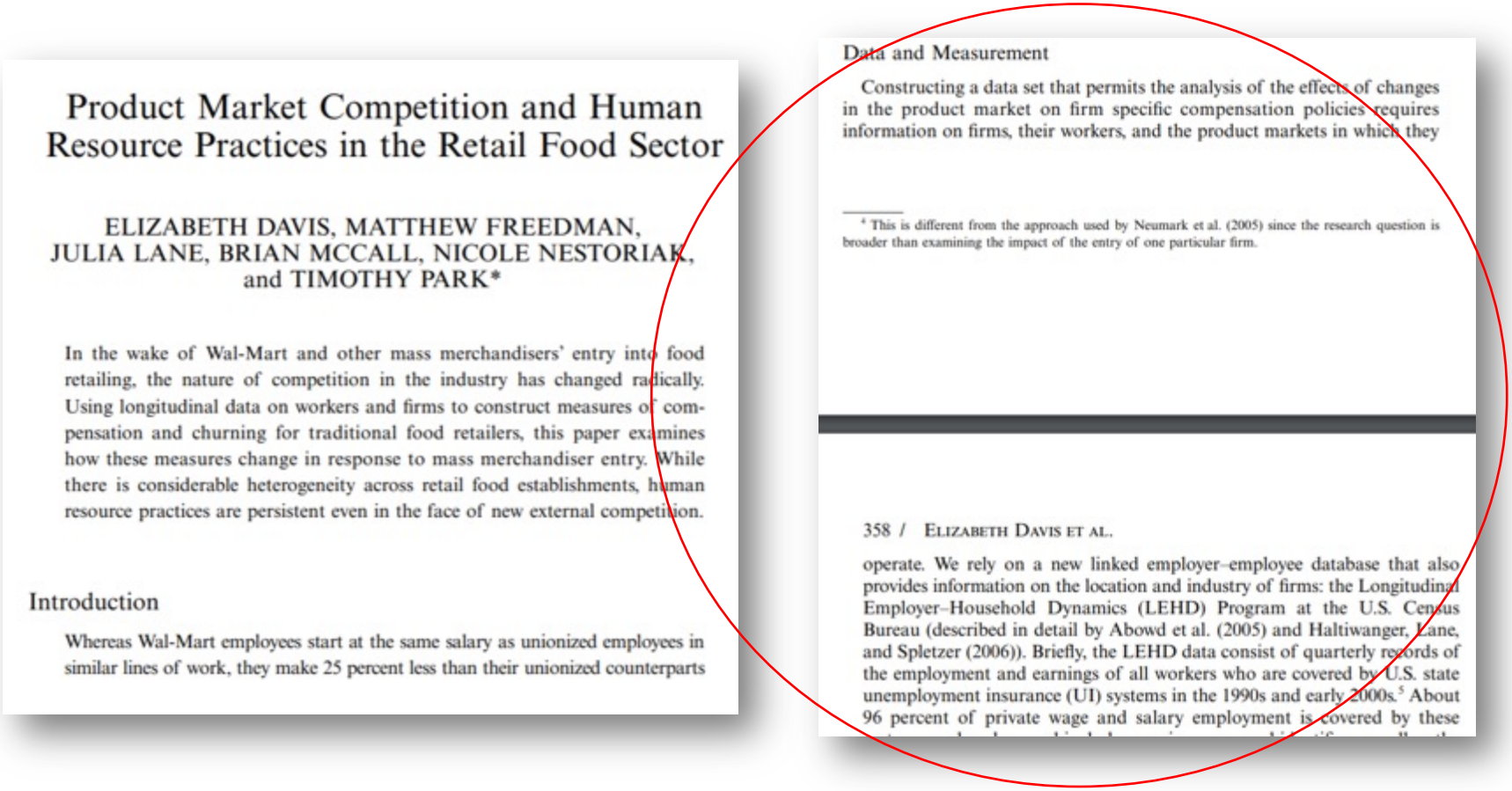
COLERIDGE INITIATIVE

Improve model through AI and community



Credit: Julia Lane, Coleridge Initiative

Approach: Deploy new technology?



Product Market Competition and Human Resource Practices in the Retail Food Sector

ELIZABETH DAVIS, MATTHEW FREEDMAN, JULIA LANE, BRIAN MCCALL, NICOLE NESTORIAK, and TIMOTHY PARK*

In the wake of Wal-Mart and other mass merchandisers' entry into food retailing, the nature of competition in the industry has changed radically. Using longitudinal data on workers and firms to construct measures of compensation and churning for traditional food retailers, this paper examines how these measures change in response to mass merchandiser entry. While there is considerable heterogeneity across retail food establishments, human resource practices are persistent even in the face of new external competition.

Introduction

Whereas Wal-Mart employees start at the same salary as unionized employees in similar lines of work, they make 25 percent less than their unionized counterparts

Data and Measurement

Constructing a data set that permits the analysis of the effects of changes in the product market on firm specific compensation policies requires information on firms, their workers, and the product markets in which they

[4] This is different from the approach used by Neumark et al. (2005) since the research question is broader than examining the impact of the entry of one particular firm.

358 / Elizabeth Davis et al.

operate. We rely on a new linked employer–employee database that also provides information on the location and industry of firms: the Longitudinal Employer–Household Dynamics (LEHD) Program at the U.S. Census Bureau (described in detail by Abowd et al. (2005) and Haltiwanger, Lane, and Spletzer (2006)). Briefly, the LEHD data consist of quarterly records of the employment and earnings of all workers who are covered by U.S. state unemployment insurance (UI) systems in the 1990s and early 2000s.[5] About 96 percent of private wage and salary employment is covered by these

Credit: Julia Lane, Coleridge Initiative

What does such a search generate?

- Dataset name (and aliases)
- Research topics
- Authors (and associated metadata)
- Related datasets
- Publication name (and associated metadata)

COLERIDGE INITIATIVE

Credit: Julia Lane, Coleridge Initiative

Usage Indicators for USDA Agricultural Resource Management Survey

Used by 232 Publications based on verified attributions for v.2, updated 01-02-2021

17	539	86	3
Topics	Authors	Journals	With Other Datasets

TOPICS	17	AUTHORS	539	JOURNALS	86	WITH OTHER DATASETS	3
Agricultural Researc...	73	Ashok K. Mishra	13	SSRN Electronic Jour...	48	Census of Agricultur...	41
Farm Household Well-...	54	William D. McBride	13	Journal of Agricultu...	27	Rural-Urban Continuu...	4
Organic Agriculture	34	Nigel D. Key	12	Agricultural and Res...	22	Survey of Industrial...	1
Poverty & Income Vol...	16	Hisham S. El-Osta	11	Agricultural Finance...	8		
Climate Change	9	James M. MacDonald	8	Sustainability	8		
Food Service Industr...	8	Ani L. Katchova	8	American Journal of ...	5		

Proof of Concept

COLERIDGE INITIATIVE

Credit: Julia Lane, Coleridge Initiative

Join Our Community!

Contact us today

Howard Ratner, Executive Director
hratner@chorusaccess.org

Mark Robertson, Director Asia-Pacific Development
mrobertson@chorusaccess.org

www.chorusaccess.org |@chorusaccess.org