### INCENTIVIZING AGRICULTURAL RESEARCHERS

### TO SHARE OPEN & INTEROPERABLE DATA

### **MEDHA DEVARE**

Senior Research Fellow | IFPRI Module Lead | CGIAR Platform for Big Data in Agriculture

CABI: Opening up Agricultural Research and Data (Nov 11, 2020)

## **A FUTURE FARM** Feeding the future. Byte by byte.

#### SURVEY DRONES

Aerial drones survey the field to map crop health and yield, and soil variation. **Cloud storage** enables faster data movement, analysis and exchange, **increasing computational capacity** while decreasing costs.

### **BIG DATA-DRIVEN SOLUTIONS**

Find, access data Interpret Aggregate Visualize Analyze

#### FARM TO CLOUD

Farms generate vast amounts of rich and varied data, which can now be stored, processed, and shared in the cloud, to be delivered back to farms as important advisory services.

Platform for Big Data in Agriculture

CGIAI



#### LIVESTOCK LIVE

Feeling a bit mad today!

> Decreasing costs of small devices, such as mobile phones and sensors, are increasing access to get rural farmers - and their livestock - online and benefiting from data exchange.

#### **UBER AGRIBOTS**

Peer-to-farmer services allow smallholder farmers access to cost-effective cutting edge technology such as agribots that can tend to or harvest crops, and provide precise application of fertiliser and water, saving economic and environmental costs. ·--((①)) •

### Hey Cigi, When should I plant my maize? How should I manage my crop?

- Real-time decision support for farmers
- Easy natural language as an interface
- Smart artificial intelligence trained by CGIAR and partners
- ...requires leveraging multiple open, harmonized, interoperable databases with quality data



We need data on crop management and yield response to validate our model and DSS. Help??

Where in Sub-Saharan Africa is fertilizer use profitable?

Where do I find data or pubs on nutrition status in rural households?

Open data... that is also Findable Accessible Interoperable Reusable Ethical Reproducible

### WHAT DO I NEED TO DO?

• Clear policy (2013; 2021); digital strategy (2021)

### HOW DO I DO IT?

 Capacity enhancement: guidelines to implement; webinars; workshops; data curathons/sprints; bespoke help sessions...



All courses > FAIR Data Management

Next Session

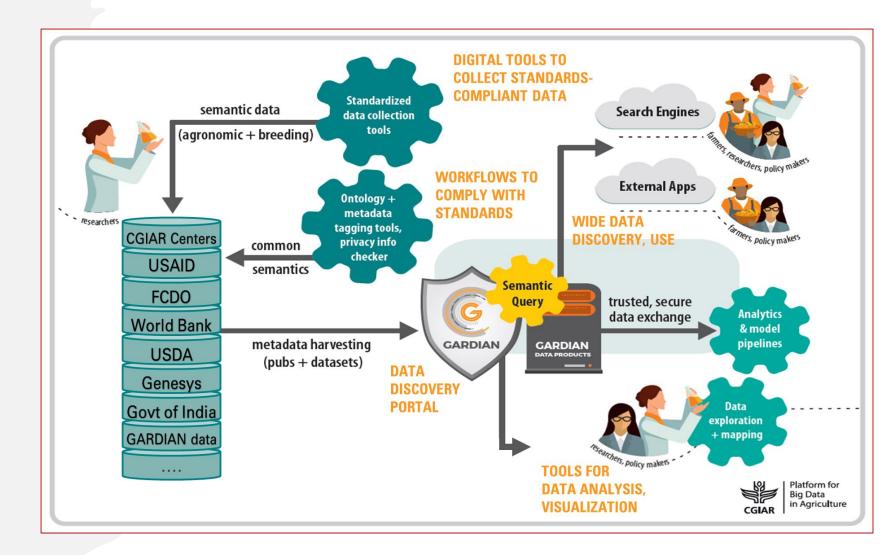
Started Jan 8, 2020 Change

JOIN COURSE

Data assets that are interpretable and interoperable for humans and machines are critical to enhance the impact of research in the agricultural domain and to catalyze innovation and transform agricultural reserch for development. This course has been developed to strengthen capacity in the CGIAR System and beyond to create, manage and share research and development data assets that are not only open (i.e., discoverable and downloadable), but also easily interpretable, interoperable, and reusable. The FAIR Principles embody the guidelines to make resources Findable, Accessible, Interoperable, and Reusable (FAIR); this course is intended to provide more concrete "how to" help in operationalizing these, and ensuring that assets are responsibly managed.

### HOW DOIDOIT?

 Tools, services, data processing scripts that are easy to use



## **GARDIAN DATA PORTAL**



Global Agricultural Research Data Innovation Acceleration Network



#### gardian.bigdata.cgiar.org

#### **EXPLORE**

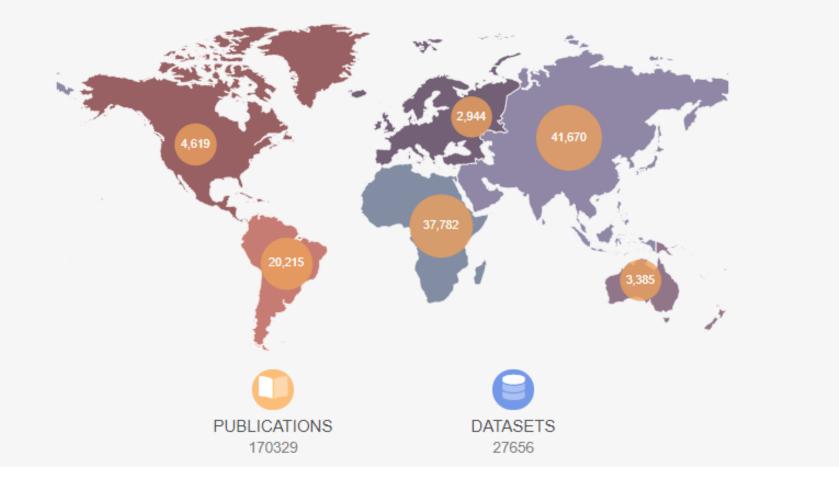
- publications and datasets
- global crop production data
- global climate model projections

#### COLLABORATE

on a data science project

#### EXPLORE GARDIAN'S ASSETS BY GEOGRAPHY

Explore data assets from across CGIAR and a growing set of institutional partners, including: USAID, the UK's Foreign, Commonwealth & Development Office, the World Bank, the US Department of Agriculture, the Indian Council for Agricultural Research, and the Open Government Data Platform India.



#### GET HELP

\_225\_

- with your metadata
- with your PII
- with your FAIRness

### HOW DOIDOIT?

• FAIR guidance and workflows



"I've been told I need to make my data FAIR. How do I do this?"



#### bigdata.cgiar.org/resources/agronomy-ontology

OUT AGRO BROWSE DO

#### ABOUT AgrO

An ontology is a formal representation of a disciplinary domain, rep standard that can be employed to annotate data where key concep as the relationships that exist between those concepts (Gruber, 20) a common language for different kinds of data to be easily interpre interoperable allowing easier aggregation and analysis.

The Agronomy Ontology (AgrO) provides terms from the agronom semantically organized and can facilitate the collection, storage and data, enabling easy interpretation and reuse of the data by humans

To fully understand the implications of varying practices within cro derive insights, it is often necessary to pull together information fro disciplinary domains. For example, data on field management, soil, a phenotypes may need to be aggregated to assess performance of p different management interventions.

However, agronomic data are often collected, described, and storer impeding data comparison, mining, interpretation reuse. The use of



www.cropontology.org



#### **GARDIAN** FAIRification Workflow



gardian.bigdata.cgiar.org/fair-workflow.php

"Are there tools to help me collect standardscompliant agronomic data?"

#### AgroFIMS Agronomy Field Information Management System to easily produce Findable, Accessible, Interoperable, and Reusable (FAIR) data Aligns a priori with Enables digital Standardizes data Allows data quality CGIAR's CG Core collection of checks, statistical collection and metadata schema agronomy trial data annotation reports AgroFIMS workflow Fieldbook creation using ontology-based variables, terminology, and units in modules representing typical cycle of operations in agronomic trials Digital data collection with KDSmart mobile app (Diversity Arrays Technology) - and soon with other digital data collection apps Data quality checks, analysis via AgroFIMS statistical scripts N (R-based) and reports

Data archiving through easy upload to institutional repository

### agrofims.org

### AgroFIMS

Agronomy Field Information Management System

Generate standardized field books to collect agronomic data that is...



### HOW DO I DO IT?

- Capacity enhancement
- Tools, services, data processing scripts that are easy to use
- FAIR guidance and workflows
- Publish datasets data journals or accompaniments to publications
- DMPs Open and FAIR from project conception (Funder requirements would help!)
- Budgetary recommendations and guidance (Funder requirements?!)



### ...WHY DO I NEED TO DO IT?

- Open, FAIR data has transformed other sectors (biomedical, environmental research...)
- Enables realization of the value of data (exemplars; data products that can be leveraged)

Develop A search of the analysis of the control of the analysis of th	All Databases  All Databases					Search)			
Resource List (A-2)     The National Centre for Blocedbrology Information advances science and health ity providing access to thomescience and presents it Stronger St		Get	test public health information from CDC tibe latest research from NH1 https://w	hites./www.coronaverus.yov. www.nih.gov/coronaverus	an survey 2				
Rescarce List (A-2)       The stateout Correct for Stateout Correct for Stateout Correct (List (A-2))       PAMMed         All Rescarces       Correction & Stateout Correct for Stateout Correct (List Rescarce As Biog       Decision Stateout Correct (List Rescarce As Biog         Data & Stateout Correct for Stateout Correct (List Rescarce As Biog       Decision Stateout Correct (List Rescarce As Biog       Decision Stateout Correct (List Rescarce As Biog         Data & Stateout Correct for Stateout Correct (List Rescarce As Biog       Decision Stateout Correct (List Rescarce As Biog       Decision Stateout Correct (List Rescarce As Biog         Operation & Stateout Correct for Stateout Correct (List Rescarce As Biog       Decision Stateout Correct (List Rescarce As Biog       Decision Stateout Correct (List Rescarce As Biog         Operation & Stateout Correct (List Rescarce As Biog       Decision Stateout Correct (List Rescarce As Biog       Decision Stateout Correct (List Rescarce As Biog         Search (P)       Data stateout Correct (List Rescarce As and Rescarce As		Welcome to NCBI				Popular	Descurras	_	_
All Recordsort Distribution   Chemical & Bichwank & Biogi   Data Bickwank & Biogi	rt (A-Z)	The National Center for Biotechnology Information advances science and health by providing access to biomedical and genomic information.							
Cherricals & Biosinayis   Data & Storware   Data Storware   Data Storware   Data Storware   Data Storware <td>1.000</td> <td colspan="4"></td>	1.000								
Data & Dothwam Submit Download Lean   Data & Dothwam Submit Download Lean   Denses & Expression Dense data or manunolish Download Lean   Denses & Expression Dense data or manunolish Download Lean   Denses & Expression Dense data or manunolish Download Lean   Denses & Expression Dense data or manunolish Dense data or manunolish Dense data or manunolish   Denses & Expression Dense data or manunolish Dense data or manunolish Dense data or manunolish   Denses & Expression Dense data or manunolish Dense data or manunolish Dense data or manunolish   Denses & Expression Dense data or manunolish Dense data or manunolish Dense data or manunolish   Denses & Expression Dense data or manunolish Dense data or manunolish Dense data or manunolish   Denses & Expression Dense data or manunolish Dense data or manunolish Dense data or manunolish   Dense data or manunolish Dense data or manunolish Dense data or manunolish Dense data or manunolish   Dense data or manunolish Dense data or manunolish Dense data or manunolish Dense data or manunolish   Dense data or manunolish Dense data or manunolish Dense data or manunolish Dense data or manunolish   Dense data or manunolish Dense data or manunolish Dense data or manunolish Dense data or manunolish   Dense data or manunolish Dense data or manunolish Dense data or manunolish Dense dat									
Constraints & Structures   Generals & Structures   Financeg & Tutoruse   Matterion   Bacchernin   Generals & Structures   Develop   Matterion   Bacchernin   Matterion   Bacchernin		Construction ( summer ) and	Personal Construction of the second			BLAST			
Derivative & Structures   Genes & Structures   Genes & Structures   Genes & Structures   Genes & Structures   Process & Mage   Honology   Jamandre   Process & Mage   Honology   Jamandre   Process & Mage   Proc		Submit	Download	Learn		Nucleotide			
Centers & Expension   Genes & Expension   Genes & Mage   Including   Jension   Seconders & Mage   Noticity   Including   Jension   Seconders & Mage   Noticity   Seconders & Mage   Seconders & Mage   Seconders & Mage </td <td>tructures</td> <td></td> <td></td> <td></td> <td colspan="2"></td> <td colspan="3"></td>	tructures								
Beretiske B. Modeline Beronnes & Moge Hennbarg Montania Monta									
Seconders Maps   Nonclogy Immunolity   Nonclogy Immunolity   Nonclogy Immunolity   Nonclogy Immunolity   Nonconvy Immunolity   Namong & Tubutak Nonclogy   Neaton Immunolity   Naton Immunolity   Immunolity Immunolity   Immuno			-						
Hondogy densitive Provides Becarino Analyze Tearing 5 Tutorials Metation	Aapa		4						
Proteins       Develop       Analyze       Research         Baconory       Lise NCBL Affins and coon       Mentify an NCBL tool tor your       Epicre NCBL research and       Gerdamit yeases 10 tolling bacepated         Materian       Materian       Materian       Materian       Baconory       Research       Beginse NCBL research and       Gerdamit yeases 10 tolling bacepated         Materian		Т	<u> </u>	1		PubChem			
Develop       Analyze       Research       Gerdank 240.0 is available and supasses 10 tillion bisepairs         haverog & Tuonals       Marking & Tuonals       Marking & Tuonals       Explore NEBI research and obstorative projects       Gerdank 240.0 is available and supasses 10 tillion bisepairs         Aration       Gerdank 240.0 is available and supasses 10 tillion bisepairs       Marking & Tuonals       Explore NEBI research and obstorative projects       Gerdank 240.0 is available and supasses 10 tillion bisepairs         Marking & Tuonals       Marking & Tuonals       Marking & Tuonals       Explore NEBI research and obstorative projects       More Tuonals         Marking & Tuonals       Marking & Tuonals       Marking & Tuonals       More Tuonals						PERCENT			_
Develop       Analyze       Research       suppases 10 billion basepaird         Namering & Tutorials       Use NCBI APIs and cool bits are to built applications       Mentify an NCBI tool for your data analysis tak       Explore NCBI research and coldboardive projects       Ger@ask remeases 2400 (1002000)         Wentsion       Bob       ABOUT       DATA       SUBMIT       TOOLS       Jump to:       Dol or ID       Ger         March       Or data       ABOUT       DATA       SUBMIT       TOOLS       Jump to:       Dol or ID       Ger         March       Or data       DATASETS 1 TO 25 OF 27,895       Soft by       Most recent						NCBI New	ws & Blog		
Data attribute       Data attribute       Data attribute       Shelby JOHNSTON. 2020. Bedrock and Detrital Muscovite 403-4356 addbc00884.       Hide Map > 100 or D       Good S         Image: Creator       Image: Creator <td>alysis</td> <td>Develop</td> <td><b>Analize</b></td> <td>Bassassi</td> <td></td> <td>Ger@ank 2</td> <td>240.0 is availa</td> <td>ble and</td> <td></td>	alysis	Develop	<b>Analize</b>	Bassassi		Ger@ank 2	240.0 is availa	ble and	
Interving & Tutontation						15 Nov 200			
Ariadon  November 19 Webburn Aree was  Provember 19 Webburn Ar	tortala				fr and				
ABOUT DATASETS 1 TO 25 OF 27,895     bearch     Bearch Phrase     I     I     I     I     I     I     I     I     I     I     I   I     I     I     I     I     I     I    I   I   I   I   I   I   I </td <td></td> <td></td> <td>3ac</td> <td>1</td> <td></td> <td>prepare ge</td> <td>nome submis</td> <td>sions using</td> <td></td>			3ac	1		prepare ge	nome submis	sions using	
Barta SE IS T 10 25 OF 27,895         Search phrase         1       2       3       1       2         1       2       3       1,116       Next       Sort by Most recent       33       3       2         Iter by:       Image: Shelby JOHNSTON. 2020. Bedrock and Detrital Muscovite 40Ar/30Ar ages in Central and Western Nepal. Knowledge Network for Biocomplexity. umaskt bolibebe6-0482-1682-1682-1682-1682-1682-1682-1682-16				SUBMIT	1988	-		Go	
1       2       3       1,116       Next       Sort by       Most recent                108       43       2         Image: Shelby JOHNSTON, 2020, Bedrock and Detrital Muscovite 40Ar/30Ar ages in Central and Western Nepal, Knowledge Network for Biocomplexity, unual tolbebee-bi58-lede-b344-36lade00884       7       66       5         Image: Data files       Image: Nicola Falco, Haruko Wainwright, Baptiste Deffion, Emmanuel Leger, Image: Nicola Falco, Haruko Wainwright, Baptiste Deffion, Emmanuel Leger, Image: Nicola Falco, Haruko Wainwright, Baptiste Deffion, Emmanuel Leger, Image: Nicola Falco, Haruko Wainwright, Baptiste Deffion, Emmanuel Leger, Image: Nicola Falco, Haruko Wainwright, Baptiste Deffion, Emmanuel Leger, Image: Nicola Falco, Haruko Wainwright, Baptiste Deffion, Emmanuel Leger, Image: Nicola Falco, Haruko Wainwright, Baptiste Deffion, Emmanuel Leger, Image: Nicola Falco, Haruko Wainwright, Baptiste Deffion, Emmanuel Leger, Image: Nicola Falco, Haruko Wainwright, Baptiste Deffion, Emmanuel Leger, Image: Nicola Falco, Haruko Wainwright, Baptiste Deffion, Emmanuel Leger, Image: Nicola Falco, Haruko Wainwright, Baptiste Deffion, Emmanuel Leger, Image: Nicola Falco, Haruko Wainwright, Baptiste Deffion, Emmanuel Leger, Image: Nicola Falco, Haruko Wainwright, Baptiste Deffion, Emmanuel Leger, Image: Nicola Falco, Haruko Wainwright, Baptiste Deffion, Emmanuel Leger, Image: Nicola Falco, Haruko Wainwright, Baptiste Deffion, Emmanuel Leger, Image: Nicola Falco, Haruko Wainwright, Baptiste Deffion, Emmanuel Leger, Image: Nicola Falco, Haruko Wainwright, Baptiste Deffion, Emmanuel Leger, Image: Nicola Falco, Haruko Wainwright, Baptiste Deffion, Emmanuel Leger, Image: Nicola Falco,		DATASETS 1 TO	25 OF 27,895			ice map .»	_	-	
Sheley JOHNSTON, 2020, Bedrock and Derital Missovite     404/30Ar ages in Central and Western Nepal. Knowledge Network     for Biocomplexity, unusid bb/bebe6-b/58-4e8e-b344-38fadbc00884.     7 66 5     Creator     Micola Falco, Haruko Wainwright, Baptiste Deffion, Emmanuel Leger,     hind Reterand at al. 2019, Remote Sensing and Geophysical	350 Q	t 2 3	1,116 Next	Sort by Most recent	-	33	3	2	
Image: Data attribute       for Biocomplexity, umsuid bb/bebe6-btttt-leBe-b344-38ladbc00884.       7       66       5         Image: Data files       Image: Data files       Image: Data files       7       66       5         Image: Data files       Image: Data files       Image: Data files       Image: Data files       7       66       5         Image: Data files       Image: Data files       Image: Data files       Image: Data files       1251       58         Image: Data files       1251       58         Image: Data files       Image: Data files <td></td> <td colspan="4"></td> <td>108</td> <td>43</td> <td>2 0</td> <td>-</td>						108	43	2 0	-
Data files     Image: Creator     7     66     5       ▲ Creator     Junce: Nicola Falco, Haruko Wainwright, Baptiste Dattion, Emmanuel Leger, Units Paterson, et al. 2019, Remote Sensing and Geophysical     +1     1251     58	attribute						2373000	Par 14	t
Nicola Falco, Haruko Wainwright, Baptiste Daffion, Emmanuel Leger,	files					7	66	5	
John Peterson, et al. 2019, Remote Sensing and Geophysical	or					+ 1	1251	58	
101				A Part of the second			7778	137	
Characterization of a Floodplain-Hillslope System in the East River     Watershed, Colorado. ESS-DIVE: Deep Insight for Earth Science Data.									100

doi:10.21952/WTR/1490867, version: ess-dive-61001536a05b57d

Jesse Coleman and Courtney Carothers, 2013. Graying of the Fleet in Alaska's Fisheries Survey Data, 2015. Research Workspace.

0.24431/rw1k45x, version: 10.24431 rw1k45x 20201030T214916Z

202011027180038552

0 1 9

6

da Taxon

Q Location

10412

128

1

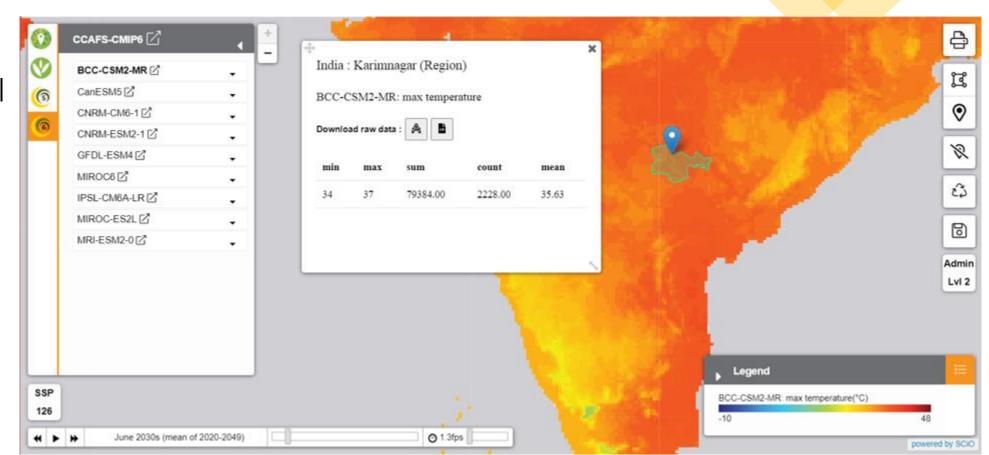
Terrain

Satellite

## "I want to be able to explore data visually!"

CLICK 1: Select CMIP6 Model CLICK 2: Pin region CLICK 3: Select time period CLICK 4:

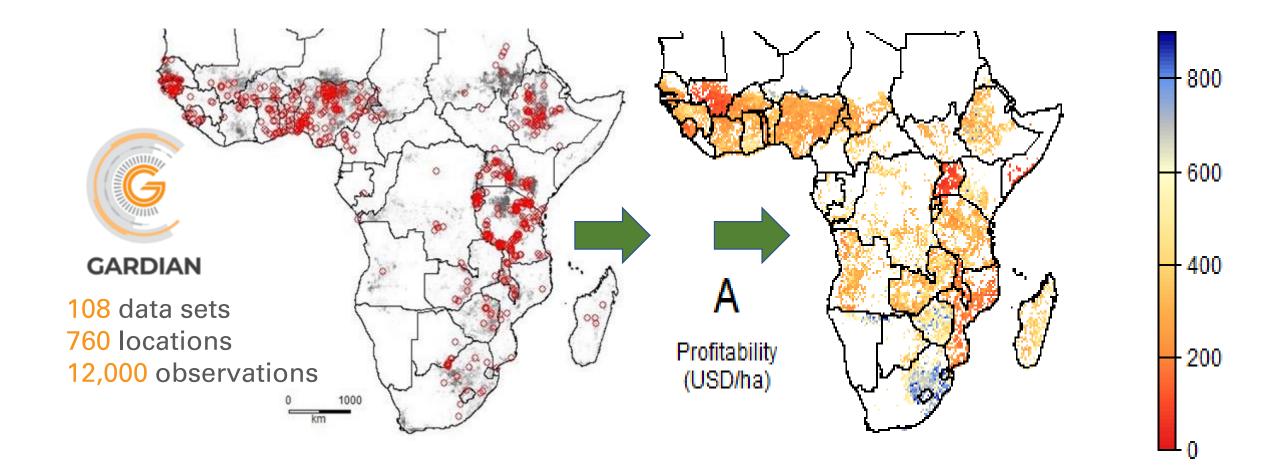
Download data as GeoJSON or CSV





#### gardian.bigdata.cgiar.org/labs.php

### "Where in Sub-Saharan Africa is fertilizer use profitable?"



Bonilla (CIAT), Chamberlin (CIMMYT), Hijmans (UC Davis): https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0227764

### WHY DO I NEED TO DO IT?

- Open, FAIR data has transformed other sectors (biomedical, environmental research)
- Enables realization of the value of data (exemplars; data products that can be leveraged)
- Annual evaluation, KPIs, project indicators (should!) demand it
- Rewards?? (salary increments; publicity)
- Project management should not close project financially w/out FAIR data
- Institutional, individual reputation, as funder OA/OD policies evolve (Funders – non-compliance??)

How We Work — BILL & MELINDA GATES FOUNDATION OPEN ACCESS POLICY

3. **Publications and Underlying Data Will Be Accessible and Open Immediately**. All Funded Research including articles accepted for publication shall be available immediately at publication, without any embargo period. Each accepted article must be accompanied by a Data Availability Statement that describes where any primary data, associated metadata, original software, and any additional relevant materials necessary to understand, assess, and replicate the reported study fundings in totality can be found.

The Foundation shall require that underlying data supporting the accepted article shall be immediately accessible and open upon article publication. Grantees are encouraged to adhere to the FAIR principles to improve the findability, accessibility, interoperability, and reuse of digital assets.

# THANK YOU!

### INFORMATION AND DATA MANAGEMENT

WELCOME ENGAGE NEWS EVENTS RESOURCES

Join our Community of Practice if you would like to know more.

MEDHA DEVARE m.devare@cgiar.org

#### WELCOME!



bigdata.cgiar.org/communities-of-practice/info-data-mgmt