

# La gestió de les dades de recerca a l'ISGlobal

Sergi Sanz ([sergi.sanz@isglobal.org](mailto:sergi.sanz@isglobal.org)). Institut de Salut Global Barcelona

Fira FAIR Data. 15 Juny 2023

**ISGlobal**  
Institut de  
Salut Global  
Barcelona



Una iniciativa de:



## Quins formats i tipus de dades tenim a ISGlobal?

- Mesures instrumentals obtingudes al laboratori (Elisa, PCRs, RDTs, ...)
- Observacions experimentals (Assaig Clínic, Estudi Cohorts, ...)
- Imatges, vídeos i àudios (Estudis qualitatiu)
- Documents de text, fulls de càlcul, bases de dades
- Resultats d'enquestes i transcripcions d'entrevistes
- Dades de simulació, models i programació
- Mostres



INFECTIOUS DISEASES, CHAGAS, PARASITIC AND IMPORTED DISEASES, CHAGAS  
Chagas, Parasitic and Imported Diseases



NON-COMMUNICABLE DISEASES & ENVIRONMENT, CLIMATE AND HEALTH  
Climate and Health



NON-COMMUNICABLE DISEASES & ENVIRONMENT, AIR POLLUTION & URBAN ENVIRONMENT  
Air Pollution and Urban Environment



NON-COMMUNICABLE DISEASES & ENVIRONMENT, NCDs AND ENVIRONMENT  
Non-communicable Diseases and Environment



NON-COMMUNICABLE DISEASES & ENVIRONMENT, CHILDHOOD & ENVIRONMENT  
Childhood and Environment



INFECTIOUS DISEASES, VIRAL AND BACTERIAL INFECTIONS  
Viral and Bacterial Infections



INFECTIOUS DISEASES, MALARIA  
Malaria

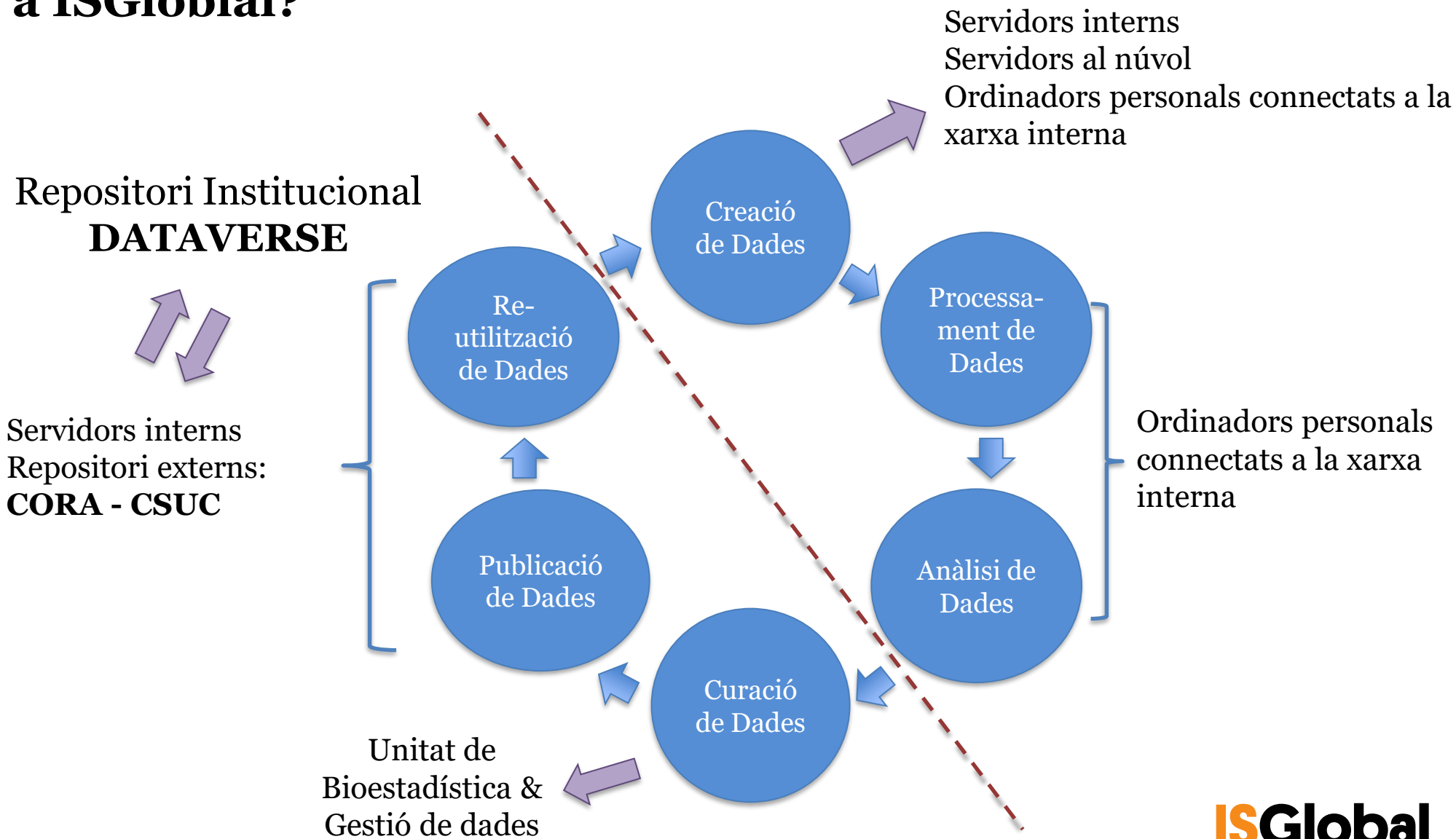


NON-COMMUNICABLE DISEASES & ENVIRONMENT, RADIATION  
Radiation



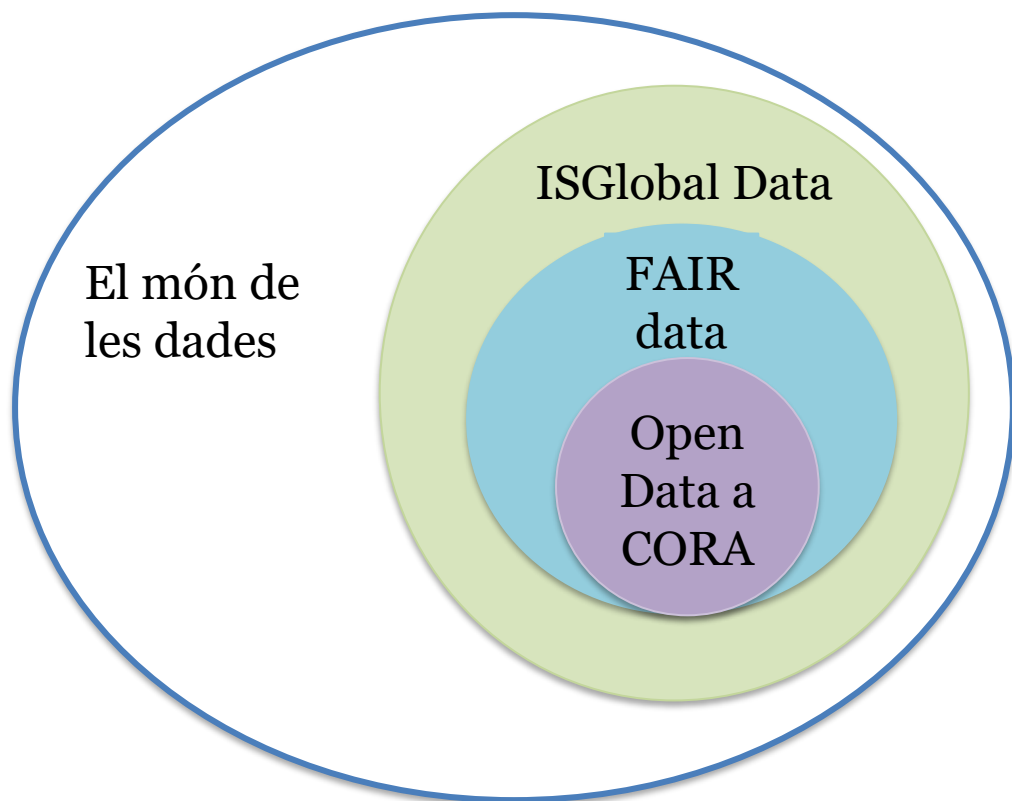
INFECTIOUS DISEASES, MATERNAL, CHILD AND REPRODUCTIVE HEALTH, MATERNAL, CHILD AND REPRODUCTIVE HEALTH  
Maternal, Child and Reproductive Health

# Com organitzem les dades a ISGlobal?



## Per què hem creat un Repositori institucional amb l'eina DATAVERSE?

1. Ens permet ordenar les dades de forma clara i eficient.
2. Podem curar les dades i convertir-les en format FAIR per a ús intern
3. Ens Permet gestionar el nivell de obertura de los dades amb CORA.



**as open as possible,  
as closed as necessary**

dcc.ac.uk

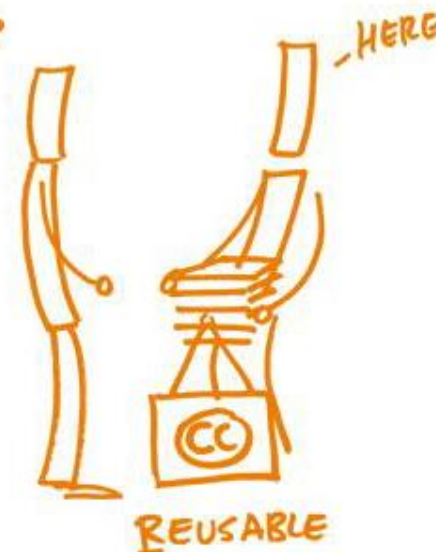
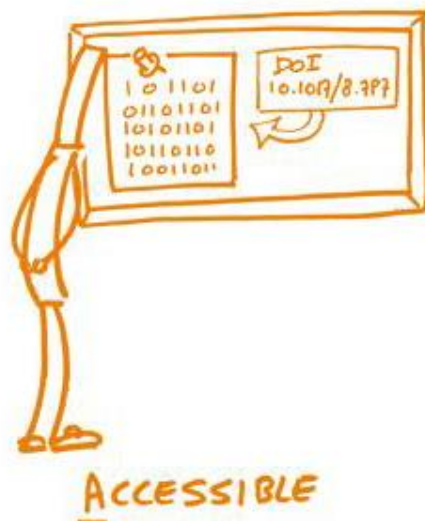


Imagen: 'Balancing rocks'  
by Viewminder CC-BY-SA-ND



# Com fem la gestió de les “Dades de recerca” per transformar-les FAIR?

## FAIR DATA PRINCIPLES



- F**
- Metadades
  - Identificadors persistents (DOI)
  - Vocabulari estàndard (MESH)

- A**
- Repositori CORA

- I**
- Formats estàndards / Protocols oberts (CSV, TXT, ODT,... )

- R**
- Llicències (CC BY, CC BY-NC, CC BY-NC-SA)
  - Diccionari descriptiu de les dades (format, rang, definicions,...)
  - Pla de gestió de dades
  - Readme.txt

## Anàlisi de les Fortaleses, Debilitats, Oportunitats i Amenaces

### Fortaleses

- Recolzament institucional (Acció estratègica en Gestió de dades)
- Formar part del projecte CORA (CERCA / CSUC)
- Acció recollida en el pla estratègic institucional (2018-2022 i 2023-2027)
- Petició del comitè avaluador extern

### Oportunitats

- Augment dels requisits en la gestió de dades per part de finançadors i revistes
- No estar sols i aprendre d'experiències existents: CORA, Open Aire, Harvard
- Optar a la captació de recursos en projectes de investigació com programa Horizon

### Debilitats

- Coneixement insuficient de la realitat en ISGlobal
- Insuficient recursos humans per curar totes les dades institucionals, noves i **ANTIGUES**
- Establir una cultura institucional en tots els grups de recerca

### Amenaces

- Cultura científica “tradicional”
- Por a la pèrdua del control de les dades per part dels investigadors
- No fer un sistema atractiu i pràctic per als investigadors
- Existència dels recursos necessaris pel manteniment del sistema.

	A	B	C
1	<b>Param</b>	<b>Value</b>	
2	<b>title</b>	Demographic survey	
3	<b>dsDescription</b>	Data corresponding to a demographic survey conducted in Mopeia district (Zambezia province, Mozambique) tr	
4	<b>mainResearchProgramme</b>	Malaria	
5	<b>secondaryResearchProgramme</b>		
6	<b>project</b>	BOHEMIA	
7	<b>projectDescription</b>	The main goal of BOHEMIA is to contribute to malaria transmission i	
8	<b>subProject</b>	WP1	
9	<b>subProjectDescription</b>	Work package 1 description	
10	<b>kindOfData</b>	Census/enumeration data	
11	<b>language</b>	English	
12	<b>startTimePeriodCovered</b>		
13	<b>endTimePeriodCovered</b>		
14	<b>linkToDataverse</b>		
15	<b>license</b>	CC BY-NC-SA 4.0	

	A	B	C	D
1	<b>authorName</b>	<b>authorAffiliation</b>	<b>authorIdentifie</b>	<b>authorIdentifier</b>
2	Paula Ruiz-Castillo	Barcelona Institute for Global Health	ORCID	0000-0002-5654-503X
3	Saimado Imputiua	Manhiça Health Research Centre		
4	Kexin Xie	Virginia Tech		
5	Eldo Elobolobo	Manhiça Health Research Centre		
6	Patricia Nicolas	Barcelona Institute for Global Health		
7	Julia Montana	Barcelona Institute for Global Health		
8	Edgar Jamisse	Manhiça Health Research Centre		
9	Humberto Munguambe	Manhiça Health Research Centre		
10	Felisbela Materrula	Manhiça Health Research Centre		
11	Aina Casellas	Barcelona Institute for Global Health	ORCID	0000-0003-4939-8411
12	Xinwei Deng	Virginia Tech		
13	Achla Marathe	Virginia Tech		
14	Regina Rabinovich	Barcelona Institute for Global Health		
15	Francisco Saute	Manhiça Health Research Centre		
16	Carlos Chaccour	Barcelona Institute for Global Health	ORCID	0000-0001-9812-050X
17	Charfudin Sacoor	Manhiça Health Research Centre		

	A	B	C
1	<b>country</b>	<b>state</b>	<b>city</b>
2	Mozambique	Zambezia	Mopeia
3			
4			
5			
6			
7			
8			
9			
10			
11			

	A	B	C
1	<b>datasetContactName</b>	<b>datasetContactAffiliation</b>	<b>datasetContactEmail</b>
2	UBIOESGD	Barcelona Institute for Global Health	ubioesdm@isglobal.org
3			

- ORCID
- ISNI
- LCNA
- VIAF
- GND
- DAI
- ResearcherID
- ScopusID

A	B	A	B	C	D
<b>keywordVocabulary</b>	<b>keywordValue</b>	<b>publicationCitation</b>	<b>publicationIDType</b>	<b>publicationIDNumber</b>	<b>publicationURL</b>
Other ▾	Mozambique				
Other ▾	cluster randomized clinical trial				
Other ▾	demographic survey				
Other ▾	mapping		ark		
Other ▾	indoor residual spraying		arXiv		
Other ▾	population structure		bibcode		
Other ▾	malaria		doi		
MeSH			ean13		
Other			eissn		
			handle		
			isbn		
<b>startTimePeriodCovered</b>					
<b>endTimePeriodCovered</b>					
<b>linkToDataverse</b>					
<b>license</b>	CC BY-NC-SA				

A	B	C	D	E	F
<b>Filename</b>	<b>Folder</b>	<b>Description</b>	<b>Restricted</b>	<b>sensitiveData</b>	<b>anonymizedData</b>
minicensus_household_final.csv	Data ▾	Household level characteristics	yes ▾	no ▾	yes ▾
minicensus_people.csv	Data ▾	Individual characteristics	yes ▾	no ▾	yes ▾
minicensus_repeat_death.csv	Data ▾	Deaths occurred during the previous 12 months	yes ▾	no ▾	yes ▾
minicensus_repeat_hh_sub.csv	Data ▾	Household head substitute characteristics	yes ▾	no ▾	yes ▾
minicensus_repeat_net.csv	Data ▾	Mosquito net details	yes ▾	no ▾	no ▾
Data_dictionary.xlsx	Data ▾	Data dictionary for all tables provided	no ▾	no ▾	no ▾
	Protocols				
	Plans				
	Technical documents				
	Scripts				
	<b>Data</b>				
	Publications				

+    ≡    **General** ▾    Geospatial ▾    Contacts ▾    Authors ▾    Keywords ▾    Publications ▾    Files ▾



# Gràcies

Sergi Sanz

Responsable de la Unitat de Bioestadística i Gestió de dades  
Responsable de la acció estratègica de Gestió de dades a ISGlobal

## Equip del projecte

Susana Méndez, Adrián Carrascosa

**ISGlobal**

---

[www.isglobal.org](http://www.isglobal.org)

Una iniciativa de:

