


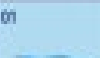






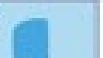














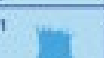

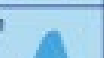
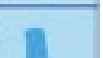














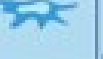

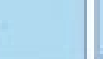




Human Cell Atlas Data Coordination, lessons learnt

A case study of data submission, standardisation and
sharing
eResearch NZ 2022

Claire Rye, Former Product Owner HCA-DCP
Ingestion Service, EMBL-EBI

Current Product Manager - Data Services, NeSI

The Human Cell Atlas (HCA): A Periodic Table for our Cells

TISSUE										Development												
01 Br Brain										01 	01 											
01 G Gut	01 K Kidney																					
CELLS																						
Brain			Eye		Gut		Lung		Skin				Liver		Kidney		Heart		Muscle		Development	
01 Lu Lung	01 Dev Development	01 	01 	01 	01 	01 	01 	01 	01 	01 	01 	01 	01 	01 	01 	01 	01 	01 	01 	01 	01 	
01 S Skin	01 H Heart	N	N	N	N	N	N	N	N	E	N	N	N	N	N	N	N	N	N	N	N	
01 L Liver	01 Rep Reprod.	01 	01 	01 	01 	01 	01 	01 	01 	01 	01 	01 	01 	01 	01 	01 	01 	01 	01 	01 	01 	
01 B Bone										01 	01 	01 	01 	01 	01 	01 	01 	01 	01 	01 	01 	

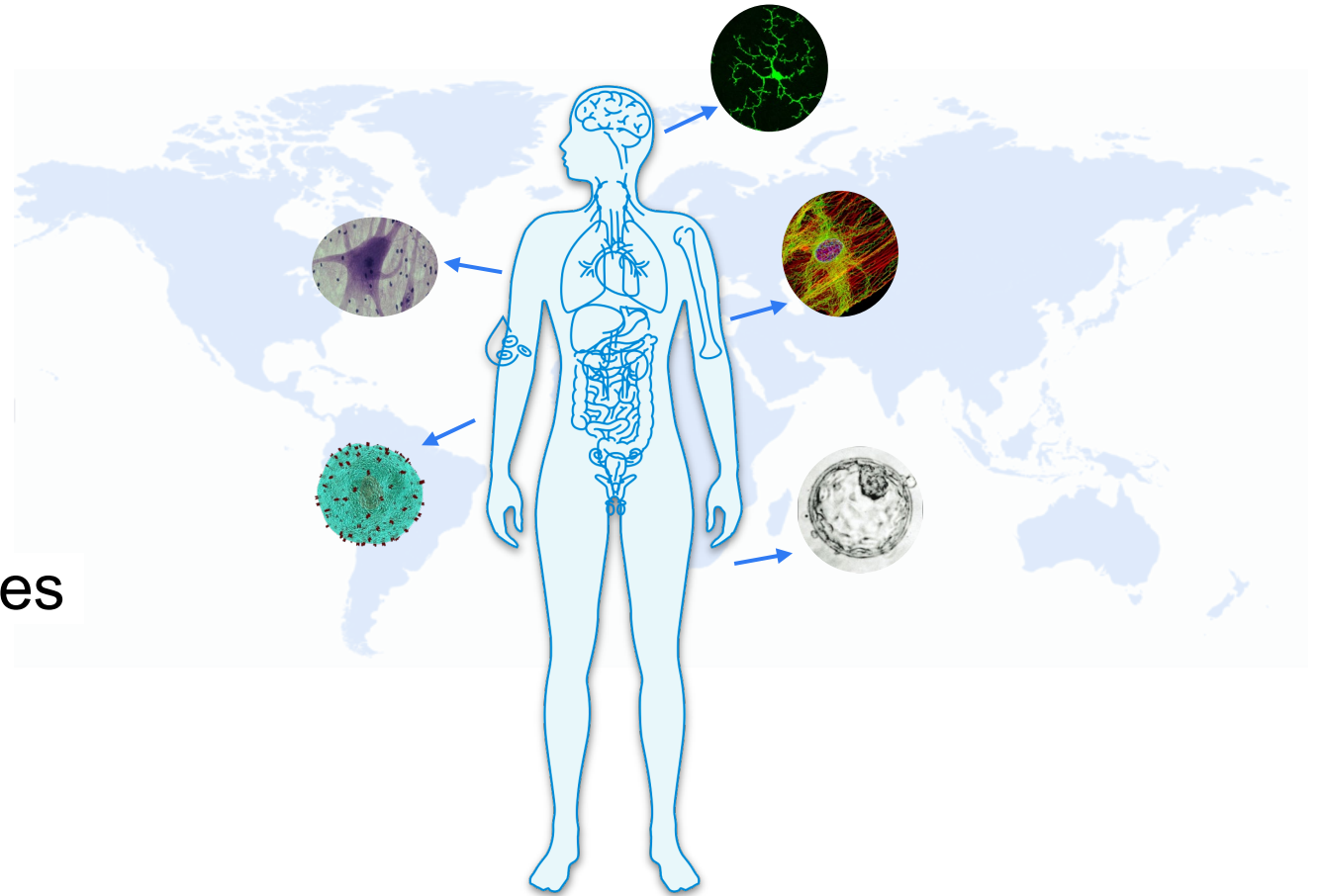
METHODS

CODEX	GeoMX DSP	tCy-CIF	ST	DNA microsc.	GeoMX DSP	MERFISH	scRNA Seq	ATAC Seq
MIBI	Multiplex IF		Slide Seq	HDST	STAR MAP	Seq FISH	snRNA Seq	DroNc Seq

The HCA community is profiling millions of diverse human cells

Global effort requiring:

- Hundreds of labs
- Organ-specific data
- Disparate experimental techniques and data types



Our community has diverse skills and data needs



Experimental Researchers

I want to:

- Find and visualize HCA data
- Contribute data



Computational Researchers

I want to:

- Find data to analyze and test a hypothesis
- Compare datasets

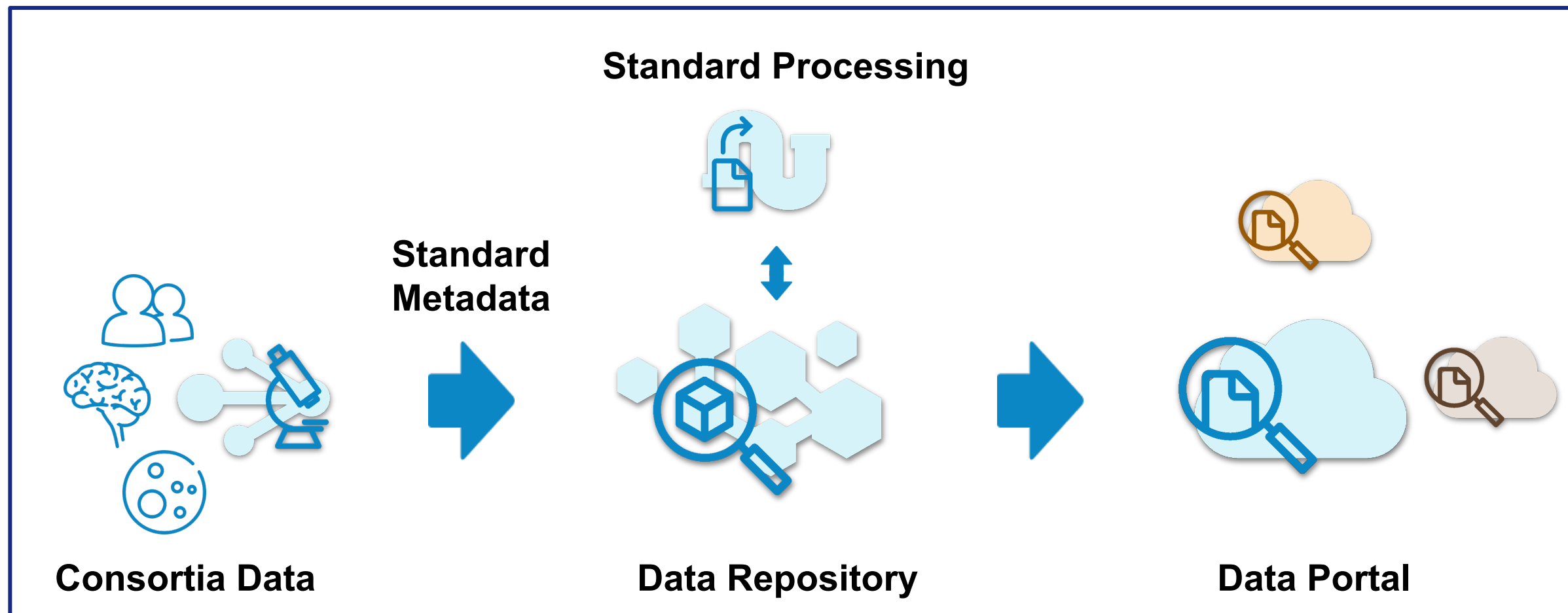


Engineers and Tool Developers

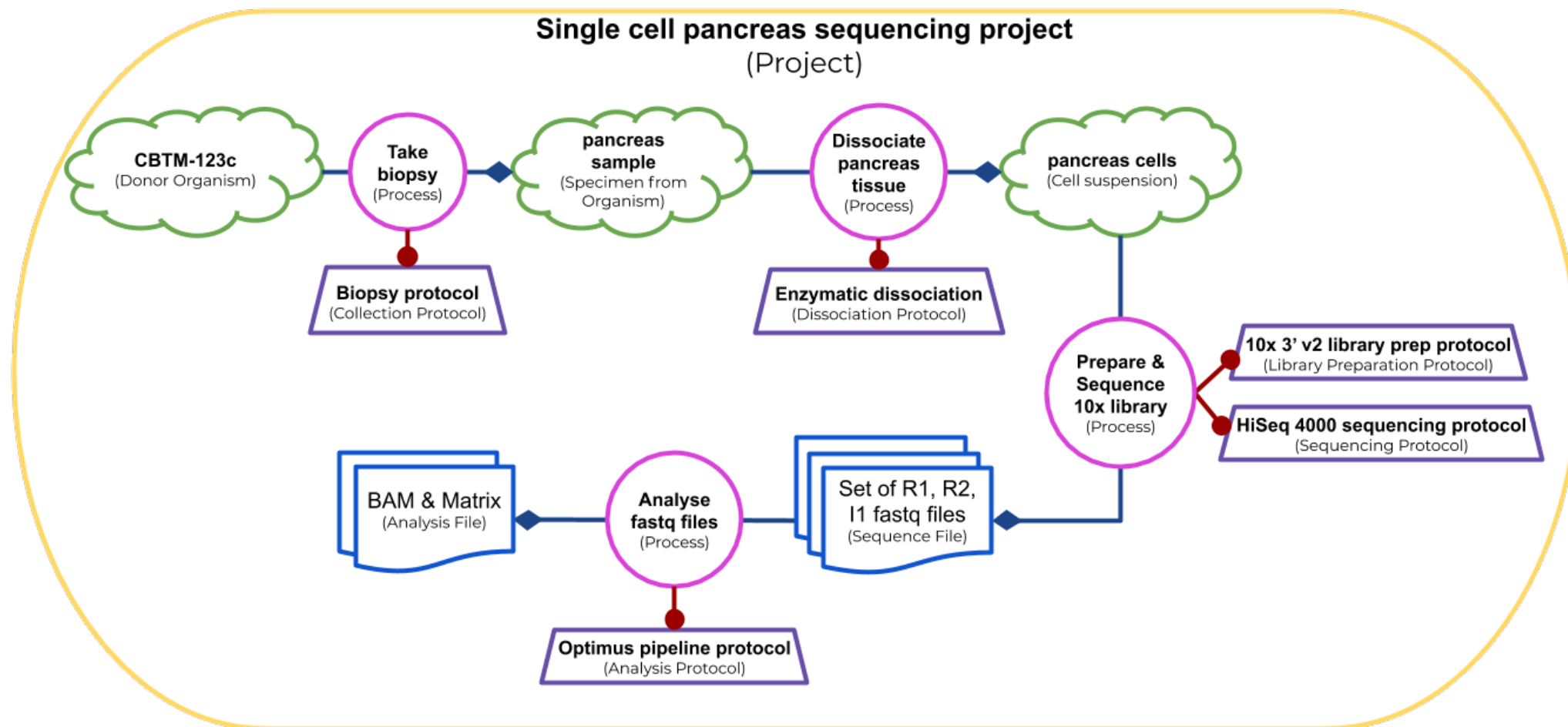
I want to:

- Bulk download HCA data to test tools
- Contribute tools

Data Coordination Platform (DCP)



Modelling the metadata

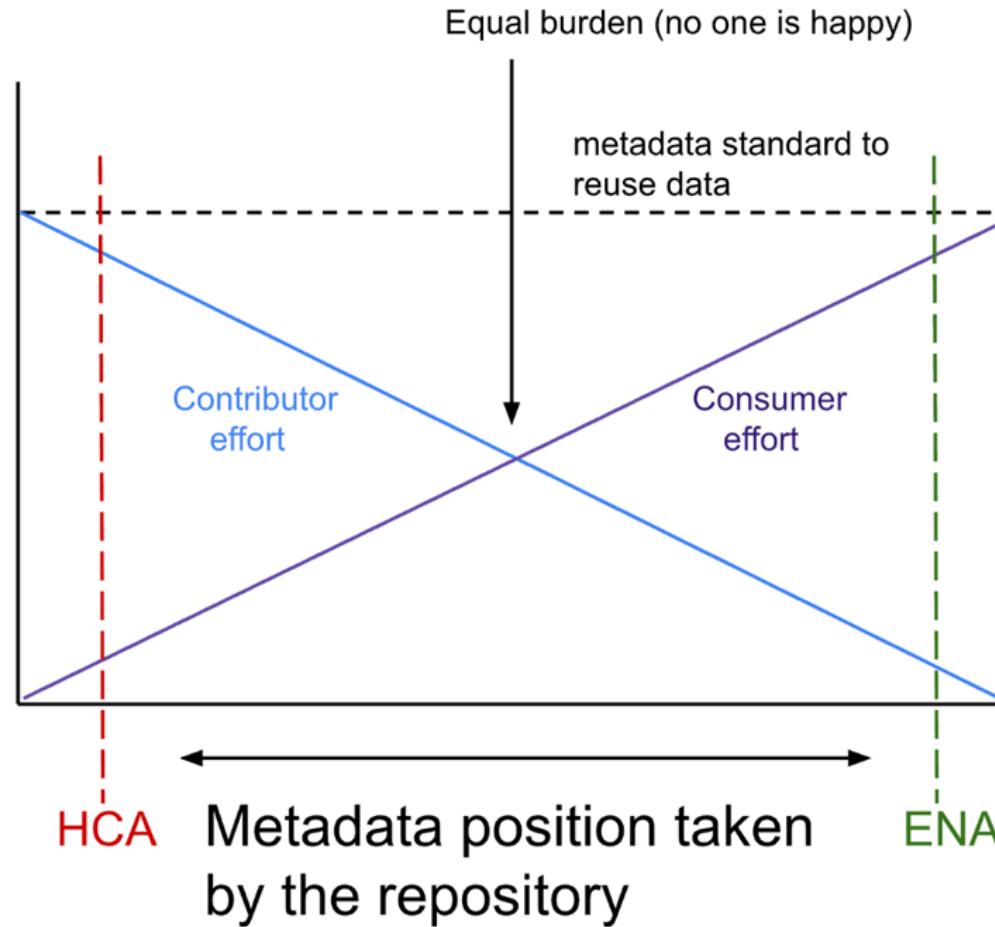


Choose your trade-off carefully



Data Contributor

User goal:
Submit data in order to
publish results.



Data Consumer

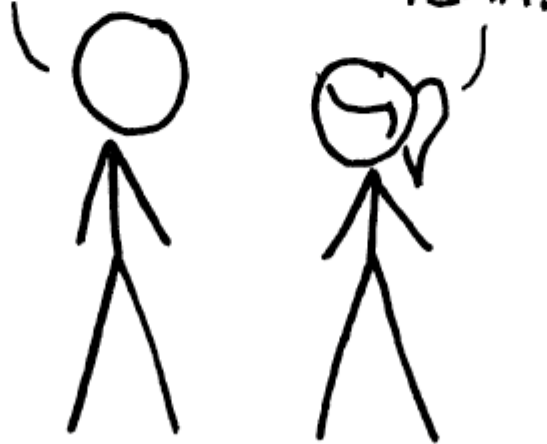
User goal:
Consume data to
compare datasets.

HOW STANDARDS PROLIFERATE:

(SEE: A/C CHARGERS, CHARACTER ENCODINGS, INSTANT MESSAGING, ETC.)

SITUATION:
THERE ARE
14 COMPETING
STANDARDS.

14?! RIDICULOUS!
WE NEED TO DEVELOP
ONE UNIVERSAL STANDARD
THAT COVERS EVERYONE'S
USE CASES.



SOON:

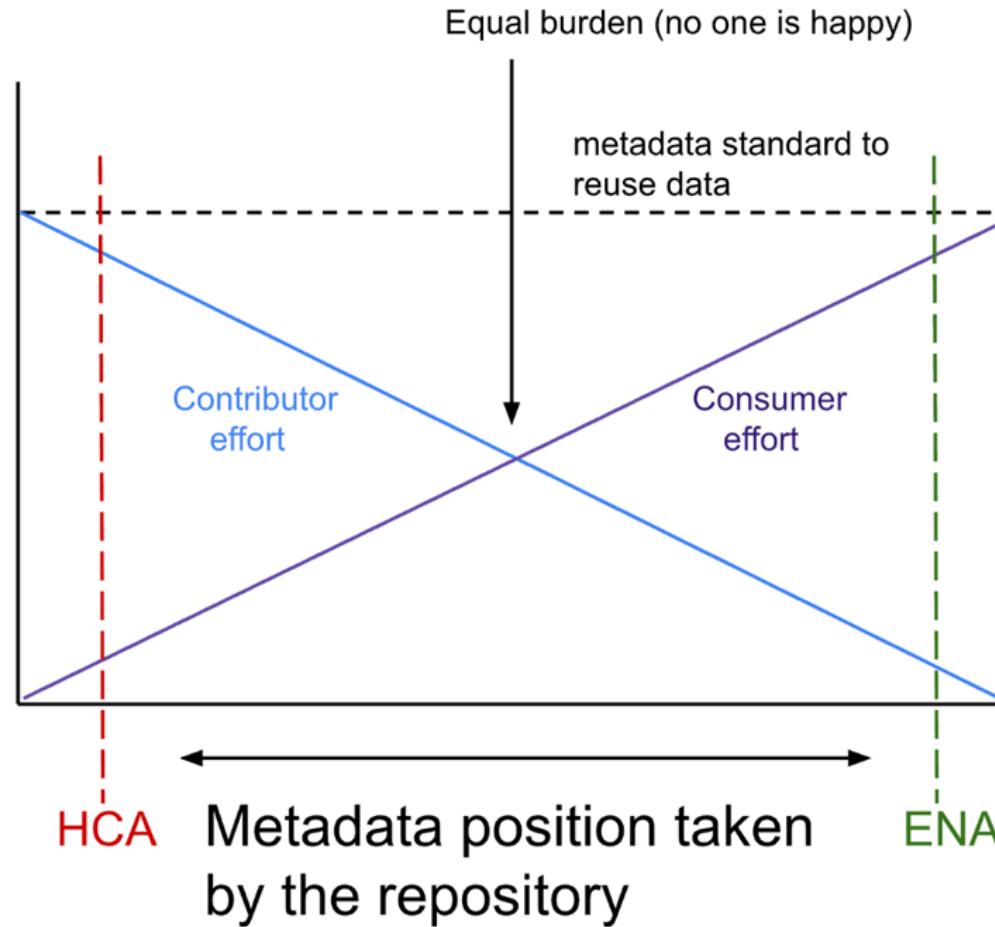
SITUATION:
THERE ARE
15 COMPETING
STANDARDS.

Choose your trade-off carefully



Data Contributor

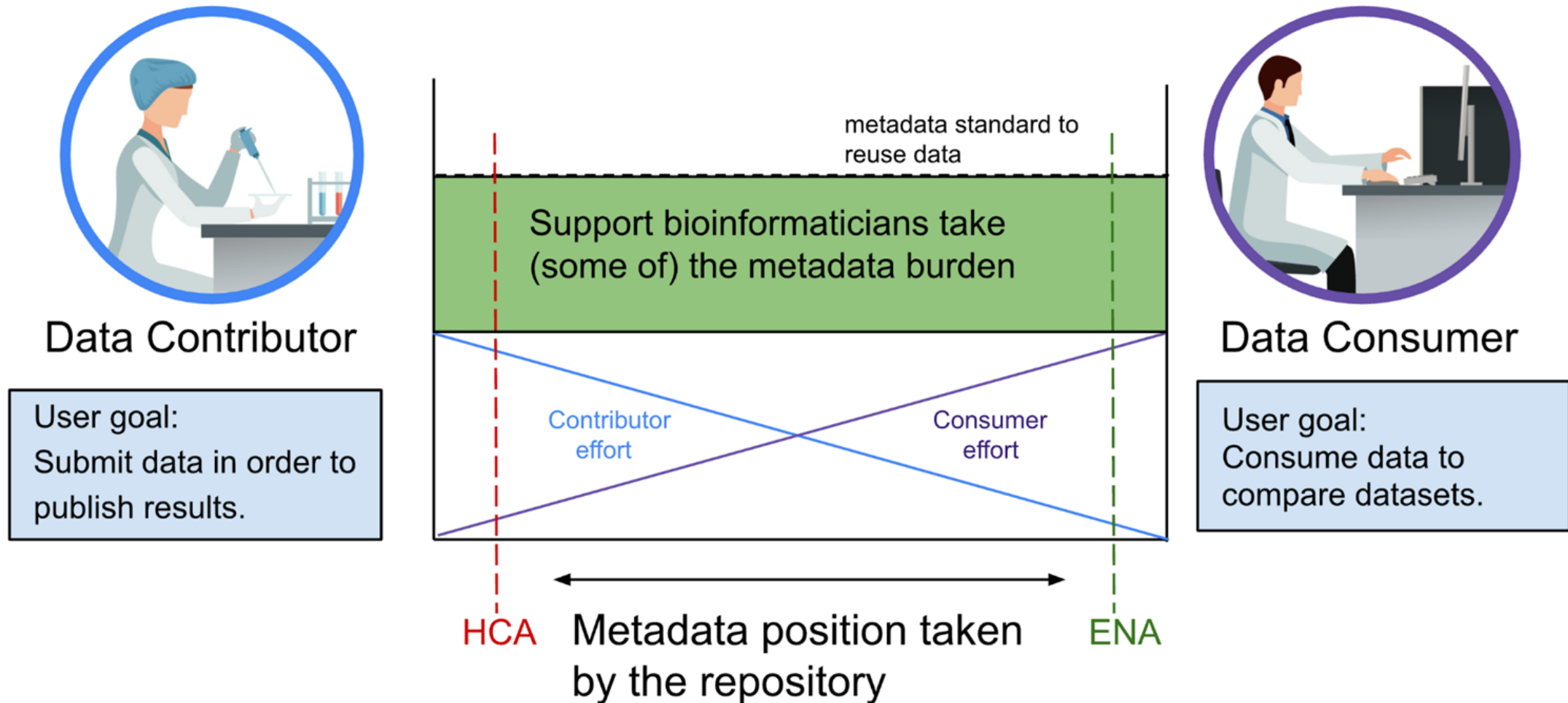
User goal:
Submit data in order to
publish results.



Data Consumer

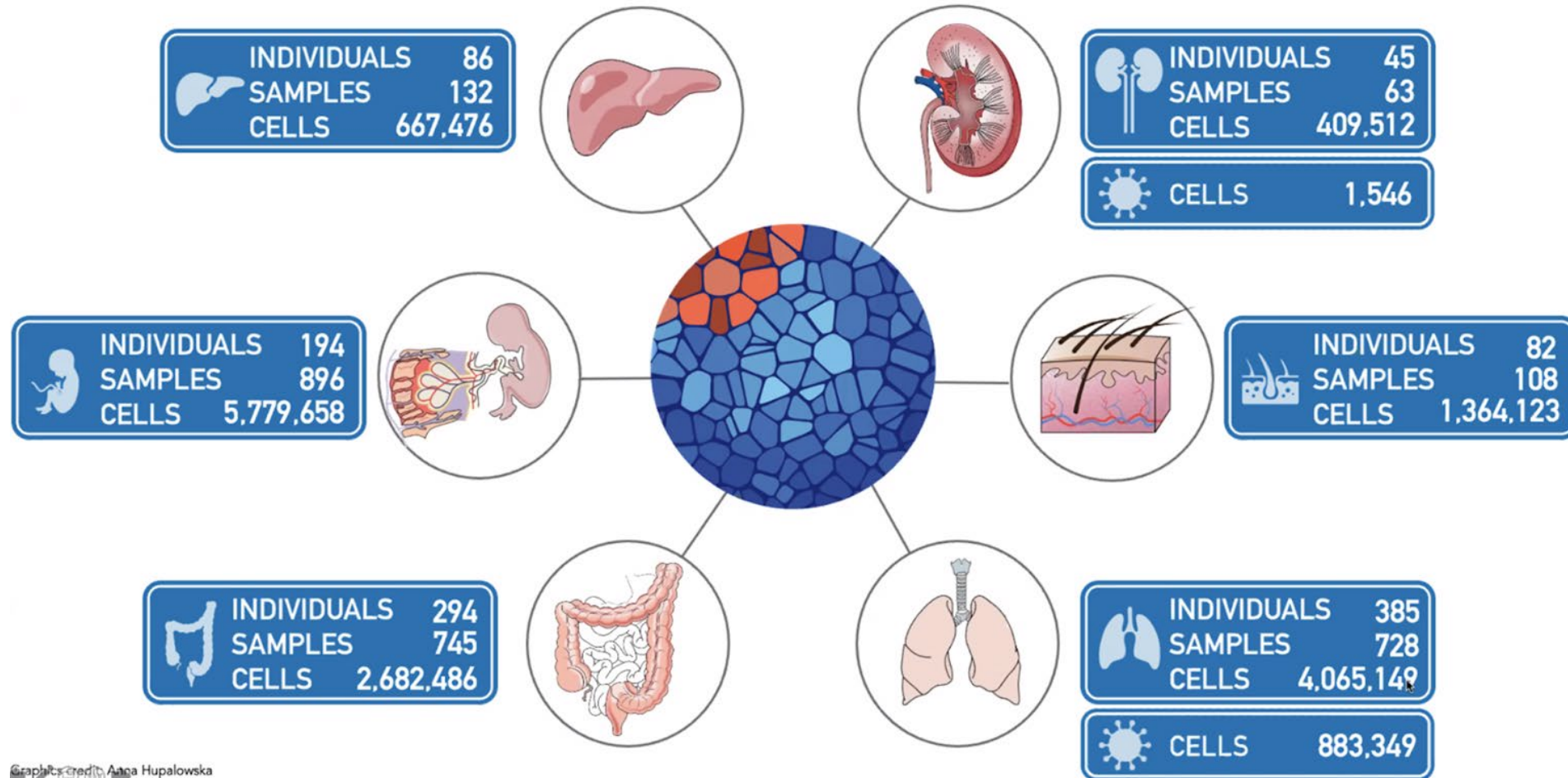
User goal:
Consume data to
compare datasets.

Can you make this trade-off easier to live with?



What does this allow us to do?

Views from the Atlases



Graphics credit: Anna Hupalowska

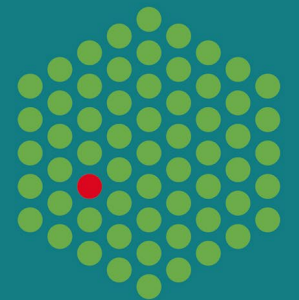
Funding and Acknowledgements

**Chan
Zuckerberg
Initiative** 

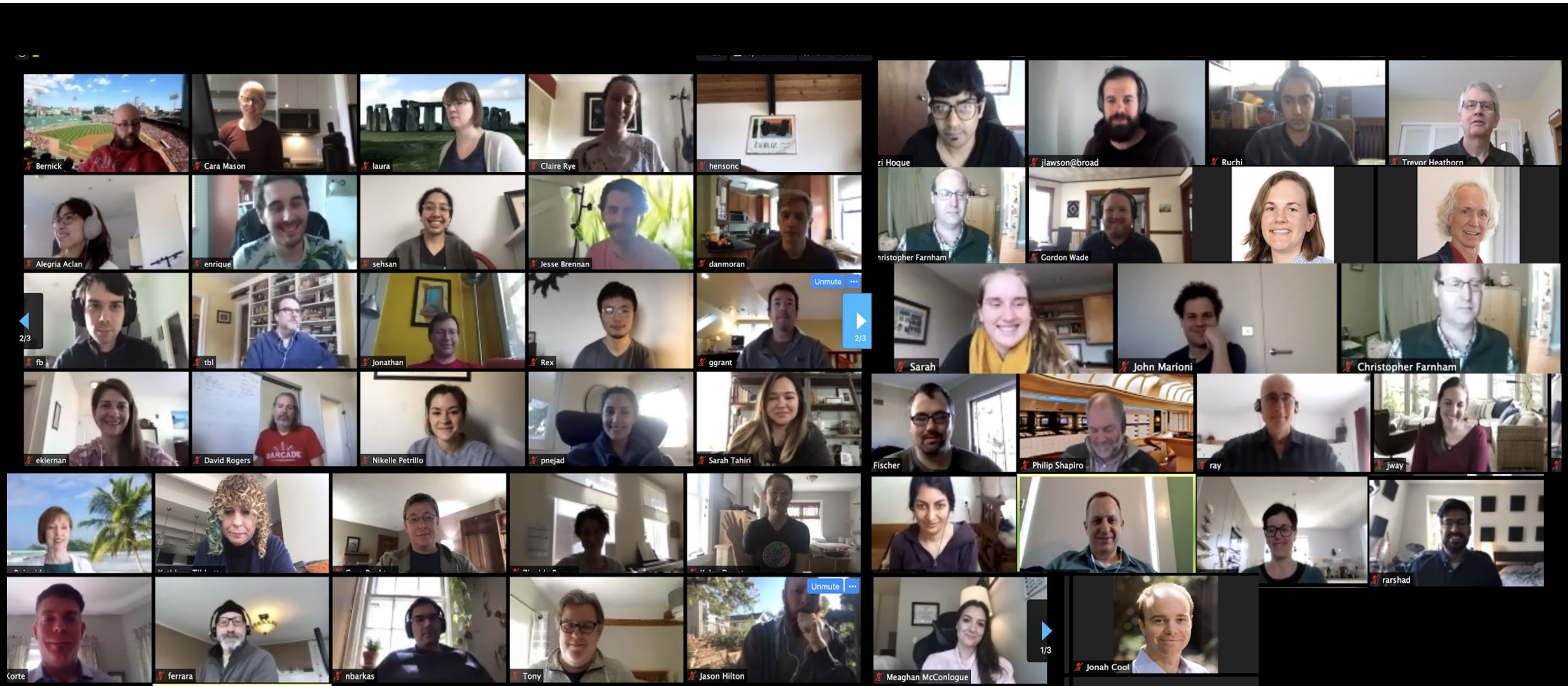


UNIVERSITY OF CALIFORNIA
SANTA CRUZ Genomics
Institute

EMBL-EBI



Acknowledgements





Data Services

- National Data Transfer Service
- Nearline storage
- Aotearoa Genomics Data Repository
- Rakeiora Pathfinder Project
- Research data management/data lifecycle

Claire Rye

Product Manager - Data services

claire.rye@nesi.org.nz