

# Where is Open Science in Cancer Research?

Carly Strasser, PhD | @carlystrasser



**FRED HUTCH**  
CURES START HERE™

CMU Open Science Symposium

November 2019

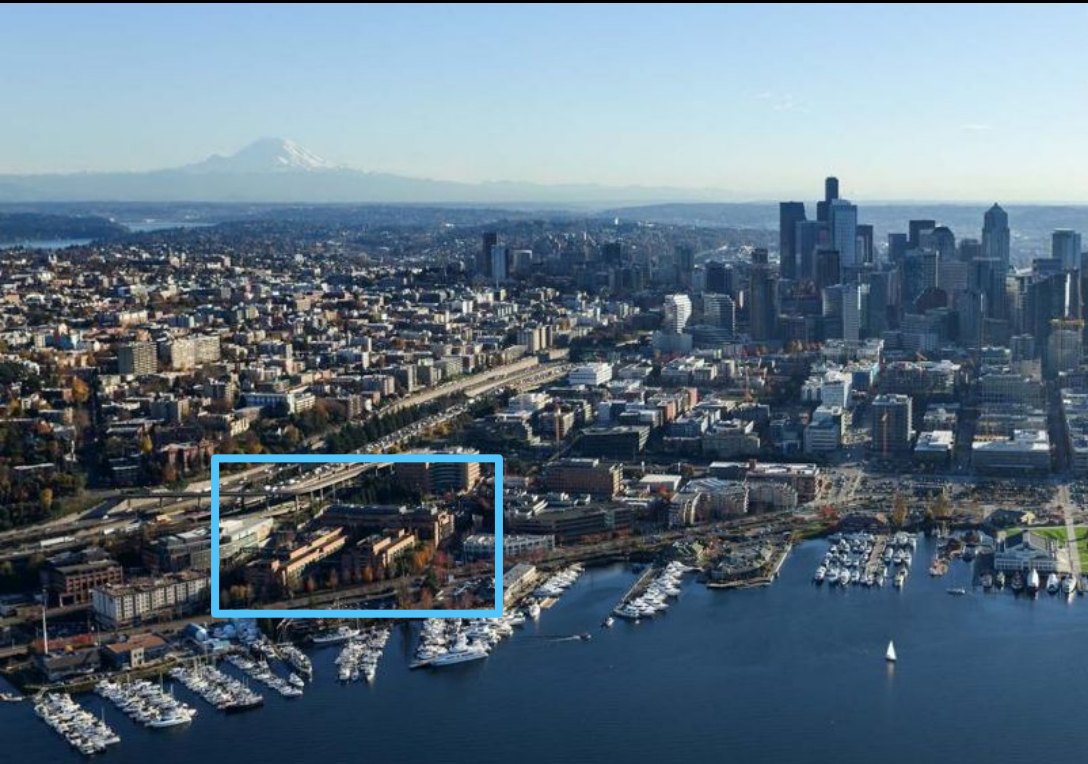




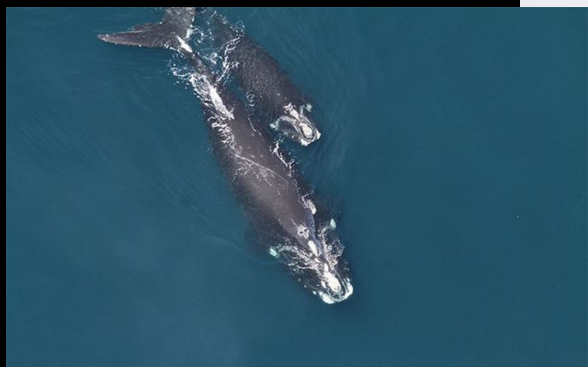
**FRED HUTCH**

CURES START HERE™

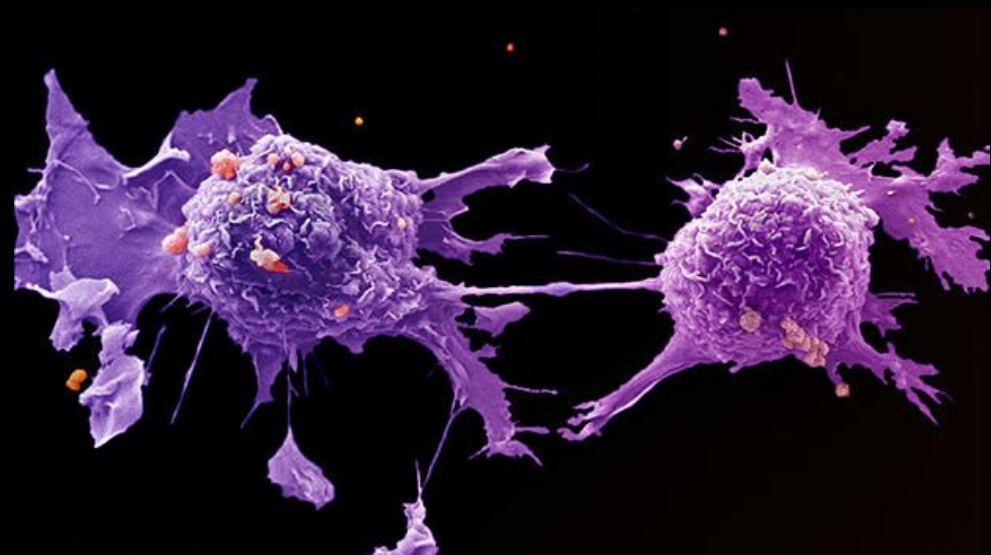
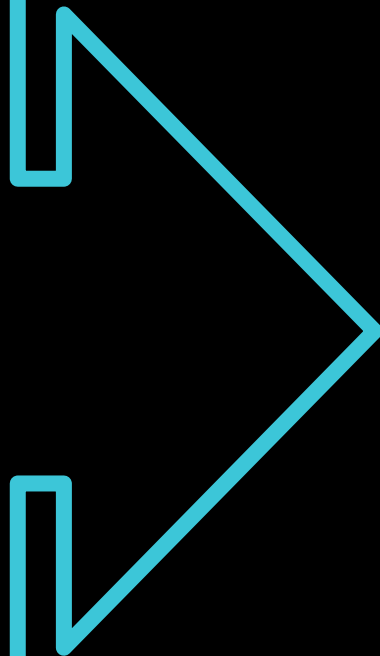
[fredhutch.org](https://fredhutch.org)



- Independent research institution
- Part of Seattle Cancer Care Consortium with University of Washington, Seattle Children's, Seattle Cancer Care Alliance
- 2,000 employees
- 238 faculty members
- Pioneered bone marrow transplants



NOAA



CC-BY LRI EM Unit

DataONE

University of California



CDL

California Digital Library

GORDON AND BETTY  
MOORE  
FOUNDATION

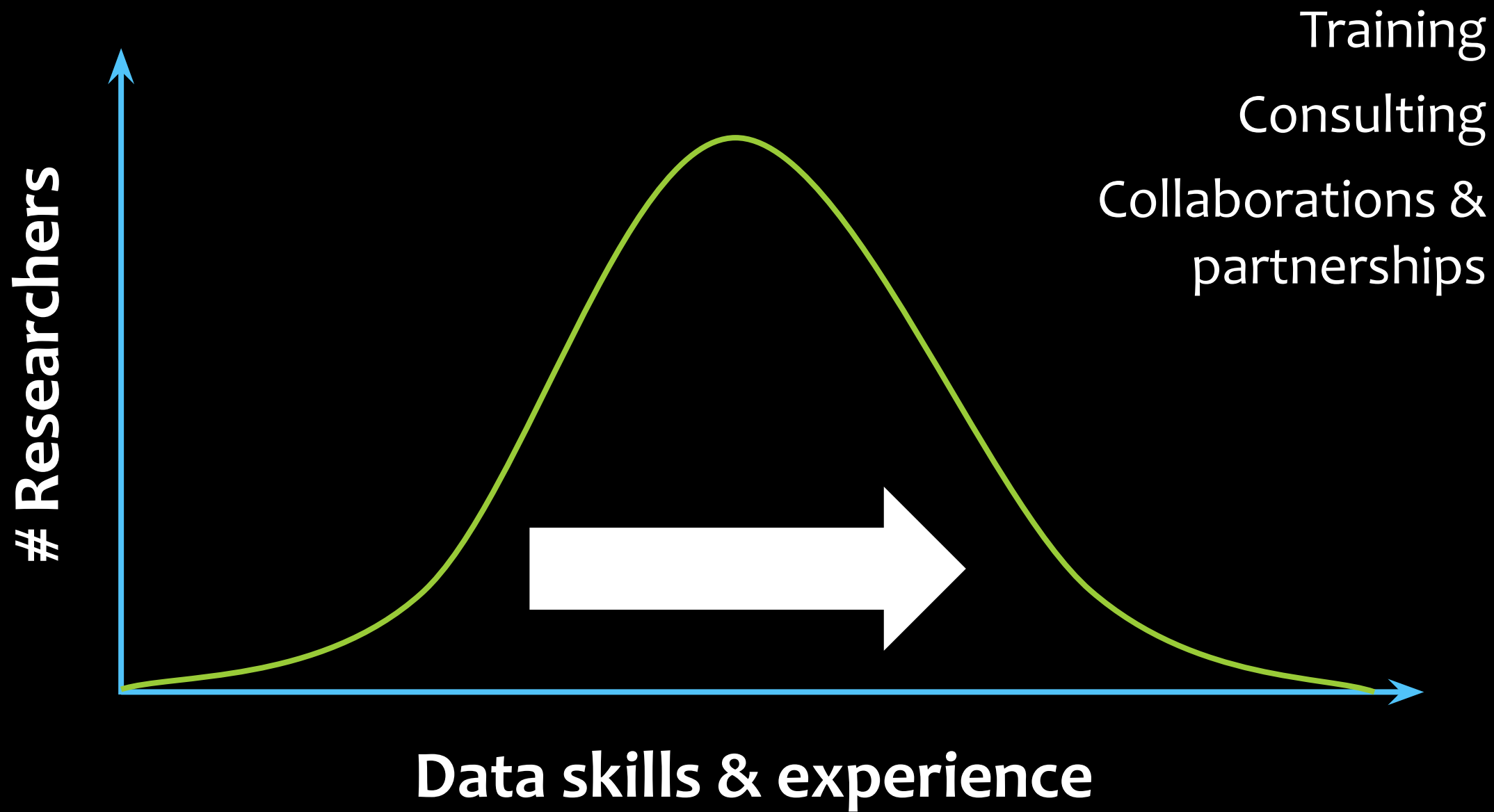
# Hutch Data Commonwealth

HDC helps build data-intensive research capabilities through  
software & data engineering | training | partnering

# Hutch Data Commonwealth

## Translational Data Science Integrated Research Center

### Bioinformatics & Data Science Cooperative (The Coop)

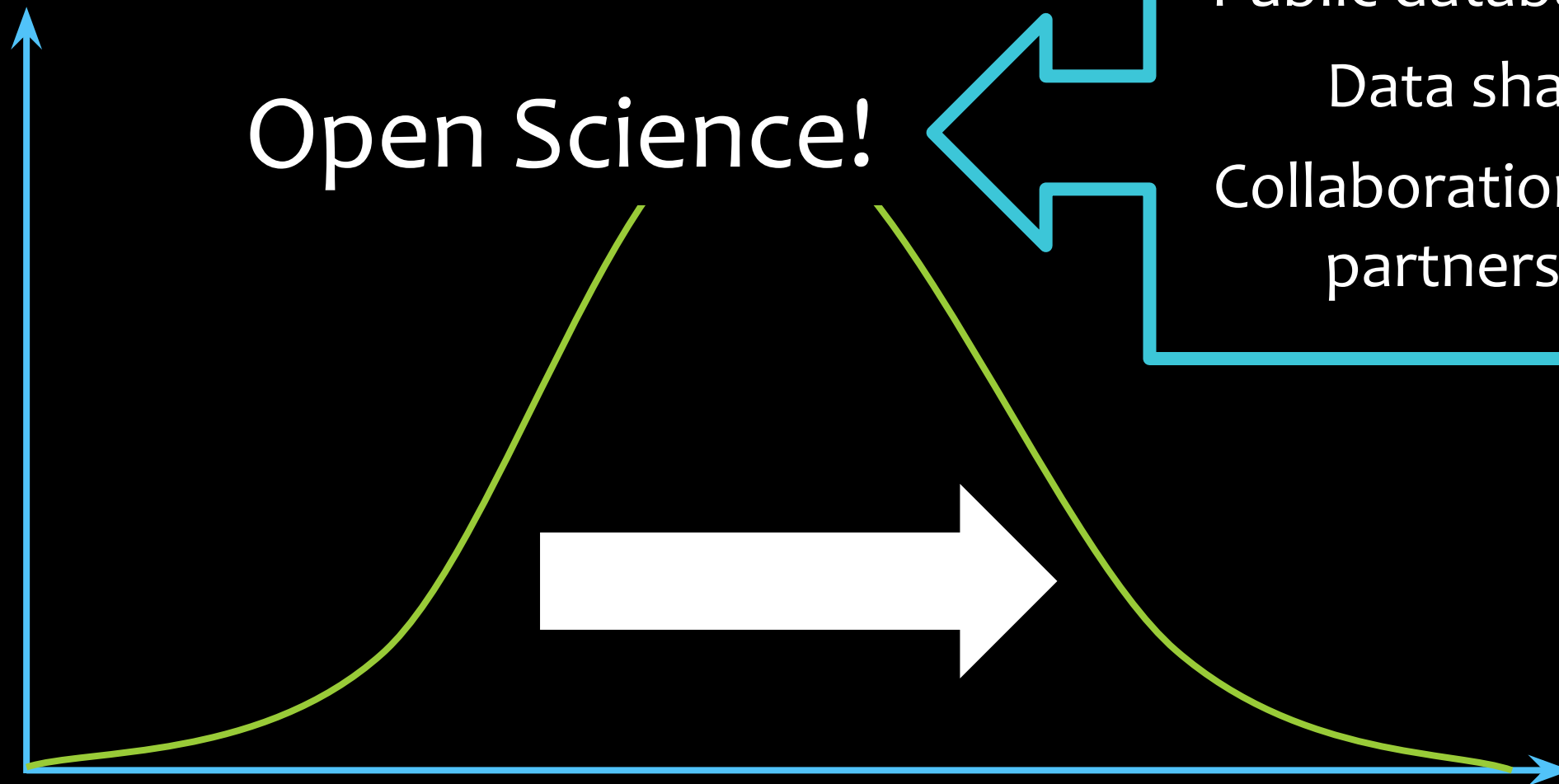


# Projects

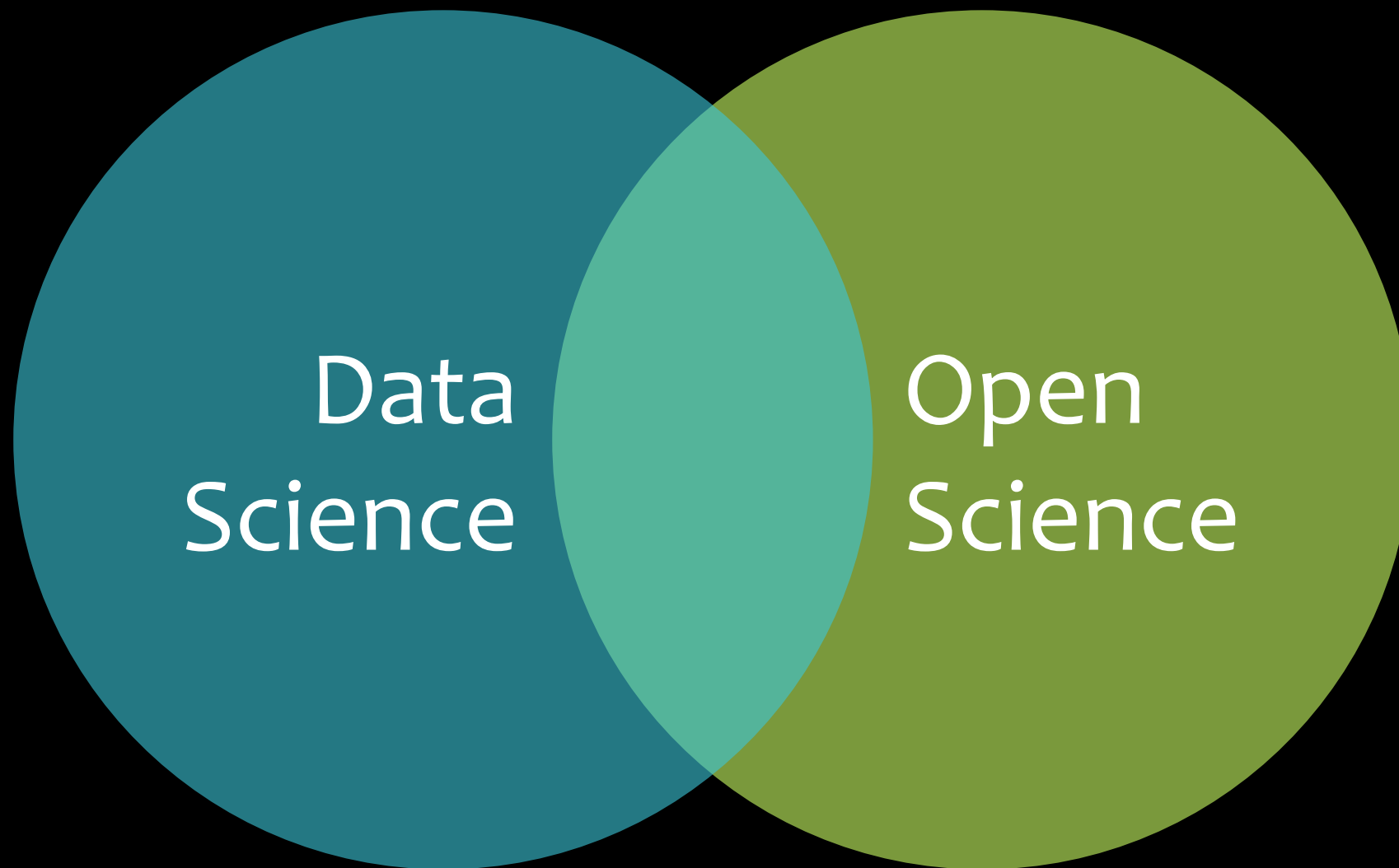
Open Science!

Public databases  
Data sharing  
Collaborations &  
partnerships

Data size & diversity









**FRED HUTCH**  
CURES START HERE™



- ✗ Open access policy
- ✗ Data sharing policy
  - ✗ OA repository
  - ✗ Data repository
- ✗ Other open science tools

# Sharing Publications

## SHARE

IN DEPTH | SCIENTIFIC PUBLISHING

# Open access takes root at National Cancer Institute

Jocelyn Kaiser

+ See all authors and affiliations

*Science* 16 Aug 2019:  
Vol. 365, Issue 6454, pp. 629  
DOI: 10.1126/science.365.6454.629

Article

eLetters


 PDF

Log in to view PDF

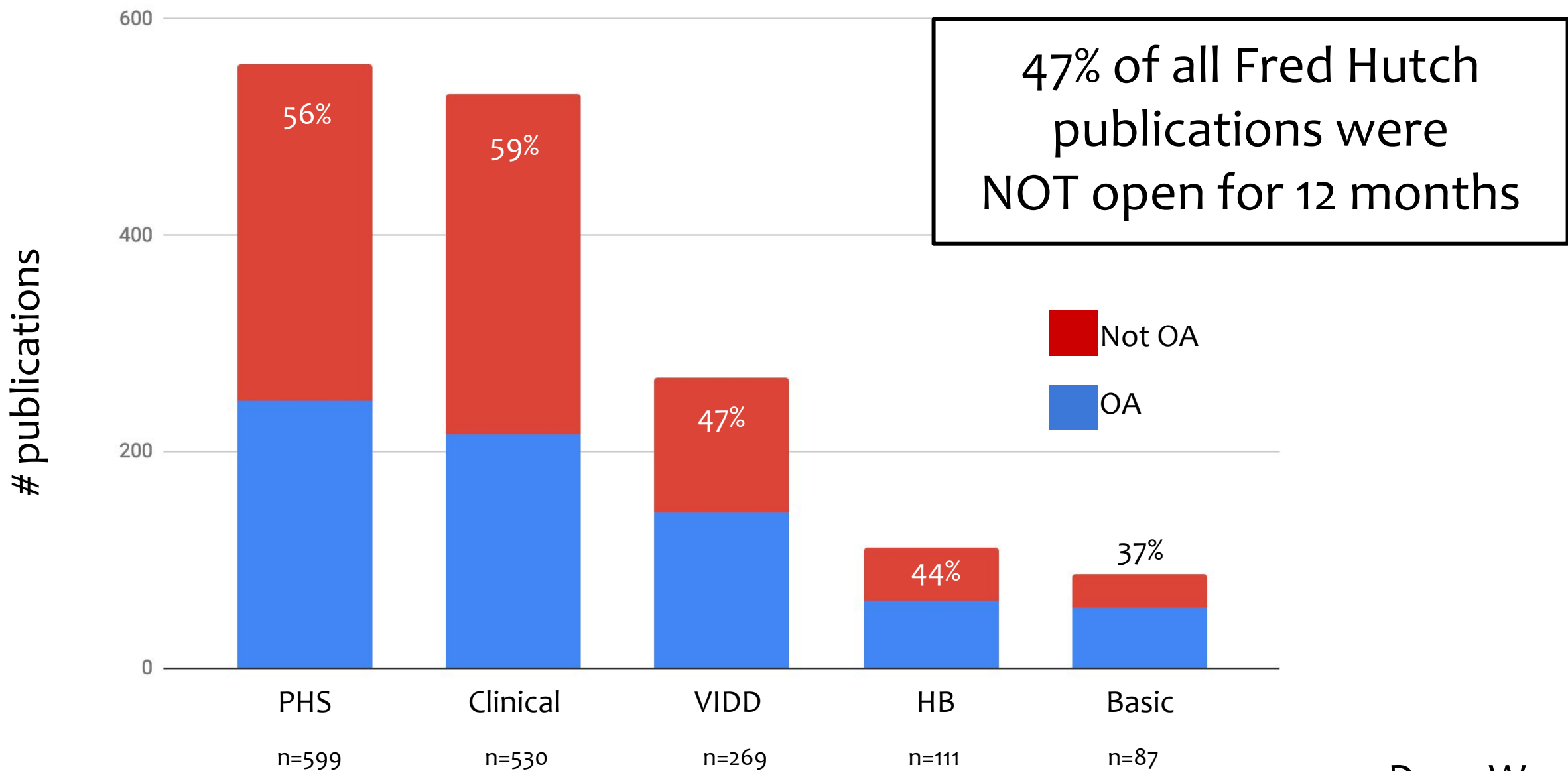
via AAAS login

NIH requirements  
are sufficient

The NIH Public Access Policy mandates that any peer reviewed, published research that stems all or in part from NIH funding must be **archived in PubMed Central** upon acceptance for publication to be made available **no later than 12 months after publication**.



71	NCI-Designated Cancer Centers
50	Comprehensive Cancer Centers
12	Have OA policies (primarily via academic affiliation)
8	Independent Cancer Centers
0	Independent Cancer Centers with OA policies



# Why Fred Hutch President Gary Gilliland is enlisting techies to defeat cancer by 2025

BY **ALAN BOYLE** on September 26, 2017 at 8:26 am

2 Comments

f Share 2.5k

🐦 Tweet

📺 Share

👍 Reddit

✉ Email

GeekWire Summit: Get your tickets here!



1,882 days!



# Sharing Data

## Science News

*from research organizations*

### How sharing cancer data can save lives

Date: May 24, 2017

Source:

Summary:

Share:

RELATED TO

Health & M

> Breast Cancer

> Diseases and Conditions

Global leaders in cancer research have called for the **worldwide sharing of cancer data** to save lives. The Global Alliance for Genomics and Health argue how the 'freeing of data' for a disease that knows no borders will enable researchers to find better treatments that increase survival and improve quality of life for cancer patients.

worldwide sharing of cancer data to save lives. The

Global Alliance for Genomics and Health argue how

# Barriers to Sharing Data

COMMENT | VOLUME 18, ISSUE 1, P15-16, JANUARY 01, 2017



Purchase



Subscribe

# Challenges of data sharing: value or liability?

Ian D Davis

Published: November 15, 2016 • DOI: [https://doi.org/10.1016/S1470-2045\(16\)30564-2](https://doi.org/10.1016/S1470-2045(16)30564-2)

References

Article Info

Linked Articles

Stephen Hawking wrote in the acknowledgments to his book *A Brief History of Time* that “Someone told me that each equation I included in the book would halve the sales.” Hawking and colleagues present one genuinely terrifying and possibly journal-subscription-halving prospect: the use of some very technical language and ideas. It would be easy to glaze over and hurriedly move on to the next pioneering piece of work with major implications for how clinical research data will be shared.

To read this article in full you will need to purchase or subscribe to *The Lancet Oncology*.

[Purchase one-time access](#)

[Subscribe to \*The Lancet Oncology\*](#)

[Or purchase The Lancet Choice](#)

[Want to read more? Claim online access](#)

Access any 5 articles from the Lancet Family of journals

Are you a current subscriber? [Sign in](#)

Register: [Create an account](#)

## NIH requirements are sufficient

“The NIH will expect investigators supported by NIH funding to make their research data available to the scientific community for subsequent analyses.”

- Grants over \$500k expected to include a data-sharing plan
- Certain institutes and data types have specific requirements

NIH requirements  
are sufficient

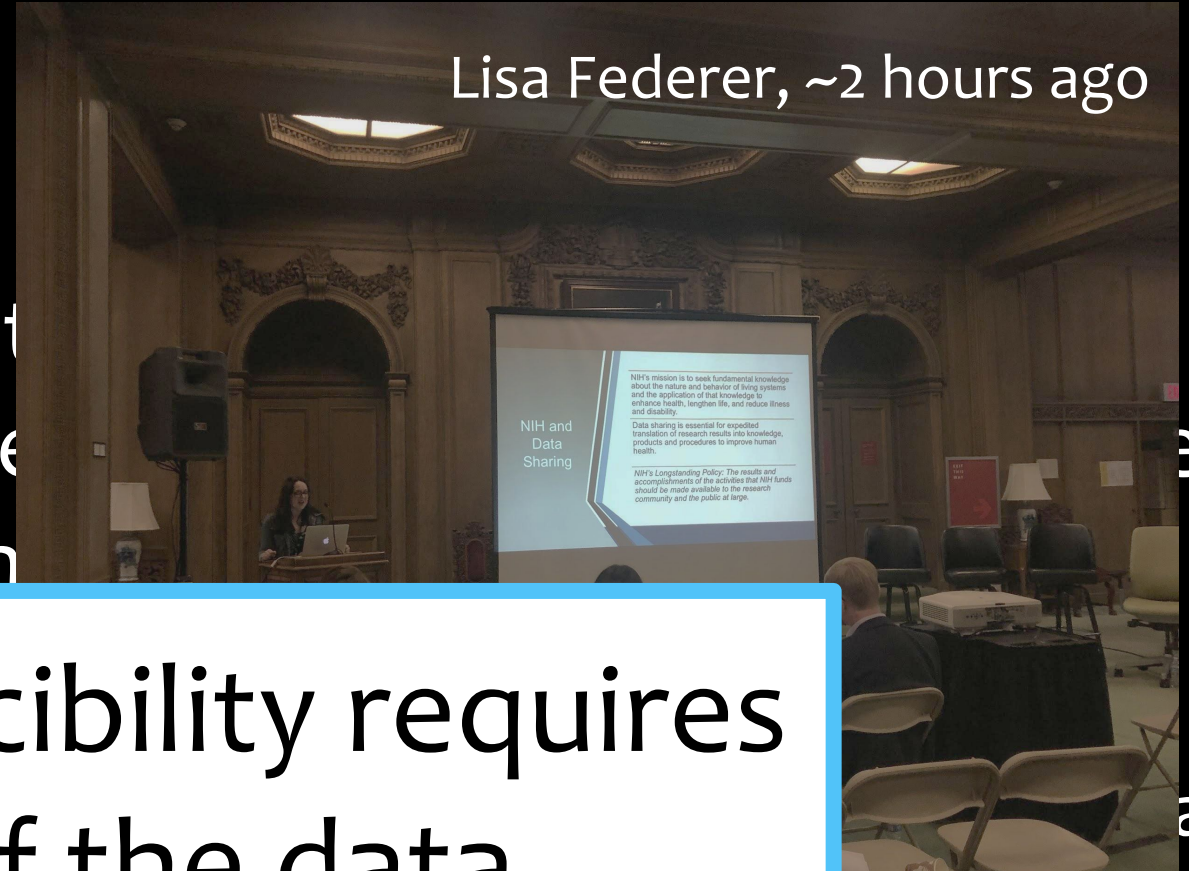
“The NIH will expect  
NIH funding to make  
to the scientific com  
analys

- Gran
- Cert

requirements

Reproducibility requires  
all of the data

Lisa Federer, ~2 hours ago



# Interoperability challenges

Billing  
codes

Data  
standards

Workflows

Metadata  
standards

Data types

File types

Analytical  
methods

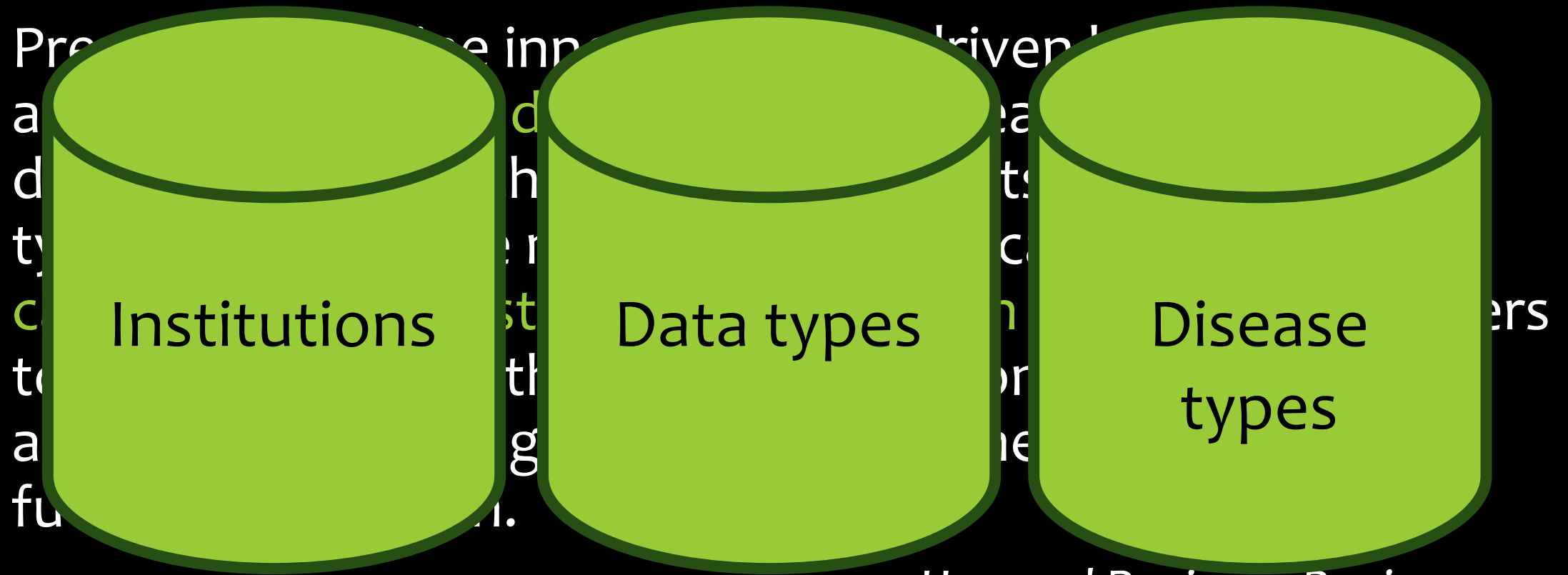
# Silos

Precision medicine innovations are driven by complex analysis of **massive data sets**, with researchers querying data derived from thousands of patients for each cancer type. For all but the most prevalent of cancers, **no single cancer center or institution has enough data** for researchers to gain insight into the genetic mutations and molecular abnormalities that ignite the development of cancer and fuel its progression.

— *Harvard Business Review* 2017



# Silos



— *Harvard Business Review* 2017

# Privacy **versus** Data Sharing

## Fighting cancer by sharing data

South Korea's National  
against global cancer in

Produced by

**nature**research  
CUSTOM MEDIA

국립  
NATIO

“One of the biggest challenges to NCC’s operations is South Korea’s **data-protection legislation**, considered one of the world’s most restrictive. Driven by respect for individuals’ privacy, it **prevents the spread of information** that could ultimately save lives.”

Trending: [Microsoft and Nokia reunite for new alliance, 5 years after failed \\$7B smartphone megadeal](#)

# Adaptive Biotechnologies stock soars 90% after ringing Nasdaq opening bell following \$300M IPO

BY [JAMES THORNE](#) on June 27, 2019 at 8:08 am

[Post a Comment](#) [f Share 60](#) [Tweet](#) [Share](#) [Reddit](#) [Email](#)



Trending: [Microsoft and Nokia reunite for new alliance, 5 years after failed \\$7B smartphone megadeal](#)

## Juno Therapeutics acquired by Celgene for \$9B in dramatic deal for rising biotech star

BY [CLARE MCGRANE](#) on January 22, 2018 at 7:52 am

[1 Comment](#) [f Share 44](#) [Tweet](#) [Share](#) [Reddit](#) [Email](#)

[The Playbook 11/7: Get tickets today!](#)



# Potential profits

# Lack of training (Researchers)

“Why aren’t people learning this in grad school?”



THE  
CARPENTRIES

FH  
.IO

# Lack of training (Clinicians)

Healthcare IT News

TOPICS

Asia P

[Analytics](#)

## Data and analytics come to the med school curriculum

'If y  
disa

By Ka

Figuring out how to access and interpret all that data is not a skill that most physicians learned in medical school. In fact, it's not even been taught in medical school, but that's changing.



# The wrong incentives

Citation!  
Tenure and promotion!  
New metrics!  
Funders and institutions!



# Signs of progress



nextflow



# Signs of progress



[Become a Member](#)

Science

[Contents](#) ▾

[News](#) ▾

[Careers](#) ▾

[Journals](#) ▾

## In departure for NIH, Cancer Moonshot requires grantees to make papers immediately free

By [Jocelyn Kaiser](#) | Aug. 14, 2019 , 5:20 PM

The long-standing debate over open access to research results has been marked by a geographic divide. In Europe, some public funders have launched a high-profile open-access initiative, dubbed Plan S, that would ultimately require grantees to publish only in journals that immediately make papers free to all. But in the United States, federal agencies have stuck to a decade-old policy that allows grantees to publish in journals that keep papers behind a paywall for up to 1 year. Now, the divide is starting to blur, with one prominent U.S. research program starting to require immediate open access to the peer-reviewed publications it funds.

The policy is part of the Cancer Moonshot program at the National Cancer Institute (NCI) in Bethesda, Maryland, the 7-year, \$1.8 billion research initiative spearheaded in 2016 by then–Vice President Joe Biden after his son Beau died of brain cancer. Biden felt that broader data sharing



SHARE

EDITORIAL



## Data sharing for pediatric cancers

Olena Morozova Vaske<sup>1,\*</sup>, David Haussler<sup>2,†</sup>

To advance the sharing culture, along with the NIH, pediatric cancer foundations such as the St. Baldrick's Foundation and **Alex's Lemonade Stand Foundation** have incorporated genomic data-sharing requirements into their grants processes.



PDF

ers

th and Human Services and  
of President Trump's pledge  
ough the United States has  
ly more resources and access  
e that the spotlight on  
pediatric cancer community to

elevate genomic data sharing to a level where its full potential can finally be realized

## A Future for NIH in Data Science and Open Science

### What we need to do to get there...

- **Policy and implementation**
- **Citation and incentivization**
  - National Academies of Sciences, Engineering, and Medicine (NASEM) Roundtable on Aligning Incentives for Open Science
  - **Value assessment**
    - 2017 NIH- NSF Science of Science Innovation Policy (SciSIP) Workshop on The Value of Data Sharing
    - New NASEM study on "Forecasting Costs for Preserving, Archiving, and Promoting Access to Biomedical Data: A Study and Workshop for the National Library of Medicine."
- **At-scale curation**
- **Coordinate and partner** with other funders and organizations



U.S. National Library of Medicine

website

[strasser.github.io](https://strasser.github.io)

email

[carlystrasser@gmail.com](mailto:carlystrasser@gmail.com)

[carly@fredhutch.org](mailto:carly@fredhutch.org)

twitter

[@carlystrasser](https://twitter.com/carlystrasser)



**FRED HUTCH**  
CURES START HERE™