MNI Open Science Symposium

John Wilbanks
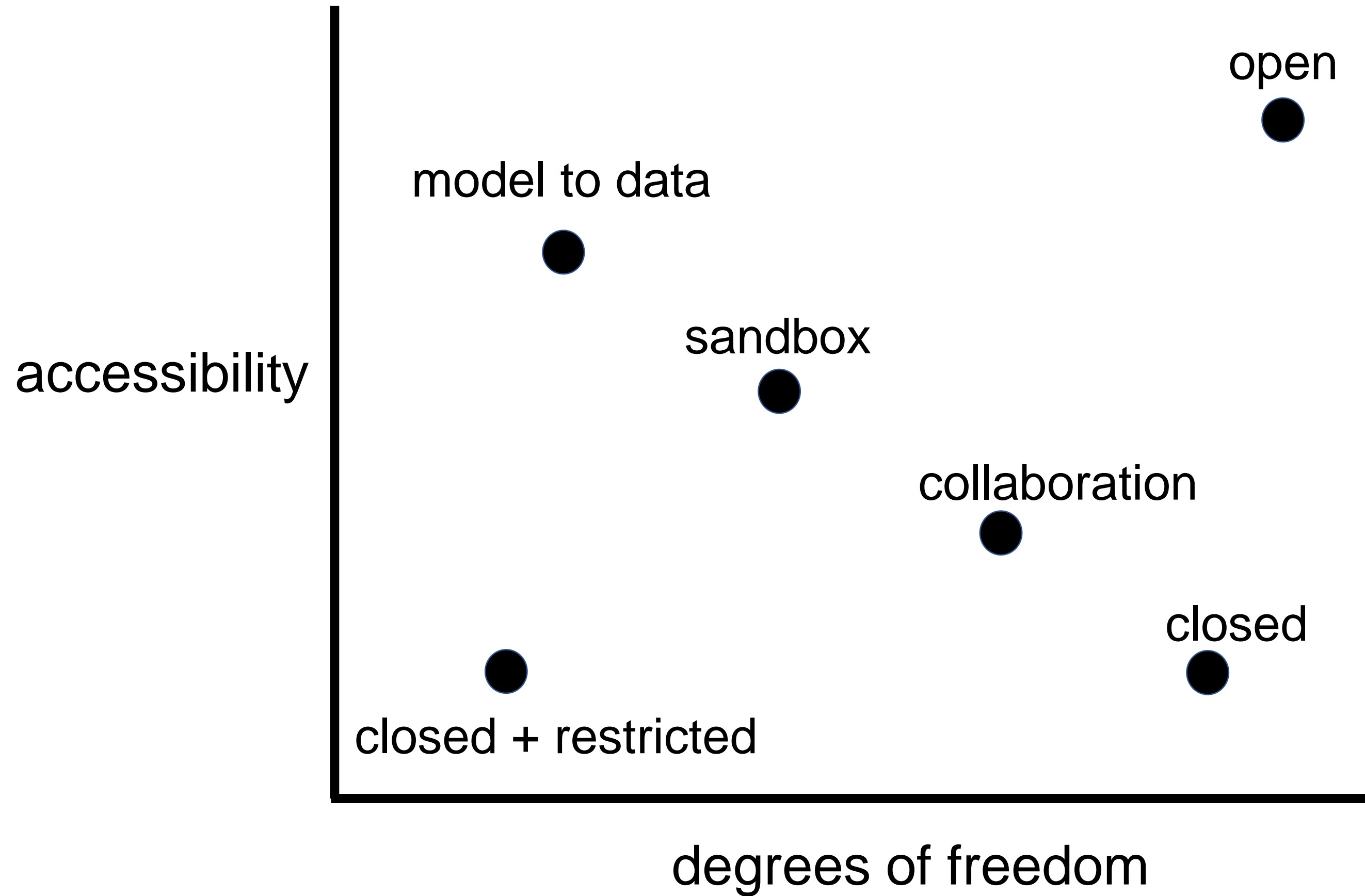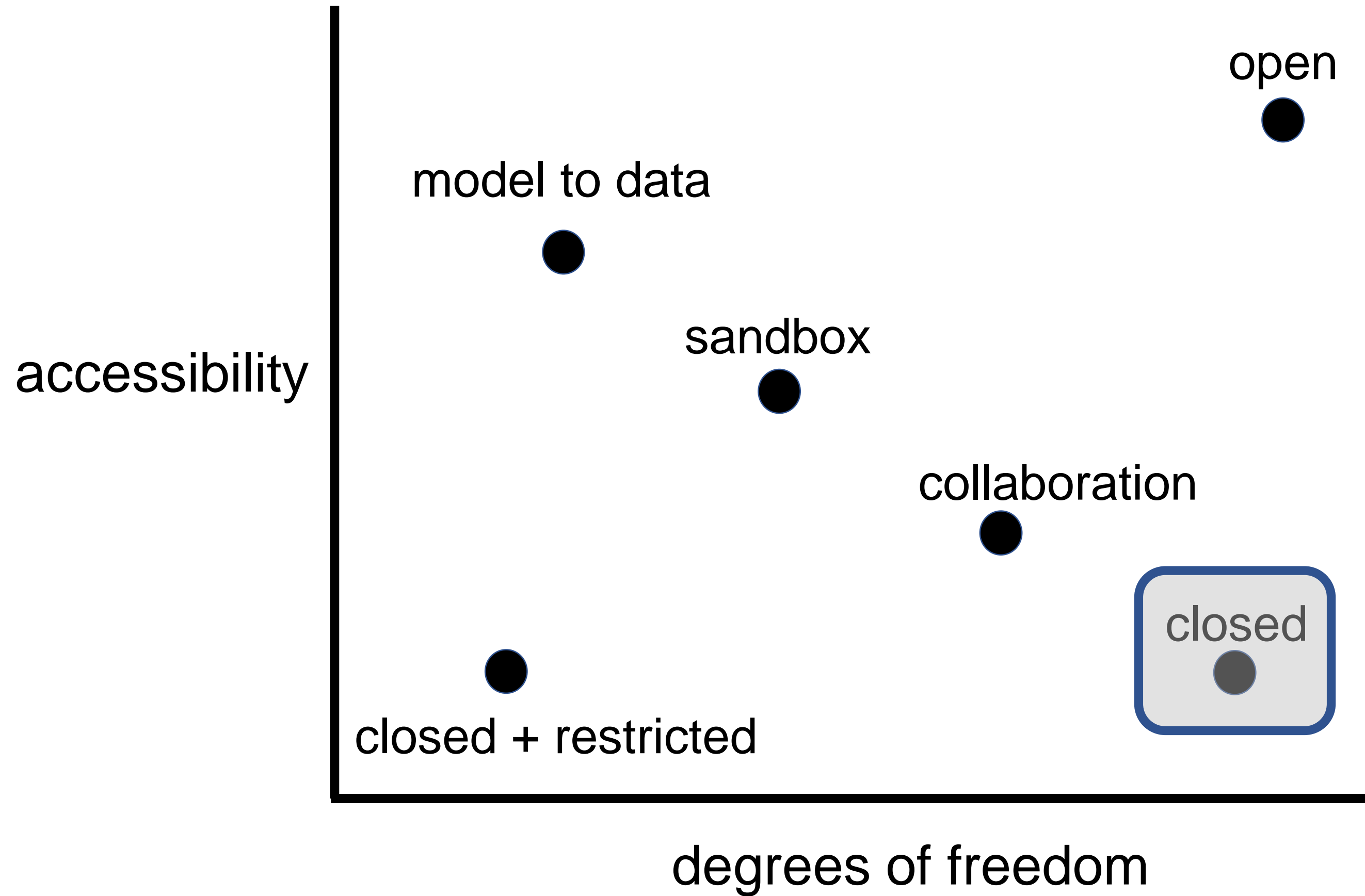
"massive incrementalism"

heather piwowar
@researchremix

Between 0.6 and 5
published papers per
$100k funding

Detailed Indexing Statistics: 1965-2017

MEDLINE® consists of completed citations indexed with MeSH® (Medical Subject Headings®).

| Fiscal Year (Oct. 1-Sep. 30)	Number of Journals Indexed in *Index Medicus*	Number of Journals in MEDLINE	Number of Citations in MEDLINE	Total Citations[1]
2017	5,150	5,617[3]	813,598	24,335,332
2016	5,136	5,623[3]	869,666	23,531,948
2015	5,123	5,618[3]	806,326	22,391,870
2014	5,118	5,647[3]	765,850	21,582,742
2013	5,067	5,640[3]	734,052	20,695,240
2012	5,025	5,633	760,903	19,974,272
2011	4,946	5,559	724,831	19,155,303
2010	4,866	5,484	699,420	18,340,055
2009	4,759	5,394	712,675	17,641,559
2008	4,660	5,319	671,904	16,888,640
2007	4,530	5,194	670,943	16,113,221[2]
2006	4,416	5,020	623,089	14,103,589
2005	4,279	4,928	606,000	13,476,222
2004	4,189	4,839	571,000	

propagation of uncertainty

within a statistical experiment...

Example formulae [edit]

This table shows the variances of simple functions of the real variables A, B, with standard deviations σ_A, σ_B, covariance σ_{AB} and exactly known real-valued constants a, b (i.e., $\sigma_a = \sigma_b = 0$).

Function	Variance	Standard Deviation
$f = aA$	$\sigma_f^2 = a^2 \sigma_A^2$	$\sigma_f = \|a\|\sigma_A$
$f = aA + bB$	$\sigma_f^2 = a^2 \sigma_A^2 + b^2 \sigma_B^2 + 2ab\,\sigma_{AB}$	$\sigma_f = \sqrt{a^2 \sigma_A^2 + b^2 \sigma_B^2 + 2ab\,\sigma_{AB}}$
$f = aA - bB$	$\sigma_f^2 = a^2 \sigma_A^2 + b^2 \sigma_B^2 - 2ab\,\sigma_{AB}$	$\sigma_f = \sqrt{a^2 \sigma_A^2 + b^2 \sigma_B^2 - 2ab\,\sigma_{AB}}$
$f = AB$	$\sigma_f^2 \approx f^2 \left[\left(\frac{\sigma_A}{A}\right)^2 + \left(\frac{\sigma_B}{B}\right)^2 + 2\frac{\sigma_{AB}}{AB} \right]$ [11][12]	$\sigma_f \approx \|f\| \sqrt{\left(\frac{\sigma_A}{A}\right)^2 + \left(\frac{\sigma_B}{B}\right)^2 + 2\frac{\sigma_{AB}}{AB}}$
$f = \frac{A}{B}$	$\sigma_f^2 \approx f^2 \left[\left(\frac{\sigma_A}{A}\right)^2 + \left(\frac{\sigma_B}{B}\right)^2 - 2\frac{\sigma_{AB}}{AB} \right]$ [13]	$\sigma_f \approx \|f\| \sqrt{\left(\frac{\sigma_A}{A}\right)^2 + \left(\frac{\sigma_B}{B}\right)^2 - 2\frac{\sigma_{AB}}{AB}}$
$f = aA^b$	$\sigma_f^2 \approx \left(abA^{b-1}\sigma_A\right)^2 = \left(\frac{fb\sigma_A}{A}\right)^2$	$\sigma_f \approx \left\|abA^{b-1}\sigma_A\right\| = \left\|\frac{fb\sigma_A}{A}\right\|$
$f = a\ln(bA)$	$\sigma_f^2 \approx \left(a\frac{\sigma_A}{A}\right)^2$ [14]	$\sigma_f \approx \left\|a\frac{\sigma_A}{A}\right\|$
$f = a\log_{10}(bA)$	$\sigma_f^2 \approx \left(a\frac{\sigma_A}{A\ln(10)}\right)^2$ [14]	$\sigma_f \approx \left\|a\frac{\sigma_A}{A\ln(10)}\right\|$

at the "claim" level…
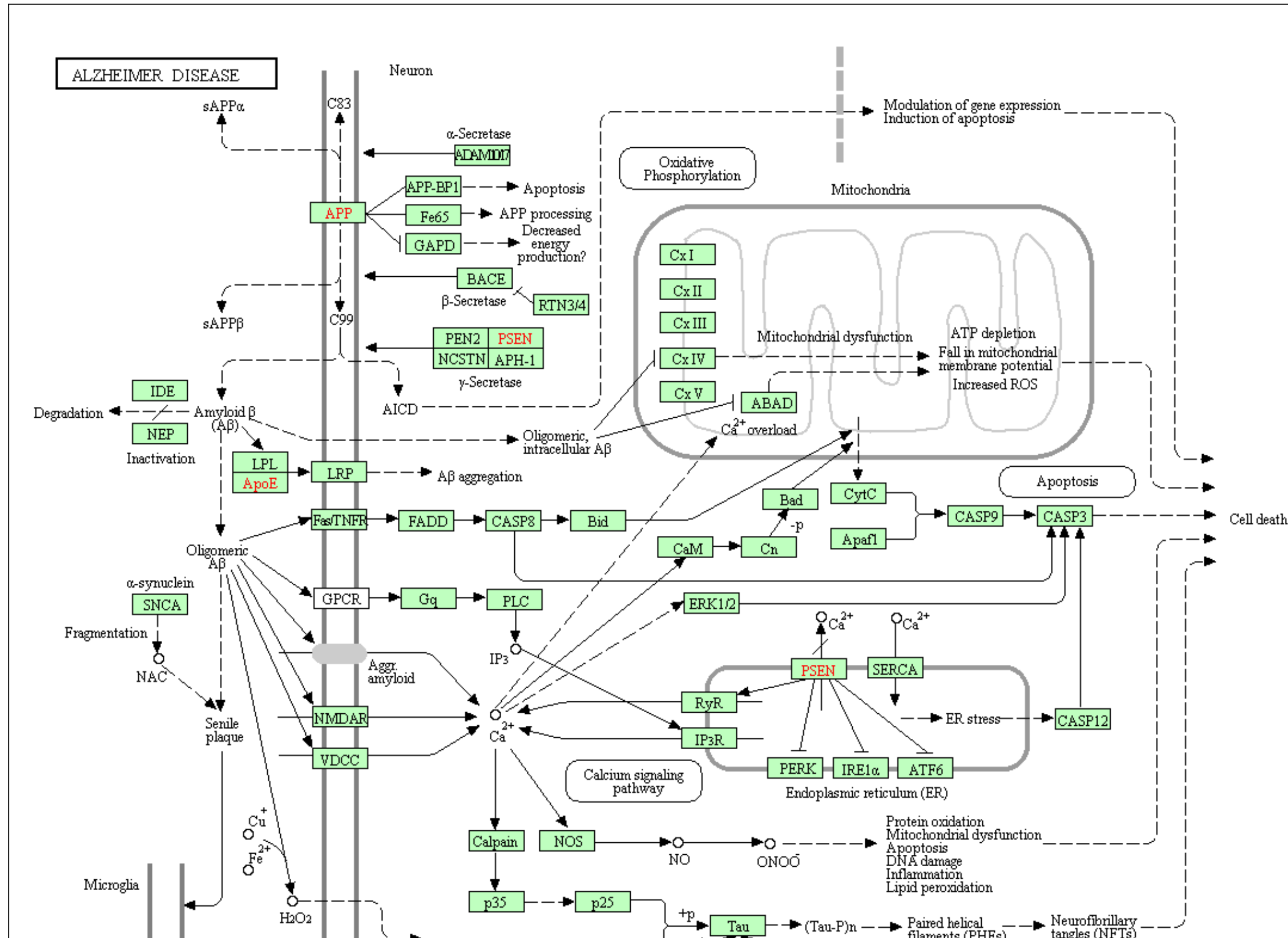
ALZHEIMER DISEASE

Search Results   7,501 items

Summative Effects of Vascular Risk Factors on the Progression of Alzheimer Disease.

Lee WJ, et al. J Am Geriatr Soc. 2019

Full text

Genetic and epigenetic study of an Alzheimer's disease family with monozygotic triplets.

Zhang M, et al. Brain. 2019

Full text

Associations Between Midlife but not Late-Life, Elevated Coronary Heart Disease Risk and Lower Cognitive Performance: Results From the Framingham Offspring Study.

Armstrong NM, et al. Am J Epidemiol. 2019

Full text

APOE ε4, white matter hyperintensities, and cognition in Alzheimer and Lewy body dementia.

Mirza SS, et al. Neurology. 2019

Full text

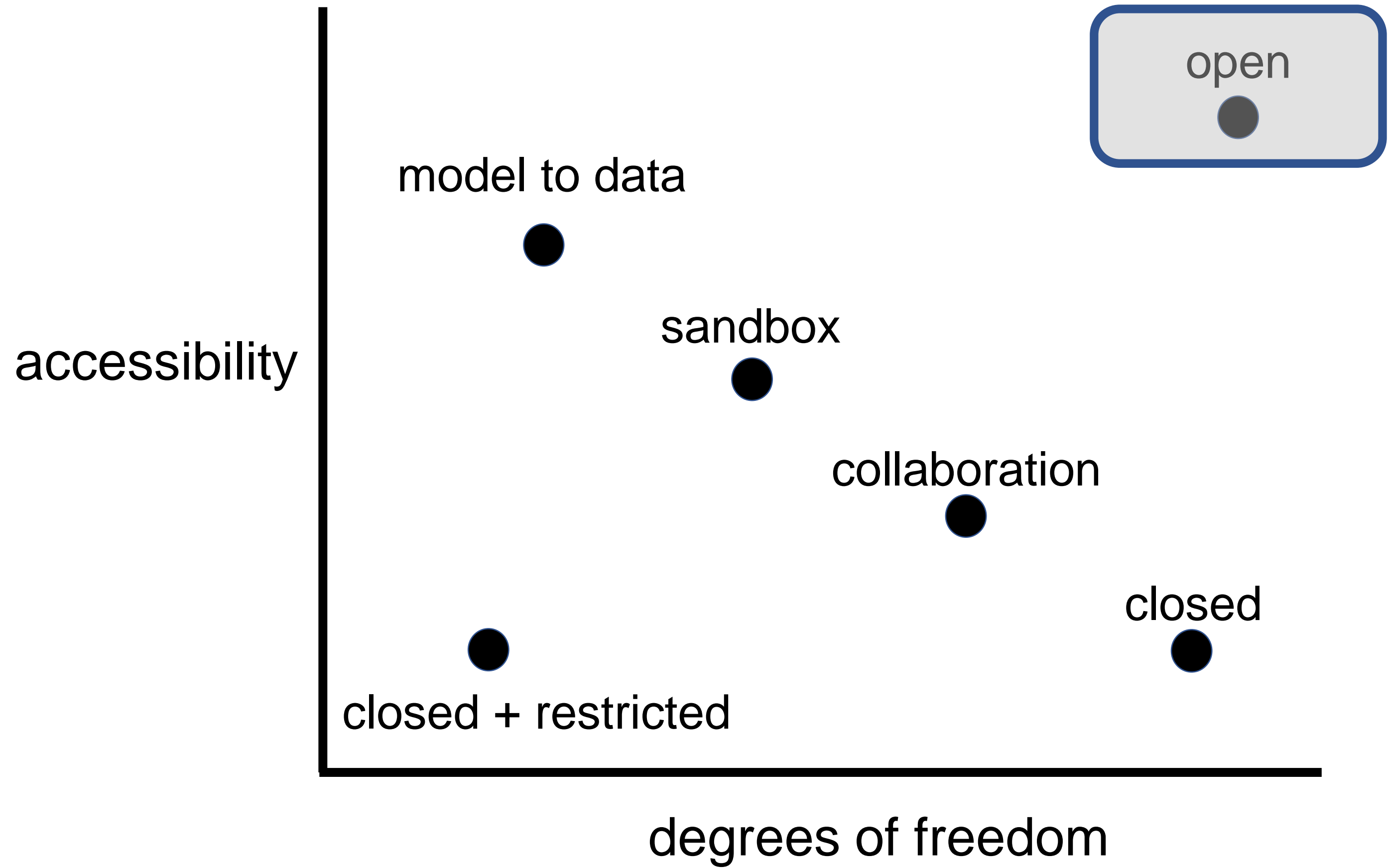**Apolipoprotein E ε4 allele effects on longitudinal cognitive trajectories are sex and age dependent.**

Williams OA, et al. Alzheimers Dement. 2019

Full text

APOE genetic variants and apoE, miR-107 and miR-650 levels in Alzheimer's disease.

Open Science Principles

Open Science Platforms

A Generous Donation

Measuring Open Science

The Road to Open Science

Our People

Open Science In The Press

Open Science at The Neuro

In 1934, Dr. Wilder Penfield had the unique vision to advance medicine through patient-centred science. Today, we continue to lead scientific innovation as **the first academic institute to develop and adopt Open Science**.

Patrons of the library at the University of California, Berkeley, will no longer have easy access to journals from the publisher Elsevier. CHUCKSTOCK/SHUTTERSTOCK.COM

University of California boycotts publishing giant Elsevier over journal costs and open access

By **Alex Fox, Jeffrey Brainard** | Feb. 28, 2019 , 7:00 PM

The mammoth University of California (UC) system announced today it will stop paying to subscribe to journals published by Elsevier, the world's largest scientific publisher, headquartered in Amsterdam. Talks to renew a collective contract broke down, the university said, because Elsevier refused to strike a package deal that would provide a break on subscri[ ] Screenshot

Availability of data, materials, code and protocols

An inherent principle of publication is that others should be able to replicate and build upon the authors' published claims. A condition of publication in a Nature Research journal is that **authors are required to make materials, data, code, and associated protocols promptly available to readers without undue qualifications**. Any restrictions on the availability of materials or information must be disclosed to the editors at the time of submission. Any restrictions must **also** be disclosed in the submitted manuscript.

After publication, readers who encounter refusal by the authors to comply with these policies should contact the chief editor of the journal. In cases where editors are unable to resolve a complaint, the journal may refer the matter to the authors' funding institution and/or publish a formal statement of correction, attached online to the publication, stating that readers have been unable to obtain necessary materials to replicate the findings.

See sections below for details on:

- reporting requirements
- availability of data
- availability of materials
- availability of computer code
- experimental protocols
- clinical trials
- futher reading

The preregistration revolution

Brian A. Nosek[a,b,1], Charles R. Ebersole[b], Alexander C. DeHaven[a], and David T. Mellor[a]

[a]Center for Open Science, Charlottesville, VA 22903; and [b]Department of Psychology, University of Virginia, Charlottesville, VA 22904

Progress in science relies in part on generating hypotheses with existing observations and testing hypotheses with new observations. This distinction between postdiction and prediction is appreciated conceptually but is not respected in practice. Mistaking generation of postdictions with testing of predictions reduces the credibility of research findings. However, ordinary biases in human reasoning, such as hindsight bias, make it hard to avoid this mistake. An effective solution is to define the research questions and analysis plan before observing the research outcomes—a process called preregistration. Preregistration distinguishes analyses and outcomes that result from predictions from those that result from postdictions. A variety of practical strategies are available to make the best possible use of preregistration in circumstances that fall short of the ideal application, such as when the data are preexisting. Services are now available for preregistration across all disciplines, facilitating a rapid increase in the practice. Widespread adoption of preregistration will increase distinctiveness between hypothesis generation and hypothesis testing and will improve the credibility of research findings.

Progress in science is marked by reducing uncertainty about nature. Scientists generate models that may explain prior observations and predict future observations. Those models are approximations and simplifications of reality. Models are iteratively improved and replaced by reducing the amount of prediction error. As prediction error decreases, certainty about what overconfidence in post hoc explanations (postdictions) and inflate the likelihood of believing that there is evidence for a finding when there is not. Presenting postdictions as predictions can increase the attractiveness and publishability of findings by falsely reducing uncertainty. Ultimately, this decreases reproducibility (6–11).

Mental Constraints on Distinguishing Predictions and Postdictions

It is common for researchers to alternate between postdiction and prediction. Ideas are generated, and observed data modify those ideas. Over time and iteration, researchers develop understanding of the phenomenon under study. That understanding might result in a model, hypothesis, or theory. The dynamism of the research enterprise and limits of human reasoning make it easy to mistake postdiction as prediction. The problem with this is understood as post hoc theorizing or hypothesizing after the results are known (12). It is an example of circular reasoning—generating a hypothesis based on observing data, and then evaluating the validity of the hypothesis based on the same data.

Hindsight bias, also known as the I-knew-it-all-along effect, is the tendency to see outcomes as more predictable after the fact compared with before they were observed (13, 14). With hindsight bias, the observer uses the data to generate an explanation, a postdiction, and simultaneously perceives that they would have anticipated that explanation in advance, a prediction. A common case is when the researcher's prediction is vague so that many possible outcomes can be rationalized after the fact as supporting the prediction. For example, a biomedical researcher might predict that a

Vice President Joe Biden meets with his moonshot federal task force earlier this month. ASSOCIATED PRESS

Biden's moonshot cancer plan calls for more data sharing

By **Jocelyn Kaiser** | Oct. 17, 2016 , 2:15 PM

Vice President Joe Biden today released his vision for doubling progress against cancer over 5 years. It includes numerous policy recommendations and a laundry list of projects by the National Cancer Institute (NCI) and other federal agencies that would require additional funding.

Biden and his wife, Jill, have met with thousands of experts and patient advoca   Screenshot   ain in

Expanding Public Access to the Results of Federally Funded Research

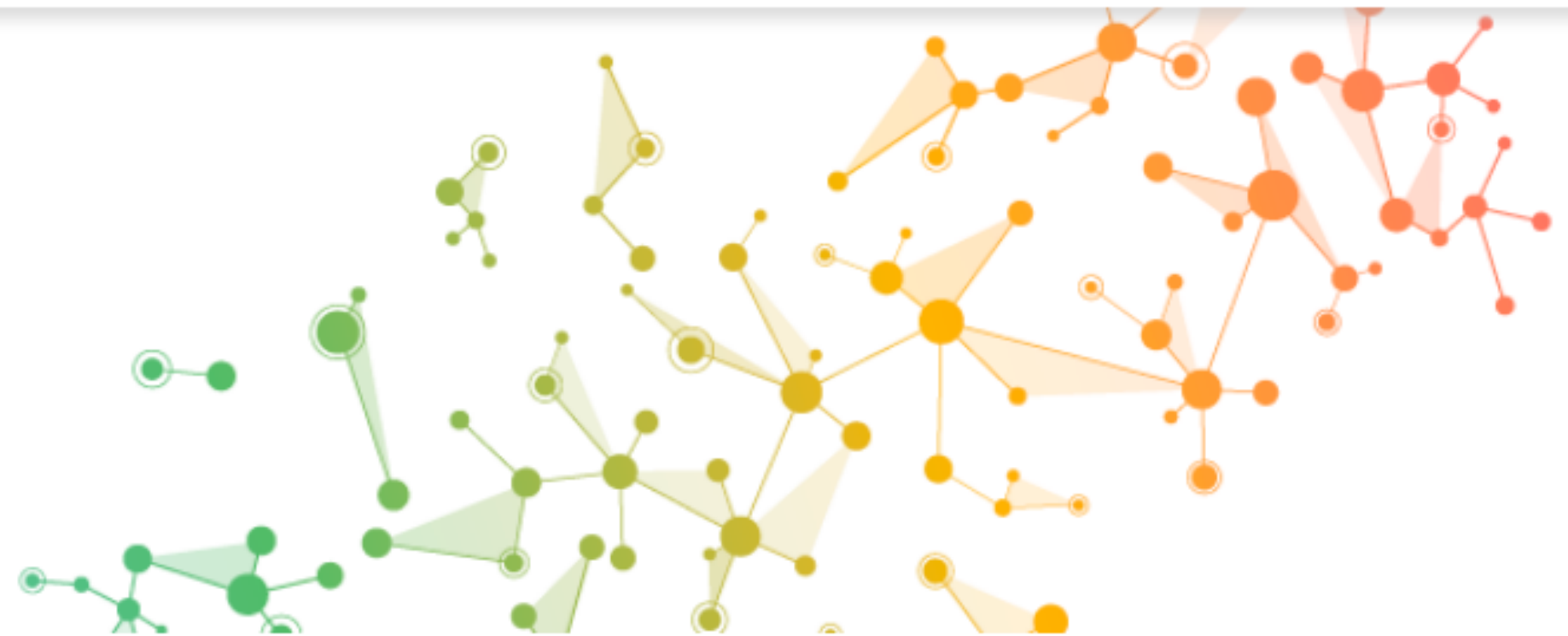FEBRUARY 22, 2013 AT 12:04 PM ET BY MICHAEL STEBBINS

Summary: The Obama Administration is committed to the proposition that citizens deserve easy access to the results of research their tax dollars have paid for. That's why, in a policy memorandum released today, OSTP Director John Holdren has directed Federal agencies with more than $100M in R&D expenditures to develop plans to make the results of federally funded research freely available to the public—generally within one year of publication.
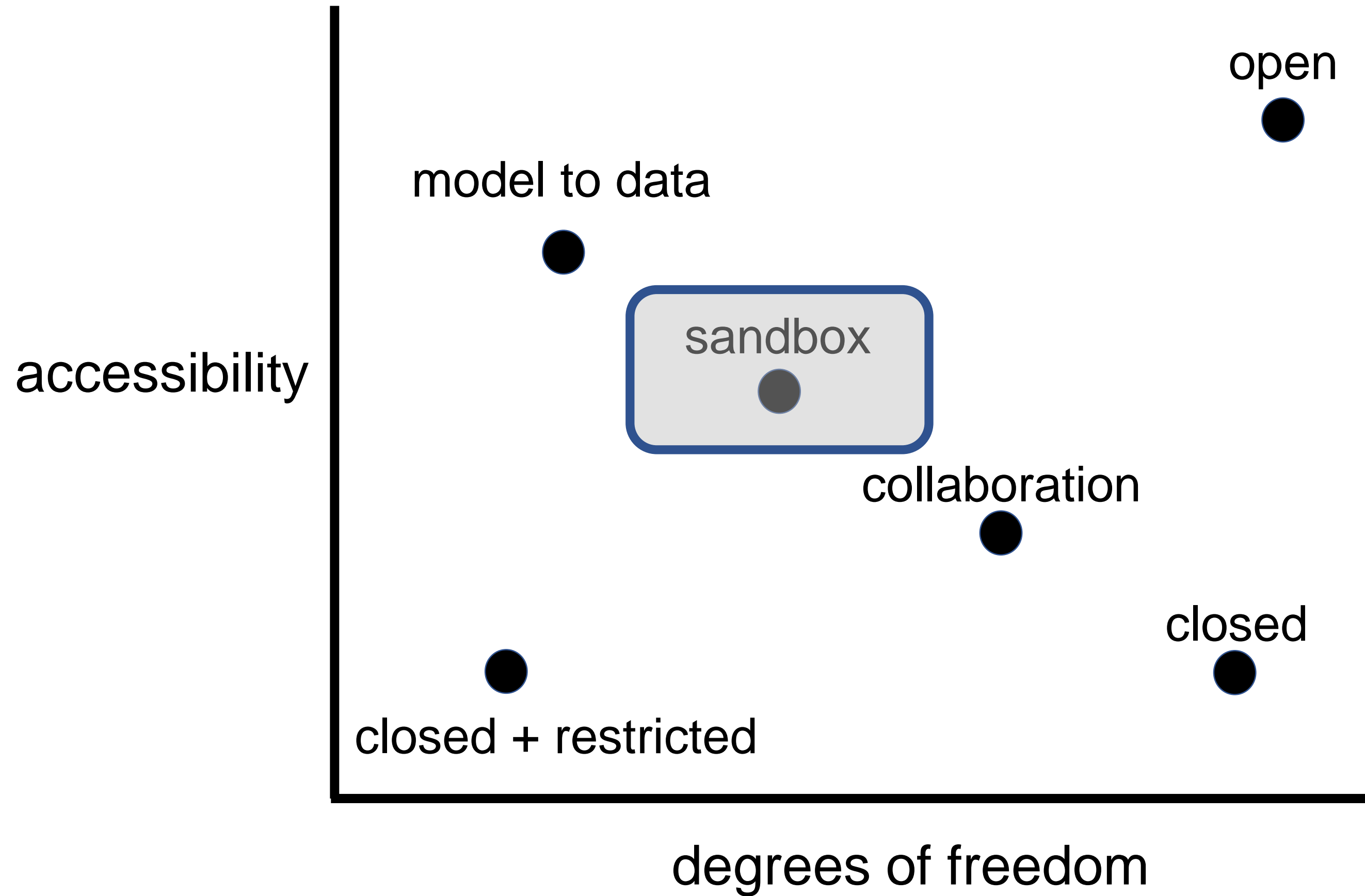
The Obama Administration is committed to the proposition that citizens deserve easy access to the results of scientific research their tax dollars have paid for. That's why, in a policy memorandum released today, OSTP Director John Holdren has directed Federal agencies with more than $100M in R&D expenditures to develop plans to make the published results of federally funded research freely available to the public within one year of publication and requiring researchers to better account for and manage the digital data resulting from federally funded scientific research. OSTP has been looking into this issue for some time, soliciting broad public input on multiple occasions and convening an interagency working group to deve[...] [...]e final policy reflects substantial inputs from scientists and scientific organizations, publishers, men[...]ers of Congress, and other

Screenshot

open

model to data

sandbox

accessibility

collaboration

closed

closed + restricted

degrees of freedom
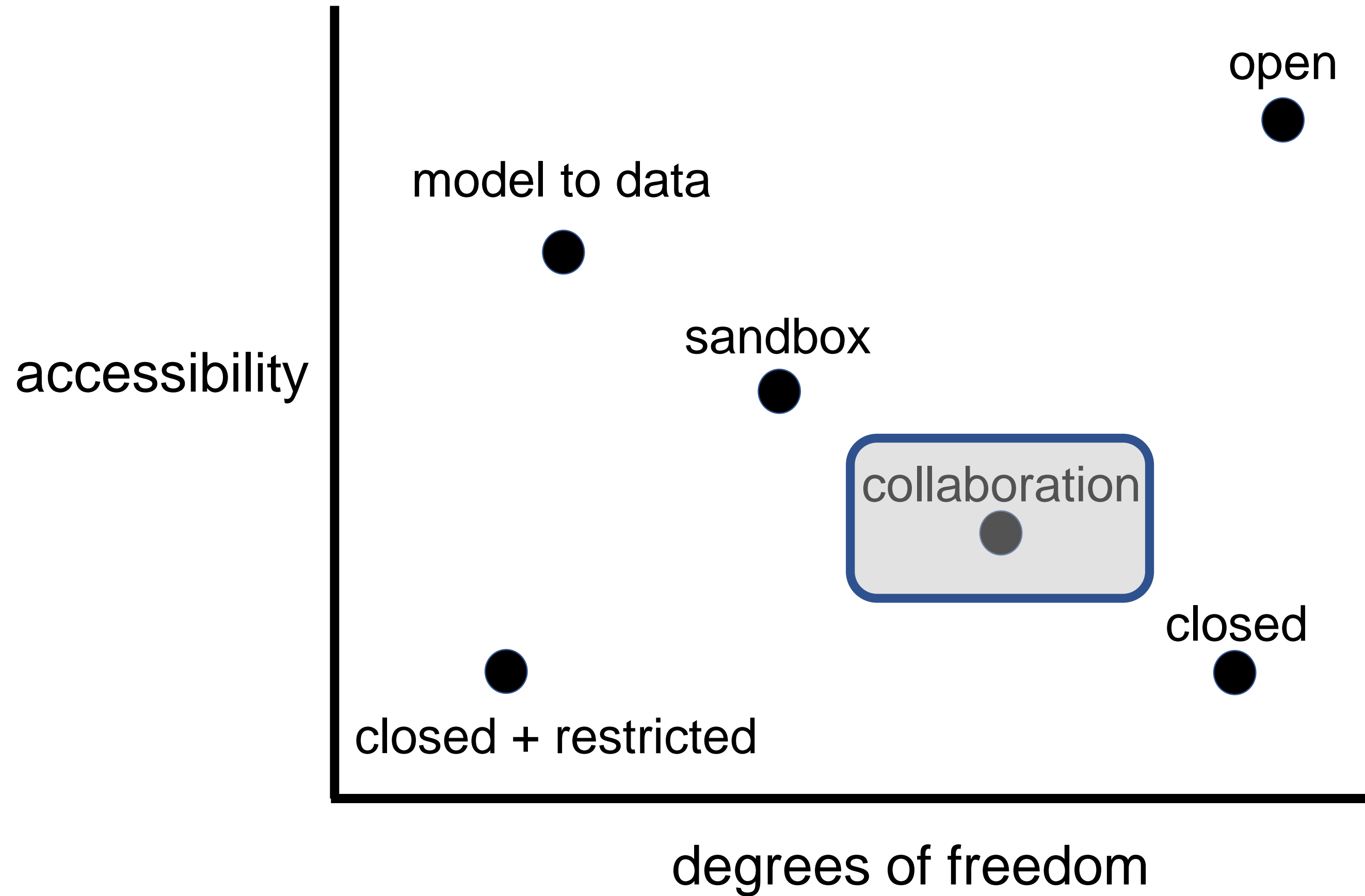
Better Science Together

At Sage Bionetworks, we believe that we can learn more by learning from each other. We partner with researchers, patients, and healthcare innovators to drive collaborative data-driven science to improve health. We advance biomedicine by making science more open, collaborative, and inclusive.

open

model to data

accessibility

sandbox

collaboration

closed

closed + restricted

degrees of freedom

crc subtyping consortium

Subtypes of primary colorectal tumors correlate with response to targeted treatment in colorectal cell lines

Andreas Schlicker[1], Garry Beran[3], Christine M Chresta[3], Gael McWalter[4], Alison Pritchard[3], Susie Weston[4], Sarah Runswick[4], Sara Davenport[3], Kerry Heathcote[3], Denis Alferez Castro[3], George Orphanides[3], Tim French[4]* and Lodewyk FA Wessels[1,2,5]*

A colorectal cancer classification system that associates cellular phenotype and responses to therapy

Anguraj Sadanandam, Costas A Lyssiotis, Krisztian Homicsko, Eric A Collisson, William J Gibb, Stephan Wullschleger, Liliane C Gonzalez Ostos, William A Lannon, Carsten Grotzinger, Maguy Del Rio, Benoit Lhermitte, Adam B Olshen, Bertram Wiedenmann, Lewis C Cantley, Joe W Gray & Douglas Hanahan

Gene Expression Classification of Colon Cancer into Molecular Subtypes: Characterization, Validation, and Prognostic Value

Laetitia Marisa, Aurélien de Reyniès, Alex Duval, Janick Selves, Marie Pierre Gaub, Laure Vescovo, Marie-Christine Etienne-Grimaldi, Renaud Schiappa, Dominique Guenot, Mira Ayadi, Sylvain Kirzin, Maurice Chazal, Jean-François Fléjou, Daniel Benchimol, Anne Berger, Arnaud Lagarde, Erwan Pencreach, Françoise Piard, Dominique Elias, Yann Parc, Sylviane Olschwang, Gérard Milano, Pierre Laurent-Puig ✉, Valérie Boige [view less]

Gene expression patterns unveil a new level of molecular heterogeneity in colorectal cancer
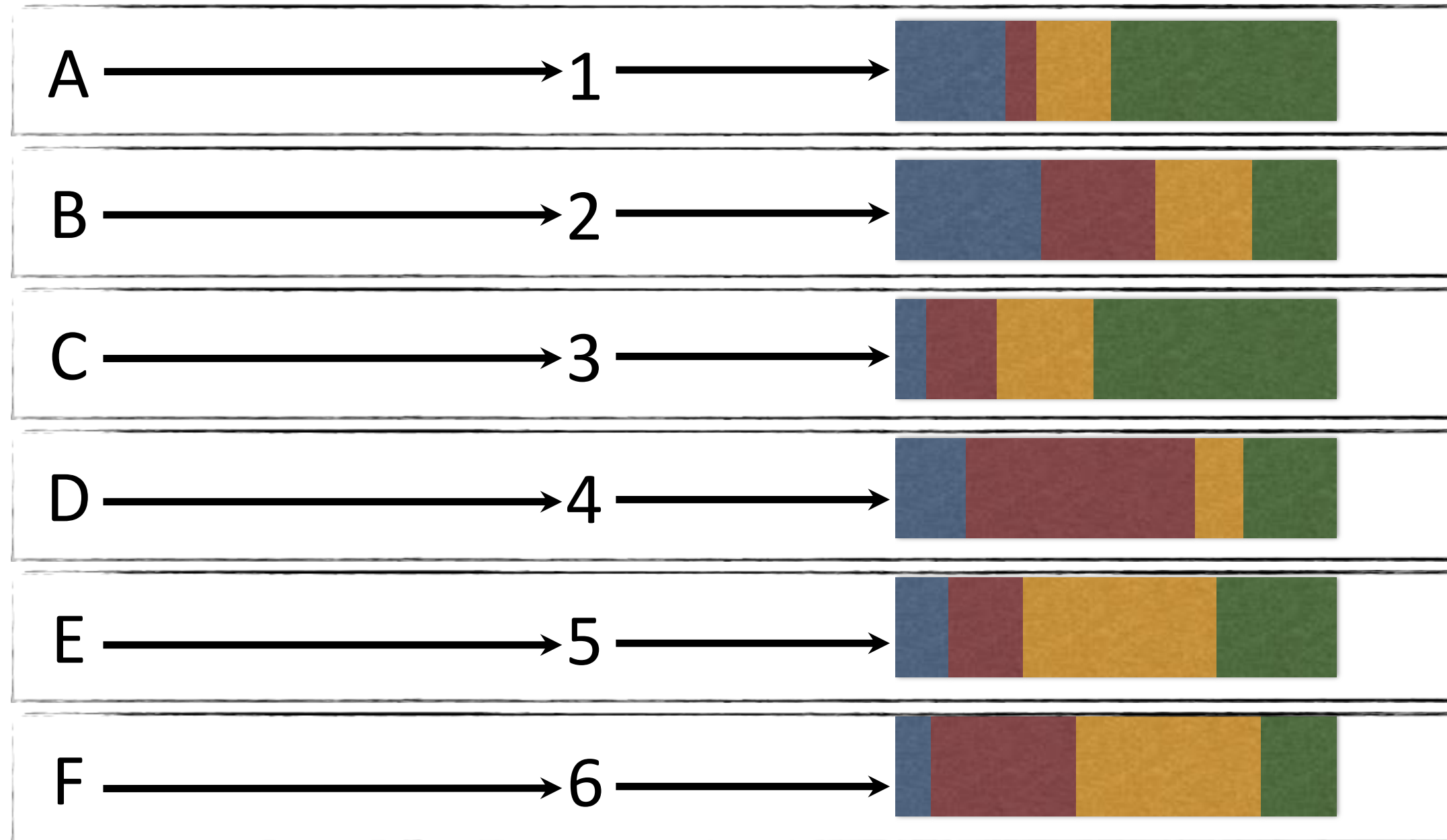
Eva Budinska,[1,2]* Vlad Popovici,[1,2] Sabine Tejpar,[3] Giovanni D'Ario,[1] Nicolas Lapique,[1] Katarzyna Otylia Sikora,[1] Antonio Fabio Di Narzo,[1] Pu Yan,[4] John Graeme Hodgson,[5] Scott Weinrich,[5] Fred Bosman,[5] Arnaud Roth[6,7] and Mauro Delorenzi[1,8]

crc subtyping consortium

expert team	data	subtype

A ⟶ 1 ⟶

B ⟶ 2 ⟶

C ⟶ 3 ⟶

D ⟶ 4 ⟶

E ⟶ 5 ⟶

F ⟶ 6 ⟶

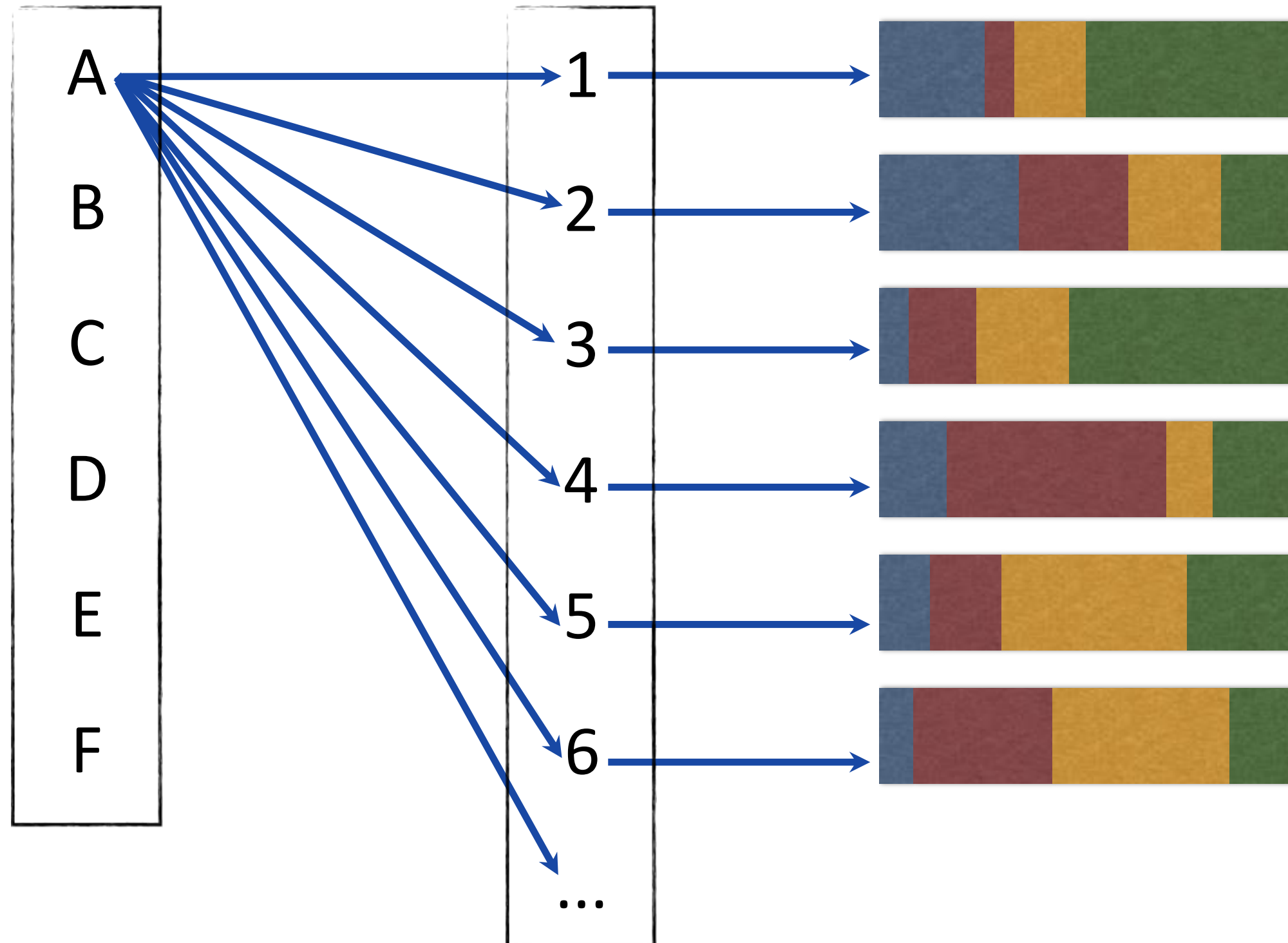# crc subtyping consortium

expert team

data

subtype

crc subtyping consortium

CMS1 MSI immune	CMS2 Canonical	CMS3 Metabolic	CMS4 Mesenchymal
14%	37%	13%	23%
MSI, CIMP high, hypermutation	SCNA high	Mixed MSI status, SCNA low, CIMP low	SCNA high
BRAF mutations		*KRAS* mutations	
Immune infiltration and activation	WNT and MYC activation	Metabolic deregulation	Stromal infiltration, TGF-β activation, angiogenesis
Worse survival after relapse			Worse relapse-free and overall survival

CIMP, CpG island methylator phenotype; MSI, microsatellite instability; SCNA, somatic copy number alterations.

accelerating medicines partnership

HEALTH

An Unusual Partnership to Tackle Stubborn Diseases
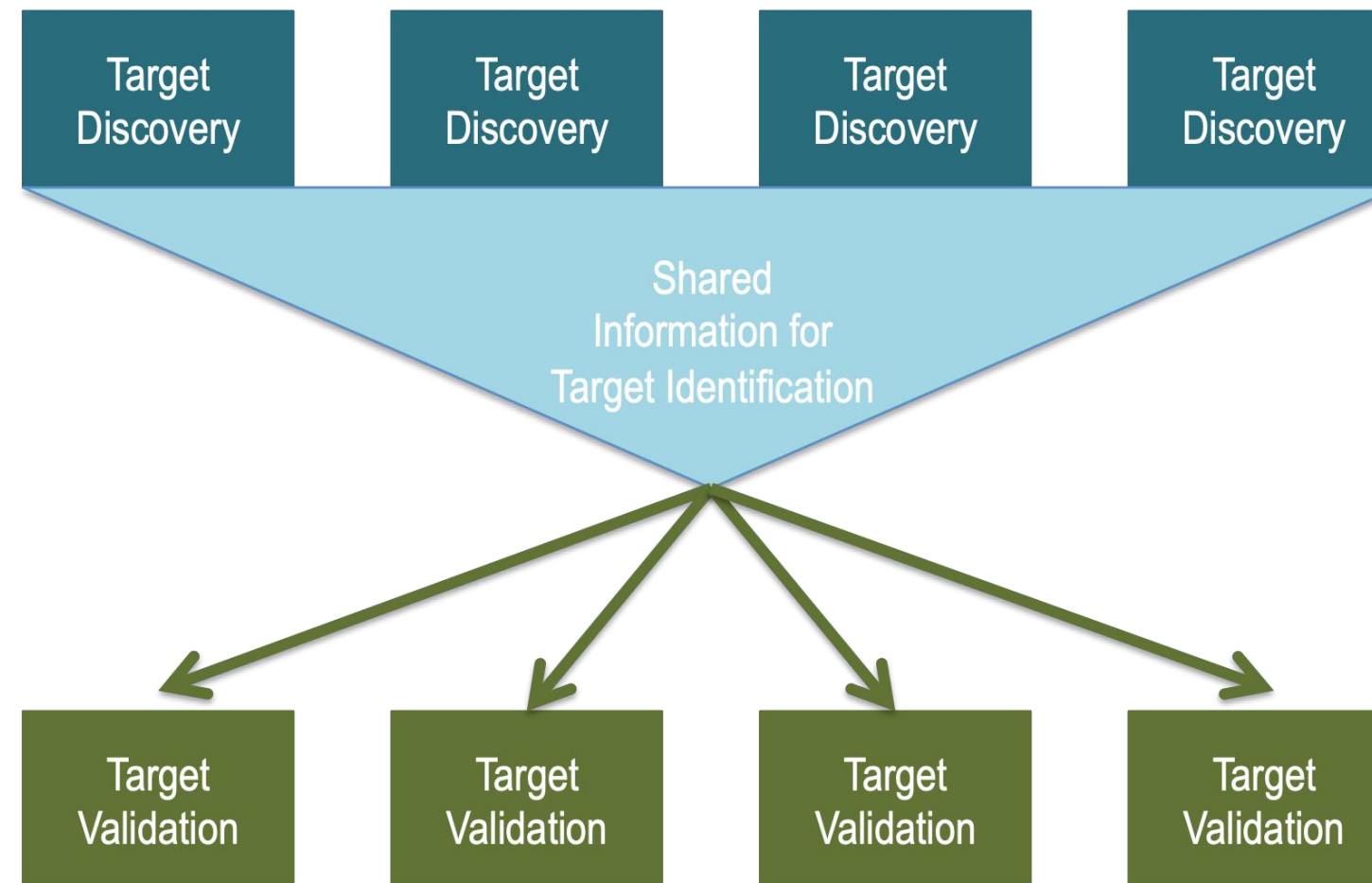
By GINA KOLATA   FEB. 4, 2014
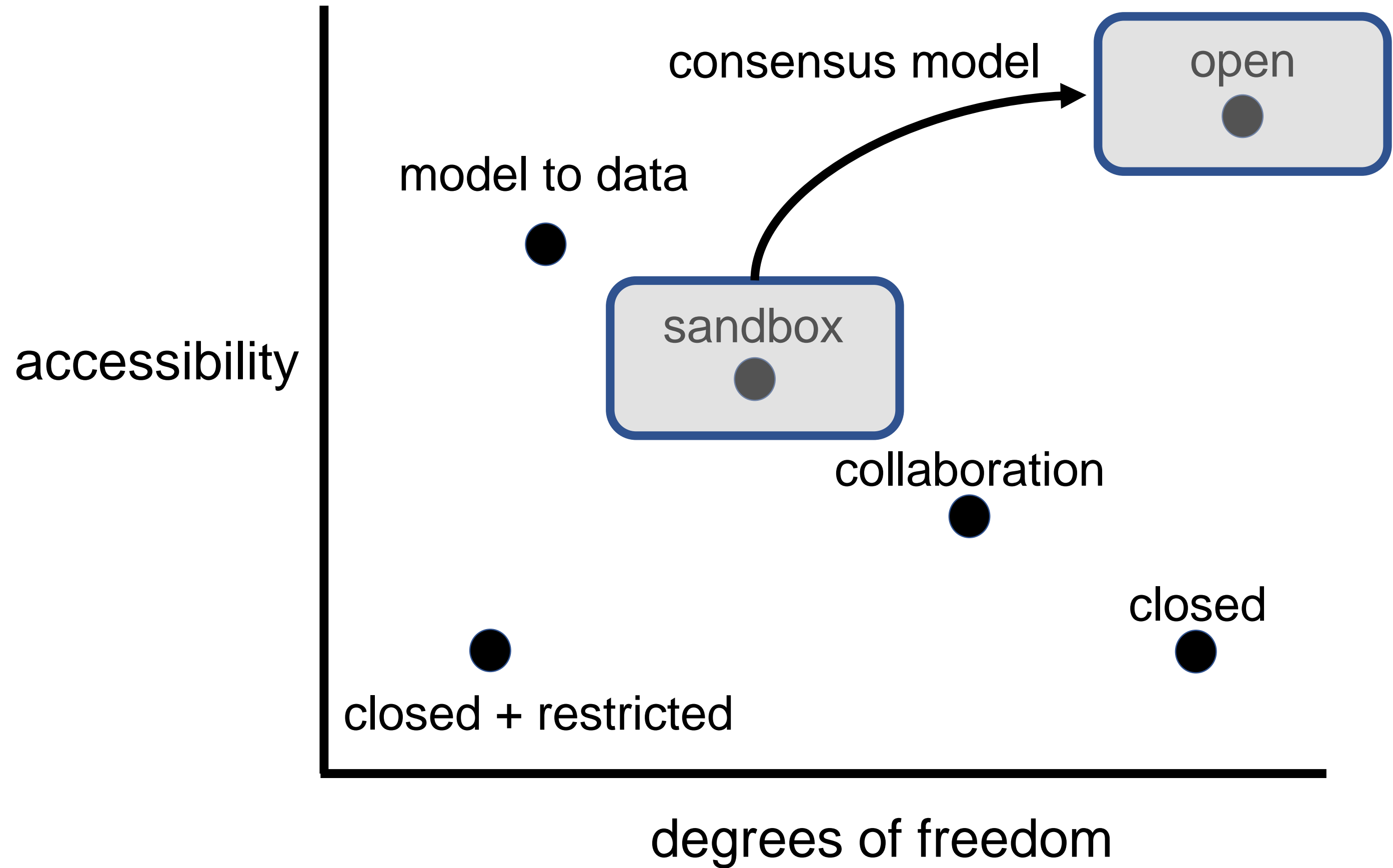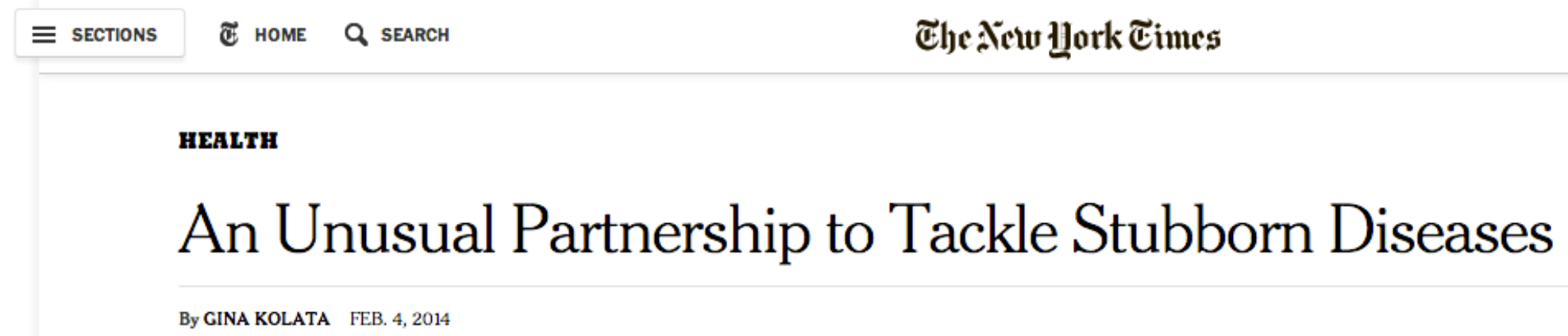
public / private partnership between NIH,
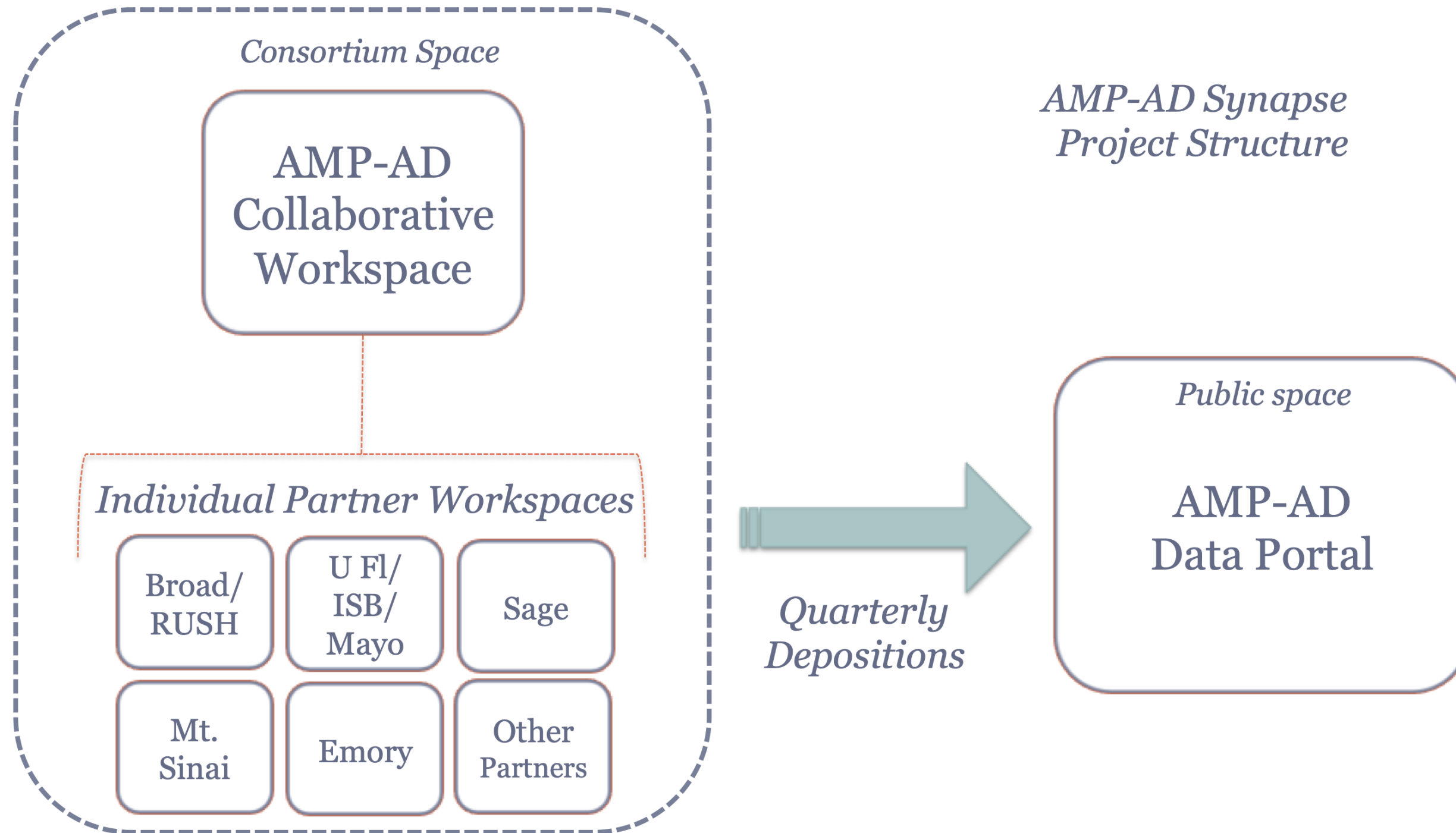10 biopharmaceutical companies
and several non-profit organizations

accelerating medicines partnership

coordinate sharing of early-phase
target identification insights

accelerating medicines partnership

AMP-AD Synapse Project Structure

Consortium Space

AMP-AD Collaborative Workspace

Individual Partner Workspaces

- Broad/RUSH
- U Fl/ISB/Mayo
- Sage
- Mt. Sinai
- Emory
- Other Partners

Quarterly Depositions

Public space

AMP-AD Data Portal

RESEARCH TOOLS

These research tools have been generated by AMP-AD, MODEL-AD, M2OVE-AD, and Resilience-AD teams. Follow the links in the summaries below to learn how to obtain these resources for use in your own research.

Recombinant Adeno Associated Virus Vector (rAAV) Tools

Description: An extensive library of rAAV tools for the study of neurodegeneration.

Resource Contact: Todd Golde

Provided By: The UFL-Mayo-ISB AMP-AD Grant

More Information: This documentation provides an overview of promoters, capsids, fluorescent proteins, organelle markers and neurodegenerative disease relevant proteins that have been encoded in the available rAAV vectors and a list of the various rAAV constructs that have been generated. The documentation also describes any Material Transfer Agreements that are required to obtain the rAAV vectors.
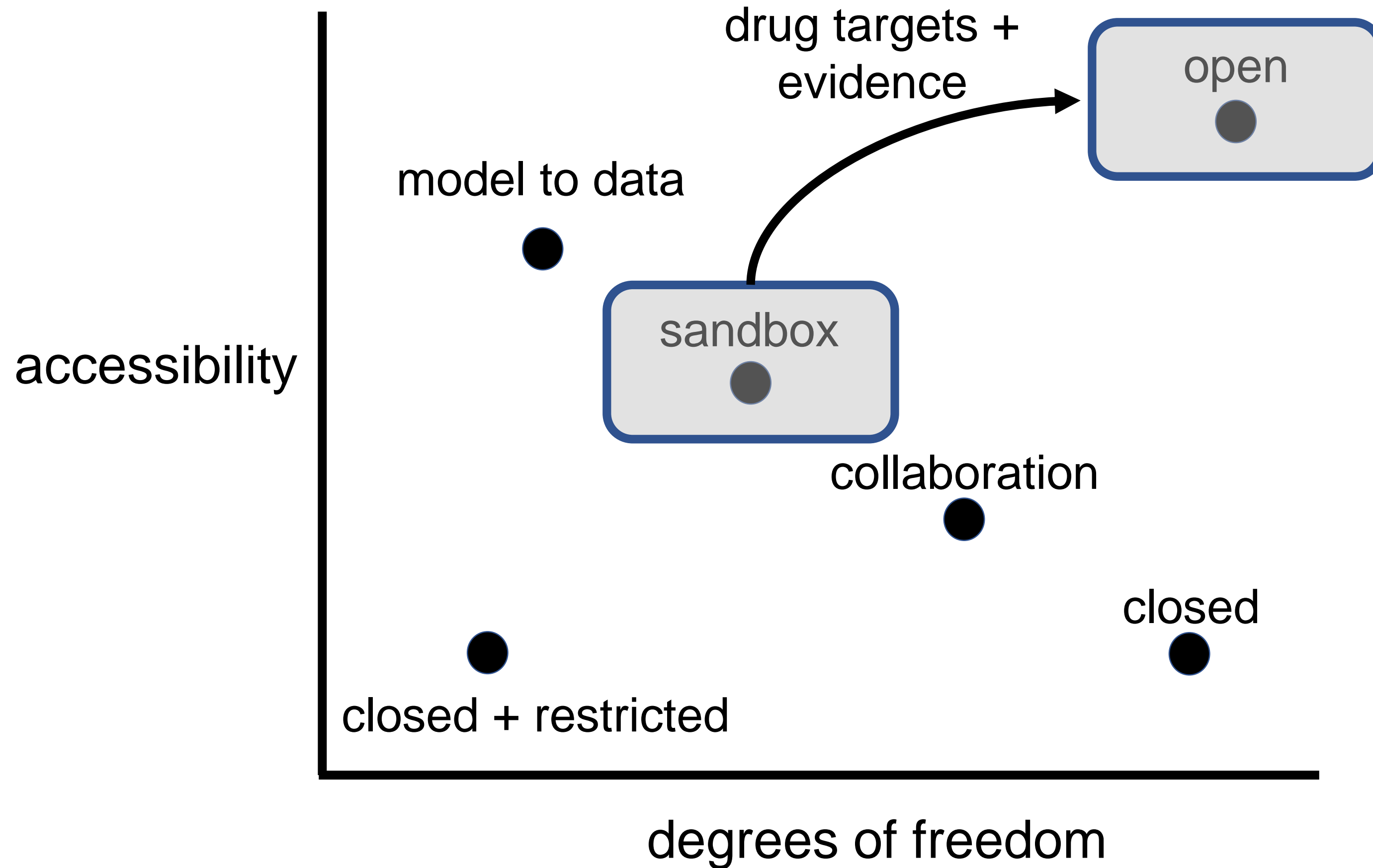
Mouse Models: University of Florida

Description: Amyloid precursor protein (APP) and microtubule-associated protein tau (MAPT) mouse transgenic models.

Resource Contact: Todd Golde

Provided By: The UFL-Mayo-ISB AMP-AD Grant

More Information: See the table below for a list of models. RNAseq data and other results from the CRND8, PS1/APP, MAPT_P301L, and rTG4510 models are available through The TAUAPPms study.

Model Type	Model Name	Transgenes	Transgene Technology	Model Phenotypes	Contact/Access Requirement
APP Mouse TG	CRND8	APP695 SWE Indiana	Hamster Prion Promoter	Amyloid plaques by 2 months of age, Beh... ...f ... months	MTA from U. Toronto Contact Todd Golde UF or http://www.informatics.jax.org/allele/MGI:3589475

epistemically diverse findings

triangulate to better bets

Alzheimer Centers for Discovery of New Medicines (U54 Clinical Trial Not Allowed)

U54 Specialized Center- Cooperative Agreements

New

- **November 26, 2018** - NIH & AHRQ Announce Upcoming Updates to Application Instructions and Review Criteria for Research Grant Applications. See Notice NOT-OD-18-228.
- September 21, 2018 - Notice of Pre-Application Webinar for RFA-AG-19-010. See Notice NOT-AG-18-028.
- August 10, 2018 - Notice of Correction to Eligibility Information in RFA-AG-19-010. See Notice NOT-AG-18-025.

RFA-AG-19-010

None

See Section III. 3. Additional Information on Eligibility.

93.866

This Funding Opportunity Announcement (FOA) invites U54 Cooperative Agreement applications aiming to establish multi-component Alzheimer Centers for the Discovery of New Medicines. The overarching purpose of this Centers program is to improve, diversify and reinvigorate the Alzheimer's disease (AD) drug development pipeline by accelerating the characterization and experimental validation of next generation therapeutic targets and integrating the targets into drug

triangulate to better bets

Part 2. Full Text of Announcement

Section I. Funding Opportunity Description

Purpose

This Funding Opportunity Announcement (FOA) invites applications to establish multi-component Alzheimer Centers for Discovery of New Medicines. The overarching objective of this Centers program is to improve, diversify and reinvigorate the Alzheimer's disease (AD) drug development pipeline by accelerating the characterization and experimental validation of next generation therapeutic targets and integrating the targets into drug discovery campaigns. More specifically, each funded Center will 1) design, develop and disseminate tools that support target enabling packages (TEPs) for the experimental validation of novel next generation therapeutic targets, including those emanating from the NIA-funded target discovery programs such as AMP-AD/AMP-AD Wall of Targets, and 2) initiate early stage drug discovery campaigns against the enabled targets. To achieve these goals, it is expected that each Center will be staffed by a multi-disciplinary team of scientists with combined expertise in data science, computational biology, network biology, disease biology, structural genomics, biostatistics, assay development, medicinal chemistry, pharmacology, and clinical science. Central to this initiative is the open-access, rapid dissemination of data, methods, and computational and experimental tools generated by the Centers to all qualified researchers for their use in advancing AD drug discovery and AD disease biology.
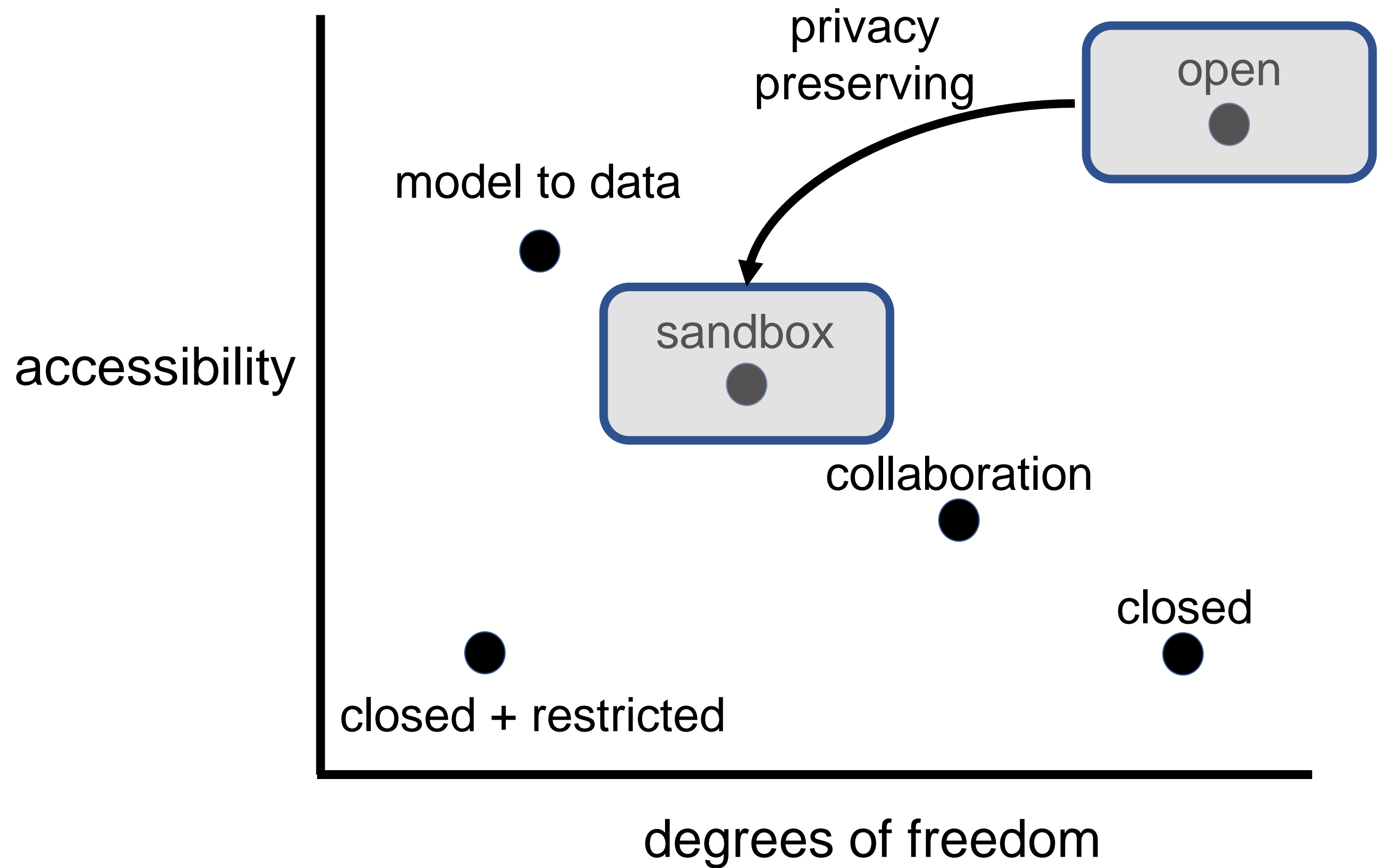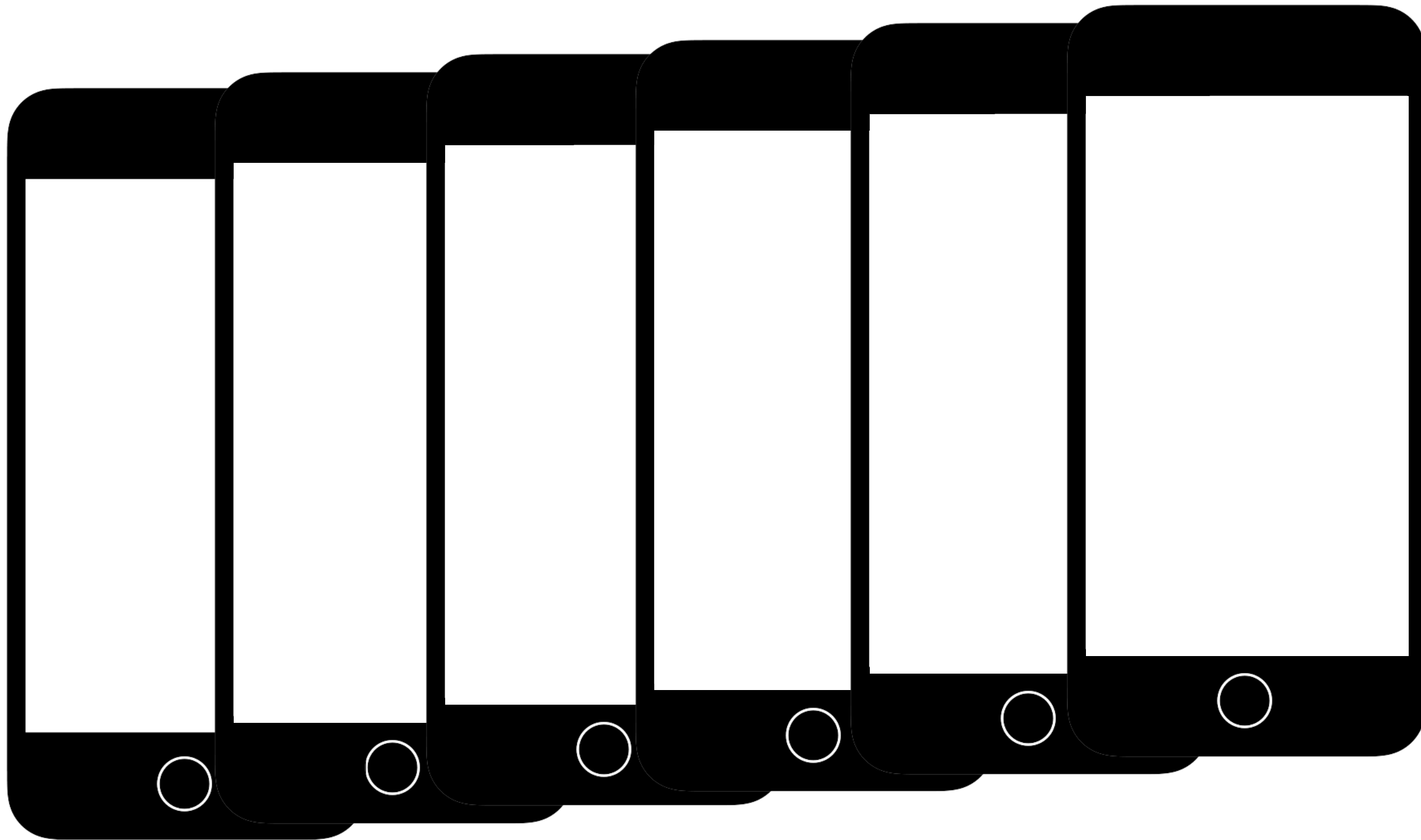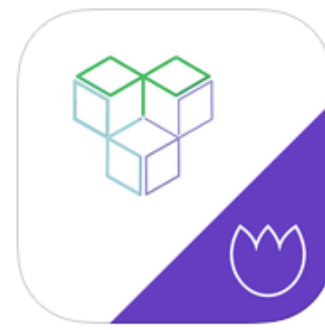
New Open-AD Drug Discovery Center

SGC and collaborators to create open research tools for Alzheimer's disease research funded by National Institute on Aging
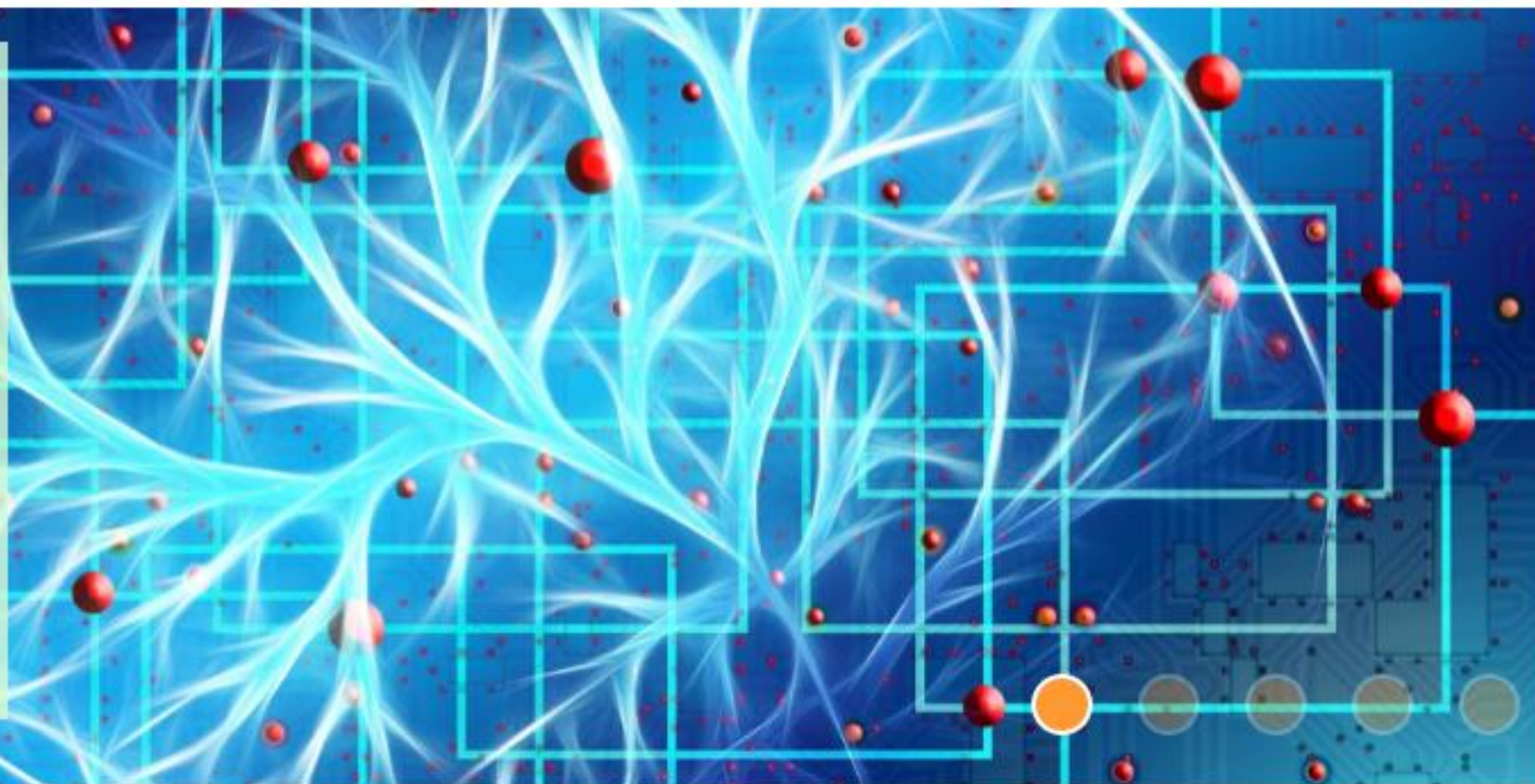
field without built-in inertia

mPower

mPower helps decipher Parkinson's disease.

The variability in Parkinson's disease symptoms has left many questions unanswered. So the University of Rochester and Sage Bionetworks created the mPower app to precisely measure data such as dexterity, balance, memory, and gait. This information could help researchers better understand how various symptoms are connected to Parkinson's disease. In turn, participants could start to recognize their own signs and symptoms.

first 6 months

16,585 participants consented

14,684 participants enrolled

9,520 agreed to 'share broadly'

1,087 self reported PD diagnosis

Sage BIONETWORKS

Robert Wood Johnson Foundation

ParkinsonNet

THE MICHAEL J. FOX FOUNDATION FOR PARKINSON'S RESEARCH

dimensionality of lived experience

Traditional Measures	First-order Features
Number of Taps	Number of taps, Mean tapping interval, Median tapping interval, Minimum tapping interval, maximum tapping interval, Standard deviation of tapping interval, Kurtosis of tapping interval, Interquartile range of tapping interval, Interquartile range of right button X, Range right button X ,Standard deviation right button X, Interquartile range of left button X, Range left button X ,Standard deviation left button X, Interquartile range of right button Y, Range right button Y ,Standard deviation right button Y, Interquartile range of left button Y, Range left button Y ,Standard deviation left button Y, Correlation X and Y, Skew tapping interval, No-button tapping frequency

invisible impacts made visible
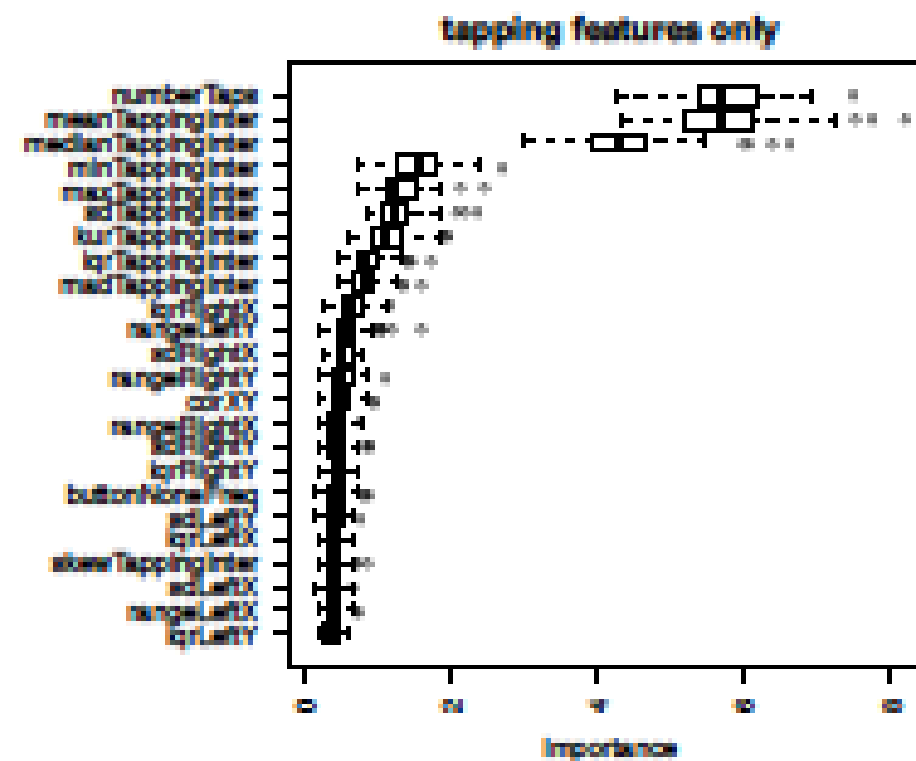
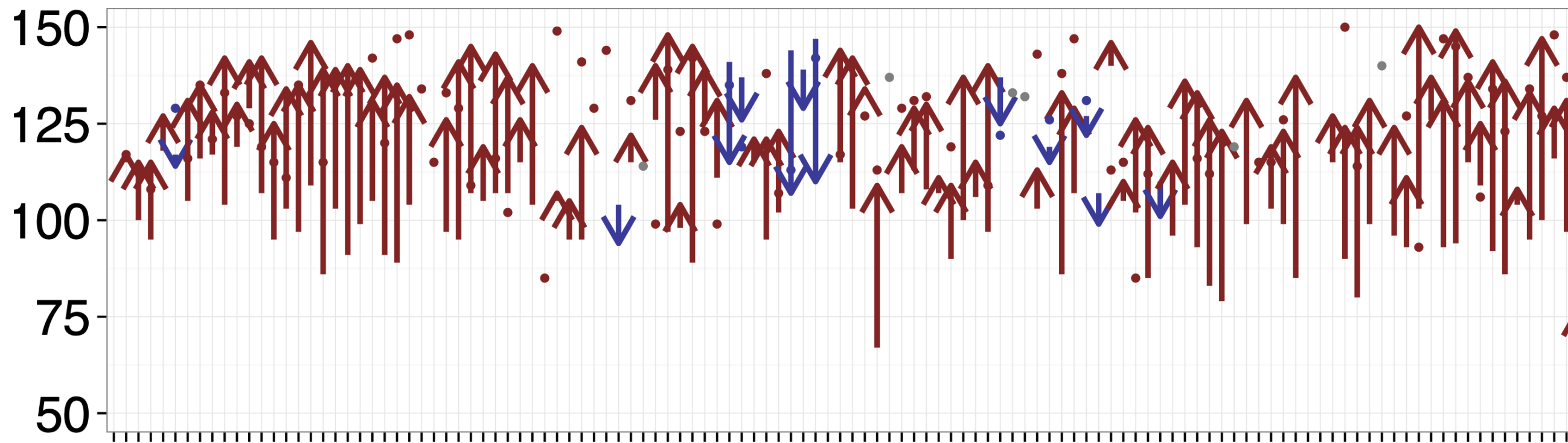Number of Taps
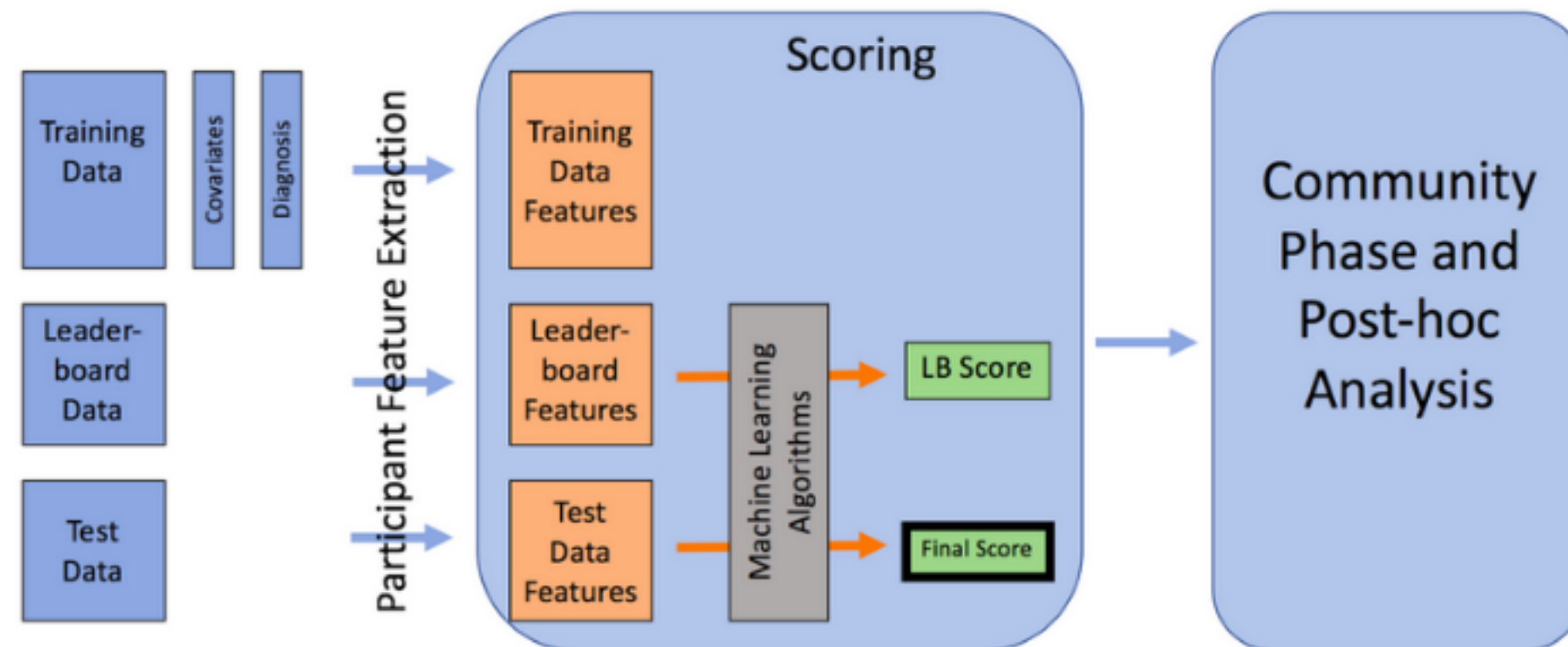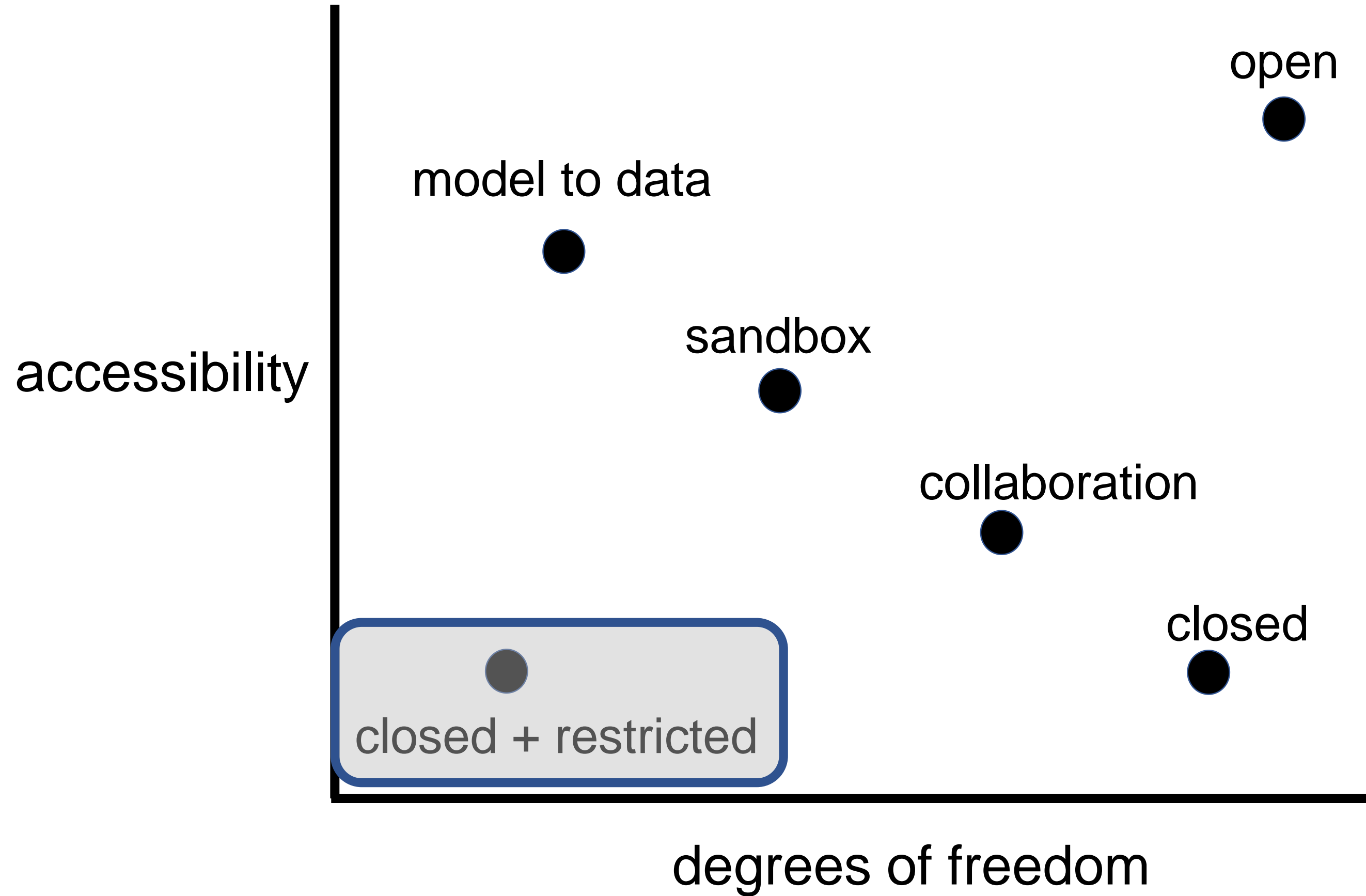Mean Tapping Interval
Median Tapping Interval
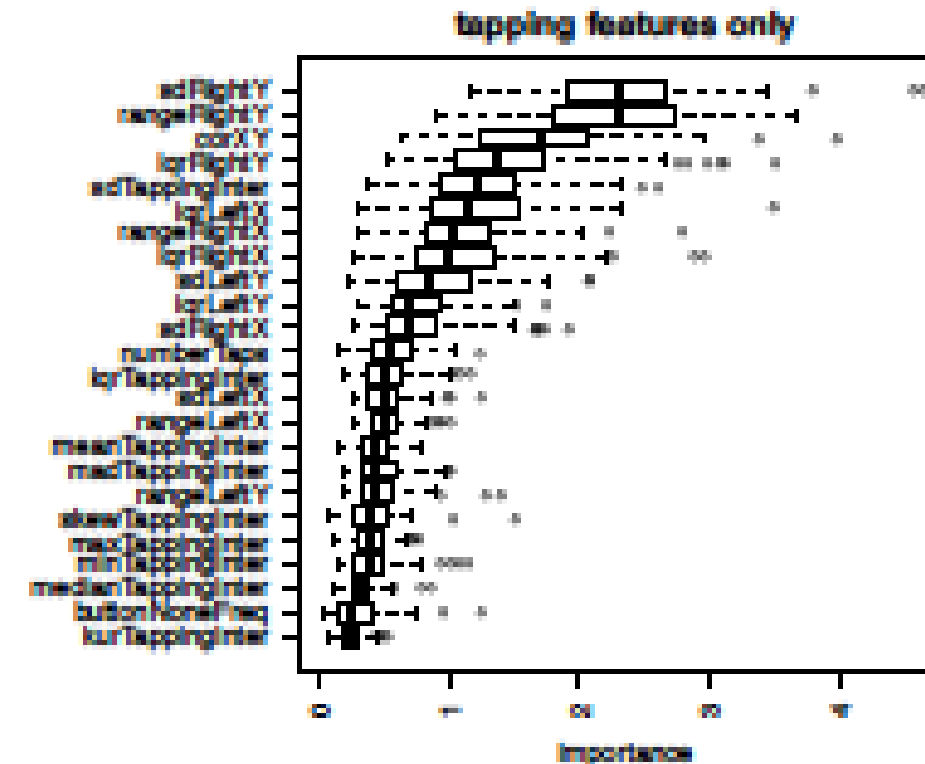
Standard Deviation R Y
Range Right Y
Correlation X Y

the reality of personal health

benchmark as peer review

accessibility

open

model to data

sandbox

collaboration

closed

closed + restricted

degrees of freedom

Why a DNA data breach is much worse than a credit card leak

You can't change your DNA

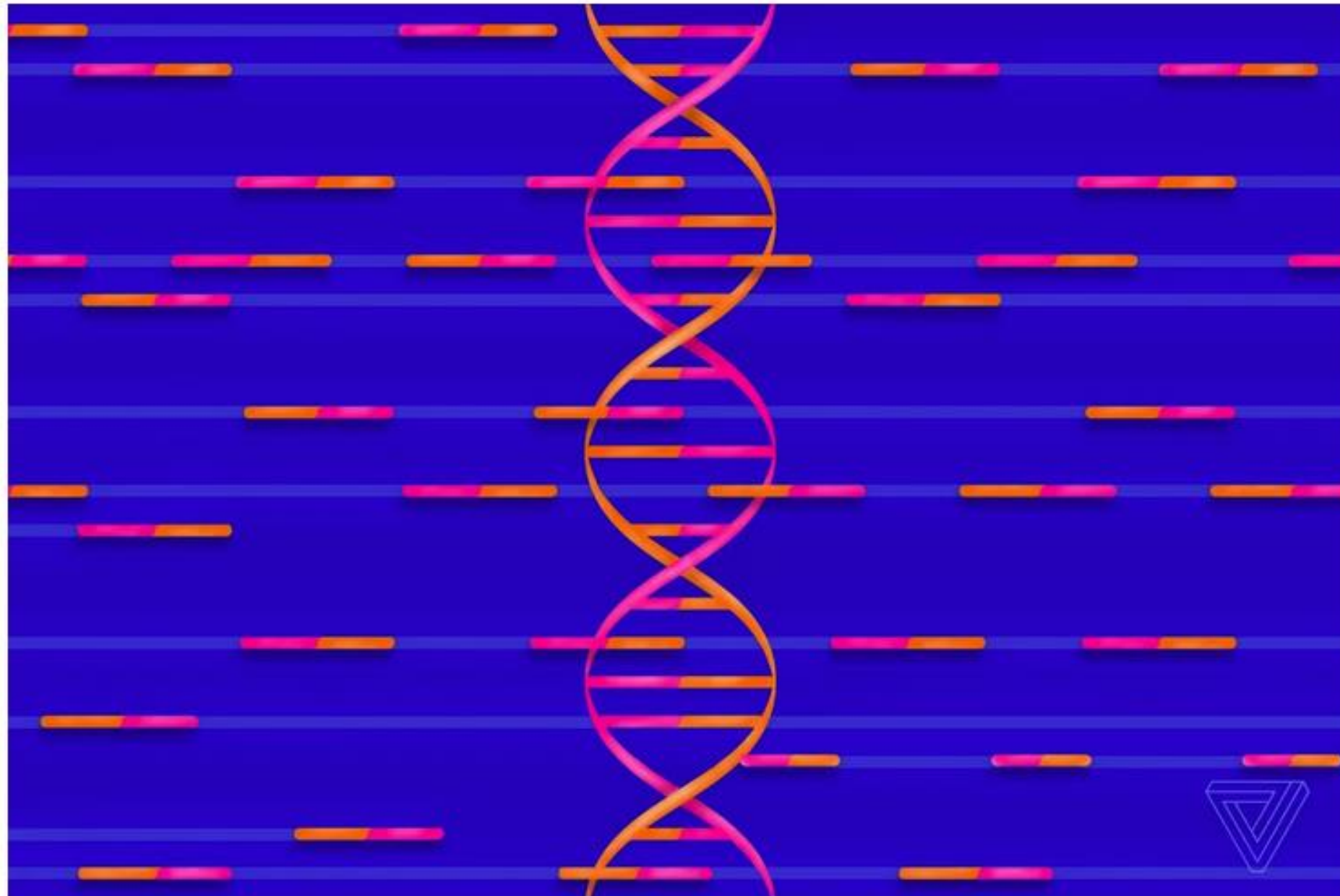By Angela Chen | @chengela | Jun 6, 2018, 3:54pm EDT

f 🐦 ↗ SHARE

Illustration by Alex Castro / The Verge

Genomic Justice for Native Americans: Impact of the Havasupai Case on Genetic Research

Nanibaa' A. Garrison[1]

[1]Center for Integration of Research on Genetics & Ethics, Stanford Center for Biomedical Ethics, Stanford, CA, USA

Abstract

In 2004, the Havasupai Tribe filed a lawsuit against the Arizona Board of Regents and Arizona State University (ASU) researchers upon discovering their DNA samples, initially collected for genetic studies on type 2 diabetes, had been used in several other genetic studies. The lawsuit reached a settlement in April 2010 that included monetary compensation and return of DNA samples to the Havasupai but left no legal precedent for researchers. Through semistructured interviews, institutional review board (IRB) chairs and human genetics researchers at US research institutions revealed their perspectives on the Havasupai lawsuit. For interviewees, the suit drew attention to indigenous concerns over genetic studies and increased their awareness of indigenous views. However, interviewees perceived no direct impact from the Havasupai case on their work; if they did, it was the perceived need to safeguard themselves by obtaining broad consent or shying away from research with indigenous communities altogether, raising important questions of justice for indigenous and minority participants. If researchers and IRBs do not change their practices in light of this case, these populations will likely continue to be excluded from a majority of research studies and left with less access to resources and potential benefit from genetic research participation.

Keywords

Open Peer Commentaries

Check for updates

The Unbearable Requirement of Informed Consent

Ellen Wright Clayton, Vanderbilt University Medical Center

In the spirit of full disclosure, I have been a member of the Delphi panels discussed in this article (Beskow and Weinfurt 2019) since their inception and was one of the people who was recently interviewed for this study (not that I know which, if any, of the published comments were mine). I have thought about informed consent for a long time. I took Jay Katz's year -long seminar on that topic during my last year in law school in 1978–1979 and was a research assistant for his book *The Silent World of Doctor and Patient* (Katz 2002). I have learned through experience that the kind of shared decision making he advocated is extremely difficult to achieve in both the clinic and research (Clayton 2006).

Nonetheless, reading Beskow and Weinfurt's article in this issue has given me pause. Despite the authors' best efforts in a prior study to create a truly effective consent process, it is not surprising that some of their participants could not demonstrate total comprehension on a quiz even after review and retesting (Beskow et al. 2015). After all, their subject population was representative of the U.S. population at large and hence was highly diverse. Most people from all walks of life had trouble getting a perfect score on their first try, and participants with low education, with low income, who lacked Internet access, and who were ethnic or racial minorities had more trouble. But these issues are complicated— even the highly educated participants in the Personal Genome Project have had a hard time passing that project's test (Angrist 2009).

No, what was disquieting for me was reading the enormous diversity of the Delphi panelists' responses— our responses—to the finding that many people could not

thereby acknowledging them as potential participants. Maybe acceding to apparently poorly informed decisions to participate honors autonomy. Maybe we don't see participating in a biobank as all that risky, especially in light of other safeguards, so it isn't all that important to understand what is involved. Intriguingly, many of us would have allowed people who lacked complete understanding to enroll in Phase III oncology trials, a more risky endeavor. Maybe broad-based participation even by those who are less than fully informed is necessary for all of society to benefit from the fruits of research.

One might argue that the Delphi process and telephone interviews did not capture the full nuance of our thoughts, and that may be true to some extent. But one thing is quite clear from this study—for most of these respondents, truly informed decision making is not always ethically required for research participation. This finding stands in stark contrast to the literature over the last 50+ years that points to the pivotal role of informed consent. To make matters worse, the significantly increased disclosure requirements embodied in the recent changes to the Common Rule may well decrease comprehension even further.

This study has done a great service by highlighting some crucial questions. It is time to think more clearly about what we hope to achieve with informed consent. If something other than knowledgeable, well-considered agreement is required to participate, the goals of the process need to be specified. Perhaps different words that are more descriptive of what is at stake would encourage more honest discussion. Acknowledging the individual and his or her potential contribution to the research pro-
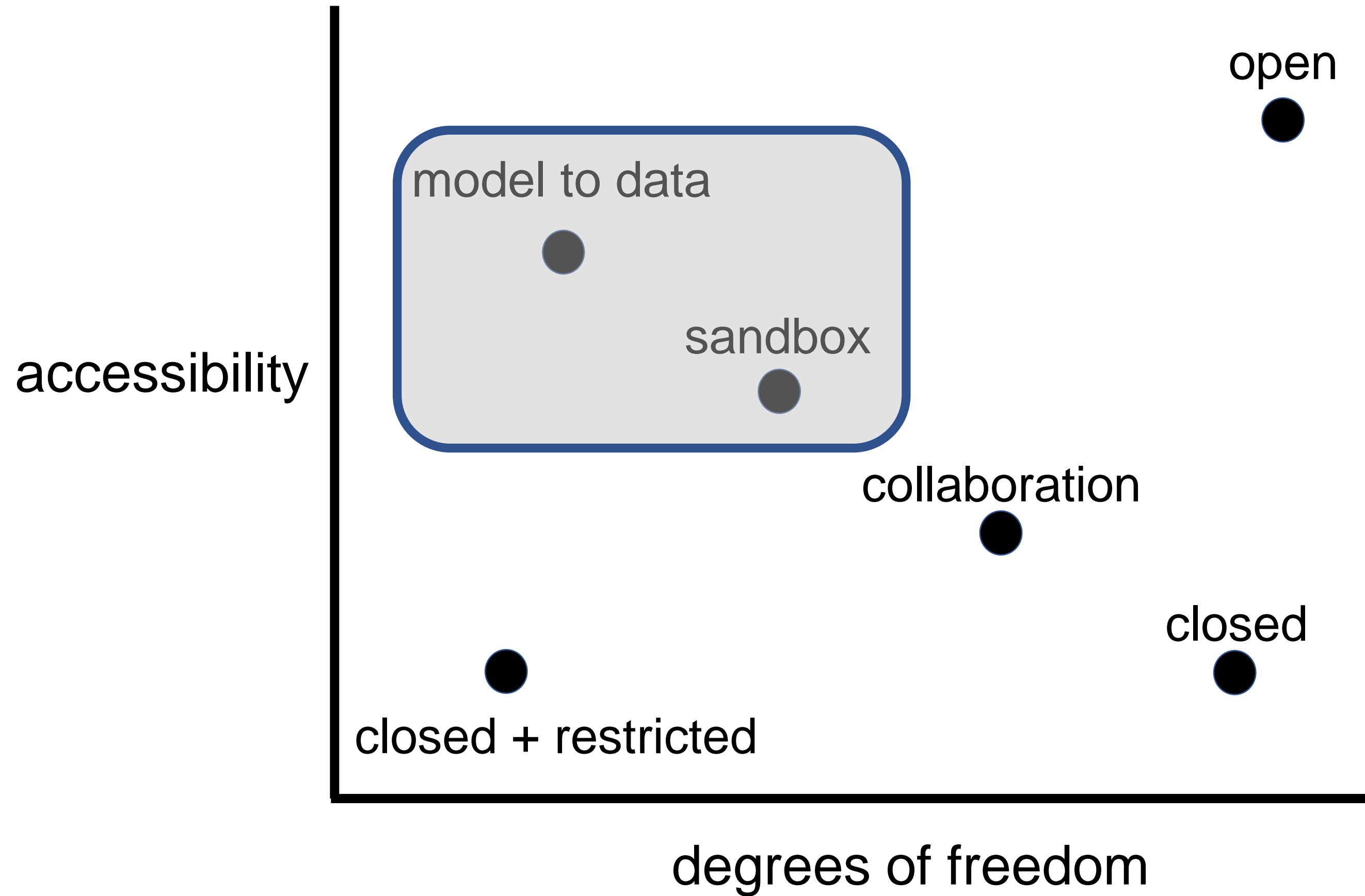
Welcome, *All of Us* Researchers.

The *All of Us* Research Program, part of the National Institutes of Health, is building one of the largest biomedical data resources of its kind. The *All of Us* Research Hub will store health data from one million or more diverse participants in the *All of Us* Research Program.

Here in the Research Hub, everyone can learn more about the types of data participants are providing and how approved researchers can use our data and tools to conduct studies that may speed up medical breakthroughs.

SEE THE DATA

Scroll Down

Data Snapshots

model-to-data for _necessarily closed_ data

Digital Mammography Dream challenge

Multiple Myeloma Dream Challenge

nature biotechnology

Correspondence | Published: 09 May 2018

Alternative models for sharing confidential biomedical data

Justin Guinney ✉ & Julio Saez-Rodriguez

Nature Biotechnology **36**, 391–392 (2018) | Download Citation ⬇

Cancer Research

search 🔍

Advanced Search

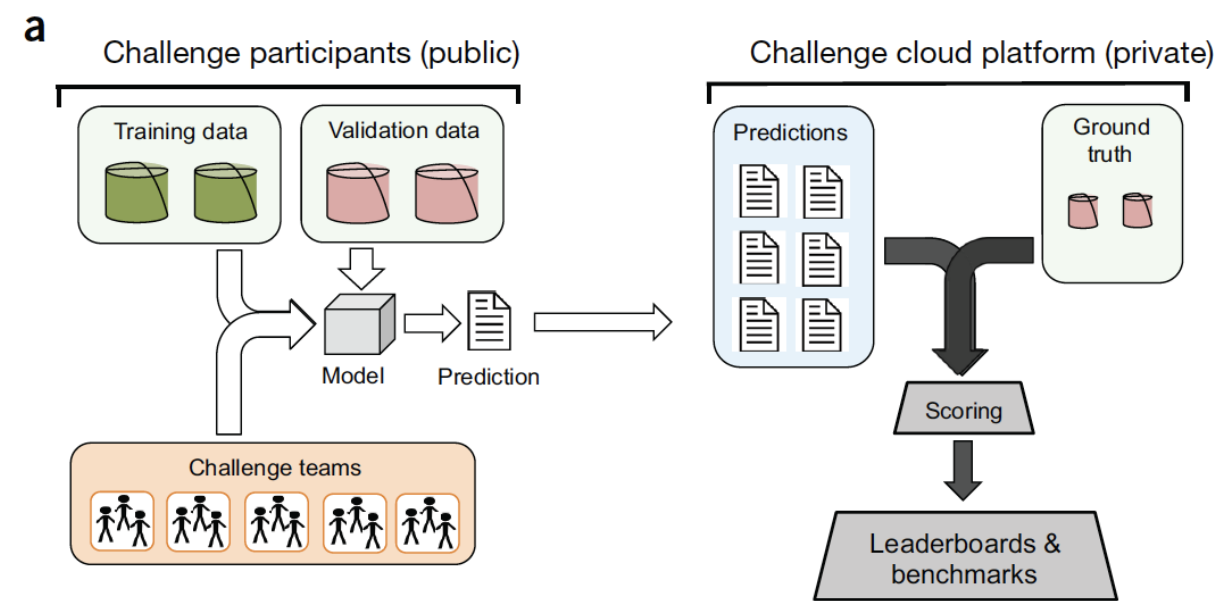Home   About   Articles   For Authors   Alerts   News

Clinical Research (Excluding Clinical Trials)

Abstract 4725: Multiple Myeloma DREAM Challenge: A crowd-sourced challenge to improve identification of high-risk patients
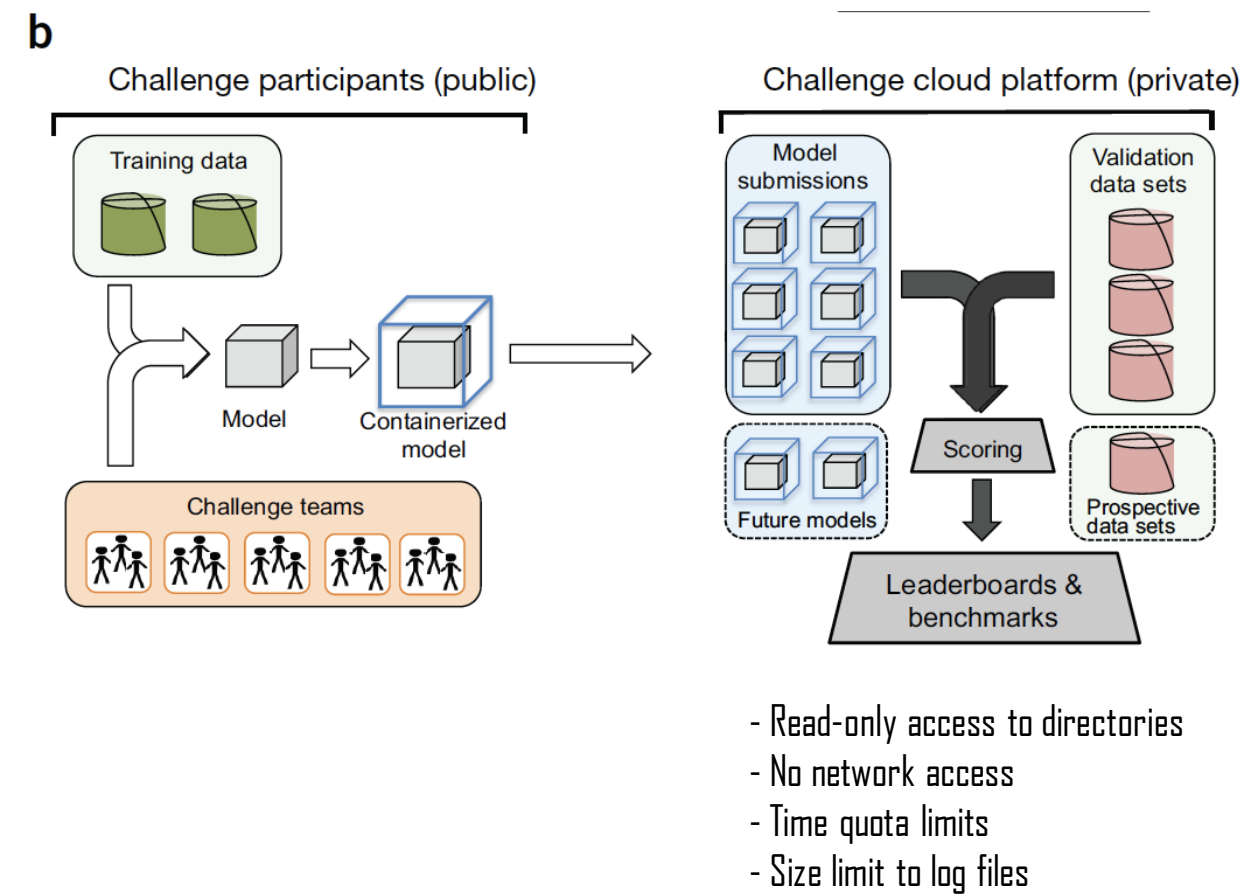
Michael Mason, Michael Amatangelo, Daniel Auclair, Doug Bassett, Hongyue Dai, Andrew Dervan, Erin Flynt, Hartmut Goldschmidt, Dirk Hose, Konstantinos Mavrommatis, Gareth Morgan, Nikhil Munshi, Alex Ratushny, Dan Rozelle, Mehmet Samur, Frank Schmitz, Ken Shain, Anjan Thakurta, Fadi Towfic, Matthew Trotter, Brian Walker, Brian S. White, Thomas Yu, and Justin Guinney

DOI: 10.1158/1538-7445.AM2017-4725 Published July 2017   Check for updates

Traditional Data to Modeler

Model to Data

a

Challenge participants (public)

Training data

Validation data

Model

Prediction

Challenge teams

Challenge cloud platform (private)

Predictions

Ground truth

Scoring

Leaderboards & benchmarks

b

Challenge participants (public)

Training data

Model

Containerized model

Challenge teams

Challenge cloud platform (private)

Model submissions

Validation data sets

Future models

Scoring

Prospective data sets

Leaderboards & benchmarks

- Read-only access to directories
- No network access
- Time quota limits
- Size limit to log files

Guinney and Saez-Rodriguez Nature Biotechnology volume 36, pages 391–392 (2018)

a.
more massive incrementalism.

b.
the stacks.

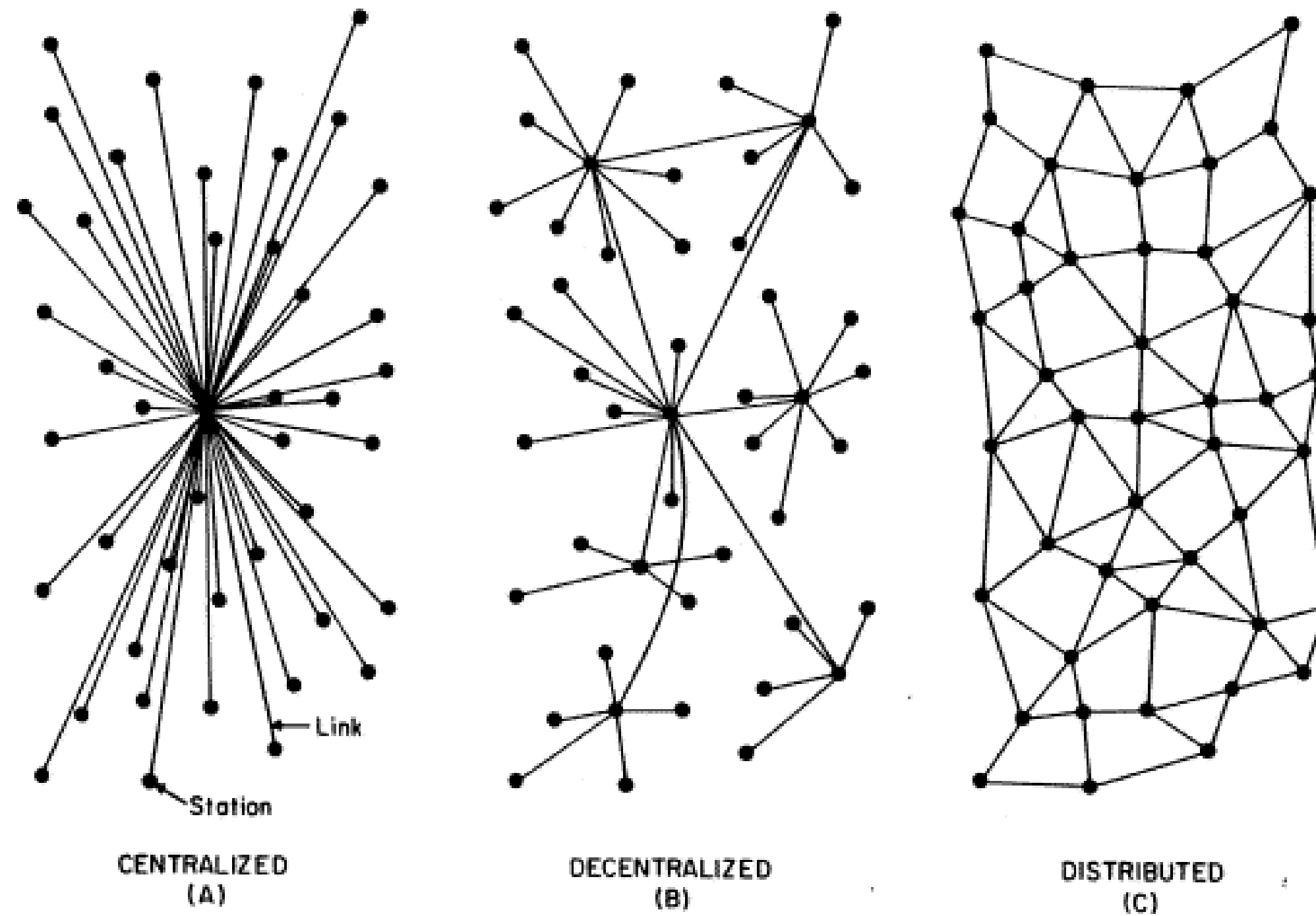CENTRALIZED
(A)

DECENTRALIZED
(B)

DISTRIBUTED
(C)

FIG. I — Centralized, Decentralized and Distributed Networks
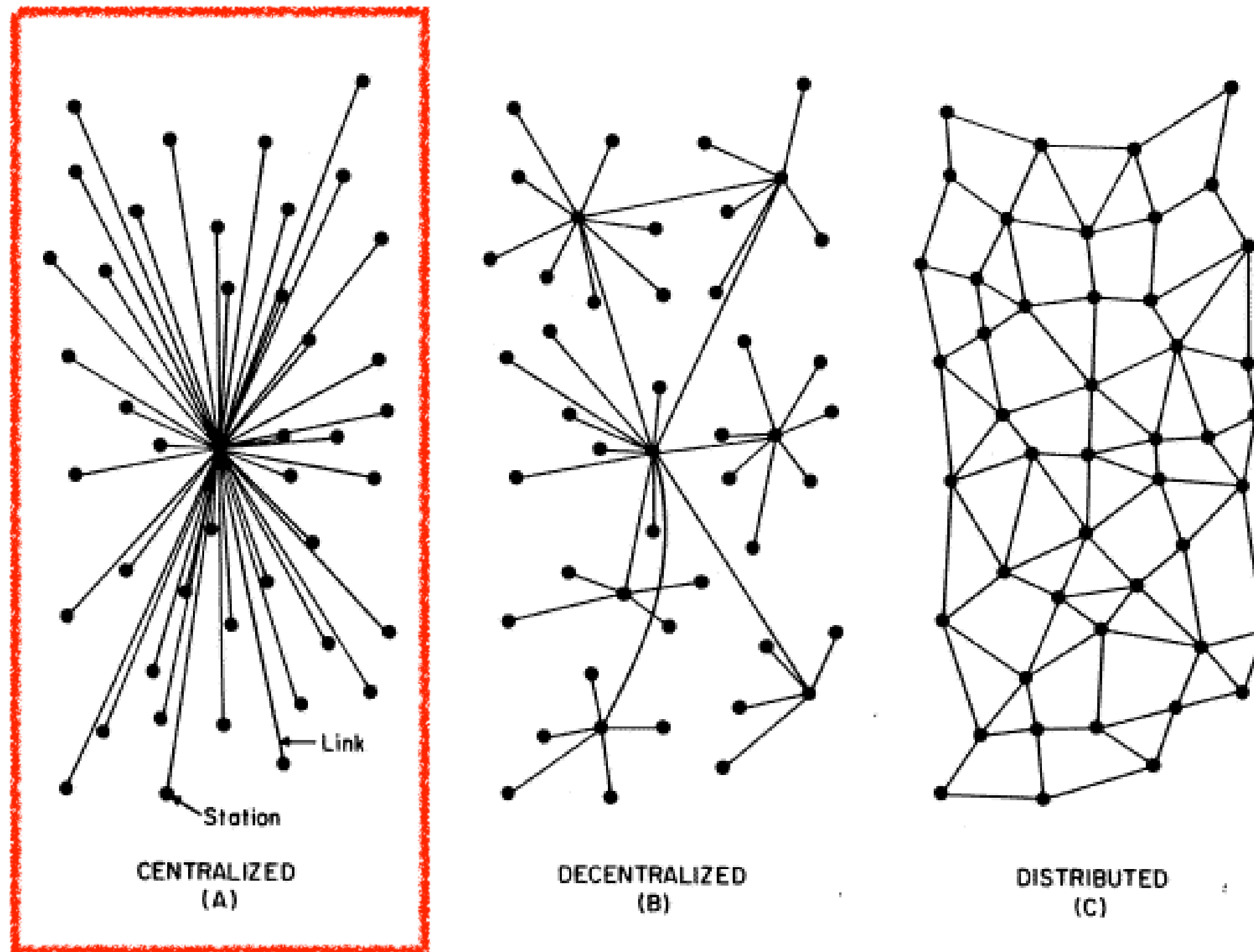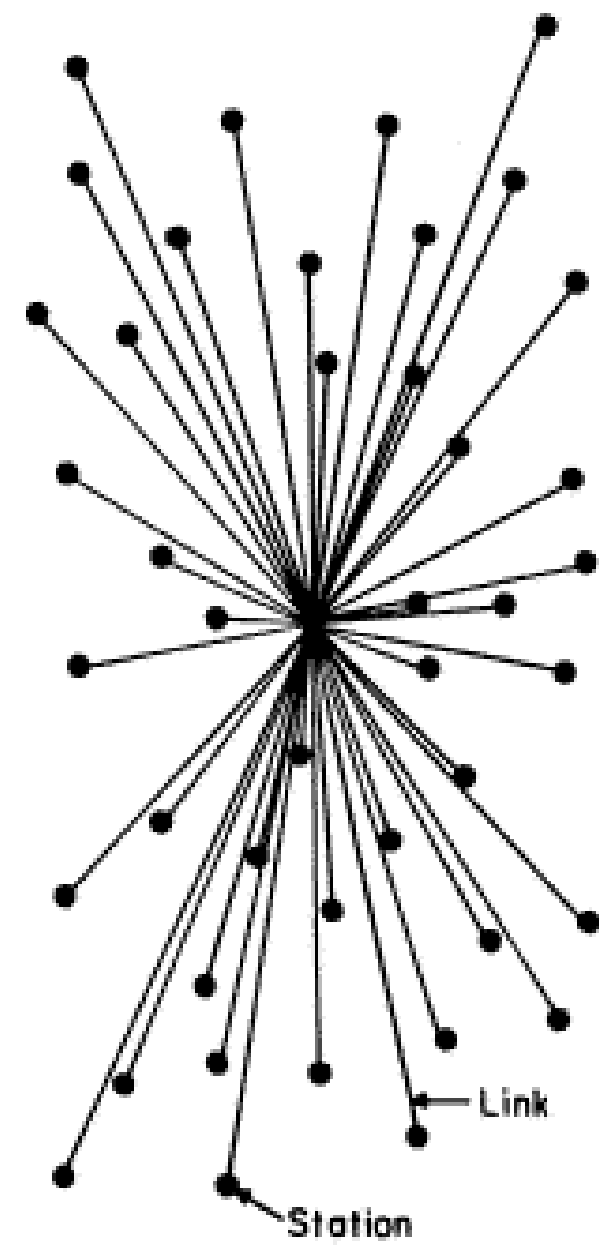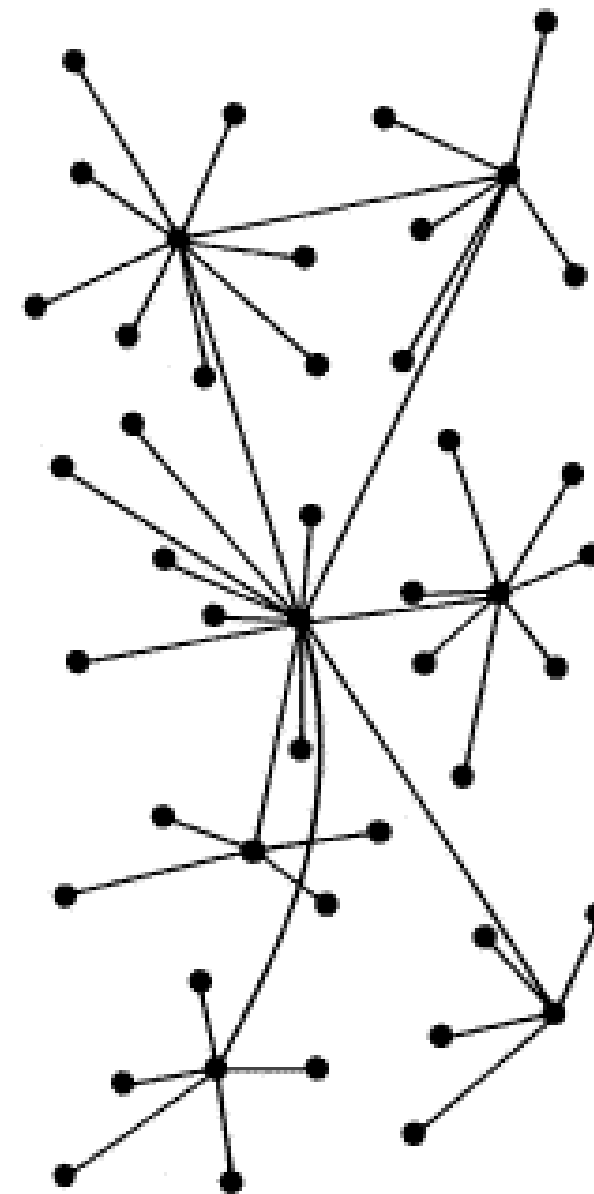
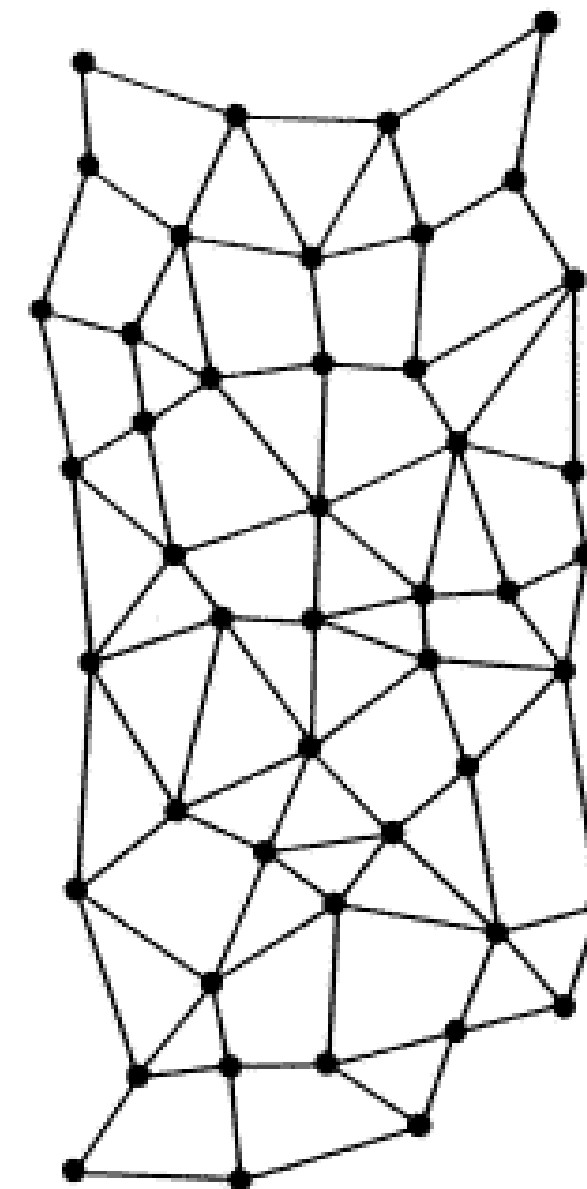FIG. I — Centralized, Decentralized and Distributed Networks

c.
institutional change

Link →

Station →

CENTRALIZED
(A)

DECENTRALIZED
(B)

DISTRIBUTED
(C)

FIG. I — Centralized, Decentralized and Distributed Networks

better science, together

thank you

@wilbanks
john.wilbanks@sagebase.org