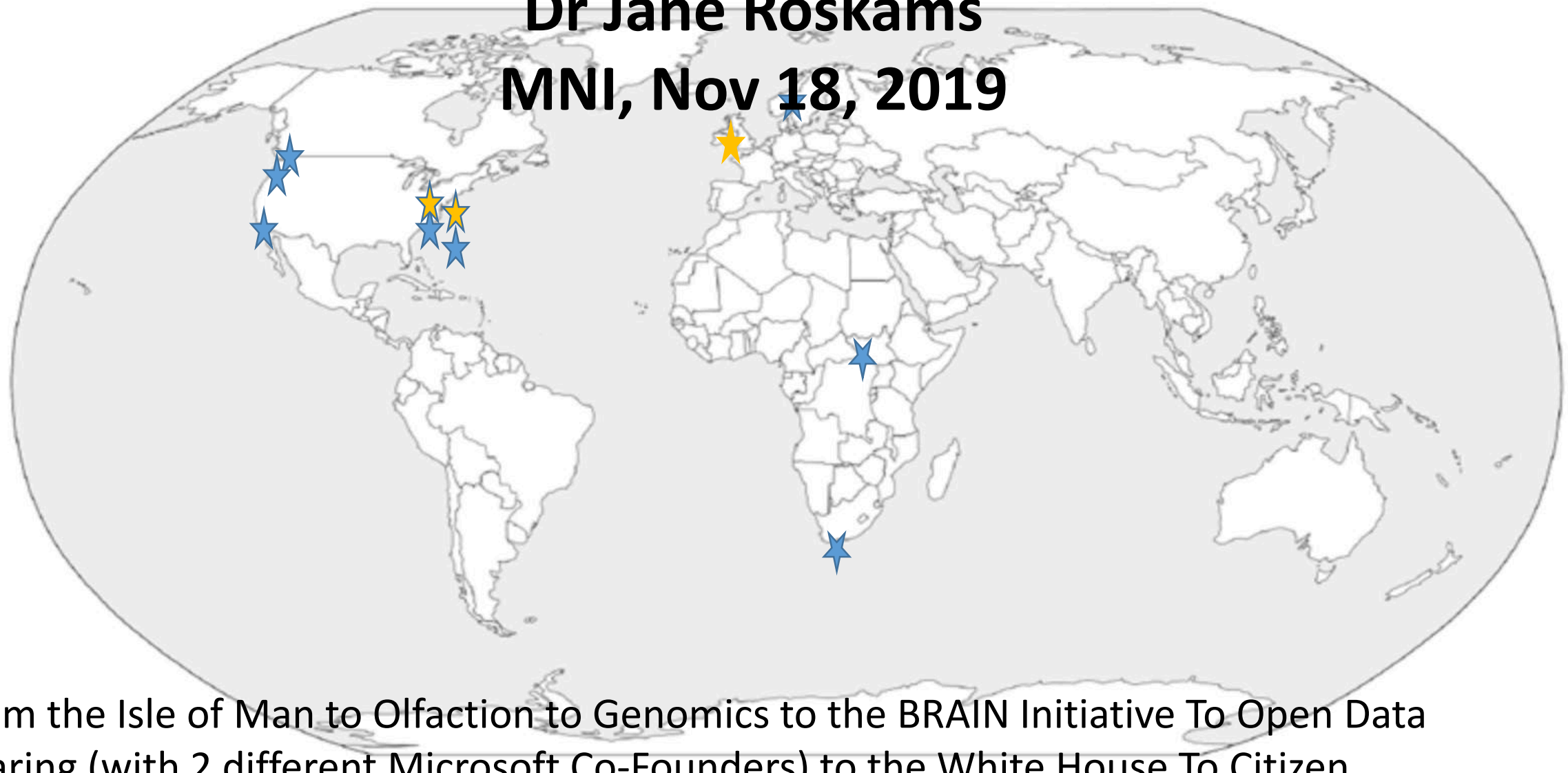
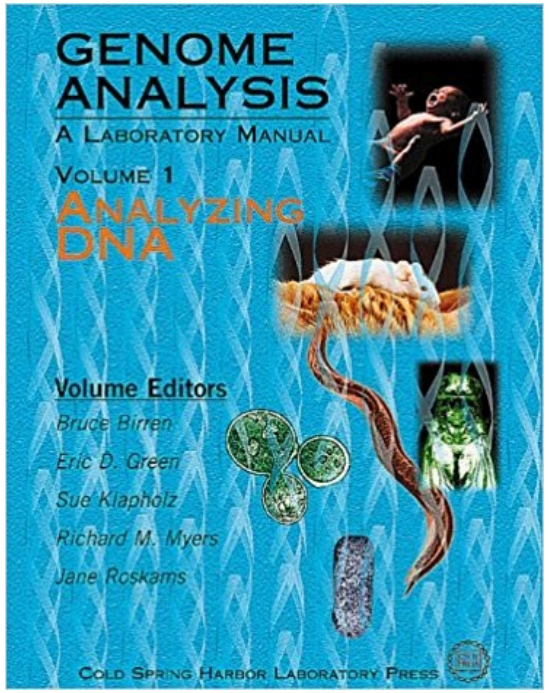
# My Rather Unexpected Open Science Journey

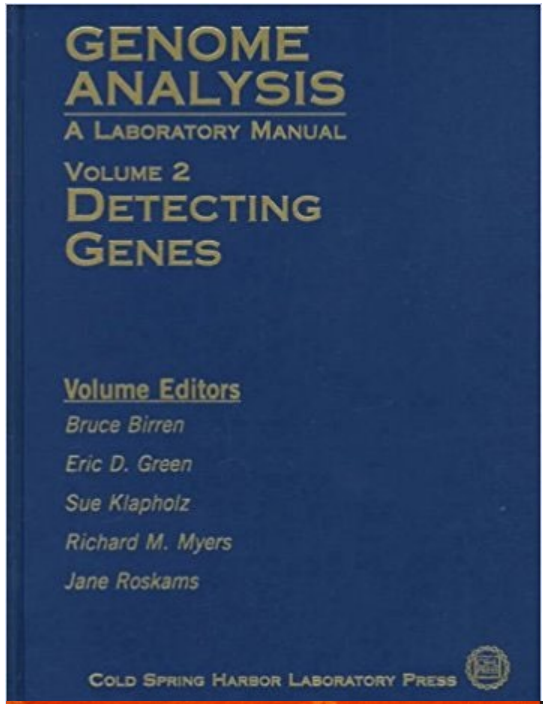
Dr Jane Roskams

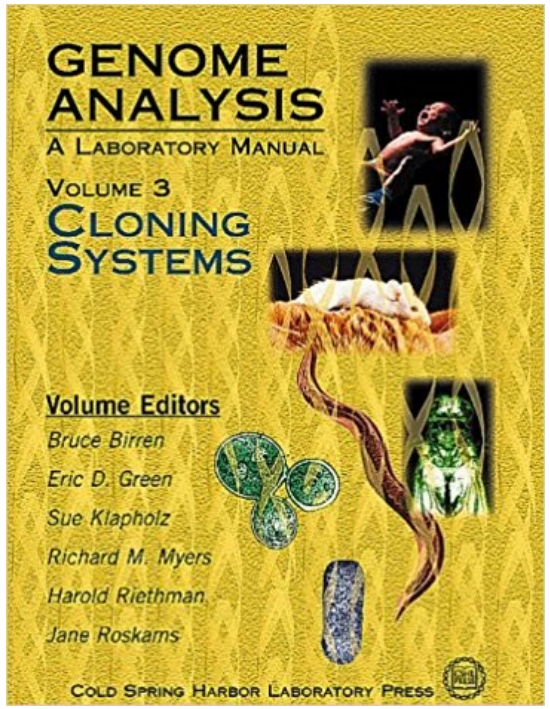
MNI, Nov 18, 2019

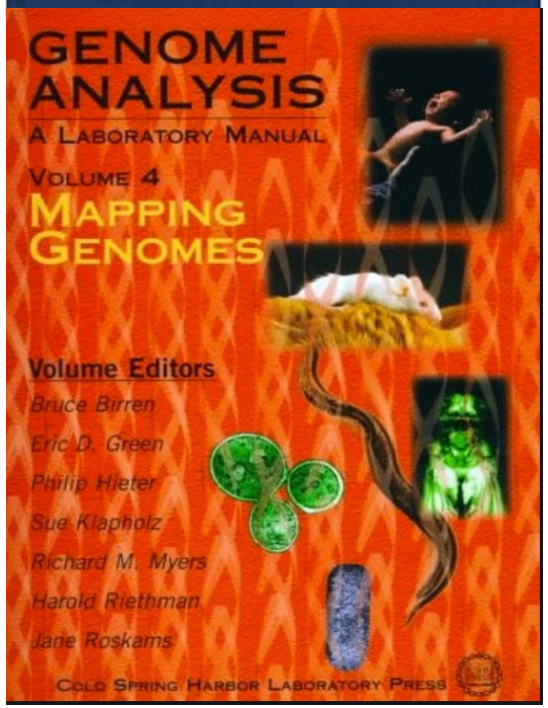


From the Isle of Man to Olfaction to Genomics to the BRAIN Initiative To Open Data Sharing (with 2 different Microsoft Co-Founders) to the White House To Citizen NeuroScience Game Design, and International Brain Data Analytics and Training

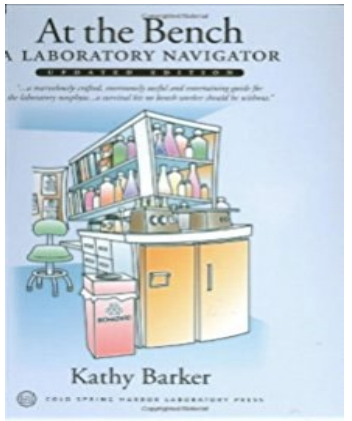


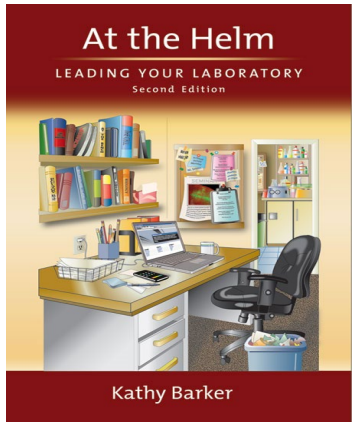



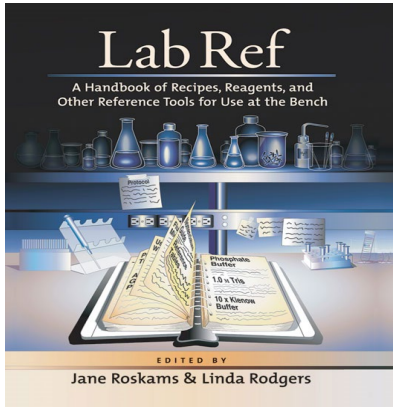




# 1995/6: My Beginning in Open – A Journey into the Human Genome (CSHL via NCBI)




- *"Why don't we ever show what it looks like when we screw up?"*
-
- *-Life Lessons in epigenetics, and Game-Changing Collaboration across Genomics communities ->Human Genome Project*

# Understanding and Unlocking Human/Cell Potential

OPEN DNA/Chromatin and its Many Impactors…

# What happens in Bermuda (1996-7)….

Sirs John Sulston and Sydney Brenner

1st International Strategy Meeting on Human Genome Sequencing

Hamilton Princess Hotel Bermuda

February 25th-28th, 1996

HGP-funded sequence assemblies 1 Kb or greater released to public domain ASAP, ideally daily.

"In order to encourage further research and development and to maximise its benefit to society."[a]

2nd Bermuda Meeting

February 25th-March 2nd, 1997

1996 Principles re-affirmed.

Scientists and funding agencies asked to ensure that daily release was occurring.[b]

3rd Bermuda Meeting

February 27th-March 1st, 1998

2 Kb assemblies are the new release trigger.

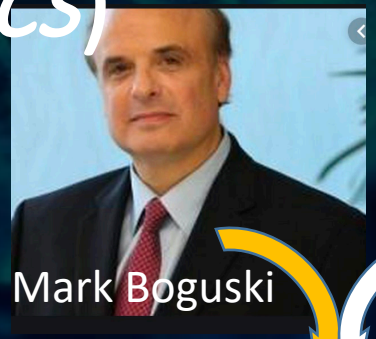
Bermuda Principles extended to mouse.[c]

Francis Collins and J Craig Venter

- Quality Standards
- Sequence Submission and Annotation
- Sequence Claims (IP) and Etiquette
- Increased funding of open tools for data exchange and technology development
- Coordination network/governance

# 1996-97: The Exponential Growth of NCBI (*and birth of my education in informatics*)

- Sequence deposited from all over the world, and coordinated through NCBI
- Huge amount of brain data coming off the pipeline
- dbEST, xrefdB, etc
- New tools for alignment
- The birth of NLP, genome-style
- Computational modeling across genomes
- And MUCH more to take to my new UBC lab!



Stephen Friend

ROSETTA INPHARMATICS

ALLEN INSTITUTE for BRAIN SCIENCE
Fueling Discovery

SageBionetworks

# JaneLab Open for Business - at UBC, 1997-2014)

- Lab pioneered the development of in vitro, lesion and ko/ transgenic approaches to understand olfactory system regeneration and degeneration mechanisms.
- Openly Published active protocols on our lab website; shared mice
- Developed embryo-derived “stem cell” lines - openly shared with dozens of labs worldwide
- Open collaboration with Don Nicholson, Merck-Frosst on neuronal cell death inhibitors
- Welcomed trainees from many collaborator’s labs (across Canada, US, UK, S Africa, Germany and Holland) to openly train in lesion, stem cell culture techniques we developed

- Jumped collaboratively (with W Tetzlaff, UBC; others) into the spinal cord injury world
- Co-funded for research on regeneration, remyelination and stem cell transplants with colleagues at UBC, UCSD, U of Utah, Miami Project, UCI, Univ of Cambridge, UCL, ISB
- Invited to participate in NIH SCI replication studies (2004-2008)

# WHY OPEN? 2004-2014: THE ENORMOUS CHALLENGE OF Spinal Cord Injury REPRODUCIBILITY!

- Different injury models
- rat vs. mouse (vs. primate?)
- behavioral assessment

Transgenic mice-genetic variability across strains

- immune variables
- How to transplant ??
- Secondary Damage?
- What to assess?
- How?

ELSEVIER

Experimental Neurology

Volume 198, Issue 2, April 2006, Pages 483-499

A re-assessment of the consequences of delayed transplantation of olfactory lamina propria following complete spinal cord transection in rats

Oswald Steward [a], Kelli Sharp [a], Gowri Selvan [a], Anthony Hadden [a], Maura Hofstadter [a], Edmund Au [b], Jane Roskams [b]

Show more

https://doi.org/10.1016/j.expneurol.2005.12.034

Get rights and content

ELSEVIER

Experimental Neurology

Volume 233, Issue 2, February 2012, Pages 597-605

Editorial

Replication and reproducibility in spinal cord injury research

Oswald Steward [a, b, c, d], Phillip G. Popovich [e, f], W. Dalton Dietrich [g, h], Naomi Kleitman [i]

Show more

https://doi.org/10.1016/j.expneurol.2011.06.017

Get rights and content

JNeurosci
THE JOURNAL OF NEUROSCIENCE

An Official Journal of SfN SOCIETY for NEUROSCIENCE

1969–2019 CELEBRATING 50 YEARS

search

Advanced Search

Submit a Manuscript

HOME CONTENT ALERTS FOR AUTHORS EDITORIAL BOARD ABOUT

Previous Next

Articles, Development/Plasticity/Repair

Olfactory Ensheathing Cells Do Not Exhibit Unique Migratory or Axonal Growth-Promoting Properties after Spinal Cord Injury

Paul Lu, Hong Yang, Maya Culbertson, Lori Graham, A. Jane Roskams, and Mark H. Tuszynski

Journal of Neuroscience 25 October 2006, 26 (43) 11120-11130; DOI: https://doi.org/10.1523/JNEUROSCI.3264-06.2006

NIH-Funded – Me + collaborators = 2 “failures to replicate” (and new insight!)

# SCI is Highly Complex: Gazillions of tiny measures of biofunction


Adam Ferguson
UCSF

2008 Asilomar Regeneration Conference - Principal Component Analysis!

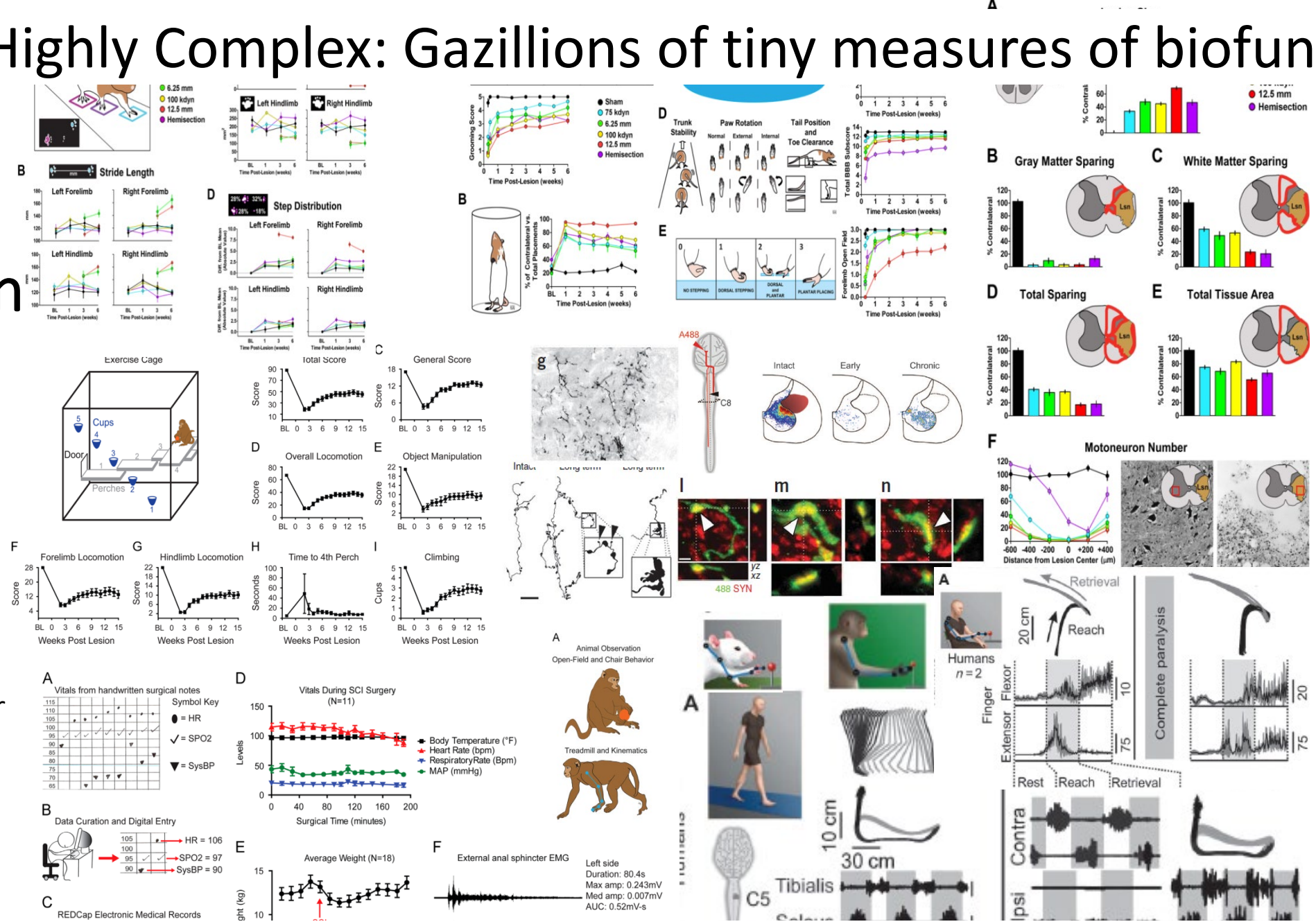



Figure Sources: Rosenzweig et al., *2010; 2019 Nat Neurosci;* Ferguson et al., 2013 *PloS One*; Nielson et al., 2014, *J Neurotrauma*; Nielson et al., 2015, *Brain Res.;* Friedli et al., 2015 *Science TM;* Rosenzweig et al., 2019 *Nat Medicine*

# WHERE DO ALL THESE DATA COME FROM?

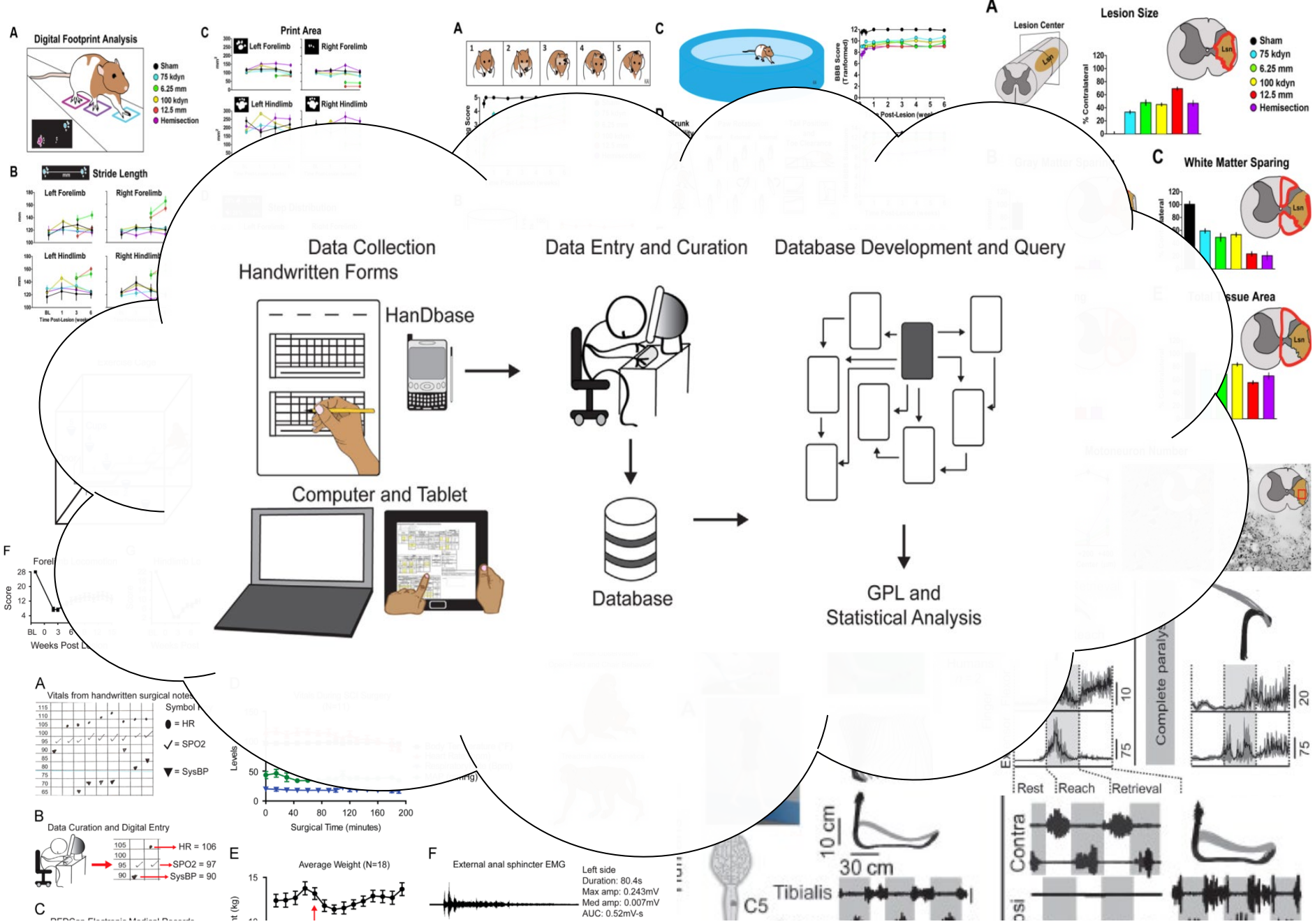



Figure Sources: Rosenzweig et al., *2010; 2019 Nat Neurosci;* Ferguson et al., 2013 *PloS One*; Nielson et al., 2014, *J Neurotrauma*; Nielson et al., 2015, *Brain Res.;* Friedli et al., 2015 *Science TM;* Rosenzweig et al., 2019 *Nat Medicine*

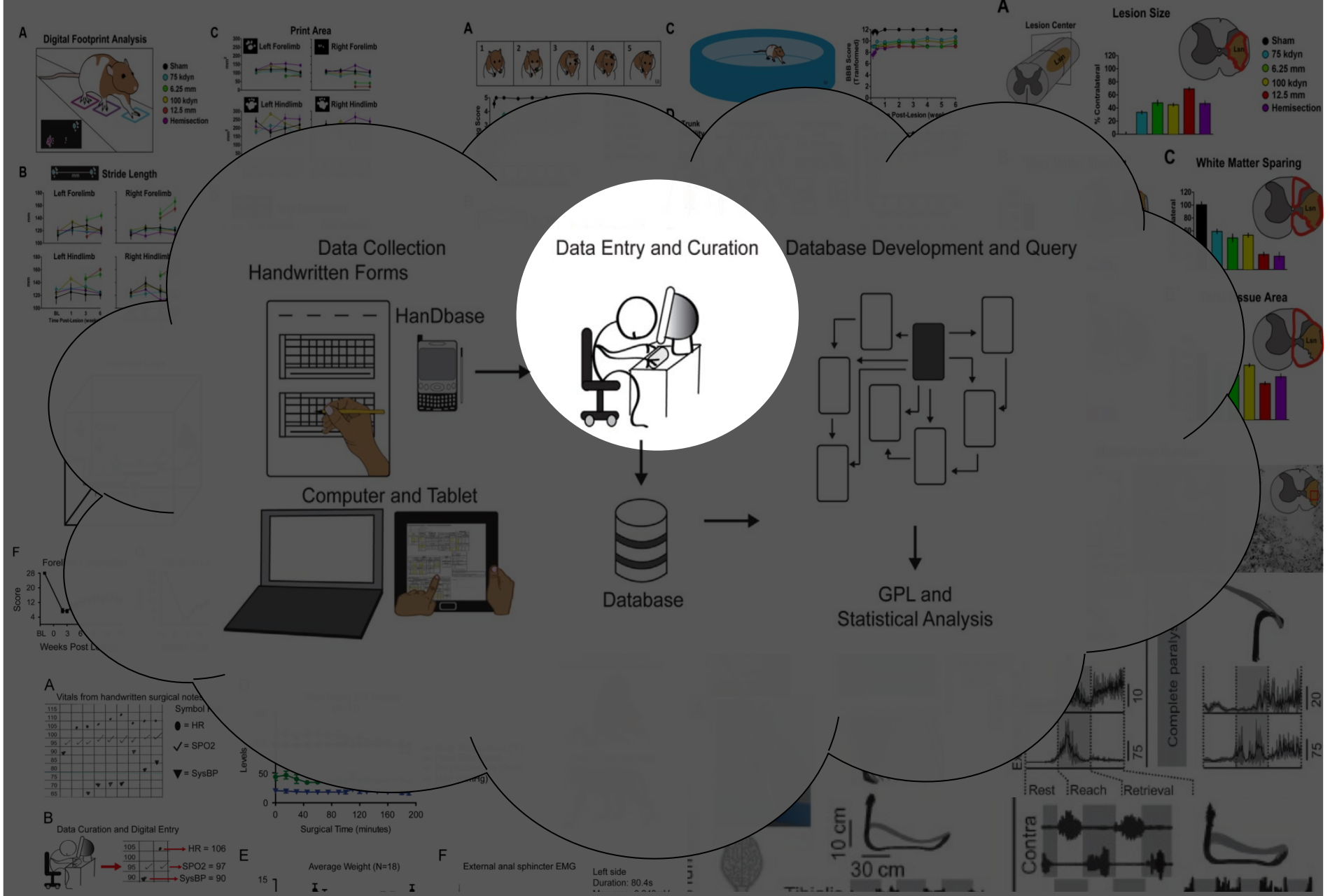


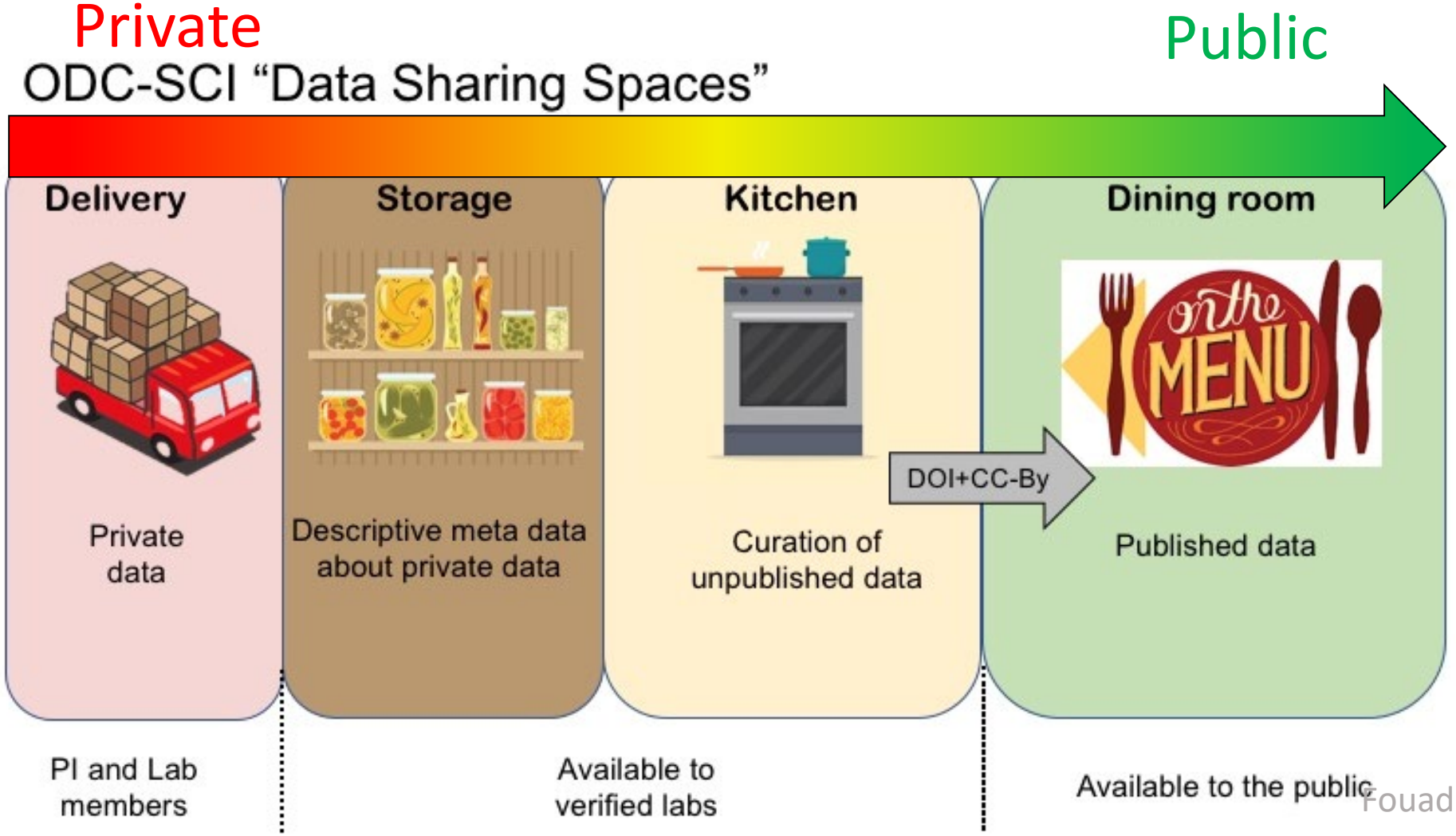

Figure Sources: Rosenzweig et al., *2010; 2019 Nat Neurosci;* Ferguson et al., 2013 *PloS One*; Nielson et al., 2014, *J Neurotrauma*; Nielson et al., 2015, *Brain Res.;* Friedli et al., 2015 *Science TM;* Rosenzweig et al., 2019 *Nat Medicine*

# 2016 - ?: The Spinal CORD *Community* – hoping to be a working model for basic-clinical FAIR Data Sharing, Standards and Multi-Modal Analysis

Spinal Cord Open Data Sharing (multiple International Foundations), Replication (NIH)

# 2019: Democratization of SCI Data Science – Community-Driven Open Data Sharing => Hope for Reproducibility

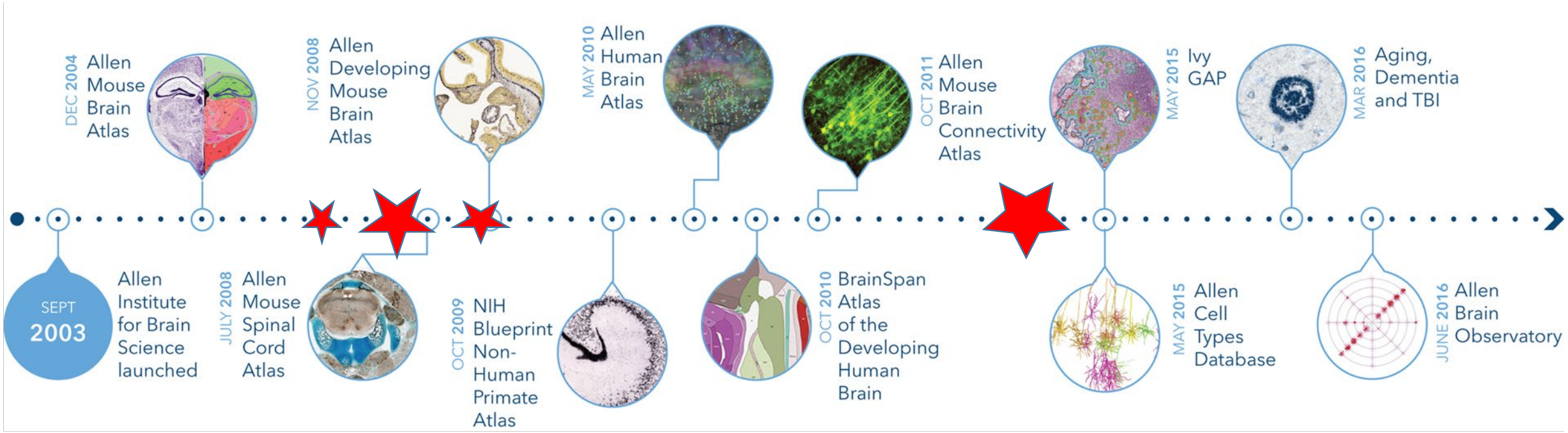
HTTP://ODC-SCI.ORG




Fouad et al., submitted

In 2014, I decided to jump off the academic lab merry-go-round and commit to the world of Open Science

# 2014 Destination Seattle: Pioneering Open Science & Data - The Allen Institute for Brain Science




- Since 2003, their product portfolio includes downloadable data repositories, software applications, reference standards and training toolkits – open to all

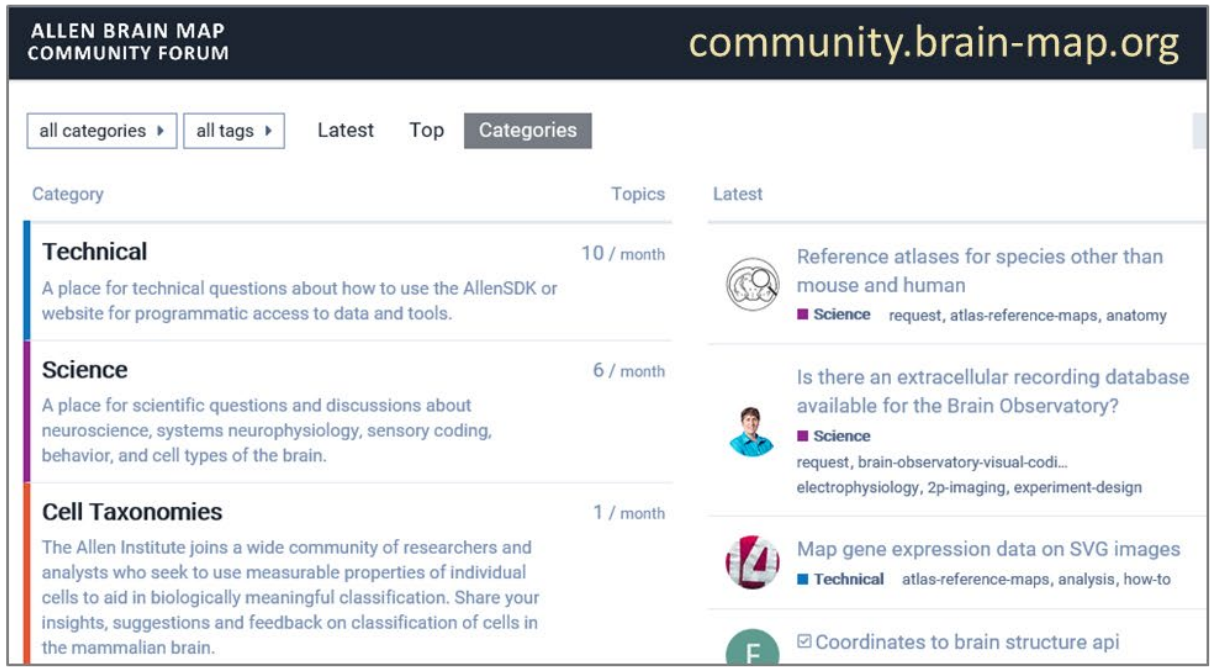
**2003-2012:** 50,000 visitors/month

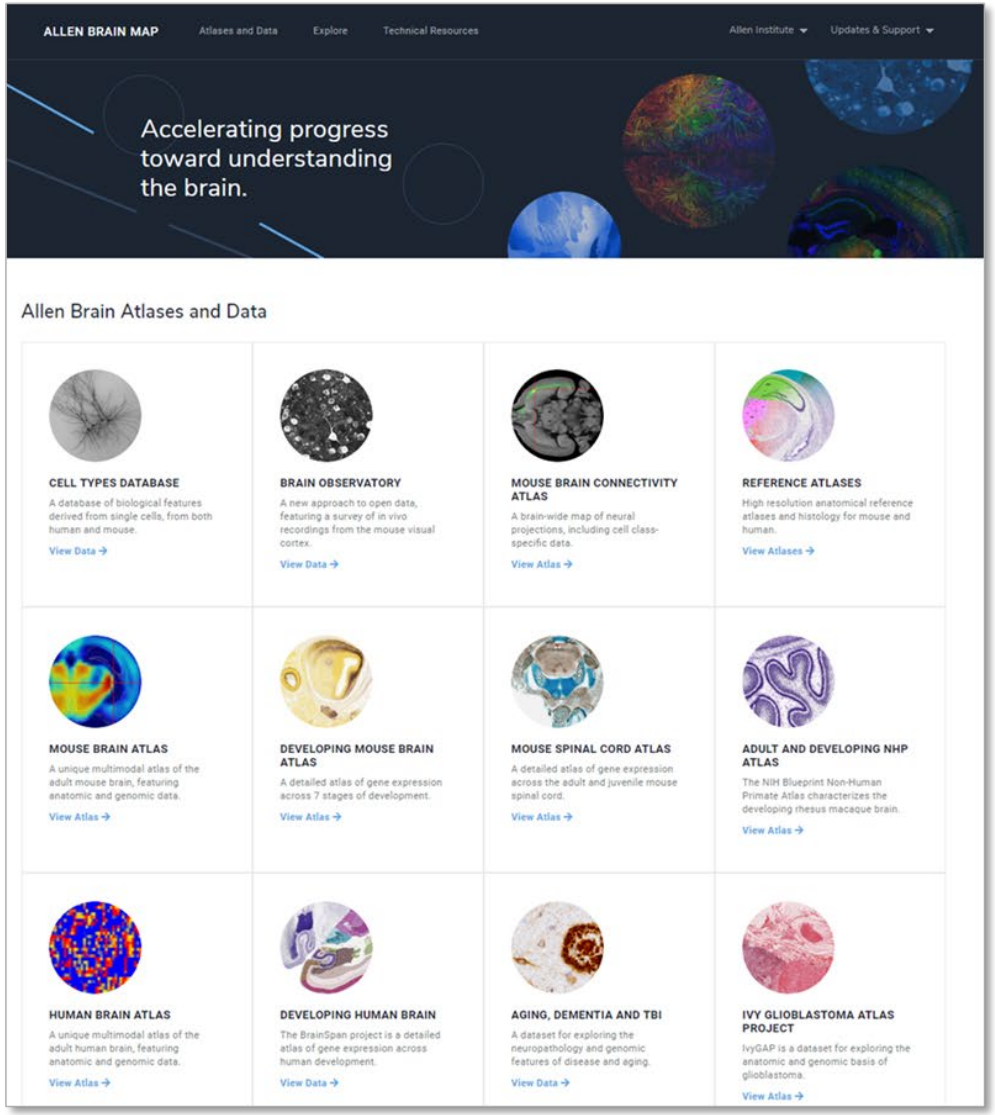
- Users from around the globe
- >**3 Petabytes** data generated
- >5000 engineered mice
- > 1 million microscope slides

**Over 35 million page visits through brain-map.org by 2019**

# Data & tools are available at brain-map.org

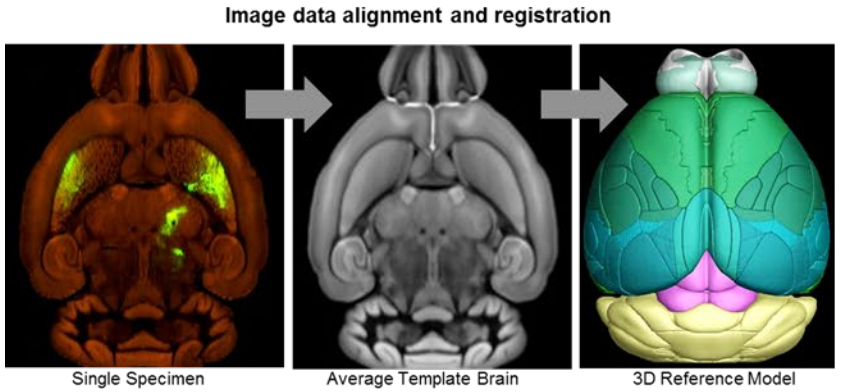
- Research tools and rich reference datasets covering human brain in health & disease, model organisms and standards for scientists
- Resources for brain, eye, spinal cord research
- Access to publications, laboratory resources, workshops & training toolkits



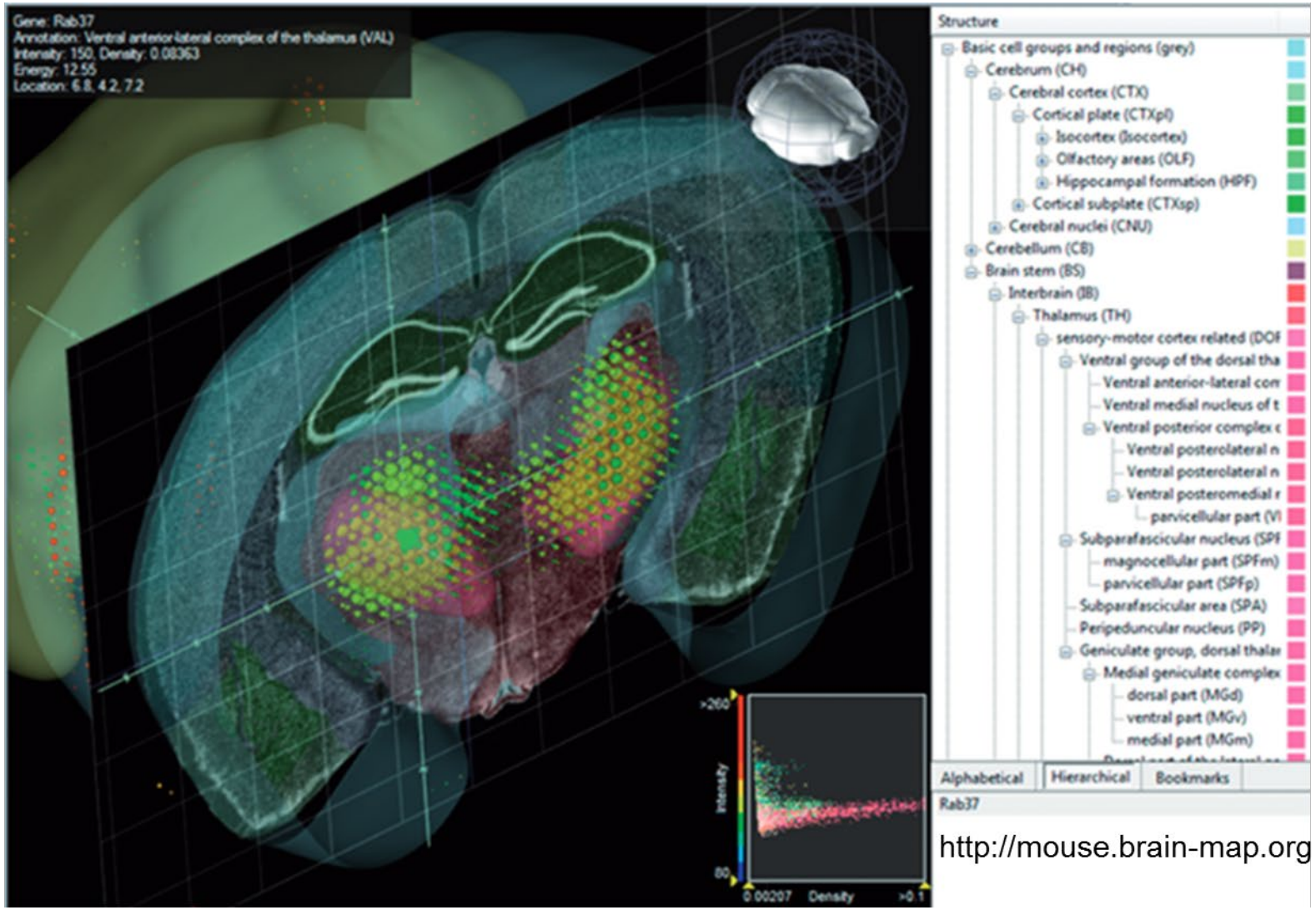



# Creation of standardized frameworks: Mouse brain reference maps

- The Allen Reference Atlas is now used to map multimodal data and cells to a common space

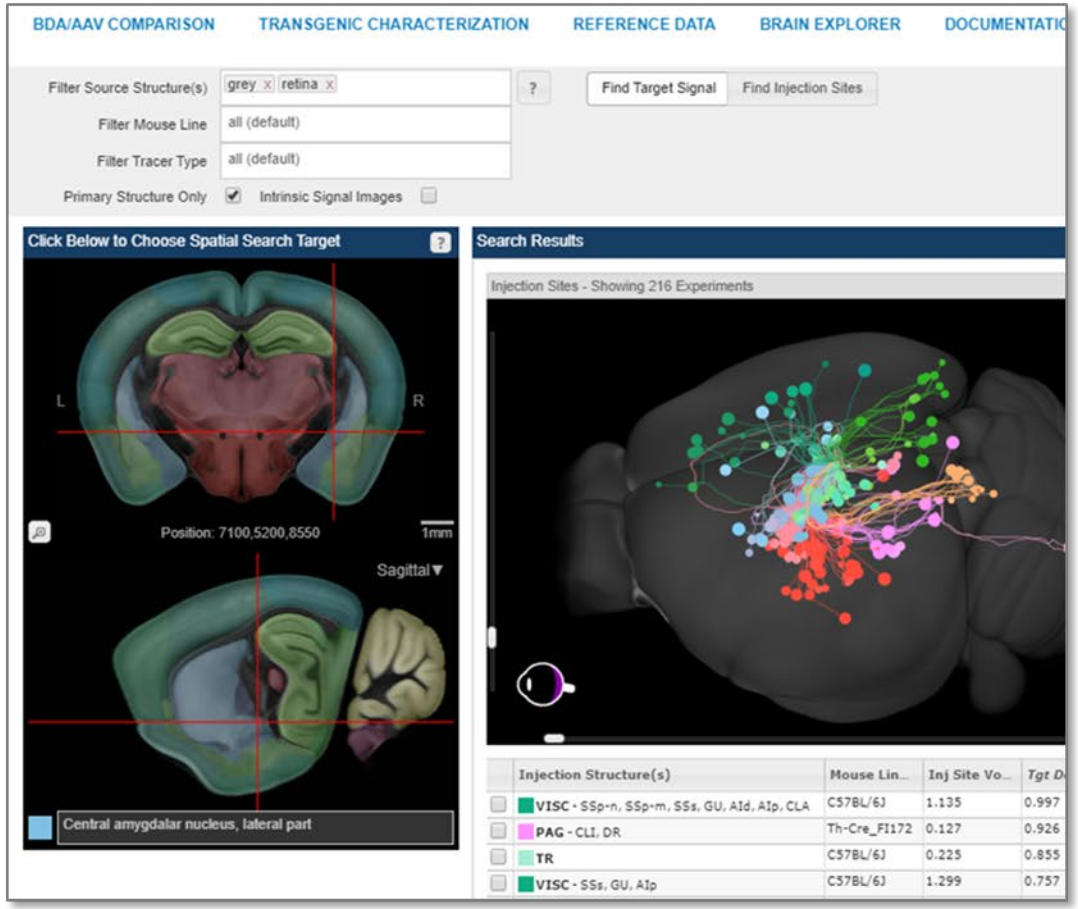


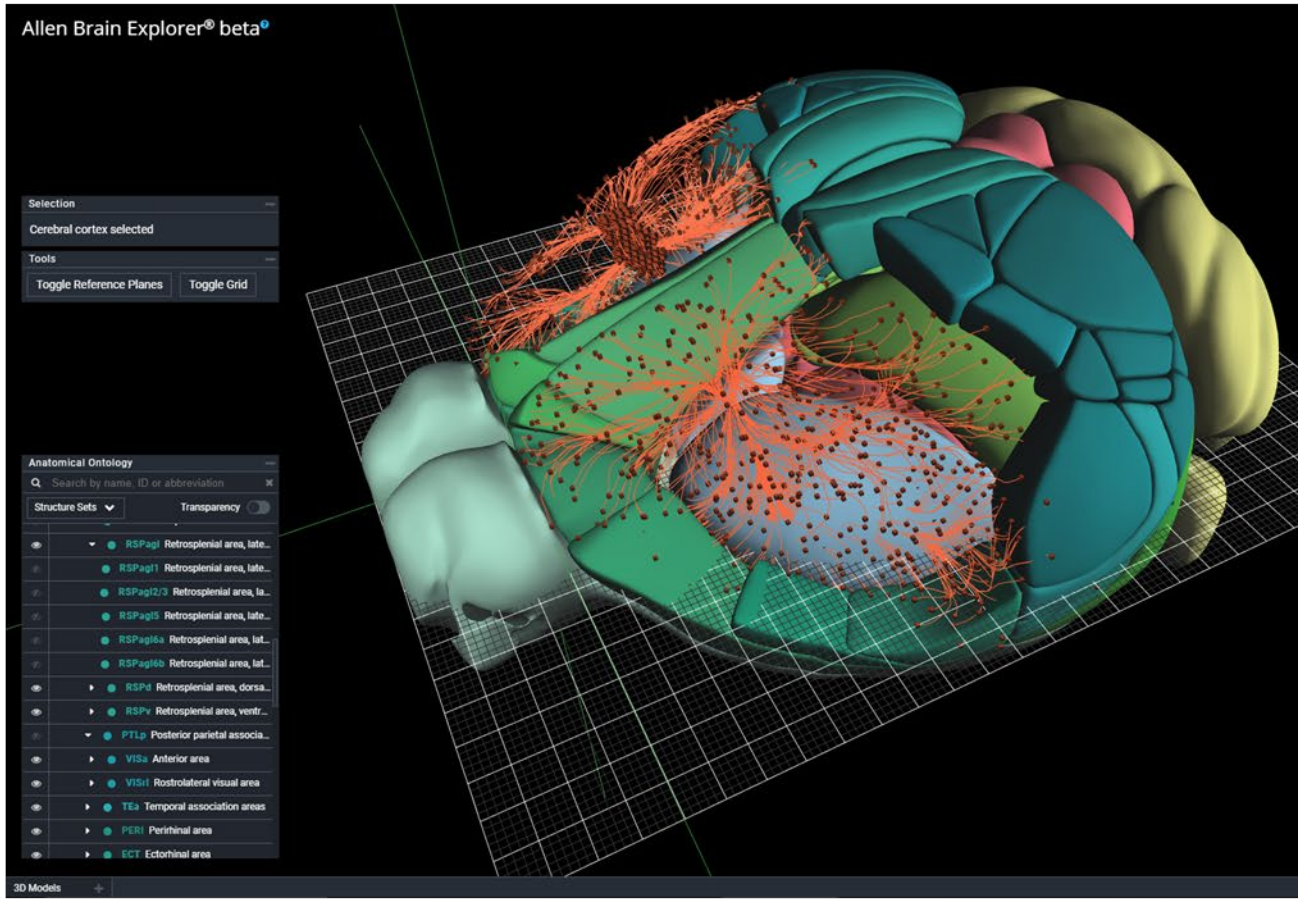

www.brain-map.org



# Allen Brain Explorer: View anatomy & connectivity in the mouse

connectivity.brain-map.org/3d-viewer

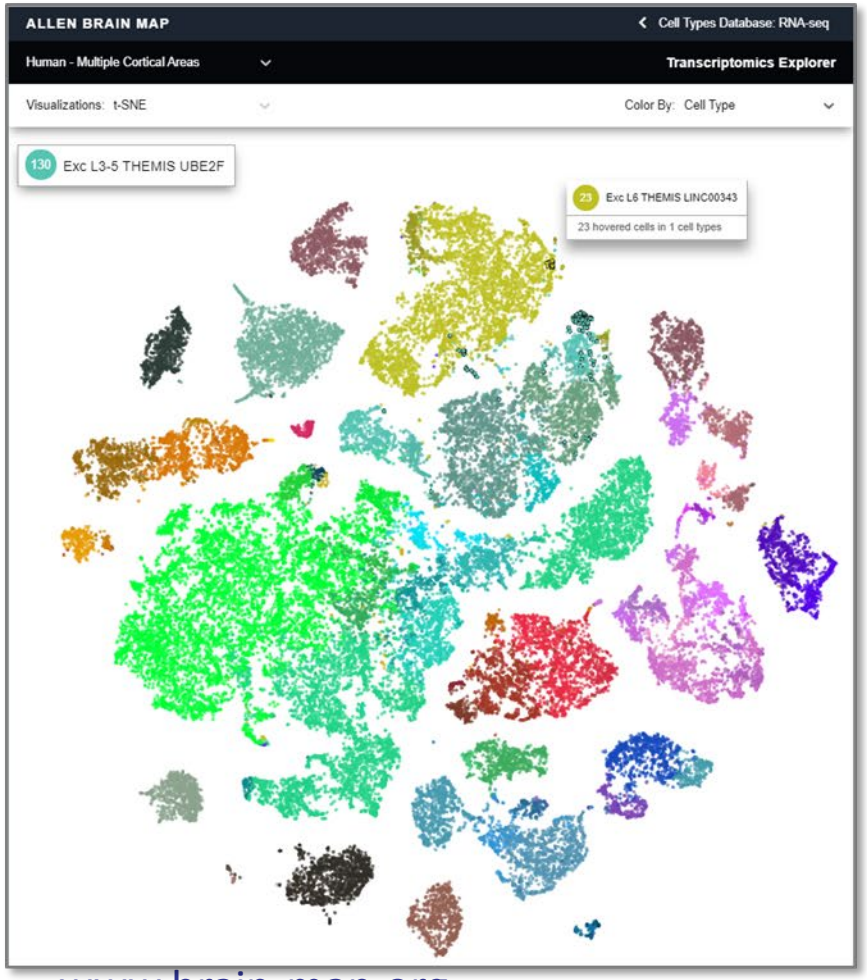



www.brain-map.org

# Allen Cell Types Database - Transcriptomics: RNA-Seq Data Navigators

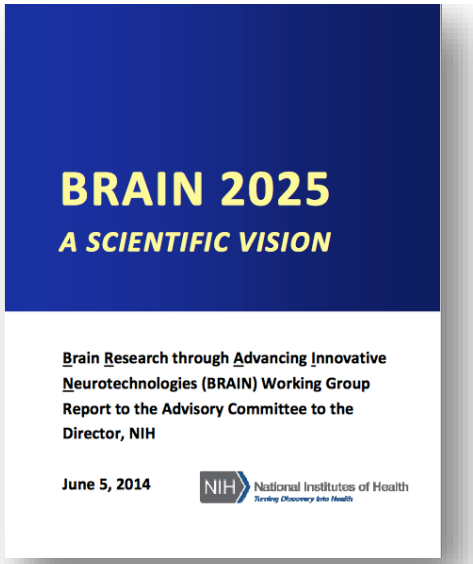
celltypes.brain-map.org/rnaseq




www.brain-map.org

# THE BRAIN INITIATIVE®

# Focus on Circuit Structure and Function

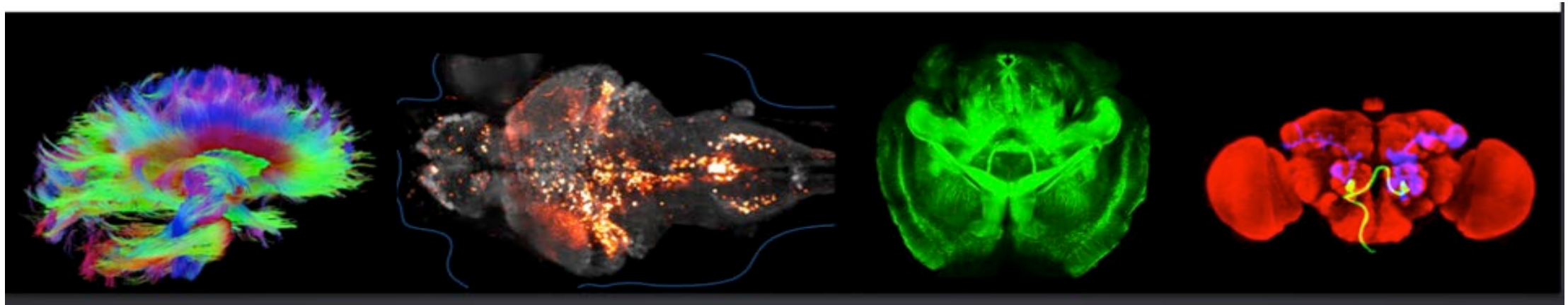



**Goal**: See circuits in action to understand the Brain

- How the brain moves, plans, executes
- How to monitor/manipulate circuits for improved function
- That disordered brain circuits cause neuro/mental/substance use disorders

**Long-term goal:** Make circuit abnormalities the basis of diagnostics, and normalization of circuit function the target of intervention

- 7 PRIORITY research areas - **Discovering diversity, Maps at multiple scales, the brain in action, demonstrating causality, Identifying fundamental principles, Advancing human neuroscience, integration**

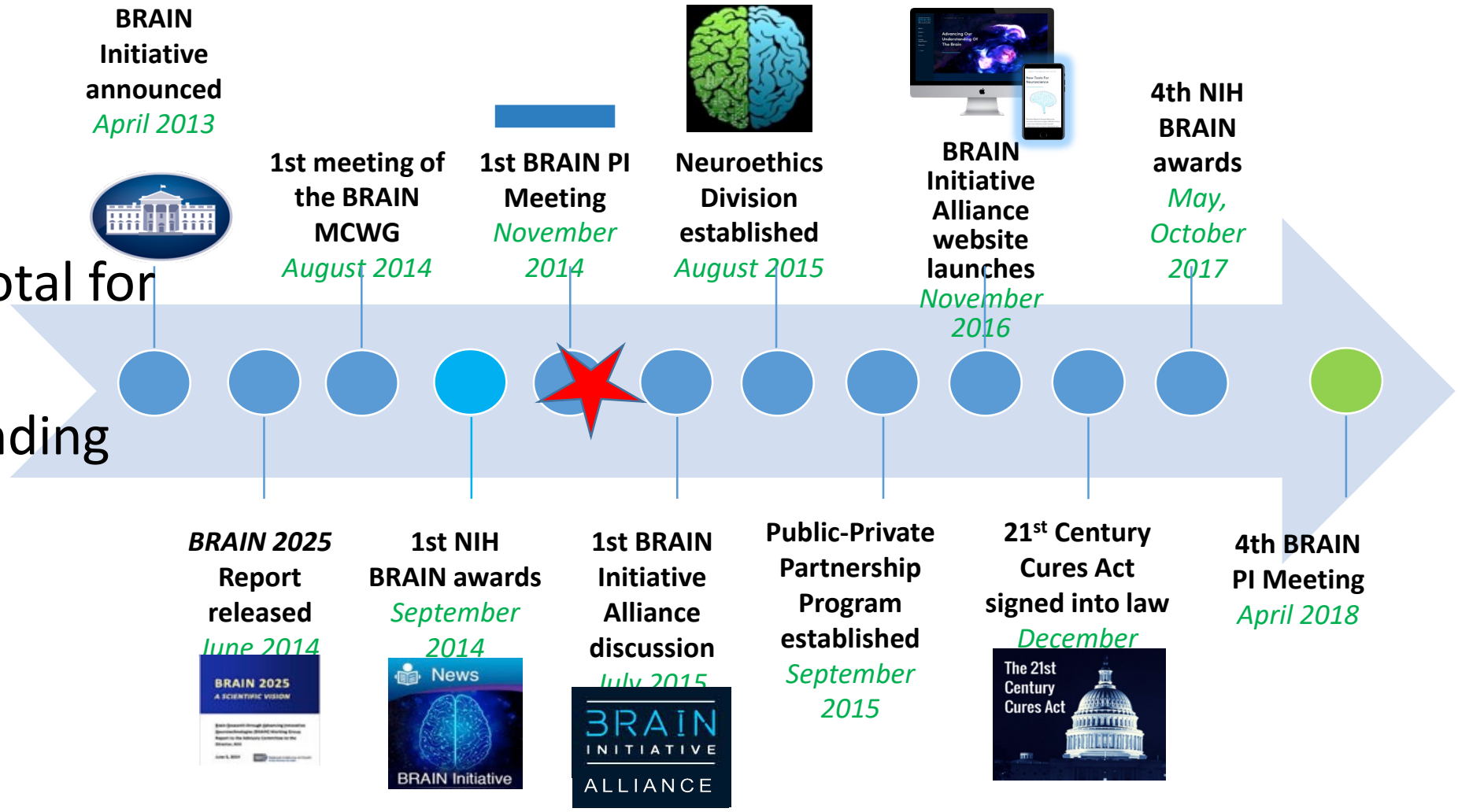


THE BRAIN INITIATIVE®

# Focus on Circuit Structure and Function

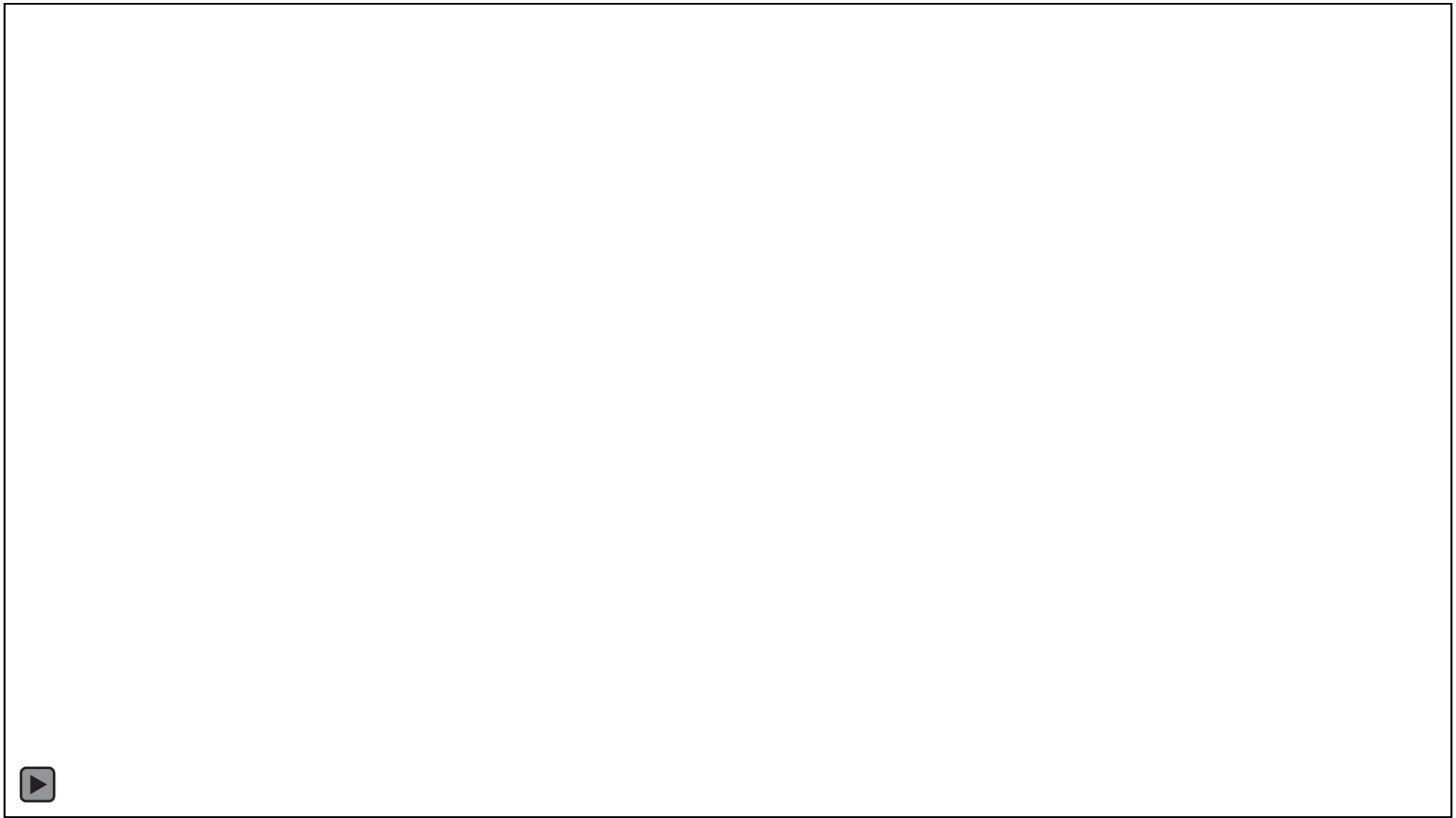
**$4.2B:** Projected total for BRAIN

**$550M:** BRAIN Funding through 2017




There's going to be so many piles of Data! What to do with it all?

# So Many Neurons In the Brain – how will we understand the form and function of all of them?



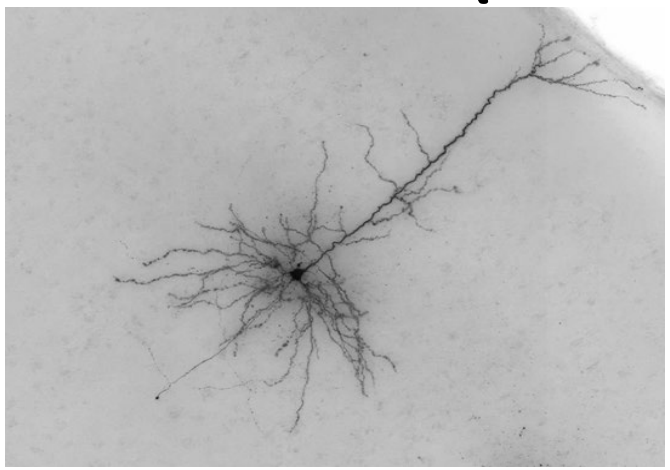
# Mozak: Democratizing Neuronal Reconstruction - Gamifying to Make the World a Neuroscience Lab

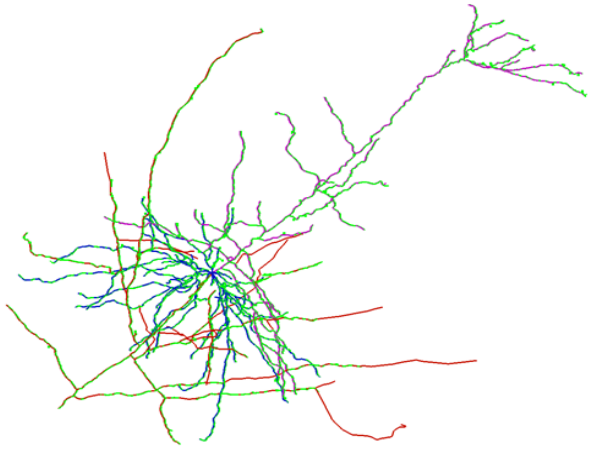
UNIVERSITY of WASHINGTON

With Zoran Popovic, UW, *Foldlt*

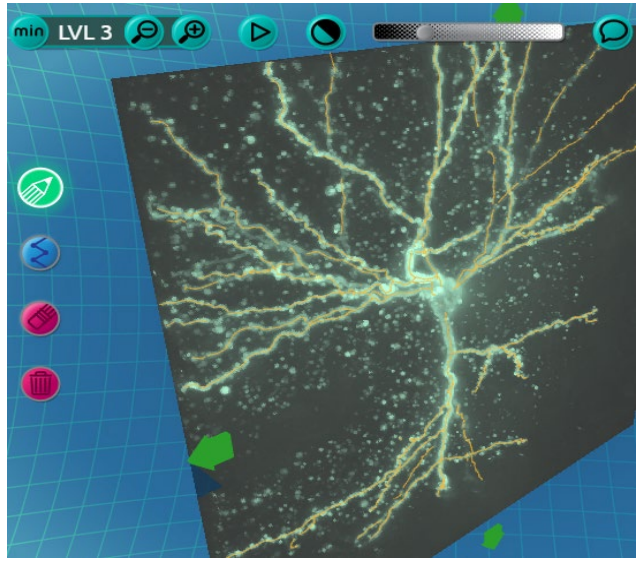
# Mozak – Create a Game to Rebuild the Brain one Neuron at a time (NSF-pilot, NIH Funded)

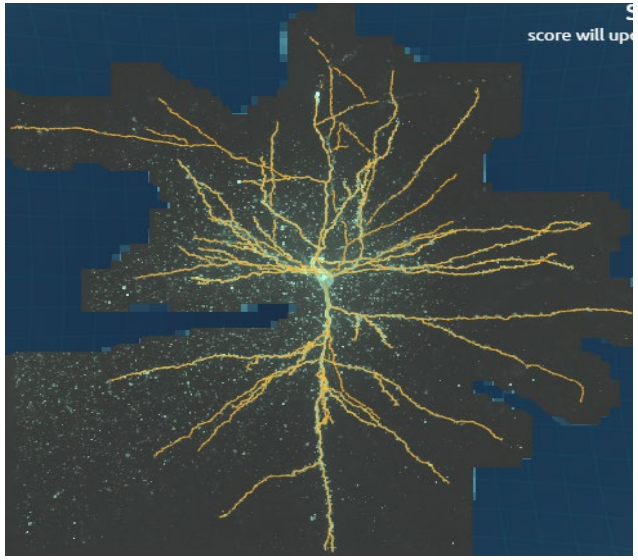
What
We
Need



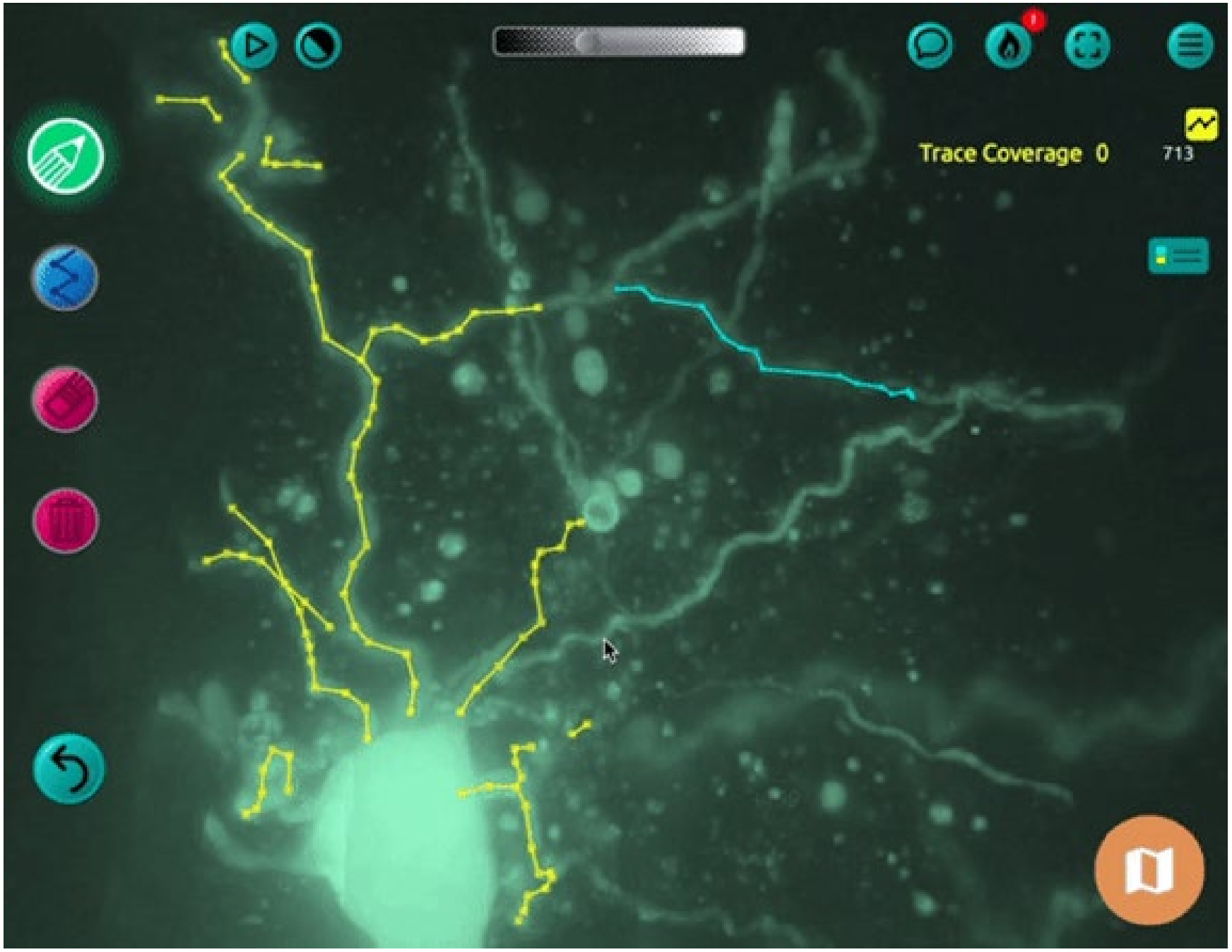


**What**
**We**
**Get**




***https://www.mozak.science/landing***




mozak brainbuilder

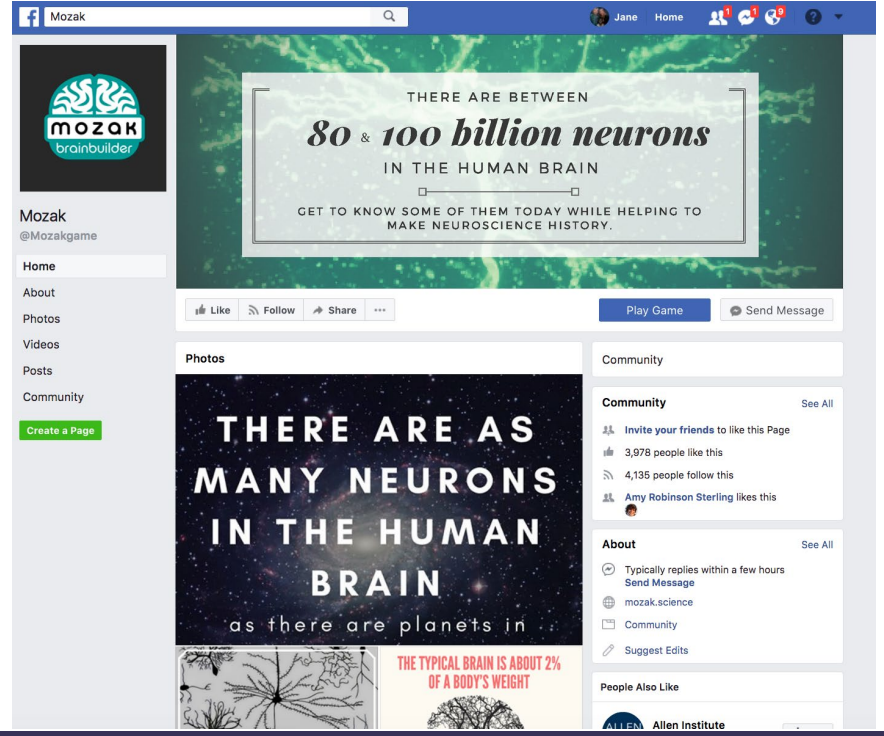
**Player Actions**:
- Create neuron edges
- Delete edges
- Connect edges
- 3D: Rotate, pan, zoom
- Navigate 2D map
- Adjust colors/contrast

**Automated Actions:**
- Calculate consensus
- Add/remove edges from consensus

# Mozak – Building a Global Community of Discoverers

- The good side of social media – FB, Twitter, NNR (Neuro News and Research)
  - and the NYT Science Section

- **100K players by 2018**
  - Increased output >4 fold over “experts” at AIBS
  - >90% accuracy on gold standard reconstructions
  - Already enhanced algorithm development
  - Now funded by NIH through 2024



# 2016: A year of Open Science, Citizen Science, 21st Century Cures, Hope….

the WHITE HOUSE
PRESIDENT BARACK OBAMA

BRIEFING ROOM | ISSUES | THE ADMINISTRATION | 1600 PENN

HOME · BLOG

## Open Science and Innovation: Of the People, By the People, For the People

PRESIDENT OBAMA'S FINAL
STATE OF THE UNION
JANUARY 12, 9PM ET

the WHITE HOUSE
PRESIDENT BARACK OBAMA

BRIEFING ROOM | ISSUES | THE ADMINISTRATION | 1600 PENN

HOME · BLOG

## Accelerating Citizen Science and Crowdsourcing to Address Societal and Scientific Challenges

SEPTEMBER 30, 2015 AT 6:00 AM ET BY TOM KALIL AND DAVE WILKINSON

Summary: Today, the White House is hosting a forum on citizen science and crowdsourcing.

- and heart-break

BILL & MELINDA GATES *foundation*

# HEALTHY BIRTH, GROWTH & DEVELOPMENT

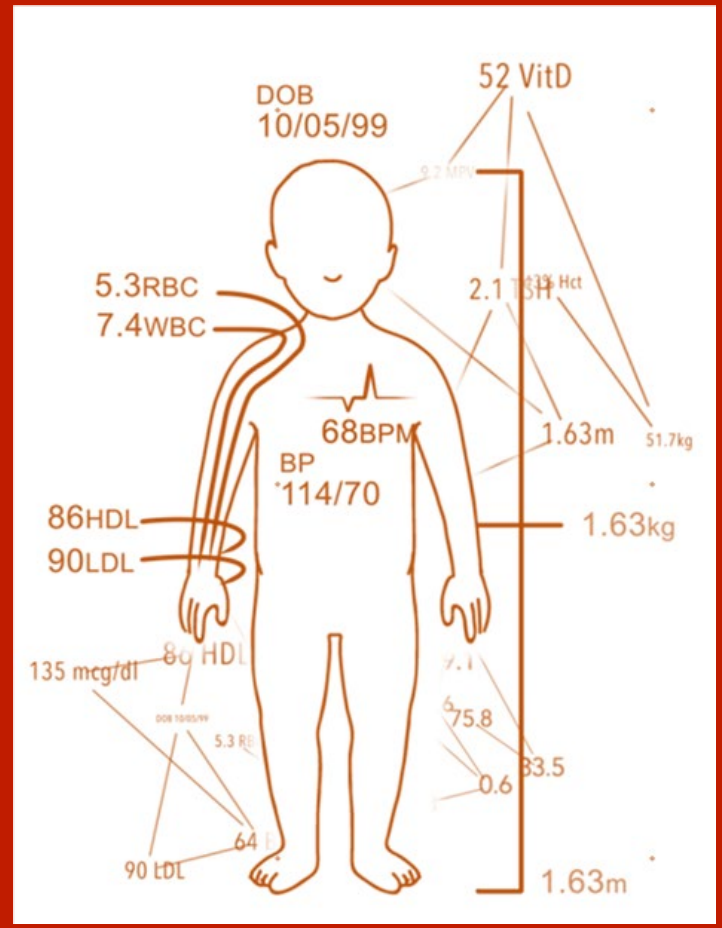
## *knowledge integration*

N. L'ntshotsholé “Shasha” Jumbe

# HBGDki What Do we Really Need in our First 1,000 days?

## *OVERALL GOALS- STUNTING (AND BEYOND?)*

- HBGDki GOAL - to determine the **right** interventions in the **right** dosage to get the **right** response and avoid the **wrong** outcomes
- Integrating knowledge to quantify drivers of variability, determine effect size and enumerate interactions
- While focus is on poor communities with greatest need, counterfactual data collected to accelerate learning on intervention effects



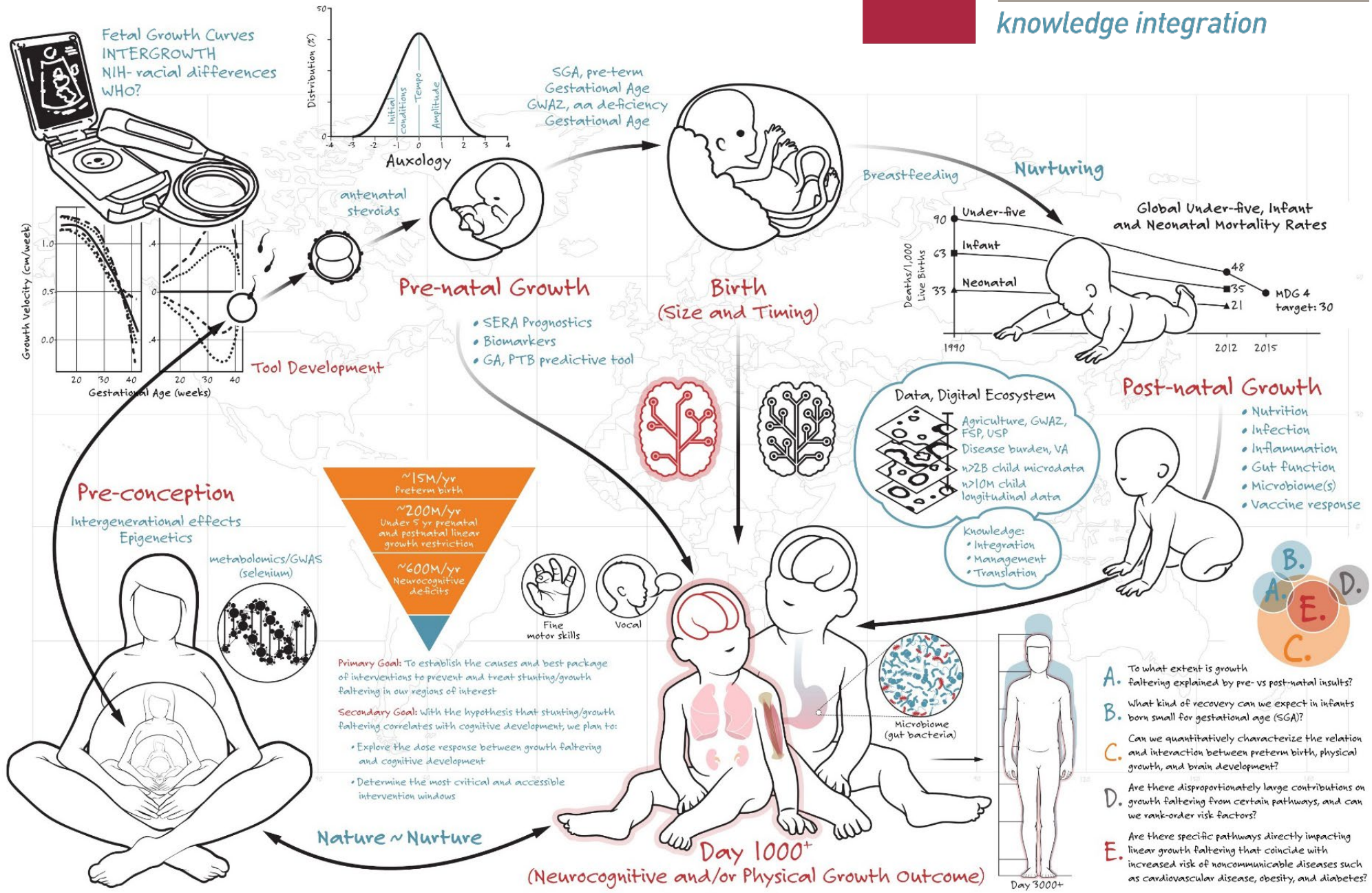

## *WHAT I LEARNED – THE GOOD, BAD, CHALLENGING*

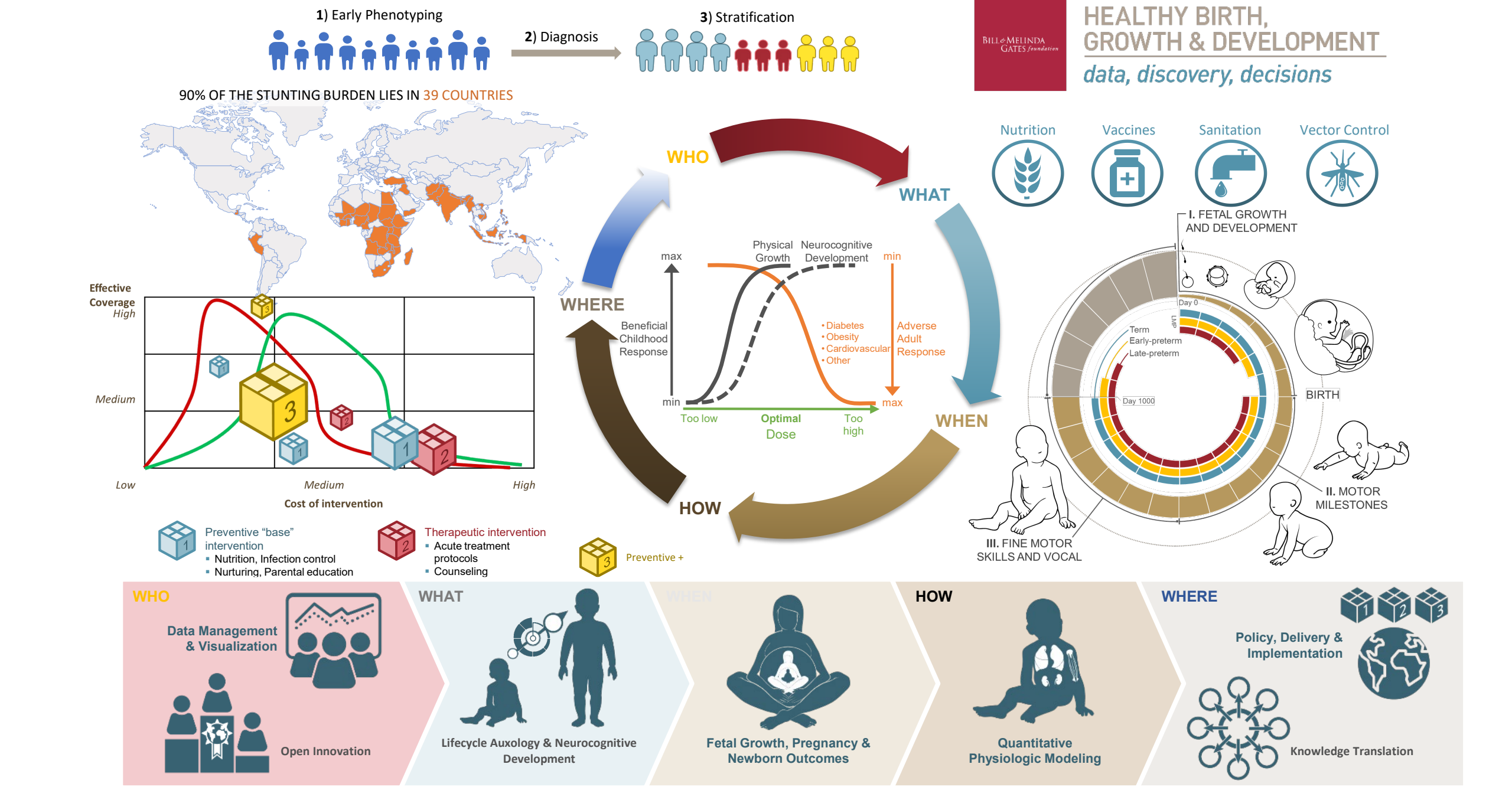
# What Do we Need in our First 1,000 days?

BILL & MELINDA GATES foundation

HEALTHY BIRTH, GROWTH & DEVELOPMENT

*knowledge integration*



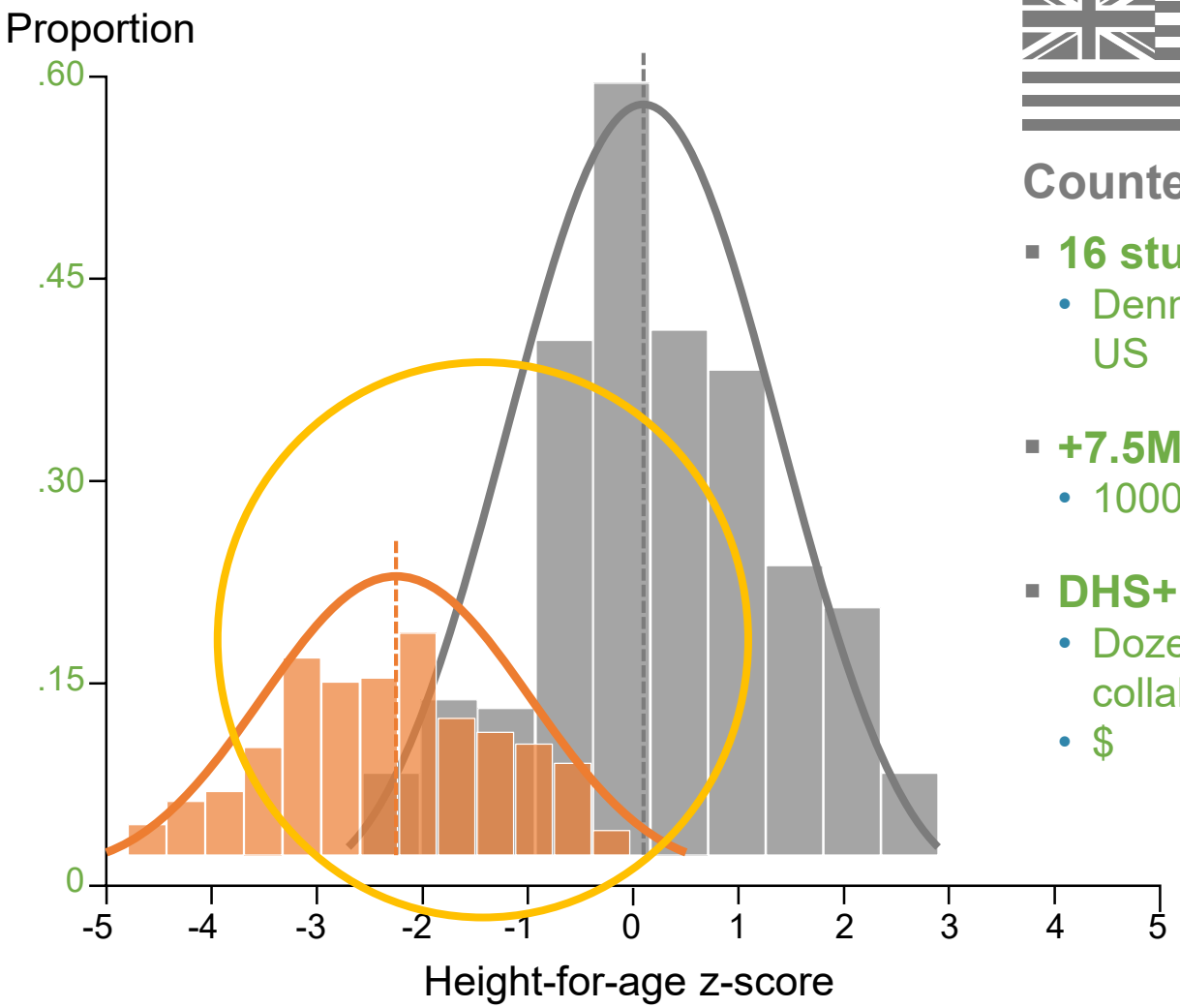



# Early Development is Highly Complex: Gazillions of tiny variables

# The Hbgd*ki* Knowledge base contains LMIC and counterfactual HIC data as natural experiments

HBGD*ki*

- **Reality: LMIC**
- **153 studies:**
  - From most high stunting burden countries
- **~10.5M subjects data:**
  - 100s of covariates
- **DHS+:**
  - Dozens of prospective collaborations
  - $$$




**Counterfactual: HIC**

- **16 studies:**
  - Denmark, Netherlands, Singapore, US
- **+7.5M individual subject data:**
  - 1000s of covariates
- **DHS+:**
  - Dozens of prospective collaborations
  - $


© Bill & Melinda Gates Foundation |

# What was different about HBGD*ki*?

- Integrating knowledge to quantify and understand drivers of variability

## OPEN DATA and ANALYSIS - DRIVING CATALYTIC CHANGE…

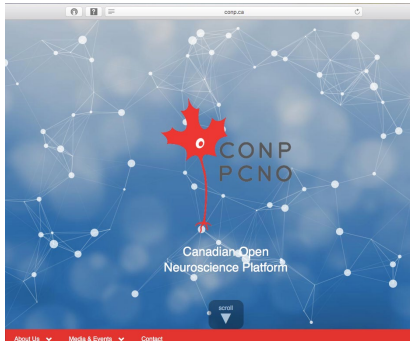
Aim to achieve ***greatest impact*** in the ***shortest period*** of **time** with the ***least amount*** of human and financial resources.

# What is a BRAIN Commons?

- A scalable, centralized big data cloud-based platform for computational innovation and data driven discovery for brain diseases
- Integrates individual level data across data types *(genomics, biomics, imaging, wearable, etc.)*
- Able to scale to work with large quantities of data (begin with PTSD)
- Equipped with data-analysis and systems biology tools
- FISMA Moderate Security & HIPAA Compliant & Controlled Data-sharing

Regulation, Governance, legal obstacles across studies, importance of core analysis

# CANADIAN OPEN NEUROSCIENCE PLATFORM




www.conp.ca

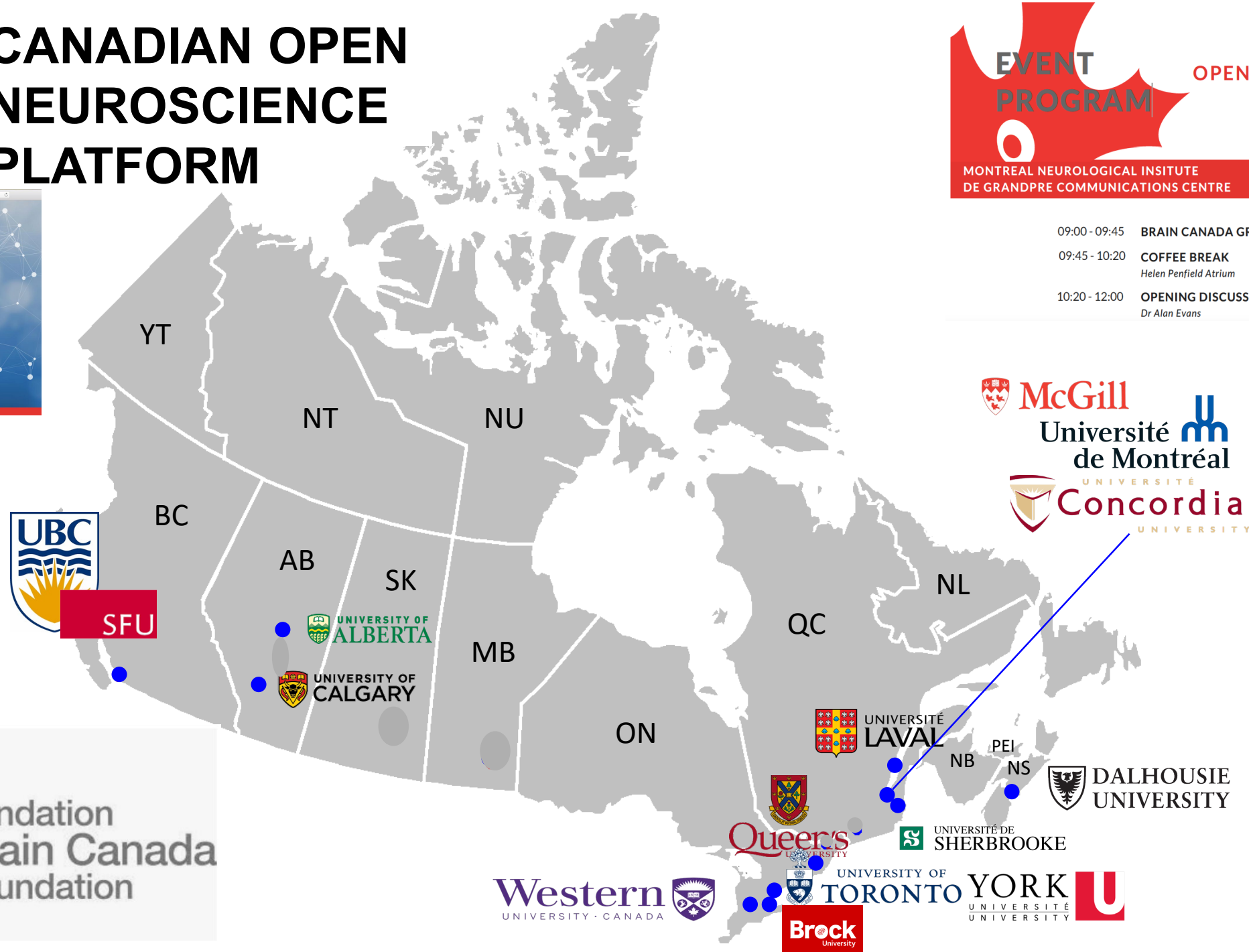
EVENT PROGRAM

CANADIAN OPEN NEUROSCIENCE PLATFORM
PLENARY MEETING

MONTREAL NEUROLOGICAL INSITUTE
DE GRANDPRE COMMUNICATIONS CENTRE

FEBRUARY 19, 2018

| | |
|---|---|
| 09:00 - 09:45 | **BRAIN CANADA GRANT ANNOUNCEMENT** |
| 09:45 - 10:20 | **COFFEE BREAK**<br>*Helen Penfield Atrium* |
| 10:20 - 12:00 | **OPENING DISCUSSION**<br>*Dr Alan Evans* |




Fondation Brain Canada Foundation

# The International Neuroinformatics Coordinating Facility

## What is INCF?

### A standards organization for open and FAIR neuroscience

The mission of INCF is to develop, evaluate, and endorse standards and best practices that embrace the principles of Open, FAIR, and Citable neuroscience. INCF also provides training on how standards and best practices facilitate reproducibility and enables the publishing of the entirety of research output, including data and code.

#### Standards

INCF serves as a standards organization dedicated to open and FAIR neuroscience by vetting, endorsing, and promoting the use of community standards and best practices.

#### Training

INCF provides informatics educational resources for the global neuroscience community both online and through in-person courses and workshops at the INCF Assembly.

#### Working groups

In our Working Groups and Special Interest Groups, users and developers work collaboratively to develop community standards and best practices and implement them in tools and resources.

#### INCF Assembly

The INCF neuroinformatics conference provides a forum for researchers, infrastructure providers, and developers to connect and train or be trained in neuroinformatics.

# The International Brain Lab (IBL), Intern'l Brain Initiative (IBI)

# Lessons Learned

- OPEN = data and software and AI and ideas and protocols and minds (and hearts)
- Egos cannot drive open, but smart, ethical people reaching consensus can
- Without acceptable standards, data is not FAIR
- Without open software sharing (models, analytics), data are just data

- Listen to and understand the people who collect, and users who need to use, the data
- Use Cases will get you closest to the answers you are looking for
- UX Design (User interface) is critical for engagement
- Track users - where do data/software go, who is using them, how?
- Understand and respect (and change?) the ethical and legal

restrictions for “open” use of data, software, AI, publishing

The AI revolution is Coming (Open AI Commons)!

# It takes an entire Network of Communities to Make Open Science Happen

- The Chemical Senses  community
- The Spinal Cord Community/Neurotrauma community
- The Stem Cell Community (Stem Cell Network of Canada)
- The International Brain Org Community – IBRO, INCF
- Global Teaching and Outreach – SfN, BrainFacts.org, Dana Foundation
- Open Data Sharing and Open Science Community (including Sage Bionetworks, CONP, ODSC)
- The Citizen Science Community  (inc. DREAM Challenges)
- The Patient Communities (and advocates) who let us use their data

The Philanthropy Community–Brain Canada, BRAIN Initiative, Reeve Foundation (CDRF), Spinal Research (ISRT), Wings For Life, NMSS, MSHRF, MS Canada, Kavli, Gatsby, Simons Foundation, UsAgainst Alzheimer's (George Vradenburg), OneMind, Xprize, Wellcome Trust, Paul and Jody Allen, Bill and Melinda Gates Foundation, Ludmer Foundation, Larry Tannenbaum (TOSI).

UBC

UNIVERSITY of WASHINGTON

NIH National Institutes of Health

NSF

Natural Sciences and Engineering Research Council of Canada | Conseil de recherches en sciences naturelles et en génie du Canada | Canada

CIHR IRSC