

# National Cybersecurity Center of Excellence

## NCCoE Virtual Workshop on Cybersecurity of Genomic Data

Wednesday, January 26, 2022, 11:00 AM – 4:30 PM (ET)

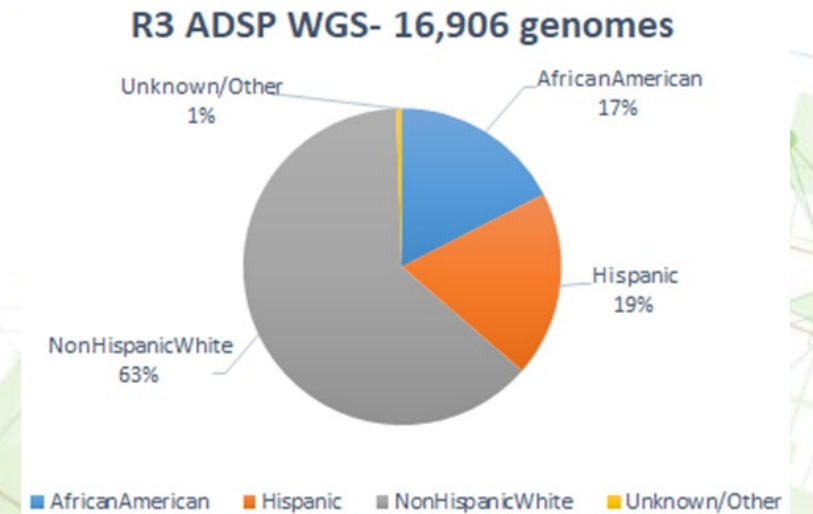
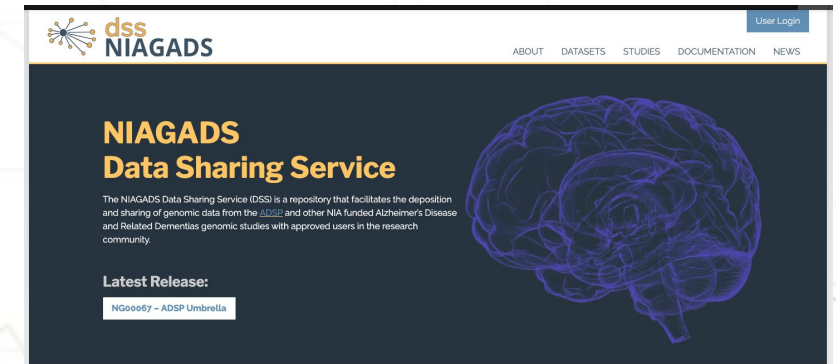
# Lessons Learned from NIAGADS and ADSP Data Sharing

Wan-Ping Lee & Li-San Wang

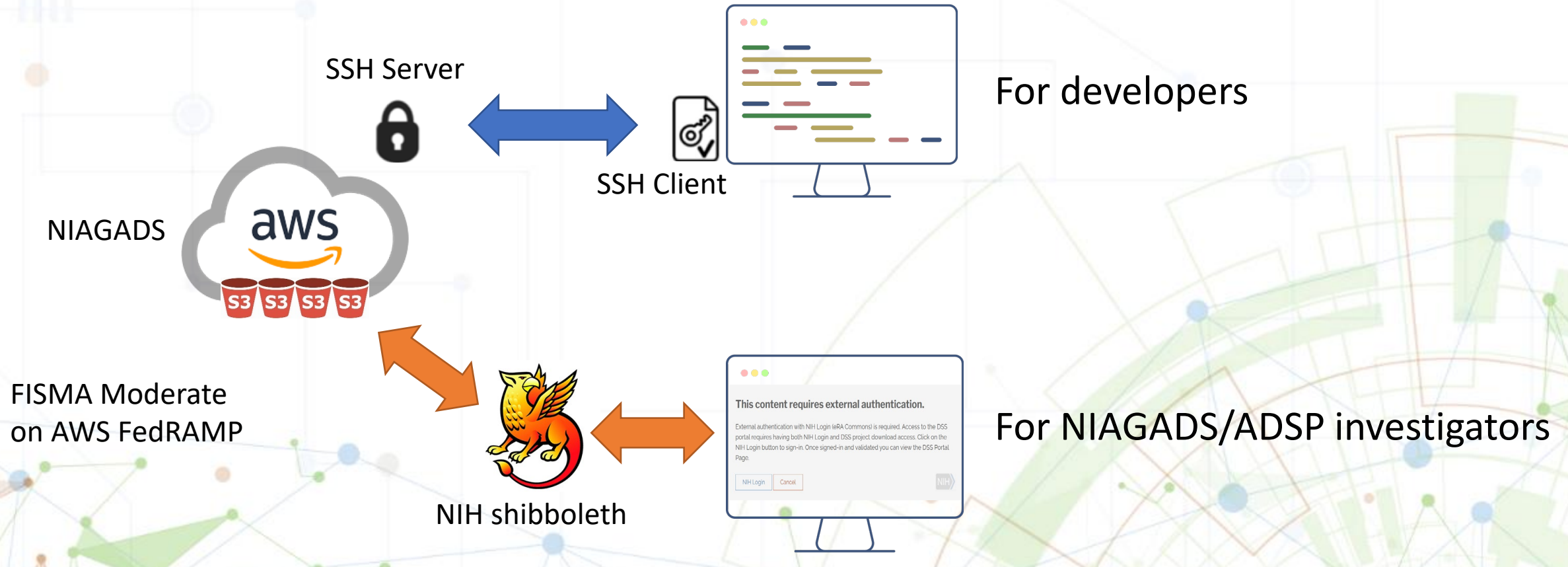
University of Pennsylvania Perelman School of Medicine

# Alzheimer's Disease Sequencing Project (ADSP) and NIAGADS

- 17k complete genomes released in 2021
- 36k (2022), 75k (2023) planned
- Data releases are managed by NIAGADS
  - National Institute on Aging Genetics of Alzheimer's Disease Data Storage Site at University of Pennsylvania

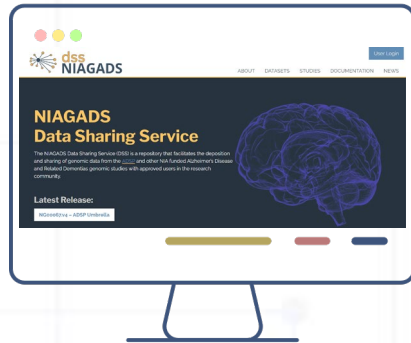


# NIAGADS DSS Security Setup

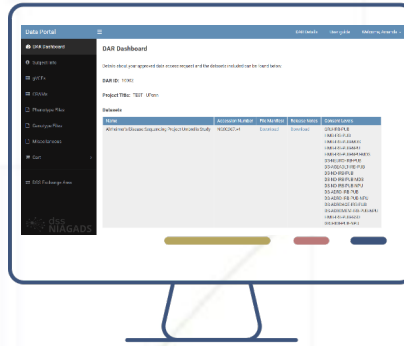
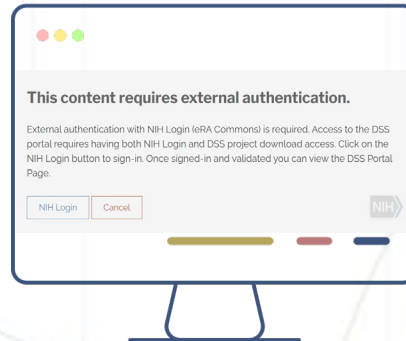


# Authentication of data requesters

Step 1: <https://dss.niagads.org>



Step 2: NIH eRA Authentication



Step 3: Data manifest browsing on NIAGADS DSS



Step 4: Data accessing



# Learning curve for FISMA is steep

- Scope
- Amount of work
  - 10 months preparation
  - 1 year for 3 times external assessment
- Regulations to be met
- Cost benefit analysis

# Respecting Informed Consent

- NIH Genome Sharing Policy
- Institutional certification to capture informed consent conditions
- Data access committee
- How to split data based on informed consent

# De-identification of genomic data

- Genome sequencing is identifiable
- What about functional genomics data? Theoretically RNA-Seq is identifiable because it carries variants. What other types of sequencing data?
- What does it take to de-identify data?



# What will be helpful if we do it again

- Guidelines for FISMA requirements that are specific to human genome data
- Tutorials and FAQs on how to set up a FISMA compliant framework: timeline, amount of work, cost (budget and staff)