ODSS AI PI Meeting
Slido Highlights

What award are you associated with?
Multiple Choice Poll — 83 votes — 83 participants

FY22 AI Ethics (NOT-OD-22-065) - 17 votes — 20%

FY22 AI Readiness (NOT-OD-22-067) - 33 votes — 40%

FY23 AI Readiness (NOT-OD-23-082) - 27 votes — 33%

Observer - 6 votes — 7%

What areas of expertise did you bring to your team?
Wordcloud Poll — 132 responses — 79 participants

Law · Data integration · statistical modeling · Time series · clinical knowledge · Statistical data analysis · statistics · Bioethics · social sciences · Data Science · Qualitative research · Machine learning · Genomics · science · Spine · rehabilitation · Bioinformatics · anthropology · modeling · EHR · ethics · social · AI and medical imaging · Proteomics · Data privacy · Data Science, AI Model De

What best practices did you use to test/validate efforts to mitigate biases in your project (e.g., in the data, processes)?
Open text poll — 38 responses — 35 participants

Anonymous
Retrain available models with data actually from target populations

Anonymous
Random generated parametric test data

Anonymous
Feedback from people being represented

Anonymous
NA... animal data, collected from multiple species

Anonymous
Participatory action research foundation

Anonymous
Expert curated data

Anonymous
Statistical assessment of patient demographics

Anonymous
qualitative; engaging with original data users

Anonymous
investigation of data collection practices

Anonymous
Preprocessing components that influence the biases

Anonymous
Uniforming data distribution

Anonymous
Single cell data and optimal transport

Anonymous
Test for heterogeneity of performance across different populations in validation

Anonymous
external validation

Anonymous
Tools to test biases like Fairlearn

Anonymous
blinded study design

Anonymous
Expert review

Anonymous
Single secure collaboration environment--no silos. Regular check-ins across diverse team members.

Anonymous
review of data collection processes from many different individuals/stakeholders

Anonymous
Uniforming data distributions, data transformations
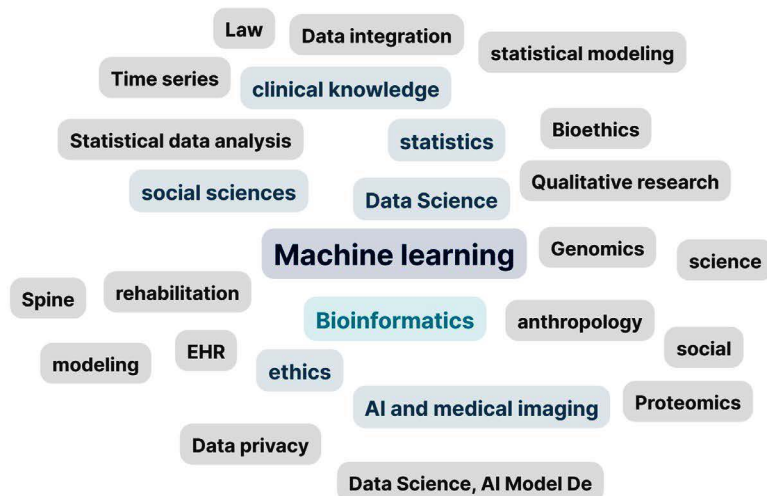
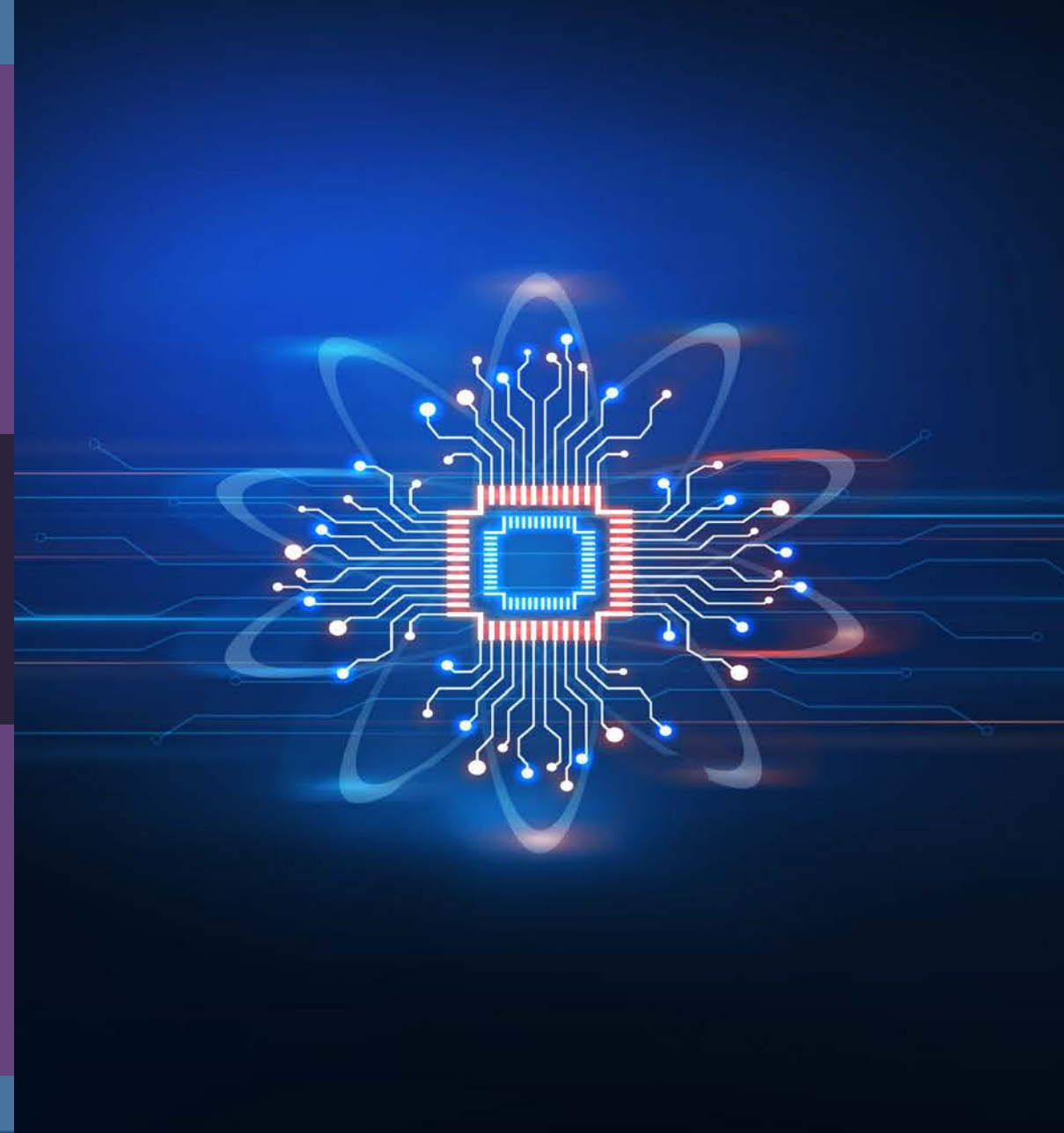# ODSS AI PI Meeting 2024
Example Highlights

- **Many groups encountered challenges and surfaced gaps/needs in generating 'AI ready' data**
  - Data harmonization – aggregating data / challenging data types / combining multimodal data
  - Disparate standards, formats, schema
  - Quantifying data collection bias
  - Missingness & balancing representation
  - Removing noise while preserving information from data
  - Comprehensive data dictionaries
  - AI readiness – can mean different things at different stages of development pipeline

- **Sharing methods/reporting/documentation around…**
  - Models with incomplete data
  - Approaches for improving 'noisy'/ 'dirty'/ 'historical' data
  - Quantifying and improving representativeness
  - Challenges & strengths of synthetic data
  - Findability of/access to data/datasets
  - Reproducible pipelines for data processing/sharing/FAIR
  - Appropriate metadata & provenance
  - Uncertainty quantification
  - Continual monitoring/validation
  - Measuring, identifying, and mitigation biases – metrics/tools
  - Ethical principles and approaches that evolve with the field / needs & concerns of communities

- **Ethical approaches & needs**
  - Privacy standards & reuse and consent limitations/concerns
  - Biases encountered across the board
  - Gauging community perceptions for AI use in healthcare
  - Improving training on the ever-evolving landscape of AI and AI-Ethics
  - Intersectional approaches are key
  - AI Ethics experts on appropriate review panels (e.g., IRBs)/collaborative projects
  - Tackling problems relevant to disadvantaged populations – e.g., through SDOH; improved approaches to these challenges
  - Community engagement strategies and processes are key
  - Critical to understand & respect autonomy/needs of communities affected or targeted
  - Ethical and responsible AI requires commitment to evaluation – will change over time and requires many stakeholders
  - Further questions surrounding generative AI/LLMs

- **Training**
  - Culture change needed for inherently multidisciplinary space – need for support for cross-disciplinary work and supporting the next gen of researchers on this path
  - Collaboration gaps, spurring interdisciplinary collaborations
  - Enhanced workshops/trainings etc., to teach researchers in quickly moving space – best practices, current state, leading edge of AI/AI Readiness/AI Ethics
  - Access to compute & data infrastructure is challenging / expansive

Bridge2AI 2024 Open House

April 18-19, Rockville, MD

NIH Neuroscience Center

Learn More:

http://bit.ly/4c5fV9x

National Institutes of Health
Office of Strategic Coordination–The Common Fund

Bridge2AI Open House

April 18-19th, 2024 | NIH Neuroscience Building

- Join Bridge2AI for the 1st Open House, featuring highlights from the four Grand Challenge Data Generation Projects (DGPs) with information on new biomedical datasets.

- Attendees will have an opportunity to meet the DGPs and learn about the launch of Bridge2AI's Data Challenges and our road ahead with biomedical AI.

- More information on the Open House and registration may be found: http://bit.ly/4c5fV9x

Grand Challenges:

- Voice AI
- Salutogenesis
- Clinical Care
- Functional Genomics

Call for Submission

NIH Special Track – ISMB 2024 (July 12-16, 2024; Montreal)

https://www.iscb.org/ismb2024/home

One-day special conference track on NIH funded projects focusing on:

Artificial Intelligence & Machine Learning (AI)
Cloud Research
Research Software Development

Awardees from the following funding opportunities are invited to submit an abstract for selection:

- AI/ML Readiness, Ethics, Bias, Transparency, Workforce (NOT-OD-21-094, NOT-OD-21-079, NOT-OD-22-065, NOT-OD-22-067, NOT-OD-23-082)
- AIM-AHEAD (OTA-21-017)
- Exploratory Cloud Research (NOT-OD-23-070)
- Software Tools for Open Science (NOT-OD-20-073, NOT-OD-21-091, NOT-OD-22-068, NOT-OD-23-073)
- NIH-NSF Smart Health Program (SCH) (NOT-OD-21-011, NOT-OD-23-165)

Stay tuned for information on how to submit an abstract not to exceed **300 words** on project accomplishments by end of **April 2024**

National Institutes of Health
Office of Data Science Strategy

NIH Office of Data Science Strategy

datascience.nih.gov

A modernized, integrated, FAIR biomedical data ecosystem

@NIHDataScience

linkedin.com/showcase/
nih-office-of-data-
science-strategy

datascience@nih.gov