Cancer Biomarkers AI and Bioinformatics Workshop Planning Call

Date: November 30, 2023

Time: 9:00 AM - 10:00 AM PST

Attendees:

- Daniel J. Crichton (NASA JPL)
- Ziding Feng (Fred Hutchinson Cancer Research Center)
- Matthew B. Schabath (Moffitt Cancer Center)
- Anirban Maitra (MD Anderson Cancer Center)
- Sudhir Srivastava (NIH/NCI)
- Christos Patriotis (NIH/NCI)
- Natalie Abrams (NIH/NCI)
- Jennifer Ellen Beane-Ebel (Boston University)
- Zhen Zhang (Johns Hopkins University)
- Eugene Koay (MD Anderson Cancer Center)
- Ashish Mahabal (Caltech)
- Heather Kincaid (NASA JPL)

Agenda:

1. Workshop Vision and Objectives Overview
2. Workshop Goals Discussion
3. Audience and Program Committee Proposals
4. Workshop Format and Session Structure
5. Hackathon Planning
6. Workshop Preparation Next Steps

Action Items:

- ☐ Send additional Program Committee member suggestions to Dan or Heather
- ☐ Feedback on workshop dates
- ☐ Create Google Doc to shape high-level topics/sessions with definitions and get feedback
- ☐ Feedback on Audience
- ☐ Post Workshop info on EDRN Portal
- ☐ Determine date/time for a regular Program Committee Planning Call

Meeting Notes:

- The workshop will be held at Caltech. Possibly June 2024
- 3 day workshop and ½ day Hack-a-thon
  - Keynotes
  - Call for talks
  - Call for posters
- Goals of meeting
  - Define use cases
    - Include discovery phase of research
    - Lead with science questions and how AI can help
    - Include Disparity and FSI (Foreign Service Institute)
  - Discuss the state-of-the-art and define current gaps
  - Discuss challenges around reproducibility, interpretability, and explainability
  - Discuss current informatics capabilities and infrastructure needs to support use cases
    - Include data sharing/data transfer/federated learning
    - Include legal, security, privacy, DUAs, non tech issues
  - Develop a set of goals and recommendations to address these needs in cancer biomarker research for the next few years
- Audience
  - EDRN
  - NCI
  - Academic partners
  - Others mentioned:
    - Invite other agency programs
    - ARPA-H
    - DARPA
    - Cooperative Agreements
    - Industry partnerships
- Sessions
  - Proposed high-level sessions
    - Session 1: Cancer biomarker discovery -- use cases and opportunities for AI
    - Session 2: AI and Bioinformatics in biomarker validation
    - Session 3: Data preparation and design
    - Session 4: Application of AI, ML and bioinformatics algorithms and methods
    - Session 5: Infrastructure support for data capture, sharing, and computation
    - Session 6: Recommendations and next steps
  - Feedback on call
    - Create sessions after open calls?
    - Maybe give structure to calls and then use that info to create detailed sessions
    - Think about focusing sessions to include the following:

- Imaging
- Large — and EHR data
- High dimension data including omics, biomarker discovery
- Maybe dedicated topics?
  - Include types of substrates
    - Blood
    - Tissue
    - Imaging
    - Radiomics
  - Also include utilizing data that contains PII, how to deal with (privacy / policy). Cooperative Groups have to deal with this, so may have good use cases.
  - Focus sessions with engaging titles that are very clear.
  - Invite a speaker to address the bias that can occur with AI - Challenges of Bias and Ethical Considerations.
  - UCSF just had a good conference on this
  - Possible suggestion on topics from Sudhir
    - 1. in Silico Biomarkers's Discovery - methodology, application and verification
    - 2. Data Integration: Integration of imaging, radiomics, high dimensional data (Proteomics's, genomics, epidemics, etc.)
    - 3. Application of AI/ ML: Data Preparation, Harmonization and standardization
    - 4. Case Studies for AI/ML (see attached reference)
    - 5. Academic-Industry Partnerships: what to expect? PreCompetitive Collaboration?
    - 6. Showcasing Datasets for AI/ML - real world data, synthetic data, simulated data
- Hack-a-thon
  - ½ day is not very long
  - We want to use this to understand and gather ideas
    - Gaps
    - Bringing data together
    - Diverse technology and diverse data
    - Should this be a linked to a set of questions that motivate the hack-a-thon
    - Should provide a structure so it is productive
- Other questions
  - Funding to support speakers and early stage investigators to attend?
  - Reach out to Data Commons folks
  - Publish workshop output