## CHoRUS DGP

| Module | Challenges | Proposed solutions |
|---|---|---|
| Data Acquisition | Assessing level of readiness and specific site challenges<br><br>Identifying key line of communication with sites<br><br>Responsiveness to tactical obstacles encountered at sites<br><br>Defining core data elements<br><br>Identifying a common data platform*<br><br>Willingness of sites to contribute data<br><br>*Shared with Tools module | Initial detailed survey<br>Individual site exploratory meetings<br>Identification of site experts for procurement of domain-specific data<br><br>Broad site representation and designation of a main contact<br><br>Open office weekly hours<br>Targetted Module PIs-sitePI and relevant site representives meetings<br><br>Survey of existing critical care datasets and case report forms of large trials<br>Open collaboration with the Society of Critical Care Medicine<br>Consensus development process among site PIs and CHoRUS MPIs<br><br>Broad survey of potential vendors<br><br>Early engagement of key site leaders |
| Ethics | Clarify control, stewardship of gathered data and de-risk institutions from IP and privacy loss<br><br>Data colonialism | Meet with hospital CIOs<br>Prevent re-identification through Diff Privacy<br>Consider approach to manage “dual uses”<br><br>Focus groups |

| | | |
|---|---|---|
| | Vulnerable populations<br><br>Public-Private Partnership<br><br>Exemption from consent and highly complete and sensitive data elements (e.g., neurologic disease)<br><br>Trustworthy pathways for data-sharing and use of CHoRUS data for possible future, non-research uses (e.g., public health, regulatory, treatment uses of CHoRUS data)<br><br>Representative datasets to ensure models calibrate/validate<br><br>Representing race/ethnicity/sex/gender in convenience samples<br><br>Modeling SDoH from geocoding is potential privacy risk; lack of completeness of notes; lack of reliability of public datasets (undomiciled)<br><br>Federal regulatory policy is chilling clinical decision support | Following indigenous protocols<br>Ensuring representation of the study team<br>Focus groups<br><br>Two-stage IRB/contract process to allow for ETAI focus groups<br><br>Legal and compliance requirements (federal/state/institutional) and additional ethical protections over and above the minimal legal protections<br><br>Diversity in patients, diversity in the hospitals/practice<br>Model cards<br><br>Sophisticated sampling<br><br>Contextual SDoH - literature review, offboard code to prevent passing addresses to web<br>Individual SDoH<br><br>Review federal legal landscape<br>Review state legal and institutional policy landscapes |
| SWD | - Different training requirements for trainees from different backgrounds and at various stages of their career<br>- The lack of a structured “AI in Critical Care” curriculum based on a single comprehensive dataset<br>- Lack of culturally-aware curriculum in the area<br>- Incorporating feedback from | - Developing microlearning modules on AI in Critical Care<br>- Organzing themed datathons and workshops for early career scientists<br>- Travel support for underrepresented trainees, career mentoring, networking opportunities<br>- Developing a culturally-aware tribal college curriculum for |

| | | |
|---|---|---|
| | community stakeholders | medical AI<br>- Including citizen scientists, high schools students, and scientists at various stages of their career |
| Standards | | |
| Teaming | - Multi-project, module, multi-center program is a source of inherent complexity in milestones<br><br>- Numerous discussion points for data acquisition centers have the risk to occupy large portion of all-hands meetings.<br><br>- Strength of institution diversity is also a challenge with a variety of approaches to regulatory review and issuing agreements as well as personnel and infrastructure.<br><br>- Historic lack of diversity in AI initiatives | - Centralize communication for each module at the module lead level. Program manager network critical to guide each module and project active and upcoming milestones via module meetings, project newsletters, organized collaborative document repository, collaborative project management software, and collaborative project management software<br><br>- Dedicating time at module-specific all-hands meetings for dedicated discussion time and feedback from each module.<br><br>- Readiness dashboards; iterative transparent review of each site's progress to enable each site's team to have visibility to processes at others; teaming meeting together with site Chief Informatics Officers and privacy officers where needed; explicit survey of privacy officers to utilize teaming as a learning opportunity<br><br>- Plan for Enhancing Diverse Perspectives with explicit goals and review of progress in promoting diversity of the team.<br>- Teaming module assistance with provisioning equitable |

| | | |
|---|---|---|
| | | process for evaluating applications to CHoRUS AI Bootcamp travel grants |
| | - Multidimensional external collaboration with multiple academic societies, BRIDGE Center, NIH, FDA, and industry | - Embed plan for internal and external collaboration into BRIDGE Center Workgroups and Committees to ensure harmonization of efforts. |
| | - Collaboration in big teams tends to silence conflict and average opposing perspectives | - Set aside divergent thinking/brainstorming time that promotes idea sharing without critique |
| | - Knowledge management and knowledge integration at the team level can be accomplished only manually. I.e., knowledge graphs do not exist at the team level | - Stay diligent with minute-capturing and recording meetings |
| | - Teaming for the B2AI means different things to different people, from PEDP to PM and other acronyms in between | - Continue cross-DGP dialogues around the science of team science |
| | - Teaming protocols for role clarity, debriefing, onboarding, etc. are being developed de-novo but at times come too late as the onboarding of data acquisition sites has started | - For the future, teaming modules may have to start earlier, similar to how the Bridge center and core started their work a few months before the DGPs |
| | - Teaming with site investigators and personnel is challenging, as it requires effort which is not always funded at the site level (i.e., legal survey work) | - Explore motivations and “value” added for the sites beyond direct effort funding. |

| | | |
|---|---|---|
| Tools | Dependency on “seed” data that already conform with CHoRUS standard to start designing and developing community-user facing tools and hitting milestones | Start with data from a subset (even N=1 site) of sites. |
| | Many tool development groups at multiple institutions: Emory, MGH, MIT, Duke, UT Health, Mayo, Tufts and UFL - hard to track progress as a whole for module-leadership | Have more efforts from the project director allocated to the tooling module to help track progress; make sure each tool development team has a rep at the weekly standing meeting |
| | Communication with data acquisition and standard modules, e.g., to understand needs of some of data acquisition sites | Organize demos of our tools and 3rd party ones to data engineering personnel from the data acquisition sites and get direct feedback from them |
| | Some tool development effort depends on the selection of CHoRUS cloud data platform partner and their cloud services | Decide to pursue a short-term cloud solution while completing a comprehensive process in selecting and designing our long-term cloud solution and strategy |
| | The need to establish win-win principles– we’d like to centralize all the data, but sites want autonomy to do their own transformation, which benefits their local data environments and analytics | Working with local sites to establish capabilities and match those capabilities to solutions that can be generalized within the hybrid-cloud structure; e.g. in some cases CHoRUS developed tools may be used, while in other causes local toolsheds may be used with a wrapper for standardizing it to CHoRUS harmonizations |
| | Development of common tools that works across various sites requires adherence to basic common platforms, tools, vendors, etc. This is often easier in practise than in the real-world, where each institution may have relationships and contracts that have been developed with specific external agents; thus, the challenge for tools is how do you build something for 20% of the needs which address 80% of the | Engage local data infrastructure leadership to identify key risks and security concerns, work to address how such security concerns can be mitigated through established standards and a goal towards minimal viable products, without trying to sell loafty goals and their bells and whistles |

| | | |
|---|---|---|
| | usecases<br><br>Moving away from academic to deployable tools– challenges in efficient computational products that can be scaled and implemented in the real-world | Partnerships with industry partners who can develop extensions or mirror solutions based on the templates derived within the consortium efforts. The consortium can act as a standards institutions like the “Bluetooth Special Interest Group (Bluetooth SIG)” which industry can develop unique products on |

## Salutogenesis DGP:

- Based on our discussions last time, I asked the Salutogenesis modules to share the challenges they identified when writing their proposals for the DGP FOA and what solutions they proposed to address them.

| Module | Challenges | Proposed solutions |
|---|---|---|
| Data Acquisition | AI algorithms can be extremely sensitive to many biases that can originate from training datasets (e.g. unbalancing of demographic characteristics).<br><br>AI algorithms often require a large sample size which can be cost prohibitive. | Need data sharing solutions to allow training datasets from various sources<br><br>Continue to develop training methods that can rely on small sample sizes and/or synthetic data. |
| Ethics | Dynamic consent for evolving research<br><br>Legal and Policy alignment with FDA and HIPAA<br><br>Data ownership; FAIR vs. CARE<br><br>Extended identification (family members/ isolated populations)<br><br>Prioritization of CABs<br><br>How to mitigate and measure bias and discrimination<br><br>Can AI be used to improve health equity?<br><br>Privacy. Deciding how the data should be disseminated and who should have access.<br><br>Create a culture of ethical inquiry within large interdisciplinary projects<br><br>Legal liability | Build a platform for data and study progress tracking so participants can affirm or revoke their consent in real-time.<br><br>Consents need to be reviewed by ELSI experts (CEER http://www.genome.gov/Funded-Programs-Projects/ELSI-Research-Program/Centers-of-Excellence)<br><br>Re-consideration of Bayh-Dole and America Competes act with regard to ML/AI DGPs<br><br>Better and more precision definitions of “community”<br><br>CABs input have to be formalized by contract or some other means of “partnership” and control.<br><br>Prioritization of Ethics in all modules. Pre-agreed, formalized methodology and commitments to incorporate ethical recommendations<br><br>Needs assessment of ELSI awareness within and across modules. Results will drive consultation priorities and capacity |

| | | |
|---|---|---|
| | | building education<br><br>Planned onboarding of community advisors that includes education about ML/AI, AI READI, data collection/management, risks/benefits for downstream data use. |
| SWD | The AI workforce lacks diversity, and this may hinder recognition of biases and/or identification of unintended consequences that have downstream negative effects for certain populations.<br><br>There can be a disconnect between individuals working in the health/biomedical sciences domain and those coming from a computational background, resulting in AI algorithms that may not translate well into clinical practice. | By developing a training program and engaging in widespread recruitment efforts, we hope to train a more diverse AI workforce.<br><br>The multi-disciplinary make-up of the project team will foster synergistic collaborations among individuals from a wide range of backgrounds. Also, we plan to develop educational modules and training tools that will be broadly accessible to people from different disciplines. Our interns can also serve as tangible “bridges” between different groups as they engage in interdisciplinary projects. |
| Standards | A critical ingredient for the success of robust AI algorithms is the availability of data in a consistent representational standard such that not only the trained AI models work across a multitude of disparate datasources but also generalize to other clinical infrastructures. | We propose to build a common pathway for each data type to have a consistent representational standard relevant for each domain of data type and to encourage industry and the wider international standards community to adopt any extensions that are developed. |
| Teaming | AI algorithms often reflect the interests of the computer scientists who create them and not those of a broader range of stakeholders, including the communities that contribute data and those that will use the algorithms. | By iteratively generating AI algorithms in collaboration with members of the communities that contribute data and those that will use them for research purposes, the algorithms are more likely to be useful in practice. |
| Tools | FAIR data practices are widely promoted and encouraged to optimize the reusability of data by humans and machines (AI/ML) but researchers collecting data | We believe automated tools that assist researchers can address these challenges. Accordingly, we propose to develop a web platform called fairhub.io that will simplify data management, curation, and sharing |

| | | |
|---|---|---|
| | typically lack knowledge, training, and support for making their data FAIR and AI-ready | according to FAIR guidelines through intuitive user interfaces and automation tools. Our goal is to make fairhub.io usable by anyone irrespective of background, training, knowledge, or expertise. |

- When Dataset will be made available? We plan on releasing a first set of data from our pilot study at the end Year 1 (i.e. by 08/31/2023)

## VOICE DGP

| Module | Challenges | Proposed solutions |
|---|---|---|
| Data Acquisition | Many limitations to Voice AI research remain including limited size of datasets available of questionable acoustic quality, lack of data labelling for more than one condition, lacking data source and data modality diversity.<br><br>Moreover, Voice is considered a biometric identifier subject to HIPAA regulation, limiting multi-institutional collaborations due to ethical considerations – ultimately hampering the creation of accessible, robust, and diverse voice datasets. | For voice to emerge as one of the biomarkers of health, there is a pressing need for large, high quality, multi-institutional and diverse voice databases linked to other health biomarkers from various data of different modality (demographics, imaging, genomics, risk factors, etc.) to fuel voice AI research and answer tangible clinical questions. Such an endeavor is only achievable through multi-institutional collaborations between voice experts and AI engineers, supported by bioethicists and social scientists to ensure the creation of ethically sourced voice databases representing our populations. |
| Ethics | 1. Privacy: How to protect confidential information of research participants; how to ensure they are informed about privacy and identifiability concerns when they consent; how to ensure data sharing is done responsibly and who should have access.<br><br>2. How to mitigate and measure bias and discrimination | 1. Governance framework that ensures data protection in terms of privacy, as well as responsible trustworthy data sharing approaches. Ethics module input into design of consent.<br><br>2. Engage stakeholders to identify best practices to enhance equity diversity and inclusion in recruitment and in AI application, to ensure voice as a biomarker is non-discriminatory and enhances equity. |

| | | |
|---|---|---|
| SWD | 1. The workforce does not represent the communities affected well, especially in Voice AI, so increasing representation and diversity is crucial for ethical and fair AI<br><br>2. AI affects many people and groups, but developing the skills needed to understand it – especially potential biases and increasing trust – can be challenging for key constituencies | 1. Partnering with our ethics team, with AIM AHEAD, and other initiatives, we will work to engage key communities to improve recruitment from those communities of learners at all levels.<br><br>2. We have a competency framework that is adapted to different learner needs and goals and plan to use it to identify key gaps and assemble flexible curricula that can better meet these needs. |
| Standards | 1. No standard format exists for storing voice waveforms which facilitates modern artificial intelligence approaches.<br><br>2. Medical data is inherently multimodal, and clinically deployable algorithms will need to acquire and process this data. | 1. Build a new standard which optimizes for efficient storage, access, and learning from waveform data.<br><br>2. Adopting the Fast Healthcare Interoperability Resources (FHIR) standard to interact across modes of data (imaging, genomics, waveforms) with a consistent protocol while enabling storage of each mode of data in a domain specific optimal format |
| Teaming | 1. There is a huge gap between existing technology developed by tech industry or AI scientists and solutions that are actually used clinically to help clinicians and patients.<br><br>2. Many teams looking into voice biomarkers do similar work in silos. They are confronted to the same issues in terms of lack of standards, resource intensiveness of data collection, ethical issues, legal implications | 1. Pair up tech developers with clinician scientists and favor collaborations BEFORE products are developed. Ones products are developed, fund projects who aim to integrate technology into clinical care as opposed to focusing only on development<br><br>2. Develop a network to foster collaborations and sharing of ideas and resources between teams with a common goal |
| Tools | 1. Although smart devices (phones, tablets & wearables) have become ubiquitous nowadays, there is a lack of reliable and clinical grade tools to utilize their capabilities to collect | 1. We are building standardized tools for smart devices that will support study enrollment, consent and data collection on a multitude of devices that can run on a clinical trial setting |

| | | |
|---|---|---|
| | various data from sensors (i.e. voice, fitness, EHR stored via FHIR) and transmit it securely supporting clinical studies.<br><br>2. Although medicine is experiencing its AI renaissance at the moment, there exist no standardized tools/platforms to train or validate AI models without having explicit access to the datasets creating a privacy barrier that hinders AIs true potential. | and collect data while in clinic or remote. We also plan to connect these tools with industry standards that will support scaling of trials but also extending these tools beyond the scope of this project.<br><br>2. We are building an open-source Federated Learning platform that will remove the privacy barrier allowing models to be trained or validated without sacrificing on privacy or on creativity. The platform will empower researchers to train AI models across different datasets essentially breaking silos but not privacy. |

## CM4AI DGP

| Module | Challenges | Proposed solutions |
|---|---|---|
| Data Acquisition | AI algorithms can be sensitive to many biases that can originate from training datasets. We're generating three different datatypes for the same cells in three different sites. | Generate large and standardized datasets that will be integrated by the tools module for robust CM4AI maps |
| Ethics | CM4AI opens a host of scientific and medical possibilities by leveraging fine-grained understanding of fundamental biology and genomics, but also through advanced ML/AI techniques. As a result, advances in cell maps analyzed by visible ML offer possibilities that could not be expected until recently. This increased level of precision in the understanding of biological mechanisms underlying predictions by medical AI systems, notably through the crossing of techniques and data-types, also generates entirely new Ethical, Legal and | A mixed-method ethical inquiry that includes sound conceptual and participatory works is being employed to guide the development of CM4AI. By engaging with various stakeholder groups and affected communities early on, specific ethical concerns that arise from functional genomics, visible ML and the development of n-of-1 medicine can be identified and addressed. The aim is to offer normative and design solutions that provide informed ethical guidance, tailored to the unique realities of stakeholders, throughout the entire continuum of CM4AI development, from basic biology to downstream clinical applications. Ethical decision-making, trustworthiness, and social responsibility are |

| | | |
|---|---|---|
| | Social Issues (ELSI) that we will tackle in an anticipatory manner and anchored in participatory engagement. | promoted through this approach. |
| SWD | There are a variety of challenges to developing skills for biomedical ML/AI applications. Because of the inherent multi-disciplinary nature of the work, individuals require an understanding of the data, data standards, and domain-specific tools in addition to having ML/AI development expertise. In addition, having an understanding of the ethical considerations regarding data acquisition and the implementation of ML/AI algorithms is critical when using these tools in biomedicine. | To succeed in developing the next generation of biomedical ML/AI researchers, it will be necessary to recruit diverse groups of individuals and provide them with the resources needed to drive discovery. The development and release of high-quality datasets with associated metadata annotations, tools, and documentation are needed to provide a strong foundation that can be used by trainees and investigators to create high-quality output that will improve health and health care. |
| Standards | Balancing: robust implementation of the most relevant standards, supporting data and software FAIRness across the pipeline from data acquisition to data engineering and analysis to final results - while maintaining flexibility and agility to align with and support an experimentalist culture in the large and geographically diverse labs. | Staying aligned with and appreciating the lab culture while showing how standardization can help them. Frequent meetings with key contributors in other modules. Build the lightest-weight tools possible to do the job. Document progress and strategy and share with other groups. |
| Teaming | We are working in a large team with different geographical, cultural, institutional, disciplinary, and personal contexts. It is challenging to collaborate across so many boundaries and work productively from day one. | Need to learn to recognize our differences and adapt a common set of working etiquette, communication protocols, and organizational structures driven by a core set of values. We focus on a science-driven culture and a diverse inclusive meeting environment. |
| Tools | Our tools will only become widely used if we have identified the communities who would benefit, understand their needs, and design to meet those needs.<br><br>To satisfy the users in those communities, the tools must be | Tools Module Challenges and Approa…<br><br>Briefly, we will adopt a structured, needs-driven development process that includes early prototypes and internal use.<br><br>In a second phase, we will work with Workforce and Teaming to define our target |

| | | |
|---|---|---|
| | robust and easy to use (including well-targeted documentation.)<br><br>To reach those users, the tools must be findable and “advertised” to the target communities.<br><br>To increase the likelihood of adoption of our tools by the community, we need to demonstrate their utility by application to example biological problems. | users and find early partners/adopters. One part of this strategy is to find synergy with our DEI efforts, reaching out to researchers at HBCUs for potential student training and projects or broader collaborations.<br><br>Apply our tools to example biological problems, and document this effort in papers and by dissemination of detailed descriptions (eg, via papers, conferences, web sites, databases, youtube, …). |

## Ethical and Trustworthy AI Core:

| **Challenge** | **Proposed Solutions** |
|---|---|
| Confusion around when ethical challenges should be define | Define an AI/ML development and use lifecycle. Provide illustrations of where and why ethical issues arise as well as how to address them before it is too late. |
| Disconnect between AI/ML researchers and the people to whom data corresponds and the users of the technology | Develop a principled approach to human-centered AI (HCAI), which integrates diversity in team members, cultural perspectives, and values of those who will benefit from the resulting technologies |
| As data becomes increasingly more detailed, there are increasing privacy concerns | Investigate how new approaches to distributed learning, such as federate learning, split learning, and multiparty computation can enable AI/ML without sharing individual level records. |
| How do you establish consent for data collection if use cases are unknown? | Investigate potential ways to integrate data contributors or their representatives into the AI/ML research and development process. |

## Tools Core:

| **Challenge** | **Proposed solutions** |
|---|---|
| Existing tools do not always generalize well to new data, compute environments or problem domains | ● A cross-functional group of algorithmic, software and benchmarking experts will provide "consulting" support<br>● Community-driven articulation and dissemination of best-practice guidelines<br>● RFPs to provide focused tool development support |
| Tool development is diffuse, with duplicative efforts across consortia and different scientific communities | ● Systematically survey, annotate and publicize tool-chains for each DGP within Bridge2AI<br>● Identify and foster partnerships across consortia<br>● Establish best-practices and incentives for containerization |
| Tool selection is made difficult by a unknown performance in most problem domains | ● Design and execute regular benchmarking exercises on tool-chains critical to Bridge2AI<br>● Link benchmarking directly to tool-chain documentation<br>● Establish benchmarking best practices |
| Some important areas of tool development are not considered "high-impact" | ● Identify gaps in tool-development through a intra-Bridge2AI tooling survey and analysis of public tool development & support patterns |

| | |
|---|---|
| or “exciting” and receive less attention. | ● Deploy target resources to directly support critical and poorly maintained or architected tools |
| Tool developers are often unaware on how to maximize the utility and performance of the tools they create. | ● Create a standardized tool registry linked to benchmarking, software, documentation, etc.<br>● Coordinate best-practices with other major consortia<br>● Establish workshops and training material focusing on both best practices in tool development and on maximizing utilization of optimal tool-chains |

## Standards Core:

| **Challenge** | **Proposed solutions** |
|---|---|
| Data standards are not one-size-fits-all, with multiple standards that may be used for the same data in different contexts | ● Form cross-consortium work groups specific to common data domains to survey and compare standards<br>● Collaboratively develop recommendations for standard selection within a data domain<br>● Identify inter-standard conversion tools for cross-standard data exchange (when needed) |
| Data standards are not inherently interoperable, with differences in data representation, permissible values, and semantic context | ● Promote use of a consistent schema definition language<br>● Establish common vocabularies for common data types<br>● Reference or define permissible data values within schemas and vocabularies<br>● Work with standard developers to understand schema contexts and create semantic mappings |
| Data standard selection is driven by consideration of whether data fits, instead of how data is to be used | ● Develop standard selection guides for different data use cases (e.g. storage, readability, machine learning)<br>● Summarize and document key differences between standards within a data domain |
| Data standards may not exist for a particular type of data, or may be a poor fit for data within a particular application | ● Identify gaps in existing standards by creating and evaluating data testbeds<br>● Evaluate consistency and fidelity of data expressed in available data standards<br>● Use standardized schema definition framework to construct new data schemas as needed. |
| Data standards are misused by implementations to make data fit | ● Define criteria for data suitability in AI/ML applications<br>● Create a standards evaluation dashboard for scoring how well these datasets meet data suitability criteria |

## Teaming Core:

| Challenge | Proposed solutions |
|---|---|
| Teaming & transdisciplinary science at scale is the way forward, yet people remain attached to structures and ways of working that do not serve that purpose | • Interesting that this was just released by NAP: https://nap.nationalacademies.org/catalog/26863/transforming-research-and-higher-education-institutions-in-the-next-75-years<br>• Meet people where they are and advance their willingness to engage in teaming activities by solving their problems. Sometimes this requires waiting until a problem is recognized.<br>• Engage in teaming activities and structures that respond to data-driven needs. Collect these data from the Consortium and show these data to the Consortium to support your suggestions. |
| Operations of large, distributed consortia quickly become overwhelming. | • Take advantage of existing tools<br>• Automate as much as possible<br>• Understand the needs of the consortium and design processes that support existing workflows and goals<br>• Develop a communication plan that clearly articulates who needs to know what<br>• Transparent task assignment and tracking<br>• Develop clear roles, responsibilities, and process before work begins<br>• Face-to-face kickoff and social time - builds trust (this helps with many of the listed challenges here)<br>• Dedicated staff for intersections of groups |
| Decision-making and discussion must be carefully balanced so that voices are heard and decisions can be made in a timely fashion. | • Collaboratively establish SOPs on decision-making and timing of discussion will provide structure and manage expectations.<br>• See some solutions for operating large team size<br>• Form smaller groups with representatives that have decision-making power to perform specific tasks. Be transparent about the decisions these smaller groups will be making. Be sure that the size of the group matches the complexity of the task and the time remaining to achieve consensus.<br>• Create the expectation that all feedback will be considered, but may not be reflected in the final decision. |
| It is easy for quieter individuals to get lost in a large consortium and only a few loud voices are heard.<br>Achieving inclusion at scale. | • Create protocols and formalized structures that create space for everyone's contribution and reflection<br>• Make space for verbal, written, or asynchronous feedback to capture all voices<br>• Articulate workflows so people can visualize how, when and where they can intersect–either to contribute or to learn<br>• Begin with smaller groups so people can engage in meaningful conversations, questioning and planning. People want to hear others' ideas as a way of forming relationships to |

| | |
|---|---|
| | novel work, and to new collaborators. Building self-confidence in this world of digital transformation is supported when people can hear and see others negotiating it as well. |
| Teaming roles and responsibilities are undefined and the impact of Teaming activities on scientific output is unrecognized; thus, Teaming activities are deprioritized. | ● Define Teaming roles and responsibilities<br>● Find ways to build “the R’s” - Roles & Responsibilities, Relationships, Resilience (which includes Response to conflict and setbacks), Results<br>● Focus on easy wins early in the collaboration that can demonstrate teaming impact<br>● Leadership can create the expectation and build the incentive structures needed to prioritize the initial teaming work needed to lay the foundation for a smooth and fruitful collaboration |
| Large consortia contain a diversity of communication styles and preferences that may require additional work to mesh. | ● Support multiple methods of synchronous and asynchronous communication<br>● Set the expectation that communication happens freely as needed but within the scope of the roles, responsibilities, and process<br>● Appreciate that multicultural, personality, family backgrounds, and prior work experiences all shape how people communicate – create ways of surfacing these |
| Innovative research requires an atmosphere of trust and this takes time to build, making it difficult for new collaborators to quickly engage in impactful work. | ● Create a space for people to articulate the conditions and behaviors they need to practice, and see in others, in order to feel comfortable, and willing to take some risks, try new things, and work with unfamiliar people.<br>● Create opportunities to work together in-person in low risk settings |
| Diverse teams create better results, but require greater operational support to overcome the additional barriers to communication and trust | ● Be intentional about building communication and trust<br>● It will take time. Make the time.<br>● Highlight the importance of diversity for strong teams: different perspectives enhance creativity and problem-solving<br>● Create a culture that supports hearing multiple voices (create psychological safety) |

General “solutions”
*address challenges early and often
*build transparency to your tools
*clear effective communication systems
*consist messaging about overarching goals and reasons for teaming
*You can lead from anywhere on the team
*infrastructure is key: offload burden on researchers

## Skills and Workforce Development Core:

| **Challenge** | **Proposed Solutions** |
|---|---|
| The field of AI/ML is moving forward at such a pace that many individuals involved in biomedical sciences haven't had the necessary education to understand the promise, the potential pitfalls, and the appropriate usage | well organized curriculum with individual educational modules covering ethical, technical and practical aspects of AI/ML from ethical considerations, to data generation and evaluation, to potential health care and biomedical use cases. The videos of these lectures will be archived and disseminated widely online. |
| Lack of engagement between the individuals generating and developing algorithms from datasets (data scientists) and the individuals using the algorithms in practice (clinicians) | targeted mentoring programs for physicians who are leaders in transforming AI-supported clinical practice. These mentoring programs are specifically designed to help them understand how to maneuver the interface of these tools and to gain their feed as well as involvement in actual design and generation of AI/ML workflows/tools. |
| The current AI/ML field lacks diversity. Diversity is vital to ensuring widespread applicability and equity in the algorithms developed. | a team based mentoring program targeting diverse individuals in medicine. This program will offer mentoring teams consisting of a data scientist, a clinician involved with AI/ML, and an diverse leader in a structured curriculum. |