

Advancing method benchmarking & data sharing through crowd-sourced competitions

James A. Eddy,
Director of Informatics & Biocomputing
Data & Tooling @ Sage Bionetworks

SageBionetworks
Better Science Together

Our ITCR Project at a Glance

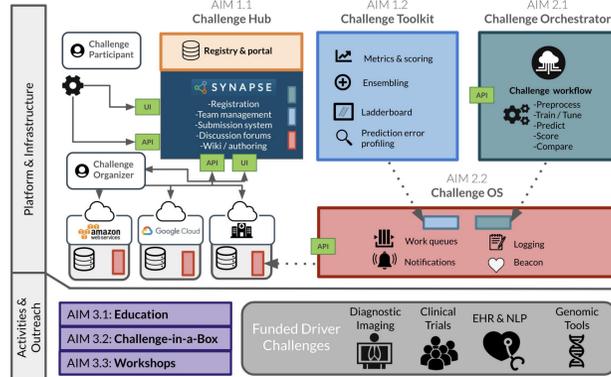
Crowd-sourced competitions have transformed biomedical research by incentivizing the coalescence of communities around timely and difficult problems. These communities have created new standards and benchmarks, and developed innovative solutions to dozens of pressing research problems. To enable the next generation of Challenges, we aim to expand our existing infrastructure into an innovative platform for rapid, rigorous and scientifically-valid Challenge-based assessments.

We aim to create a Challenge framework that (1) provides a discovery engine for biomedical Challenges and their outputs, (2) contains innovative tools that enhance and streamline Challenges, and (3) is scalable and supports distributed assessments using private and sensitive data. These will be married to a suite of educational tools, best-practices guidelines, and workshops and conferences to expand the solver community and support the organization of independent challenges to maximize impact on biomedical research.

AIM 1: Develop a community hub and benchmarking toolkit for biomedical challenges.

AIM 2: Develop portable software, services for distributed benchmarking on sensitive and protected data.

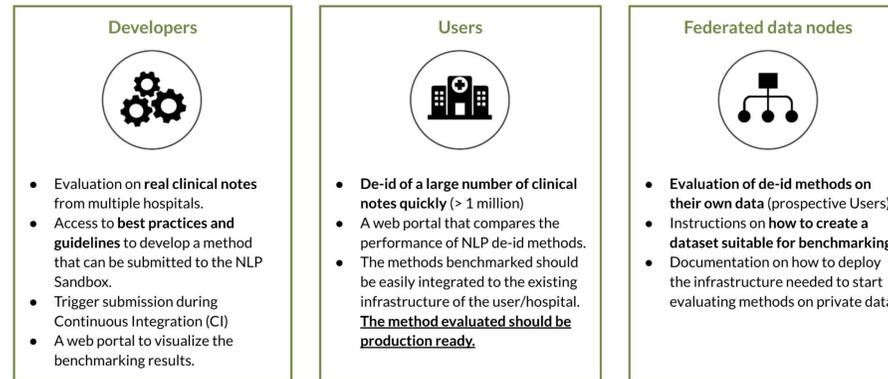
AIM 3: Expand the biomedical challenge community through improvements in education, outreach, and empowering the organization of independent challenges and benchmarking projects.



Cloud Sandbox Framework for NLP Tools

Natural language processing, or NLP, is a technology used in many ways to help computers understand human language. This is particularly impactful in biomedical research, where hospitals have millions of unstructured notes they need to de-identify before sharing with researchers. Manually de-identifying them would put significant strain on healthcare systems, presenting an excellent use case for the application of NLP.

[NLP Sandbox.io](https://nlp.sagebionetworks.io) is one of the first tool-benchmarking platforms that securely connects developers to healthcare data providers. The platform streamlines your development process and the assessment of tools that are re-usable, reproducible, portable and cloud-ready. The NLP Sandbox adopts the model-to-data architecture to enable NLP developers to assess the performance of their tools on public and private datasets. When a developer submits a tool, data partners automatically download the tool and evaluate its performance against their private data. This architecture enables our partners to fully control their data and ensure no sensitive information leaves their secure environment.



EHR DREAM Challenges

Healthcare institutions are attempting to move away from a rules-based approach to clinical care, toward a more data-driven model of care. To achieve this, machine learning algorithms are being developed to aid physicians in clinical decision making. However, a key limitation in the adoption and widespread deployment of these algorithms into clinical practice is the lack of rigorous assessments and clear evaluation standards. A framework for the systematic benchmarking and evaluation of biomedical algorithms – assessed in a prospective manner that mimics a clinical environment – is needed to ensure patient safety and clinical efficacy.



The rapid rise of COVID-19 has challenged healthcare globally. Due to the importance and emergent need for better understanding of the condition and the development of patient specific clinical risk scores and early warning tools, we have developed a platform to support testing analytic and machine learning hypotheses on clinical data without data sharing as a platform to rapidly discover and implement approaches for care. This Challenges aims to identify risk factors that lead to a positive test utilizing electronic health recorded data mapped to the OMOP Common Data Model.



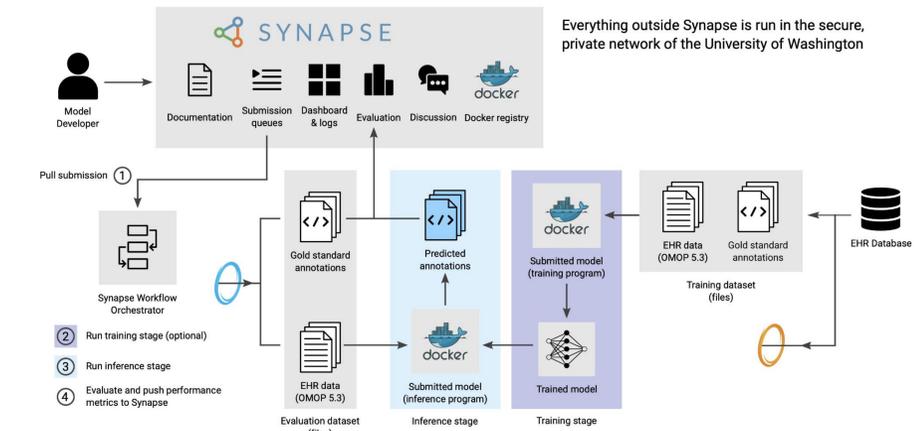
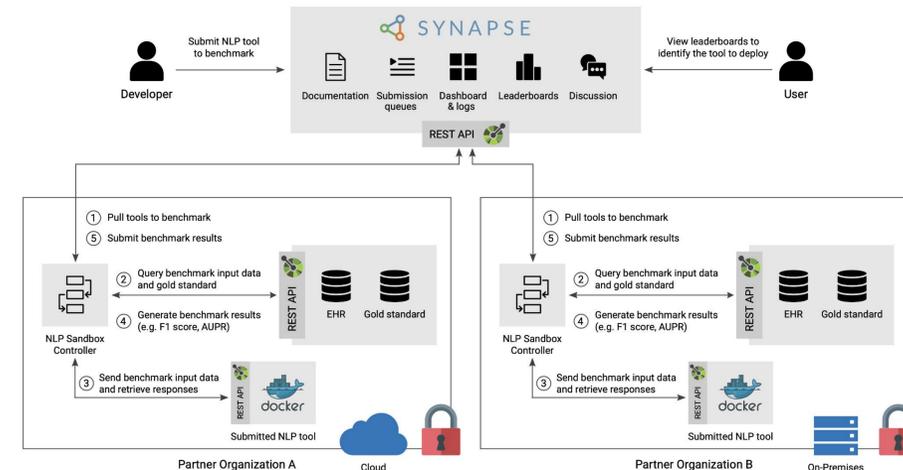
Design of a Next-Gen Challenges Platform

We undertook a principled design process, following the NIST model for enterprise architecture and partnering with the core engineering and user experience (UX) design teams at Sage, in order to develop a roadmap and specifications for a next-generation challenges and benchmarking platform. With this roadmap established and the proposed architecture documented, we have started prototyping individual services and components towards a minimum viable product.

Persona Type	Challenge Participant
Quote	"This challenge aligns with the work I'm conducting in my PhD/PostDoc or during my free time. I will participate in this challenge to 1) evaluate objectively the performance of my method and 2) earn the challenge incentives (publication, bragging rights, cash award)"
Goals	<ul style="list-style-type: none"> Win the challenge Advance science Learn (e.g. new technology, gain experience with specific data)
Needs for Success	<ul style="list-style-type: none"> Clear instructions (e.g. challenge questions, timeline, incentive eligibility) Submission example Basic requirements for joining Communication channel with organizers and other participants Way to get help with clear expectations on support Access to data and/or compute resources
Pain Points	<ul style="list-style-type: none"> Synapse (not primarily designed for participating in challenge) Not having clear picture of what and how to do it Lack of standardization for challenge Complexity of submission (e.g. Docker in model-to-data challenges; but today's participants start to get used to it) Technical issues/limitations (e.g. submissions run slower than expected due to incomplete CPU isolation, delayed submissions when submission rate exceeds submission processing capacity) Changing challenge timeline (e.g. delayed launch of the challenge or release of the results because the organizers underestimated the workload)
Motivators	<ul style="list-style-type: none"> Being on challenge manuscript (publication) Cash award Being part of new scientific communities Show off your skills / build resume Explore new types of data and technologies (learner) Sometimes asked by their company to participate Free compute resources
Software Use	<ul style="list-style-type: none"> Docker (model to data), Python, R, C/C++, deep learning framework (Caffe, Keras, Tensorflow)
Outputs for Success	<ul style="list-style-type: none"> Publication Increased visibility on professional platforms (e.g. LinkedIn) and receive job propositions Made the best-performing method(s) widely available (e.g. new reference solution in the community, commercialization)
Years of Experience	<ul style="list-style-type: none"> All levels (post docs, undergrads, retirees) - Not a key factor for participation
Education	<ul style="list-style-type: none"> BA+, high schoolers (need to be 13+ years old to use Synapse)
Example	<ul style="list-style-type: none"> PhD student, PostDoc, researcher, engineer, retiree Lack of qualifications having submitted a model in the final phase of the Patient Mortality Prediction DREAM Challenge

We designed a common development model for building NLP tools and benchmarking their performance on public and private datasets. This is designed to be a stepping stone towards the development of Sage Bionetworks next-gen challenge platform.

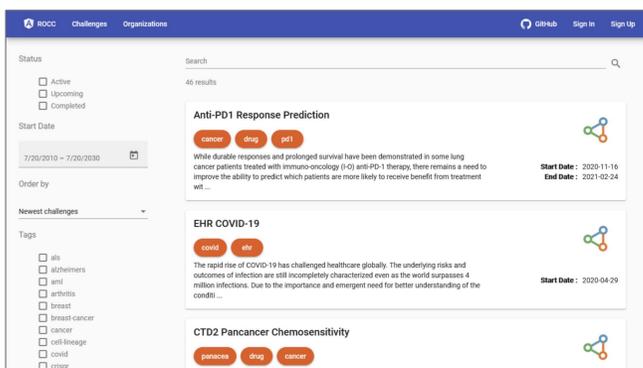
The benchmarking infrastructure developed (below) is deployed by partner organizations to benchmark the performance of NLP tools on their own private data (currently EHR). The only information leaving the secure environment controlled by partner organizations is the performance of the tools, which is then returned to the developer and published in leaderboards.



Registry of Open Community Competitions

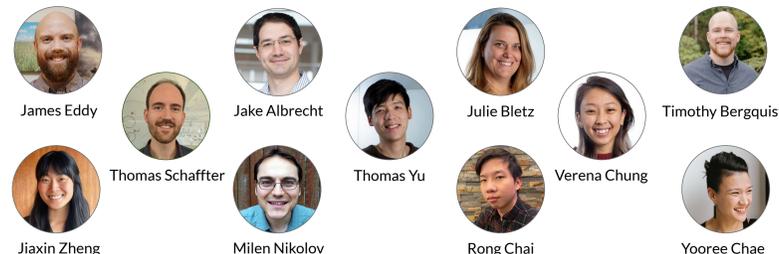
The Challenge experience is currently cumbersome for all stakeholders involved. Participants encounter a huge pain-point in discovering Challenges they are interested in, while Organizers struggle to find a streamlined solution to reaching a wide audience.

The Registry of Community Challenges (ROCC) strives to address some of the friction in the Challenge experience by providing a centralized hub to reduce participant search cost and increasing Challenge visibility. Think of the ROCC like an EventBrite for the scientific benchmarking space, empowering both participants with the most up-to-date information about relevant Challenges, as well as organizers with standardized Challenge event templates.



Meet the Team

Sage Bionetworks is a nonprofit biomedical research organization with a vision to promote innovations in personalized medicine by enabling a community-based approach to scientific inquiries and discoveries. We strive to activate patients and to incentivize scientists, funders, and researchers to work in fundamentally new ways in order to shape research, accelerate access to knowledge, and transform human health.



RSNA-ASRN-MICCAI BraTS Challenge

The Brain Tumor Segmentation (BraTS) Challenge celebrates its 10th anniversary, and this year is jointly organized by the Radiological Society of North America (RSNA), the American Society of Neuroradiology (ASNR), and the Medical Image Computing and Computer Assisted Interventions (MICCAI) society.

The RSNA-ASNR-MICCAI BraTS 2021 challenge utilizes multi-institutional multi-parametric magnetic resonance imaging (mpMRI) scans, and focuses on (Task 1) the evaluation of state-of-the-art methods for the segmentation of intrinsically heterogeneous brain glioblastoma sub-regions in mpMRI scans. Furthermore, this year's challenge also focuses on (Task 2) the evaluation of classification methods to predict the MGMT promoter methylation status at pre-operative baseline scans. Participants are free to choose whether they want to focus only on one or both tasks.



RSNA ASNR MICCAI SageBionetworks