

Sage Bionetworks teams up with Amazon Web Services Open Data to enable access to in-demand biomedical datasets

The collaboration accelerates research by combining high-capacity data storage with Sage Bionetworks' reliable governance and scientific expertise.



Sage Bionetworks



SEATTLE , WASHINGTON , UNITED

STATES , May 15, 2024 /EINPresswire.com/ -- While scientists are now encouraged, and in many cases, required to share research data, the safe transfer and use of this data by others remains significantly limited - an issue that some in the research community refer to as Open Data in Appearance Only.



Our top priority is to provide scientific communities with access to high-quality biomedical data for reuse in exciting ways."

Sarah Chan, PhD, Principal of Strategic Partnerships at Sage Bionetworks

Sage Bionetworks and Amazon Web Services (AWS) have initiated a collaboration that will increase public access to popular datasets so that they can be readily reused to accelerate biomedical research. The new agreement will see enhanced data infrastructure support for Sage Bionetworks' flagship data sharing and management platform, Synapse, allowing users to access open data in their workflows.

Specifically, the enhanced infrastructure will support the

new Synapse Basic Hosting Plan, as well as the Self-Managed Plan, which is designed to serve researchers adhering to the [National Institute of Health's Data Management and Sharing Policy](#).

"Our top priority is to provide scientific communities with access to high-quality biomedical data for reuse in exciting ways," says Sarah Chan, PhD, Principal of Strategic Partnerships at Sage Bionetworks. "But data management does come at a cost. Therefore, we actively seek out partners who can help advance our unique vision of open science."

The collaboration is part of the AWS Open Data Sponsorship Program, which provides secure storage and egress for publicly available, high-value and cloud-optimized datasets. Sage

Bionetworks then curates these datasets and offers technical expertise to scientists, ensuring that the data in Synapse is findable, accessible, interoperable and reusable (F.A.I.R).

While these high-value datasets are listed through the Registry of Open Data on AWS, Sage Bionetworks provides governance controls, ensuring the data can be shared ethically and responsibly. With these safeguards in place, Synapse users can then access and download data files for open reuse and transformation.

Early Success Stories

The first dataset to benefit from the collaboration was the [UK Biobank Pharma Proteomics Project \(UKB-PPP\)](#), published in the journal Nature in October 2023. This landmark initiative between the UK Biobank and 13 biopharmaceutical companies characterized the plasma proteomic profiles of over 54,000 participants to advance the development of biomarkers, predictive models and therapeutics.

The Nature article has been accessed over 70,000 times since its publication, ranking in the 99th percentile for all journal articles published around the same time. This dataset alone has driven 900 terabytes (TB) of data egress from the Synapse platform, suggesting that scientists are eager to access this open-source dataset to advance ongoing research.

The UKB-PPP dataset listing on Synapse was closely followed by the addition of a comprehensive, [single-cell atlas dataset](#) of human blood, spanning ~2 million cells from healthy individuals across the human lifespan. The study, published in the journal Immunity in December 2023, has had its dataset accessed by more than 200 unique Synapse users, providing the scientific community with novel insights into the age-dependence of immune cell populations.

“This collaboration has expanded access to large datasets, and we are optimistic that our data will encourage a novel exploration of human aging,” says Marina Terekhova, MD, staff scientist at Washington University School of Medicine in St. Louis and first author of the single-cell atlas study.

These milestone datasets make up a fraction of the more than 2 petabytes (PB; equal to 1,000 TB) of data already supported on the Synapse platform, which has served over 100,000 data users across multiple disciplines.

“Research participants place deep trust in the scientists, data brokers, and developers who use their data to advance scientific knowledge,” says Ann Novakowski, MPH, Associate Director for Governance Innovation at Sage Bionetworks. “We take very seriously our obligation to promote safe, ethical, and responsible use of that data. We hope that this data reuse can bring scientific communities together, driving meaningful discoveries that can transform patient lives.”

Explore Synapse by heading to synapse.org

ABOUT SAGE BIONETWORKS

Sage Bionetworks is a non-profit health research organization based in Seattle, Washington. We guide responsible data sharing and reuse, benchmark scientific methods and results, and empower participants to be active partners in research. Learn more at sagebionetworks.org

To actively partner with Sage Bionetworks, contact us at partnerwithus@sagebionetworks.org

Media contact:

Drew Duglan, PhD

Principal, Communications and Marketing, Sage Bionetworks

drew.duglan@sagebase.org

Drew Duglan

Sage Bionetworks

+1 858-247-9110

[email us here](#)

Visit us on social media:

[Facebook](#)

[Twitter](#)

[LinkedIn](#)

[Instagram](#)

[Other](#)

This press release can be viewed online at: <https://www.einpresswire.com/article/711254713>

EIN Presswire's priority is source transparency. We do not allow opaque clients, and our editors try to be careful about weeding out false and misleading content. As a user, if you see something we have missed, please do bring it to our attention. Your help is welcome. EIN Presswire, Everyone's Internet News Presswire™, tries to define some of the boundaries that are reasonable in today's world. Please see our Editorial Guidelines for more information.

© 1995-2024 Newsmatics Inc. All Right Reserved.