



Data Sharing Agreement

Last updated: December 10, 2025.

Project Description

SecureBio is building a *Predictive Biology Benchmark* to test how well AI models can generate hypotheses and predict biological experiment outcomes. You are contributing data to help us build and evaluate these benchmarks. **This document explains how your data will be used and protected.**

We recognize that the data you collected are extremely valuable, and keeping them private is essential for academic research to function. The **measures** below are designed so that the benchmark can be used to measure AI advancement, while maintaining the novelty of your data so that you can report on new discoveries at the time of your choosing.

1. What counts as “Contributed Materials” and “Derived Evaluation Datasets”

“Contributed Materials” means any datasets, manuscripts (including drafts), grant applications, metadata, and related documents provided by you to SecureBio under this Agreement. **This does not include any datasets or materials that are publicly available.**

We will use these Contributed Materials to construct “Derived Evaluation Datasets” — structured summaries of the experiments and their measured outputs (for example, in the form of question-answer pairs). We will then use the Derived Evaluation Datasets to evaluate AI model performance.

2. You keep ownership

You retain full ownership and control of everything you share. This agreement does not transfer intellectual property rights to SecureBio.



3. How your data will (and won't) be shared

We will use Contributed Materials only to create, share and publish **aggregated benchmark performance metrics**. Performance metrics will be reported at a **level of generality that prevents inference or reconstruction of unpublished scientific conclusions** contained within the Contributed Materials.

SecureBio will not publish or disclose to any third party:

- (a) the Contributed Materials,
- (b) Derived Evaluation Datasets, or
- (c) scientific conclusions or specific per-experiment results that could reasonably be used to identify or infer the underlying scientific findings of individual Contributors

4. You notify us if/when data becomes public

Please let us know if any of your Contributed Materials are published or otherwise made publicly available, so we can exclude or flag them appropriately in future evaluations.

5. Credit and Authorship

If an academic paper or report is published using your Contributed Materials, you will be listed as a co-author. If no academic paper or report is published, you will still be credited in presentations or other venues where benchmark results that made use of the Contributed Materials are disclosed (e.g. on our website or social media).

6. Data Security

SecureBio shall implement reasonable administrative, technical, and physical safeguards to protect the confidentiality and integrity of the Contributed Materials, including access restrictions to authorized personnel only

7. Use Limitation

The Contributed Materials shall be used exclusively for the Predictive Biology Benchmark project described in this Agreement. SecureBio will not repurpose, transfer, sublicense, or use the Contributed Materials for any other research project, commercial application, or purpose without the Contributor's prior written consent.



SecureBio

8. Entire Agreement

This Agreement constitutes the full understanding between the Contributor and SecureBio regarding the use of the “Contributed Materials”. The terms here supersede all prior or contemporaneous agreements.

By signing this Agreement and/or submitting Contributed Materials through SecureBio’s designated form, you acknowledge that you have read, understood, and accepted the terms of this Agreement, agree to collaborate with SecureBio by providing materials for use in this project in accordance with the terms set forth herein, and represent that you have the authority to share the Contributed Materials described in Annex A.

Name: _____

Role: _____

Institution: _____

Date (dd/mm/yyyy): _____

Annex A. ID: _____

Signature: _____



Annex A. Description of Contributed Materials

Annex ID: _____

This Annex forms an integral part of the Data Sharing Agreement between the Contributor and SecureBio. By completing this form, the Contributor confirms that the materials described below constitute the "Contributed Materials" referenced throughout the Agreement.

1. Description of Contributed Materials

Please describe the dataset(s), manuscript(s), or other materials being shared. Be as specific as possible.

1.1. Type of material (check all that apply)

- Tabular dataset (e.g., CSV, TSV, Excel)
- Manuscript (published / draft)
- Grant application
- Figures / images (e.g., western blots, microscopy)
- Other: _____

1.2. Brief description of the data (e.g., binding affinity measurements of X against Y; viral mutant replication assay; proteomics dataset).

1.3 Biological system / organism - *if applicable* (e.g., SARS-CoV-2, E. coli, human cell line HEK293)



1.4 Experimental conditions / key parameters - *if applicable*

(e.g., temperature, cell type, treatment, timepoints — anything needed to interpret or replicate the data)

1.5 Approximate size / scope

(e.g., ~500 data points, 3 replicates, 12 samples, 20-page manuscript)

1.6 File format(s) (e.g., .csv, .xlsx, .pdf, .tiff)

2. Publication Status

2.1 Current status of this data (select one)

- Fully unpublished
- In preparation / not yet submitted
- Under review
- Published

2.2 Do you anticipate publishing this data in the future?

- Yes Expected Timeframe:
- No
- Unsure

If yes, please notify SecureBio as outlined in Section 4 of this Agreement.



3. Authorization

3.1 I confirm that I have the right and authority to share the materials described above, and that doing so does not violate any institutional policy, funding agreement, or third-party obligation.

Yes

No