

THE AD WORKBENCH DATA REPOSITORY

As of January 25, 2023, the NIH's new Data Management and Sharing Policy is in effect. To help researchers share and access data, the NIH maintains a list of recognized data sharing repositories. This includes domain-specific repositories, which are limited to certain data types or disciplines.

Researchers benefit from a robust selection of NIH-supported repositories. For those working in Alzheimer's disease (AD) and related dementias, the **AD Workbench Data Repository** from the Alzheimer's Disease Data Initiative (ADDI) can meet this need.

ADDI accepts human data from studies in AD, related neurodegenerative disorders, aging, and healthy controls. While ADDI does not accept bio-samples, information about how they can be accessed from a data contributor may be included with a dataset listing.

Data contributors can store, organize, validate, archive, preserve, and distribute data in compliance with FAIR Data Principles (Findable, Accessible, Interoperable, Reusable). Data in the AD Workbench Data Repository is securely sharable, citable, and reusable.

Features of The AD Workbench Data Repository

NIH-supported data repositories exemplify these desirable characteristics, which are shared by the AD Workbench Data Repository:

- [Unique Persistent Identifiers](#)
- [Long-term Sustainability](#)
- [Metadata](#)
- [Curation and Quality Assurance](#)
- [Free and Easy Access](#)
- [Broad and Measured Access](#)
- [Clear Use Guidance](#)
- [Security and Integrity](#)
- [Confidentiality](#)
- [Common Format](#)
- [Provenance](#)
- [Retention Policy](#)

Unique Persistent Identifiers

A Digital Object Identifier (DOI) can be assigned to each study (including associated datasets, preprints, or code) in the AD Workbench Data Repository. If supporting documentation is updated, it

can track dataset versions. If a dataset is de-accessioned or no longer available, the DOI can also point to a persistent landing page. In that case, ADDI can create a tombstone page for withdrawn datasets and ensure the data is no longer findable. An ORCID ID, a unique author identifier, can also be added to each user profile in **AD Connect**, ADDI's online community platform.

Long-Term Sustainability

ADDI launched in late 2020, after a coalition of nine partners recognized that data silos are a key barrier in accelerating new discoveries. ADDI is supported by long-term commitments from donors who are interested in improving data sharing in AD and related dementia. Our current donors have indicated willingness to extend their funding to support ADDI's future success. We are also cultivating a broader range of prospective donors, which improves our likelihood of securing additional funding.

Metadata

For all datasets, ADDI offers metadata that is displayed in a common catalog format. This information is easily accessible to users. For datasets with a DOI, metadata can be viewed with [DataCite](#) and translated into common metadata formats such as [DublinCore](#).

Curation and Quality Assurance

Currently, ADDI only hosts human data that is properly anonymized or pseudonymized (i.e., it excludes personally identifiable information or personal health information). While ADDI does not offer comprehensive curation services, it does provide a comprehensive step-by-step [guide](#) to help data contributors prepare data for their platform. For more information on anonymization best practices, see: [UK Data Service guide for data management](#).

Data contributors are responsible for the accuracy and integrity of their data and ADDI encourages sharing high quality research outputs that are well documented. ADDI encourages data contributors to use the help guides and other resources that it offers. Contributors can also email ADDI with any questions about the accuracy and integrity of their data or metadata, at:

addi.support@alzheimersdata.org.

To help data contributors identify legal and policy issues related to data sharing, ADDI offers a [Data Sharing Toolkit](#). It includes information on legal and regulatory policies and best practices from trusted partners.

Free and Easy Access

One of ADDI's goals is to remove barriers to data sharing. To help achieve that, we offer data sharing and analysis tools that are easy to access and use. These tools are available at no charge to the research community.

Researchers can visit AD Connect to read high level descriptions about the datasets that may be permissibly accessed on the AD Workbench. These summaries also include a list of manuscripts that cite the dataset, instructions on how to request access, and other helpful details.

If a researcher wants to explore a dataset and field-level metadata within the FAIR dataset listings, they must sign up for an ADDI account, which is free of charge. To sign up, a user must provide some basic information and agree to ADDI's [Terms of Use](#). An ADDI account provides access to both the AD Workbench and AD Connect. Within the AD Workbench, a user can also submit a data access request or request a free workspace. Within AD Connect, a member can discover an array of resources, including discussion forums, live and recorded events, an AD Workbench technical library, and more.

Broad and Measured Reuse

ADDI maintains a log of submitted dataset requests. ADDI uses this information to learn which data types are of greatest interest to the research community and to identify issues/barriers in processing data access requests. [Matomo](#) is ADDI's privacy preserving web metrics tracking solution. While these metrics are not publicly available, we can provide anonymous, summary-level information to data contributors as requested.

Clear Use Guidance

To become an ADDI registered user, a researcher must accept ADDI's [Terms of Use](#). Each data contributor may also include their own Data Use Agreement (DUA) with additional terms related to

the use of their dataset. These customized DUA terms (if any) are integrated into the data access request process within the AD Workbench. Data contributors can define their own terms, but ADDI provides an example as a courtesy. See [Alzheimer's Disease Data Use Agreement template](#).

Data contributors can use the [ADDI Data Preparation Guide](#) when seeking to onboard a dataset. This guide is used in conjunction with the Data Prep Templates that ADDI also provides. ADDI team members are available to support data contributors upon request.

[AD Connect](#) provides a resource library to help users access and use the datasets that are hosted on ADDI. The resource library includes user guides, tutorials, FAQs, discussion forums, and more. An ADDI registered user can ask questions and get prompt answers from the community and ADDI team members.

AD Connect also hosts a growing number of events. If a researcher wants to talk to a data science expert, our team's *Office Hours*, which is a regular drop-in session – a place to ask questions live and learn more about ADDI's tools and resources.

Security and Integrity

The AD Workbench is a secure platform. Its security certifications include:

- ISO27001
- ISO27701
- HITRUST CSF
- Cyber Essentials
- ICO registration certificate
- NHS Data Security and Protection Toolkit assessment

Security protocols include: OWASP Top 10, OS patching, nightly backups, and quarterly BCP/ER exercises. Data and access management include HTTPS-TLS 1.2+, encryption in transit and at rest (FIPS 140-2), login with required two-factor authentication, intrusion protection-security alerts, and regular audits. For more information, see: [ADDI Privacy Policy](#), [Aridhia Security and Compliance](#), and [Azure Compliance](#).

Confidentiality

Our [Privacy Policy](#) describes what personal information we collect, how we use it, and what we share. Our [GDPR FAQ Guide](#) describes our privacy practices in Europe.

Common Format

ADDI does not mandate that data be harmonized to any specific standard or ontology. We do, however, have a standard set of variables (with code) that we encourage data contributors to use. We offer data contributors a comprehensive [guide](#), to be used along with our Data Prep Templates, to help prepare data tables and metadata for the AD Workbench. ADDI team members are available to provide additional guidance.

ADDI encourages data contributors to use non-proprietary formats. If proprietary formats are used, we work with data contributors to identify resources for researchers to access the data easily.

Provenance

Datasets may change over time, especially those from active clinical programs. All ADDI datasets and metadata have a full audit trail. This records what changes are made, when they are made, and who makes them. Administrators (or those with appropriate permissions) can [access these details](#). Workspace administrators can also track and monitor their workspace activity using the [Audit tab](#). The activity log breaks down all individual actions in the workspace in granular detail.

Data users can be notified (via email) when a given dataset has been updated by using our [dataset subscription](#) feature.

Retention Policy

Generally, there is no expiration on ADDI's storage of datasets. Data contributors, however, may ask for a specific retention period. Details about ADDI's retention of personal information can be found in our [Privacy Policy](#). Additional details, as it pertains to ADDI, can be found in Aridhia's [Privacy Policy](#).

Additional Considerations for Human Data

The NIH outlines desired data repository characteristics for human participant data, including de-identified human data. These characteristics are shared by the AD Workbench Data Repository:

- [Fidelity to Consent](#)
- [Restricted Use Compliant](#)
- [Privacy](#)
- [Plan for Breach](#)
- [Download Controls](#)
- [Violations](#)
- [Request for Review](#)

Fidelity to Consent

ADDI's platform allows for data to be sensitively handled:

- Data must be de-identified before it is delivered to the platform.
- Datasets can be held in private mode if needed.
- Data access can include permissioning steps and criteria, as well as additional terms of use for potentially sensitive data.

Before datasets can be shared, data contributors are responsible for determining if the datasets they wish to share are subject to any restrictions based upon participant consents. Consent language often changes over time, so it is necessary for data contributors to analyze each relevant consent form separately and determine which data is subject to which version of their consent forms. ADDI provides a [decision tree](#) resource to help contributors in determining whether sharing is permitted.

Restricted Use Compliant

[Data Use Agreements](#) (DUAs) outline a data contributor's terms regarding data sharing on ADDI's platform. While ADDI encourages data contributors to publicly list their datasets on our platform, we offer several controls to help meet additional control or transparency needs. Contributors can:

- Temporarily embargo a dataset.
- Hold a dataset in private mode for a period. A dataset in private mode is only discoverable and accessible by those who are proactively invited by the data contributor. For more information, see: [Dataset Forum](#).

- Use distributed access: data is hosted on premises by the data contributor, with the metadata queried remotely. Selected data may be transferred to an approved user's workspace.
- Use a [federated solution](#): data is hosted on premises by the data contributor, record-level data are queried and analyzed remotely, and only the approved results are transferred back to an approved user's workspace.

The ADDI platform includes features for requesting and providing restricted access to certain logged-in users. The “role-based access model” manages access to workspace resources based on several predefined user roles. Assigning a role to a user allows them to use specific features of the workspace and manage files. This permits workspace administrators to ensure that the principle of least privilege is followed. For a summary of the roles and their privileges, see our [Role-based access control](#) guide.

Privacy

ADDI's platform offers data contributors several tools to help maintain control and transparency over permissioning. Our permissioning workflows are flexible, which allows each contributor to decide what information must be provided by requesters, and to whom those requests will be routed. Data contributors can use this information to approve or reject requests at their own discretion. For more information on how data contributors can tailor data requests forms, see: [Custom DAR Forms Guide](#).

Users control who has access to their private workspaces. Within the workspaces, they can manage access based on several predefined user roles. Assigning a role to a user grants them permission to use specific features of the workspace and manage files. For more information on the various roles and privileges available, see: [Role-based access control](#) guide.

Anyone who requests access to data must have an ADDI account. This requires users to consent to ADDI's [Terms of Use](#). Beyond that, ADDI is not involved in reviewing or approving data access requests. Data contributors may entirely waive the need for permissioning of their datasets, in which case the platform automatically approves all requests on the data contributor's behalf.

ADDI accepts neuroimaging data and emerging forms of digital biomarkers such as voice data. Any human data ADDI accepts must be de-identified. For more information, see: [Privacy Policy](#).

Plan for Breach

All datasets are de-identified prior to posting on ADDI's platform. ADDI also employs several security controls and monitoring processes to protect data privacy and security. For a list of these measures, see: [Aridhia's Security and Compliance](#).

Download Control

Anyone requesting access to data, or a workspace must first be a registered ADDI user. This includes agreeing to our Terms of Use. When a data access request is approved by the data contributor, the data (or results) are transferred into a private workspace. , Data download is not permitted by default. Data may be downloaded from this workspace, provided that it is allowed by the Data Use Agreement (if any) created by the data contributor. For more information, see: [Downloading Files from Within Your Workspace](#).

Violations

ADDI's [Terms of Use](#) includes procedures to address any user violations and mismanagement of data.

Request Review

AD Connect has a [Dataset Forum](#), which includes a brief description of the access request process for each dataset. Some datasets may be automatically approved, while most require the data contributor to review and approve each access request. The review and approval timeline will vary, depending on each data contributor's governance and operations.