Aridhia TRE

Service Description

Authors: **Aridhia PO Team**
Date: **13 November 2023**
Version: **1.0**

Table of Contents

Executive Summary

The Aridhia TRE provides secure collaborative cloud-hosted workspaces for individual research teams and projects. Aridhia's TRE workspaces allow the collaborative analysis of healthcare research data. They can be safely accessed from multiple locations, allowing secure, advanced project collaboration across sites.

Workspaces can only be accessed by authorised users who can upload and analyse data, collaborate on data or other project resources, follow project activity and export data for publishing or further research.

A Workspace can be accessed via a web browser and provides access to a large range of analytical and data tools. Workspaces can be provisioned with a virtual machine where the user can also 'bring their own' analysis tools or software for data manipulation or analysis.

Workspace activity is audited and traced for maximum security and compliance.

The full feature list and user guide is available at: https://knowledgebase.aridhia.io

Scope of document

This document describes the features and functionality of Aridhia's TRE service version 3.27.

TRE Workspaces Features

Key Features

COLLABORATION AT THE HEART

- Worldwide access via common web browsers
- Bring all project members to the data
- Collaborate with users from outside of the organisation
- Built in private Git available for collaborative development via Linux VMs

SCALABILITY

- Access up to five unique workspaces
- Request virtual machine access

BUILT-IN DATA ANALYSIS

- Native support for R, shiny and SQL.
- Built-in R-Studio® and Jupyter Notebooks with Git support
- Built-in statistics modules for point and click data analysis

SECURE DATA STORAGE

- Workspace is audited
- Data can be imported by the user and exported or shared via Airlock service

PRIVACY BY DESIGN

- Secure data access and management via MFA, RBAC, encryption and secure key management
- ISO 27001, ISO 27701, Cyber Essentials Plus and HITRUST accredited.

CLOUD-NATIVE SERVICE

- Developed and hosted on the Microsoft Azure cloud.
- Integrates with and improves on cloud technologies.

Service Overview

Workspace Provisioning

Workspaces provided by Aridhia are cloud based, hosted on Microsoft's Azure cloud, and delivered in a 'Software as a Service' (SaaS) model. This allows a Workspace to be scalable and adjustable for the user's needs and accessible globally from authorised locations, in line with organisational security policies. Compute power can be added for more complex data analysis and more users can easily be added for larger collaboration projects. The length of time a Workspace is provided can vary depending on the scope of the project.

Users can register with Azure B2C any time after their hub has been set up.

- A Tenant Administrator can be assigned, they can create Workspaces and set the membership for each one.
- A Tenant Administrator can pre-register users who will receive an email inviting them to sign up to the platform. Up to 10 users can be added to a workspace.
- The Workspace Administrator can manage which users have access to their Workspace.
- Users have a defined role associated with their Workspace account through the built-in role-based access control.
- Users can sign in using their e-mail address and a password which they set. All users must use second factor authentication using a mobile or a landline phone to log in.
- Organisations also have the option to allow access to workspaces using a login service federated to their own Active Directory.
- The Workspace Administrator and Manager can invite registered users into the Workspace and users can accept or reject the invitation.
- When adding users to a workspace, they can be identified by name and e-mail address.
- The Tenant Administrator can Hibernate a workspace and "wake up" a hibernated workspace
- A Tenant Administrator can apply restriction descriptions to a workspace that members must accept prior to being able to access the workspace. Acceptance is audited.
- Tenant administrators can suspend and re-enable suspended users.
- Tenant administrators can order workspaces by 'last activity'
- Tenant administrators can export a CSV of all workspaces, members, last activity and state

User Management and Workspace Access

Aridhia's Workspaces have privacy built in by design. The Tenant Administrator and administrator of a Workspace can add registered users to a Workspace and can assign them one of the following roles:

The **Standard** user profile applies to the majority of Workspace users. The Standard user can:

- upload files and data into the workspace
- edit and update files, blobs and database tables
- create, edit and collaborate on database tables and their associated metadata and delete work files
- create, preview, publish, delete and run R Shiny apps
- install, run and delete software in virtual machine
- run containerised Apps
- create notes
- create comments (but can only update or delete their own)
- generate Upload API access tokens
- request an Airlock of files and/or data
- view a list of Workspaces that they have been granted access to, and the associated members.

The **Administrator** is the user in charge of the Workspace. They have all the same rights as the Standard user above, but they can also:

- add or remove members to the Workspace(s)
- set the role for members

- access the audit of their Workspace(s)
- approve/deny Airlock requests to export results and files out of the workspace.

The **Manager** can also manage the workspace and its users. A Workspace manager has the same rights as a standard user and can also:

- add or remove members to the Workspace(s). They cannot add Workspace Administrators.
- set the role for members
- access the audit of their Workspace(s)

Contributor users can add data to a workspace via Workspace to Workspace Airlock. They have no access to the workspace UI.
Contributors can upload files and data using the upload API but must have a token generated by a standard user or workspace administrator.

Supporting the Research Lifecycle

The Aridhia Workspace is built to support day to day analytics and the lifecycle of a research project. It aids collaboration by allowing users to access shared projects and data in a cloud hosted environment. The Workspace is designed to allow easy analysis of data using the built-in tools but also offers flexibility for the more advanced researcher.

Once a user has been granted access to a Workspace and has accepted the invitation, they are able to carry out all or some of the following actions, depending on their assigned role.

LOG IN

In order to access their Workspaces, a user must first log into the Aridhia DRE. The user can:

- authenticate using Multi-factor Authentication (MFA) via TOTP
- see a list of their Workspaces and the role which they have in that Workspace
- see a summary of each Workspace
- access documentation.

The user can select the Workspace they want to work in and are redirected into that Workspace. Once logged into a Workspace the user can:

- see a project homepage (if it is set)
- see the activity summary
- see notifications in a service banner.

The activity page provides a timeline of the activity within the Workspace so that users can see what each other have been working on. Users can:

- comment on work or activity
- leave notes for other users
- see project highlights and insights
- promote notes to an insight to highlight them to the other workspace users.

IMPORT OF DATA

Workspace users can import data into the Workspace for analysis. They can:

- upload items directly in the Workspace UI using drag and drop or file selection
- upload items using an API
- upload items to an Inbox or
- upload data as a CSV or add the TDF and have data tables created automatically

- select where the upload should be stored within the Workspace

All uploaded data is automatically scanned for viruses and malware before being moved into the workspace.

Data can be stored in a file or blob store.

WEB UI

The Aridhia Workspace provides the following functions for the analysis and manipulation of the uploaded data:

- View dataset summaries
- View the data in read only mode in a table format
- Edit the table data in a specialist editor
- Edit the data using SQL scripts
- Run SQL queries to find and analyse data
- Quickly visualise data using built in analytics tools
- Run analysis modules on random sample sets of very large datasets
- Access uploaded files and blobs
- Edit and execute R scripts and other code using a syntax-highlighting editor (supported file extensions are sql, csv, r, rnw, ipynb and txt)
- View a rendered version of markdown files
- Install R packages into a console session from a library of resources
- Create and edit documents to support the project
- Generate PDF reports
- Search through uploaded and created documents and files
- Create and use R Shiny 'mini-apps', providing targeted, interactive visualisation and analytics.
- Use Built-in Apps such as R-studio and Jupyter notebooks.

All the project's essential documents and information can be safely held in one system which saves users from switching between contexts to find information; it also means that all project users can have easy access to the information that they need.

VIRTUAL MACHINE

Virtual Machines can be added to Workspaces, they can be used to access additional tools and analytical software.

- Linux or Windows virtual machines.
- Default of 4 vCPU and 16GB Memory
- All VMs in a workspace are in an isolated subnetwork specific to the workspace for enhanced security and to enforce workspace separation
- Access to the internet outside of the workspace subnet is controlled via a proxy allowed domains list
- Workspace files and data tables can be written to and read from the VM and files from the VM can be saved back to the Workspace.
- Users can install their own software including analytical or graphical tools.
- Use of VMs is audited at the workspace level
- VMs are automatically patched on a regular basis

OUTBOUND AIRLOCK

At any time in the project, data or project artefacts can be exported from the Workspace through the Airlock. The user can:

- create a request in the Workspace to export some data or artefact that they have been working on
- send the request to the Workspace Administrator who must approve the export of the data.

The Airlock approval step ensures that the Workspace Administrator knows where the data is being used and provides data security.

Airlock requests can also be made to send data and artefacts to other workspaces which the user is a member of using the **Workspace to Workspace Airlock** feature.
Here, the user can

- create a request in their workspace and select one of their other workspaces as a destination
- send a request to the administrator to have their transfer approved
- receive the zipped Airlock package in the Inbox of their other workspace

Approved Airlock requests are packaged in a zipped archive for convenient download or transfer.

AUDIT

Many of the steps described above include entries into an Audit log which can be seen by the Workspace administrator. The administrator can export this log into a CSV format so that it can be stored or used elsewhere.

Audit logs include the:

- date and time of the action
- user name and role of the user who carried out the action
- type of event they carried out
- resource affected
- events including data manipulation, uploading of new data and exports are reported in the Audit log.

The audit log is available for the lifetime of a workspace although the workspace administrator may only be able to see the most recent portion of this. The full audit log can be requested form Aridhia at any time.

Service Level Agreement (SLA)

Aridhia offers availability at an SLA of 99.5%, excluding planned downtime.

Standard support SLAs are described in the Aridhia Service Desk SLA document.

A database and all the data in the workspace are backed-up on a nightly basis (every 24 hours). Back-ups are retained for fourteen days on a rolling basis and saved in a local Azure region or an alternative region on client request. Virtual Machines' local disks are not backed-up.

Terms and conditions of the service can be found in the End User License Agreement (EUA).

Technical Information

Technical Requirements

AZURE SERVICES USED

Workspaces are hosted on Microsoft Azure platform. DRE currently uses the following Azure services:

Tenant/ Subscription level Resource Group

- Azure Defender
- Storage (ARM templates)

Hub Resource Group

- Virtual networks
- Azure Kubernetes Service
- IPS
- Azure Service Bus
- Azure Container registry

Workspace Resource Group

- Application Insights
- Log Analytics
- Container Registry
- Storage Account
    - Blob container
    - Azure Files share
- Azure Key Vault
- (Optional) Azure Machine Learning

Backup Resource Group

- Recovery Vault

BROWSER SUPPORT

The current list of supported browsers are described in the Aridhia DRE Browser Support Knowledge base article.

Deployment

Aridhia TRE is developed upon and hosted on Microsoft Azure. Aridhia TRE is a multi-tenanted SaaS service. Each unique organisation making use of the TRE has a self-contained data processing environment while sharing underlying services infrastructure services with all other TRE users.

Security

Aridhia has implemented the DRE to ISO 27001 standards and has been ISO27001 certified by external auditors. This covers Aridhia's processes and technology (see certificate).

Aridhia also holds the following security certificates:

- ISO27701,
- HITRUST and
- Cyber Essentials

Aridhia self-certifies to the NHS Digital Data Security and Protection Toolkit assessment.

The DRE is deployed on Microsoft Azure which is also accredited to ISO27001 and numerous other certifications such as CSA STAR.

Aridhia security and privacy by default approach is reflected in the following security measures:

- Web standard authentication using OAuth2, MFA, TOTP and role-based access
- Regular penetration testing by third parties
- Secure Development Life Cycle
- Data never leaves the platform except through Airlock approval
- Comprehensive Audit log
- Virus scan on upload.

Azure Defender is used to protect the DRE against threats.

ENCRYPTION AND KEY MANAGEMENT

Aridhia follows best practice guidelines for data encryption and secure key management for all cryptographic operations:

- All user and API communication is encrypted using HTTPS
- Encryption of data at rest and in flight.

MFA AND ROLE-BASED ACCESS

Multi-Factor Authentication (MFA) is enforced for all user accounts of the service. The principle of least privilege is employed via role-based user access. Newly signed up users are given no permissions until assigned by an Administrator.

Shared Responsibilities

This section outlines the responsibilities for Aridhia and the customer in order to successfully deploy and operate the service.

The customer shall:

- Submit any change requests, feature requests, incidents or general support questions to the Aridhia Service Desk
- Identify a customer service owner(s) for the purpose of being a primary point of contact for the service
- Ensure security packages are up to date on Workspace VMs.

Aridhia shall:

- Provide access to a dedicated organisation TRE environment
- Normally deploy a new version once every two weeks
- Provide access to training material on how to utilise the service
- Provide support to the customer via the Aridhia Service Desk adhering to defined support SLAs
- Continue to provide updates to the service where necessary (e.g. features, security) and notify the customer of such updates and any impacts to the running of the service (e.g. downtime).
- Act only as a processor of data and **not** a data owner/controller.

Appendix A – All Workspace Features

The list below contains a detailed description of all features available in Workspaces on Azure.

User Sign up (registration)

Users can be invited to register via automated e-mail
Users are authenticated through Azure Active Directory B2C
Users are authorised through workspace membership rules

Workspace Administration

Workspace Administrator can invite users to their Workspace
Users are identified using their name and e-mail address
Workspace Administrator or Manager can remove users from their workspace
Administrator, Manager, Standard and Contributor user roles are available - see section on roles
Hibernate workspaces and store their content in Azure Cold Storage

Workspace Management

Workspace owner can be set
Workspace description can be set
Workspace logo can be set
Hubs can be provisioned in many Azure regions
Workspaces can be hibernated to save Hub resources
Restrictions can be added to workspaces

Login

Multi Factor Authentication (MFA) is required
Second factor authorisation TOTP via Microsoft Authenticator App

Workspace list is accessible from a known domain name

Workspace List

A user can only see their own Workspaces
List of Workspaces is searchable and filterable
Each Workspace has its own URL, which can be shared

Homepage

The Workspace homepage is customisable by a member using HTML
It can include links to the "outside world"
The homepage can include links to other areas of the Workspace
Images can be included in the homepage
Users can see notifications in a service banner.

Workspace Activity

Users are identified by name
Users can comment on other user's actions
Links to files can be used
Airlock requests are reported in the summary
The Workspace administrator can approve or deny Airlock requests
A user can download or transfer their Airlock package
Comments and Notes can be promoted to Insights
The activity list can be filtered for insights
Users can add formatting to a note

Auditing

The Workspace audit log is available to the Workspace administrator (not standard users)
Each audit record reports Time, User, Event, Resource and Detail parameters of the audit event
The audit columns are filterable
The audit log can be exported (downloaded) by the Workspace Administrator
The full audit history of a workspace can be requested from Aridhia

Apps

New apps can be created in the workspace using R Shiny
R Shiny Apps can be uploaded
R Shiny Apps can be edited in files
R Shiny Apps can be edited in R Studio in VM
Apps are hosted in dedicated Kubernetes pod and can access Workspace resources
Jupyter Notebook
R-Studio

Database Table

Data can be imported from a CSV file
CSV can be transformed into a database table via wizard
Metadata can be added to tables
Metadata can be imported from a file (.tdf) or copied from another table
Metadata can be exported to a file
Database tables and views can be previewed
Users can create custom views
Table data can be visualised
Data can be edited inline
Analytics module with >20 commonly used methods in medical statistics (see Appendix B for details)

Data can be exported as a datafile (csv)
Datasets can be deleted
Datasets can be accessed from the VM
PostgreSQL DB (enables R Studio/Jupyter)

Files

Files can be stored in Files and Blob storage
Users can create subfolders
Users can create new files
Users can upload files
Users can set a file type
Files can be moved
Files can be renamed
Common code and file formats can be opened in an editor (see Appendix D)
Files can be uploaded to an Inbox
Users can view rendered markdown

Other file types may be opened as text files

Markdown files open as HTML or can be edited
Executable files can be run from the editor or file list

Scripting

Syntax highlighting (.r and .sql)
Run file from the editor

Datafiles

CSVs can be viewed and analysed as data tables
Create and Edit data via data table
Very large CSVs will be randomly sampled in the data table

R Console

Start a new session in the console
Create new scripts in the console
Select R Version
Open scripts from the sidebar
Access Workspace files and data
Import packages from approved CRAN mirrors
View history, plots and reports

Virtual Machines

Virtual machines can be provisioned by Aridhia
VMs can access workspace files
VMs can access workspace DB
Outbound network connectivity is limited by a hub-wide standard allowed domains list
(access to internet is limited to whitelisted sites)
Virtual machine sizes can be chosen from approved list
Windows and Ubuntu Linux operating system options available
Linux VM users have unique user identity scoped home directories
Linux VMs have access to Gitea collaborative development and versioning
VMs can be set to auto-shutdown at a predefined time of day
VMs are automatically updated unless otherwise agreed.

Airlock (Data Export)

Standard users can make a request to remove items from the Workspace
Workspace Administrators can self-approve Airlock requests
Workspace Administrators must approve the Airlock request
Workspace Administrators can reject an Airlock request
Airlock requests are highlighted in the summary tab
All users must submit a reason for requesting an Airlock Package
Airlock Packages can be downloaded or transferred directly to other workspaces
An Airlock can be requested from the side bar
Airlock can be requested from the selected file's option menu

Appendix B – Medical Statistics Modules

Below is a list of available Workspace methods. For a full description see the Medical Statistics Core section of the Aridhia DRE Knowledge Base.

Descriptive statistics & visualisation

- Density plot
- Histogram
- Frequency Bar Chart
- Column Bar Chart
- Box/Whisker plot
- Scatter plot
- Bubble chart
- Line chart
- Area chart
- Heatmap
- Radar Chart
- Word Cloud
- Timeline

Tests

- Single mean
- Compare means
- Crosstab
- Student's T-test (with visualisation)
- Mann–Whitney U test (with visualisation)
- Pearson's chi-squared test

Modelling

- Correlation
- Analysis of variance (ANOVA)
- Linear Regression
- Logistic Regression

Appendix C: Outbound Connectivity Allow list

By default, the following destinations are whitelisted for HTTPS proxy access from a Workspace VM. The list can be configured for customer requirements. For example, connectivity to APIs, compute or storage facility hosted elsewhere can be enabled.

FOR R PACKAGE MANAGEMENT (CRAN)

- cloud.r-project.org
- .rstudio.org
- rstudio.com
- r.aridhia.net
- www.stats.bris.ac.uk
- cran.ma.imperial.ac.uk
- cran.cnr.berkeley.edu
- ftp.osuosl.org

FOR PYTHON PACKAGE MANAGEMENT

- pypi.python.org
- pypi.org
- files.pythonhosted.org
- anaconda.org
- repo.anaconda.com

FOR UBUNTU LINUX PACKAGE MANAGEMENT

- ftp.heanet.ie
- nl.archive.ubuntu.com
- azure.archive.ubuntu.com
- archive.ubuntu.com
- security.ubuntu.com
- ppa.launchpad.net

ARIDHIA SERVICES & AZURE SERVICE DEPENDENCIES

- knowledgebase.aridhia.io
- .azure-automation.net
- .ods.opinsights.azure.com
- .oms.opinsights.azure.com
- For Windows patching
- file.core.windows.net
- packages.microsoft.com

MISCELLANEOUS

- (For GPU/CUDA drivers) developer.download.nvidia.com

Appendix D – Natively Supported File types

The following files types can be opened in the Workspace file editor without a warning popping up to make the user confirm that the file is safe to open as a text file:

- SQL (.sql)
- R files (.r)
- LaTeX markup (.rwn)
- Jupyter Notebook (.ipynb)
- plain text (.txt)
- markdown (.md)
- R-Shiny app version (.version)
- YAML (.yml/.yaml)
- JSON (.json)
- Shell scripts executable (.sh)
- JSON files (.json)
- Markdown (.md)

Appendix E – R-Studio