

# Open-science solutions enabled within a digital research environment to support MIDD (Model Informed Drug Development) and Precision Dosing

Jeff Barrett<sup>1</sup>, Scott Russell<sup>1</sup>, Janani Murthy<sup>1</sup>

<sup>1</sup>Aridhia Bioinformatics, Glasgow, Scotland G2 4SQ

## INTRODUCTION

A Digital Research Environment (DRE) is a secure end to end biomedical research facilitation platform, connecting data owners to researchers in a controlled and audited manner, ensuring data owners are always in control of how their data is used, who it can be used by and what it can be used for. The Data services component of the DRE provides meta-data catalogue, search, summary cohort analysis and orchestrated access request facilities in line with FAIR data principles [Wilkinson 2016]. It allows data controllers to openly publish metadata about their datasets, control and approve access to those datasets and allow registered and approved researchers to query and build cohorts of interest. The Data Service can act as a repository for data or can connect remotely to established data repositories in a federated network.

Once approved, researchers make use of a Workspace, otherwise known as a Trusted Research Environment (TRE), to work with data, build machine learning models, develop models, curate novel datasets and author collaborative documentation to support their study, all with fully audited and traceable activity and code and data version control. A workspace can be pre-loaded with tooling relevant to the research intentions, as well as providing standard and popular data science development and visualization tools. Researchers have access to compute on demand and can, upon approval, upload their own tools to be deployed and made available to all other users of the workspace. The Workspace provides a safe haven for clinical researchers, bioinformaticians and pharmacologists to analyze and develop models on sensitive data with the confidence that the data and models developed are secure and protected. Workspace architecture and “out-of-the-box” analytics are depicted in Figure 1.

## OBJECTIVES

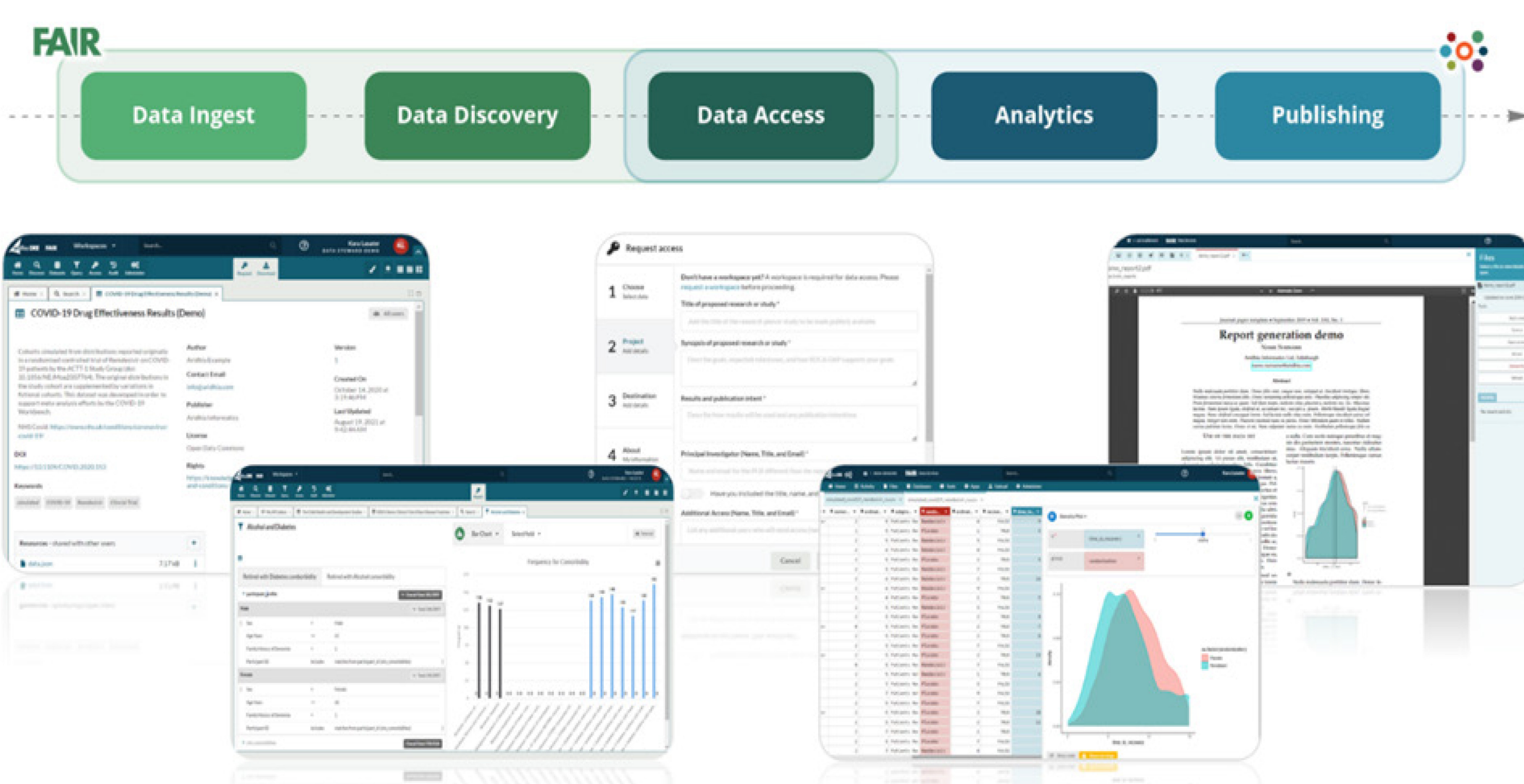
- To enhance the outreach and reliability of current open-science solutions supporting MIDD and precision dosing as part of a broader precision medicine strategy.
- Collaborative efforts to support expanded open-science solution functionality include the addition of nlmixr2 with the nlmixr team and Great Ormond Street Hospital (GOSH) and others for nonlinear mixed effect modeling and r-based system pharmacology capabilities in collaboration with ESQ Labs.
- Commercial vendor solutions are also in scope for workspace integration and agreements with ICON have already enabled NONMEM implementation within a TRE Workspace.
- The status and functionality of these and other solutions for R&D are described along with general guidance on the various general mechanisms by which such solutions can be exploited in a Digital Research Environment (DRE) workspace.

## RESULTS

Common components of a workspace currently include:

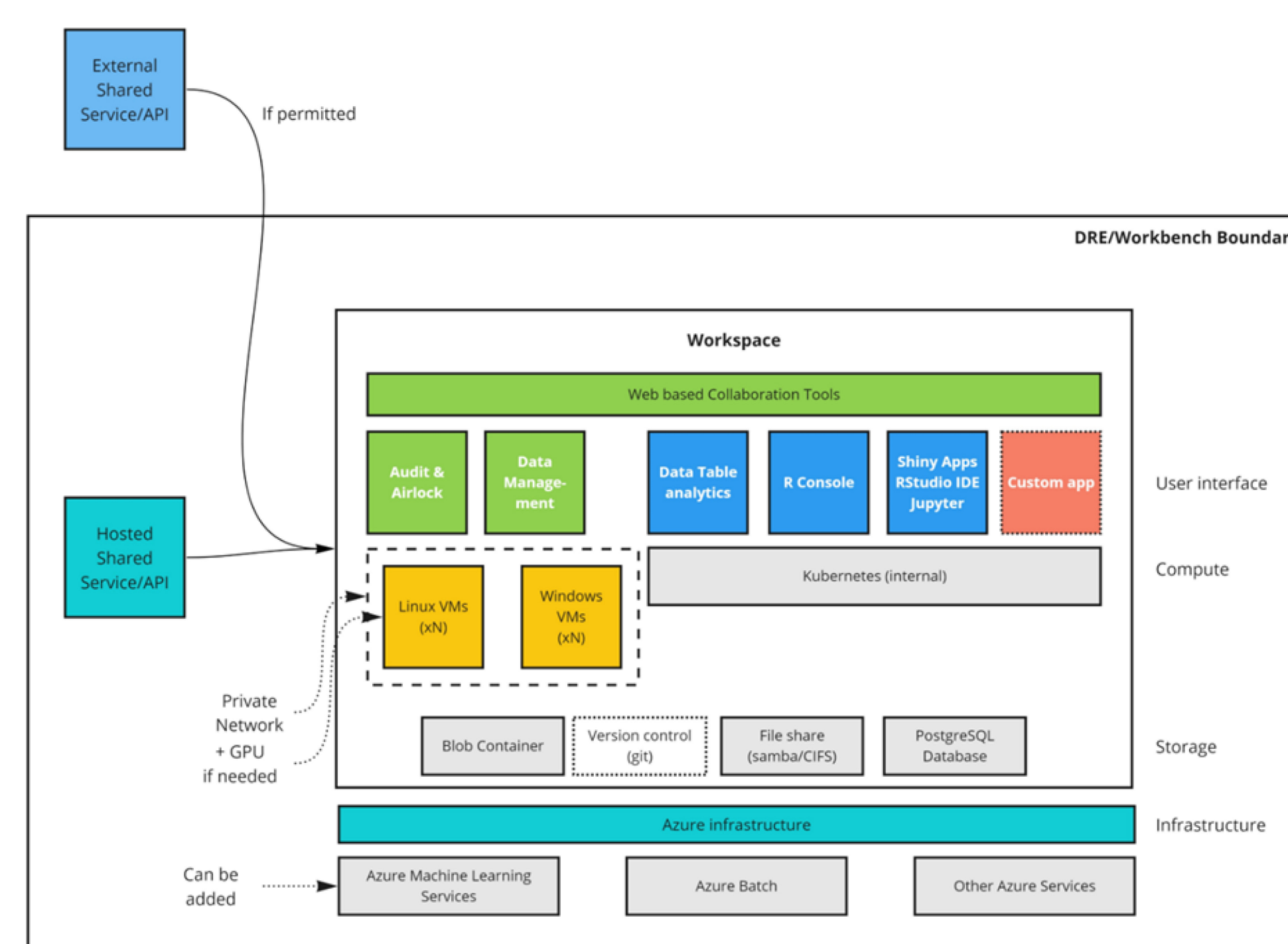
- A shared file system for structured and unstructured data to allow workspaces members to work on a common set of data and code files
- A relational database that can be used to contain original data or build novel datasets and provide SQL access from workspace tools
- A Kubernetes cluster for hosting customized containerized applications.
- A secure token-based file upload mechanism to allow users to bring their own code and data into the workspace
- An inbound ‘airlock’ to scan incoming files for viruses and provide a means for administrators to review incoming files prior to general access
- R and Python coding environments to build apps and models
- Data exploration tools such as table sorting, filtering and pre-defined biomedical analysis modules
- Git, to provide the ability to collaborate and version code and data within a workspace
- Access to virtual machine compute, including GPU backed machines for large scale modelling purposes.

Figure 1: Workspace architecture and “out-of-the-box” analytics



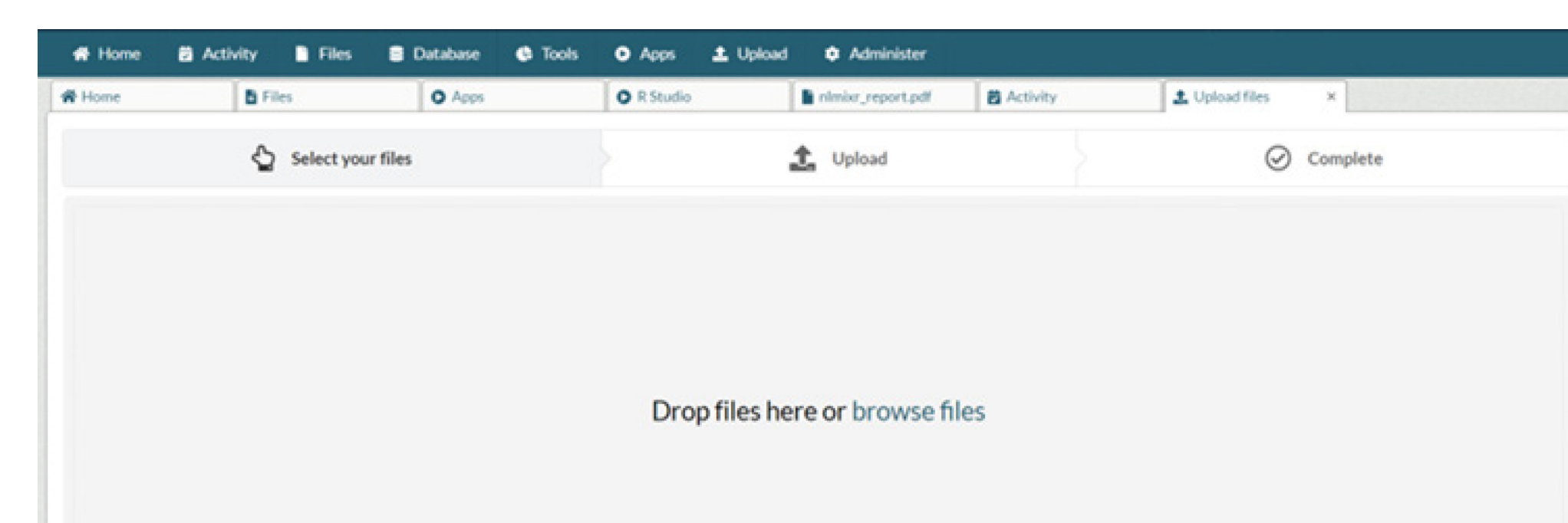
## RESULTS

Figure 2. Anatomy of a trusted research environment



The below example demonstrates the steps required to use nlmixr2 inside a DRE with an in-built Rstudio app:

1. Upload the dataset required for the nlmixr2 model using the upload option in workspace.



2. Once the dataset is loaded view and analyze the data using analysis tools in DRE.



3. Open Rstudio app.
4. Install nlmixr2 package using install.packages(“nlmixr2”) command.
5. Create a new project, write a R script to read the dataset, create the nlmixr model with required parameters based on the data analysis and place to store the model outputs.
6. Run the nlmixr models inside RStudio.

```
RStudio
# R script
1 library(nlmixr2)
2
3 pheno_sd <- read.csv("~/files/nlmixr2modelling/pheno_sd_dataset/pheno_sd.csv")
4
5 ##### Source the functions from last week #####
6 source("gof_plots.R")
7
8
9 ##### 1. compare the model with and without correlation #####
10 # run 1 is without correlation
11
12 # STEP 1: make a function defining the model. We will call it run1.mod
13 run1.mod <- function() {
14   # The ini block is used to specify initial estimates.
15   # Note we will log transform parameters - this has some advantages for
16   # source the functions from last week :
17 }
18
19 ##### Background jobs #####
20 R 4.3.1 .../files/NonmemModelling/
21 497: -5.000075 0.347633 0.244464 0.156375 2.833022
22 498: -5.000080 0.347613 0.244478 0.156338 2.832842
23 499: -5.000085 0.347613 0.244490 0.156350 2.833099
24 500: -5.000094 0.347599 0.244438 0.156325 2.833501
25
26 ##### Calculating covariance matrix #####
27 ##### Loading into symengine environment... #####
28 ##### Pruning branches ('if/else') of same model... #####
29 ##### Finding duplicate expressions in saem predOnly model... #####
30 ##### Finding duplicate expressions in saem predOnly model 1... #####
```

## REFERENCES

- Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018 (2016). <https://doi.org/10.1038/sdata.2016.18>
- Barrett JS. Nonlinear Mixed Effect Modelling – The use of nlmixr2 in a DRE workspace. <https://www.aridhia.com/blog/nonlinear-mixed-effect-modelling-the-use-of-nlmixr2-in-a-dre-workspace/>
- Russell, S. SATRE: Standardised Architecture for Trusted Research Environments – Supporting Capabilities. <https://www.aridhia.com/blog/satre-standardised-architecture-for-trusted-research-environments-supporting-capabilities/>
- Fidler M, Wilkins JJ, Hooijmaijers R, Post TM, Schoemaker R, Trame MN, Xiong Y, Wang W. Nonlinear Mixed-Effects Model Development and Simulation Using nlmixr and Related R Open-Source Packages. CPT Pharmacometrics Syst Pharmacol. 2019 Sep;8(9):621-633. doi: 10.1002/psp4.12445. Epub 2019 Jul 16. PMID: 31207186; PMCID: PMC6765694.
- <https://en.wikipedia.org/wiki/NONMEM>

## CONTACT

To schedule a demo, please contact Jeff Barrett ([jeff.barrett@aridhia.com](mailto:jeff.barrett@aridhia.com)) or Scott Russell ([scott.russell@aridhia.com](mailto:scott.russell@aridhia.com)). Check out our website for video-based tutorials and future webinars with open-science collaborators (<https://www.aridhia.com/>).