**MetWorx\GitHub**

**Solution Blueprint**

**Template Version Number: 8.0**

|  |  |  |  |
| --- | --- | --- | --- |
| **Document Approval / Review** | | | |
|  | **Name and Role** | **Signature** | **Date** |
| Author | Jim J Fedoryczuk , Solution Architect |  | 18/20/21021 |
| Reviewed | <reviewer>, <role> |  | dd/mm/yyyy |
| Reviewed | <reviewer-information-architecture>, <role> |  | dd/mm/yyyy |
| Reviewed | <domain-lead>, <role> |  | dd/mm/yyyy |
| Approved | <approver>, <role> |  | dd/mm/yyyy |
| Approved | Cyber Security Architecture |  | dd/mm/yyyy |
| Approved | CFIT Head of Architecture /  SSIT Head of Domain Architecture |  | dd/mm/yyyy |
| Approved | Architecture Review Board / Enterprise Architecture Review Board |  | dd/mm/yyyy |
| Approved | Service Lead / Platform Lead /  Capability Service Manager |  | dd/mm/yyyy |
| Status | <Not yet approved/Approved\*> |  | dd/mm/yyyy |

|  |  |  |  |
| --- | --- | --- | --- |
| **Version Control** | | | |
| **Version** | **Date** | **Description of Changes** | **Author** |
| 0.1 | 18/10/2021 | Draft | JJF |

|  |
| --- |
| **Disclaimer** |
| If you are reading a copy of this document (i.e. print out of controlled electronic document or an electronic copy of a wet ink signed paper document) you may be reading an uncontrolled version. Before making updates, please ensure that you are using the latest version of this document and the associated template. |

Table of Contents

[1 Introduction 7](#_Toc64964996)

[1.1 Purpose of Document 7](#_Toc64964997)

[1.2 Architecturally Significant Assessment 7](#_Toc64964998)

[1.3 Data Classification Assessment 8](#_Toc64964999)

[1.4 Document Scope 8](#_Toc64965000)

[2 Business Architecture Overview 10](#_Toc64965001)

[2.1 System Purpose 10](#_Toc64965002)

[2.2 Business Context 10](#_Toc64965003)

[2.3 Intended Value for Customer/User 11](#_Toc64965004)

[2.4 Business Capabilities Provided 11](#_Toc64965005)

[2.4.1 Business Capabilities Supported 11](#_Toc64965006)

[2.4.2 Value Streams Supported 12](#_Toc64965007)

[2.4.3 Impact on Application Estate 13](#_Toc64965008)

[2.5 Business Process Overview 15](#_Toc64965009)

[2.6 User / Roles Overview 16](#_Toc64965010)

[2.7 Regulatory Position 17](#_Toc64965011)

[2.7.1 Sarbanes-Oxley (SOx) Position 17](#_Toc64965012)

[2.7.2 Data Privacy and General Data Protection Regulation (GDPR) Position 18](#_Toc64965013)

[2.7.3 Good Pharmaceutical Practice (GxP) Position 18](#_Toc64965014)

[2.7.4 Information Management Position 19](#_Toc64965015)

[2.7.5 eDiscovery Position 20](#_Toc64965016)

[2.7.6 Medical Device Classification 20](#_Toc64965017)

[2.7.7 Other Applicable Regulations 21](#_Toc64965018)

[2.8 Criticality Statement 21](#_Toc64965019)

[2.9 Resiliency Statement 22](#_Toc64965020)

[2.10 Externalisation Statement 23](#_Toc64965021)

[3 Architecture Alignment 24](#_Toc64965022)

[3.1 Alignment to EA Principles 24](#_Toc64965023)

[3.2 Alignment to Strategies and Roadmaps 25](#_Toc64965024)

[3.3 Alignment to Standards and Reference Architectures 26](#_Toc64965025)

[3.4 Development and Delivery Model 26](#_Toc64965026)

[3.4.1 Development Lifecycle Model 26](#_Toc64965027)

[3.4.2 Development Model 27](#_Toc64965028)

[3.4.3 Delivery Model 27](#_Toc64965029)

[4 Solution Architecture 29](#_Toc64965030)

[4.1 Conceptual and Logical Architecture 29](#_Toc64965031)

[4.1.1 Conceptual Solution Architecture 29](#_Toc64965032)

[4.1.2 Logical Solution Architecture 31](#_Toc64965033)

[4.2 Technical Solution Architecture 33](#_Toc64965034)

[4.2.1 Solution Technologies and Lifecycle Status 33](#_Toc64965035)

[4.2.2 Enterprise Middleware Consumption Overview 34](#_Toc64965036)

[4.2.3 Communication Flow between components 35](#_Toc64965037)

[4.2.4 Software of Unknown Provenance (SOUP) / Software as a Medical Device (SaMD) 35](#_Toc64965038)

[4.3 Infrastructure 36](#_Toc64965039)

[4.3.1 Hosting Summary 36](#_Toc64965040)

[4.3.2 Clustering / Containerisation / Replication 37](#_Toc64965041)

[4.3.3 Implementation Architecture Diagram 37](#_Toc64965042)

[4.3.4 Network Diagram 38](#_Toc64965043)

[4.3.5 System Component Table 39](#_Toc64965044)

[4.4 Environments 40](#_Toc64965045)

[4.4.1 Environment Overview Diagram 42](#_Toc64965046)

[4.5 Integrations 44](#_Toc64965047)

[4.5.1 System Context / Interfaces Diagram 44](#_Toc64965048)

[4.5.2 Interface Table 46](#_Toc64965049)

[4.5.3 Integrations Table 46](#_Toc64965050)

[4.6 Support and Governance Model Statement 48](#_Toc64965051)

[4.7 User Experience Overview 49](#_Toc64965052)

[4.7.1 User Experience Measures 50](#_Toc64965053)

[4.7.2 User Experience Design 50](#_Toc64965054)

[4.7.3 User Experience Journey Maps 50](#_Toc64965055)

[4.8 Data Refresh and ETL Overview 51](#_Toc64965056)

[4.9 e-Discovery Response Overview 51](#_Toc64965057)

[4.10 Scalability and Elasticity Overview 52](#_Toc64965058)

[4.11 System Lifecycle Plan 52](#_Toc64965059)

[4.12 Component Lifecycle Overview and Evergreen Statement 53](#_Toc64965060)

[4.13 High Availability / Resiliency / COOP / BCP / DR Description 53](#_Toc64965061)

[4.14 License Model Statement 53](#_Toc64965062)

[4.15 Funding Model 54](#_Toc64965063)

[4.16 Sustainability Statement 54](#_Toc64965064)

[5 Security Overview 56](#_Toc64965065)

[5.1 Application Security 56](#_Toc64965066)

[5.1.1 Data Centre and Infrastructure Security 56](#_Toc64965067)

[5.1.2 Shared Hosting Platform Security 57](#_Toc64965068)

[5.1.3 Third Party Access Security 58](#_Toc64965069)

[5.1.4 Network Security 58](#_Toc64965070)

[5.1.5 Mobile Platform Security 60](#_Toc64965071)

[5.2 Data Security 60](#_Toc64965072)

[5.2.1 Level of Assurance 61](#_Toc64965073)

[5.2.2 Data Encryption 61](#_Toc64965074)

[5.2.3 Auditing Assurance and Compliance 62](#_Toc64965075)

[5.2.4 Information Lifecycle Management 62](#_Toc64965076)

[5.2.5 Other Data Security Controls 62](#_Toc64965077)

[5.3 API Security 62](#_Toc64965078)

[5.4 Identity and Access Management 63](#_Toc64965079)

[5.4.1 Access Channels 63](#_Toc64965080)

[5.4.2 Identity, Access and Authorisation 64](#_Toc64965081)

[5.5 Logging, Monitoring and Alerting 65](#_Toc64965082)

[5.5.1 System Activity Monitoring 66](#_Toc64965083)

[5.5.2 Security Incident Management, eDiscovery and Cloud Forensics 67](#_Toc64965084)

[5.6 Threat, Vulnerability and Patch Management 67](#_Toc64965085)

[5.6.1 Software Maintenance and Security Patching 67](#_Toc64965086)

[5.6.2 Threat and Vulnerability Management 67](#_Toc64965087)

[5.6.3 Configuration Change Management 67](#_Toc64965088)

[5.7 Business Continuity 67](#_Toc64965089)

[5.7.1 Data Backup and Recovery 68](#_Toc64965090)

[5.7.2 High-Availability / Failover 68](#_Toc64965091)

[5.7.3 Disaster Recovery (DR) 68](#_Toc64965092)

[6 Application Management 70](#_Toc64965093)

[6.1 Application Code Controls 70](#_Toc64965094)

[6.2 Application Performance 70](#_Toc64965095)

[6.3 Application Defect Management 71](#_Toc64965096)

[6.4 Application Release Management 71](#_Toc64965097)

[6.5 Application Recovery Plan 71](#_Toc64965098)

[7 Data and Analytics Architecture 72](#_Toc64965099)

[7.1 Data Architecture 72](#_Toc64965100)

[7.1.1 Data Model Summary 72](#_Toc64965101)

[7.1.2 Model Diagram 72](#_Toc64965102)

[7.1.3 Major Entities 74](#_Toc64965103)

[7.1.4 Master and Reference Data 74](#_Toc64965104)

[7.1.5 Metadata Architecture 75](#_Toc64965105)

[7.1.6 Data Specification Chart 75](#_Toc64965106)

[7.2 External Data Standards 76](#_Toc64965107)

[7.3 Other Data and Analytics Considerations 76](#_Toc64965108)

[8 Appendix 77](#_Toc64965109)

[8.1 Pre-requisites and Dependencies 77](#_Toc64965110)

[8.2 Definitions 77](#_Toc64965111)

[8.3 Architecture Decision Log (including Solution Options decisions) 78](#_Toc64965112)

[8.3.1 Decision: <decision-title> 78](#_Toc64965113)

[8.4 Legacy Solution Diagrams or Descriptions 79](#_Toc64965114)

[8.5 References to other documents table 79](#_Toc64965115)

# Introduction

## Purpose of Document

This document describes the High-Performance Computing platform in MetWorx that will be used by the Clinical Pharmacology & Quantitative Pharmacology (CPQP) group to perform pharmacometrics (PK/PD) analysis on clinical data from EntimICE. The platform is used to run computational analysis tools, e.g., NONMEM, R/Rstudio, MATLAB, Monlix, etc. Analysis data, models and results are used for clinical trial reporting and submission. An Enterprise GitHub repository will be integrated with the MetWorx platform in the AstraZeneca VPC to store and track data for the business process.

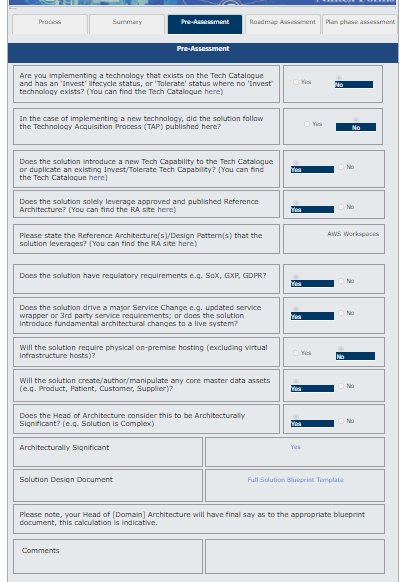
Included in the document is the rationale for the solution architecture and the design decisions and assumptions that led to its selection. The blueprint is kept live and updated with high level system changes and is retained according to system lifecycle.

## Architecturally Significant Assessment

The outcome of the Enterprise Architecture Review Board (EARB) Pre-Assessment is that the solution is architecturally significant. This is recorded in the Enterprise Solution Review Register with Id 846.

The basis for this outcome is <STATE the determining factor(s) for the outcome>.

The Pre-Assessment tab is reproduced below.



## Data Classification Assessment

The AstraZeneca information classification for the data used by the solution is: **Strictly Confidential**.

## Document Scope

* The scope of the document includes the following:
* Hosting and infrastructure.
* Identity, access management and security.
* Accommodation of non-functional requirements (resilience, scalability, performance etc).
* Technology platforms, frameworks and software components.
* Alignment to business capabilities and processes.
* Business criticality and service needs.
* Alignment to architecture standards.
* Design decisions taken and the assumptions on which the decisions and the architecture are based.

# Business Architecture Overview

## System Purpose

This Solution will provide a platform for the CPQP scientist to perform complex computational calculations with multiple AWS EC2 Linux instances that will be created as needed and terminated after computational analysis is completed. Each EC2 Instance will be created through the MetWorx dashboard directly onto a dedicated AZ AWS VPC. Additional storage for the EC2 instances will be provided by a GitHub Enterprise server also hosted on the AWS VPC, to provide the CPQP team the ability to check in and check out working data and patient data from the GitHub repository.

**Business Service**

The solution requires a new Business Service entry in ServiceNow.

In progress contacting CSM

ServiceNow Id: <snow-service-id>

ServiceNow Name: <snow-service-name>

## Business Context

The proposed solution was selected to address GxP compliance issues associated with using clinical data in the current non-validated solution using the Scientific Computing Platform (SCP) for both storage and computation. Metrum (vendor) provides an out of the box platform that will create EC2 instances with prequalified set of software. Since Metrum is an AWS partner, AWS was selected as the hosting platform. Outside of the EC2 instance storage, a GitHub Enterprise server will be hosted on the AWS VPC to provide the CPQP team to store working data and provide version control throughout the entire business process.

## Intended Value for Customer/User

The solution will provide improved GxP compliance over the current SCP application, particularly with its use of clinical data from EntimICE. MetWorx is GxP application fully supported and validated by the vendor making support by AstraZeneca IT and CPQP easier. It is also expected to deliver the following benefits:

* Improved GxP compliance and data integrity
* Improved efficiency and productivity for the end-to-end modelling/pharmacometrics activities
* Improved tracking and control of data in and out of the application and platform
* Improved tracking and metrics of computational parameters for each job
* Complete functional validation package from vendor

## Business Capabilities Provided

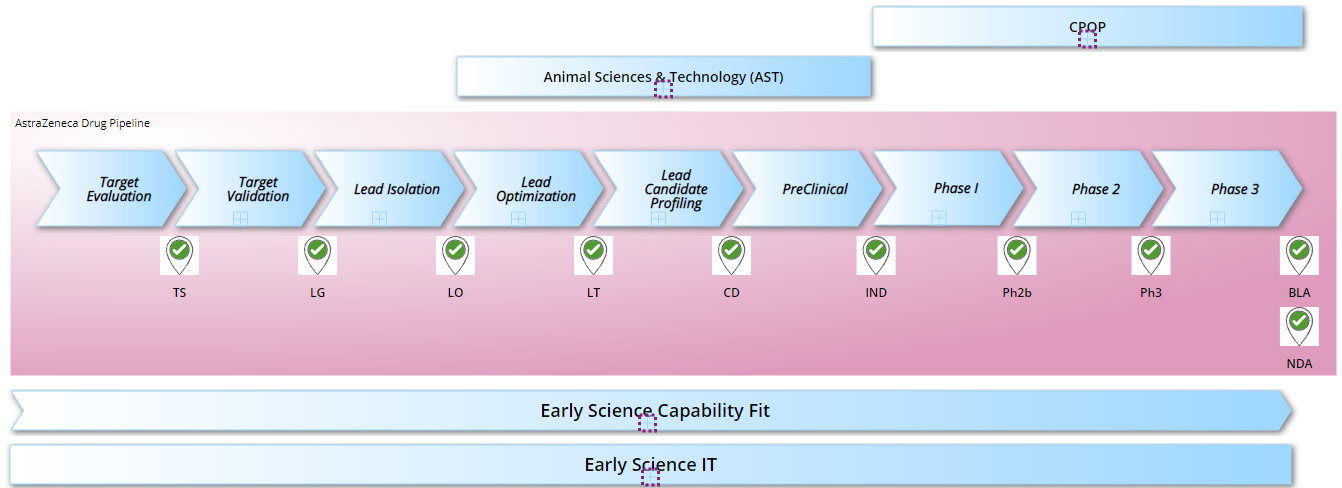
### Business Capabilities Supported

The MetWorx platform is a key component in attaining our Operational and Transformational IT Capabilities for CPQP. This business function plays a critical role in delivering the right dose to the right patient with the right safety. This effort requires robust end to end IT solutions for data storage, compliance, data analytics and data visualization. Our IT Capabilities need to be both Operational and Transformative for the business, i.e.,

* Compliant access for all current and evolving clinical data
* Accessible to the Big Data world
* Integrated quantitative analytics
* Metadata management
* Collaborative and flexible platforms in the cloud

### Value Streams Supported

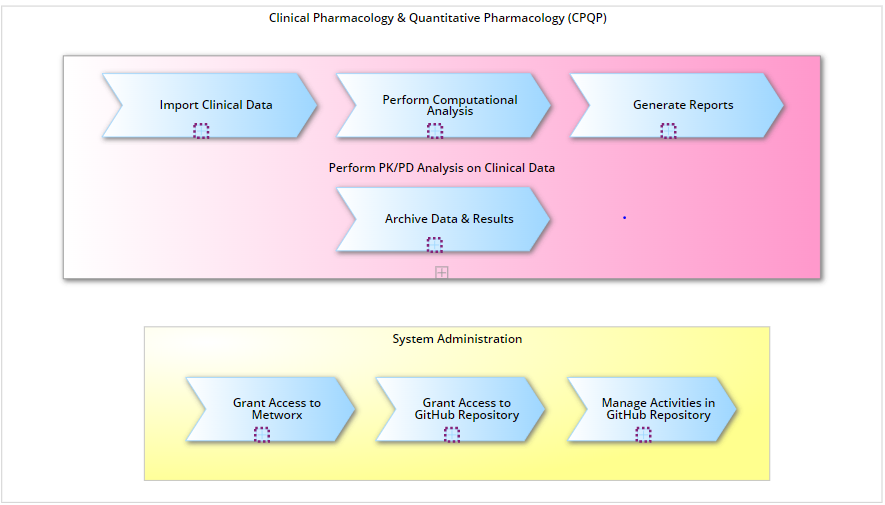
The MetWorx platform plays a critical role in supporting Exposure-Response (PK/PD) analysis for Phase 1 through Phase 3 clinical trials. Activities for this value stream include dose selection and optimization, drug efficacy versus safety and quantitative approaches using clinical trial simulations to develop clinical models for the drug in patients. The IT Pre-clinical capability in Early Science supports these activities for CPQP.



### Impact on Application Estate

Refer to Section 2.3

## Business Process Overview



**Business Process**

The business process can be classified into the following ‘Themes’ or ‘Use Cases’:

* **Import**: The system will capture and import the clinical data from EntimICE and manage the clinical data and analyzed results within the GitHub repository.
* **Analyses**: The system will perform complex computational analyses (PK/PD) of clinical trial data for regulatory submissions.
* **Report**: The system will support the generation of analysis reports that allow the data stored in the centralized data repository to be queried, viewed, plotted or extracted in a user specified manner.
* **Archive**: The system will support the archival of data stored in the GitHub repository back to EntimICE as per the AZ GRAD code for PK/PD clinical data.
* **System Administration**: The system will support administration of access to the MetWorx platform and the GitHub repository in the AZ VPC.

## User / Roles Overview

The solution accommodates the following user types:

|  |  |  |
| --- | --- | --- |
| Application Role/Group | Organisation Role(s) | Location (#Users in Application Role at Location) |
| GitHub Administrator | Administrator | 1 admin in GB |
| AWS Administrator | Administrator | 1 admin in GB |
| MetWorx Administrator | Administrator | 1 admin in GB |
| GitHub\MetWorx User | User | 100 users globally |
| GitHub Browser | User | 100 users globally |

Estimated average user concurrency is 30. This is the average number of users expected to be active in the system over time.

Estimated maximum user concurrency is 100. This is the maximum number of users expected to be active at peak times, i.e. the high watermark.

## Regulatory Position

**GAMP Category**

The solution has been assigned GAMP Category 4 as per RID (AZDoc0203162).

### Sarbanes-Oxley (SOx) Position

The system does not have a Sarbanes-Oxley regulatory impact as per RID (AZDoc0203162).

### Data Privacy and General Data Protection Regulation (GDPR) Position

As per the RID (AZDoc0203162) and the Privacy Impact Assessment (PIA), PIA-5812, the system does have a data privacy impact since it used clinical data from EntimICE.

### Good Pharmaceutical Practice (GxP) Position

The Risk Impact Determination (RID) assessment (AZDoc0203162) determined that GxP applies system. As far as 21 CFR Part 11, electronic records apply to the system, but excludes electronic signature applicability.

### Information Management Position

The RID assessment determined that the solution requires provision for the retention and disposal of company information under the AstraZeneca GRAD schedule.

### eDiscovery Position

The outcome of the RID is that the solution is not subject to eDiscovery controls because none of the eDiscovery criteria apply.

### Medical Device Classification

The outcome of the RID is that the solution does not fall within the classification of a medical device, as defined by the Medical Devices Regulation, because none of the criteria apply

### Other Applicable Regulations

None identified.

## Criticality Statement

The business criticality of the service provided by the solution is Somewhat Critical

The determining factors for this are:

* Violation of regulatory compliance

## Resiliency Statement

The business requirements for resilience and business continuity are:

* **Recovery Point Objective (RPO):** Greater than 24 hours data updates are to be lost as a result of an unplanned service outage.
* **Recovery Time Objective (RTO):** Between 4 and 24 hour  RTO  for service is to be restored to normal operation following an unplanned service outage
* **Backup and Recovery:** GitHub Enterprise server will utilize AWS snapshot daily with a retention of seven days and weekly backups for at least six months
* **Disaster Recovery (DR):** No provision is made for disaster recovery because the output of the BIA does not require DR

## Externalisation Statement

The MetWorx\GitHub solution will not be accessed by external users.

# Architecture Alignment

## Alignment to EA Principles

The proposed solution addresses the AstraZeneca Enterprise Architecture principles as follows:

* Enterprise First: The MetWorx\GitHub solution will be used to provide a more stable and flexible cloud based environment to CPQP scientist perform analysis.
* **Steward the Enterprise:** This solution uses existing AZ enterprise solutions (AWS- GitHub) that can be incorporated with existing frameworks, processes and collaboration models.
* **Foster Reuse:** The MetWorx\GitHub solution provides a specific platform for the CPQP team, it cannot be reused.
* **Align to Business Strategy and Value:** The MetWorx\GitHub has been developed to meet validation and GxP requirements for use of clinical data.
* **Design for Change:** The MetWorx\GitHub solution has its own lifecycle from the vendor (Metrum). AstraZeneca has a service agreement to provide and support version upgrades.
* **Customer-Centric Design:** Both MetWorx and GitHub both provide a dashboard and CLI interface to provide an intuitive platform to conduct analysis
* **Data is a Competitive Strategic Asset:** GitHub provides a check-in\check-out approach to accessing data to provide version control and data protection
* **Access Anytime, Anywhere from any Device:** MetWorx\GitHub can be access via any browser and any AZ approved devices
* **Open but Compliant and Secure:** MetWorx\GitHub will support AD based roles
* **Operational Readiness by Design:** Short term support will be delivered by the project team until it is transitioned to the R&D IT application support team

## Alignment to Strategies and Roadmaps

The solution supports the Capability Roadmap for CPQP within Pre-clinical

## Alignment to Standards and Reference Architectures

The solution aligns to the following standards:

* Cloud computing standards
* GitHub enterprise
* Buy before build

## Development and Delivery Model

### Development Lifecycle Model

The system development lifecycle method used to develop the solution is Agile

### Development Model

MetWorx and GitHub platform are COTS developed by Metrum and GitHub. Once configuration is complete no customization of the software will need to be performed.

### Delivery Model

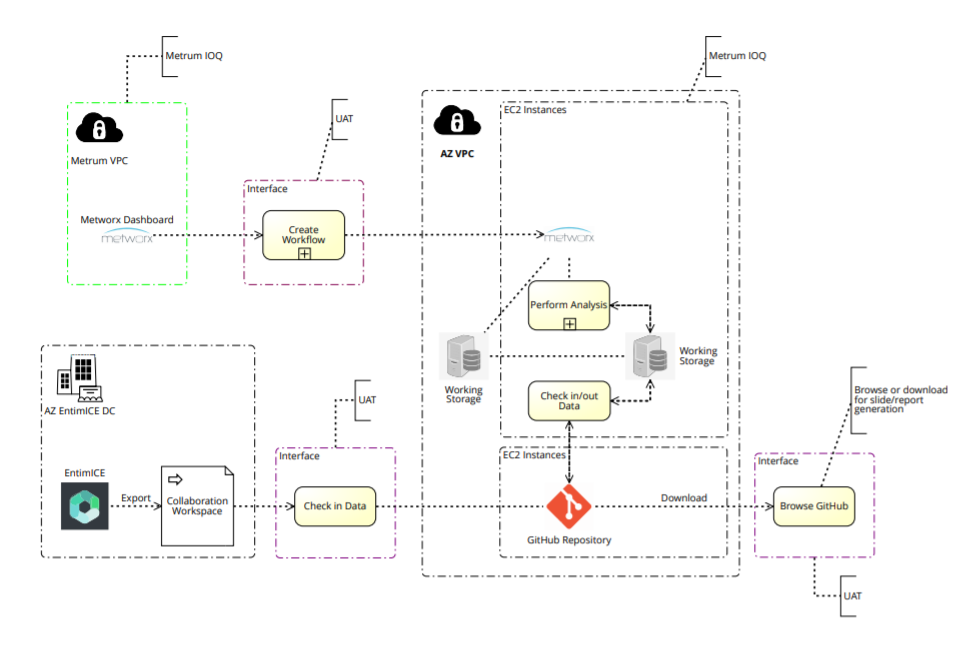
The service is delivered to users through via a web based dashboards. The MetWorx dashboard will be externally hosted, but the GitHub enterprise server will be hosted on an AZ AWS VPC. All MetWorx instances will be hosted on the same AZ AWS VPC.

# Solution Architecture

## Conceptual and Logical Architecture

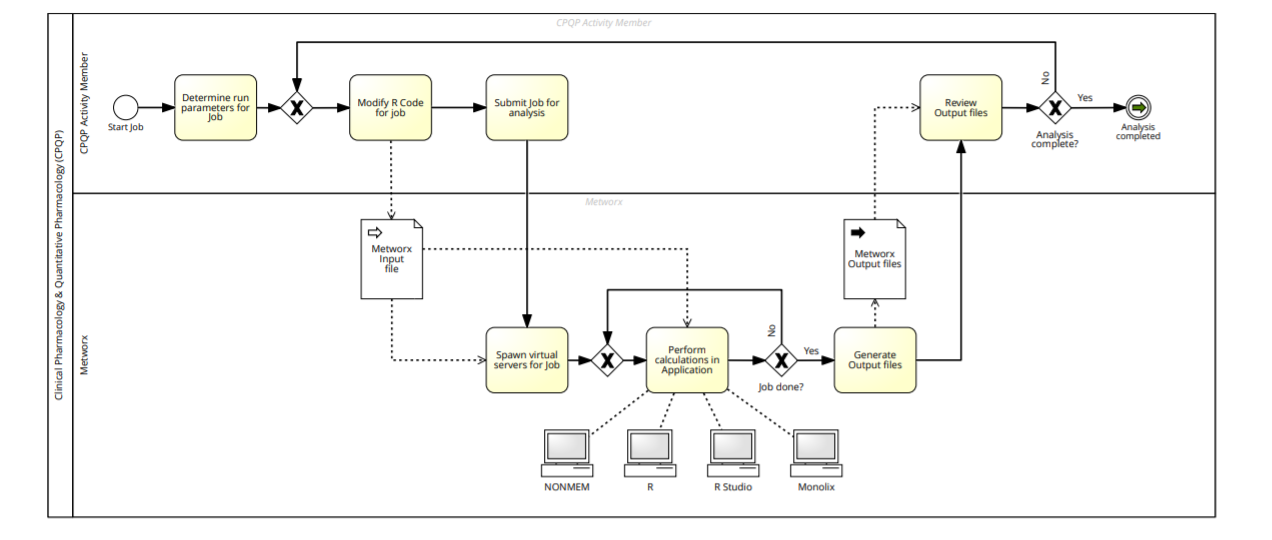
### Conceptual Solution Architecture

The Following solution will allow CPQP scientist to create workflows for analyses from the MetWorx Dashboard (top left corner). After the workflow is submitted via the dashboard, a Linux based EC2 instance will be launched on an AZ AWS VPC according to the specs specified through the workflow request submitted by the CPQP scientist. Once the EC2 instance is running, the CPQP scientist can logon to the instance via Rconsole. CPQP users can also check data into GitHub via a browser which will them be used during CPQP analysis that will be performed on the MetWorx instance



### Logical Solution Architecture

Logical Solution Architecture



## Technical Solution Architecture

### Solution Technologies and Lifecycle Status

The technologies used in the solution are listed in the table below, including their lifecycle status, as specified in the AstraZeneca Enterprise Architecture [Technology Catalogue](https://azcollaboration.sharepoint.com/sites/AA311/Lists/Technology%20Catalogue/Default.aspx).

| Technology | Domain Team | Lifecycle Status |
| --- | --- | --- |
| AWS VPC | Infrastructure | Invest |
| AWS EC2 (Linux RHE 7) (supplied by MetWorx) | Infrastructure | Invest |
| GitHub Enterprise sever | Infrastructure | Invest |
|  |  |  |

**Exceptions**

No exceptions were raised for this solution.

### Enterprise Middleware Consumption Overview

The Following middleware service will be used:

Active Directory – GitHub and MetWorx will both utilize ADFS for SSO services

Enterprise Messaging- Email services will used for reporting and alerts

### Communication Flow between components

*Refer to Conceptual Diagram 4.1.1*

### Software of Unknown Provenance (SOUP) / Software as a Medical Device (SaMD)

None of the components of the solution fit the definition of Software as a Medical Device (SaMD).

## Infrastructure

### Hosting Summary

The solution is hosted externally in an AstraZeneca AWS cloud account.

The cloud infrastructure level is platform-as-a-service (PaaS) using AWS services.

The hosting region is US-East-1

### Clustering / Containerisation / Replication

No components of the solution use clustering, containerisation or data replication technologies.

### Implementation Architecture Diagram

Implementation Architecture

Refer to diagram in section 4.1.1

| Technology | Server name \ID | Platform |
| --- | --- | --- |
| AZ VPC | az-metworx-poc \ 780013771074 | AWS |
| GitHub Enterprise server (Dev) | uscvalgheapdv01 | EC2 Linux AWS |
| GitHub Enterprise server (Prod) | Uscvalgheappd01 | ECS Linux AWS |
| MetWorx Instances | Dynamic servers – created and terminated as needed | ECS Linux AWS |
|  |  |  |
|  |  |  |

### Network Diagram

Refer to diagram 4.3.3

### System Component Table

Refer to diagram 4.3.3

## Environments

The solution comprises of one environment for MetWorx, supported by the vendor. The internal GitHub Enterprise server will consist of one Dev Server and one Prod server

All environments will be hosted on the AWS cloud services platform.

The purpose of the GitHub Development environment is to enable testing of new releases and configurations, and will provide and platform to develop future processes and procedures.

The Dev environment comprises of one server and will be sized at according to vendor specification

The Production GitHub environment is comprised of one GitHub enterprise. The sizing is specified in the System Component Table.

### Environment Overview Diagram

Refer to diagram 4.3.3

## Integrations

### System Context / Interfaces Diagram

Refer to 4.3.3

### Interfaces Exposed

No interfaces are exposed by this solution for other systems and users consume.

### Interfaces Consumed

No integrations exposed by other solutions are consumed by this solution.

#### Data Management and Integration Technologies

The table below shows the integration technologies used by the solution.

| Integration Technology | Domain | Lifecycle Status | Standard or Non-Standard? | Used  (Y/N)? | Review Outcome |
| --- | --- | --- | --- | --- | --- |
| PingFederate | Security | Invest | Standard | Y |  |

Integration Technologies

## Support and Governance Model Statement

**Support and Governance Model**

The service provided is not in the scope of a platform. We are following a standard ATS process and will have a combination of support from the BAU support team, Metrum and GitHub as needed

**System Owner**

The System Owner for the solution is Chris Penland, System Owner and Process Owner.

**Service Hours**

The hours during which the production system will be available are as follows:

| Region | Day(s) | Hours |
| --- | --- | --- |
| US EST | Sun – sun | 24 hours |
|  |  |  |

Production Service Hours

**Support Hours**

The hours during which support will be available for the production system are as follows:

| Region | Day(s) | Hours |
| --- | --- | --- |
| US EST | M-F | 5:00 – 18:00 |
|  |  |  |

Production Support Hours

## User Experience Overview

### User Experience Measures

User experience measurements were taken as part of the POC process

### User Experience Design

COTS product user experience was evaluated during the POC

### User Experience Journey Maps

N/A

## Data Refresh and ETL Overview

The solution has no specific requirements for data ingestion, refresh or bulk data load processes.

No Operational Level Agreements (OLA) are in place for the availability of data from source systems

## e-Discovery Response Overview

The outcome of the RID is that the solution is not subject to eDiscovery controls.

## Scalability and Elasticity Overview

The solution’s capacity management and monitoring approach, including how the solution is equipped to respond to increases and/or decreases in demand for compute resources.

Resources are increased and decreased according the CPQP users demands, programable through the MetWorx dashboard.

## System Lifecycle Plan

This is the first implementation of the solution. Metrum research is coupling the on demand GxP qualified compute environment.

## Component Lifecycle Overview and Evergreen Statement

Refer to section 4.11.

## High Availability / Resiliency / COOP / BCP / DR Description

The entire environment is hosted in AZ AWS VPC. Since it is hosted in AWS we will benefit from the standard AWS resilience practices. In addition we will take daily AWS snapshots of the GitHub enterpriser server to ensure that we meet the required recovery point. The CPQP platform is not classified as a critical system.

## License Model Statement

The licensing model for the solution components is as follows.

* GitHub Enterprise
* Metrum MetWorx
* Application licenses if needed that run on MetWorx Instances ( e.g. NonMem, RStudio, Monolix…)

## Funding Model

The Project was funded as part of the IT Pre-Clinical Portfolio. The business will fund the ongoing MetWorx\AWS\GitHub usage and licences.

## Sustainability Statement

By using compute as needed we are reducing the footprint. When analysis are complete EC2 instances are terminated

# Security Overview

The security controls applied to the solution are described below.

## Application Security

### Data Centre and Infrastructure Security

The solution is hosted externally by AWS See “Hosting Summary” in section 4.3 for further details.

Refer to AstraZeneca’s service agreement for further details re regarding infrastructure security.

### Shared Hosting Platform Security

The solution is hosted on a shared, multi-tenanted hosting platform.

A dedicated AWS VPC has been created to host the MetWorx\GitHub solution

### Third Party Access Security

Access to AstraZeneca hardware and software assets by a third party service provider is not required for this solution.

### Network Security

**Network Protocols**

The solution uses AstraZeneca standard network protocols.

**Network Access Controls**

The solution uses the standard facilities provided by the AstraZeneca network and the hosting provider.

**Access from Outside the AstraZeneca Network**

The solution will not be accessible from outside the AstraZeneca network

### Mobile Platform Security

No part of the system is delivered on a mobile device platform.

## Data Security

### Level of Assurance

The AstraZeneca Information Classification assigned to the information assets used by the solution generally is: Strictly Confidential. This requires Level of Assurance 3 protection.

The data being copied out of EntimICE to the GitHub Enterprise Server is encrypted

### Data Encryption

**Encryption at Rest**

Data stored by the system is protected as follows:

When an analysis is started, data is manually copied from EntimICE collaboration workspace (existing architecture) directly to a folder on the GitHub server hosted on the dedicated AZ AWS VPC. During the analysis data will be checked in and checked out of GitHub Enterprise server by the CPQP team.

EAS 256

**Encryption in Transit**

Data transferred by the system is protected as follows:

*TLS 1.2*

### Auditing Assurance and Compliance

The auditing requirements for the solution have been addressed as follows: GitHub Enterprise server has an audit trail that tracks data sets that are checked in and checked out. The MetWorx computational tools are tracked as each user starts and ends a workflow.

### Information Lifecycle Management

The requirements for tracking and managing information throughout its lifecycle have been addressed as follows: Once the computation modelling is completed all data will be checked back in to EntimICE.

### Other Data Security Controls

No other data security controls are defined.

## API Security

The solution does not provide or consume any Application Programming Interfaces (API).

## Identity and Access Management

### Access Channels

Users and Admins will access the solution using the corporate desktop and browsers

### Identity, Access and Authorisation

#### Authentication

The system will be accessed only by AstraZeneca employees.

**Internal Users**

AstraZeneca users authenticate to the system as follows: PRID

#### Single Sign-on

Single Sign-On (SSO) will be enabled, via GitHub and MetWorx SAML SSO

#### Authorisation

Each user has their own MetWorx on-demand workflow. Credentials for GitHub Enterprise will be passed through the MetWorx workflow.

#### User Access Management

User access will be controlled through Active Directory and IDM process.

#### Access Key Management

N\A

## Logging, Monitoring and Alerting

### System Activity Monitoring

The requirements for logging and monitoring of system activity and alerting of events and incidents are as follows:

MetWorx sends a daily report of all user activity and systems resources consumed

**ServiceNow Integration**

An interface to ServiceNow is not available yet

### Security Incident Management, eDiscovery and Cloud Forensics

We will follow standard AWS and AstraZeneca policies

## Threat, Vulnerability and Patch Management

### Software Maintenance and Security Patching

All software components of the solution and threat management will be updated and kept secure by the vendor (Metrum) through our service agreement.

### Configuration Change Management

Metrum performs all update and manages all change management processes

ETS will manages all AWS VPC changes

## Business Continuity

### Data Backup and Recovery

Refer to Section 4.13

### High-Availability / Failover

HA and failover not required

### Disaster Recovery (DR)

Refer to Section 4.13

# Application Management

## Application Code Controls

N/A

## Application Performance

Resources are scalable by design according to user specs using MetWorx instances in AWS

## Application Defect Management

Provided by the vendor (Metrum), since they supply the qualified compute environment

## Application Release Management

Provided by the vendor (Metrum), since they supply the qualified compute environment

## Application Recovery Plan.

Provided by the vendor (Metrum), since they supply the qualified compute environment

# Data and Analytics Architecture

## Data Architecture

### Data Model Summary

Diagram

Description automatically generated with medium confidence

### Master and Reference Data

The master and reference data entities used by the solution are listed in the table below, including a description of the solution’s alignment to the applicable data standards.

|  |  |  |  |
| --- | --- | --- | --- |
| Conceptual Data Entity | Standard or Ontology | Master Reference Source | Master Data Guidance |
| [CLINICAL STUDY](https://astrazeneca.collibra.com/asset/f501f6d8-bfe3-4fbe-bce0-a452a3597064) |  | EntiMICE | Process is tightly coupled with EntiMICE hence it should apply EntiMICE identifiers. If there any issues it needs to be addressed by EntiMICE not by this solution. |
| MODEL |  | Locally Managed | Strategically identifiers should be generated using a PID/PURL capability. |
| MODEL VERSION |  | Locally Managed | Strategically identifiers should be generated using a PID/PURL capability. |
| [DATA ANALYTICS EXPERIMENT](https://astrazeneca.collibra.com/asset/50ee96f5-3d3b-483c-801e-a73ab2cc6d96) |  | Locally Managed | Strategically identifiers should be generated using a PID/PURL capability. |

Master and Reference Data: Alignment to Data Standards

### Metadata Architecture

N/A

## External Data Standards

N/A

# Appendix

## Pre-requisites and Dependencies

No pre-requisites or dependencies are identified for the solution.

## Definitions

| Term | Definition |
| --- | --- |
|  |  |
|  |  |
| External access | A means of connecting to a system from outside the AstraZeneca network and which does not rely on “onboarding” the user by providing them with a PRID, i.e. registering the user’s identity in the AstraZeneca Identity and Access Management (IAM) system / Active Directory. |
| External user | A non-AstraZeneca employee. This includes marketing company and partner representatives. |
| Externalisation | The provision of external access to a system’s services. See “External access”.  Note: A system that provides access to external users exclusively by onboarding them into the AstraZeneca Active Directory, i.e. allocating them with a PRID, and providing some means of connection to the AstraZeneca network, e.g. using an AstraZeneca laptop or a Remote Access Platform (RAP) account or the installable operating system image (TPIC), is not an externalised system. |
| GAMP | Good Automated Manufacturing Practice. A set of guidelines issued by the technical subcommittee of the International Society for Pharmaceutical Engineering for manufacturers and users of automated systems in the pharmaceutical industry. |
| IMDRF | International Medical Device Regulators Forum (IMDRF). The body responsible for defining the SaMD regulation, compliance to which is required by the FDA. |
| RPO | Recovery Point Objective. The maximum elapsed time for which the business can tolerate a loss of data updates as a result of a system failure or data corruption. Factors include the criticality of the data and the transaction frequency. |
| RTO | Recovery Time Objective. The maximum period of downtime of a system that can be tolerated without adverse impact on business operations. |
| SaMD | Software as a Medical Device. A standard defined by the IMDRF as: Software intended to be used for one or more medical purposes that perform these purposes without being part of a hardware medical device. Compliance is required by the FDA. |
| SSO | Single sign-on. An authentication scheme that allows a user to log in with a single ID and password to any of several related, yet independent, software systems. |
| SOUP | Software of Unknown Provenance. A standard defined by the International Electrotechnical Commission (IEC). It refers to a generally available off-the-shelf software product that is used in relation to a medical device, but has not been developed specifically for the purpose of being incorporated into it. |

Definitions (Glossary of Terms)

## Architecture Decision Log (including Solution Options decisions)

No decisions or exceptions are recorded for this solution.

## Legacy Solution Diagrams or Descriptions

This solution replaces the use of SCP by CPQP

## References to other documents table

| Ref | Document | Location |
| --- | --- | --- |
| 1 | <document title> | <location name with embedded link> |
|  |  |  |

Document References