| Simulators Team – N x simulators access the API | | API definition??? | Data Lake Team - Cloud | | | Analytics Env. |
|---|---|---|---|---|---|---|
| Data Generators | Data Formatter | API | Data Ingest and ETL | Persistent Common Data Model | Data Mart Generation (ETL) | TRE |

Integration Engine (e.g. Rhapsody or Mirth) to generate HL7 messages

Rest API Oauth2 or WebAuth

Accepts interoperable standards only

Defined through OpenAPI

Integration Engine (e.g. Rhapsody or Mirth) to read HL7 messages

Pentaho, Talend, Nifi, Databricks for scheduled 'jobs' and management console.

API

This component should be interchangeable based on well defined interfaces

API

Pentaho, Talend, Nifi, Databricks

Sets of datamarts, provisioned in TRE.

Consideration of cost of TRE access – should a copy be created for each provisioned environment to limit data traffic?

TRE may be hosted in private cloud with a user subscription model.

The Data Lake and the TRE data marts should be 'air gapped' for security reasons



| Simulation team | API | Data ingest | Data warehouse | Data analytics | Security, monitoring and compliance |
|---|---|---|---|---|---|

(Oauth2)
(Token)

Amazon **Cognito**

(auth)

(HL7 FHIR)

Amazon **API Gateway**

AWS **Glue** (data catalog)

Amazon **Lambda**

(crawlers)

Amazon S3

databricks

Amazon **Redshift**
- Query and write in open formats or ANSI SQL
- Federated query

Amazon **KMS** (encryption)

AWS **EMR**
- Run and scale big data frameworks
- Notebooks

Machine Learning

Amazon SageMaker

Business Intelligence

amazon **QuickSight**
(visualisations & dashboards)

Amazon CloudWatch (logs)

Amazon IAM (roles)

AWS **Control Tower**

AWS **Security Hub**

AWS **CloudTrail**

**Simulation team**

**API**

**Data ingest**

**Data warehouse**

**Data analytics**

**Security, monitoring and compliance**

(Oauth2)

(Token)

Amazon **Cognito**

(auth)

(HL7 FHIR)

Amazon **API Gateway**

AWS **Glue** (data catalog)

databricks

Amazon **Lambda**

(crawlers)

Amazon S3

Amazon **Redshift**
- Query and write in open formats or ANSI SQL
- Federated query

Amazon **KMS** (encryption)

Amazon CloudWatch (logs)

Amazon IAM (roles)

AWS **EMR**
- Run and scale big data frameworks
- Notebooks

Machine Learning

Amazon SageMaker

Business Intelligence

amazon QuickSight (visualisations & dashboards)

AWS **Control Tower**

AWS **Security Hub**

AWS **CloudTrail**