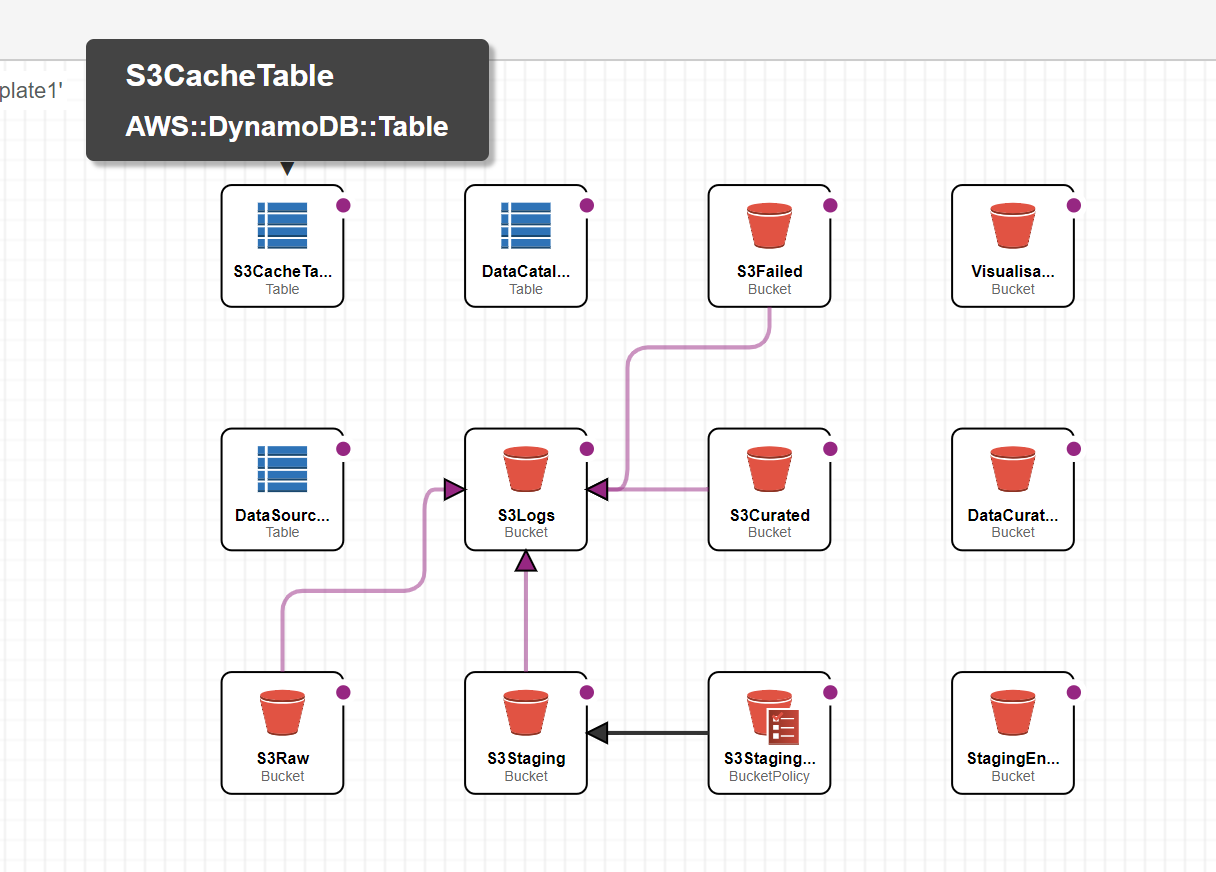
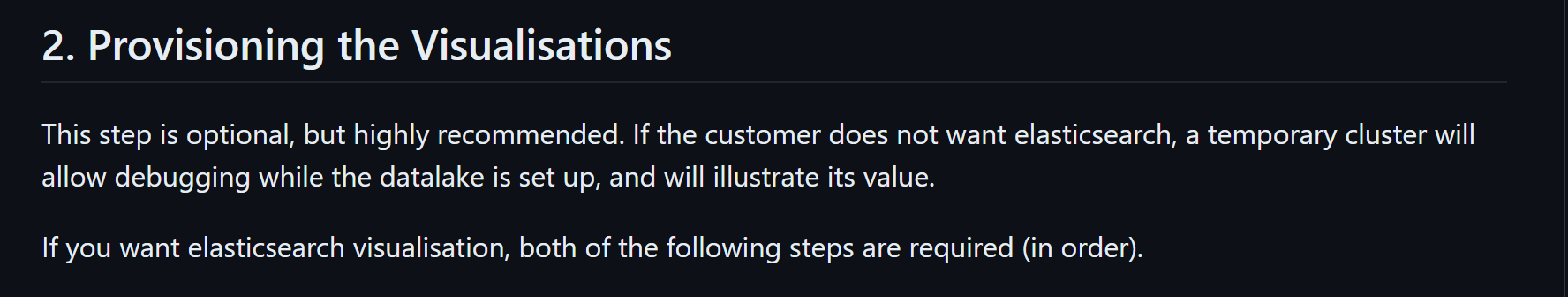
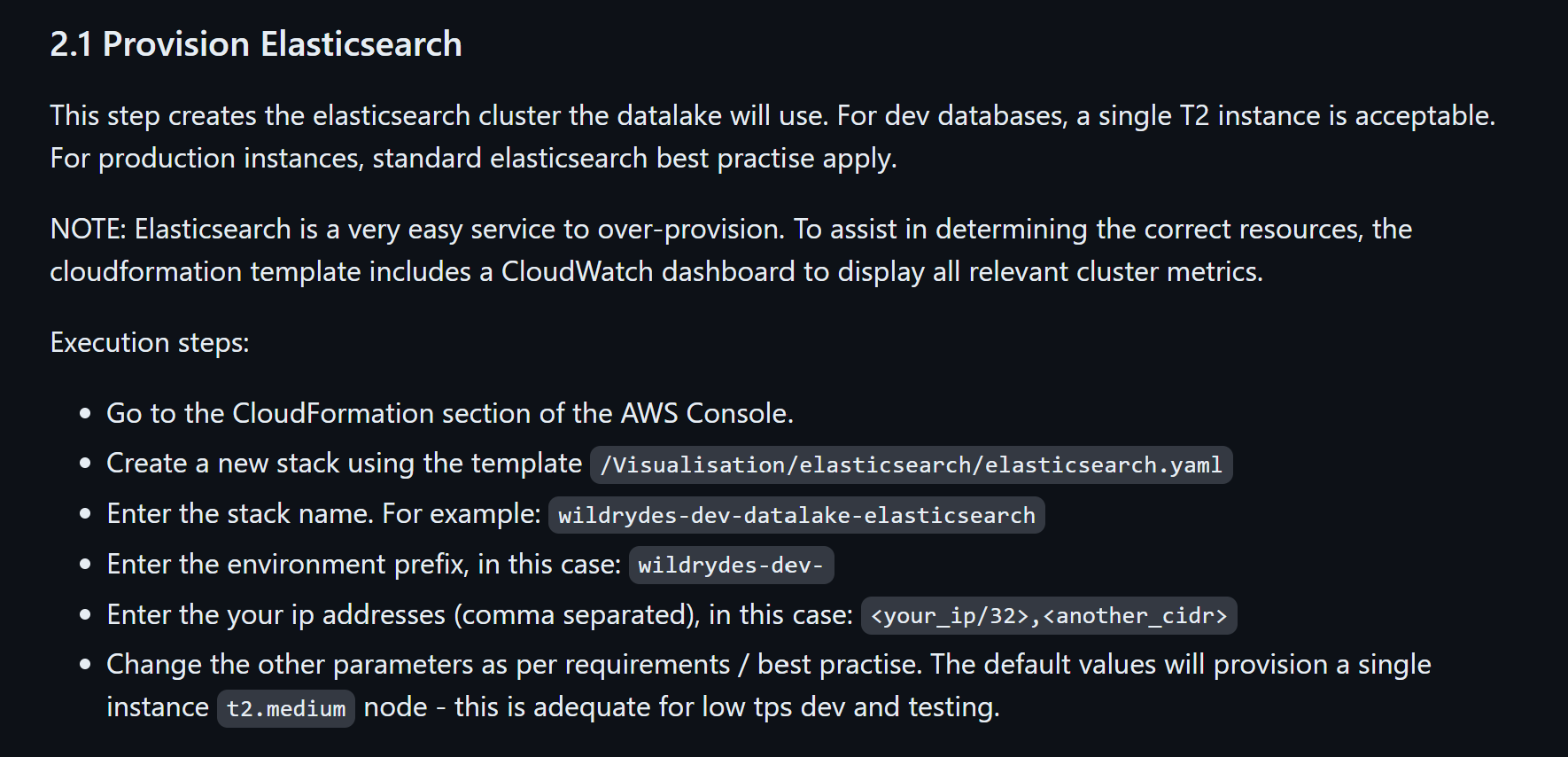
|  |  |  |  |
| --- | --- | --- | --- |
|  |  | [accelerated-data-lake](https://github.com/aws-samples/accelerated-data-lake) | Nov 2023 |
|  |  | <https://github.com/aws-samples/accelerated-data-lake> | |
| CloudFormation   * wildrydes-dev-datalake-structure * wildrydes-dev-datalake-elasticsearch * wildrydes-dev-datalake-elasticsearch-lambdas * wildrydes-dev-datalake-staging-engine |  | **STEPS:** | |
|  |  |  | |

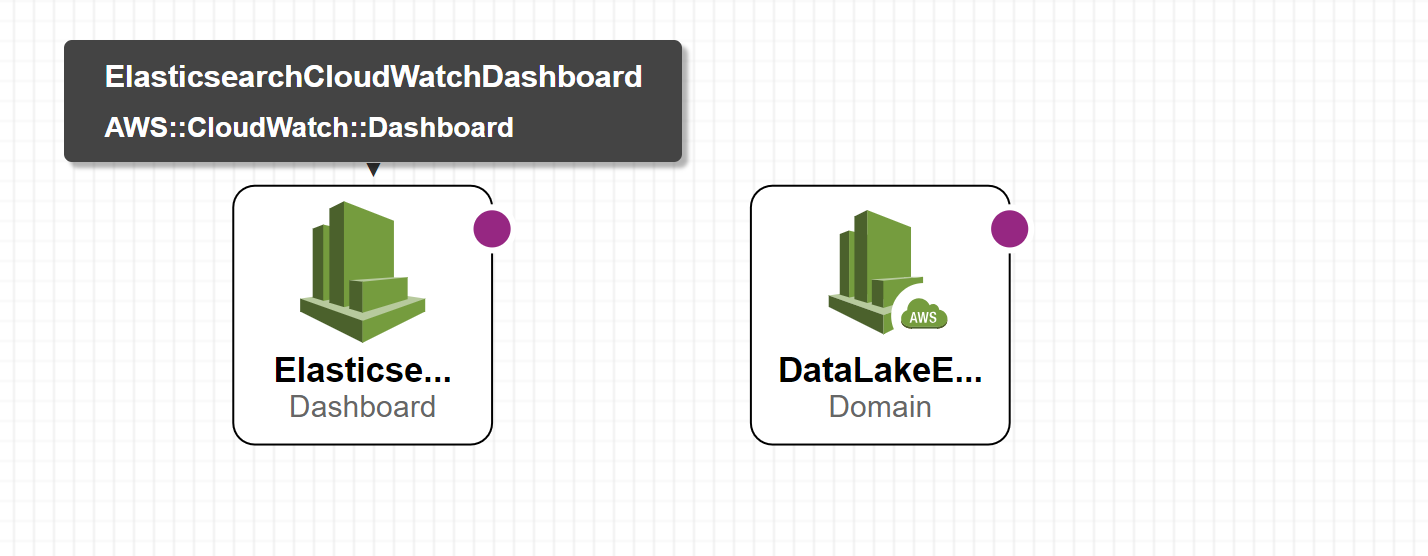
|  |  |  |
| --- | --- | --- |
| **wildrydes-dev-datalake-structure (CloudFormation)** |  |  |
|  |  | Resources  * DataCatalogTable * DataCurationCodePackages * DataSourceTable * S3CacheTable * S3Curated * S3Failed * S3Logs * S3Raw * S3Staging * S3StagingBucketPolicy * StagingEngineCodePackages * VisualisationCodePackages |

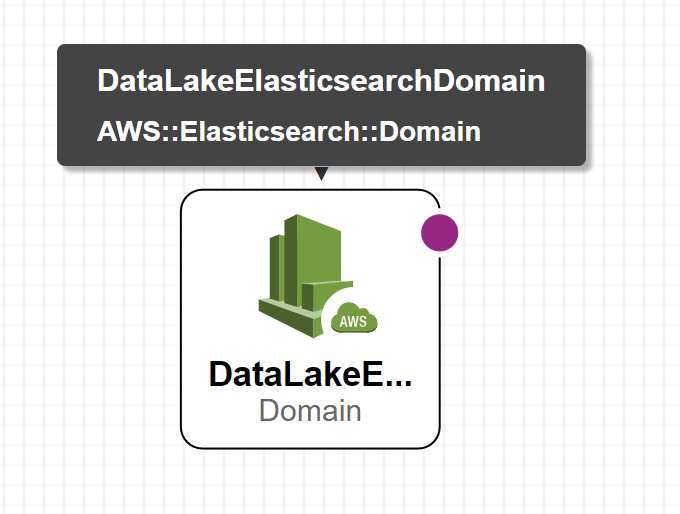


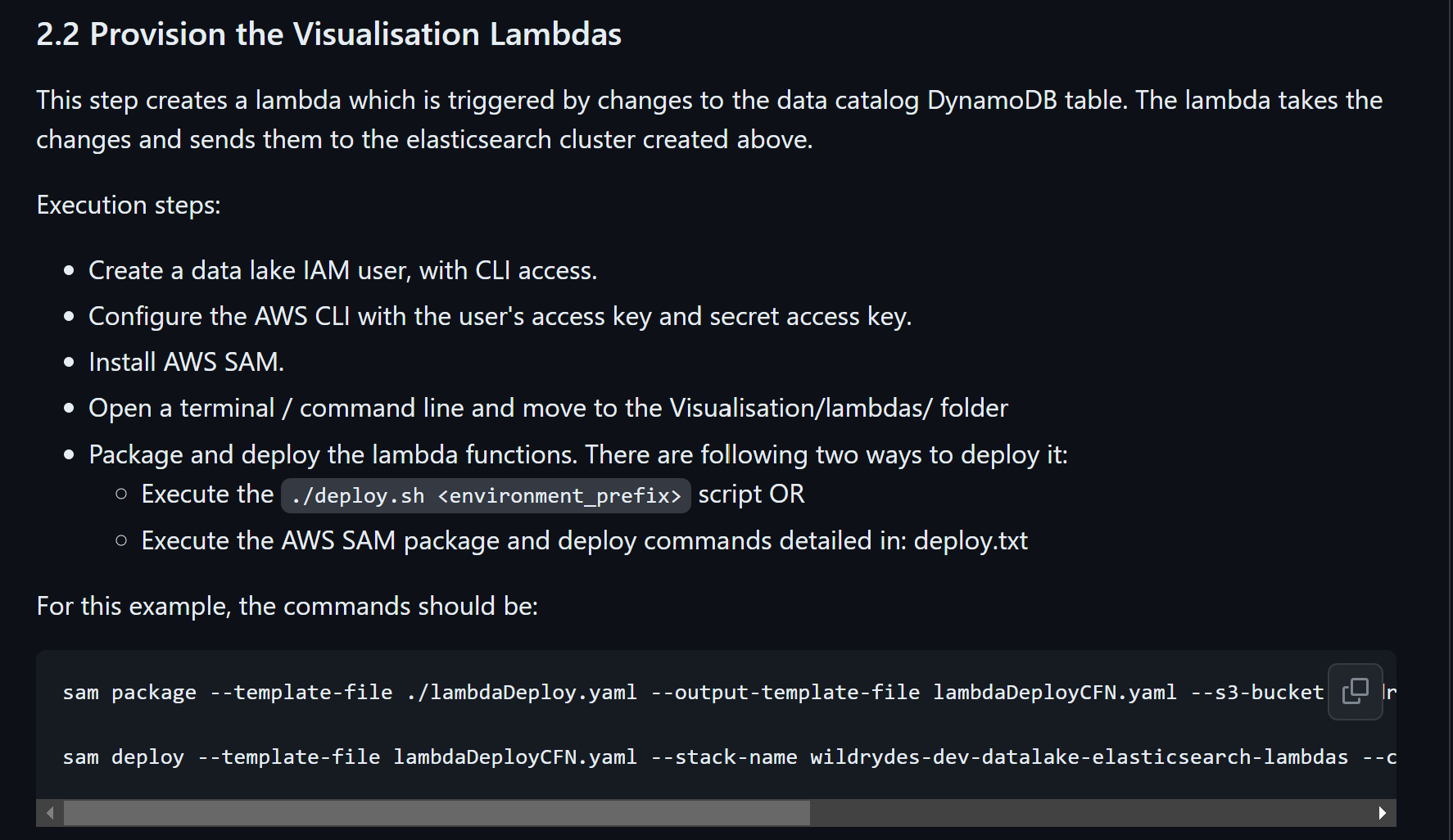




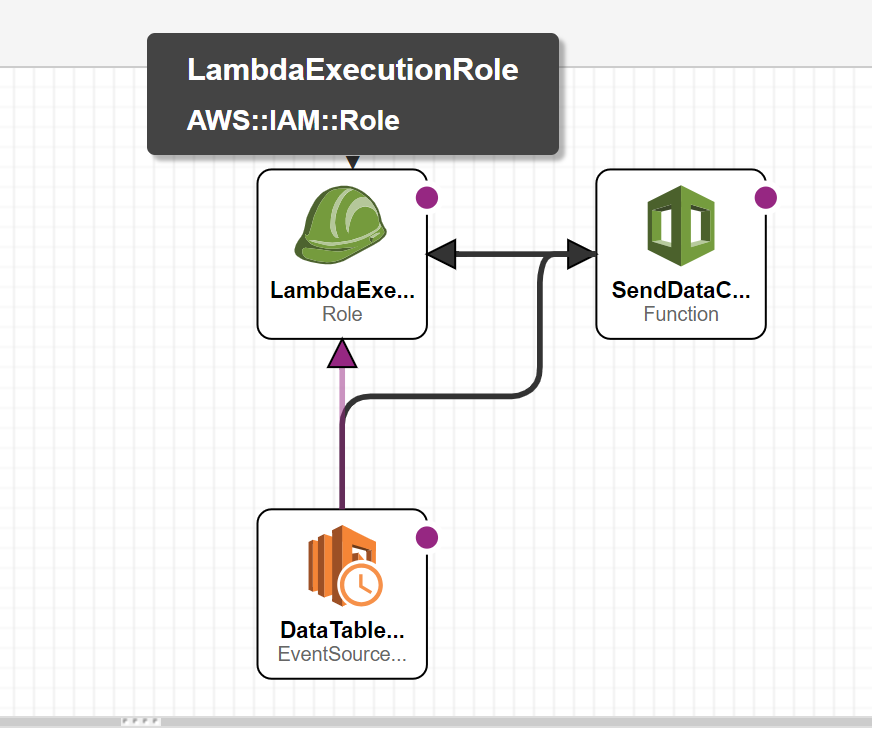
|  |  |  |
| --- | --- | --- |
| **wildrydes-dev-datalake-elasticsearch (CloudFormation)** |  |  |
|  |  | Resources  * DataLakeElasticsearchDomain * ElasticsearchCloudWatchDashboard |

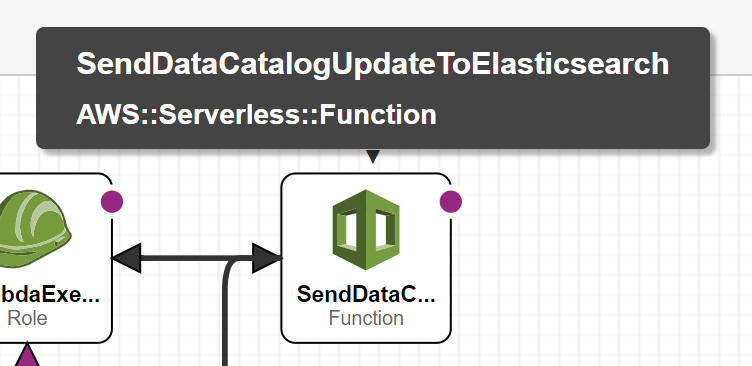




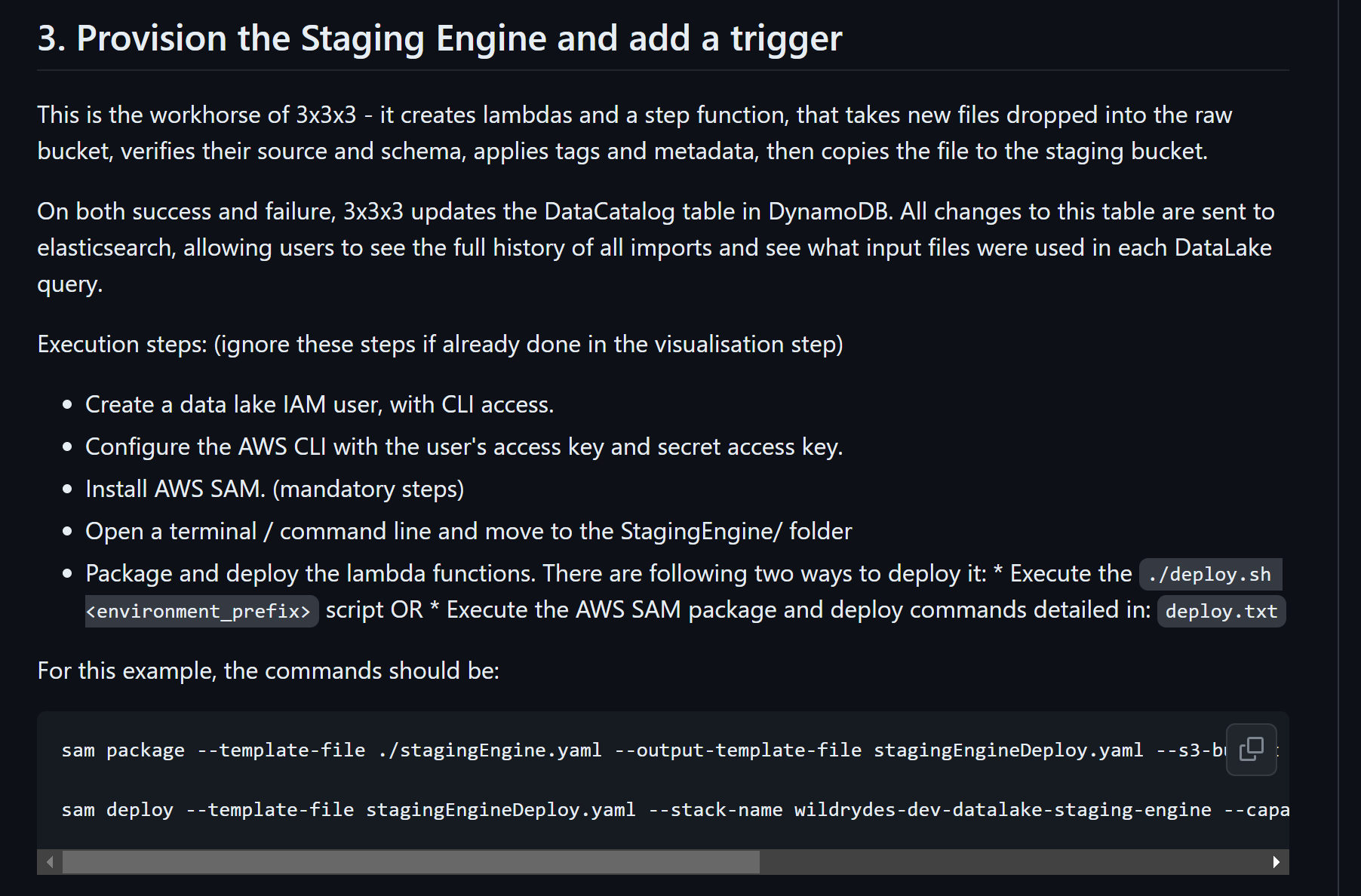


|  |  |  |
| --- | --- | --- |
| **wildrydes-dev-datalake-elasticsearch-lambdas (CloudFormation)** |  |  |
|  |  | Resources  * DataTableStream * LambdaExecutionRole * SendDataCatalogUpdateToElasticsearch |

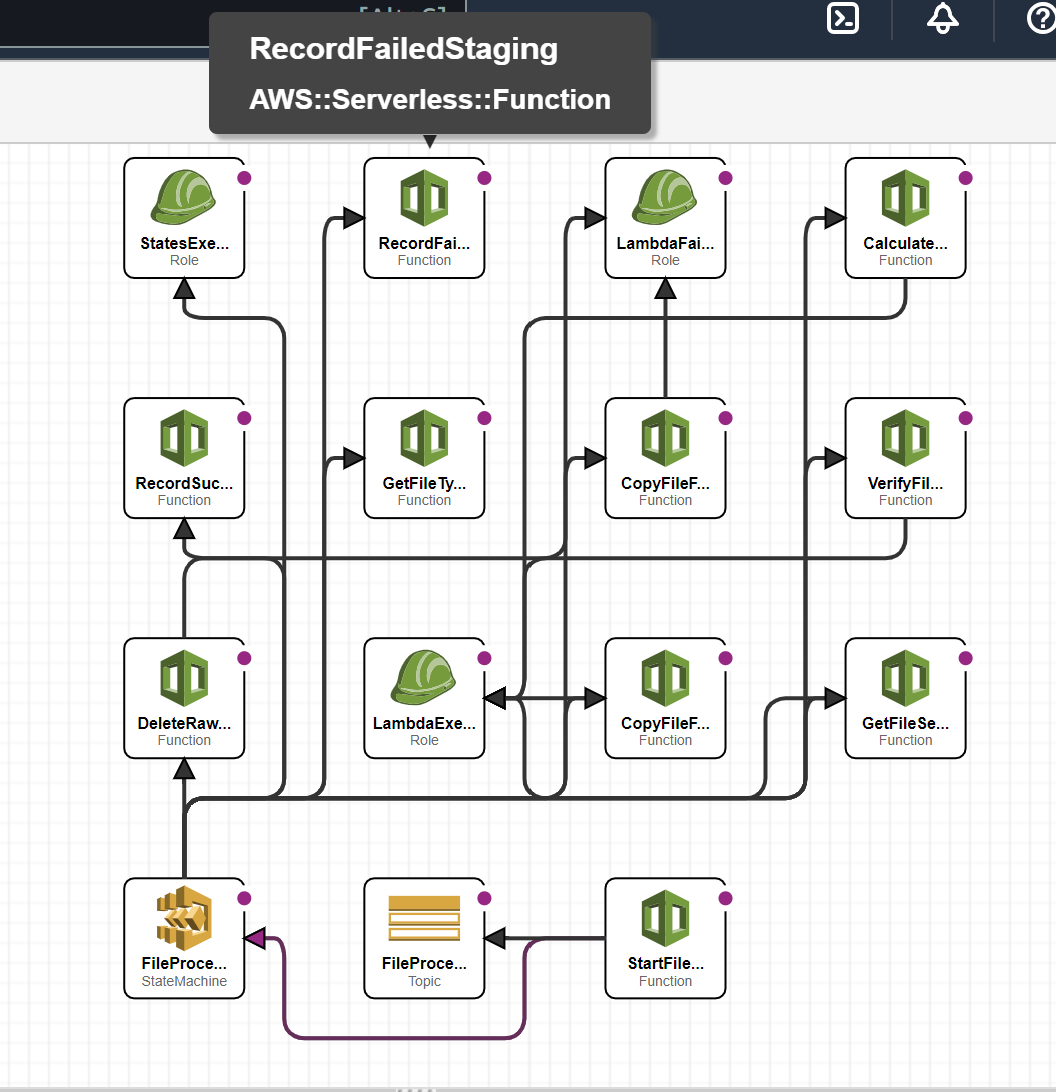


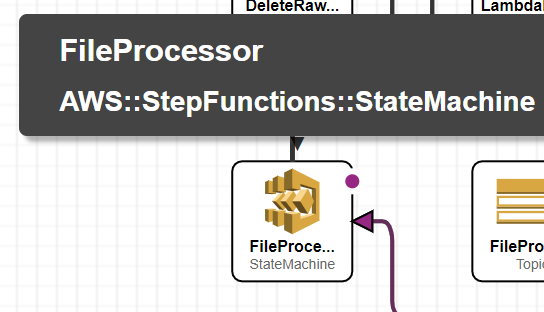


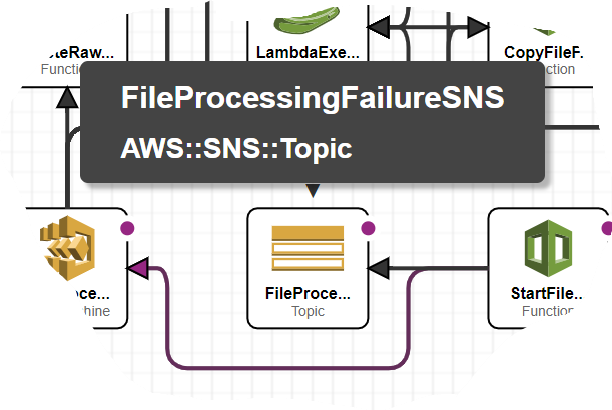


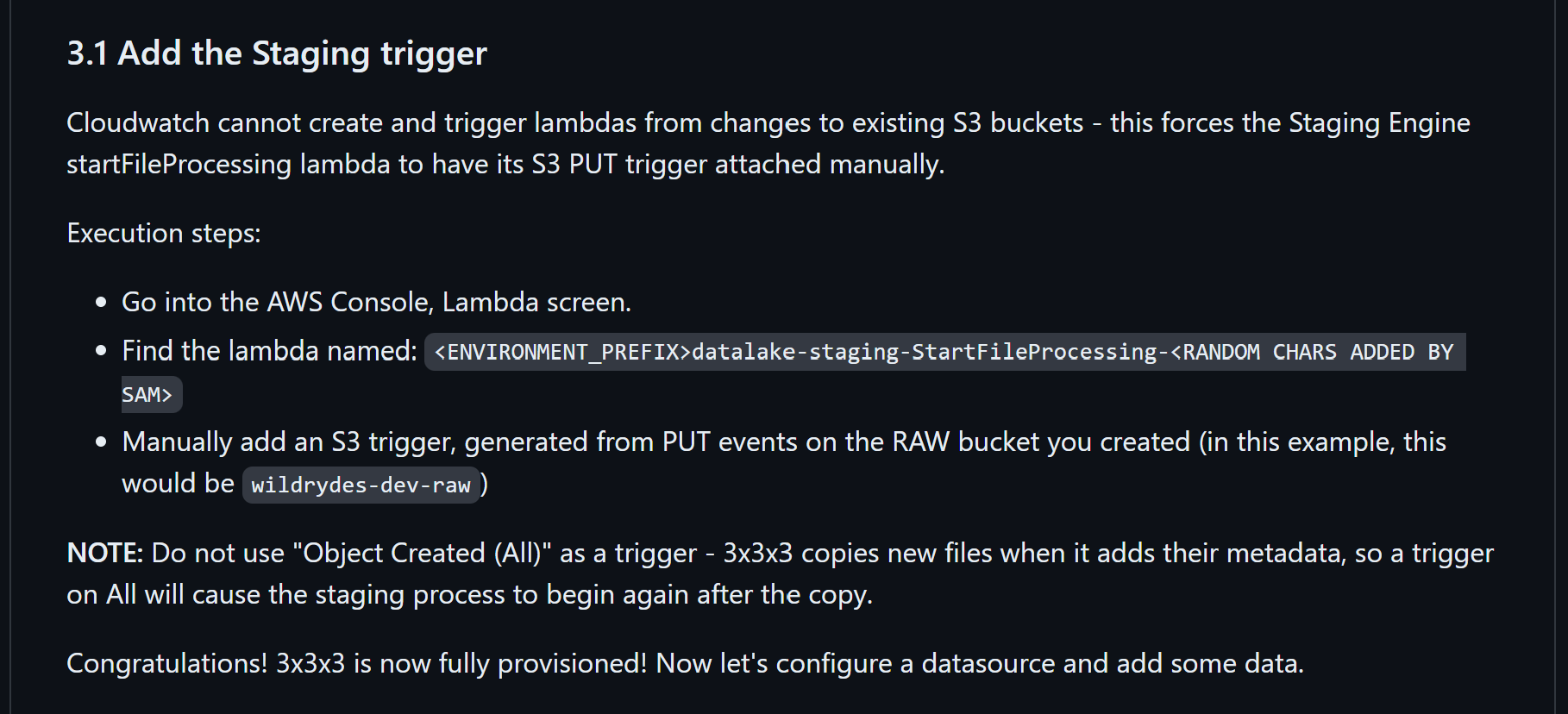


|  |  |  |
| --- | --- | --- |
| **wildrydes-dev-datalake- staging-engine (CloudFormation)** |  |  |
|  |  | Resources  * CalculateMetaDataForFile * CopyFileFromRawToFailed * CopyFileFromRawToStaging * DeleteRawFile * FileProcessingFailureSNS * FileProcessor * GetFileSettings * GetFileType * GetFileTypeRole * LambdaExecutionRole * LambdaFailedFileProcessorRole * RecordFailedStaging * RecordFailedStagingRole * RecordSuccessfulStaging * RecordSuccessfulStagingRole * StartFileProcessing * StartFileProcessingRole * StatesExecutionRole * VerifyFileSchema |

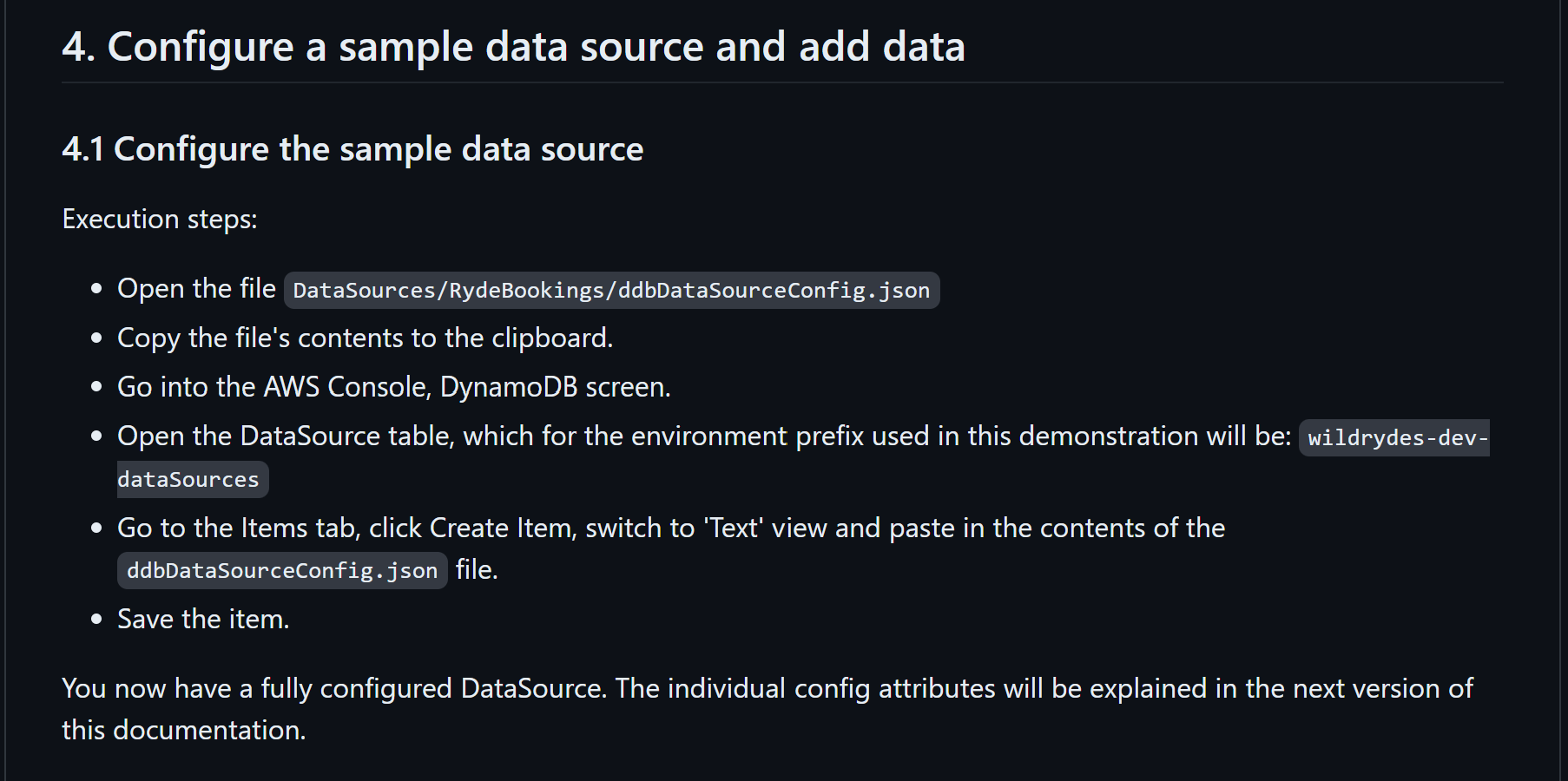


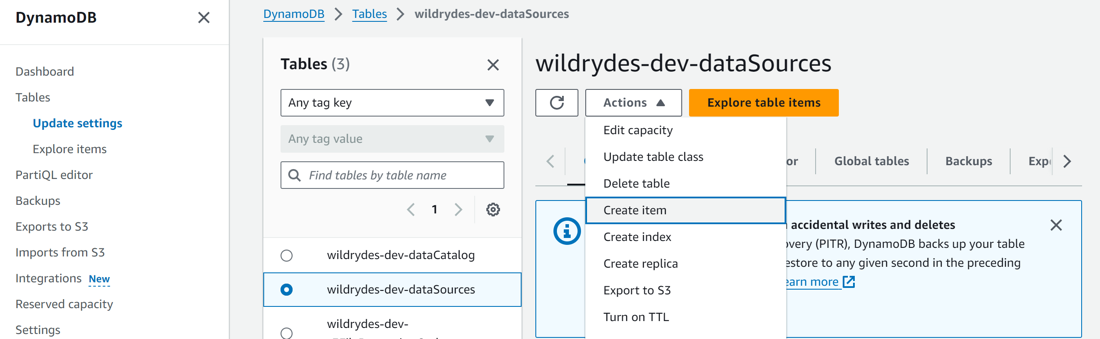


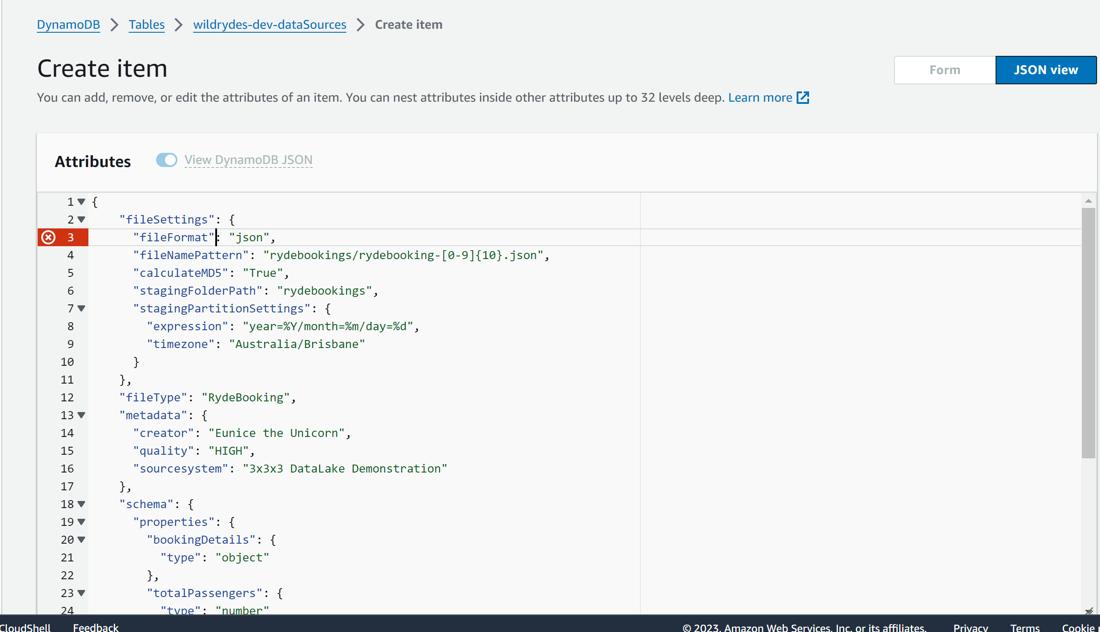




|  |  |  |
| --- | --- | --- |
| **wildrydes-dev-datalake-structure (CloudFormation)** |  |  |
|  |  | Paramaters  * DataCatalogTableName * DataSourceTableName |



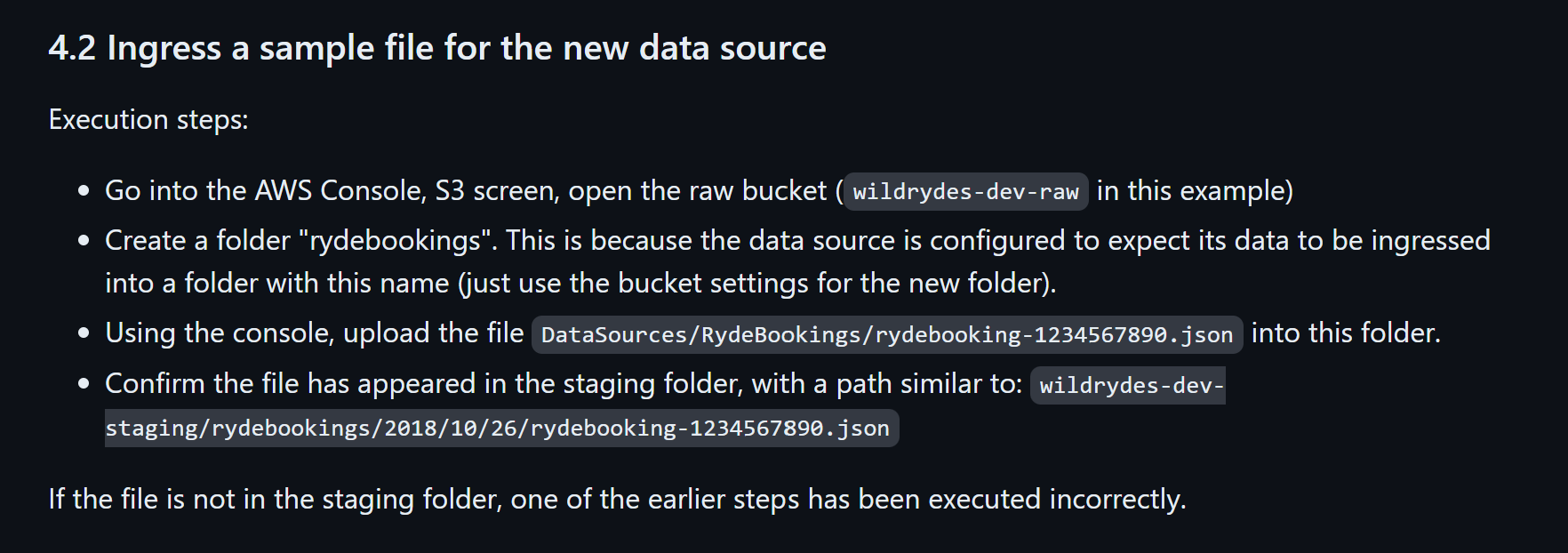


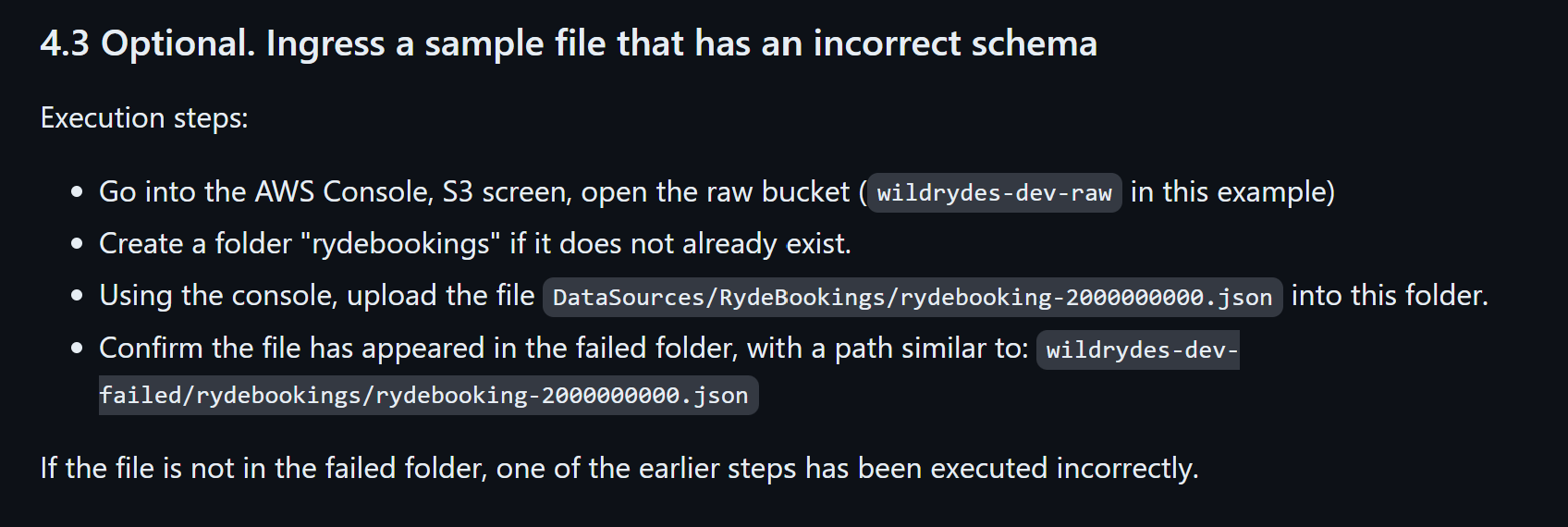
r

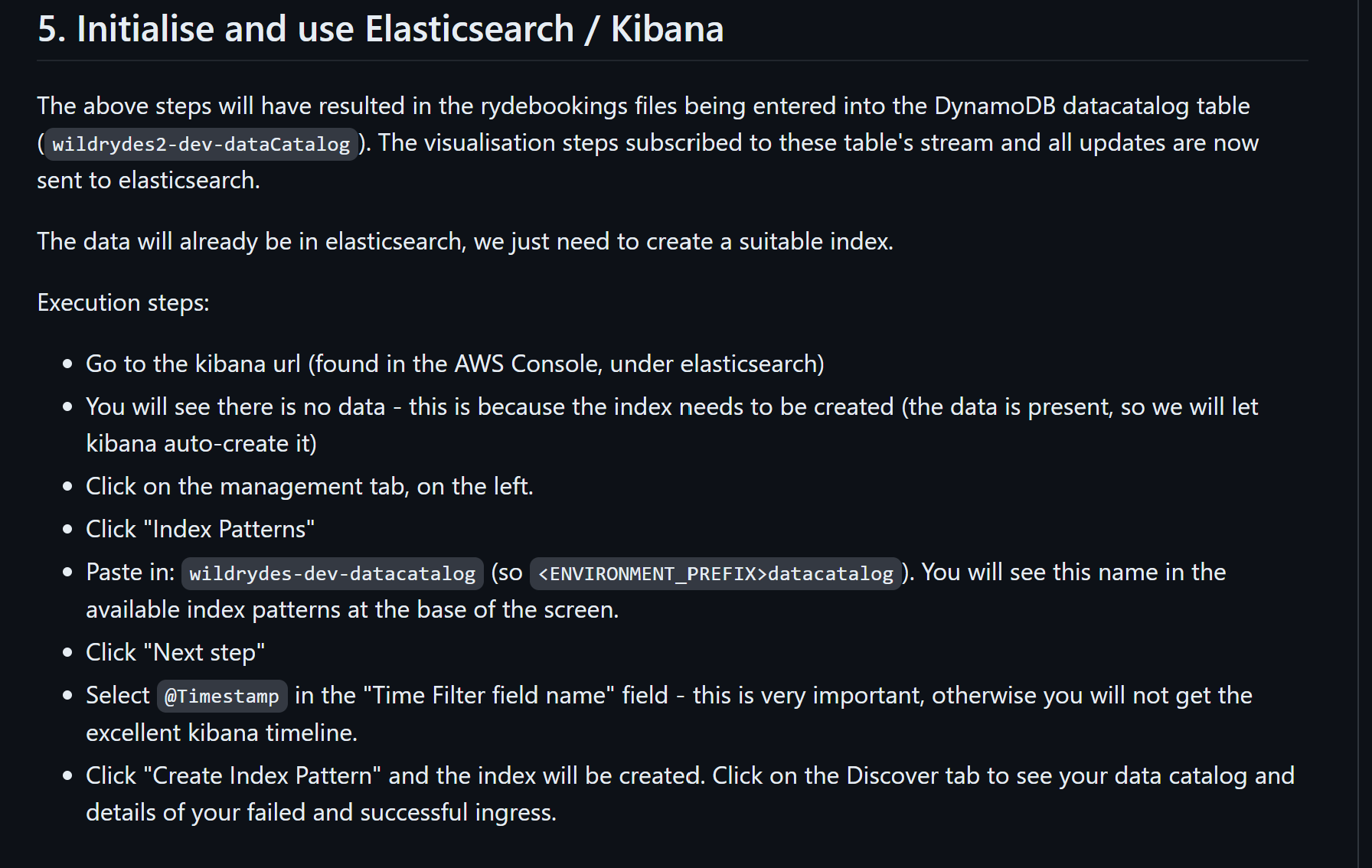


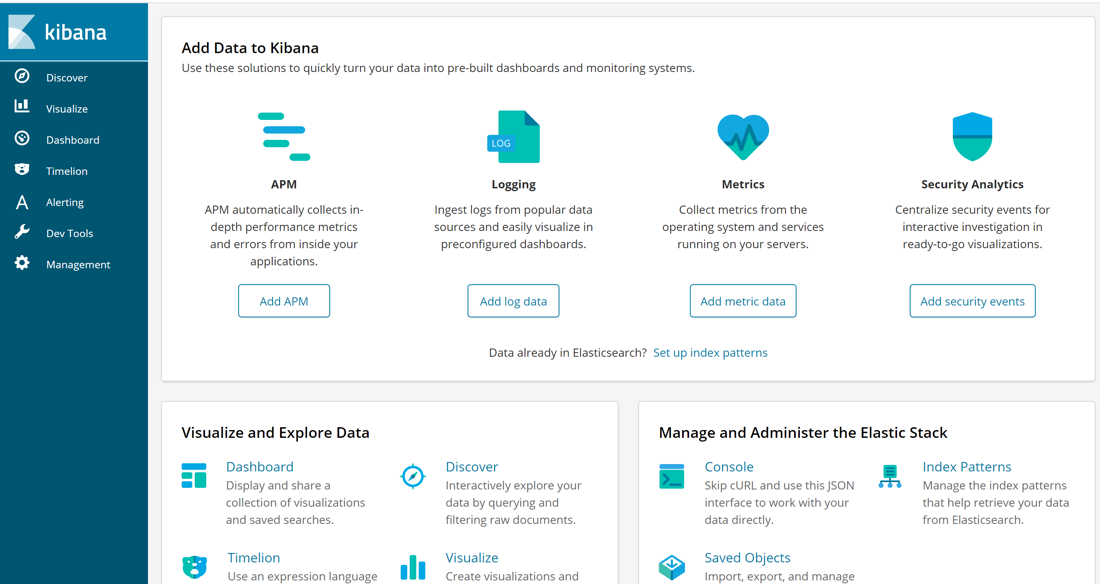
**GOT ERROR**

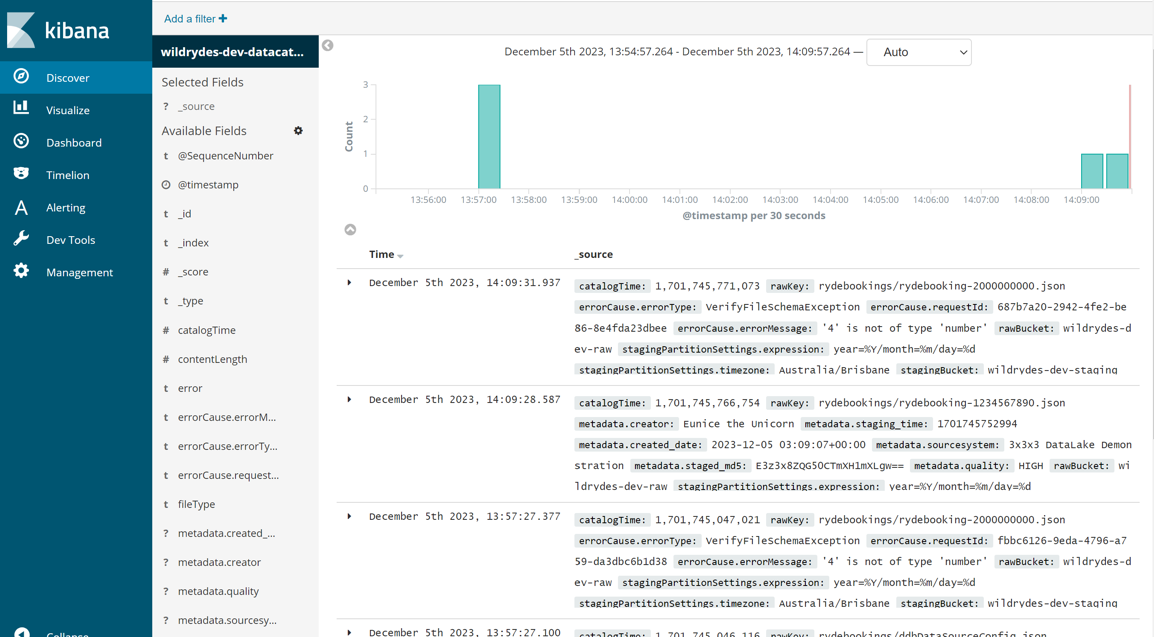
**XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX**





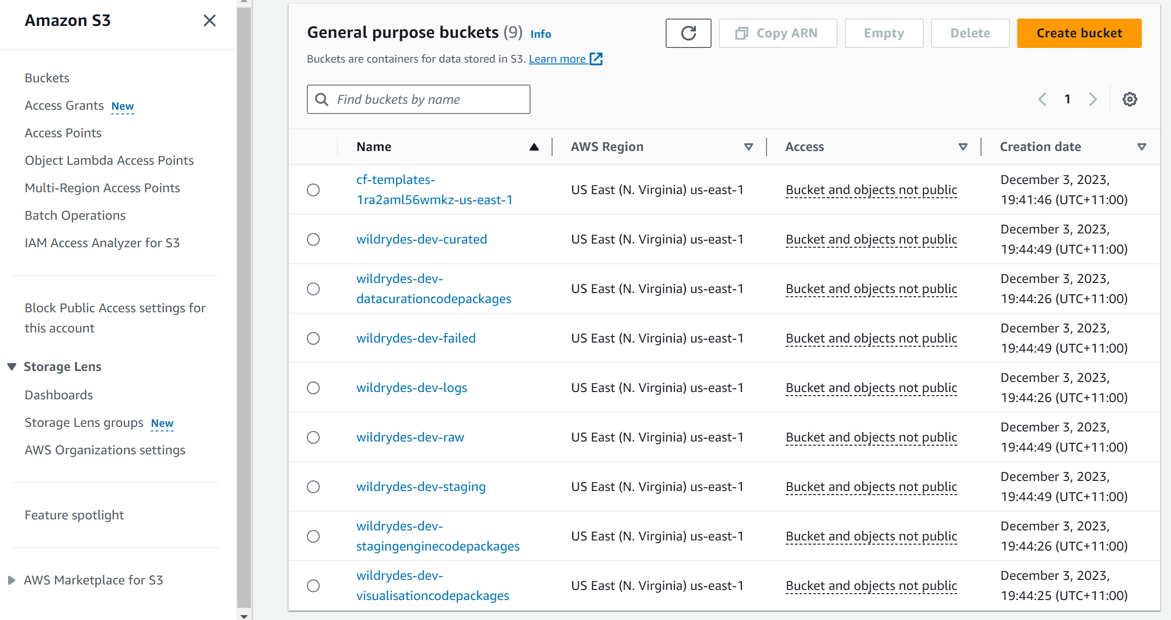




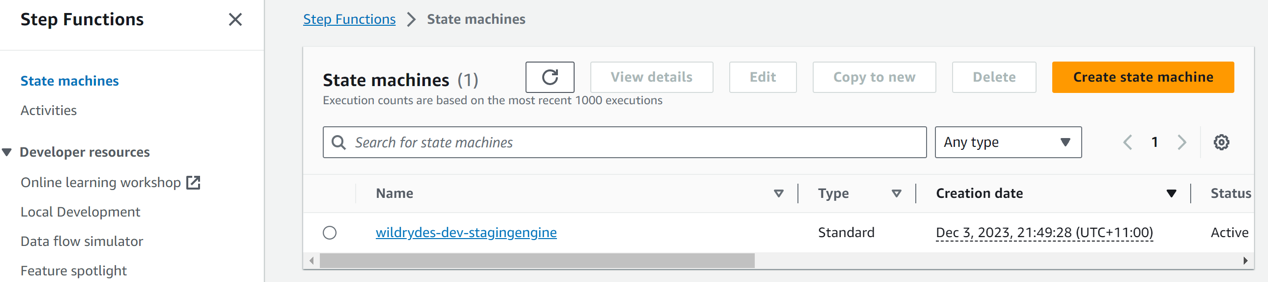


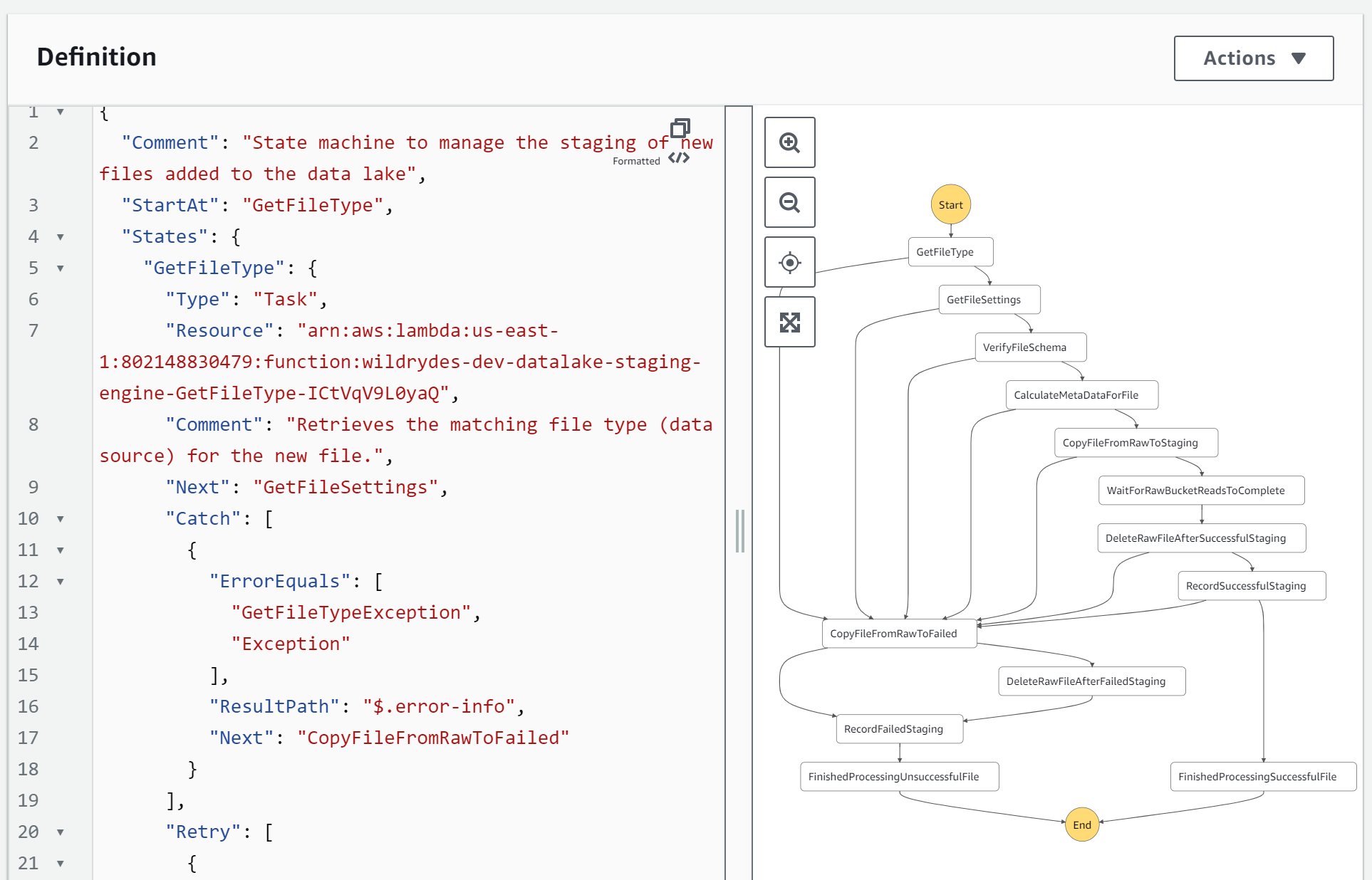
**XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX**

**S3 Bucket**



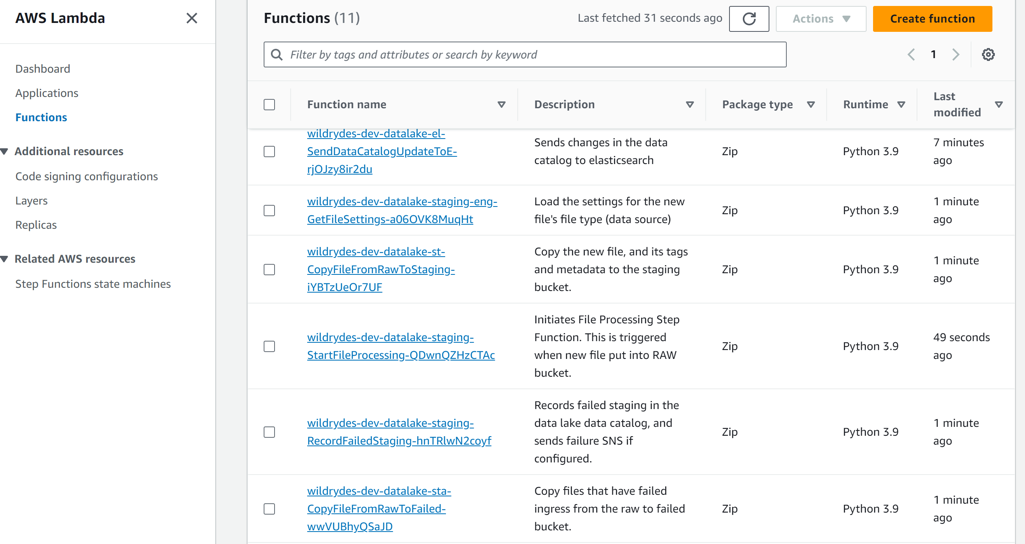
**STEP FUNCTION**





All these boxes are **States** of the file(data) that are obtained by passing through every single **Lambda Function**.

**LAMBDAS**





**GetFileType**

1. **GetFileSettings**
2. **VerifyFileSchema**
3. **CalculateMetaDataForFile**
4. **CopyFileFromRawToStaging**
5. **WaitForRawBucketReadsToComplete**
6. **DeleteRawFileAfterSuccessfulStatging**
7. **RecordSuccessfulStaging**
8. **CopyFileFromRawToFailed**
9. **DeleteRawFileAfterFailedStaging**
10. **RecordFailedStaging**