Introductions

Saturday, August 17, 2024 7:46 AM

RAD - Reporting Analytics and Data - support corporate services - BI dev shop

Client stakeholders:

Alex Schmidt - alex_schmidt@uhg.com
Patrick Whitmer - patrick_whitmer@uhc.com, Cell: (563) 349-3425
Grant D Haugan - grant.haugan@uhg.com

++

RAD team distribution list: TeamRAD@ds.uhc.com

++

Project: 27105 - RAD Data Modernization Solution Definition & De >>>
Customer: UHG Corporate

+++

AIDE IDs: https://aide.optum.com/

ASK Id	AIDE Link	Name	Category	Lifecycle Stage	Decommission Status	Service Level Owner	Technical Owner	Business Owner	GL Mapping?
UHGWM110-028139	AIDE Link	UHG Corporate Services (Orbit Alteryx)	Application	Production		WHITMER,PATRICK	HAUGAN,GRANT D	WHITMER,PATRICK	Unassigned

++

Khurana, Preeti 8:06 PM

Sharing on Chat for your Ref, will add the file on teams channel also

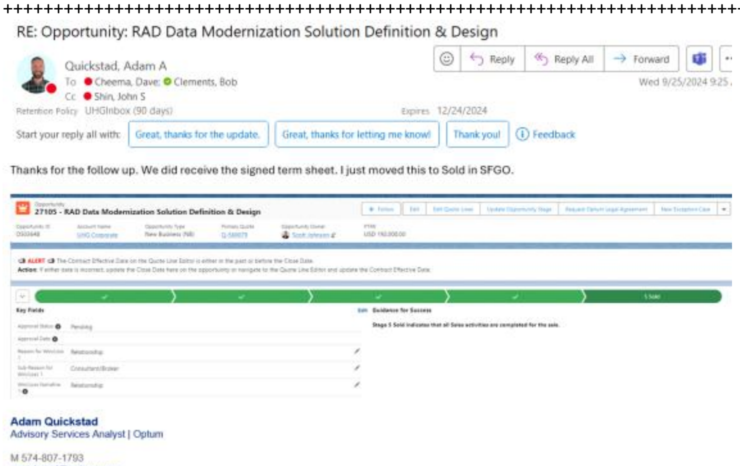
Dave Cheema	US	714-925-8990
Dijendra Kumari	US	+1 317 777 2526
Preeti Khurana	Gurgaon	91 9899962134
Muquaddar Baig	Nagpur	91 9391556350
Mahesh Kumar	Bangalore	91 7892685215
Sujit Kumar	Kolkata	91 9804158093
Asit K Sahoo	Hyderabad	91 8553326015

++

Opportunity ID: O503648
Opportunity: RAD Data Modernization Solution Definition & Design
Account Name: UHG Corporate
Opportunity Type: New Business (NB)
Primary Quote: Q-560079
Opportunity Owner: Scott Johnson
Project Owners: John S Shin and Dave Cheema
PTRR: USD 192,000.00

Where are we in the term sheet process and when do we expect it to be ready

++



++

Accounting site: AOA - Request Site - Home

ECIF

Thursday, May 15, 2025 9:38 PM

Supplier Web: https://supplier.microsoft.com/App

POE (Proof of Execution): Signature requested on "POE_Supplier_3067563_OptumInsight Inc_CAS-1737011-Q2V5H9_-993209687"

One Ask requests your signature on

POE_Supplier_3067563_OptumInsight Inc_CAS-1737011-Q2V5H9_-993209687

Review and sign

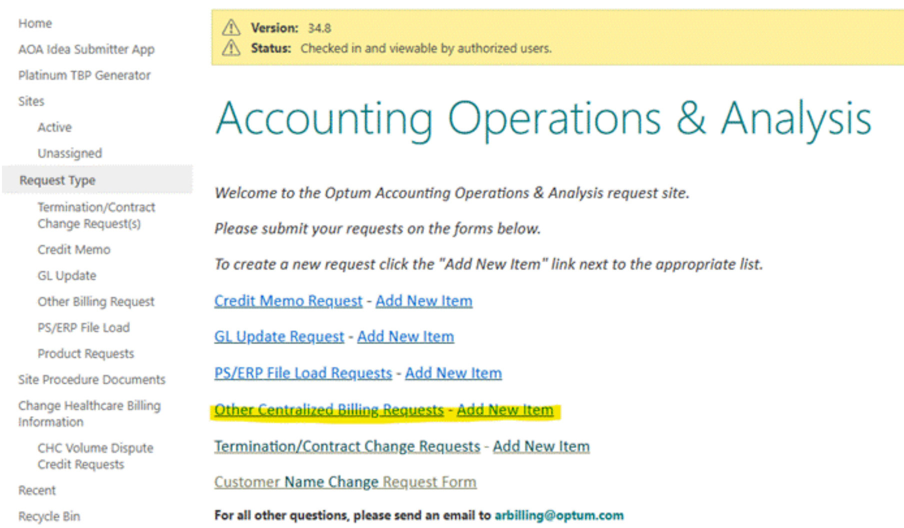
If you are an ECIF supplier please confirm completion of milestone(s) using the drop-down selection in the milestone table, enter the **start and completion dates (MMM DD, YYYY format)**, click on the box to attach additional Proof of Execution (POE) documentation.
oneask@microsoft.com

Request for off cycle billing request: AOA - Request Site - Home

+++ +++++++++++++++
From: Shelburne, Douglas A
Sent: Thursday, May 15, 2025 9:32 AM
To: Malhotra, Dinesh <dinesh.malhotra@optum.com>
Cc: Cheema, Dave <dave.cheema@optum.com>; Shin, John S <john.shin@optum.com>; Welter, Brett <brett.welter@optum.com>
Subject: RE: SF Opp ID O714981 - RAD Data Modernization Milestone 3 Invoice for submission

Please put in an off cycle billing request in our share point site.

AOA - Request Site - Home



From: Malhotra, Dinesh <dinesh.malhotra@optum.com>
Sent: Wednesday, May 14, 2025 10:17 PM
To: Shelburne, Douglas A <douglas.shelburne@optum.com>
Cc: Cheema, Dave <dave.cheema@optum.com>; Shin, John S <john.shin@optum.com>; Welter, Brett <brett.welter@optum.com>; Malhotra, Dinesh <dinesh.malhotra@optum.com>
Subject: SF Opp ID O714981 - RAD Data Modernization Milestone 3 Invoice for submission

Hi Doug – For SF Opportunity O714981 (Microsoft PO # 101279849), the project team has delivered the 3rd milestone and has secured the Client Certification as well.

- Kindly generate the invoice for $94,301 and send it over to Dave Cheema for submission to Microsoft

Thank you!

Best,
Dinesh Malhotra
M: +1 773-398-7713

From: Cheema, Dave

+++ +++++++++++++++
From: One Ask via Adobe Acrobat Sign <**adobesign@adobesign.com**>
Sent: Thursday, May 15, 2025 1:47 PM
To: Malhotra, Dinesh <**dinesh.malhotra@optum.com**>
Subject: Signature requested on "POE_Supplier_3067563_OptumInsight Inc_CAS-1737011-Q2V5H9_-993209687"

One Ask requests your signature on
POE_Supplier_3067563_OptumInsight Inc_CAS-1737011-Q2V5H9_-993209687

Due by June 14, 2025

Review and sign

If you are an ECIF supplier please confirm completion of milestone(s) using the drop-down selection in the milestone table, enter the start and completion dates (MMM DD, YYYY format), click on the box to attach additional Proof of Execution (POE) documentation. Recommendations for this additional documentation are outlined in this template. If you worked with a customer on this ECIF project, the traditional customer POE is NOT considered acceptable documentation.

Once this POE is completed, the form will be routed back to the ECIF Operations team for review.Your immediate action is greatly appreciated.

** Reminder: If you are not the supplier contact no action is required by you on this request.**

One Ask
oneask@microsoft.com

After you sign **POE_Supplier_3067563_OptumInsight Inc_CAS-1737011-Q2V5H9_-993209687**, all parties will receive a final PDF copy.
As you prepare to sign this agreement, please note that the document asks for the following document to be attached:

- additional document

Don't forward this email: If you don't want to sign, you can **delegate** to someone else.

Subset Implementation Plan

Wednesday, January 15, 2025 7:31 PM

Item #	Activity	Responsible Party	Comments
1.	Configure the Azure Cloud env. for app. needs	Asit K. Sahoo	Complete remainder tasks on Azure to process remaining API Data sources
2.	Configure Snowflake for application needs	Asit K. Sahoo	Complete remainder tasks on Snowflake to process remaining API Data sources
3.	Implement batch data loading framework for WebAPI data sources	Muquaddar & Dijendra	Review and revise batch data loading framework for remaining WebAPI data sources
4.	Framework Parameters	Sujit P.	Identify, define and implement the ingestion framework parameters and stored procedures to ingest, ETL and migrate remainder API data sources
5.	Data Pipelines	Mahesh	Setup ingestion data pipelines to ingest remaining API based data sources
6.	Refine Functional Specifications (where needed)	Muquaddar & Dijendra	Review and revise functional specifications, only on as needed basis
7.	Refine Detail design (where needed)	Muquaddar & Dijendra	Review and revise detail design specifications, only on as needed basis
8.	Ensure Database, schemas and tables definitions availability (*create Silver layer table definitions for the already created Bronze layer tables, before start creating new table definitions for the remaining Bronze tables*)	Patrick & RAD Team	Work with Patrick and team to get database, schema and table definitions to create Bronze and Silver layer tables. Request them to create database, schema, and table definitions for an API data source, then the next data source, and so on. Don't create all Bronze layer tables first and then all Silver layer tables. Work closely with Patrick & team to plan your work horizontally, not vertically. For example: identify a data source, create its Bronze and Silver tables, setup a process to ingest it, Read it from Blob Storage, ETL to Bronze layer, ETL to Silver layer, test it and then move on to the next API data source.
9.	Create schemas(where needed)	OA team	Follow the recommendation approach detailed in item #8 above
10.	Create tables (Bronze & Silver)	OA team	Follow the recommendation approach detailed in item #8 above
11.	Read API data source into Azure Blob Storage	OA team	Follow the recommendation approach detailed in item #8 above
12.	Read Blob Storage account and populate into Bronze table	OA team	Follow the recommendation approach detailed in item #8 above
13.	Read Bronze table and ETL into Silver table	OA team	Follow the recommendation approach detailed in item #8 above
14.	API data source to Silver layer testing	OA team	Follow the recommendation in item #8 above
15.	Migrate this process to production (time permitting)	RAD Team & OA team	
16.	Conduct PAT (time permitting)	RAD Team & OA team	

Notes

Saturday, August 17, 2024 7:47 AM

USE SECONDARY ROLES ALL;

AZU_SDRP_AKWUR_DEV_DEVELOPER_ROLE

Default Date format to be used across all date fields: yyyy-mm-dd

++ ++++++++++

Gold layer views

EP_GOLD_DEV.SERVICENOW_MERCURY_REQUEST_DETAIL_V
EP_GOLD_DEV.SERVICENOW_MERCURY_REQUEST_DETAIL_OPTUM_V
EP_GOLD_DEV.SERVICENOW_MERCURY_REQUEST_DETAIL_UHC_V

++

RAD Machine Learning users:

- Sam Bailey
- Mark Pedigo
- Jason Gehlen

+++

Snowflake Service account id: rad_snwflk_ts@optum.com

+++

Hi Dave. Please update me on status of staffing for RAD

and what do you think it the expected start date for the project we can communicate to client

Yes, John. We have lined up interviews with Sairam (Data Engineer) for today and with Prem for tomorrow.

Sairam is a contractor?

Yes, John.

from LTIM or Persistent?

LTIM.

How did convo with Prem go?
less than thrilling. I'll send you my comments after this meeting.
I see

you realize he is an FTE that was hired in the role we are seeking to fill, right?

John, I'll provide you my feedback and then I'll seek your guidance as to which way to proceed. Thank you.

John, the following is my feedback on Prem:

His resume looks very impressive and I was excited to speak with him. However, when I started asking him questions, his answers were less than thrilling. The following are a few examples:

1. He mentioned Snowflake at least 10 times, but it is not anywhere on his resume
2. I asked him how do you create a stored procedure in Snowflake? His response was I only created views in Snowflake
3. I asked him if he had done any performance tuning, he said he has not done it
4. I asked which database types he has experience with, It was only MS. SQL Server
5. He could not articulate end-to-end process for a data migration project
6. He does not have any experience with SHIR
7. I asked him what is his sweet spot, his response was BSA (Business Solution Analyst) and ADF
8. He does not have any experience with Azure Functions and light experience with Azure Logic Apps
9. He does have good experience with ADF, he was able to articulate the ADF pipeline
10. I had also spoken to Rajesh Konetti, he provided lukewarm endorsement of him
11. Gaurav said that he would work hard

My recommendation would be to pass on him. But I'm very interested in your feedback. Thank you.

It's an FTE that has been hired into the role we are looking for. The expectation would be that we use FTE resources and help them grow into the role. The decision to do that is up to you. However, part of expectations on leaders like you are to help grow our resources.

Dave - We need to communicate a start date for this project to the client. We need to at least start this project by early October. It takes 3 weeks+ to onboard a contractor once identified. Don't expect to have the time to weed through dozens of candidates to get to that "perfect team" to work with. that never happens.

You have to be the difference maker to bring the team together and ensure the goals are achieved

many times, that team won't be the A or B team

+++

09/15/2024 5:45 PM CST

All,

We're starting a new project that entails the on-premise data migration to the Azure/Snowflake cloud, which is very similar to the MRIS Modernization project. Since this project is so similar to the MRIS Modernization project, I would like to take advantage of the hard work you've done, experience you have gained and lessons you have learned. To leverage all that, I would like to request the following: architecture diagrams, code samples, documentation, or any other pertinent supporting materials for the following data sources:

- API data source
- Outlook Email data source
- SQL/Oracle data source
- Internal files data source
- External files

- ADF Workflow Management Framework to Control Execution of Future Data Pipelines

Note: These documents don't have be fully polished and published. We'll accept whatever form you may have. Also, if you don't have something documented, but you can articulate it, we'll also accept that as well.

I know you guys are very busy, but your timely attention to this request will be greatly appreciated. Thanks in advance.
++
Tenant administrators create a separate database https://docs.hcp.uhg.com/strategic-data-repository-platform-(sdrp)/storage-tenant
++
To create non-prod subscription: 1 - Set up a Non-Production Cloud Account - Public Cloud | HCP Docs (uhg.com)
++
some useful links: https://docs.hcp.uhg.com/public-cloud/public-cloud-subscription-strategy https://docs.hcp.uhg.com/public-cloud/azure-subscription-model
+++
Not much data transformation
Majority - gold layer request
30 Tables are important

Mercury tables
SQL table are aggregated into larger table
Download data
Requirement scope is changed
+++
Global Groups: UHGRG_rg_rad_azurestack_08a9b07_ADMIN
Resource Group SDRP groups:
UHGRG_rg_rad_sdrp_snowflake_a1fbf6e_CONTRIBUTOR
Snowflake Roles: Developer - AZU_SDRP_AKWUR_PRD_DEVELOPER_ROLE & AZU_SDRP_AKWUR_DEV_DEVELOPER_ROLE
+++

RE: Questions about Data Catalog

Koy, Kate
To Barker, Scott D; Cheema, Dave; Turk, Brian J (he/him); Motsi, Fennie Nomathamsanqa; Nair, Rajesh G; Shin, John S
Cc Arora, Anjali; Nguyen, Vuong Q; Nowak, Bobbie-Jean; Bhalala, Payal K; Nagpal, Ankita
Reply Reply All Forward
Mon 10/7/2024 5:49 PM
Retention Policy UHGInbox (90 days) Expires 1/5/2025

Hi – Here are some preliminary answers to your questions below (red).

Kate

-----Original Appointment-----
From: Barker, Scott D <scott_barker_@uhc.com>
Sent: Friday, October 4, 2024 1:46 PM
To: Barker, Scott D; Cheema, Dave; Turk, Brian J (he/him); Motsi, Fennie Nomathamsanqa; Koy, Kate; Nair, Rajesh G; Shin, John S
Subject: Questions about Data Catalog
When: Tuesday, October 8, 2024 12:00 PM-12:30 PM (UTC-05:00) Eastern Time (US & Canada).
Where: Microsoft Teams Meeting

Hi all,

As we migrate to HCP, we have some questions about Data Catalog and how it fits into our metadata strategy. We know we must use Data Catalog to support the Data Product registration and subscription processes for HCP. We also have 3 varieties of business and technical metadata in Excel spreadsheets, and want explore what our options might be for having that metadata in HCP, we also have legal restrictions from AARP.

These are the questions we want to answer. These answers will inform our strategy for using Data Catalog.

1. What types of metadata can Data Catalog store? We have 3 different varieties of metadata listed below:

RE: Questions about Data Catalog

Koy, Kate
To Barker, Scott D; Cheema, Dave; Turk, Brian J (he/him); Motsi, Fennie Nomathamsanqa; Nair, Rajesh G; Shin, John S
Cc Arora, Anjali; Nguyen, Vuong Q; Nowak, Bobbie-Jean; Bhalala, Payal K; Nagpal, Ankita
Reply Reply All Forward
Mon 10/7/2024 5:49 PM
Retention Policy UHGInbox (90 days) Expires 1/5/2025

b. Specific DBMS or all DBMS's – is vision for Data Catalog to include only Snowflake and Azure, or could it include metadata for other DBMS's like Oracle and DB2?
The grand vision is to depict the entire UHG ecosystem of data – across all types, technologies, etc. It is the preferred method of enterprise data cataloging. Sources available on the Data Catalog is dependent on the source technical teams emitting the metadata.

3. Is there a plan to put in protection or access control in the future? Ideally, MRIS would like to be able to limit access to MRIS resources and by request
Access control related to viewing **metadata** is not strategic, as it's meant to support a view into the entire landscape of data. Access control to the **real** data is supported through the Catalog subscription process for Snowflake, and will continue to include other data stores.

4. What would be the process to import existing Excel based metadata? And then maintain it.
To load technical metadata (does not exist in a source where it can be programmatically accessed): Loading Technical Metadata via CSV - Data Catalog | HCP Docs (uhg.com)
To load metadata enrichments, once technical metadata already exists in catalog: Dataset / Field Enrichments File upload using Data Catalog UI - Data Catalog | HCP Docs (uhg.com)
If your data is on Snowflake, the technical metadata ingestion is already taken care of – and it may be worth it to load business metadata directly into snowflake where possible, to make use of the automated ingestion (Snowflake → data catalog). For anything additional, the Data Catalog has functions to upload via UI.

Thanks,
Scott

Microsoft Teams Need help?

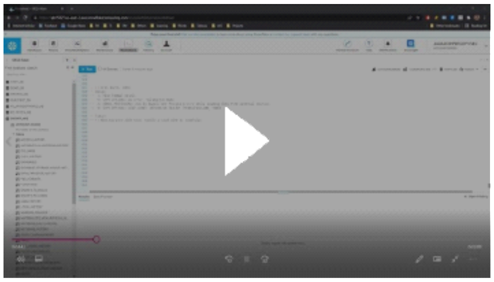
Join the meeting now

Meeting ID: 246 735 700 788

Passcode: 4djB7Y

++

Snowflake parsing JSON data: Snowflake : Parsing JSON Data



Snowflake Flatten JSON :[Snowflake Tutorials: Flatten JSON in Snowflake]

Snowflake parse json data:
Parse multi-level json using Snowflake SQL
++

RAD Data Modernization GitHub Repos:

- uhg-internal/rad-platform (github.com)
- uhg-internal/rad-adf-analytics (github.com)
- uhg-internal/rad-sf (github.com)

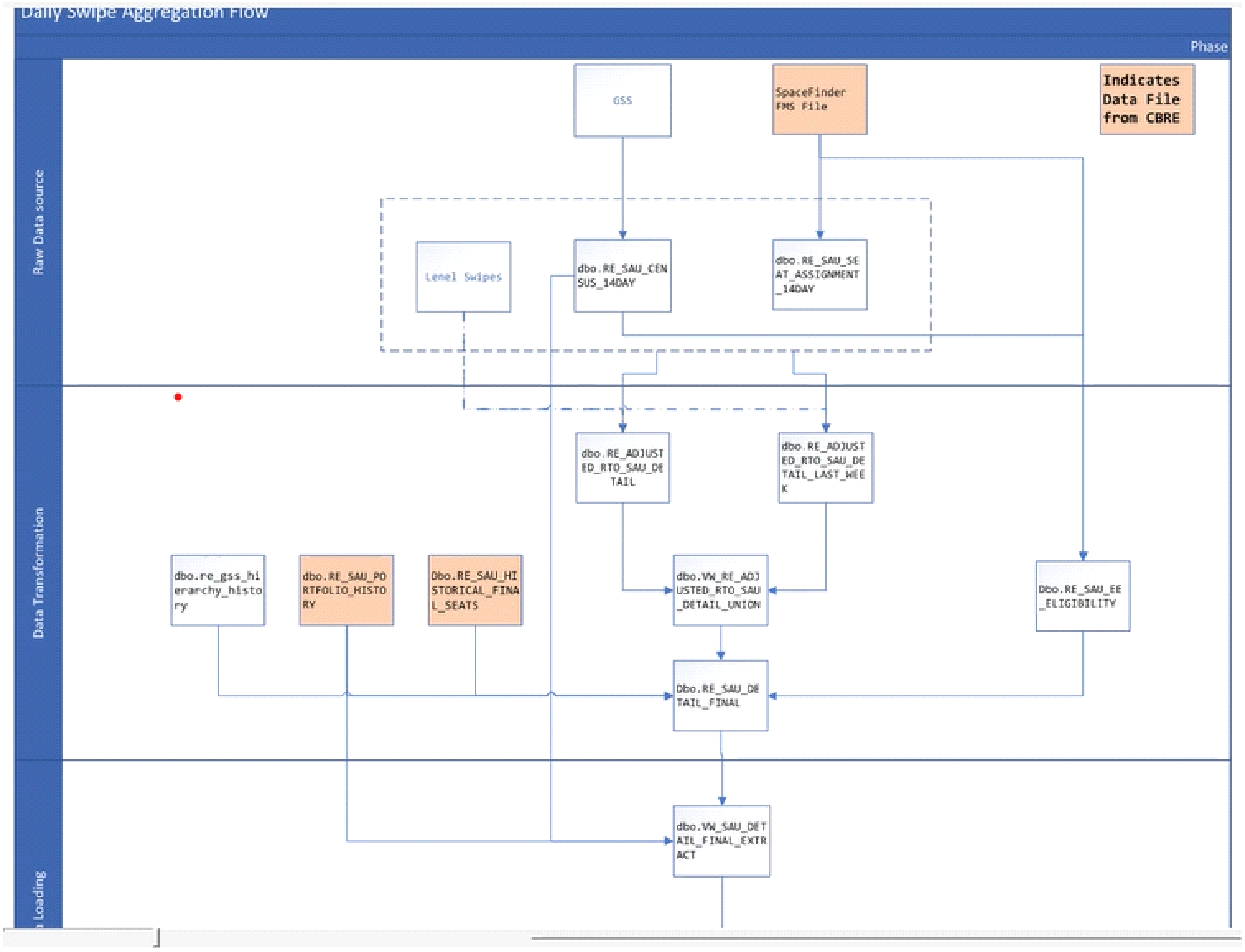
++

The real estate utilization data set for SQL and flat file inputs.
We get a flat file from a vendor and we store that in the SQL table
The genesis of the data are actual SQL tables
The real estate data set is different, completely different on the other data sets
Easy and straight forward to connect, but there's lot to it

We use the utilization real estate data to make enterprise level decisions on, e.g, what buildings should we invest and renovate; what buildings are not be used and should be closed
We have this concept of utilization and eligibility.

To determine if someone is eligible at a building, we look at a couple things:
1. are they in our space Finder file (master seating chart)? It contains information like, Grant Haugen sits in Minnesota 8, his actual space number on this floor and other metdata, such as HRML code; 2. The other component is the actual swipes - where are you swiping your data at. For that, we have a connection into Lenel. Lanel is the enterprise preferred solution for physical access control system. If there's a building that's on our IT network, it's part of the Lenel access control system.
We have a partnership with the corporate security. That's where we have access to their SQL database. When you swipe your card on a card reader, a record gets added into their SQL database.
I get card swipe data for the last 6 days, even though the table has hundreds of millions of records. it It is the SAU (site access utilization) data set. It is a well done schema.
Every day we will go and see what's the most current date or like the last date. I have in a couple tables and determine what Lenel gave us everything from this swipe date onward, right, is how we do.
It.
There's a little bit of nuance to it how the security technology works, because it's more hardware based than software.
When you swipe your card against an access reader, behind the scenes, an access reader on a door is connected to a panel - a very small server box
Those rooms on floors that have like little servers in them.
A panel is gonna have many card readers to it.

When you swipe your card at the reader card slot goes to the panel panel, then sends that back to the database in the event a panel goes offline, called offline mode. You will still get in all that data still stored on the panel. Corporate security has a 72 hour SLA to get panels back online and when they come back online they dump all their data in the backend SQL server. So we have a process that stores the last six days worth of swipes.
We re-aggregate every time because if a panel's offline today and maybe tomorrow when it comes back online on the 16th, we've already pulled the day on the 14th on the 15th.



++
SQL
Real Estate dataset
Straight forward
Utilization & Eligibility

Lenel is physical access system
Physical Access Security
Corporate security has 72 hours SLA

12-15 tables
200 million rows
Flat file?
Need to migrate historical data

100,000 swipes for 6 days
Seat assignment 100,000
There is 1+ month lag
Get detailed data
truncate and reload last 14 days data
No overlap of data

How many datafeeds:
HR People platform dataset
Notification: get email when the data load is available

Engage Brad and Jason
Credentials and server address for SQL server

If you're looking at the left side, there's what I consider the raw data source, the data transformation, and then the data loading.
We have one SQL view and you can see the name here.
Sau detail final extract that's connected on top of Tableau.
Tableau is a bunch of tableau reporting on this.

Brown tables on the left: Portfolio History, Final seats contain metadata on sites.
Bottom center has Historical data table
We aggregate week days card swipe data
Metadata changes month-to-month

This is called portfolio history; Final seats. Treat these as SQL tables that house metadata on sites.

But they're all historical. Metadata does change on a month to month basis.for example, at the division level, at a site, how many sites or seats, how many offices

What type of like agile space are assigned to a building - utilization?
There's also a cost component of this on the real estate side. The final seat state is also used on our real estate finance side so that they know that Division One because you have X amount of seats which means.

They want to see a current value; we wanna see historically what was this division seats capabilities capacity in September.

Going up through here, right, there's this GSS historical hierarchy (diagram)
A monthly process that goes to GSS slash people platform like what's the current division?
What's the unique combinations of division, market, group segment group, etc.

Eligibility is based off where you sit and what buildings you're eligible for.
These two tables within here in this view, this is called the adjusted RTOSAO detail.

If we were to remove this then this data set would really just look like employee ID building. This this detail final is where everything is.

You can see from this chart it's aggregated into it in one form or another

What the eligibility likes for a site is. We have all those rules.
it's used at the highest of levels across the enterprise for important decisions.

We connect to this application which is Lenel to pull in swipe data.
We store that in some staging tables. We get flat files from CBRE or vendor. We store in staging tables and we combine everything into one.
Report off that.

Migrate hitorical data: Yes, we would.
Yep, unlike Mercury or Mercury gets re aggregated every day, this does not. Would like to know how we should best migrate this into a snowflake.

There are about 12 SQL tables in scope

This is the application database, we just need to connect to it
This is a view, so we're worried about that

The eligibility is going to be from an input table standpoint. The eligibility one will be the largest data set.
It'll be close to about 200 million rows. The final seats in the portfolio history - maybe 100,000.
This just keeps six days worth of swipes, maybe 100,000 records at any given time.
Adjusted detail - This is the historical. This is probably like 40 million rows
These two, the 14 days, we only keep 14 days worth of census data in here, 14 days a seat assignment in here.
Census is probably about 3 million. I want to say and see the sign is like man, maybe 100,000 thought that.

So we capture key swipes data and store it. When we get data prior to that data, everything from that point onward get cleared out and populated with the dataload. These data loads get aggregated at the end of the month on the CB JLL side.

Brad, our real estate RAD partner, can speak better to the lag in this process. it's likely they do a bunch of data checks, data aggregations , etc. before they release these. The reason we push the previous data forward is that there's data points that need to be in the data set for the reporting team to actually have their calculations work. We could only get data when they release to us.

It requires dynamic parameters

The 14 days - the table is getting truncated everyday reloaded on 14 days worth of data.

There is no overlapping concern - delete everything and then re-aggregated everything

What is the frequency of these jobs?
The whole process refer to aggregating the detail final table.
It runs daily.
There are some timestamps on the side here.
Final daily aggregator runs at 4:30 AM CST.
The census 14 day gets calculated. This needs to be changed. This is actually at 8:30 AM.
They run daily
Monthly these three, everything else appears daily.

How do these files delivered?
The space Finder comes through ECG automatically
CBRE is emailing us.
we save it to our NAS and then we have a workflow that writes that file to a SQL table.
It is a manual process
How do they notify that the data load is ready?
They email
++

Today

11:57 AM

Patrick, Grant, got a question for you: since the Outlook Email data source (design doc and development) is no longer needed. Can you please help me understand the reason? And what is the alternative? Thank you.

Last read

Haugan, Grant D 12:14 PM

We only have 1 or 2 sources that send flat files via email and those are working to be remediated. The files will be automatically captured and moved to the NAS via ECG or other mechanism. When we first started the OA process at the beginning of the year, there was a wider range of flat files coming via email, but those processes have been sunset or migrated to something more manageable where the files get automatically to our NAS.

++

RE: RAD Data Migration project Update

WP Whitmer, Patrick

To Cheema, Dave; Haugan, Grant D

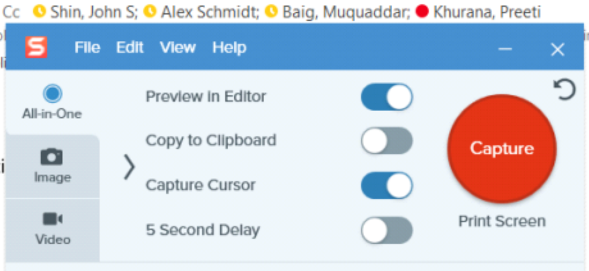
Cc Shin, John S; Alex Schmidt; Baig, Muquaddar; Khurana, Preeti

Reply | Reply All | Forward

Thu 10/17/2024 8:13 AM

Retention Po ... res 1/15/2025

You repli



Hi Dave,

Confirmati

Thanks,
Patrick

From: Cheema, Dave <dave.cheema@optum.com>
Sent: Wednesday, October 16, 2024 7:43 PM
To: Whitmer, Patrick <patrick_whitmer@uhc.com>; Haugan, Grant D <grant.haugan@uhg.com>
Cc: Shin, John S <john.shin@optum.com>; Alex Schmidt <alex_schmidt@uhg.com>; Baig, Muquaddar <muquaddar.baig@optum.com>; Khurana, Preeti <preeti_khurana@optum.com>
Subject: RAD Data Migration project Update

Patrick, Grant,

Just want to confirm that the following decisions were made in our today morning's working session:

- The Outlook Email data source type design specs (and solution) are no longer needed, since the attachments will be automatically captured and moved to a NAS drive. The same architecture and process flow will work as the Internal File data source
- The External file data source migration design (and development) will also be no longer needed since the External file(s) is dropped on to a NAS drive. We will need to develop one process for the Internal files, External files, and the Outlook attachment files

In summary:

- we will develop a pilot project for the API data sources
- We'll create design docs for the SQL and Internal File data source types.
- We'll expand the pilot project design doc to accommodate the remaining data sources.

Please feel to correct me if I have misstated anything. Please do not hesitate to contact me, should you have a question/concern.

Thank you.
Regards,
Dave Cheema

+++

Weekly state update report must show progress for: 1. Pilot project; 2. Design docs for SQL and Outlook Email, External file and internal file data source types

+++

RE: best practices for assign domains and subdomains to data products

Koy, Kate

Reply | Reply All | Forward

To Barker, Scott D; Shin, John S; Cheema, Dave; Huff, Kimberlie R.

Cc Shekhtman, Felix; Motsi, Fennie Nomathamsanqa; Ferguson, Kimberly; Turk, Brian J (he/him); Rajasekharan, Ranjith; Paramatmuni, Srilatha

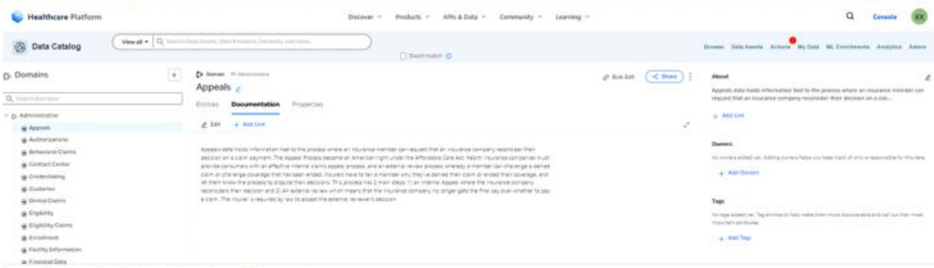
Fri 10/25/2024 4:36 PM

Retention Policy UHGInbox (90 days) Expires 1/23/2025

Hi Scott – My Chief Data Office colleague, @Huff, Kimberlie R. worked on defining these domain/subdomains. @Huff, Kimberlie R. Do you have any resources or guidance for Scott and team?

FYI Scott here are the domain definitions if those help – you can click through domains and subdomains and view definitions on the 'documentation' tabs.

[HCP Data Catalog (uhg.com)]



From: Barker, Scott D <scott_barker_@uhc.com>
Sent: Friday, October 25, 2024 4:00 PM
To: Koy, Kate <katlyn.koy@uhc.com>; Shin, John S <john.shin@optum.com>; Cheema, Dave <dave.cheema@optum.com>
Cc: Shekhtman, Felix <felix_shekhtman@uhc.com>; Motsi, Fennie Nomathamsanqa <fennie_motsi@uhc.com>; Ferguson, Kimberly <kim_ferguson@uhc.com>; Turk, Brian J (he/him) <brian_turk@optum.com>; Rajasekharan, Ranjith <ranjith_rajasekharan@uhc.com>; Paramatmuni, Srilatha <sparamatmuni@uhc.com>
Subject: best practices for assign domains and subdomains to data products

Hi all,

Do you know if there are any documentation, best practices, guidance for assigning domains and sub domains to data products? We'll look, but I thought I'd check with you as well.

Have a great weekend.

Thanks,
Scott

++

File types in scope: XLS, CSV, XML, JSON, SQL

RE: Input File Type is needed

Haugan, Grant D

To Cheema, Dave; Whitmer, Patrick; Khurana, Preeti; Baig, Muquaddar; Kumar, Mahesh; Sahoo, Asit K; Pal, Sujit

Reply | Reply All | Forward | ...

Fri 11/1/2024 9:35 AM

Retention Policy UHGInbox (90 days) Expires 1/30/2025

You replied to this message on 11/1/2024 9:44 AM.

Start your reply all with: Okay, thank you! | Perfect, thank you! | No problem, thank you! | Feedback

Yes – sorry, IT Reseller is XML. I misspoke on that piece

Grant Haugan
Senior Manager – Data Engineering
UHG Corporate Services
E grant.haugan@uhg.com

unitedhealthgroup.com

Our United Culture The way forward
Integrity | Compassion | Relationships | Innovation | Performance
Need RAD support? Submit a request in the new RAD Support Desk customer portal!

From: Cheema, Dave <dave.cheema@optum.com>
Sent: Friday, November 1, 2024 9:31 AM
To: Haugan, Grant D <grant.haugan@uhg.com>; Whitmer, Patrick <patrick_whitmer@uhc.com>; Khurana, Preeti <preeti_khurana@optum.com>; Baig, Muquaddar <muquaddar.baig@optum.com>; Kumar, Mahesh <mahesh_kumar1@optum.com>; Sahoo, Asit K <asit_sahoo@optum.com>; Pal, Sujit <sujit.pal@optum.com>
Subject: RE: Input File Type is needed

Grant,

I know there are some XML files (External File data source). Could we proceed with the understanding that we would only have XLS, XML, and JSON file types? Could you please confirm that? Thank you so much for the fast response.

Regards,
Dave Cheema

From: Haugan, Grant D <grant.haugan@uhg.com>
Sent: Friday, November 1, 2024 7:27 AM
To: Cheema, Dave <dave.cheema@optum.com>; Whitmer, Patrick <patrick_whitmer@uhc.com>; Khurana, Preeti <preeti_khurana@optum.com>; Baig, Muquaddar <muquaddar.baig@optum.com>; Kumar, Mahesh <mahesh_kumar1@optum.com>; Sahoo, Asit K <asit_sahoo@optum.com>; Pal, Sujit <sujit.pal@optum.com>
Subject: RE: Input File Type is needed

With the exception of IT Reseller (JSON), the rest of the flat files are xslx.

Thanks,

Grant Haugan
Senior Manager – Data Engineering
UHG Corporate Services
E grant.haugan@uhg.com

unitedhealthgroup.com

Our United Culture The way forward
Integrity | Compassion | Relationships | Innovation | Performance
Need RAD support? Submit a request in the new RAD Support Desk customer portal!

From: Cheema, Dave <dave.cheema@optum.com>
Sent: Thursday, October 31, 2024 9:06 PM
To: Whitmer, Patrick <patrick_whitmer@uhc.com>; Haugan, Grant D <grant.haugan@uhg.com>; Khurana, Preeti <preeti_khurana@optum.com>; Baig, Muquaddar <muquaddar.baig@optum.com>; Kumar, Mahesh <mahesh_kumar1@optum.com>; Sahoo, Asit K <asit_sahoo@optum.com>; Pal, Sujit <sujit.pal@optum.com>
Subject: Input File Type is needed

Grant, Patrick,

Could we please get the **file types** for the files listed in the **Raw Inventory** tab of the attached workbook? We need that piece of information for the implementation details. Below a sample of files that we do not know how to process them:

Adobe Sign	Signed Contract information	Flat File
	PeoplePlatform	REST API
	OT Productivity	REST API
	OT Department Breakdown	REST API
	Reports	REST API
	Misc	Flat File

Thank you.
Regards,
Dave Cheema

++

Service account: rad_dw_dev@optum.com
Password : peX7CF9gZLxJK
++

Data Product Documentation: MRIS-BILLING NON-PROD

BN Nowak, Bobbie-Jean

Reply | Reply All | Forward

To Ubele, Douglas J; Majorossy, Edward; Cheema, Dave; Saripirala, Sreenadha Reddy; Barker, Scott D; Turk, Brian J (he/him); Smith, Christopher B

Tue 11/5/2024 2:34 PM

Retention Policy UHGInbox (90 days) Expires 2/3/2025

Hello,

You are the owners listed on the MRIS-BILLING NON-PROD data product.

The product documentation will undergo changes soon, as mentioned in this post. The *Business Value Proposition / Justification, Current Business Use cases sections are currently absent from this data products. No immediate action is required from you. However, please be aware that when the migration to the new documentation section takes place, these sections will be visible within the product documentation for this product. The sections marked with an asterisk will be mandatory, so any future edits to your data product will require information in these sections. The migration is scheduled for the evening of November 6th.

Please reach out if you have any questions or concerns. Thanks!

Bobbie-Jean Nowak
HCP Data Catalog - Product | Optum Technology

M 1-612-751-4677
bobbie-jean.nowak@optum.com

++

PowerBI to Snowflake Connectivity documentation: https://engage.cloud.microsoft/main/org/optum.com/groups/eyJfdHlwZSI6Ikdyb3VwIiwiaWQiOiI0NTM4ODI1MTEzNiJ9/all

+++

Mercury API product documentation: https://www.servicenow.com/docs/bundle/xanadu-api-reference/page/integrate/inbound-rest/concept/c_TableAPI.html

+++

Link to Data Catalog Team Support team Channel: HCP Data Platform User Group | Data Catalog Support | Microsoft Teams

Data Catalog support personnel: Bobbie-Jean Nowak, Raghuram Challagundla, Kate Koy, Rajesh G Nair

++

Notes for Sahoo, Asit K:

Can you please work on below checklist

- Raise connectivity request for RAD databrick and snowflake. RAD Data Science team require this for their work.
- Update on excel python library installation in databricks cluster.
- Prepare step by step document for Azure Kye Vault creation and shared with Grant and Patrick.

+++

Notes from Dinesh related to ECIF Intake form:

1. Architecture review only
2. One page view of current state
3. Future state one pager
30. Modernize and Migrate on-prem legacy workload Azure, remove governance
37. TBD
38. co9mmunicated plan would be designed to customize for stakeholders
41. will a part of the migration, refer to geographic in Azure
42. depends on the mission critical

Approvals

Tuesday, October 15, 2024 4:51 PM

Corporate Services Reporting Analytics and Data (ORBIT PowerBI) - ECIF Architecture APPROVED

Reply | Reply All | Forward

FA Fakharzadeh, Al

To Whitmer, Patrick; Cheema, Dave; Haugan, Grant D; Shin, John S; Malhotra, Dinesh; Blomquist, Kathy; Mitchell, Lawrence E; Erickson, Brian B.; Ramsey, Jordan O; **+1 other**

Tue 1/14/2025 4:25 PM

Cc Prissovsky, Michael; Kruse, Michael D; Gesior, Luke M; Yasoju, Harish Kumar; Sri Ram, Jayadeep; Russ Jury; Gladson, Timothy S; Dubin, Corey M

Retention Policy UHGInbox (90 days) Expires 4/14/2025

ASKID: AIDE_0083635
Application Name: Corporate Services Reporting Analytics and Data (ORBIT PowerBI)

Thank you all for meeting to review the Corporate Services Reporting Analytics and Data architecture today. After reviewing the proposed architecture, we found no dependencies or risks associated with proceeding with this project, as it adheres to all current public cloud standards. The CMM team adds their formal approval to move forward in the ECIF funding process.

As mentioned on the call, the next step in the ECIF process is to work with your Strategic Vendor Partner, Optum Advisory Services, to finalize and submit the Statement of Work to Microsoft.

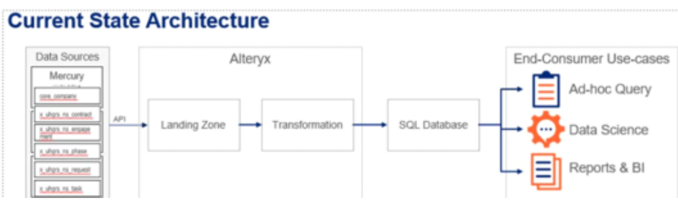
If you have any additional questions about the cloud migration, please don't hesitate to reach out to the CMM team.

Thank you,
Al

-----Original Appointment-----
From: Fakharzadeh, Al
Sent: Friday, January 10, 2025 2:16 PM
To: Fakharzadeh, Al; Whitmer, Patrick; Cheema, Dave; Haugan, Grant D; Dubin, Corey M; Gladson, Timothy S; Russ Jury; Sri Ram, Jayadeep; Yasoju, Harish Kumar
Cc: Prissovsky, Michael; Kruse, Michael D; Gesior, Luke M; Shin, John S; Malhotra, Dinesh
Subject: Corporate Services Reporting Analytics and Data (ORBIT PowerBI) - ECIF Architecture Review
When: Tuesday, January 14, 2025 4:00 PM-4:30 PM (UTC-05:00) Eastern Time (US & Canada).
Where: Microsoft Teams Meeting

Agenda:
High level architectural overview
Q&A

Please forward this meeting to anyone who can speak to the public cloud design.



++
+++++++++++++

RE: Reporting Analytics and Data(RAD) Future-State Architecture Approval

Reply | Reply All | Forward

Tue 10/15/2024 4:45 PM

WP Whitmer, Patrick

To Cheema, Dave; Sahoo, Asit K; Alex Schmidt

Cc Haugan, Grant D; Shin, John S; Baig, Muquaddar; Khurana, Preeti; Kumar, Mahesh

Retention Policy UHGInbox (90 days) Expires 1/13/2025

RAD Data Modernization - Future State Diagram v.0.1.pptx
308 KB

Start your reply all with: I approve. | Approved. | Thank you! | Feedback

Hi Dave,

Future state diagram is approved from my perspective.

@Alex Schmidt – let us know if any other concerns, otherwise please also include you approval.

Thanks,
Patrick

++
++++++++++

RE: Reporting Analytics and Data(RAD) Future-State Architecture Approval

Reply | Reply All | Forward

Alex Schmidt

To Whitmer, Patrick; Cheema, Dave; Sahoo, Asit K

Tue 10/15/2024 4:54 PM

Cc Haugan, Grant D; Shin, John S; Baig, Muquaddar; Khurana, Preeti; Kumar, Mahesh

Retention Policy UHGInbox (90 days) Expires 1/13/2025

You replied to this message on 10/15/2024 4:55 PM.

Good with me!

Thanks,
Alex

From: Whitmer, Patrick <patrick_whitmer@uhc.com>
Sent: Tuesday, October 15, 2024 4:45 PM
To: Cheema, Dave <dave.cheema@optum.com>; Sahoo, Asit K <asit_sahoo@optum.com>; Alex Schmidt <alex_schmidt@uhg.com>
Cc: Haugan, Grant D <grant.haugan@uhg.com>; Shin, John S <john.shin@optum.com>; Baig, Muquaddar <muquaddar.baig@optum.com>; Khurana, Preeti <preeti_khurana@optum.com>; Kumar, Mahesh <mahesh_kumar1@optum.com>
Subject: RE: Reporting Analytics and Data(RAD) Future-State Architecture Approval

Hi Dave,

Future state diagram is approved from my perspective.

@Alex Schmidt – let us know if any other concerns, otherwise please also include you approval.

Thanks,
Patrick

++
+++++++++++++

SQL Design Doc approval

Whitmer, Patrick
To Cheema, Dave
Cc Haugan, Grant D
Retention Policy UHGInbox (90 days) Expires 1/20/2025

Reply | Reply All | Forward
Tue 10/22/2024 9:42 AM

Start your reply all with: Thank you! | Thank you for the approval. | Thank you for your confirmation. | Feedback

SQL design document approved.

Thanks,
Patrick Whitmer | Director, Reporting, Analytics & Data, UHG Corporate Services
patrick_whitmer@uhc.com

Need RAD support? Submit a request in the new RAD Support Desk customer portal!

++

RE: SQL Design Doc approval

Whitmer, Patrick
To Cheema, Dave
Cc Haugan, Grant D
Retention Policy UHGInbox (90 days) Expires 1/20/2025
You replied to this message on 10/22/2024 1:02 PM.

Reply | Reply All | Forward
Tue 10/22/2024 1:01 PM

Start your reply all with: Thank you! | Thank you for your approval. | Great, thanks for the update. | Feedback

Hi Dave –

Grant has reviewed/updated the flat file design document and I am providing my approval.

Thanks,
Patrick

+++

RE: Request for Review and approval of Detailed Design document for the SQL Data Source type

Whitmer, Patrick
To Cheema, Dave; Haugan, Grant D
Cc Baig, Muquaddar; Khurana, Preeti; Sahoo, Asit K; Pal, Sujit; Kumar, Mahesh
Retention Policy UHGInbox (90 days) Expires 1/26/2025
You replied to this message on 10/28/2024 9:18 AM.

Start your reply all with: Thank you! | Thank you for the approval. | Great, thank you so much! | Feedback

Hi Dave – this is approved on my end.

Thanks,
Patrick

+++

RE: Review of API Data source LLD

Whitmer, Patrick
To Cheema, Dave; Kumar, Mahesh; Haugan, Grant D
Cc Baig, Muquaddar; Khurana, Preeti; Sahoo, Asit K; Pal, Sujit
Retention Policy UHGInbox (90 days) Expires 1/26/2025
You replied to this message on 10/28/2024 9:26 AM.

Reply | Reply All | Forward
Mon 10/28/2024 9:15 AM

Start your reply all with: Thank you! | Thank you for the approval. | Great, thank you so much! | Feedback

Hi Dave – this is approved on my end.

Thanks,
Patrick

+++

RE: Request for Review and approval of Detailed Design document for the Internal File Source type

Whitmer, Patrick
To Cheema, Dave; Haugan, Grant D
Cc Baig, Muquaddar; Khurana, Preeti; Sahoo, Asit K; Pal, Sujit; Kumar, Mahesh
Retention Policy UHGInbox (90 days) Expires 1/26/2025

Reply | Reply All | Forward
Mon 10/28/2024 9:34

Start your reply all with: Thank you! | Thank you for the approval. | Great, thanks! | Feedback

Hi Dave – approved on my end.

Thanks,
Patrick

+++

RE: Request for review and approval of ADF Workflow Management Framework

Whitmer, Patrick
To Cheema, Dave; Haugan, Grant D
Cc Baig, Muquaddar; Khurana, Preeti; Kumar, Mahesh; Pal, Sujit; Shin, John S
Retention Policy UHGInbox (90 days) Expires 1/29/2025
You replied to this message on 10/31/2024 8:57 AM.

Reply | Reply All | Forward
Thu 10/31/2024 8:23 AM

Start your reply all with: I can do it. | Works for me. | Either way works for me. | Feedback

\pproved on my end.

@Haugan, Grant D – do you want to do a run through on this doc internally on our own? Or do you want to do this on the Friday call with Dave
ınd team?

+++

POC acknowledgment

Whitmer, Patrick
To Cheema, Dave
Cc Haugan, Grant D
Retention Policy UHGInbox (90 days) Expires 3/19/2025

Reply | Reply All | Forward
Thu 12/19/2024 9:36 AM

Hey Dave,

Just confirming that we have seen everything we need to see for the POC and we are good with moving forward to next phase once we have confirmed with Alex that he has all his detailed questions worked out.

Thanks,
Patrick

RE: RAD Data Modernization Project Weekly Status

Whitmer, Patrick
To Alex Schmidt; Shin, John S
Cc Haugan, Grant D; Cheema, Dave; Malhotra, Dinesh
Retention Policy UHGInbox (90 days) Expires 3/19/2025

Reply | Reply All | Forward
Thu 12/19/2024 10:23 AM

Prior to tomorrow's meeting, adding my acknowledgment that everything for the technical end-to-end process has been demonstrated for the POC pilot and we are comfortable with the infrastructure setup.

Will discuss any other outstanding issues tomorrow.

Thanks,
Patrick

RE: RAD Data Modernization project weekly status update report 11/23/2024 - 11/29/2024

Alex Schmidt
To Malhotra, Dinesh; Shin, John S; Cheema, Dave
Cc Whitmer, Patrick; Haugan, Grant D; Pal, Sujit; Khurana, Preeti; Baig, Muquaddar; Kumar, Mahesh; Sahoo, Asit K
Retention Policy UHGInbox (90 days) Expires 3/20/2025
You replied to this message on 12/20/2024 12:44 PM.
This message was sent with High importance.

Reply | Reply All | Forward
Fri 12/20/2024 12:13 PM

Hi Dinesh, John and Dave,

I know I speak for my team when I share our sincerest gratitude to you and the whole team for executing this project so incredibly well.

Timing, Scope, Support, all truly has been A1.

I wish you a wonderful holiday season; and look forward to continuing building something incredible for UHG with you in 2025.

Sincerely,
Alex

+++
++++++++++++
EIS Approval Link: https://optum.service-now.com/now/nav/ui/classic/params/target/sc_req_item.do%3Fsys_id%3D343cc140934ada1805f4bb6a7bba10a8%26sysparm_domain%3Dnull%26sysparm_domain_scope%3Dnull%26sysparm_view%3D
+++
+++++++++++++

RE: RAD Data Modernization Phase 1 - Weekly Status Report

Whitmer, Patrick
To Cheema, Dave; Alex Schmidt; Haugan, Grant D
Cc Khurana, Preeti; Baig, Muquaddar; Pal, Sujit; Kumar, Mahesh; Pushadapu, Dijendra Kumari; Sahoo, Asit K
Retention Policy UHGInbox (90 days) Expires 6/2/2025
You forwarded this message on 3/4/2025 9:56 AM.

Reply | Reply All | Forward
Tue 3/4/2025 9:26 AM

Hi Dave,

Final deliverables received and approved.

Thanks,
Patrick

From: Cheema, Dave <dave.cheema@optum.com>
Sent: Saturday, March 1, 2025 6:59 PM
To: Alex Schmidt <alex_schmidt@uhg.com>; Whitmer, Patrick <patrick_whitmer@uhc.com>; Haugan, Grant D <grant.haugan@uhg.com>
Cc: Khurana, Preeti <preeti_khurana@optum.com>; Baig, Muquaddar <muquaddar.baig@optum.com>; Pal, Sujit <sujit.pal@optum.com>; Kumar, Mahesh <mahesh_kumar1@optum.com>; Pushadapu, Dijendra Kumari <dijendra_pushadapu@optum.com>; Sahoo, Asit K <asit_sahoo@optum.com>
Subject: RE: RAD Data Modernization Phase 1 - Weekly Status Report

Alex, et al,

Please find attached the final Weekly Status update Report for the RAD Data Modernization Phase 1 project for the week of 02/22/2025 – 02/28/2025 for your review and reference. Included in this email is the **Final Deliverables** presentation for same project as well.

Patrick, please review the Final Deliverables presentation, review it closely. If you agree, please respond with "Received and approved".

Please do not hesitate to contact me, should you have any questions/concerns. Thank you.

Regards,
Dave Cheema
+++
+++++++++++

RE: 0$ SOW Approval requested: RAD Data Modernization Implementation

Alex Schmidt
To: Cheema, Dave; Whitmer, Patrick; Baxter, Ken S

Reply | Reply All | Forward

Retention Policy: UHGInbox (90 days) Expires: 8/12/2025

Fri 3/14/2025 2:37 PM

You replied to this message on 3/14/2025 3:08 PM.

Reviewed and approved.

From: Cheema, Dave <dave.cheema@optum.com>
Sent: Friday, March 14, 2025 11:14 AM
To: Alex Schmidt <alex_schmidt@uhg.com>; Whitmer, Patrick <patrick_whitmer@uhc.com>; Baxter, Ken S <ken_s_baxter@uhc.com>
Subject: FW: 0$ SOW Approval requested: RAD Data Modernization Implementation
Importance: High

Gentlemen,
Could you please provide an approval for the 0$ SOW shown below, please make it a priority, we need to send an invoice to the Microsoft ECIF team. Thanks in advance.

Regards,
Dave Cheema

From: Cheema, Dave
Sent: Wednesday, March 12, 2025 9:16 PM
To: Alex Schmidt <alex_schmidt@uhg.com>; Whitmer, Patrick <patrick_whitmer@uhc.com>
Cc: Shin, John S <john.shin@optum.com>; Johnson, Scott D <scott.d.johnson@optum.com>; GoPayerTechnology <GoPayerTechnology@uhgazure.onmicrosoft.com>
Subject: 0$ SOW Approval requested: RAD Data Modernization Implementation

RAD Data Modernization Implementation
0$ SOW & Pricing Proposal

Hello Alex, Patrick,

Thank you for choosing the Optum Advisory– Digital Transformation team to help support for your continued effort to modernize RAD Data Platform. Attached you will find the 0$ SOW & Pricing Proposal for this initiative. Please let me know if you have questions. I appreciate your trust in our team to support you.

Snapshot:

- Ask: **RAD Data Modernization Implementation**
- Duration: **1/24/2025 - 9/26/2025**
- Total Cost Estimate: **$725,002**

Next Steps:

1. **Review** the attached SOW for scope, pricing, terms, and conditions.
2. If you agree with the terms, simply respond to this email with **"Reviewed and agree."**
3. You will receive an email notification from Optum's 'iTRAC Tool' for your **formal online approval** of this SoW
4. The **Finance team** member too need to provide their approval via 'Intake Tool'.
   Please confirm the name of your Finance Team member along with the GL string for billing.
   a. GL String
      - BU: 20020
      - OU: 01000
      - Dept: 100505
      - Location: USASS800
   b. Finance Contact: Ken Baxter | ken_s_baxter@uhc.com

OAS Digital Transformation Key Business and Technical Contacts:

- John Shin
- Dinesh Malhotra
- Dave Cheema

Optum

RE: ECIF Milestone Completion Sign-off: Please Update & Respond

Whitmer, Patrick
To: Shin, John S
Cc: Malhotra, Dinesh; Cheema, Dave

Reviewed, approved, and signed below.

Thanks,
Patrick

From: Shin, John S [illegible]
Sent: Friday, March 14, 2025 [illegible]
To: Whitmer, Patrick [illegible]
Cc: Malhotra, Dinesh [illegible]; Cheema, Dave [illegible]
Subject: ECIF Milestone Completion Sign-off: Please Update & Respond

RAD Implementation - ECIF Milestone Completion Sign-off

Dear Patrick – Please complete the placeholders highlighted in blue to send it back as your approval / certification for the completion of the RAD Implementation Milestone per details below.

Microsoft Supplier Number: [illegible]

Microsoft SoW ID: [illegible]

Microsoft PO #: [illegible]

Project Name: [illegible]

Milestone: [illegible]

Milestone Invoice Amount: [illegible]

Milestone Description:

[illegible]

Milestone Completion Date: [illegible]

Client Approval:

[illegible]

Client Name: Patrick Whitmer

Client Title: Director, Business Analysis Services, UHG Corporate Services

Client Company / LoB: UHG Corporate Services

Client Email: [illegible]

Client Signature: Patrick Whitmer

Certification Date: [illegible]

Thanks,

John Shin
[illegible]

Optum

Integrity • Compassion • Inclusion • Relationships • Innovation • Performance

++
++++++++++++

RE_
Request f...

RE: Request for Approval and Receipt of milestone 2 completion

Whitmer, Patrick
To Cheema, Dave
Cc Alex Schmidt; Shin, John S
Retention Policy UHGInbox (90 days) Expires 7/8/2025
You replied to this message on 4/9/2025 3:00 PM.

Reply | Reply All | Forward
Wed 4/9/2025 2:59 PM

Received and Approved

From: Cheema, Dave <dave.cheema@optum.com>
Sent: Tuesday, April 8, 2025 2:41 PM
To: Whitmer, Patrick <patrick_whitmer@uhc.com>
Cc: Alex Schmidt <alex_schmidt@uhg.com>; Shin, John S <john.shin@optum.com>
Subject: Request for Approval and Receipt of milestone 2 completion

Patrick,
Please find attached the RAD Data Modernization – Milestone 2 completion deliverables presentation. **Note**: this is an actual project progress report. Please review it closely and let me know if you have any questions/concerns. If everything looks okay and acceptable, please respond with "Received and approved". Thanks in advance.
Regards,
Dave Cheema

++
++++++++++++
Client approval for milestone 2:

RE_ ECIF Milestone...

++
++++++++++++
RAD Data Modernization Bridge project refund:

Contract Change 1...

Contract Change 1950 Has Been Completed

Microsoft Power Apps and Power Automate <microsoft@powerapps.cc
To Cheema, Dave
Cc Nelson, Alexander S
Retention Policy UHGInbox (90 days) Expires 7/27/2025

Caution: External email. Do not open attachments or click on links if you do not recognize the sender.

Hello,

The following item has been completed:

Title: Termination of existing contract
Status: Submitted
Date Booked in ERP: 2025-04-28
Description: Charges to be refunded to UHG to same GL that was initially charged.
Assigned Biller: {"@odata.type":"#Microsoft.Azure.Connectors.SharePoint.SPListExpandedUser","Claims":"i:0#.f|membership|alexander.nelson@optum.com","DisplayName":"Nelson, Alexander S","Email":"alexander.nelson@optum.com","Picture":"https://uhgazure.sharepoint.com/sites/OCBR/_layouts/15/UserPhoto.aspx?Size=L&AccountName=alexander.nelson@optum.com","Department":"Optum Finance & Accounting","JobTitle":"Sr Acctnt"}

If you have any questions, please reach out to arbilling@optum.com.

To view the item, click here: [illegible]

If you want to unsubscribe from these emails, please use this form.
++
+++++++++++

RE_ Request f...

RE: Request for Approval and Receipt of milestone 2 completion

Whitmer, Patrick
To Cheema, Dave
Cc Alex Schmidt; Shin, John S; Malhotra, Dinesh; Kalidindi, Venkata

Reply | Reply All | Forward

Received and approved

From: Cheema, Dave <dave.cheema@optum.com>
Sent: Monday, May 5, 2025 9:25 PM
To: Whitmer, Patrick <patrick_whitmer@uhc.com>
Cc: Alex Schmidt <alex_schmidt@uhg.com>; Shin, John S <john.shin@optum.com>; Malhotra, Dinesh <dinesh.malhotra@optum.com>; Kalidindi, Venkata <venkata.kalidindi@optum.com>
Subject: RE: Request for Approval and Receipt of milestone 2 completion

Patrick,
Please find attached the RAD Data Modernization – Milestone 3 completion deliverables presentation.
Note: this is an actual project progress report. Please review it closely and let me know if you have any questions/concerns.

If everything looks okay and acceptable, please respond with "Received and approved". Thanks in advance.
Regards,
Dave Cheema

++
++++++++++++

Your workshee...

Your worksheet approval has been processed

ARBilling@Optum.com
To Cheema, Dave
Retention Policy UHGInbox (90 days)

Reply | Reply All | Forward

Customer Name:	UHG Corporate
Project Name:	27105 - RAD Data Modernization Implementation 2025 - EC >>>
Total Invoice Amount:	$-138,500.00
Bill Cycle ID:	Optum Insight Consulting
Billing Worksheet:	TMP-0000093731
Contract:	CON000000020407

++
++++++++++++
Milestone #3 approval email:

RE_ ECIF Milestone...

RE: ECIF Milestone Completion Sign-off: Please Update & Respond

Whitmer, Patrick
To Cheema, Dave
Cc Malhotra, Dinesh; Shin, John S
Retention Policy UHGInbox (90 days) Expires 8/12/2025
You replied to this message on 5/14/2025 4:44 PM.

Reply | Reply All | Forward

Wed 5/14/2025 4:42 PM

Start your reply all with: Got it, thanks! | Thank you for sharing. | Thank you! | Feedback

See below.

Thanks,
Patrick

From: Cheema, Dave <dave.cheema@optum.com>
Sent: Wednesday, May 14, 2025 11:18 AM
To: Whitmer, Patrick <patrick_whitmer@uhc.com>
Cc: Malhotra, Dinesh <dinesh.malhotra@optum.com>; Shin, John S <john.shin@optum.com>
Subject: ECIF Milestone Completion Sign-off: Please Update & Respond

RAD Implementation - ECIF Milestone Completion Sign-off

Dear Patrick – Please complete the placeholders highlighted in blue to send it back as your approval / certification for the completion of the RAD Implementation Milestone per details below.

Microsoft Supplier Number: 3067563

Microsoft SoW ID: SOW_3067563_Optum Insight Inc_CAS-1737011-Q2V5H9_724702146

Microsoft PO #: Microsoft PO #: 101279849 - CAS-1737011-Q2V5H9 / BDI-C – United Health Group – FY25-26 (Optum Insight, Inc. $725,002.00)

Project Name: POE_Supplier_3067563_OptumInsight Inc_CAS-1737011-Q2V

5H9_1044905659

Milestone: 3 of 8

Milestone Invoice Amount: $94,301

Milestone Description:
Development & Unit Testing:
- Migrate remainder complex 9 API datasets from data source Mercury into Bronze and Silver layers of Medallion architecture
- WEBAPI, ADF, Azure Blob Storage, Custom Framework
- Migrate 4 peoplesoft datasets into the Bronze and Silver layers.
- SHIR, ADF, Azure Blob Storage, VNET, Custom Scripting

Milestone Completion Date: May 15, 2025

Client Approval:

I, Patrick Whitmer, Director, Business Analysis Services, UHG Corporate Services, hereby acknowledge that the above-mentioned milestone has been completed and delivered to my satisfaction. By signing this document, I certify that the deliverable meets all the agreed-upon requirements and standards. I approve the above-mentioned milestone for invoicing to Microsoft.

Client Name: Patrick Whitmer

Client Title: Director, Business Analysis Services, UHG Corporate Services

Client Company / LoB: UHG Corporate Services

Client Email: Patrick Whitmer

Client Signature: **Patrick Whitmer**

Certification Date: **5/14/2025**

Thanks & regards,
Dave Cheema

Optum

Integrity • Compassion • Inclusion • Relationships • Innovation • Performance

++
++++++++++++

RE_
Request f...

RE: Request for review and Approval for RAD Data Modernization project Weekly Status Milestone 4...

Whitmer, Patrick
To Cheema, Dave
Retention Policy UHGInbox (90 days) Expires 9/2/2025
You replied to this message on 6/4/2025 10:28 AM.

Reply | Reply All | Forward

Wed 6/4/2025 10:22 AM

Reviewed and Accepted

From: Cheema, Dave <dave.cheema@optum.com>
Sent: Tuesday, June 3, 2025 10:11 AM
To: Whitmer, Patrick <patrick_whitmer@uhc.com>
Subject: Request for review and Approval for RAD Data Modernization project Weekly Status Milestone 4 Completion report

Patrick,
Please find attached Status update report for the RAD Data Modernization project Weekly Status - Milestone 4 Completion report (milestone focused report, which ended on 06/03/2025).

Please review it closely. If everything looks acceptable, please respond with "Reviewed and Accepted". Thanks in advance.
Regards,
Dave Cheema

++
+++++++++++++

RE_
Request f...

RE: Request for Approval and Receipt of milestone 4 completion

Whitmer, Patrick
To Cheema, Dave
Cc Alex Schmidt; Shin, John S; Malhotra, Dinesh; Kalidindi, Venkata
Retention Policy UHGInbox (90 days) Expires 9/2/2025
Wed 6/4/2025 10:17 AM
You replied to this message on 6/4/2025 10:26 AM.

Received and Approved

From: Cheema, Dave <dave.cheema@optum.com>
Sent: Tuesday, June 3, 2025 10:03 AM
To: Whitmer, Patrick <patrick_whitmer@uhc.com>
Cc: Alex Schmidt <alex_schmidt@uhg.com>; Shin, John S <john.shin@optum.com>; Malhotra, Dinesh <dinesh.malhotra@optum.com>; Kalidindi, Venkata <venkata.kalidindi@optum.com>
Subject: RE: Request for Approval and Receipt of milestone 4 completion

Sorry, trying to do too many things at once.

Regards,,
Dave Cheema

From: Cheema, Dave
Sent: Tuesday, June 3, 2025 9:39 AM
To: Whitmer, Patrick <patrick_whitmer@uhc.com>
Cc: Alex Schmidt <alex_schmidt@uhg.com>; Shin, John S <john.shin@optum.com>; Malhotra, Dinesh <dinesh.malhotra@optum.com>; Kalidindi, Venkata <venkata.kalidindi@optum.com>
Subject: Request for Approval and Receipt of milestone 4 completion

Patrick,

Please find attached the RAD Data Modernization – Milestone 4 completion deliverables presentation.
Note: this is an actual project progress report. Please review it closely and let me know if you have any questions/concerns.

If everything looks okay and acceptable, please respond with "Received and approved". Thanks in advance.

Regards,
Dave Cheema
++
+++++++++++

RE_ ECIF Milestone...

RE: ECIF Milestone Completion Sign-off: Please Update & Respond

Whitmer, Patrick
To Cheema, Dave
Cc Malhotra, Dinesh; Shin, John S
Retention Policy UHGInbox (90 days) Expires 9/14/2025
Mon 6/16/2025 9:46 AM
You replied to this message on 6/16/2025 9:48 AM.

Hi Dave, see updated form below.

Thanks,
Patrick

From: Cheema, Dave <dave.cheema@optum.com>
Sent: Friday, June 13, 2025 1:35 PM
To: Whitmer, Patrick <patrick_whitmer@uhc.com>
Cc: Malhotra, Dinesh <dinesh.malhotra@optum.com>; Shin, John S <john.shin@optum.com>
Subject: ECIF Milestone Completion Sign-off: Please Update & Respond

RAD Implementation - ECIF Milestone Completion Sign-off

Dear Patrick – Please complete the placeholders highlighted in blue to send it back as your approval / certification for the completion of the RAD Implementation Milestone per details below.

Microsoft Supplier Number: 3067563

Microsoft SoW ID: SOW_3067563_Optum Insight Inc_CAS-1737011-Q2V5H9_724702146

Microsoft PO #: Microsoft PO #: 101279849 - CAS-1737011-Q2V5H9 / BDI-C – United Health Group – FY25-26 (Optum Insight, Inc, $725,002.00)

Project Name: POE_Supplier_3067563_OptumInsight Inc_CAS

-1737011-Q2V5H9_1044905659

Milestone: 4 of 8

Milestone Invoice Amount: $ 93,838

Milestone Description:
Development & Unit Testing:
• Transform and migrate silver layer tables, created in the milestone #3 into ~52 Gold layer views of the medallion architectures.
• ADF, Azure Blob Storage, ETL, Custom Framework, Custom Scripting
• Migrate 20 ESG datasets into the Bronze and Silver Layers of Medallion architecture.
• AZCopy, Azure Blob Storage, ADF, ETL, Custom Framework, Custom Scripting

Milestone Completion Date: June 15, 2025

Client Approval:

I, Patrick Whitmer, Director, Business Analysis Services, UHG Corporate Services, hereby acknowledge that the above-mentioned milestone has been completed and delivered to my satisfaction. By signing this document, I certify that the deliverable meets all the agreed-upon requirements and standards. I approve the above-mentioned milestone for invoicing to Microsoft.

Client Name: Patrick Whitmer

Client Title: Director, Business Analysis Services, UHG Corporate Services

Client Company / LoB: UHG Corporate Services

Client Name: Patrick Whitmer

Client Title: Director, Business Analysis Services, UHG Corporate Services

Client Company / LoB: UHG Corporate Services

Client Email: Patrick_whitmer@uhc.com

Client Signature: Patrick Whitmer

Certification Date: 6/16/2025

Thanks & regards,
Dave Cheema

Optum

Integrity • Compassion • Inclusion • Relationships • Innovation • Performance

Access Information

Saturday, August 17, 2024 7:47 AM

Hcc-intake: hcc-intake/ISDC/ https://github.com/optum-eeps/hcc-intake/tree/main/ISDC

RAD Snowflake: https://app.snowflake.com/uhg/uhgdwaas

Azure cloud user ID: dcheema1@optumcloud.com
https://docs.hcp.uhg.com/strategic-data-repository-platform-(sdrp)/storage-tenant

ASK ID: AIDE_0083635

Azure:
Cloud account name: CorpSvcs_RAD
Account ID: 58a5fc58-629a-4f16-9e2d-7b94d2afef5d
Global Groups:
UHGRG_rg_rad_azurestack_08a9b07_ADMIN
UHGRG_rg_rad_azurestack_08a9b07_CONTRIBUTOR
UHGRG_rg_rad_azurestack_08a9b07_READ

Resource Group SDRP groups:
UHGRG_rg_rad_sdrp_snowflake_a1fbf6e_ADMIN
UHGRG_rg_rad_sdrp_snowflake_a1fbf6e_CONTRIBUTOR
UHGRG_rg_rad_sdrp_snowflake_a1fbf6e_READ

Patrick,

Thank you for meeting with me on such a short notice. I would like to discuss the following items:

- Azure Cloud subscription
- SDRP tenant
- Secure roles to access Azure cloud subscription and SDRP tenant
- Access to current solution & codebase
- Available specifications for the existing data processes
- SMEs & their availabilities
- Secure roles to get access to the current solutions, processes, architecture diagrams, and workflows
- Login credentials and service endpoints to the in-scope data sources
- Resource gap

+++

USE SECONDARY ROLES ALL;

Tenant: RAD
prefix: AWKUR
Environment: Dev & Prod
Snowflake Roles:
Developer - AZU_SDRP_AKWUR_PRD_DEVELOPER_ROLE & AZU_SDRP_AKWUR_DEV_DEVELOPER_ROLE
AZU_SDRP_AKWUR_DEV_DEVELOPER_ROLE

Support - AZU_SDRP_AKWUR_PRD_SUPPORT_ROLE & AZU_SDRP_AKWUR_DEV_SUPPORT_ROLE
Analyst - AZU_SDRP_AKWUR_PRD_ANALYST_ROLE & AZU_SDRP_AKWUR_DEV_ANALYST_ROLE
Secure Global Groups (roles)
Prod/Dev:
RAD Storage Tenant > Tenant Analyst (AKWUR); Tenant Developer (AKWUR); Tenant Support (AKWUR)

- Access to current solution & codebase
  - You can request access to the following global groups in Secure to get access to our SQL databases:
  - We currently use Alteryx as our ETL tool. We can send you the workflows if you'd like.
- Available specifications for the existing data processes.
  - Is this referring to our data transformation rules?
- SMEs & their availabilities
  - Melissa Lanham – West Coast
  - Grant Haugan – Central time
  - Nathan Drapeau – Central Time
  - Lakshay Munjal – India

+++

From: Haugan, Grant D <grant.haugan@uhg.com>
Sent: Wednesday, October 2, 2024 11:30 AM
To: Whitmer, Patrick <patrick_whitmer@uhc.com>; Cheema, Dave <dave.cheema@optum.com>
Subject: RE: OAS info for pilot

Here's what I have Dave -

Tenant: RAD
prefix: AWKUR
Environment: Dev & Prod
Snowflake Roles:
Developer - AZU_SDRP_AKWUR_PRD_DEVELOPER_ROLE & AZU_SDRP_AKWUR_DEV_DEVELOPER_ROLE
Support - AZU_SDRP_AKWUR_PRD_SUPPORT_ROLE & AZU_SDRP_AKWUR_DEV_SUPPORT_ROLE
Analyst - AZU_SDRP_AKWUR_PRD_ANALYST_ROLE & AZU_SDRP_AKWUR_DEV_ANALYST_ROLE

Secure Global Groups (roles)
[DC] - Could you please provide me with the Global Groups (roles)?
Prod/Dev:
RAD Storage Tenant > Tenant Analyst (AKWUR); Tentant Developer (AKWUR); Tenant Support (AKWUR)

- Access to current solution & codebase
  - You can request access to the following global groups in Secure to get access to our SQL databases:

    [DC] - Which global groups in Secure do I need to request? Is the Secure group for the database going to be Corpvcs_ESP_Spend_DbReadOnly?

    Which database type is this going to be – ORACLE, SQL, or both?

    To extract DDLs, you would need db_ddladmin or some other similar role

  - We currently use Alteryx as our ETL tool. We can send you the workflows if you'd like.

  [DC] - I meant access to platform where current solution runs. And the codebase that is running for this solution. Since there are the following data source types are in-scope: API data source, Outlook Email data source, Oracle/SQL data source, Internal File data source, and External data source, I would like to know what the address and which credentials is to use to get access to each of those solutions and their codebases.

- Available specifications for the existing data processes.
  - Is this referring to our data transformation rules?

[DC] – No. By that I meant, any available requirements, workflow diagrams, architecture diagrams, and documentation.

- SMEs & their availabilities
  - Melissa Lanham – West Coast
  - Grant Haugan – Central time
  - Nathan Drapeau – Central Time
  - Lakshay Munjal – India

  [DC] – Which area(s) are these folks SMEs of, e.g., Melissa Lanham – Oracle/SQL data source, etc.?

+++

From: Haugan, Grant D <grant.haugan@uhg.com>
Sent: Wednesday, October 2, 2024 1:19 PM
To: Cheema, Dave <dave.cheema@optum.com>; Whitmer, Patrick <patrick_whitmer@uhc.com>
Subject: RE: OAS info for pilot

Hey Dave –

All good and appreciate the follow-up. I've added my comments my below inline as it pertains to the Mercury dataset work we're looking to do with your team.

Tenant: RAD
prefix: AWKUR
Environment: Dev & Prod
Snowflake Roles:
Developer - AZU_SDRP_AKWUR_PRD_DEVELOPER_ROLE & AZU_SDRP_AKWUR_DEV_DEVELOPER_ROLE
Support - AZU_SDRP_AKWUR_PRD_SUPPORT_ROLE & AZU_SDRP_AKWUR_DEV_SUPPORT_ROLE
Analyst - AZU_SDRP_AKWUR_PRD_ANALYST_ROLE & AZU_SDRP_AKWUR_DEV_ANALYST_ROLE
Secure Global Groups (roles)
[DC] - Could you please provide me with the Global Groups (roles)?
GH: I think we'd need to engage Secure team for this. Those roles are requested in Secure via Application > DWaaS> Then navigating to the RAD Storage Tenant.
Prod/Dev:
RAD Storage Tenant > Tenant Analyst (AKWUR); Tenant Developer (AKWUR); Tenant Support (AKWUR)

- Access to current solution & codebase
  - You can request access to the following global groups in Secure to get access to our SQL databases:

    [DC] - Which global groups in Secure do I need to request? Is the Secure group for the database going to be Corpvcs_ESP_Spend_DbReadOnly? – for read only access yes.

    Which database type is this going to be – ORACLE, SQL, or both?: SQL (Optum Azure Cloud)

    To extract DDLs, you would need db_ddladmin or some other similar role
    GH: We do have a global group that my team is part of which grants DBO access. If that's something you need I can pass it along. Just want to be cautious with it as it provides DBO access. Otherwise we could work with our Business Segment DBA for getting DDL files.

  - We currently use Alteryx as our ETL tool. We can send you the workflows if you'd like.

  [DC] - I meant access to platform where current solution runs. And the codebase that is running for this solution. Since there are the following data source types are in-scope: API data source, Outlook Email data source, Oracle/SQL data source, Internal File data source, and External data source, I would like to know what the address and which credentials is to use to get access to each of those solutions and their codebases.
  GH: For this effort we're focusing on our Mercury dataset. Mercury is an application that sits on the Optum ServiceNow platform and was/is built/maintained by the Optum Tech group under Scott Greggor. Today we use a REST API to get the majority of the data out via our Alteryx workflows – within Alteryx we build an end point URL, download the data, parse it and then transform before loading into SQL - along with 1 flat file they send to us via ECG on a daily basis. Not sure if you're looking for access to the codebase for Mercury, if so we can engage the Mercury Dev group (OT developers). If you're looking for how the Alteryx ETL workflows run I'm more than happy to show you that. We have some documentation that shows which Mercury API tables and their fields get pulled and the SQL tables they're stored in. I would be more than happy to send that over.

- Available specifications for the existing data processes.
  - Is this referring to our data transformation rules?

[DC] – No. By that I meant, any available requirements, workflow diagrams, architecture diagrams, and documentation.
GH: Great question and this is where the challenge is. There's a lot of data transformations (some complex, some not) that we do within our Alteryx workflows for the Mercury data. Unfortunately we don't have much documentation on it as the work has been built out over a time prior to our new process which better documents them. We have some documentation that shows which Mercury API tables get pulled and stored (along with their fields) into certain SQL tables.

- SMEs & their availabilities
    - Melissa Lanham – West Coast
    - Grant Haugan – Central time
    - Nathan Drapeau – Central Time
    - Lakshay Munjal – India

[DC] – Which area(s) are these folks SMEs of, e.g., Melissa Lanham – Oracle/SQL data source, etc.?

GH: These are the SME's/Data Engineers from my team for the various Mercury Data sets.

Grant Haugan
Associate Director of Reporting, Analytics, & Data
UHG Corporate Services
E grant.haugan@uhg.com

++

SECURE ROLE for RAD Data Modernization:

- Global Groups: **UHGRG_rg_rad_azurestack_08a9b07_ADMIN**
- Resource Group SDRP groups: **UHGRG_rg_rad_sdrp_snowflake_a1fbf6e_CONTRIBUTOR**
- Snowflake Roles: Developer - **AZU_SDRP_AKWUR_PRD_DEVELOPER_ROLE** & **AZU_SDRP_AKWUR_DEV_DEVELOPER_ROLE**

++

ROLE: AZU_SDRP_AKWUR_DEV_DEVELOPER_ROLE

Miscellaneous

Saturday, August 17, 2024 7:47 AM

Connecting On-Premise API to Azure Functions: A Comprehensive Guide

Understanding the Challenge

Connecting an on-premise API to Azure Functions introduces complexities due to network isolation. This guide will explore potential solutions and best practices to bridge the gap effectively.

Key Considerations:

- **Network Connectivity:** How will your Azure Function access the on-premise API?
- **Data Security:** How will you protect sensitive data transmitted between the two environments?
- **Performance:** What is the expected traffic and latency?
- **Reliability:** How will you handle failures and retries?

Potential Solutions:

1. Using Azure Functions with the Office365 Connector:

- Create an Azure Function.
- Trigger the function based on new email arrivals in your Outlook mailbox.
- Use the Office365 Connector to access the email content and attachments.
- Extract the attachments and process them as needed (e.g., save to Blob storage, send to another service).

2. Using Azure Logic Apps:

- Create an Azure Logic App.
- Trigger the Logic App based on new email arrivals in your Outlook mailbox.
- Use the Office365 Connector to access the email content and attachments.
- Extract the attachments and process them as needed (e.g., save to Blob storage, send to another service).

3. Using Azure Data Factory:

- Create an Azure Data Factory pipeline.
- Use the Copy activity to copy emails from your Outlook mailbox to a storage account.
- Use the Derived Column transformation to extract attachments from the email body.
- Process the attachments as needed (e.g., save to Blob storage, send to another service).

4. Using Azure Blob Storage and Azure Functions (for large attachments):

- Create an Azure Function triggered by new email arrivals.
- Use the Office365 Connector to access the email content and attachments.
- For large attachments, upload them directly to Azure Blob Storage using the Blob Storage SDK.
- Process the attachments in the Azure Function or trigger another function to process them.

5. Using third-party tools and services:

- There are many third-party tools and services available that can extract Outlook email attachments. Some popular options include:
  - **Mailparser.io:** A cloud-based email parsing API that can extract attachments and other information from emails.
  - **ParseHub:** A web scraping tool that can also be used to extract attachments from emails.
  - **Zapier:** A popular automation tool that can integrate with Outlook and other services to extract attachments.

Key considerations when choosing a method:

- **Attachment size:** If you're dealing with large attachments, using Azure Blob Storage and Azure Functions might be more efficient.
- **Complexity:** Some methods, like using Azure Data Factory, might be more complex to set up and manage.
- **Cost:** Consider the costs associated with each method, including Azure Function usage, storage costs, and third-party tool fees.
- **Scalability:** Ensure that the chosen method can scale to handle your expected workload.

++

A few options to extract outlook email attachment in Azure:

1. Azure Logic Apps: You can use Azure Logic Apps to create a workflow that monitors your Outlook email for new messages, and then extracts the attachments to a desired location, such as Azure Blob storage or Azure Data Lake Storage. You can use the "Get emails" trigger in Logic Apps to retrieve email messages, and then use the "Get attachment content" action to extract the attachments.
2. Azure Functions: You can create an Azure Function that connects to your Outlook mailbox using the Microsoft Graph API. The function can be triggered by a new email event, and then use the API to retrieve the email attachments and save them to a storage account or any other desired location.
3. Azure Data Factory: Azure Data Factory is a cloud-based data integration service that can be used to extract data from various sources, including Outlook. You can create a pipeline in Azure Data Factory that connects to your Outlook mailbox and retrieves the email attachments. The attachments can be saved to Azure Blob storage or any other supported storage location.

++

How to create an on-premises API that triggers an Azure Function:

1. **Set up an on-premises API**:
   - Install Azure API Management on your on-premises environment.
   - Create a new API in Azure API Management and define the desired endpoints and operations.
   - Configure the API to use the on-premises data source or backend system.
2. **Create an Azure Function**:
   - Go to the Azure portal and create a new Azure Function.
   - Choose the appropriate runtime and programming language for your function.
   - Write the code for your function to perform the desired actions or operations.
   - Make sure your function is set up to receive HTTP requests.
3. **Connect the on-premises API to the Azure Function**:
   - In Azure API Management, go to the API you created earlier.
   - Add a new operation that corresponds to the Azure Function you created.
   - Configure the operation to use the Azure Function as the backend.
   - Define the request and response mappings as needed.
4. **Test the integration**:
   - Use a REST client or any other tool to send a request to the on-premises API.
   - Verify that the request is properly routed to the Azure Function.
   - Check the response from the Azure Function and ensure that it meets your expectations.

++

Work Plan

Scope - API data sources, Outlook Email data sources, SQL/Oracle data sources, Internal file data sources, External File

Meet with RAD SME, request all available artifacts
Request APIs patterns, sources and target systems, and technology being used
Variety, Volume, Velocity
Identify initial dataload and transactional dataloads
Collate all material and data points
Analyze data, document findings and opportunities

Create solution pattern(s) for each unique workload type
Identify tech stack and usage of each tool
Build out cloud infrastructure
Build pilot prototypes
Create detailed solution to populate Bronze, Silver & Gold tables
Create detailed estimates of Dev, Test, Deployment

Repeat above steps for all unique data source types

++

Solution Definition & Design	TBD - Lead Cloud Data Engineer		Contractors		36	170	170	340		
Solution Definition & Design	TBD - Cloud Data Engineer		Contractors		36	170	170	340		
Solution Definition & Design	TBD - Cloud Data Engineer		Contractors		36	170	170	340		
Solution Definition & Design	TBD - Data Platform Engineer		Contractors	27		85	85	170		
Program Management & Governance	TBD - Scrum Master		Contractors	27		85	85	170		
Program Management & Governance	Dave Cheema - Solution Delivery Lead		Technology - Digital/Config	30		160	160	320		
Program Management & Governance	John Shin - Delivery Oversight		Technology - Digital/Config	31		20	20	40		
Program Management & Governance	Dinesh Malhotra - Executive Oversight		Technology - Digital/Config	32		5	5	10		
						865	865	1730		

1. Would it be RAD responsible for getting the Azure Subscription or OAS? If RAD, what will be our role in it?
    1. RAD
2. There are two stand-up calls at the same time - MRIS and Persistent (for RAD Modernization). Which one should I attend? Please advise.
    1. Find resources, do whatever is necessary
3. Are all patterns in scope or just one or two? If just one or two, which ones? One from each type seems way too much.
    1. Deliver whatever is possible, but solutioning will be all data source types

4. Review the detailed project plan and create a pragmatic plan
    1. Work with RAD folks and take on whatever is possible. Be sure to deliver whatever is started.
5. Below a table of data sources in scope by type for reference purpose.

API	Outlook Email	SQL/Oracle	Internal Files	External Files
Mercury OT DataHub Field Glass M&A Masterbrain SESH (Salesforce) Employee Learning & Development	CBRE Hubstar HR/Peoplesoft/AP (xls) ESG HireRight	LENEL M&A HR/Peoplesoft/AP (SQL/Oracle) One-Click Concur Cority	Ariba (flat) Noosh (flat) AdobeSign (xls) eGRC (xls)	IT

++

There are several frameworks that can be used to collect the current state of a system. Here are a few popular options:'

SWOT	SWOT (Strengths, Weaknesses, Opportunities, and Threats) analysis is a commonly used framework to assess the current state of a system. It involves identifying and analyzing the internal strengths and weaknesses of the system, as well as the external opportunities and threats it faces.
Gap Analysis	Gap analysis is a framework used to identify the gaps between the current state and the desired state of a system. It involves comparing the current performance or capabilities of the system with the desired goals or standards to determine areas of improvement.
PESTEL	PESTEL (Political, Economic, Social, Technological, Environmental, and Legal) analysis is a framework used to assess the external factors that may impact a system. It helps identify and understand the current state of the system by analyzing the broader macro-environmental factors.
Balanced Scorecard	The Balanced Scorecard is a strategic management framework that provides a comprehensive view of the current state of a system. It incorporates multiple perspectives, such as financial, customer, internal processes, and learning and growth, to evaluate the performance and progress of the system.
Root Cause Analysis	Root cause analysis is a systematic approach to identifying the underlying causes of problems or issues in a system. By analyzing the current state and investigating the root causes of any challenges or shortcomings, this framework helps to understand the system's current state more comprehensively.

Outlook Email:

What are we trying to migrate to the cloud

If only attachments, then the following data sources are all attachment of type XLS

- CBRE
- Hubstar
- HR/Peoplesoft/AP
- ESG
- HireRight

Are there any other data sources?

Are there any other attachment content types?

What is the frequency of these files?

What is the size of these files?

Can there be multiple attachments in an email?

Are XLS format consistent or they can change from workload to workload?

Do they need to be migrated to the cloud immediately or they can be batched?

How are they expected to be migrated - store them as they are in a blob storage or they need to be parsed and stored into a data warehouse?

Can the same attachment be received multiple times by multiple recipients?

If yes, how are they expected to be migrated?

Could you please provide us samples of each data source type file received?

Could you please provide information of the FROM outlook senders, sending relevant emails?

Could you please provide us the receiving Outlook mailbox credentials and Exchange server info?

SQL/Oracle:

Could you please help us understand the scope of SQL/Oracle data source migration in -scope

Which RDBMS is in-scope:

SQL:

Oracle:

Both:

Which tables/ views/database objects in scope

What is the Volume, Velocity & Variety in-scope

Any existing solution patterns

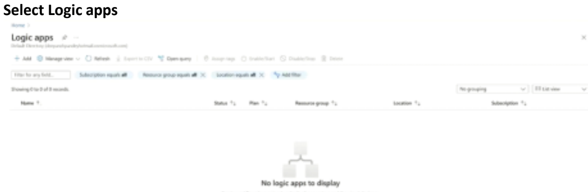
Any issues with data access and/or migration

Could you please provide access credentials

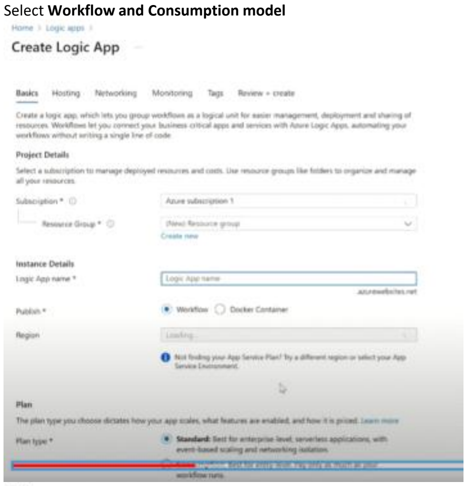
++

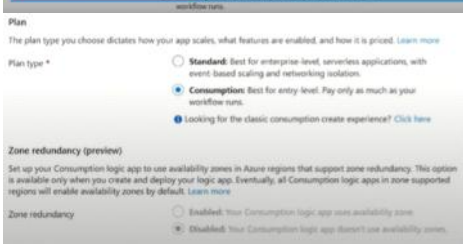
Main Sample using Azure Logic Apps: [How to save email attachments in Azure Storage with Logic Apps? - Azure 2023]

Select Logic apps

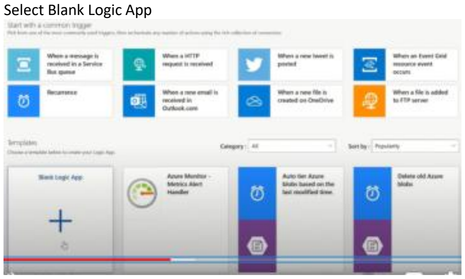


Select **Workflow and Consumption model**

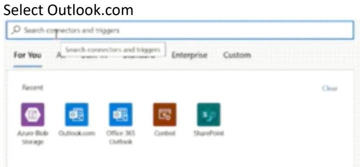




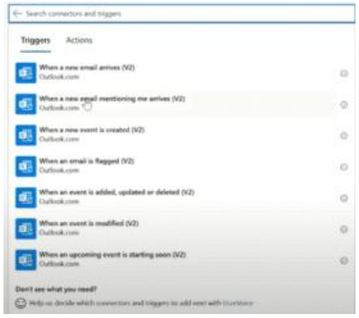
Select Blank Logic App



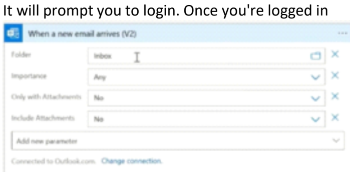
Select Outlook.com



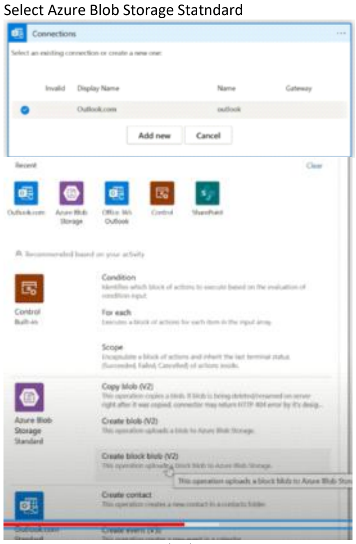
Select When a new email arrives (V2)



Folder: Inbox, Only with Attachments: Yes, Include Attachments: Yes
Click on Change connection
Add new (connection)
It will prompt you to login. Once you're logged in

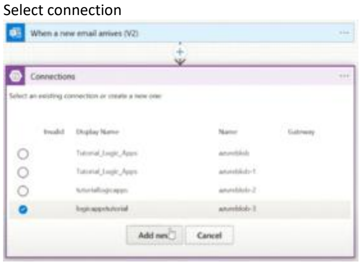


Click on Add new (to add a new connection)
Select Azure Blob Storage Statndard

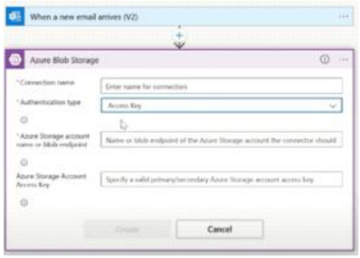


select Create block blob (V2)
Provide all storage account details and Access keys

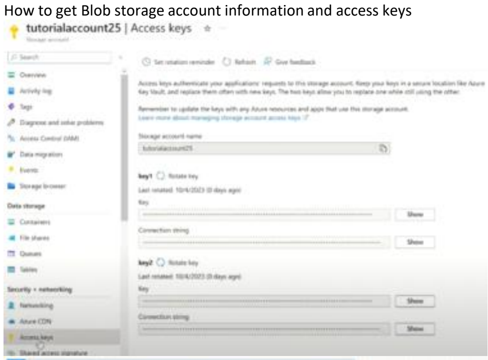
Select connection

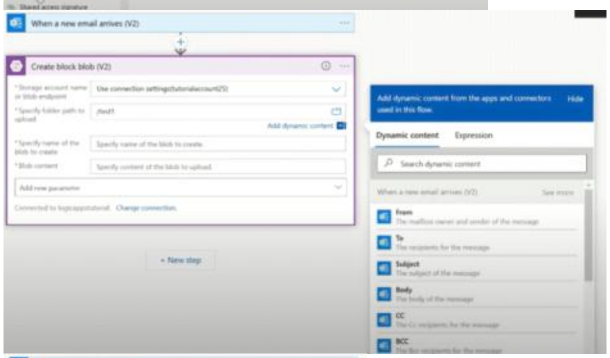


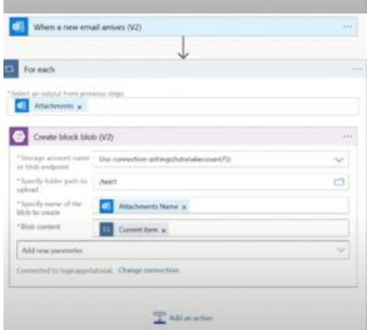
Provide Azure Blob Storage specifics when a new email arrives

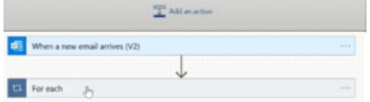


How to get Blob storage account information and access keys









Click on Run Trigger and it will start the attachment extraction process
Validate it.

Sample 2: https://www.mssqltips.com/sqlservertip/6492/azure-logic-app-to-extract-and-save-email-attachments/
Sample 3: https://learn.microsoft.com/en-us/answers/questions/1343849/logic-app-get-attachment-from-email-and-put-it-on

Using ADF pipeline: Trigger the Execution of an ADF Pipeline from an Email

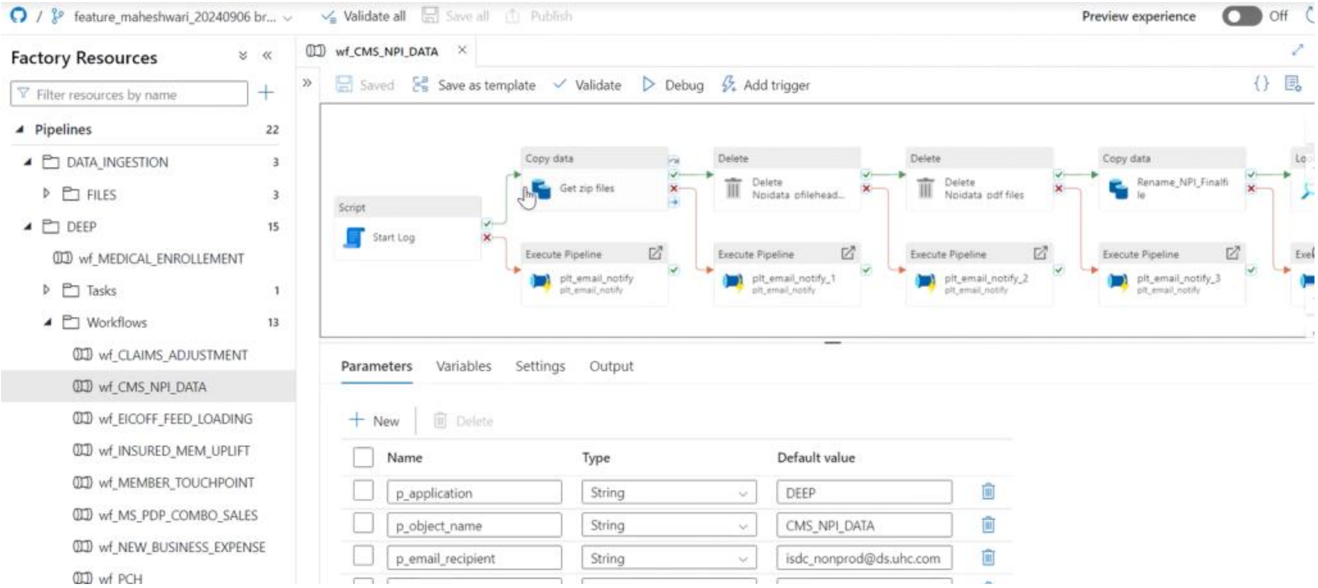
++

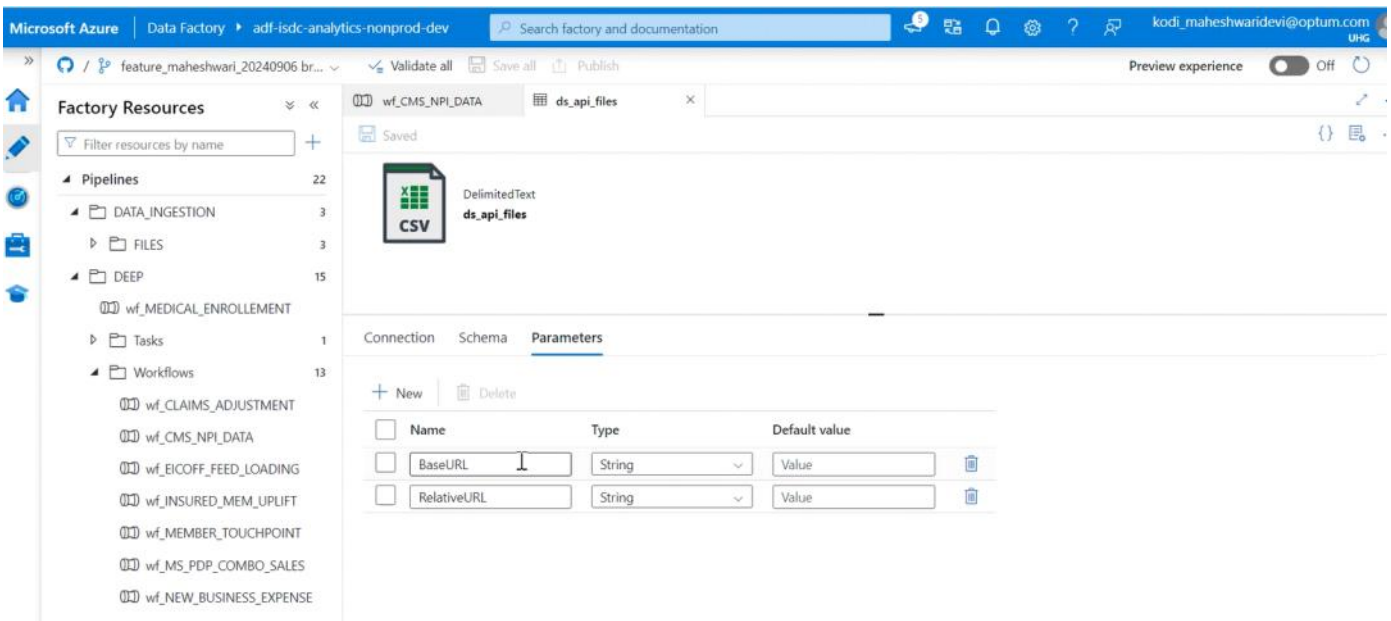
Both Azure Function Apps and Azure Logic Apps can be used to read Outlook attachments. Here's a high-level overview of how you can achieve this using each approach:

1. Azure Function Apps:
   - Create an Azure Function App and choose the appropriate trigger, such as an HTTP trigger or a Timer trigger.
   - Use the Microsoft Graph API or Outlook REST API within your function to authenticate and connect to the Outlook mailbox.
   - Retrieve the required email with the desired attachments using the API.
   - Extract and process the attachments as per your requirements, such as saving them to Azure Blob Storage or performing further actions.

2. Azure Logic Apps:
   - Create an Azure Logic App and choose the appropriate trigger, such as the "When a new email arrives" trigger from the Outlook connector.
   - Configure the necessary credentials and permissions to access the Outlook mailbox.
   - Once an email arrives, you can use the available actions and conditions to extract the attachments, such as the "Get attachments" action from the Outlook connector.
   - Perform any desired processing on the attachments, such as saving them to OneDrive or sending them to other systems via connectors.

++





++

Presentation:

- Our understanding
- Approach
  - Equal partnership
- Dependencies
- SMEs
- SME Availabilities
- Weekly status

++

[C&E@Work Lesson - Leading with Integrity [OLC1824520]]

++

Resource Name	Skillset	Permissions/ Roles
Preeti Khurana	Scrum Master	RALLY, Office Tools
Asit Sahoo	Azure Admin	RALLY, Azure, Snowflake, SDRP, CICD, GitHub, Terraform
Muquaddar Baig	Lead Cloud Data Engineer	RALLY, Azure, Snowflake, SDRP, current state

Mahesh Kumar	Could Data Engineer	RALLY, Azure, Snowflake, SDRP, current state
TBD	Cloud Data Engineer	RALLY, Azure, Snowflake, SDRP, current state

Azure:
Cloud account name: CorpSvcs_RAD
Account ID: 58a5fc58-629a-4f16-9e2d-7b94d2afef5d
Global Groups:
UHGRG_rg_rad_azurestack_08a9b07_ADMIN
UHGRG_rg_rad_azurestack_08a9b07_CONTRIBUTOR
UHGRG_rg_rad_azurestack_08a9b07_READ

Resource Group SDRP groups:
UHGRG_rg_rad_sdrp_snowflake_a1fbf6e_ADMIN
UHGRG_rg_rad_sdrp_snowflake_a1fbf6e_CONTRIBUTOR
UHGRG_rg_rad_sdrp_snowflake_a1fbf6e_READ

Azure:
Cloud account name: CorpSvcs_RAD
Account ID: 58a5fc58-629a-4f16-9e2d-7b94d2afef5d
Snowflake: https://uhg-uhgdwaas.snowflakecomputing.com/console/login#/

Request access to the below listed Permissions/Roles in SECURE

Resource Name	Skillset	Permissions/ Roles
Preeti Khurana	Scrum Master	Secure --> Request For: Application; Application: Rally, Environment: Prod , Office/Office 365 Tools , GitHub
Asit Sahoo	Azure Admin	Secure --> Request For: Application; Application: Rally, Environment: Prod , Global Groups: UHGRG_rg_rad_azurestack_08a9b07_ADMIN , SDRP Groups: UHGRG_rg_rad_sdrp_snowflake_a1fbf6e_ADMIN , CI/CD , GitHub , Terraform
Muquaddar Baig	Lead Cloud Data Engineer	Secure --> Request For: Application; Application: Rally, Environment: Prod , Global Groups: UHGRG_rg_rad_azurestack_08a9b07_ADMIN , SDRP Groups: UHGRG_rg_rad_sdrp_snowflake_a1fbf6e_ADMIN , UHGRG_rg_rad_sdrp_snowflake_a1fbf6e_CONTRIBUTOR GitHub current state credentials: to be provided
Mahesh Kumar	Could Data Engineer	Secure --> Request For: Application; Application: Rally, Environment: Prod , Global Groups: UHGRG_rg_rad_azurestack_08a9b07_CONTRIBUTOR , UHGRG_rg_rad_sdrp_snowflake_a1fbf6e_CONTRIBUTOR GitHub current state credentials: to be provided
TBD	Cloud Data Engineer	Secure --> Request For: Application; Application: Rally, Environment: Prod , Global Groups: UHGRG_rg_rad_azurestack_08a9b07_CONTRIBUTOR , UHGRG_rg_rad_sdrp_snowflake_a1fbf6e_CONTRIBUTOR GitHub current state credentials: to be provided

++
Concern:
EIS review and approval is required for any new component added to the cloud solutions, even for non -prod environment
To get an EIS review you need to submit solution Architecture diagram, network diagram and Business Unit architect's approval
So far we have created solution Architecture diagram, network diagram and have received Business Unit architect's approval
Now we're going to submit the EIS approval before we can stand up any infrastructure for the RAD Data Migration project
EIS approval takes time.

In-Progress:
We're taking KT from RAD SMEs
So far we have covered two data source types
API Data Source
Current state template - created and in approval stage
Future state template - created and in approval stage
Oracle/SQL data source type
Document current state
Create future state
Still to do:
Receive KT for:
Outlook Email
Document current state
Create future state
Internal Data source
Document current state
Create future state
External data source
Document current state
Create future state

Pilot project
Develop
Validate
Create Estimates for
Development
Test
Deploy

We have Muquaddar, Mahesh and Suman cloud data engineers
++
Key Outcomes:
- Define cloud solution patterns identifying the cloud tech stack and appropriate usage of each tool - done
- Build out cloud infrastructure to support pilot prototype development - in-progress
- Build pilot prototypes for cloud solution patterns based on tech stack definition - in-progress
- Detailed solution design for data acquisition & data integration to populate Broze, Silver, and Gold data into RAD Data Platf orm for all data sources
- Detailed estimation of effort for Dev, Test, and Deployment phase of program

++

Low Level Design Guideline:

1. Introduction

1.1 Scope of the document

- *This section will cover details regarding scope of the document*
- *Low level design document will be at component level i.e. for Website portal there will be One LLD*

Sample Content:

This document outlines the Low level design of tax payer registration functionality. It highlights/refers the Use cases in registration process and low level design details of components along with the rationale for the same. This document is reference design document for developers/implementers so that they can develop system with minimal effort

1.2 Intended Audience

- *This section will cover categories of audiences who will be referring/reviewing this document*

Sample Content:

- *Web portal Development Team*
- *Architecture Team*
- *Quality Team*
- *IRD Department*

1.3 System overview

- *This section will capture overview of system application i.e. for what system is being developed*
- *Who are the stake holders of system*
- *What are other external systems through which this will be interacting*

2. Low Level System Design

- *Low level design/implementation approach will be captured*

2.1 Sequence Diagram

- *This section will capture flow of information between all application layers required t*
  meet the requirement
- *Sequence diagram is input for low level design of components*

Sample Sequence Diagram:

2.2 Navigation Flow/UI Implementation

- *This section will capture the navigation flow i.e. How different screens/pages are being displayed depending upon different conditions involved in implementation*
- *This will provide visualization to developer/implementer for integration of different views/pages In the flow*
- *This section will also include/reference the screen layouts/wireframes*

2.2.1 Screen Validations, Defaults and Attributes

This section will capture the validations, default values of fields which has not been covered in SRS

2.3 Client Side Validation Implementation

- *This section will capture the implementation approach for handling client side validations*
- *Some code snippets for implementation of client side validation i.e.JS frameworks based implementation can be captured In this section*

2.4 Server Side Validation Implementation

- *Server side validations should also be there in any web portal apart from client side validation as attacker can disable java script in the browser*
- *This section will captured the implementation approach for server side validations*

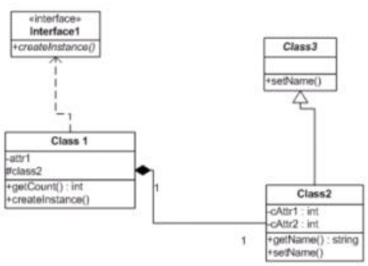
2.5 Components Design Implementation

- *This section will capture low level design of the components(in different layers) involved in sequence diagram*
- *Here components are being referred for set of independent implementations which can be integrated to achieve complete functionality. For example: <> Web portal can have user authentication component, digital signing component, component for caching data, email/SMS messaging client component, components for validation of user data from CBDT database, MCA DB and UAID DB*

2.5.1 Component 1

A. Class Diagram

- *This will list down all the classes involved for implementing the component*
- *This also depicts relationship of different classes*
- *More details on classes for particular component can be found in java document*



B. Pseudo Code

This section will capture the pseudo code for algorithms (if any) involved in implementation of component.

C. Other Implementation Details

This section will capture other implementation details which has not been covered under above mentioned heads

2.5.2 Component 2

2.6 Configurations/Settings

This section will capture the details of different settings/configurations being used in system application

Sample Content:

All the configuration of different parameters will be stored in properties file in the form of key value pair. This property file will be loaded to server during start-up

2.7 Interfaces to other components

This section will capture the design details of interfaces to other components

3. Data Design

3.1 List of Key Schemas/Tables in database

This section will capture the details for list of key schemas/tables through which module in scope is interacting.

3.2 Details of access levels on key tables in scope

This section will capture the access level i.e. Read, insert, update etc. details of module on key tables/schemas

3.3 Key design considerations in data design

This section will capture the key design considerations taken while designing the data model/selecting the data access mechanism etc.

4. Testing

- *This section will capture the details for list of test cases being used for testing particular component*
- *Reference of unit test case document can be given in this section.*

++

Hcc-intake: hcc-intake/ISDC/ https://github.com/optum-eeps/hcc-intake/tree/main/ISDC

ASK ID: AIDE_0083635

Azure:
Cloud account name: CorpSvcs_RAD

Account ID: 58a5fc58-629a-4f16-9e2d-7b94d2afef5d
Global Groups:
UHGRG_rg_rad_azurestack_08a9b07_ADMIN
UHGRG_rg_rad_azurestack_08a9b07_CONTRIBUTOR
UHGRG_rg_rad_azurestack_08a9b07_READ

<u>Resource Group SDRP groups:</u>
UHGRG_rg_rad_sdrp_snowflake_a1fbf6e_ADMIN
UHGRG_rg_rad_sdrp_snowflake_a1fbf6e_CONTRIBUTOR
UHGRG_rg_rad_sdrp_snowflake_a1fbf6e_READ

++
Tenant: RAD
prefix: AWKUR
Environment: Dev & Prod
Snowflake Roles:
Developer - AZU_SDRP_AKWUR_PRD_DEVELOPER_ROLE & AZU_SDRP_AKWUR_DEV_DEVELOPER_ROLE
Support - AZU_SDRP_AKWUR_PRD_SUPPORT_ROLE & AZU_SDRP_AKWUR_DEV_SUPPORT_ROLE
Analyst - AZU_SDRP_AKWUR_PRD_ANALYST_ROLE & AZU_SDRP_AKWUR_DEV_ANALYST_ROLE
Secure Global Groups (roles)
Prod/Dev:
RAD Storage Tenant > Tenant Analyst (AKWUR); Tenant Developer (AKWUR); Tenant Support (AKWUR)

++
Prod/Dev:
RAD Storage Tenant > Tenant Analyst (AKWUR); Tentant Developer (AKWUR); Tenant Support (AKWUR)
++

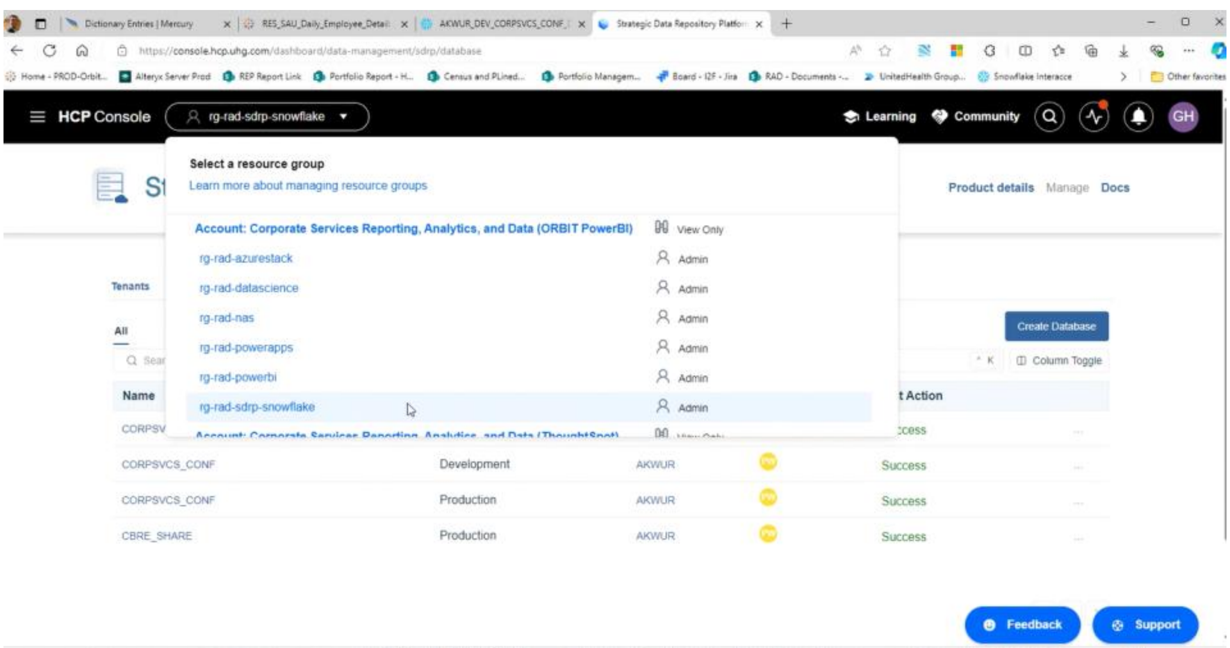
Tenant: RAD
prefix: AWKUR
Environment: Dev & Prod
Snowflake Roles:
Developer - AZU_SDRP_AKWUR_PRD_DEVELOPER_ROLE & AZU_SDRP_AKWUR_DEV_DEVELOPER_ROLE
Support - AZU_SDRP_AKWUR_PRD_SUPPORT_ROLE & AZU_SDRP_AKWUR_DEV_SUPPORT_ROLE
Analyst - AZU_SDRP_AKWUR_PRD_ANALYST_ROLE & AZU_SDRP_AKWUR_DEV_ANALYST_ROLE
Secure Global Groups (roles)

Prod/Dev:
RAD Storage Tenant > Tenant Analyst (AKWUR); Tenant Developer (AKWUR); Tenant Support (AKWUR)

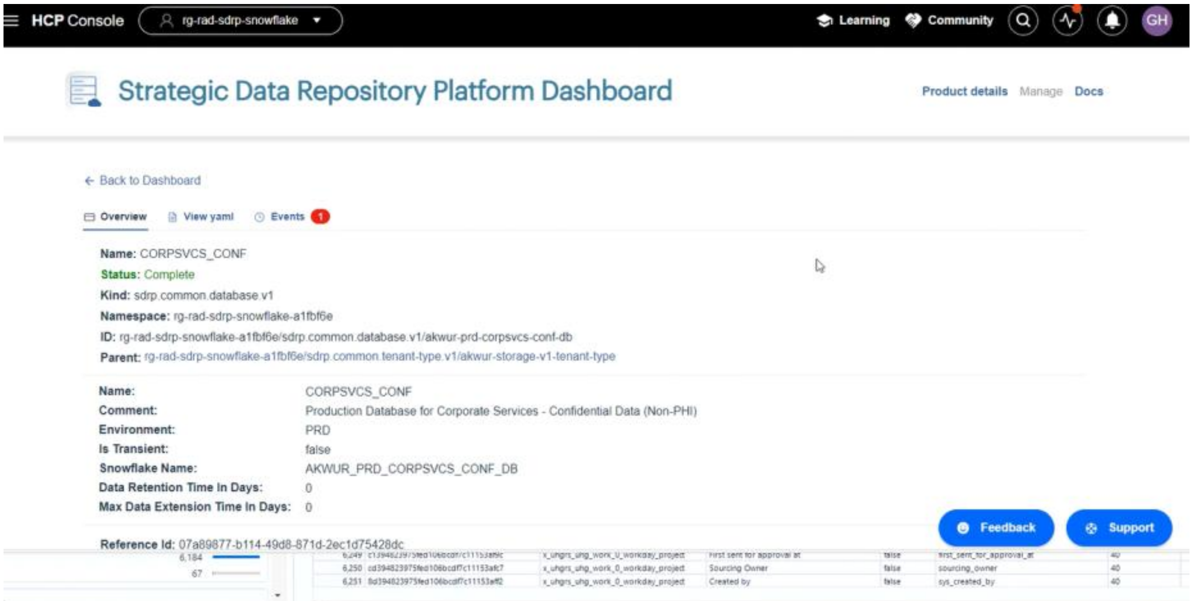
SECURE ROLE for RAD Data Modernization:

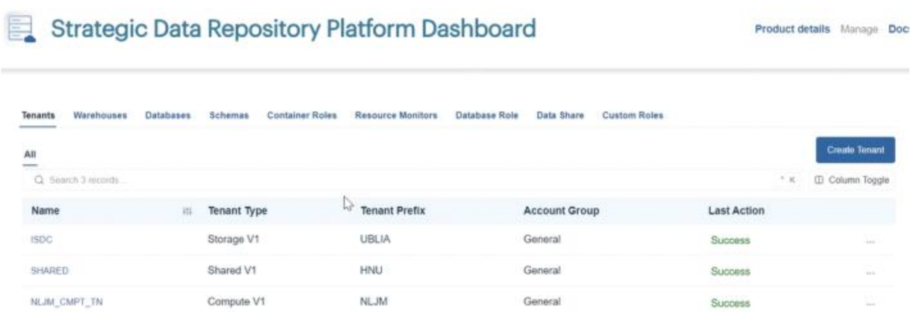
- Global Groups: **UHGRG_rg_rad_azurestack_08a9b07_ADMIN**
- Resource Group SDRP groups: **UHGRG_rg_rad_sdrp_snowflake_a1fbf6e_CONTRIBUTOR**
- Snowflake Roles: Developer - **AZU_SDRP_AKWUR_PRD_DEVELOPER_ROLE** & **AZU_SDRP_AKWUR_DEV_DEVELOPER_ROLE**

++
For more information, please contact: **Patrick Whitmer or Grant D Haugan**



++





++

ETL_BATCH_OBJECT_CONTROL_ANALYTICS & ETL_BATCH_OBJECT_CONTROL_LOG_ANALYTICS

wf_ACCOUNTABLE_CALLS
UTIL.PROGRAM_LIST_ANALYTICS PGM
UTIL.PROCESS_LIST_ANALYTICS
UTIL.PROCESS_RUN_LOGS_ANALYTICS

ETL_BATCH_OBJECT_CONTROL_ANALYTICS & ETL_BATCH_OBJECT_CONTROL_LOG_ANALYTICS
++
Start the Process: kick off the process using start_time parameter
Start Log - read acess creadentials and insert a row in <schema>.PROCESS_RUN_LOGS_ANALYTICS table
Get_Object_Metadata: get job run metdata info. from UTIL.PROGRAM_LIST_ANALYTICS and UTIL.PROCESS_LIST_ANALYTICS for active process and based on the pipeline name
If Start Log does not start correctly, send notificaiton email

Get metadata details using stored procedure SP_ACCOUNTABLE_CALLS and trigger pipeline by provideing pipeline info.

On success, end process and send a successful run email, otherwise send error email notification
++
rg-rad-sdrp-snowflake

rg-rad-azurestack
++
Total: 77
Rest API count: 34
- CSV count: 30
- PeoplePlatform: 1
- OT Productivity: 1
- OT Department Breakdown: 1
- Reports: 1

Flat File count: 36
- ECG count:
- Outlook email attachments
- XML count:

Integration count: 1
SharePoint List: 1
SQL/Oracle count: 5

CIDR Block setup
add specific user, assign CIDR block

[HCP Data Platform User Group | Strategic Data Repository Platform | Microsoft Teams]()
++
Flat file (XLSX/CSV) - 37
- Adobe Sign - 1
- Ariba - 2
- CBRE - 3

eGRC - 4
ESG - 20
PeoplePlatform
- 3
Hubstar - 1
Mercury - 1
Noosh - 1
EOHS (SharePoint)- 1
REST API (CSV) - 34
XML (5 suppliers) - 5
SQL/Oracle - 4
Concur Travel & Expense
- 1
EOHS Cority - 1
Lenel Badge Swipe data pull
- 1
One-Click E&P Mercury Surveys
- 1
++
Unique file types: XLSX/CSV, XML, SQL,
++
Each Data Source Type

- Create & populate medallion architecture tables
  - Bronze Layer
    - Create table
    - Read data from Integration storage
    - Populate data into table
  - Silver Layer
    - Create table
    - Extract and transform bronze layer data
    - Populate data into silver layer table
  - Gold Layer
    - Create table
    - Extract and transform silver layer data
    - Populate data into gold layer table
  - Data Product Creation
    - Work with RAD SMEs to design Data Products
    - Create Data Products
    - Work with RAD SMEs and Tier 1 Governance team to seek approvals
  - Data Product Consumption
    - Work with RAD SMEs and business stakeholders to subscribe Data Products for consumption
    - Work with RAD SMEs and Tier 1 Governance team to secure subscription approvals
    - Validate end-to-end data pipeline
- Repeat the same process for all data source types

API Data Source

- Bronze Layer
  - Create table
  - Read data from Storage
  - Populate table
- Silver Layer
  - Create table
  - Read data from bronze layer
  - Transform & populate silver table
- Gold Layer
  - Create table
  - Read data from silver layer
  - Transform & populate gold table
- Data Product Creation with RAD SMEs
  - Design Data Products
  - Create Data Products
  - Secure approvals
- Data Product Consumption with End Consumers
  - Subscribe Data Products
  - Secure subscription approvals
  - Validate end-to-end data pipeline

Repeat the same process for all data source types
++
https://uhgazure.sharepoint.com/:x:/r/teams/OAS-MRISAssessment/_layouts/15/Doc2.aspx?action=e...

https://uhgazure.sharepoint.com/:x:/r/teams/OAS-MRISAssessment/_layouts/15/Doc2.aspx?action=e...
EDP Generic Estimation
++
File sources:
Batch Files

- Ariba
- CBRE
- eGRC
- ESG
- HR/Peoplesoft
- Hubstar
- IT Reseller (XML)
- Mercury
- Noosh
- System
- OLC Groups (Sharepoint list)

API data source

- Field Glass
- PeoplePlatform
- Mercury
- OT Datahub

SQL data source

- SQL/Concur
- EOHS/Cority
- Lenel/Badge Swipe

One-click/E&P Mercury Survey
++
Copying should be different, e.g., different endpoints, but dropping into the Blob Storage should be the same/similar
+++
Source file types:
CSV/XLSX
XML
SQL
SharePoint list
++
Components
WebAPI
ADF
Blob Storage
Datasets
VNET
SHIR
scripts/configurations
Snowflake Stored procedures
Snowflake Platform
Snowflake ETL
Snowflake Warehouses
Pipelines

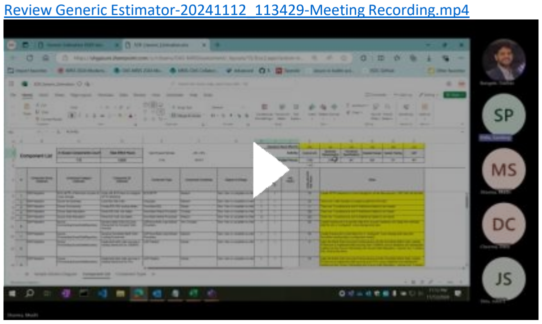
WebAPI
Blob Storage
ADF
Datasets
Linked services
Private Links
SHIR
VNET
Data Pipelines

Snowflake
Tables
Stored procedure
ETL via stored procedures
Virtual Warehouses

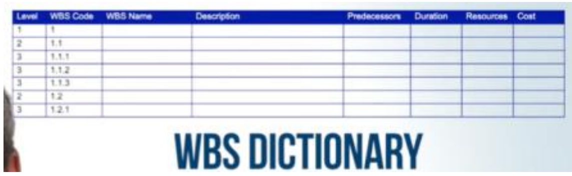
Access & Policy controls
Azure monitoring and alerting services
Azure Active Directory
GitHub
Policy based archival
Move inactive data to cold storage
Buy services in advance to get significant savings
Current data size 800 GB - 1 TB
Expected growth 20%

Amount of data transfer
Initial ~1 TB
Daily: 10 GB
Weekly: 5 GB
Monthly: 5 GB

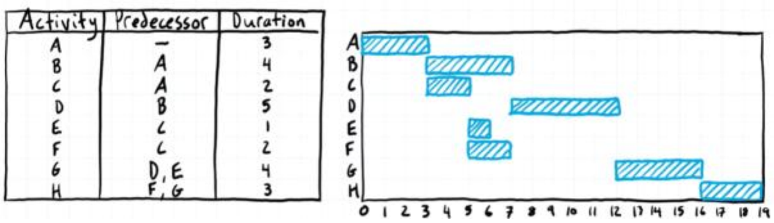
+++
Infrastructure components
WebAPI
Service Account
ADF
SHIR
VNETS/Subnets
Private Links
Approvals
EIS
GitHub
User Groups
Data Products
Creation
Design
Creation
Approval
Subscription
Search & subscribe
Approval
Validation
[Review Generic Estimator-20241112_113429-Meeting Recording.mp4]

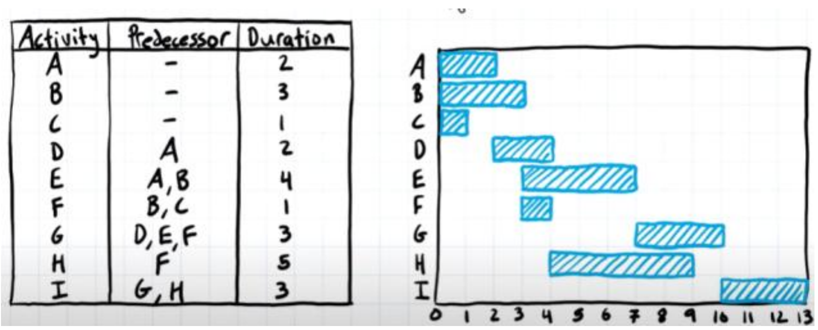


++
Work Breakdown Structure (WBS)

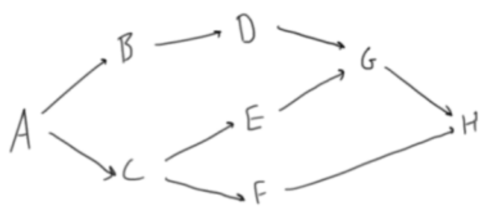


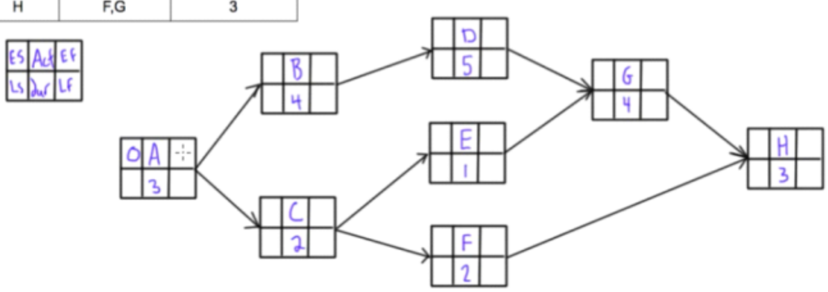
++

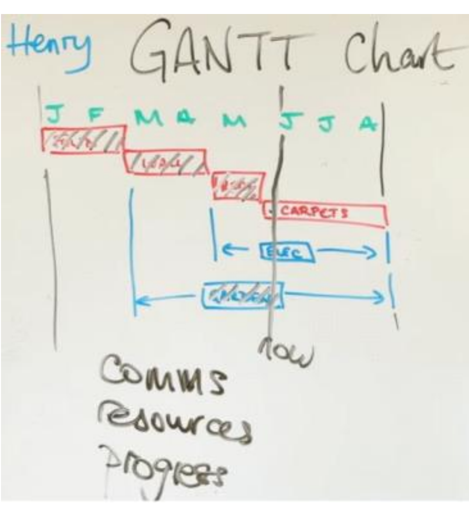




Activity	Predecessor	Duration (days)
A	-	3
B	A	4
C	A	2
D	B	5
E	C	1
F	C	2
G	D,E	4
H	F,G	3







++
https://docs.hcp.uhg.com/strategic-data-repository-platform-(sdrp)/storage-tenant

intended audience of that data is. To know more about permissions that will be granted to the functional roles depending on the below schema type, click here.

Schema Type	Analyst	Developer	Support	Use Case
Raw	No access	Read/write	Read	Original data, unprocessed and usually loaded from a source system
Target	No access	Read/write	Read	Refined datasets built from raw data; intermediate data used in later pipelines
Access	Read	Read/write	Read	Final (gold) datasets intended for consumer use

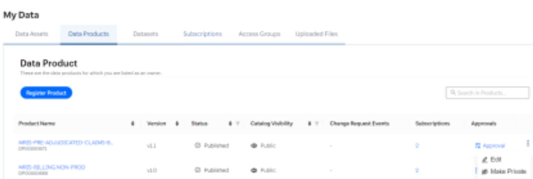
To create a schema, see the Schema section in the HCP Console guide.

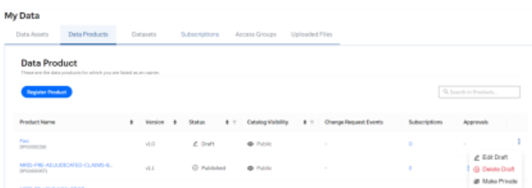
++

- Cloud solution patterns
- solution design
- Detailed data acquisition & data integration to populate Broze, Silver, and Gold data into RAD Data Platform for all data sources
- Detailed estimation

- Build out cloud infrastructure to support pilot prototype development
- Build pilot prototypes for cloud solution patterns based on tech stack definition
- ADF Workflow Management Framework

6. How are we going to migrate API data source Historical data?





++

- Define cloud solution patterns identifying the cloud tech stack and appropriate usage of each tool

- Build out cloud infrastructure to support pilot prototype development
- Build pilot prototypes for cloud solution patterns based on tech stack definition - Give them a demo

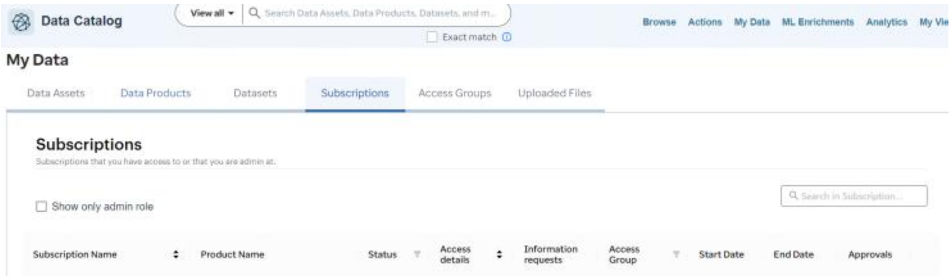
- Detailed solution design for data acquisition & data integration to populate Broze, Silver, and Gold data into RAD Data Platform for all data sources
- Detailed estimation of effort for Dev, Test, and Deployment phase of program

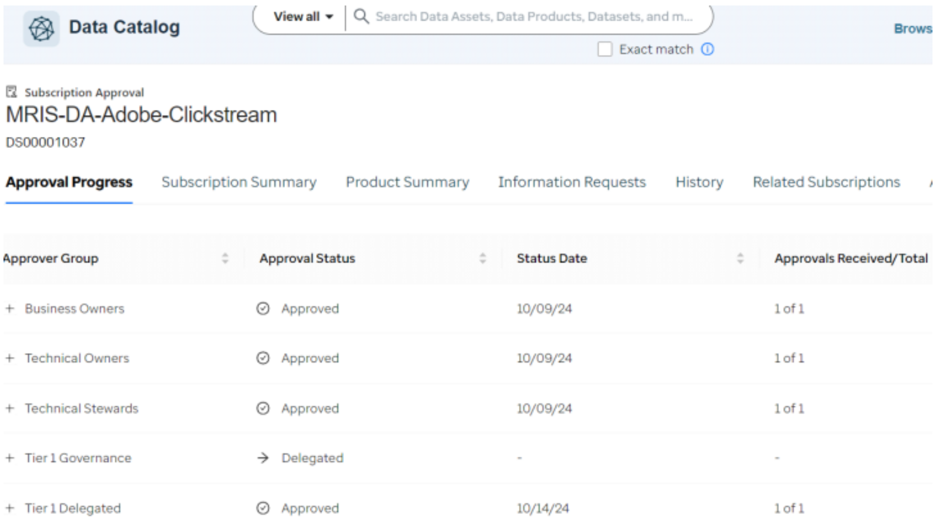
Deliverables
- ADF pipeline framework
- Overall architecture
- IaC and CICD automations
- Solution Designs
  - API Data Source
    - Architecture
    - HLD
    - LLD
  - Batch Data Source
    - Architecture
    - HLD
    - LLD
  - SQL Data Source
    - Architecture
    - HLD
    - LLD

Corporate Services Reporting Analytics and Data (ORBIT PowerBI)
https://aide.optum.com/applications/browse/AIDE_0083635
Data Warehousing and Reporting Environment for Corporate Services - Used for EOHS, Enterprise Procurement, and Real Estate Service in support of data, reporting, and analytics.

From <https://ask.optum.com/search?q=AIDE_0083635#aideapplications>

++


Data Catalog
View all
Exact match
Browse
Actions
My Data
ML Enrichments
Analytics
My Data
Data Assets
Data Products
Datasets
Subscriptions
Access Groups
Uploaded Files
Subscriptions
Show only admin role
Subscription Name
Product Name
Status
Access details
Information requests
Access Group
Start Date
End Date
Approvals


Data Catalog
View all
Search Data Assets, Data Products, Datasets, and m...
Exact match
Subscription Approval
MRIS-DA-Adobe-Clickstream
DS00001037
Approval Progress
Subscription Summary
Product Summary
Information Requests
History
Related Subscriptions
Approver Group
Approval Status
Status Date
Approvals Received/Total
Business Owners
Approved
10/09/24
1 of 1
Technical Owners
Approved
10/09/24
1 of 1
Technical Stewards
Approved
10/09/24
1 of 1
Tier 1 Governance
Delegated
-
-
Tier 1 Delegated
Approved
10/14/24
1 of 1

Data Catalog

View all ▾ Search Data Assets, Data Products, Datasets, and m...

Exact match

Browse Actions My Data ML Enrichments

Subscription

MRIS-DA-Adobe-Clickstream

DS00001037

Summary Approvals **Access Details** Access Group

Platform instance general details

This Subscription grants access to a Data Product hosted within the **Uhg Dwaas Snowflake** Platform.
To start using your Subscription:

- Click the login link to navigate to the Snowflake UI in your browser
- Click the "Sign in using AzureAD" button if prompted. This will sign you in via Optum SSO. Do NOT login with manual username/password.
- In the top-right corner, click your name and select the Role named after your HCP Resource Group (see Credentials section below).
- You are now able to start executing queries against any Uhg Dwaas Snowflake data products your HCP Resource Group is subscribed to!
- For full documentation and advanced connection options, please visit the **Product Site**

Credentials

Provisioned Service Accounts:

- isdc_dev_rpt@optum.com
- isdc_prd_rpt@optum.com

Provisioned User Account:

- su_wang@uhc.com
- edward.noreika@uhc.com
- christopher_b_smith@uhc.com
- gurtej_singh@optum.com
- sonal_rana@uhc.com
- omar_deleon@uhc.com
- sai_garikipati@uhc.com
- sumanjali_mutyala@uhc.com
- mujeeb_mohammed@uhc.com
- mujeeb_g_mohammed@optum.com
- dave.cheema@optum.com
- sandeep.palla@optum.com
- veeranjaneyulu_nagulapati@optum.com
- kanchana_jagannathan@uhc.com
- mary_herries@optum.com
- meghan_gardy@optum.com
- jared_a_crowther@uhc.com
- jean.rao@uhc.com
- brian_gu@uhc.com
- samuel.neuschwander@uhc.com
- trevor_jones@uhc.com
- kevin_ahtou@uhc.com
- edward_p_kost@uhc.com
- kevin_k_ching@uhc.com

- Snowflake Role : HCDP_MRISDA_DATA_VIEWER_PRD_DAG_ROLE

Connection information

UHG DWaaS Login

- **Account Identifier:** uhg-uhgdwaas
- **Browser Login:** https://uhg-uhgdwaas.snowflakecomputing.com/console
- **Connector Path:** uhg_uhgdwaas.east-us-2.azure

Status

Active

Access Time

09/30/2024 - 09/30/2025

Application ID

AIDE_0085977

Resource Group

MRISDAISDCProduct1

Data Access Group

mrisda

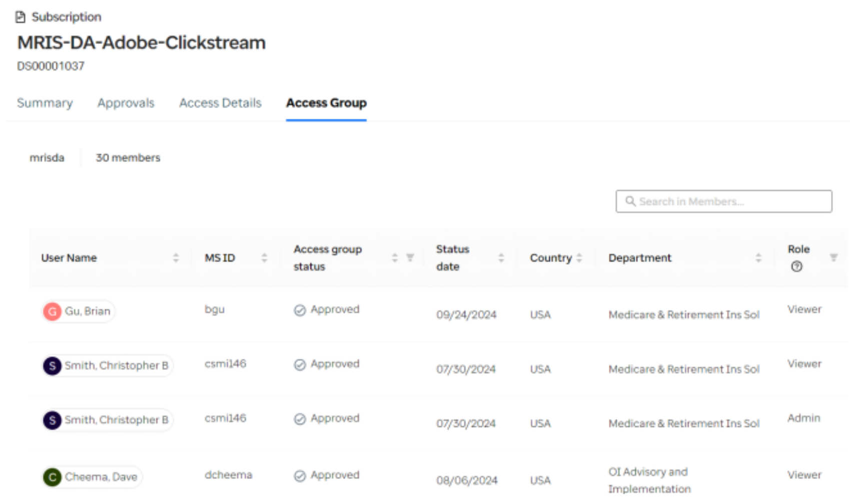
Admins

S Smith, Christopher B

Business Sponsor

S Singh, Gurtej

Feedback


Subscription

MRIS-DA-Adobe-Clickstream

DS00001037

Summary | Approvals | Access Details | Access Group

mrisda | 30 members

Search in Members...

User Name	MS ID	Access group status	Status date	Country	Department	Role
Gu, Brian	bgu	Approved	09/24/2024	USA	Medicare & Retirement Ins Sol	Viewer
Smith, Christopher B	csmi146	Approved	07/30/2024	USA	Medicare & Retirement Ins Sol	Viewer
Smith, Christopher B	csmi146	Approved	07/30/2024	USA	Medicare & Retirement Ins Sol	Admin
Cheema, Dave	dcheema	Approved	08/06/2024	USA	OI Advisory and Implementation	Viewer

++

Durga,

Thank you so much for meeting with me last Friday (11/22/2024) and discussing some challenges you're facing in attempting to onboard Jazz. It would be great if the potential risks/vulnerabilities could be identified, then we might be able to find mitigative approaches and solutions. I'll be able to provide all necessary guidance, mentoring, and needed support to see him succeed.

I greatly appreciate you trying to explore different opportunities to bring him onboard.

A few noteworthy point I'd like to make:

1. He has incredible will to succeed
2. Anytime there is something new to learn, he puts an extraordinary amount effort in it
3. He does not get fixated on just one solution, he explores all possible options and then settles on the optimal one

Somebody must have given us the opportunity to get started. Wouldn't it be great if he got his at Optum?

++Data Science Platform

1. Immediate goal: Provide you a platform to run your data science models and an external Dataset(s) to provide data;
   a. Setup Azure Databricks instance
   b. Create Azure Blob Storage
   c. Upload external dataset to Azure Blob Storage
   d. Upload Jupyter notebook
   e. Enable Databricks instance connectivity and access permissions with dataset in the Blob Storage
   f. Setup GitHub and Notebook integration
   g. Create output folder
   h. Storage Machine Learning run results in the output folder

2. Long-term goal: Provide a platform that is more robust with appropriate controls, governance, and automations. Proper data models will be designed, created, and populated with data.
   a. Determine the best platform, Choices: United AI Studio (UAIS), Azure Machine Learning Studio, Databricks

Capability	**Azure Machine Learning studio**	**Databricks**	**UAIS**
Functionality	• Build and train Azure Machine Learning model with any type of compute including Spark and GPUs for cloud-scale large AI workloads. • Run automated Azure Machine Learning (AutoML) and drag-and-drop UI for low-code Azure Machine Learning. • Implement end-to-end Azure Machine LearningOps and repeatable Azure Machine Learning pipelines. • AI dashboard for bias detection and error analysis. • Orchestrate and manage engineering and LLM flows. • Deploy models with REST API, real-time, and batch inference	• Scalable Spark-based processing for large-scale data processing and model training, handling massive datasets. • MLflow integration for model tracking, versioning, and deployment with Mlflow • Centralized repository to manage machine learning artifacts. • AutoML functionality to explore different algorithms and hyperparameters to find the best model for your problem. • Collaborative notebooks supports multiple languages and teams can work together • Databricks Runtime for Machine Learning is Preconfigured clusters optimized for machine learning • Git integration for version control and collaboration on code within the Azure Databricks workspace. • Job scheduling to schedule and automate machine learning workflows, including data preparation, training, and model deployment.	Build, deploy, manage and scale machine learning (ML) models. Easy and secure access to data needed to build the models and run analytics. Streamlines the assessment of responsible use of artificial intelligence (AI) metrics, allows deployment and sharing of approved models across UnitedHealth Group (UHG). It keeps the steps of model development contained in one place, which makes calculating costs more convenient and helps ensure accountability.

EIS approval completed

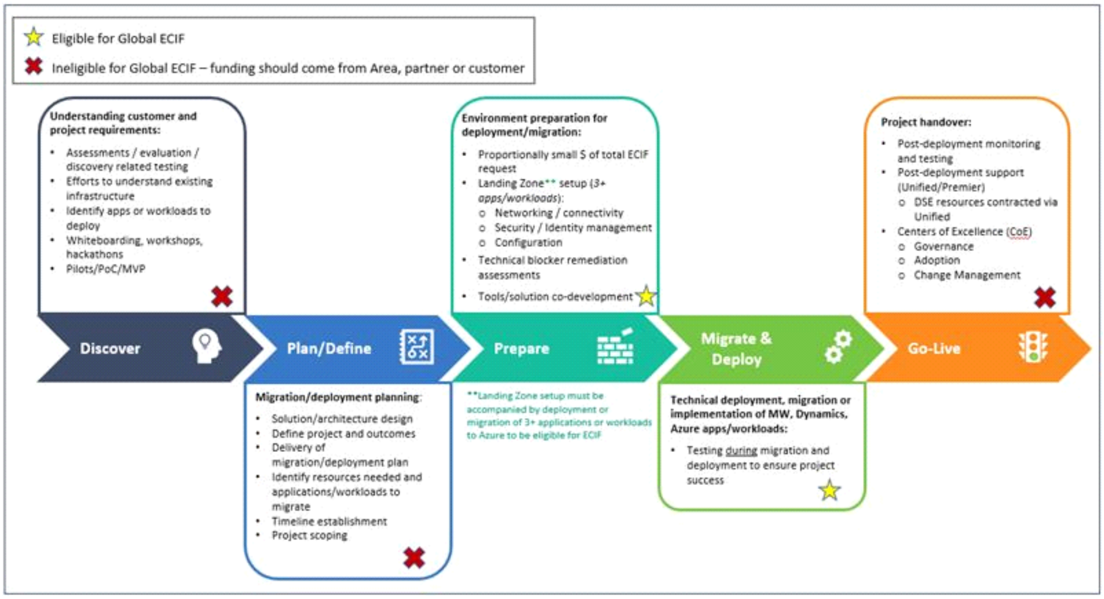
Pilot data product created and received approvals

++

Shadow Training Steps for Azure Data Factory:

- RAD team to familiarize with basic cloud data platforms and cloud data engineering concepts
- OA team will point them to official documentation sites
- OA team provides overview of data migrations, transformations, and integrations
- RAD team to shadow OA team during data migration frameworks, transformations, and data movements
- Allow them to create simple tasks such as, table/view creation, store procedure creations, etc.
- Start them with simple data pipelines to learn how compose a data pipeline
- Show them how to monitor and troubleshoot pipelines and activities
- Encourage them to ask questions

++



++

Disparate systems data in multiple data sources	Migrate all data assets into a single cloud data platform (Snowflake)
Data acquisition and integration challenges	Leverage cloud built-in capabilities such as, Azure Data Factory, WebAPI, and low-code Activities to enable data pipelines
Data Acquisition and integrations are time consuming	Cloud based flexible infrastructure enables to handle data loads of any size
No standardized data architecture	RAD data hosted in Azure, Snowflake platform using Medallion architecture, which organizes data into layers for improved quality and reliability
Lack of cohesive data governance and control strategy	Healthcare Platform (HCP) provides out-of-box standardized data governance with appropriate roles and permissions

Deliverable	Purpose	Link	Warranty Support Contact
Define cloud solution patterns identifying the cloud tech stack and appropriate usage of each tool	Solution patterns based on which the RAD Data Modernization solutions will be built	API Data Source Flat File Data Source SQL Data Source	Dave Cheema
Build out cloud infrastructure to support pilot prototype development	A clod-based set of components & services to run RAD Data Modernization solutions	Platform Infrastructure Related Documents	Dave Cheema
Build pilot prototypes for cloud solution patterns based on tech stack definition	Initial project to validate approach, patterns and solution	Pilot Project Related Documents	Dave Cheema
Detailed solution design for data acquisition & data integration to populate Broze, Silver, and Gold data into RAD Data Platform for all data sources	Detailed design patterns to be used to build RAD Data Modernization solutions	API Data Source - Detailed Design Document v0.0.docx Flat File Data Source - Detailed Design v0.0.docx SQL Data Source - Detailed Design Document v0.0.docx	Dave Cheema
ADF Workflow Management Framework	ADF Workflow Management Framework to Control Execution of Future Data Pipelines	ADF Workflow Management Framework	Dave Cheema
Detailed estimation of effort for Dev, Test, and Deployment phase of program		To be presented and provided separately	Dave Cheema

+++

This is a new request from Scott for a Alteryx Developer versus the request for SAS Develop that was submitted Monday.

Alteryx Developer Requirements:

Thank you for submitting your entry. A copy is included below for your records.

Optum Advisory Resource Request Form

Request type	Staffing/Contractor Request
Requestor	Vicky Tee
Organization	Digital Transformation
OA Hiring Manager	John Shin
Location	US
Location Justification	Offshore Data Access restriction
Work Location Type	Telecommute
Travel Requirements	No travel requirements
OI Revenue Alignment	Directly Billable
Project SalesForce ID	O679407
Project Stage	5 - Sold

Project Name	STARs UCEE Migration
Role Title	Alteryx Developer
GL	28
Quantity	1
Desired Start Date	01/01/2025
Duration	7 Months
Contractor Appetite	Yes
Contractor Cost Rate	90
Pre-identified Candidate?	No
Role Description	Description: The STARS Data Operations Program Orchestration - Manager Data Analytics will work with STARs HEIDIS measure Owners as a Senior Workflow Developer in M&R Workflow Programming Team utilizing advanced analytics. Must have solid programming skills to develop, implement, and maintain Medicare Part C/CAHPS/HOS/C&S Workflows, reporting of programs as required by STARS program orchestration team Primary Responsibilities: • Develop and maintain complex workflows using IBM Unica software. • Collaborates with business partners to design, build and code cross channel communications using campaign platforms such as U nica/IBM Campaign. • Serves as the technical lead on projects. Translates and conveys business solutions and strategies into system and database a rchitecture. • Designs and builds strategies, processes and technology solutions that support business needs and align with industry best pr actices. • Ability to Create and Translate Complex SQL Queries to support ad -hoc business requests. • Identifies data solutions to support complex campaigns including: SQL Server, Oracle, Teradata, flat files and data warehouse s. • Develops and maintains best practices and standard operating procedures for campaigns. • Serves as the primary liaison to technology partners for data requests, issue escalation and resolution. Troubleshooting tech nical issues and providing technical support to Business partners • Maintaining and updating the workflow automation environment to ensure optimization • Participating in quality assurance, software testing, and documentation efforts
Skill Requirements	Required Qualifications: • Bachelor's degree in engineering or equivalent work experience • 2+ years of Hands on development experience with Alteryx • 5+ years of Experience working with Campaign Management platforms (Adobe Campaign, Unica (preferred), SAS, Alterian, Eloqua, etc.) • Strong Experience with workflow design and development. • 4+ years of experience with SQL and RDBMS systems • 4 years of experience in a data quality review process, eg., Data QA, UAT etc. • 3+ Years of experience in Automation and optimization of workflows. • Proven experience with requirements gathering and analysis techniques • Proven ability to design, develop, deploy and maintain technical Workflows • Competent to work on complex projects independently • Ability to think strategically (outside-the-box) and provide creative solutions • Ability to frame problems and opportunities to present solutions.
Desired Skills	Preferred Qualifications: • Experience with or understanding of HEDIS Star measures • Experience with or understanding of the Medicare and Retirement and/or Community and State lines of business • Proficient in UNIX, Linux • Proficiency in Programming Languages (.Net/C++/C#/Java). • Proficiency in Scripting (Batch/Shell/Java Script). • Proficiency in Dashboard tools (Tableau/Cognos/Business Objects/Power BI). • Knowledge of ETL Application (SSIS/Alteryx/Informatica) Key CLL Skills • Understanding Goals, Priorities and Plans • Delivering Results • Resilience and Resourcefulness • Providing Support • Growth Mindset

++

- **SAS Workflow Development:**

1. Can you explain your experience in developing SAS workflows for senior-level projects?
2. How do you approach designing and building complex SAS workflows?
3. What are some challenges you have faced in developing SAS workflows, and how did you overcome them?
4. Can you describe a specific project where you had to optimize or improve an existing SAS workflow? What steps did you take, a nd what was the outcome?
5. How do you ensure the quality and accuracy of data within SAS workflows?
6. Have you worked with any SAS data integration tools or solutions? Can you provide examples of how you have utilized them in y our workflows?
7. How do you collaborate with other team members or stakeholders to gather requirements and ensure the successful implementatio n of SAS workflows?
8. Can you discuss your experience in troubleshooting and debugging issues within SAS workflows? What tools or techniques do you use?
9. Are you familiar with any industry best practices or standards for developing SAS workflows? How do you stay updated on the l atest trends and advancements in SAS technologies?
10. Can you give an example of a time when you had to prioritize multiple projects or tasks related to SAS workflow development? How did you manage your time and ensure timely delivery?

- **Workflow Design:**
  - How do you ensure the efficiency and maintainability of your workflows?
  - How do you document and communicate your workflow designs?
- **SAS Enterprise Guide:**
  - Describe your experience with SAS Enterprise Guide and its features.
  - How do you create and manage projects and folders in Enterprise Guide?
  - How do you use task flows and code tasks within a workflow?
  - How do you handle errors and exceptions within a workflow?
- **SAS DI Studio:**
  - Explain the components of a SAS DI Studio job flow.
  - How do you transform data using transformations in DI Studio?
  - How do you handle data quality issues in DI Studio?

- Describe your experience with metadata management in DI Studio.

- **Data Manipulation & Analysis:**
  - **Data Cleaning:**
    - How do you identify and handle outliers?
    - How do you deal with duplicate records?
    - How do you perform data imputation for missing values?
  - **Data Transformation:**
    - How do you create new variables based on existing ones?
    - How do you perform data aggregation and summarization?
    - How do you reshape data (e.g., pivot tables)?
  - **Data Analysis:**
    - How do you perform basic statistical analyses (e.g., t-tests, ANOVA)?
    - How do you create and interpret descriptive statistics?
    - How do you perform data visualization using SAS?
  - Problem-Solving:
    - How do you identify and resolve performance bottlenecks in your workflows?
    - How do you handle unexpected errors and exceptions?

  Communication & Teamwork:
  - How do you communicate technical information to both technical and non-technical audiences?
  - How do you contribute to and learn from others on your team?

    Experience & Projects
  - SAS Workflow Development:
  - SAS Enterprise Guide:
    - Explain the process of designing and executing a workflow.
    - How do you use task flows and code tasks within a workflow?
    - How do you handle errors and exceptions within a workflow?
  - Industry Experience:
    - Describe your experience working with specific industries (e.g., finance, healthcare, retail).
    - How has your SAS expertise been applied in these industries?
  - Data Manipulation & Analysis:
  - Data Cleaning:
    - How do you identify and handle outliers?
    - How do you deal with duplicate records?
    - How do you perform data imputation for missing values?

  Data Transformation:
    - How do you perform data aggregation and summarization?
    - How do you reshape data (e.g., pivot tables)?
  - Data Analysis:
    - How do you create and interpret descriptive statistics?
    - How do you perform data visualization using SAS?

  Problem-Solving & Design
  Behavioral Questions
  - Motivation & Interests:
    - Why are you interested in this position and company?
    - What motivates you to excel in your work?
    - How do you stay current with the latest developments in SAS and data analytics?
  - Work Ethic:
    - Describe your work style and how you approach deadlines.
    - How do you handle pressure and stressful situations?
    - How do you prioritize tasks and manage your time effectively?
  - Teamwork & Collaboration:
    - How do you build and maintain effective relationships with colleagues?
    - How do you contribute to a positive and productive team environment?
    - Describe a situation where you had to work effectively with a difficult colleague.

++

Alteryx Interview questions:

1. Can you explain what Alteryx is and how it is used in data analytics?
2. What is the difference between a tool and a macro in Alteryx?
3. How do you handle missing values or outliers in Alteryx?
4. How would you optimize a large Alteryx workflow for better performance?
5. Can you explain the concept of data blending in Alteryx and provide an example of when you would use it?
6. How do you handle errors and exceptions in Alteryx workflows?
7. Can you explain the process of data cleansing and preparation in Alteryx?
8. How do you schedule and automate workflows in Alteryx?
9. Have you worked with Alteryx Server or Alteryx Gallery? Can you explain how they are used?
10. Can you provide an example of a complex Alteryx workflow you have developed and explain the logic behind it?
11. Have you integrated Alteryx with other tools or platforms? If so, can you provide an example?
12. How do you collaborate with other team members or stakeholders when developing Alteryx workflows?
13. Can you describe a challenging problem you have faced while using Alteryx and how you resolved it?
14. Are you familiar with Alteryx Designer SDK? Have you used it to create custom tools or macros?
15. How do you stay updated with the latest features and updates in Alteryx?

Technical Skills

- **Core Alteryx Concepts:**
  - Explain the Alteryx Designer interface and its key components.
  - Describe different data input/output methods (files, databases, APIs).
  - Explain data transformation techniques (cleaning, filtering, joining, aggregating).
  - Discuss data blending and its importance in Alteryx.
- **Advanced Alteryx Tools:**
  - **Predictive Tools:** Explain how to use tools like the Decision Tree, Logistic Regression, and Time Series.
    - Discuss the importance of data preparation for predictive modeling.
    - Explain how to evaluate model performance (accuracy, precision, recall, AUC).
  - **Spatial Tools:** Describe how to work with spatial data (shapefiles, coordinates).
    - Explain common spatial operations (buffering, intersecting, distance calculations).
    - Discuss the use cases for spatial analysis in your experience.
  - **Macros:** Explain the concept of macros and their benefits.
    - Describe different types of macros (standard, batch, iterative).
    - Discuss how to create and utilize macros effectively.
- **Workflow Optimization:**
  - Explain techniques for improving workflow performance (data sampling, tool optimization, parallel processing).
  - Discuss the importance of workflow documentation and commenting.
  - Explain how to debug and troubleshoot workflow errors.

- **Alteryx Server:**
    - Describe the role of Alteryx Server in deploying and scheduling workflows.
    - Explain how to configure and manage Alteryx Server.
    - Discuss security and access control within Alteryx Server.

Business Acumen

- **Data Analysis:**
    - Describe your approach to analyzing business problems and identifying data needs.
    - Explain how you translate business requirements into actionable Alteryx workflows.
    - Discuss how to communicate data insights effectively to stakeholders.
- **Project Management:**
    - Describe your experience with project planning and execution.
    - Explain how you manage priorities and deadlines in a fast-paced environment.
    - Discuss your experience with Agile or other project methodologies.
- **Communication and Collaboration:**
    - Describe your experience working with cross-functional teams (business analysts, data engineers, etc.).
    - Explain how you communicate technical concepts to non-technical audiences.
    - Discuss your experience with mentoring or knowledge sharing.

Behavioral Questions

- **Tell me about a challenging Alteryx project you worked on and how you overcame the obstacles.**
- **Describe a situation where you had to prioritize competing demands.**
- **How do you stay current with the latest developments in Alteryx and data analytics?**
- **Why are you interested in this Senior Alteryx Developer position?**

++

Dev and Production environment Checklist	Hours/Days
Infrastructure code creation Terraform(Plan, validate and apply)	2.5 weeks
Setting up Azure virtual network, Blob storage , ADF, Role assignment, IAM role, NSG	1 week
SHIR request for VM	1 day
Grouping creation for Dev/Production in Secure	1-2 Hours
Snowflake set up and connectivity request IP whitelisting	3 days
CI CD set up for ADF and Snowflake prod environments	1.5 Week
RAD service Account Set up for Snowflake Prod environment	1-2 hours
Email notification graph API Service principal configuration	2 days
ADF blob storage connectivity set up, Linked service connection	2 days
Storage Integration set up in Snowflake for Dev/Prod	2 days
Data bricks set up via Terraform	4 days
SCIM Implementation	2 days
Snowflake connectivity, Mount blob storage	3 days
GRAPH API for email notification Testing	1 day
Linked service, Integration run time testing on ADF instance	2 days
Nonuser ID creation, ServiceNow creation Via secure for new application	1 days
ADF GitHub Integration and validation	1 days
Databricks snowflake connectivity, Mount storage connectivity	2 days
Key vault set up for data bricks credentials	1 days

++

Development & Unit Testing:
- Migrate remainder complex 9 API datasets from data source Mercury into Bronze and Silver layers of Medallion architecture.
- Migrate 4 Peoplesoft datasets into the Bronze and Silver layers.

++

HR/PeopleSoft/AP	AP Spend
HR/PeopleSoft/AP	HR Hierarchy Information
HR/PeopleSoft/AP	PS M&A Suppliers
HR/PeopleSoft/AP	PeoplePlatform

+++

Flat file (XLSX/CSV)		42		
Adobe Sign		1	Is this still needed dataset?	
Ariba		10	ACW Data, Invoice Data (transition to API), PR, PO	
CBRE		0	No longer needed (Snowflake share)	
eGRC		0	data is replace with Mercury/Aravo APIs	
Aravo		2	API connection	Set up time with Ryan Anderson on Aravo
ESG		20	Speak with Nate on necessary datasets (lower priority?)	
PeoplePlatform		2	API connection - verify how many tables?	
Hubstar		0	No longer needed	
GeoSpatial	??		API - need further investigation	
Mercury		0	no longer needed	
Noosh		4	3 flat file inputs for diversity reporting, 1 flat file for spend reporting	
EOHS (SharePoint)		1	Verify what this dataset is? Are there others?	
Site Strategy		2	SQL tables	
Golden Record exec tracker		1	Sharepoint List	
PredictX	??		REST API	
Enablement Database		1	Sharepoint List	
PeopleSoft		17	not inclusive of BDPaaS	
Snowflake share (AP)		3	replaces BDPaaS (2 of these are redundant to flat files)	
OT Data Hub		2	Ask Raffi if we still need these	
Fieldglass	??		REST API	
External File (XML, 5 suppliers)		10	Reseller data (ECG dropping to NAS today, change config to drop to Blob?)	
SQL/Oracle		4		
Concur Travel & Expense		1		
EOHS Cority		1	Verify with Holly/Manny this is only 1 table	
Lenel Badge Swipe data pull		1		
One-Click E&P Mercury Surveys		1	NPS/CSAT	
REST API (CSV)		34		
Total Count:		90		

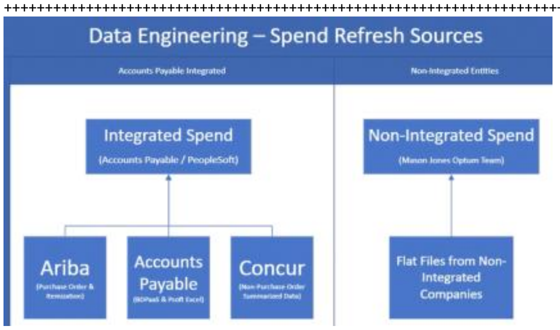
++

```
# Example Write Excel file from Databricks to Azure Blob Storage
from pyspark.sql import SparkSession
from pyspark.sql.functions import *
from io import BytesIO
import pandas as pd
from azure.storage.blob import BlobServiceClient

# Replace with your Azure Blob Storage connection details
storage_account_name = "your_storage_account_name"
storage_account_key = "your_storage_account_key"
container_name = "your_container_name"
blob_name = "output.xlsx"  # Desired name of the Excel file in Blob Storage

# Replace with your Databricks table or DataFrame
# Example: Creating a sample DataFrame
data = [("Alice", 25), ("Bob", 30), ("Charlie", 35)]
columns = ["Name", "Age"]
df = spark.createDataFrame(data, columns)
```

```
# Convert Spark DataFrame to Pandas DataFrame
pandas_df = df.toPandas()

# Create a BytesIO object to store the Excel file in memory
excel_buffer = BytesIO()

# Write the Pandas DataFrame to the Excel file in the buffer
pandas_df.to_excel(excel_buffer, index=False, engine='openpyxl') # Ensure openpyxl is installed.

# Reset the buffer's position to the beginning
excel_buffer.seek(0)

# Create a BlobServiceClient
connection_string = f"DefaultEndpointsProtocol=https;AccountName={storage_account_name};AccountKey={storage_account_key};EndpointSuffix=core.windows.net"

blob_service_client = BlobServiceClient.from_connection_string(connection_string)

# Get a BlobClient for the desired blob
blob_client = blob_service_client.get_blob_client(container=container_name, blob=blob_name)

# Upload the Excel file from the buffer to Blob Storage blob_client.upload_blob(excel_buffer.getvalue(), overwrite=True)
 print(f"Excel file '{blob_name}' uploaded to '{container_name}' in Azure Blob Storage.")
```

Development & Unit Testing:
- Migrate 20 API datasets from data source Mercury into Bronze and Silver layers of Medallion architecture.
- WEBAPI, ADF, Azure Blob Storage, Custom Framework

++



++

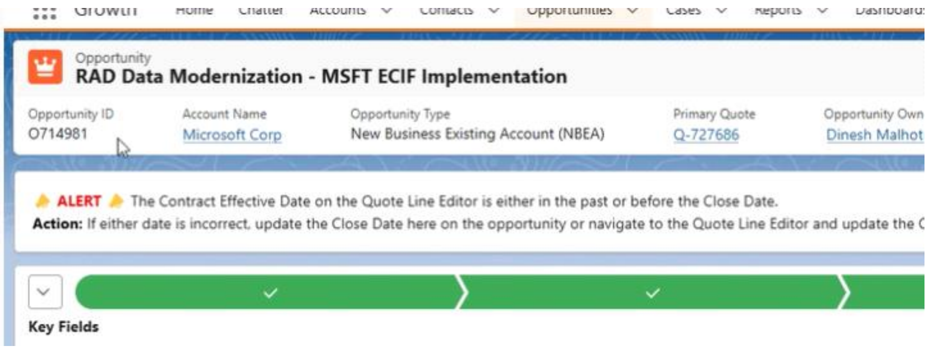
	Milestone	Details	PROGRESS	START	END
	Data Engineering				
Cloud Platform Setup	Infrastructure Provisioning	Provision Azure resources and set up the Snowflake environment (network, security, storage).	50%	3/31/25	4/21/25
	Environment Configuration & Testing	Configure backup, monitoring, and user access; perform connectivity and performance tests.	35%	4/15/25	4/28/25
Data Pipelines	Data Source Schema Design	Document key data sources (PeopleSoft, Quantum, Pam, etc.) schema details.	15%	4/17/25	5/12/25
	ETL Workflow Design	Draft initial ETL/ELT designs for each data source using Azure Data Factory; map out required transformations.	0%	4/22/25	5/15/25
	Initial Pipeline Development & Testing	Build and deploy first version of data pipelines; run initial tests to verify data flows.	0%	4/27/25	5/7/25
	Iteration 1 - Pipeline Tuning	Analyze initial test results; adjust ETL mappings and transformation logic; update pipelines accordingly.	0%	5/13/25	5/27/25
	Iteration 2 - Pipeline Refinement	Retest after adjustments; gather stakeholder feedback; further tweak parameters (e.g., handling of data delays or performance bottlenecks).	0%	5/27/25	6/12/25
	Pipeline Stabilization & Documentation	Finalize settings and document the tuned pipelines for consistent, repeatable data ingestion.	0%	6/15/25	6/22/25
Automated Data Workflows	Design Data Automation Strategy	Define scheduling and parameters for delta-loads; design an automation plan that minimizes performance impacts (noting any Data Source limitations).	0%	6/17/25	6/20/25
	Configure Automated Pipelines	Develop and deploy automated workflows in Azure Data Factory; set up incremental data load schedules, error handling, and notifications.	0%	6/21/25	6/28/25
	End-to-End Workflow Testing - Iterative Cycle	Run full-cycle tests from data ingestion to dashboard update; refine schedules and automation parameters based on iterative test outcomes.	0%	6/29/25	7/7/25

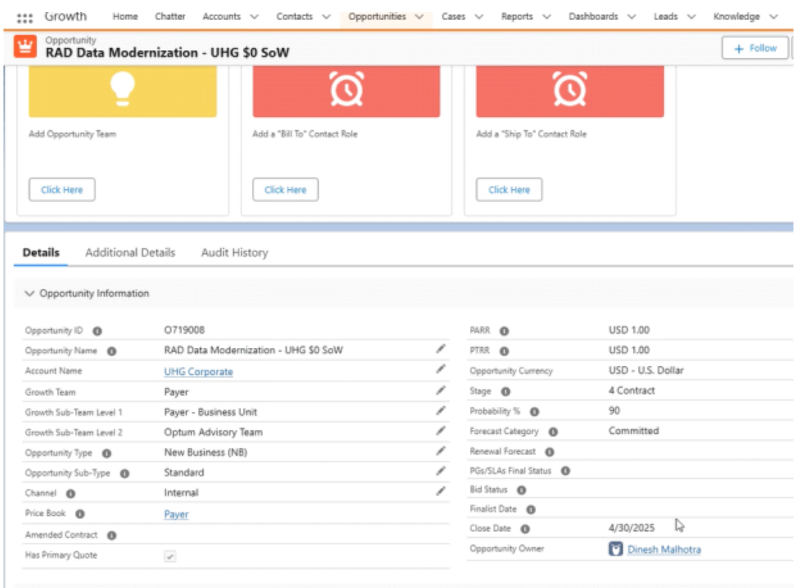
Setup production environment
Configure production environment
Migration data schema design
Load data into production environment

++

Bronze Layer Table Name	AP PeopleSoft File(s) Origin	Modernization Rename
tblSpend_BDPaas_Full2	UHPAC & UHPAC4 & Missing & Missing1	PEOPLESOFT_AP_NXEZ_SPEND
SDH_Vendors	AP_CORPSYSPULL_SDH & AP_CORPSYSPULL_SDH_INTL	PEOPLESOFT_AP_SDH_VENDORS
ApproversPS	ApproverID	PEOPLESOFT_AP_INVOICE_APPROVERS
SegmentHierarchy_PS	OU_Hierarchy	PEOPLESOFT_AP_SEGMENT_OU_HIERARCHY
Active	VENDORLIST_ACTIVE1 & VENDORLIST_ACTIVE2	PEOPLESOFT_AP_ACTIVE_VENDOR_LIST
NonActive	VENDORLIST_NONACTIVE1 & VENDORLIST_NONACTIVE2	PEOPLESOFT_AP_INACTIVE_VENDOR_LIST
Department	Department	PEOPLESOFT_AP_DEPARTMENT_DESCRIPTION
Origins	Origins	PEOPLESOFT_AP_ORIGINS_DESCRIPTION
Product	Product_Index	PEOPLESOFT_AP_PRODUCT_DESCRIPTION
PS_PO	AP_CORPSYSPULL_CONTRACTID	PEOPLESOFT_AP_PURCHASE_ORDER
ACC_PS_Spend	Spend_Query_CIDV2	PEOPLESOFT_AP_CONTRACT_VENDOR_PERSISTENCE_SPEND

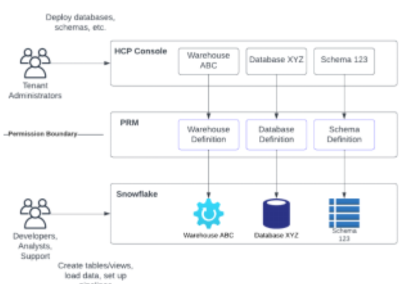
+++





++

Managing Storage Tenants



- **Tenant administrators** manage the deployment of Snowflake databases, schemas, warehouses, resource monitors and other higher-level constructs
- **Tenant users** manage Snowflake tables, views, tasks and the data used within those databases and schemas

Snowflake Roles

- Access to Snowflake objects and data is governed by roles
- Storage tenants are provisioned with three, distinct functional roles per environment: Developer, Analyst, and Support
- They are referred to as **personas** and are intended to be requested by tenant users. They follow the naming convention as shown below:

Role	**Persona**	**Description**
AZU_SDRP_{TENANT PREFIX}_{ENV}_DEVELOPER_ROLE	Developer	Provides read/write access to most objects
AZU_SDRP_{TENANT PREFIX}_{ENV}_SUPPORT_ROLE	Support	Provides mostly read-only access to data
AZU_SDRP_{TENANT PREFIX}_{ENV}_ANALYST_ROLE	Analyst	Provides read-only access to objects

Requesting Access for Snowflake Roles

- All Snowflake roles are created in Secure
- To request the access for a role, users can visit Secure and search for the roles in the DWAAS application

Container Roles

- They do **not come preconfigured** with any particular permissions, instead **allows the tenant's users to define their scope and meaning**.
- Container roles are created via HCP Console, e.g., **AZU_SDRP_{TENANT PREFIX}_{ENV}_{NAME}_CN_RL**
- They are used as membership lists to check **if a particular Snowflake user belongs to a role**.
- They can be **used in row access policies and column masking policies**

Custom Roles

- They are used **to lock down access to specific schemas**. Administrators can **create them to narrow the access scope**
- They are **provisioned in Snowflake and in Secure**, separated by logical environment
- They **appear in Secure under each storage tenant**, and **tenant users can request access** to them
- A custom role **carries a permission set similar to the functional roles and can be assigned to specific database schemas**.
- **Warehouses can be assigned to custom roles**

Custom Roles Limitations

- Administrators are **limited to a maximum of 10 custom roles** per storage tenant
- If an increase in quota is needed, a ServiceNow ticket would need to be created along with its justification

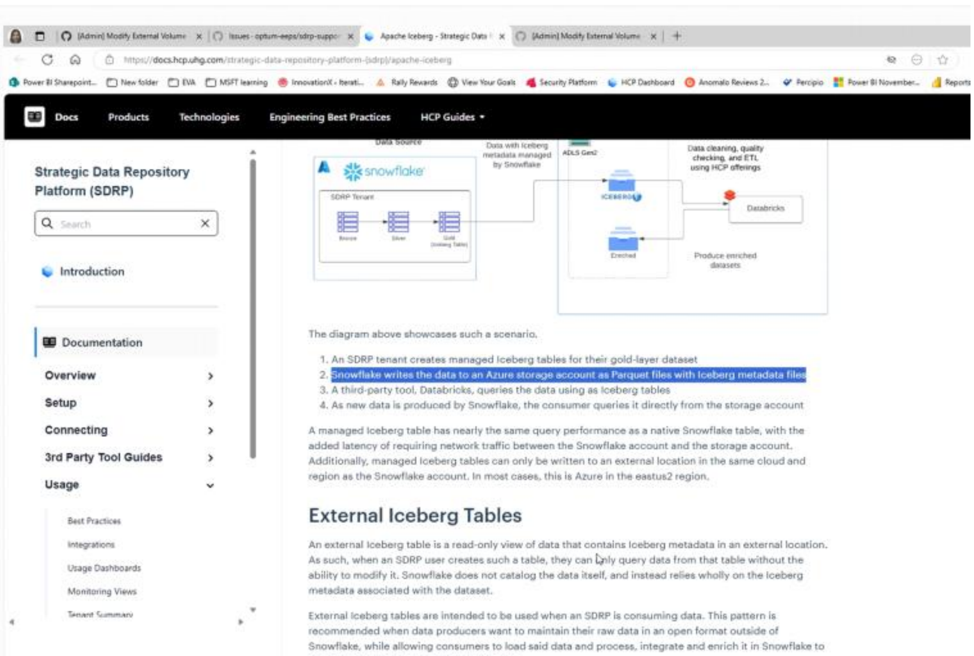
++

- Migrate remainder complex 9 API datasets from data source Mercury into Bronze and Silver layers of Medallion architecture
- Migrate 4 PeopleSoft datasets into the Bronze and Silver layers.

- Transform and migrate Silver layer tables, created in the milestone #3 into ~52 Gold layer views
- Migrate 20 ESG datasets into the Bronze and Silver Layers of Medallion architecture.
- AZCopy, Azure Blob Storage, ADF, ETL, Custom Framework, Custom Scripting

Transform and migrate Silver layer tables, created in the milestone #3 into ~52 Gold layer views
    Design 52 Gold views definitions
    Create 52 Gold Views
Migrate 20 ESG datasets into the Bronze and Silver Layers of Medallion architecture
    Create 20 Bronze layer table definitions
    Create 20 Silver layer table definitions
    Acquire data from 20 ESG datasets
    Create a stored procedure to migrate data from Landing zone to Bronze layer tables
    Create 20 stored procedures to ETL and migrate Bronze layer data to Silver layer
    Operationalize 20 ADF data pipelines
    Migrate data from Landing zone --> Bronze layer --> Silver layer
    Validate migration results

++



++

Hi Dave,

The left side of the 2nd slide (Our Understanding) should be reflective of the current situation / challenge the client team is facing and not the solution they are looking for. Solution they are looking for would be the right side (Outcomes).

WellMed Central Quality have a need to:

- Replace their current on-premise data platform with a cloud-based, highly scalable, flexible and fully secure data platform
- Retain their current infrastructure and more specifically ETL platform intact
- Implement a data architecture that incrementally improves data quality and reliability, making it more suitable for modern use cases such as, advanced business analytics and machine learning models
- They would like to serve their data applications in a subscription model

The broader question I have is what is the level of effort needed to do the assessment and what exactly do we need to learn as part of the assessment, which is what would dictate the effort.

To understand the level of effort, we would need to learn the following:

Current State
    Collect current capabilities of the data platform to identify gaps between current state and future state
    Current use cases being served to ensure business continuity
    Identification of complex business logic that must be rewritten will enable to understand the level of effort required to achieve the future state
    Document inbound/outbound integrations with applications, reporting tools and other systems to ensure business functions continue to work after conversion
    Understand the version of Alteryx being used - Alteryx On-Premise Server or Orbit Alteryx will impact the integration design of the solution
    Which capabilities work well and which ones could use improvements will enable what needs to be preserved and what must be redesigned/replaced
Security
    Current governance model, security model, protection and privacy models will help understand what needs to be maintained
    Identification of PII, PHI, or other sensitive data requiring masking or encryption to ensure the same/similar protections are maintained
    Regulatory requirements (e.g., HIPAA, GDPR, SOC 2) to preserve regulatory compliances
Data Objects
    Estimate total data size, growth rate, and frequency of updates will determine the data migration strategies
    There are total 17 databases, how many databases are in scope: 13 or 14 (RO_ORCHESTRATION is missing from the deck)?
    Are objects (databases, tables, views) in the lists provided to be migrated as-is or will require redesign will have a direct impact on the size of the effort
    If redesign, which object types and objects? Who will provide transformation definitions will bring clarity to the division of responsibility
    Document T-SQL features (e.g., triggers, cursors, CLR functions) that may not be supported or need rewriting in Snowflake streams and tasks will impact the size of the effort
    Collect details of SQL Agent Jobs, triggers, and stored procedures to determine what would need to be converted and what would need to be rewritten and the size of the effort
Performance
    Identify high-frequency and long-running queries to understand which process will require special attention during design, creation and validation phases
    Document peak usage times, concurrency, and batch processing jobs will impact configurations and infrastructure sizing
    Identify performance optimizations that may need redesign in Snowflake platform
Costs
    Document SQL Server licensing, hardware, and maintenance costs will help in understanding the cost differences between current and future states

Migration

Preferred migration approach (big bang, phased, or hybrid)Data Acquisition) to prioritize migrations
How will current data be acquired - direct query, flat file, API, etc. It will determine data acquisition design choices
Will it require ECG? If yes, does WellMed already have ECG setup? This process takes weeks to complete

Do we need 4-weeks and 3 resources or is this smaller in scope and we could do it in shorter time with less resources?

We would actually need 2 resources (Lead Cloud architect and Cloud Data Analyst) to do the assessment; Delivery Oversight will have the minimum involvement

At a high-level, after this assessment, we need to have the future state solution and an estimate of what the effort and timeline to do it. I agree - current state assessment will be the basis for the future state solution, effort size and timeline

The questions we have to answer to get that defined are what needs to be addressed as part of the assessment

Hopefully, items specified above will collectively be able to answer that.

Some questions could be answered by what was shared already – or it may lead to more concise questions.

Yes, I have factored them in while compiling the list above

++

https://app.snowflake.com/uhg/uhgdwaas/#/data/databases/AKWUR_DEV_CORPSVCS_CONF_DB/schemas/UTIL

++

From: optumbilling@optum.com <optumbilling@optum.com>
Sent: Thursday, May 22, 2025 9:33:14 PM
To: Malhotra, Dinesh <dinesh.malhotra@optum.com>
Subject: Optum Invoice 0001752849 for review and payment

Optum

A new invoice 0001752849 has been generated and is attached for review and payment along with any supporting documents.

Invoice	0001752849
Amount Due	$94,301.00 USD
Invoice Date	05/22/2025
Due Date	06/21/2025
Terms	Net 30
PO	
Optum Contract	CON000000021768
Optum Customer	0000016442
Optum Customer Name	Microsoft

Remit To:
OPTUM
PO BOX 84019
Chicago IL 60689-4019
United States

ACH/Wire Transfer:
Fifth Third
ABA-042000314; Swift Code: FTBCUS3C
Account Name: Optum
A/C #: 7026949821

Include Invoice Number(s) and Customer ID with all payments

Online Payments:

To pay for US Dollar (USD) invoices electronically Visit the website: https://billpay.optum.com
to pay online. First time user?
Login with Bill Pay Number:0000016442-06310

Optum Contacts:

Billing inquiries: optumbilling@optum.com

Payment inquiries or ACH remittances: cashap@optum.com

Include Customer ID, Contract ID, and Invoice Number on all inquiries

++

2/17/2025	3/15/2025
3/16/2025	4/15/2025
4/16/2025	5/15/2025
5/16/2025	6/15/2025
6/16/2025	7/15/2025
7/16/2025	8/15/2025
8/16/2025	9/15/2025
9/16/2025	10/17/2025

Milestone #	Item #	Task	Start Date	End Date	Status
1	0	Migrate 20 API datasets from data source Mercury into Bronze and Silver layers	2/17/2025	3/15/2025	Complete
2	0	Migrate second set of 24 API datasets from data source Mercury into Bronze and Silver layers	3/16/2025	4/15/2025	Complete
3	0	Migrate remainder 9 API datasets from data source Mercury into Bronze and Silver layers. Migrate 4 Peoplesoft datasets into the Bronze and Silver layers.	4/15/2025	5/15/2025	Complete

++

Chaos testing
Load testing
Performance testing
Stress testing

OA team will identify testing requirements, workload type, performance metrics, cloud environment, appropriate testing tools.

References

Friday, September 20, 2024 12:24 PM

https://uhgazure.sharepoint.com/teams/OAS-MRISAssessment/Shared%20Documents/General

General OAS-MRIS 2024 Modernization Program: OAS-MRIS 2024 Modernization Program | General | Microsoft Teams

MRIS 2024 Modernization Program
MRIS Documents
OAS Documents
Future State Reference artifacts: General
++
API Data Source example: adf-isdc-analytics-nonprod-dev - Azure Data Factory
+++
Snowflake Parse Jason example: https://docs.snowflake.com/en/sql-reference/functions/parse_json
+++

PM In
OpenAir

Meetings

Saturday, August 17, 2024 7:48 AM

Meeting 08/16/2024 with John
Subject: RAD Modernization
Move reports and consolidate source data to cloud
plan of the build (8 weeks) define solution patterns, pilot prototype
show the kind of solution they'd be getting
Cost estimate: ~$500K
Internalize and detail it out
staffing - 3 cloud engineers, 1 data platform engineer
part time role for 2 month (offshore) (scrum master)

Get familiar with the Burn Report

Put out a plan
Create detailed set of activities
Assign - who is going to be doing what
Use Rally for project tracking (spreadsheet if necessary)
Resources needed within 3-4 weeks
++
Meeting 10/08/2024 with Patrick and Grant
Subject: API Data Source current state
Mercury in-house developed tool
It is an end-to-end workflow procurement system
You request thru Mercury to get new contarctors
It is a process from business systems to payments
Currenbtly has 5+ million contracts
with delegation authority
Mercury is a very important dataset
We analyze bottlenecks in the process
Mercuriy has been around since June 2020
It is comprised of customized processes and has gone thru iterative efforts
It is a business to supplier tool
There are handful teams are involved
Current spend is ~$12 Billion

Mercury system is a central tool
Procurement manage spend goes thru mercury
Over 12 thousand vendors
AT&T alone has 37 profiles, we want to parent them up to AT&T
There is a Golden Record Initiative - 100 contractors to clean the data
There are 1.2 million records
very compartmentalized
30 or so API tables
REST API in Alteryx - 2:00 am 5 PM, delta pull
API calls pull directly, stored in SQL table
After do some aggregation
++
Thoughts for 10/09/2024 session
What would be the credentials to pull data?
Will share
What do you do with the data that you pull?
Pull in
Could you please describe that process in detail?
Alteryx is the documentation. Pull in delta data.
Update Bronze tables
Run multiple workflow to transform data - changing value, cleansing, table joins
Keep table separate, set up start schema
Download Alteryx desktop or complete and complex
Do you have any documentation available for that?
Tables and fields, get calculations,
What are the exceptions, alerts, notifications, retries, etc.
Rerun, have some data checks
Tasks get mapped to personas
We have our own mapping/hierarchy

Do you have the Database and schema already created in SDRP environment?
do not have in it - Have database and schema

If yes, what are database and schema names and Access Details, e.g. Web UI URL; If no, what would you like to name the database and the schema?
Do you have the data model, database schema?
If yes, could we please get it; if not, would you help with the definition and design of it?
++
Question for 10/10/2024 meeting
Are Json within Json depth known to us? Is the Json schema standardized?
Can we get detailed understanding of the JSON sample provided?
++
Questions:

Have you identified the datasets?
Yes, later
Which call to use for them?
Mercury, Service now
What are table structures?
What are the future state data model?
Grant, did you update the diagram that shows how to select

Request table and task table
++
10/15/2024
SQL data source -

- please share the diagram
- Share access credentials
- Could we get access to the source code
- Any documentation
- Grant: "**I've sent it so these two tables here. These are essentially mirror images of each other**". What is the purpose of having the mirror tables?
  - "When we get data load prior to the data stored, everything from that point onward get cleared out and populated with the dat a load."
- The transformation process is hairy. We need to detail it out to fully understand it.

Talk about dynamic parameters

VMs
Approval
SQL server permissions
++
10/22/2024

1. RAD
   a. table design for, Medallion architecture
   b. Ask if they reviewed the High Level design documents
   c. Divide & conquer

2. Offsore
   a. VDI
   b. Detail Design documents - for pilot.
   c. Test to connect to Mercury database and get result set and drop it in the Azure Blob Storage

3. Asit
   a. Help with connectivity to Mercury
   b. Key vault
   c. Need to start scripting and CICD automation

+++
10/24/2024

1. RAD
   a. Project timeline will be extended by 2 weeks due to holidays, people starting to take time off, data products and subscriptio n approvals will also stretch out
   b. The revised project completion date will be December 20th of 2024
   c. Should I notify formally or verbal communication is adequate?
   d. Detailed Design - Sample to be reviewed
   e. Table design for Medallion architecture - Patrick is almost done

2. Offshore
   a. VDI
   b. Test to connect to Mercury database and get result set and drop it in the Azure Blob Storage
   c. Feedback on SQL data type detailed design
   d. What other infrastructure do we need to stand up for the pilot project?

3. Asit
   a. Help with connectivity to Mercury
   b. Key vault
   c. Need to start scripting and CICD automation

4. API calls, storage
5. Key vault
6. Terraform scripting

++
10/25/2024

1. RAD
   a. Internal file detailed design document submitted for review.
   b. Did you get a chance to create Bronze layer tables?
   c. Table design for Medallion architecture - Patrick is almost done
   d. Are all tables from all data source types going to be this complex
   e. Discuss Grant's concern

2. Offshore
   a. Feedback on SQL data type detailed design - done
   b. Feedback on Internal File Detailed design document - done
   c. Snowflake access
   d. Key vault set
3. Asit
   a. Help with connectivity to Mercury - done
   b. Key vault

    c. Need to start scripting and CICD automation

4. API calls, storage
5. Key vault
6. Terraform scripting

++

11/11/2024

1. RAD
    a. Silver layer tables design -
        i. where are we?
        ii. Can we do one table end-to-end?
        iii. Need to start thinking about how best to create Data Products.
        iv. There are several approaches, e.g., consumer use case specific, domain specific, etc.
    b. How are we going to migrate API data source Historical data?
    c. Did you get a chance to try to link PowerBI to Snowflake?
    d. Are there any other data source type that we would need to migrate historical data for?
    e. **Date format to be yyyy-mm-dd across all date fields - 10/30/2024**
    f. When do you think we'd be able to complete the gold layer table design for just one table?

Issues:

- It is taking a bit longer because the RAD wants to redesign the data models
- 30% Data types are different than what we were provided
- JSON value size in source data is more than 16MB in some columns, whereas the Snowflake VARIANT column size limit is 16 MB. How should that be handled? What is you guidance?

1. Offshore
    a. Sujit - pipeline framework - table DDL, scripts, SPs
       Asit -

2. Asit
    a. All automations and IaC

3. API calls, storage
4. Key vault
5. Terraform scripting
6. Mahesh has set the limit to 500 instead of 3000

+++

Meeting 11/11/2024 with John Shin
Subject: Validate estimate and JSONs larger than 16 MB

1. Validate estimates, make needed suggestions
2. How do I get the project id?
3. What should be the Deal Type?
4. Medallion arch. tables are being redesigned - time consuming. Should that impact pricing or change the Deal Type?
5. They want to embed a green developer with our team. Should this impact the pricing?

Current project related:

1. Medallion arch. tables are being redesign - time consuming
2. Source data types are different than what we were provided
3. Source data JSON value size is more than 16MB, whereas Snowflake VARIANT column size limit is 16 MB. What would be your guidance to handle this situation?
   The following are the options:
    a. Parse and store data in the Bronze layer
    b. Split JSON into 16 MB chunks
    c. Store the large JSON file in an external stage and reference the file path from within a VARIANT column

++

To do list - 12/02/2024
1. Patrick & Grant
    Alex's approval is still pending after 6 days
    Complete data product subscription
    Create data products based on Views datasets
    Identify Access Groups for each data product
    Create Access group
    Create subscriptions
    1. Documentation tool for UAT out come
    2. Muquaddar
        a. Complete Data flows
2. Reach out to Veera and Santosh to complete the PowerBI and Snowflake connectivity - sent a Teams message to them
3. Reach out to Santosh and Asit regrading Azure Machine Learning Studio set up and database connectivity - sent a Teams message to them

++

To do list - 12/05/2024

Questions:

- Have you started to validate data and pilot project?
- Plan for the Access Groups for data product subscriptions
- Create data products for the remaining two views and their subscriptions

Should create the remaining two data products, and create subscription requests for them.
Alex was concerned about too many meetings, so I have sent them the decks for the Fina deliverables and the proposal as well

Key vault demo will be done next week when Asit comes back- we're going to create a list of things for Asit to complete when he comes back on Monday, e.g., KeyVault demo
Also, a library deployment in Databricks will also be completed when Asit comes back

Recreate the view with COPY GRANTS

Assumptions validations needed:

- There are total 85 datasets to be migrated
- Four views/silver table
- RAD team will be responsible for the design effort of medallion architecture tables and views. OA team will provide needed support
- RAD will own production acceptance testing (PAT) and decommissioning phases with 1 person from the OA team
- PAT will last 8 weeks and decommissioning will last 4 weeks
- RAD team will drop Outlook email attachments, internal files, and external (vendor) files to a designated Azure Blob Storage location
- RAD team will create data products and help end consumers with the subscriptions and how to consume data
- RAD team will help end consumers with existing processes (e.g., reports, etc.) to point the new database, schema, and table names
- Resources: OA will be responsible for the Dev, Test, Deploy phases. The RAD team will lead the PAT and decommissioning phases with OA support

- PowerBI to Snowflake connectivity - in progress - working with someone from the Power-BI intake team and with Victoria Kong from Orbit team
- Data Science platform Demo - scheduled for tomorrow
- Communicated to the RAD to validate data
- The team is anxious about the next phase project
- Charge anything for the Azure Data Science Platform setup
- Pricing for the next phase
- FYI, RAD team wants us to show them how to migrate SQL data

Patrick, Based on what we know so far, it looks like the development will be 7 months, production acceptance 2 months and decommissioning 1 month
For the dev, test & deploy - we would need 4 cloud data engineers, 1 cloud platform administrator, 1 scrum master, 1 delivery lead

ML Platform

Monday, November 25, 2024 4:43 PM

To setup Machine Learning platform:

1. Setup Azure Databricks instance
2. Create Azure Blob Storage
3. Upload external dataset to Azure Blob Storage
4. Upload Jupyter notebook
5. Enable Databricks instance connectivity and access permissions with dataset in the Blob Storage
6. Setup GitHub and Notebook integration
7. Create output folder
8. Output folder to store Machine Learning results
9. CICD automation

Sample - ADF Workflow Management Framework to Control Execution

Friday, October 25, 2024   12:57 PM

wf_ACCOUNTABLE_CALLS

++


```
{
  "name": "wf_ACCOUNTABLE_CALLS",
  "properties": {
    "activities": [
      {
        "name": "Start Log",
        "type": "Script",
        "dependsOn": [
          {
            "activity": "Start the Porcess",
            "dependencyConditions": [
              "Succeeded"
            ]
          }
        ],
        "policy": {
          "timeout": "0.12:00:00",
          "retry": 0,
          "retryIntervalInSeconds": 30,
          "secureOutput": false,
          "secureInput": false
        },
        "userProperties": [],
        "linkedServiceName": {
          "referenceName": "ls_sf_isdc",
          "type": "LinkedServiceReference",
          "parameters": {
            "par_sf_acct": {
              "value": "@pipeline().globalParameters.par_sf_acct",
              "type": "Expression"
            },
            "par_sf_user": {
              "value": "@pipeline().globalParameters.par_sf_user",
              "type": "Expression"
            },
            "par_sf_db": {
              "value": "@pipeline().globalParameters.par_sf_db",
              "type": "Expression"
            },
            "par_sf_wh": {
              "value": "@pipeline().globalParameters.par_sf_wh",
              "type": "Expression"
            },
            "par_sf_role": {
              "value": "@pipeline().globalParameters.par_sf_role",
              "type": "Expression"
            }
          }
        },
        "typeProperties": {
          "scripts": [
            {
              "type": "NonQuery",
              "text": {
                "value": "@concat(\n'INSERT INTO UTIL.PROCESS_RUN_LOGS_ANALYTICS\n(\nPIPELINE_ID,\nPIPELINE_NAME,\nPROCESS_NAME,\nSUB_PROCESS_NAME,
\nSTART_DATETIME,\nSTATUS,\nDATAFACTORY,\nTRIGGER_NAME,\nTRIGGER_TYPE,\nTRIGGER_TIME,\nAPP_NAME\n)\nVALUES (''',pipeline().RunId,''',''',pipeline().Pipeline,''',
\n',''',pipeline().parameters.p_application,''',''',pipeline().parameters.p_object_name,
\n''',''',pipeline().TriggerTime,''',''Running'',''',pipeline().DataFactory,''',''',pipeline().TriggerName,''',''',pipeline().TriggerType,''',''',pipeline().TriggerTime,''',''DEEP'');')",
                "type": "Expression"
              }
            }
          ],
          "scriptBlockExecutionTimeout": "02:00:00"
        }
      },
      {
        "name": "Get_Object_Metadata",
        "type": "Lookup",
        "dependsOn": [
          {
            "activity": "Start Log",
            "dependencyConditions": [
```

```
                    "Succeeded"
                ]
            }
        ],
        "policy": {
            "timeout": "0.12:00:00",
            "retry": 0,
            "retryIntervalInSeconds": 30,
            "secureOutput": false,
            "secureInput": false
        },
        "userProperties": [],
        "typeProperties": {
            "source": {
                "type": "SnowflakeV2Source",
                "query": {
                    "value": "@concat('select\nPGM.OBJECT_DB,\nPGM.OBJECT_SCHEMA,\nPGM.OBJECT_NAME,  \nPGM.OBJECT_SIGNATURE\nFROM UTIL.PROGRAM_LIST_ANALYTICS
PGM JOIN UTIL.PROCESS_LIST_ANALYTICS PRC ON\nPGM.PROCESS_NAME = PRC.PROCESS_NAME AND PGM.DATA_FACTORY_NAME = PRC.DATA_FACTORY_NAME AND
PRC.ACTIVE_INDC = ''Y''\nWHERE \nPGM.PROCESS_NAME = substr(','''', pipeline().Pipeline,'''', ',4)\n')",
                    "type": "Expression"
                },
                "exportSettings": {
                    "type": "SnowflakeExportCopyCommand"
                }
            },
            "dataset": {
                "referenceName": "ds_sf_isdc",
                "type": "DatasetReference",
                "parameters": {
                    "par_sf_acct": {
                        "value": "@pipeline().globalParameters.par_sf_acct",
                        "type": "Expression"
                    },
                    "par_sf_user": {
                        "value": "@pipeline().globalParameters.par_sf_user",
                        "type": "Expression"
                    },
                    "par_sf_db": {
                        "value": "@pipeline().globalParameters.par_sf_db",
                        "type": "Expression"
                    },
                    "par_sf_wh": {
                        "value": "@pipeline().globalParameters.par_sf_wh",
                        "type": "Expression"
                    },
                    "par_sf_role": {
                        "value": "@pipeline().globalParameters.par_sf_role",
                        "type": "Expression"
                    }
                }
            },
            "firstRowOnly": false
        }
    },
    {
        "name": "ACCOUNTABLE_CALLS",
        "type": "Script",
        "dependsOn": [
            {
                "activity": "Get_Object_Metadata",
                "dependencyConditions": [
                    "Succeeded"
                ]
            }
        ],
        "policy": {
            "timeout": "0.12:00:00",
            "retry": 0,
            "retryIntervalInSeconds": 30,
            "secureOutput": false,
            "secureInput": false
        },
        "userProperties": [],
        "linkedServiceName": {
            "referenceName": "ls_sf_isdc",
            "type": "LinkedServiceReference",
            "parameters": {
                "par_sf_acct": {
                    "value": "@pipeline().globalParameters.par_sf_acct",
```

```
                    "type": "Expression"
                },
                "par_sf_user": {
                    "value": "@pipeline().globalParameters.par_sf_user",
                    "type": "Expression"
                },
                "par_sf_db": {
                    "value": "@pipeline().globalParameters.par_sf_db",
                    "type": "Expression"
                },
                "par_sf_wh": {
                    "value": "@pipeline().globalParameters.par_sf_wh",
                    "type": "Expression"
                },
                "par_sf_role": {
                    "value": "@pipeline().globalParameters.par_sf_role",
                    "type": "Expression"
                }
            }
        },
        "typeProperties": {
            "scripts": [
                {
                    "type": "NonQuery",
                    "text": {
                        "value": "@concat(\n'CALL ',\nstring(activity(\n'Get_Object_Metadata'\n).output.value[0].OBJECT_DB), '.',
\nstring(activity(\n'Get_Object_Metadata'\n).output.value[0].OBJECT_SCHEMA),'.','SP_ACCOUNTABLE_CALLS','(','''',pipeline().RunId,'''',',','''',pipeline().Pipeline,'''',',',string(activity(\n'Ge
t_Object_Metadata'\n).output.value[0].OBJECT_SIGNATURE),')'\n\n)",
                        "type": "Expression"
                    }
                }
            ],
            "scriptBlockExecutionTimeout": "02:00:00"
        }
    },
    {
        "name": "On Complete End Log",
        "type": "Script",
        "dependsOn": [
            {
                "activity": "ACCOUNTABLE_CALLS",
                "dependencyConditions": [
                    "Succeeded"
                ]
            }
        ],
        "policy": {
            "timeout": "0.12:00:00",
            "retry": 0,
            "retryIntervalInSeconds": 30,
            "secureOutput": false,
            "secureInput": false
        },
        "userProperties": [],
        "linkedServiceName": {
            "referenceName": "ls_sf_isdc",
            "type": "LinkedServiceReference",
            "parameters": {
                "par_sf_acct": {
                    "value": "@pipeline().globalParameters.par_sf_acct",
                    "type": "Expression"
                },
                "par_sf_user": {
                    "value": "@pipeline().globalParameters.par_sf_user",
                    "type": "Expression"
                },
                "par_sf_db": {
                    "value": "@pipeline().globalParameters.par_sf_db",
                    "type": "Expression"
                },
                "par_sf_wh": {
                    "value": "@pipeline().globalParameters.par_sf_wh",
                    "type": "Expression"
                },
                "par_sf_role": {
                    "value": "@pipeline().globalParameters.par_sf_role",
                    "type": "Expression"
                }
            }
```

```
            },
            "typeProperties": {
                "scripts": [
                    {
                        "type": "NonQuery",
                        "text": {
                            "value": "@concat(\n'EXECUTE IMMEDIATE ','''','UPDATE ',\nstring(activity('Get_Object_Metadata').output.value[0].OBJECT_DB), '.',
\npipeline().parameters.p_util_schema,'.','PROCESS_RUN_LOGS_ANALYTICS SET
STATUS=','''','''','COMPLETED','''','''',',','END_DATETIME=','''','''',convertTimeZone(activity('ACCOUNTABLE_CALLS').ExecutionEndTime, 'UTC', 'Central Standard Time'),'''','''',' WHERE
PIPELINE_ID=','''','''',pipeline().RunId,'''','''',''''\n)\n",
                            "type": "Expression"
                        }
                    }
                ],
                "scriptBlockExecutionTimeout": "02:00:00"
            }
        },
        {
            "name": "Set_variable",
            "type": "SetVariable",
            "dependsOn": [
                {
                    "activity": "plt_email_success",
                    "dependencyConditions": [
                        "Succeeded"
                    ]
                }
            ],
            "policy": {
                "secureOutput": false,
                "secureInput": false
            },
            "userProperties": [],
            "typeProperties": {
                "variableName": "Completed",
                "value": "Process is completed"
            }
        },
        {
            "name": "plt_email_notify",
            "type": "ExecutePipeline",
            "dependsOn": [
                {
                    "activity": "Start Log",
                    "dependencyConditions": [
                        "Failed"
                    ]
                }
            ],
            "policy": {
                "secureInput": false
            },
            "userProperties": [],
            "typeProperties": {
                "pipeline": {
                    "referenceName": "plt_email_notify",
                    "type": "PipelineReference"
                },
                "waitOnCompletion": true,
                "parameters": {
                    "EmailSender": {
                        "value": "@pipeline().globalParameters.par_email_sender",
                        "type": "Expression"
                    },
                    "EmailRecipient": {
                        "value": "@pipeline().parameters.p_email_recipient",
                        "type": "Expression"
                    },
                    "EmailSubject": {
                        "value": "@concat('ISDC ', toUpper(pipeline().globalParameters.par_env),' - ',  pipeline().Pipeline, ' Failed')",
                        "type": "Expression"
                    },
                    "EmailContentHTML": {
                        "value": "@concat('PipelineName: ', pipeline().Pipeline, \n'<br>', activity('Start Log').error?.message)",
                        "type": "Expression"
                    }
                }
            }
        },
```

```
{
  "name": "plt_email_notify_1",
  "type": "ExecutePipeline",
  "dependsOn": [
    {
      "activity": "Get_Object_Metadata",
      "dependencyConditions": [
        "Failed"
      ]
    }
  ],
  "policy": {
    "secureInput": false
  },
  "userProperties": [],
  "typeProperties": {
    "pipeline": {
      "referenceName": "plt_email_notify",
      "type": "PipelineReference"
    },
    "waitOnCompletion": true,
    "parameters": {
      "EmailSender": {
        "value": "@pipeline().globalParameters.par_email_sender",
        "type": "Expression"
      },
      "EmailRecipient": {
        "value": "@pipeline().parameters.p_email_recipient",
        "type": "Expression"
      },
      "EmailSubject": {
        "value": "@concat('ISDC ', toUpper(pipeline().globalParameters.par_env),' - ',  pipeline().Pipeline, ' Failed')",
        "type": "Expression"
      },
      "EmailContentHTML": {
        "value": "@concat('PipelineName: ', pipeline().Pipeline, \n'<br>', activity('Start Log').error?.message)",
        "type": "Expression"
      }
    }
  }
},
{
  "name": "plt_email_notify_2",
  "type": "ExecutePipeline",
  "dependsOn": [
    {
      "activity": "ACCOUNTABLE_CALLS",
      "dependencyConditions": [
        "Failed"
      ]
    }
  ],
  "policy": {
    "secureInput": false
  },
  "userProperties": [],
  "typeProperties": {
    "pipeline": {
      "referenceName": "plt_email_notify",
      "type": "PipelineReference"
    },
    "waitOnCompletion": true,
    "parameters": {
      "EmailSender": {
        "value": "@pipeline().globalParameters.par_email_sender",
        "type": "Expression"
      },
      "EmailRecipient": {
        "value": "@pipeline().parameters.p_email_recipient",
        "type": "Expression"
      },
      "EmailSubject": {
        "value": "@concat('ISDC ', toUpper(pipeline().globalParameters.par_env),' - ',  pipeline().Pipeline, ' Failed')",
        "type": "Expression"
      },
      "EmailContentHTML": {
        "value": "@concat('PipelineName: ', pipeline().Pipeline, \n'<br>', activity('Start Log').error?.message)",
        "type": "Expression"
      }
```

```
        }
      }
    },
    {
      "name": "plt_email_notify_6",
      "type": "ExecutePipeline",
      "dependsOn": [
        {
          "activity": "On Complete End Log",
          "dependencyConditions": [
            "Failed"
          ]
        }
      ],
      "policy": {
        "secureInput": false
      },
      "userProperties": [],
      "typeProperties": {
        "pipeline": {
          "referenceName": "plt_email_notify",
          "type": "PipelineReference"
        },
        "waitOnCompletion": true,
        "parameters": {
          "EmailSender": {
            "value": "@pipeline().globalParameters.par_email_sender",
            "type": "Expression"
          },
          "EmailRecipient": {
            "value": "@pipeline().parameters.p_email_recipient",
            "type": "Expression"
          },
          "EmailSubject": {
            "value": "@concat('ISDC ', toUpper(pipeline().globalParameters.par_env),' - ',  pipeline().Pipeline, ' Failed')",
            "type": "Expression"
          },
          "EmailContentHTML": {
            "value": "@concat('PipelineName: ', pipeline().Pipeline, \n'<br>', activity('Start Log').error?.message)",
            "type": "Expression"
          }
        }
      }
    },
    {
      "name": "Start the Porcess",
      "type": "SetVariable",
      "dependsOn": [],
      "policy": {
        "secureOutput": false,
        "secureInput": false
      },
      "userProperties": [],
      "typeProperties": {
        "variableName": "start_time",
        "value": {
          "value": "@formatDateTime(utcNow(),'yyyy-MM-ddTHH:mm:ss')",
          "type": "Expression"
        }
      }
    },
    {
      "name": "End Process",
      "type": "SetVariable",
      "dependsOn": [
        {
          "activity": "On Complete End Log",
          "dependencyConditions": [
            "Succeeded"
          ]
        }
      ],
      "policy": {
        "secureOutput": false,
        "secureInput": false
      },
      "userProperties": [],
      "typeProperties": {
        "variableName": "Endtime",
```

```
                "value": {
                    "value": "@formatDateTime(utcNow(),'yyyy-MM-ddTHH:mm:ss')",
                    "type": "Expression"
                }
            }
        },
        {
            "name": "plt_email_success",
            "type": "ExecutePipeline",
            "dependsOn": [
                {
                    "activity": "End Process",
                    "dependencyConditions": [
                        "Succeeded"
                    ]
                }
            ],
            "policy": {
                "secureInput": false
            },
            "userProperties": [],
            "typeProperties": {
                "pipeline": {
                    "referenceName": "plt_email_success",
                    "type": "PipelineReference"
                },
                "waitOnCompletion": true,
                "parameters": {
                    "EmailSender": {
                        "value": "@pipeline().globalParameters.par_email_sender",
                        "type": "Expression"
                    },
                    "EmailRecipient": {
                        "value": "[{\"emailAddress\":{\"address\":\"ism_deep_de@ds.uhc.com\"}}]",
                        "type": "Expression"
                    },
                    "EmailSubject": {
                        "value": "@concat('ISDC ', toUpper(pipeline().globalParameters.par_env),' - ',  pipeline().Pipeline, '  Success')",
                        "type": "Expression"
                    },
                    "EmailContentHTML": {
                        "value": "@concat('The    ',pipeline().parameters.p_application, '  ETL has completed successfully.','</br>','   ',\n\n'</br>','Environment : ', pipeline().DataFactory,'</br>','
',\n\n'</br>','WorkFlow : ', pipeline().Pipeline,'</br>','  ',\n'</br>','Started At :',variables('start_time'),'</br>','  ',\n\n'</br>','Completed At :',variables('Endtime'),'</br>','  ',\n\n'</br>','
Please check ETL_BATCH_OBJECT_CONTROL_ANALYTICS & ETL_BATCH_OBJECT_CONTROL_LOG_ANALYTICS tables for more details.','</br>','  ','</br>','This is an automatically generated
message. Please Do not reply.','</br>','  ','</br>','Contact ISDC Support for any further information.','</br>','
','</br>','******************************************************')",
                        "type": "Expression"
                    }
                }
            }
        },
        {
            "name": "plt_email_notify_7",
            "type": "ExecutePipeline",
            "dependsOn": [
                {
                    "activity": "plt_email_success",
                    "dependencyConditions": [
                        "Failed"
                    ]
                }
            ],
            "policy": {
                "secureInput": false
            },
            "userProperties": [],
            "typeProperties": {
                "pipeline": {
                    "referenceName": "plt_email_notify",
                    "type": "PipelineReference"
                },
                "waitOnCompletion": true,
                "parameters": {
                    "EmailSender": {
                        "value": "@pipeline().globalParameters.par_email_sender",
                        "type": "Expression"
                    },
                    "EmailRecipient": {
                        "value": "@pipeline().parameters.p_email_recipient",
```

```
                    "type": "Expression"
                },
                "EmailSubject": {
                    "value": "@concat('ISDC ', toUpper(pipeline().globalParameters.par_env),' - ',  pipeline().Pipeline, ' Failed')",
                    "type": "Expression"
                },
                "EmailContentHTML": {
                    "value": "@concat('PipelineName: ', pipeline().Pipeline, \n'<br>', activity('Start Log').error?.message)",
                    "type": "Expression"
                }
              }
            }
          }
        ],
        "parameters": {
          "p_application": {
            "type": "string",
            "defaultValue": "DEEP"
          },
          "p_object_name": {
            "type": "string",
            "defaultValue": "ACCOUNTABLE_CALLS"
          },
          "p_util_schema": {
            "type": "string",
            "defaultValue": "UTIL"
          },
          "p_email_recipient": {
            "type": "string",
            "defaultValue": "ism_deep_de@ds.uhc.com"
          }
        },
        "variables": {
          "Completed": {
            "type": "String"
          },
          "start_time": {
            "type": "String"
          },
          "Endtime": {
            "type": "String"
          }
        },
        "folder": {
          "name": "Analytics/Workflows"
        },
        "annotations": [],
        "lastPublishTime": "2024-10-25T15:55:25Z"
      },
      "type": "Microsoft.DataFactory/factories/pipelines"
    }
```

Gantt Chart & WBS

Wednesday, November 13, 2024 8:49 PM

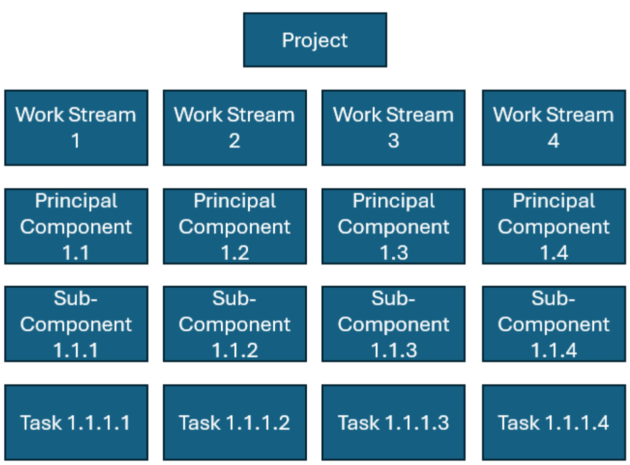
Nobody care about the activities, only deliverables and milestones matter

Gantt chart

- Plot time against the horizontal axis
- List tasks along the vertical axis
- Equal units of distance along the horizontal axis represent equal units of time
- Represent each activity by a Bar
- Length of the bar represents the duration of time
- Positioning of the bar, represent the scheduling of task
- If one task follows another, then the bar follows the previous bar
- Other tasks may follow in parallel
- A diamond shape represents and milestone
- We can put a vertical RED line thru the chart to indicate today
- A Gantt chart represents a project plan

Work Breakdown Structure (WBS)

- Work Breakdown structure (WBS) a resource management and project management tool
- It breaks the work down into a structure. It articulates a scope of work in a highly structured way
- Breakdown project into substantial chunks (work streams)
- Breakdown chunks (work streams) into principal components
- Breakdown principal components into sub-components
- Breakdown sub- components into individual tasks
- Number each level. For example:



- Work Breakdown structure should be MECE (Mutually Exclusive Completely Exhaustive)

Approach

Wednesday, August 21, 2024 6:17 PM

Execution Plan

API Data Sources

Finish API data sources for Brown and Silver layers
Create Gold layer views (Need to work with RAD team)
Migrate all API data sources to Prod environment
Kick off historical data migration

SQL Data Sources

Verify that SQL data migration (SHIR infrastructure is ready and working)
Complete migration of remaining SQL datasets (1 should already be completed). Work with RAD team and let them complete the remaining (3) tables?
Migrate all SQL Data Sources to Prod environment
Kick off historical data migration

Flat File Data Sources

Identify different file types of flat files

a. Start creating POC for each file type, start with the most files of a file type and work your way down
b. Complete remaining flat files of the same type
c. Migrate them to Prod Environment
d. Kick off historical data migration

Repeat steps a - d for the remainder file types of flat files

++

Solution Definition & Design

- **API data sources**
  - Meet with Business SME, request all available artifacts
  - Request APIs patterns, sources and target systems, and technology being used
  - Collate all material and data points
  - API Data Sources
    - Mercury (ServiceNow)
    - OT Datahub
    - FieldGlass
    - M&A Mastertrain (Smartsheet)
    - SEH (Salesforce)
    - Employee Learning & Development
  - Variety, volume, and velocity
    - Initial dataloads and ongoing dataloads
  - Collect future data points
  - Analyze data, document findings and opportunities
  - Document future state
  - Create solution pattern(s) for each unique dataload type
  - Identify tech stack and usage of each tool
  - Build out cloud infrastructure for the pilot prototypes
  - Create pilot prototypes
  - Create detailed solution to populate Bronze, Silver & Gold tables
  - Create detailed estimates of Dev, Test, and Deployment
- **Outlook Email data sources**
  - Meet with Business SME, request all available artifacts
  - Request Outlook Email types, timespan, technology being used, and available APIs
  - Collate all material and data points
  - Outlook Email data sources:
    - CBRE (XLS)
    - Hubstar (XLS)
    - HR/PeopleSfot/AP (XLS)
    - ESG (XLS)
    - HireRight (XLS)
  - Variety, volume, and velocity
    - Initial dataloads and ongoing dataloads
  - Collect future state data points
  - Analyze data, document findings and opportunities
  - Identify initial dataload and ongoing dataloads
  - Create solution pattern(s) for each unique dataload type
  - Identify tech stack and usage of each tool
  - Build out cloud infrastructure for the pilot prototypes
  - Create pilot prototypes
  - Create detailed solution to populate Bronze, Silver & Gold tables
  - Create detailed estimates of Dev, Test, and Deployment
- **SQL/Oracle data sources**
  - Meet with Business SME, request all available artifacts
  - Request database types, data types, data size, and timespan
  - Collate all material and data points
  - SQL/Oracle data sources:
    - Lenel (SQL/Oracle)
    - M&A (SQL/Oracle)
    - HR/PeopleSoft/AP (SQL/Oracle)
    - OneClick (SQL/Oracle)

        - Concur (SQL/Oracle)
        - Cority (SQL/Oracle)
    - Variety, volume, and velocity
        - Initial dataloads and ongoing dataloads
    - Analyze data, document findings and opportunities
    - Identify initial dataload and ongoing dataloads
    - Identify solution pattern(s) for each unique dataload type
    - Search for existing solutions that closely resemble solution patterns identified
    - Make necessary modifications to suit use cases
    - Identify tech stack and usage of each tool
    - Build out cloud infrastructure for the pilot prototypes
    - Create pilot prototypes
    - Create detailed solution to populate Bronze, Silver & Gold tables
    - Create detailed estimates of Dev, Test, and Deployment
- **Internal file data sources**
    - Meet with Business SME, request all available artifacts
    - Request file types, data size, timespan, and hosting technology(s)
    - Collate all material and data points
    - Internal Files data sources:
        - Ariba (Flat file)
        - Noosh (Flat file)
        - AdobeSign (XLS)
        - eGRC (XLS)
    - Variety, volume, and velocity
        - Initial dataloads and ongoing dataloads
    - Analyze data, document findings and opportunities
    - Identify initial dataload and ongoing dataloads
    - Identify solution pattern(s) for each unique dataload type
    - Search for existing solutions that closely resemble solution patterns identified
    - Make necessary modifications to suit use cases
    - Identify tech stack and usage of each tool
    - Build out cloud infrastructure for the pilot prototypes
    - Create pilot prototypes
    - Create detailed solution to populate Bronze, Silver & Gold tables
    - Create detailed estimates of Dev, Test, and Deployment
- **External File data sources**
    - Meet with Business SME, request all available artifacts
    - Request file types, data size, timespan, and hosting technology(s)
    - Collate all material and data points
    - Internal Files data sources:
        - IT Resseler (External file)
    - Variety, volume, and velocity
        - Initial dataloads and ongoing dataloads
    - Analyze data, document findings and opportunities
    - Identify initial dataload and ongoing dataloads
    - Identify solution pattern(s) for each unique dataload type
    - Search for existing solutions that closely resemble solution patterns identified
    - Make necessary modifications to suit use cases
    - Identify tech stack and usage of each tool
    - Build out cloud infrastructure for the pilot prototypes
    - Create pilot prototypes
    - Create detailed solution to populate Bronze, Silver & Gold tables
    - Create detailed estimates of Dev, Test, and Deployment

Define cloud solution patterns, identifying the cloud tech stack, and appropriate usage of each tool
Build out cloud infrastructure to support pilot prototype development
Build pilot prototypes for cloud solution patterns based on tech stack definition
Detailed solution design for data acquisition & data integration to populate Broze, Silver, and Gold data into RAD Data Platf orm for all data sources
Detailed estimation of effort for Dev, Test, and Deployment phase of program

Migrate Data Sources by Type

Monday, September 2, 2024 1:51 PM

API	Outlook Email	SQL/Oracle	Internal Files	External Files
Mercury OT DataHub Field Glass M&A Masterbrain SESH (Salesforce) Employee Learning & Development	CBRE Hubstar HR/Peoplesoft/AP (xls) ESG HireRight	LENEL M&A HR/Peoplesoft/AP (SQL/Oracle) One-Click Concur Cority	Ariba (flat) Noosh (flat) AdobeSign (xls) eGRC (xls)	IT Reseller

Scrum

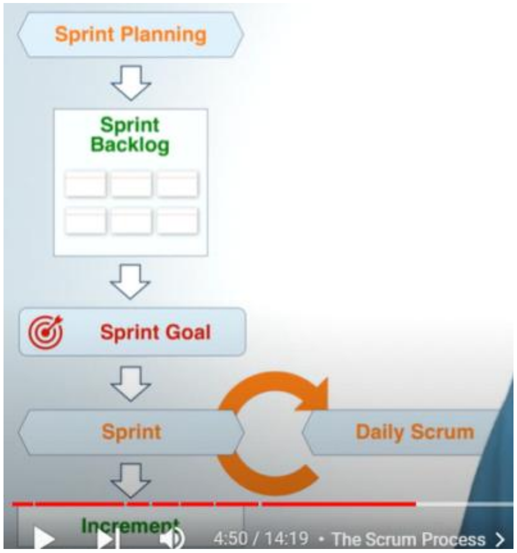
Friday, September 6, 2024 5:08 PM

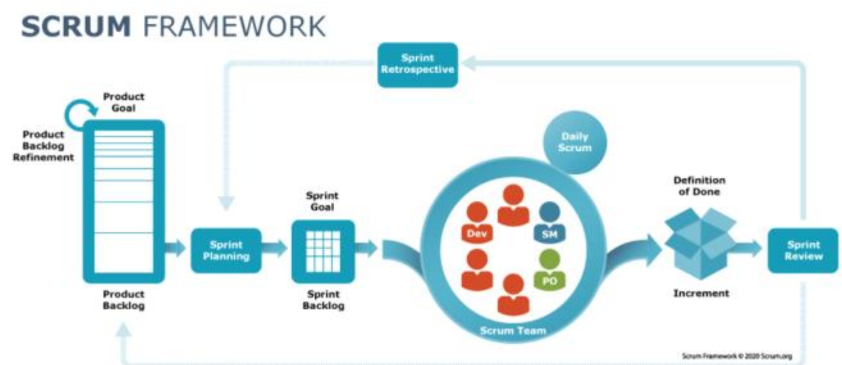
Scrum commitment
- Product goal
- Sprint goal
- Definition of done

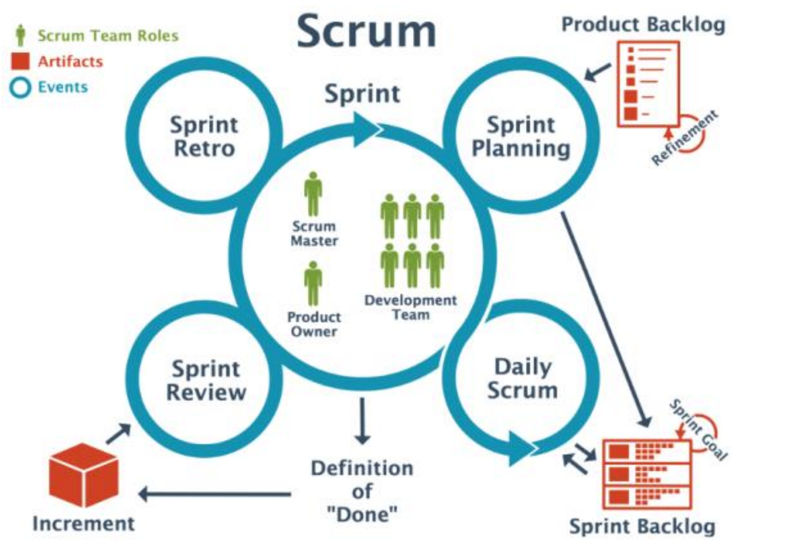
Scrum events
- Sprint planning
- Sprint
- Daily Scrum
- Sprint Review
- Sprint Retrospective

Scrum artifacts
- Product backlog
- Sprint Backlog
- Increment

++



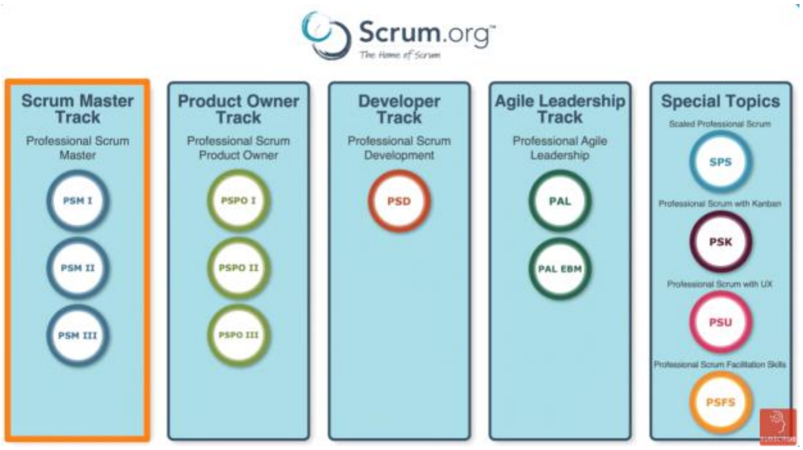


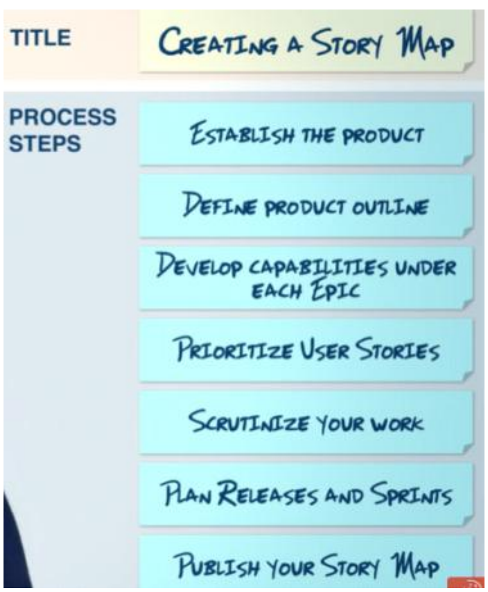

Scrum Team Roles
Artifacts
Events
Scrum
Product Backlog
Sprint
Sprint Retro
Sprint Planning
Refinement
Scrum Master
Product Owner
Development Team
Sprint Review
Daily Scrum
Sprint Goal
Definition of "Done"
Increment
Sprint Backlog


THREE PILLARS OF EMPIRICISM
TRANSPARENCY
INSPECTION
ADAPTATION


THE FIVE VALUES OF SCRUM
COMMITMENT
FOCUS
OPENNESS
RESPECT
COURAGE


Scrum Alliance®
Scrum Master Track
Certified Scrum Master
CSM
Certified ScrumMaster
A-CSM
Advanced Certified ScrumMaster
CSP-SM
Certified Scrum Professional ScrumMaster
Product Owner Track
Certified Scrum Product Owner
CSPO
Certified Scrum Product Owner
A-CSPO
Advanced Certified Scrum Product Owner
CSP-PO
Certified Scrum Professional Product Owner
Developer Track
Certified Scrum Developer
CSD
Certified Scrum Developer
A-CSD
Advanced Certified Scrum Developer
CSP-D
Certified Scrum Professional for Developers
Agile Leadership Track
CAL-E
Certified Agile Leadership Essentials
CAL-T
Certified Agile Leadership for Teams
CAL-O
Certified Agile Leadership for Organizations
CAL
Certified Agile Leadership II
Agile Guide Track
CTC
Certified Team Coach
CEC
Certified Enterprise Coach
CST
Certified Scrum Trainer
ACS-CF
Agile Coaching Skills – Certified Facilitator





Risk Identification

++

Scrum Master Sprint

Key to Success
(Agile in motion organizations)

Know your Organization / Customer

Know your Stakeholders

Know your Team

Connect with

- Product owner
- Engineers
- Stakeholders

Know your Product

- Review Current Sprint
- Review previous 2 to 3 sprints

Current Process

Process Gaps

- Conduct Retrospective meeting

KYC – Know your customer

- Learn about the company, company products, services
- Understand customer goals, vision, Mission
- Understand expectations

KYS – Know your Stakeholders

- Get to know stakeholders
- Stakeholders are within your team, outside of your team & your product consumers
- Be on top of Stakeholders communications

KYT – Know your Team

- Knowing your team is the most important thing for your journey in the organization
- It is your team who can set you and organization to succeed
- Set up 1:1 with product owner
- Set up Team Introduction session
- Set up 1:1 with engineers

KYP – Know your Product

- Unbox the product you are working on
- Learn about product benefits and value
- Know the product consumers / customers / end users

Presentation

Prepare a presentation describing

- Current state
- Process
- Proposed /future state
- Incremental changes
- Gaps
- Delivery
- Execution plan

KPI's to achieve in upcoming 3 to 4 sprints time frame

- Process adaption rate
- Healthy work environment
- High quality
- Building a self – organized team
- Velocity growth

Stakeholders

Present you stakeholders your plan of action

Engineering team

Engage with Engineering team

- Conduct sessions to discuss the proposed changes
- Present data supporting you proposed changes
- Discuss Execution plan
  - Changes incrementally
    - Immediate
    - Fast follow up
    - Time taking

++

Agile Mistakes to avoid:

1. Don't always try to help the team solve a problem.

2. Don't overload the team with work. A team under pressure will take short cuts

3. Don't commit your team to requests or deadlines without first consulting with them.

4. Don't dominate the daily scrum. These meetings are more for the development team, so a Scrum

5. Don't act like a mother hen and treat the team like children. If the team come

to rely on the Scrum Master, then they will again lose their self-organising ability.

6. Don't act like the team secretary... It is not the Scrum Master's responsibility

7. Don't deal with all impediments personally when any member of the team could deal with them.

8. Don't filter out any negative feedback from stakeholders by restricting or completely blocking

9. Don't tolerate damaging behaviour from a team member hoping that it will improve with

10. Don't prevent the team from failure,

11. single point of contact . The Scrum Master is often used as the sole point of contact

++

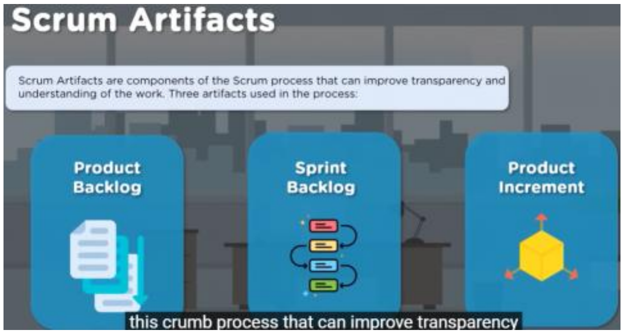
https://www.youtube.com/watch?v=SyIN_YMfoQs

- Scum is a framework that helps teams collaborate when it comes to complex project/ products
- It helps teams learn from their experiences, organize themselves while working on problems, reflect on the victories and losses to help them improve

- Scrum is a framework that enables teams to work together
- With Scrum, teams can:
  - Learn from experiences
  - Self-organize working on problems
  - To reflect on their victories
  - And their losses to improve

The Scrum master helps teams learn and apply scrum to obtain business value. He/ she helps remove impediments, protects them from interference and helps the team to adopt agile practices

The Scrum team is a collection of individuals that work together to deliver the requirements of the stakeholders



Product Backlog

The product backlog consists of a list of
- new features
- changes to made to existing features
- bug fixes,
- changes to the infrastructure and several other activities

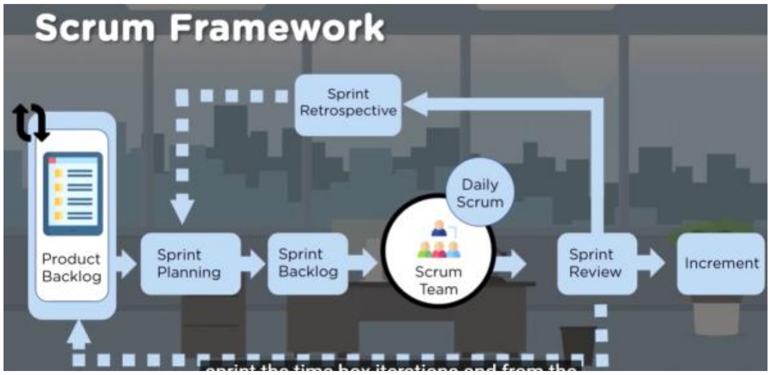
that the team needs to deliver to ensure a particular outcome

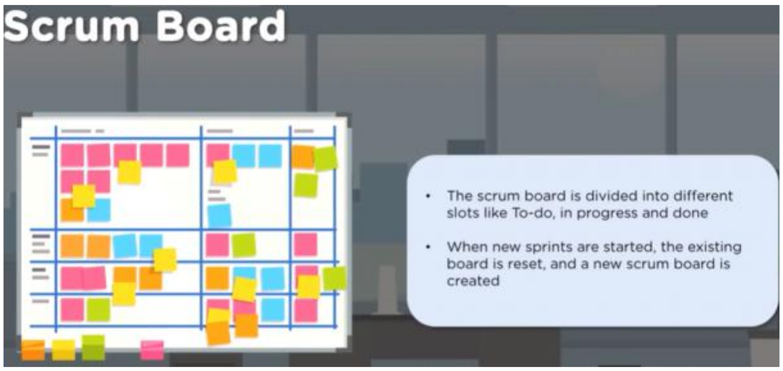
Sprint Backlog

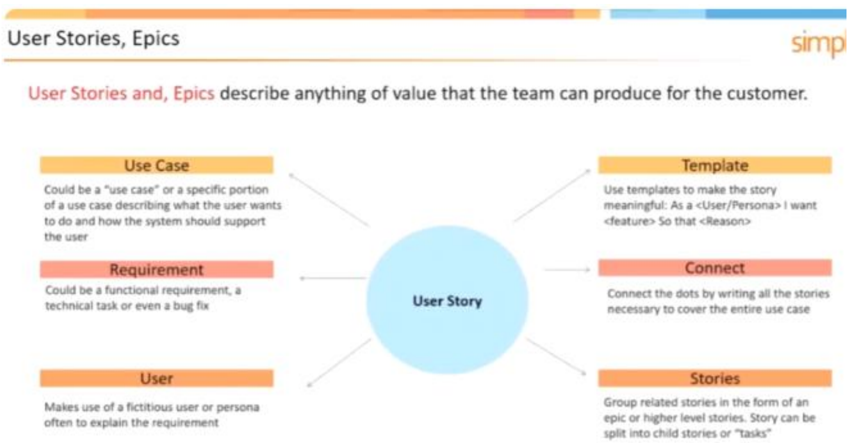
A subset of the product backlog, the sprint backlog contains tasks that the team aims to complete to satisfy the sprint goal
The team first identifies tasks from the product backlog that need to be delivered to achieve a goal. These tasks are then added to the sprint backlog

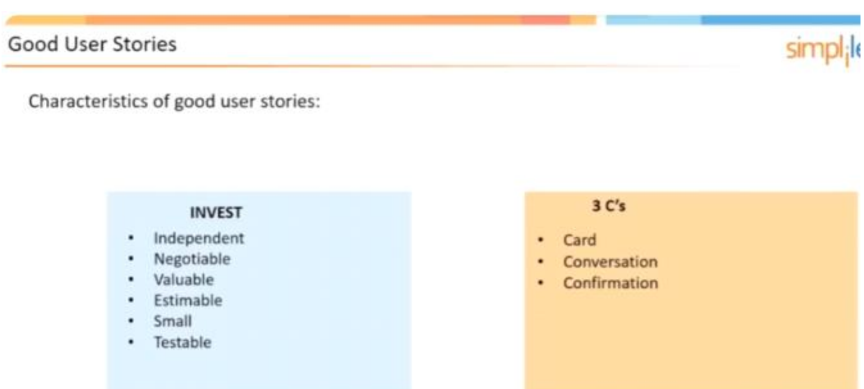
Product Increment

- The product increment is the combination of all product backlog tasks completed in a sprint and the value of increments of previous sprints
- The outcome should be in usable condition, even if the product owner doesn't decide to release it


Scrum Framework
Sprint Retrospective
Daily Scrum
Product Backlog
Sprint Planning
Sprint Backlog
Scrum Team
Sprint Review
Increment

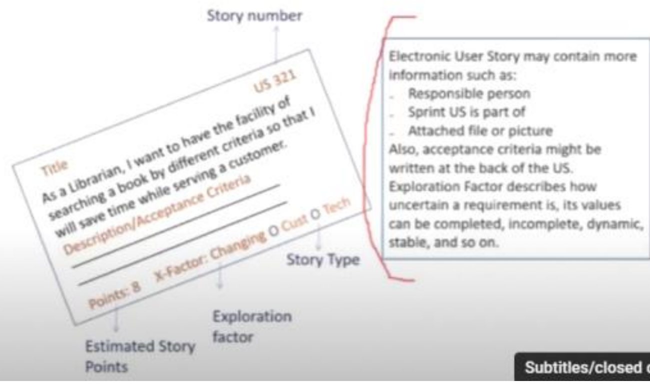

Scrum Board
• The scrum board is divided into different slots like To-do, in progress and done
• When new sprints are started, the existing board is reset, and a new scrum board is created


User Stories, Epics
User Stories and, Epics describe anything of value that the team can produce for the customer.
Use Case
Could be a "use case" or a specific portion of a use case describing what the user wants to do and how the system should support the user
Requirement
Could be a functional requirement, a technical task or even a bug fix
User
Makes use of a fictitious user or persona often to explain the requirement
User Story
Template
Use templates to make the story meaningful: As a <User/Persona> I want <feature> So that <Reason>
Connect
Connect the dots by writing all the stories necessary to cover the entire use case
Stories
Group related stories in the form of an epic or higher level stories. Story can be split into child stories or "tasks"


Good User Stories
Characteristics of good user stories:
INVEST
• Independent
• Negotiable
• Valuable
• Estimable
• Small
• Testable
3 C's
• Card
• Conversation
• Confirmation

Story Card

Here is an example of a Story Card:



Splitting Stories

Stories need to be split if they are too large, won't fit into the Sprint, or for more reliable estimation.

Splitting on operational boundaries

- As an operator, one needs to manage reservations which could be split.
- As an operator, one needs to make a reservation, cancel it, or modify it.

Separate exceptions or cross-cutting concerns

- In the beginning, develop only the main path, like Accept repayment of a loan.
- Then, address the exceptions, say, If one pays excess, then process a refund.
- Also, add on other concerns like check access restrictions or record name of the operator.

Splitting across data boundaries

- As an accountant, one needs to enter balance sheet information which could be split.
- As an accountant, one needs to enter summary information with select heads only.
- As an accountant, one needs to enter receivable details with more granular data.

Determining Value or Return on Investment (ROI)

There are four ways in which work will provide value:

New Revenue	Winning new customers who will bring in additional revenue
Retained Revenue	Stop losing customers who would otherwise leave the product or service
Incremental Revenue	Getting existing customers to buy new or more of existing services or products
Operational Efficiency	Reduce the cost of development or operation

Prioritization Models

Before we begin planning, the backlog must be prioritized.

Prioritization Models

Value-Based Prioritization

- High Value
- High Risk followed by High Value
- Low Risk followed by Low Value
- Low Risk

Kano Model

- Mandatory/Threshold
- Linear
- Exciters/Delighters

Karl Wiegers Relative Weights Method

- Calculate Value = Benefit + Penalty
- Priority = Value / (Cost + Cost Benefit + Risk + Risk Benefit)

Velocity

Velocity of the Team helps understand the productivity of the Team.

1 Definition

Velocity is the amount of work that can be completed in a Sprint. It is the sum total of the size of completed stories in a given Sprint.

2 Uses

Velocity is useful to answer questions like:

- How many stories should we accept in a given Sprint?
- How many Sprints do we need to complete a set of stories?

3 Example

A team completed 5 stories of sizes 5, 3, 8, 13, and 2. Two stories of size 5 were half complete. What is the velocity of the team?
Ans. Velocity = 5 + 3 + 8 + 13 + 2 = 31

Planning Onion

Planning happens at multiple levels:

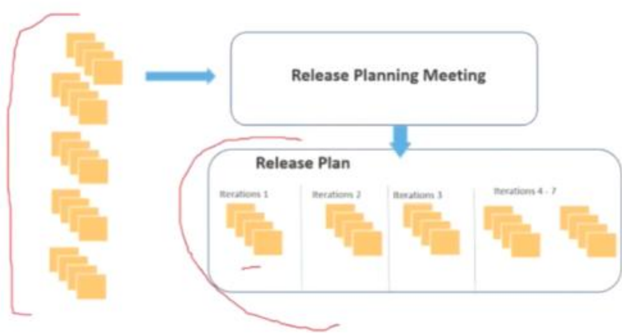
Roadmap/Release Planning

Prioritize high level epics and determine goals of releases.

- Establish goals of releases based on market demand, regulatory needs, or customer expectations.
- For each release:
  - Estimate the target stories.
  - Repeat until target stories are assigned:
    - Select an iteration length
    - Estimate velocity
    - Assign stories to iteration
  - Iterate until the stories and release date meet conditions of satisfaction.
  - Try not to pack too much into a release backlog.

Output of Release Planning

At the end of release planning, this is what you will see:



Estimation

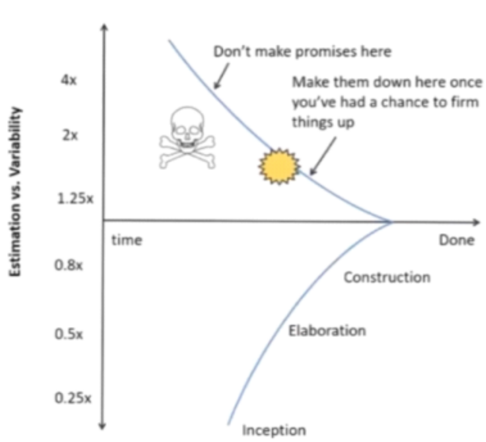
Principles behind Estimation:

Estimation

- Understand the cone of uncertainty, which is an estimate or best guess. Being accurate is neither possible nor advisable; being predictable is good enough.
- Overestimation and underestimation is present in a process; however, underestimation is more prevalent and more dangerous.
- The only estimate that matters is the one given by the Team. "Nothing is impossible for the man who does not have to do it himself." — A.H.Weiler.
- Scrum estimation is not necessarily more difficult; however, Scrum exposes bad estimates sooner, which is a good thing.

Cone of Uncertainty

Variability in estimates will narrow down as project progresses.



Estimation

Principles behind Estimation:

Estimation

- Understand the cone of uncertainty, which is an estimate or best guess. Being accurate is neither possible nor advisable; being predictable is good enough.
- Overestimation and underestimation is present in a process; however, underestimation is more prevalent and more dangerous.
- The only estimate that matters is the one given by the Team. "Nothing is impossible for the man who does not have to do it himself." — A.H.Weiler.
- Scrum estimation is not necessarily more difficult; however, Scrum exposes bad estimates sooner, which is a good thing.

Ideal Time

Ideal Time is the amount of actual work time.

- The amount of time it takes to complete a piece of work
- Converting ideal to elapsed time depends on who is doing the work, the extent of distraction present, possibility of multiple assignment, and task switching.

- Estimation is challenging because it is tough to estimate distractions
- Only considers actual work time, not the distractions

Story Points

Absolute measure of size; relative to each other. Story Point is an analogous estimation technique.

- Establish a benchmark and compare others to it:
  - Ideally, you should have more than one benchmark or triangulation like Small may be 1, Medium may be 5, and Large may be 13.
  - Benchmark works because human beings are better at comparisons than absolutes.
- Points are for the entire work.
- Use non-linear scale to pick values:
  - Modified Fibonacci: 1, 2, 3, 5, 8, 13, 20, ...
  - Doubling scale: 1, 2, 4, 8, 16, ...

Comparing Ideal Time to Story Points

Comparison between two units of size:

Ideal Time	Story Points
Uses time units like hours or days and is easier to explain outside the team	Unitless and harder to explain outside the team
Everybody's estimate may be different	Higher chances of converging estimates
Easier to do in the beginning	Faster to do, once understood
Refines understanding of wasted time	Not comparable across teams

Planning Poker

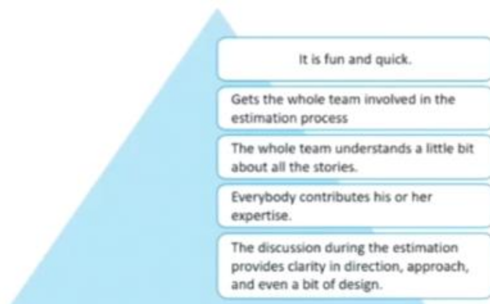
A fast, fun team approach to estimation.

1. Select the team for estimation; each gets a deck of cards with numbers printed on them.
2. Product Owner reads and quickly explains a story.
3. Team discusses the story and each member picks the card that matches his or her estimate. Each one chooses independently and does not display to others.
4. Once the Scrum Master prompts, everybody's estimates will be revealed.
5. It is expected that the numbers will vary; discuss the variations and choose cards again.
6. Repeat until estimates converge or the Team agrees on an estimate.

! For more information, check the link: https://www.planningpoker.com

Advantages of Planning Poker

Planning Poker works because:

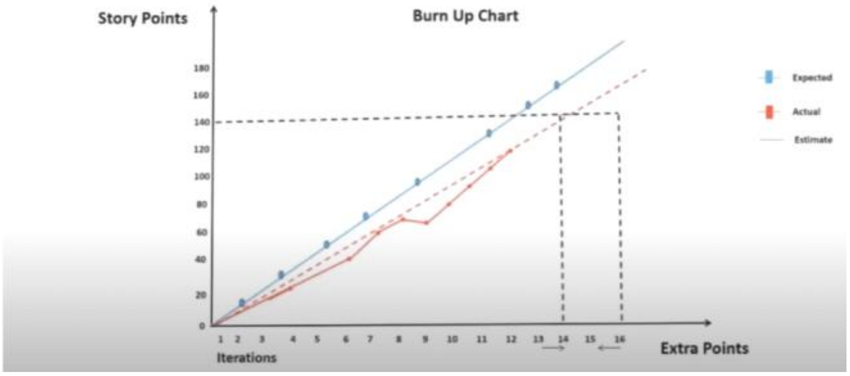


Physical Progress Chart

The simplest form of the Kanban Board:

Staying in Control

Use Burn Up and Burn Down Charts for forecasting.



Scrum Values

C-FORC abbreviates the values of Scrum.

Scrum Life Cycle

Here is a representation of the project life cycle as practised in Scrum:

Sprint

Sprint is synonymous with iteration.

Sprint

- Factors affecting Sprint duration are stability of the backlog and cost of iterating.
- Goal of Sprint is working software. The end product should be near releasable or potentially shippable.
- Sprint duration and deliverables do not change once the team has committed. Sprint can be cancelled in extreme circumstances.
- Sprint begins with planning and ends with review and retrospective.

Sprint Planning

Sprint Planning helps decide the tasks that the team has committed to in order to achieve the Sprint goal. The team should make an informed commitment about what they will deliver. A typical Sprint Planning Meeting takes anywhere between one and two hours per week of Sprint.

Sprint Planning

- Conducted right at the beginning of the Sprint.
- Attended by: Team, Product Owner, Scrum master
- Two possible approaches: Commitment- Driven, Velocity -Driven

Team buy-in is important:
- The goal should be clearly understood by all. Clearly articulated "Definition of Done."
- The entire team should agree with the goal.
- The goal should be realistic and achievable.

Daily Scrum/Standup Meeting

The Daily Scrum is a meeting of the team, by the team, and for the team.

Structure

Entire Team attends the meeting.

Duration

The meeting happens every day. The duration of the meeting is 15 minutes or less.

Agenda

The meeting has 3-point updates:
- Tasks done yesterday
- Tasks to be done today
- The blocking issues

Sprint Retrospective

simplilearn

Sprint Retrospective is a continuous improvement mechanism in a Scrum Team. It helps discuss what is working, what is not, and what could be better.

Structure

Attended by: Team, Scrum Master, or External Facilitator

Duration

The duration of the meeting is 1 to 2 hours after Sprint Review.

Agenda

Purpose:

- Makes issues visible
- Come up with improvement opportunities
- Gives the team ownership of actions

Backlog

simplilearn

Backlog item refers to any work done by the team that adds value to the customer.

1. **Product Backlog**

   The Product Backlog is the collection of all the items that will make the product better.

2. **Release Backlog**

   The Release Backlog is a subset of the Product Backlog which contains items that are targeted for the current release.

3. **Sprint Backlog**

   The Sprint Backlog is a subset of the Release Backlog that is targeted at the current sprint.

Product Backlog

Product Backlog items may be feature requests, bug fixes, technical tasks, and so on.

Product Backlog New story

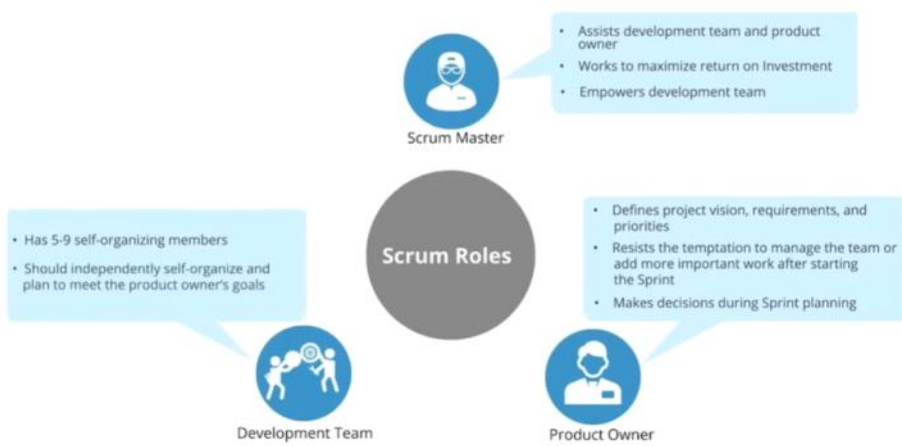
253	List tasks with no parent stories under quasi-story	New	
135	Indicate tracker type via issue id's bgcolor	New	2
145	Download pretty-fied roadmap report	New	6
55	Show/hide closed items or subtasks	New	2
40	Show/hide the issue #	New	1
135	Show/hide closed the main backlog	New	0
21	Show warning when server connection is lost	New	4
223	Support versions from parent projects	Feedback	0
213	Hint text on "talk" tab doesn't display in Safari	New	0
214	"Talk" button often fails	New	0
137	Append an email to the item's discussion section	New	6
224	Spreadsheet import for stories/tasks	New	0
225	Offer to copy tasks when copying a story	New	0
226	Time log widget	New	0

Definition of Done

"Definition of Done" is a checklist of items to be verified before marking something as Done.

- Usually prepared by the Scrum Master in consultation with the Team.
- For example, a user story would be called Done if:
  - The story is fully implemented or code completed as described.
  - Automated unit tests have been developed with at least 80% code coverage.
  - Automated unit tests and the acceptance tests in the story are passing; no Severity has one or two defects.
  - High priority test cases have been automated and added to the regression suite.
- Evolves as the team maturity increases.

Scrum Roles



Scrum Master: Roles and Commitments

A Scrum Master is:

- A "servant leader"
- A facilitator of Scrum events
- An impediment remover
- A process "coach"

A Scrum Master is not:

- A "line" manager
- A task Master
- A technical or design authority
- A decision maker

Attributes of a Scrum Master

Attributes of Scrum Master:

Be a Scrum Master

Tasks for the Scrum Master

Responsibilities of Scrum Master:

- Be the servant leader
- Ensure smooth flow of information
- Should be problem solver
- Should be the protector and peacemaker of the team
- Should be the process coach of the team and ensure that Agile principles are followed

Attributes of a Scrum Team

Desirable characteristics

- Small and Nimble
- Self-sufficient and Cross-functional
- Autonomous and Self-organizing

Forming Scrum Teams

Considerable factors while deciding Scrum team:

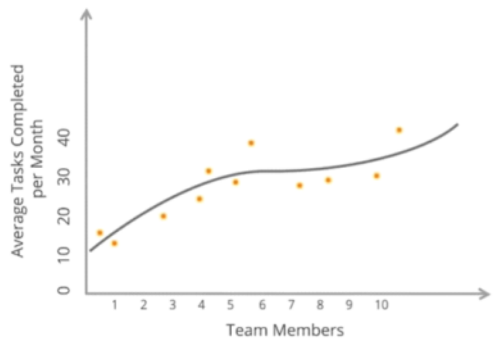
- Feature teams over Component teams

Forming Scrum Teams

Considerable factors while deciding Scrum team:

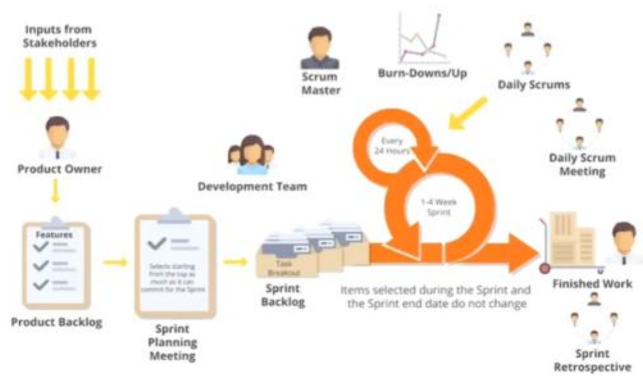
- Feature teams over Component teams
- Assemble the right people
- Distributed team

Team Size and Productivity

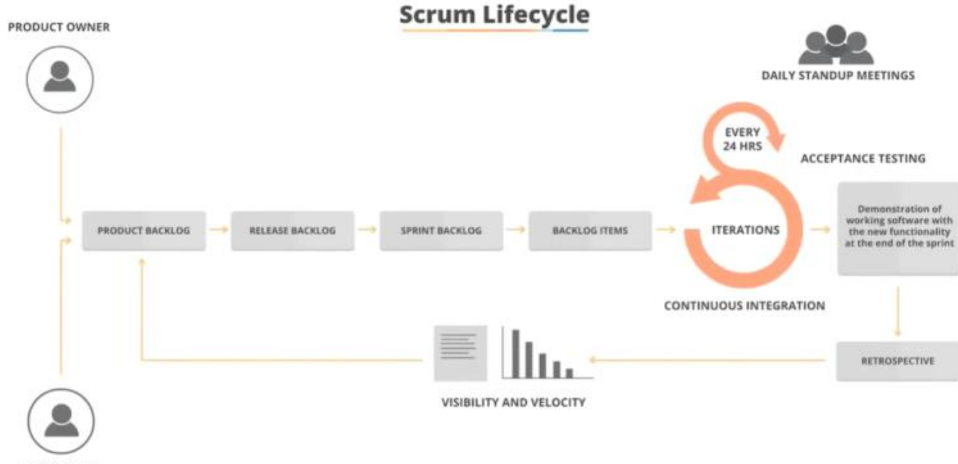


Scrum Definition

A framework within which various processes, techniques, and practices are employed



Scrum Lifecycle



Sprint Planning

- Time-boxed at 2 hours per week
- All Scrum team members should attend
- Team's capacity and defining DONE is important
- Select user stories on:
  1. Team velocity
  2. Commitment

Daily Scrum

- Time-boxed at 15 minutes – fixed
- All Scrum team members should attend
- Questions for each team member:
  1. What did I do yesterday?
  2. What am I going to do today?
  2. What are my impediments?

Sprint Review

- Time-boxed at one hour per week
- All Scrum team members including interested stakeholders should attend
- Purpose is to demonstrate the working software and obtain feedback

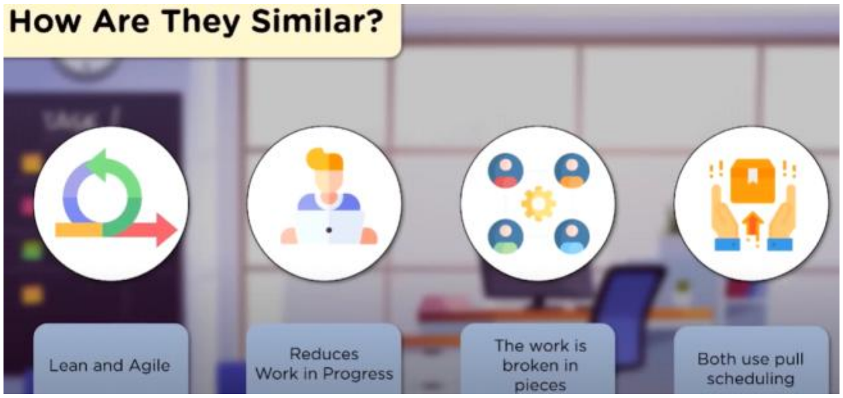
Definition of Done

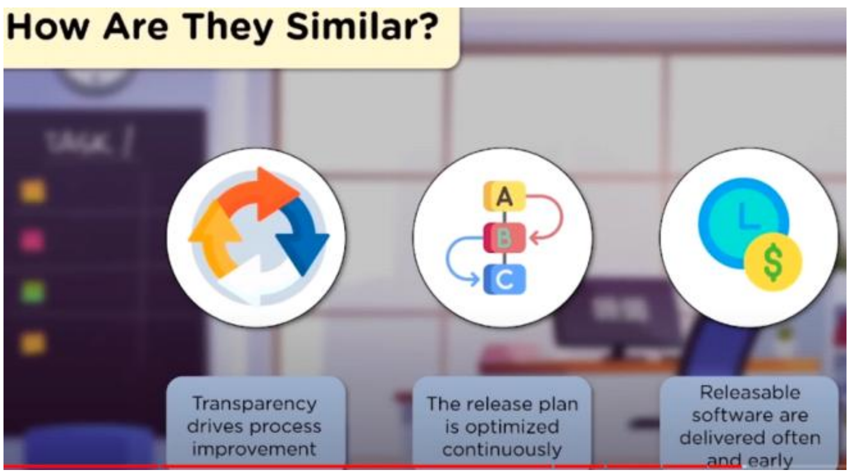
A user story would be called "Done" if:

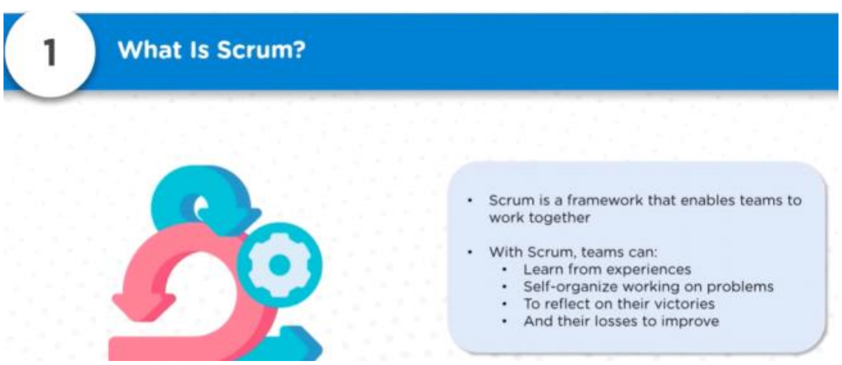
- The story is fully implemented or code completed as described
- Automated unit tests have been developed with at least 80% code coverage
- Automated unit tests and acceptance tests in the story are passing
- High priority test cases have been automated and added to the regression suite

NOTE: Each team's "Definition of Done" varies as the team matures

- Kanban or 'Kamban' in Japanese, which translates to 'signboard' is visual system that is used for managing work as it goes through the process
- Kanban can help determine bottlenecks, and then fix them in a cost-effective manner and optimal speeds. This is done with the help of a Kanban board.







2 Define the Roles in Scrum.

Scrum Master

The Scrum master helps teams learn and apply scrum to obtain business value. He/she helps remove impediments, protects them from interference and helps the team to adopt agile practices

2 Define the Roles in Scrum.

Scrum Team

The Scrum team is a collection of individuals that work together to deliver the requirements of the stakeholders

3 What are the Responsibilities of the Scrum Team?

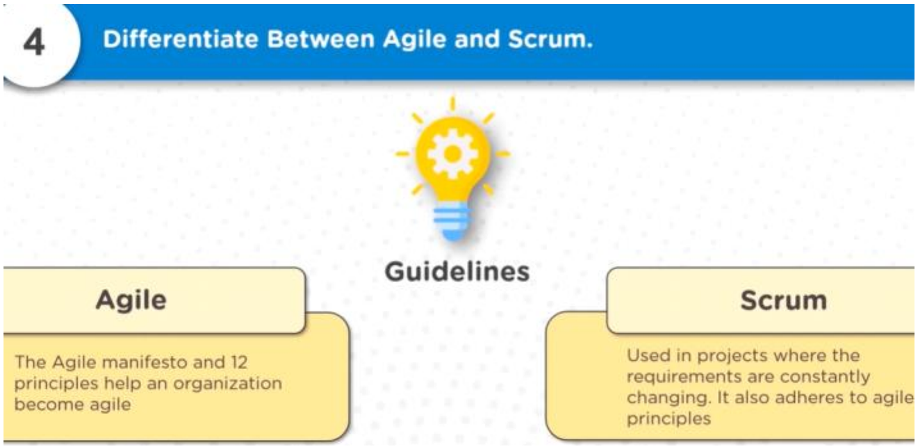
The Scrum team is a self-organizing team that has five to seven members. The team's responsibilities are:

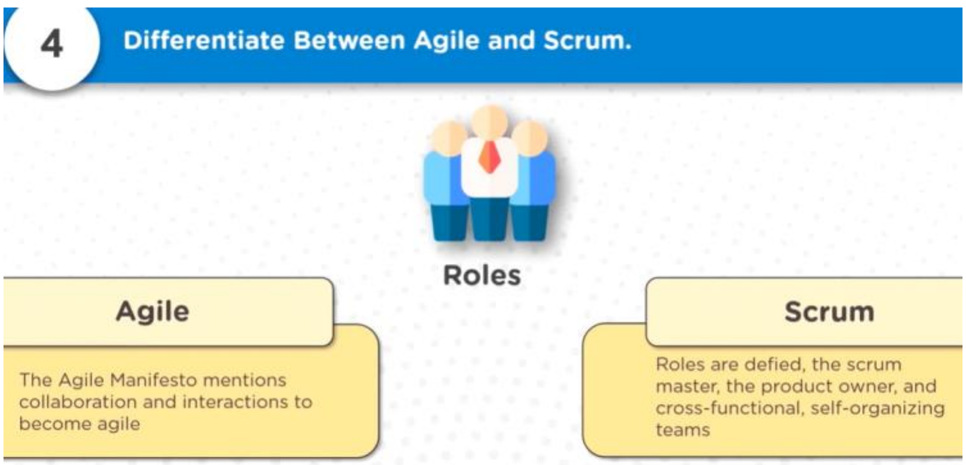
- The development and delivery of working products during each sprint
- To demonstrate ownership and transparency for the work assigned to them
- Ensuring the success of the daily scrum meeting by providing crisp and correct information
- Collaborate with the team and self

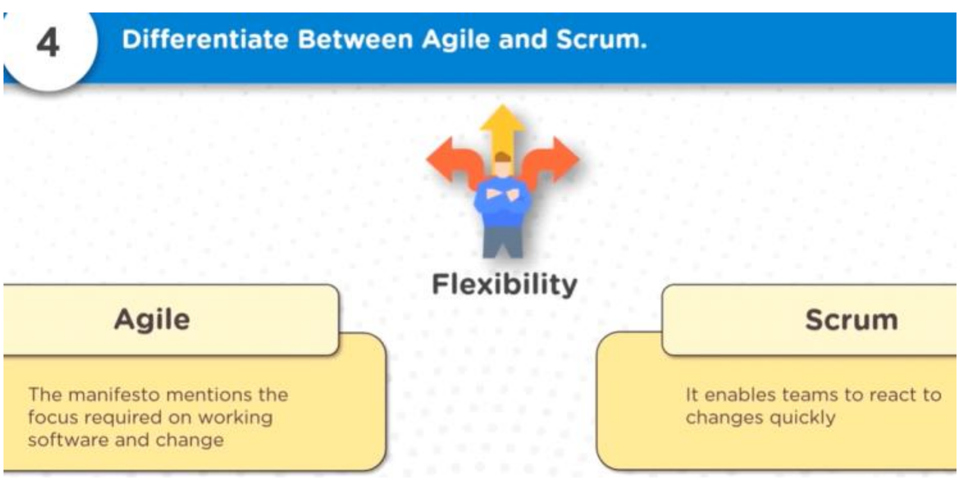
4 Differentiate Between Agile and Scrum.

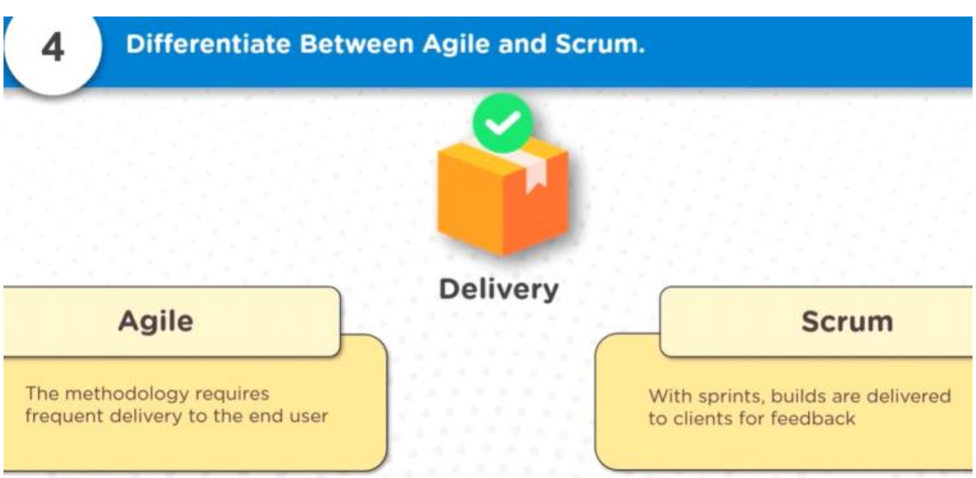
What Is It?

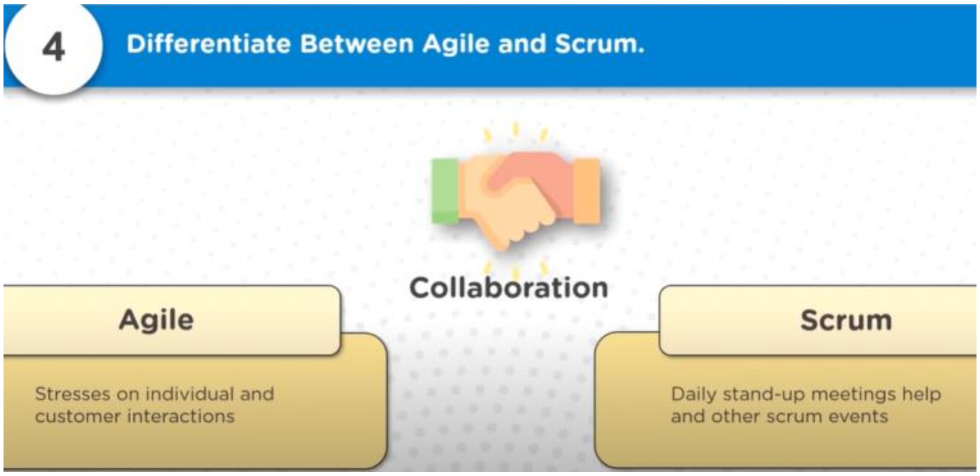
Agile	Scrum
Agile is a set of principles that's iterative and incremental in nature	Scrum is an implementation of the agile methodology

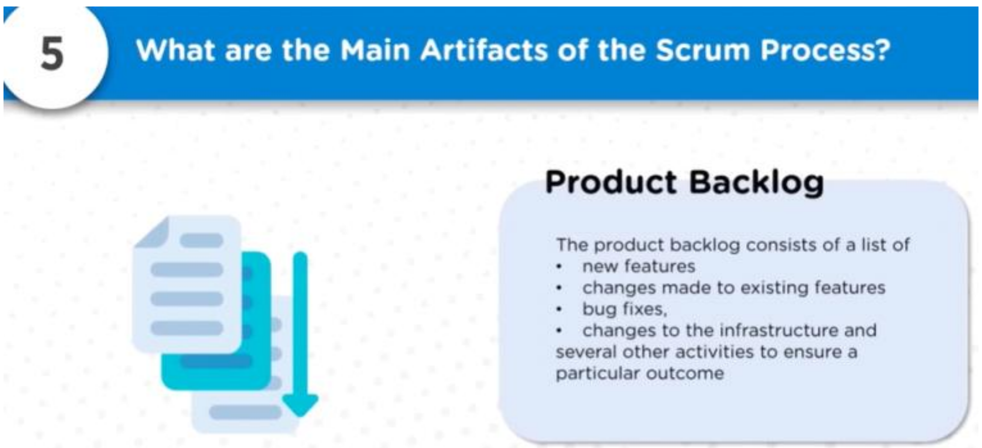

4
Differentiate Between Agile and Scrum.
Guidelines
Agile
The Agile manifesto and 12 principles help an organization become agile
Scrum
Used in projects where the requirements are constantly changing. It also adheres to agile principles

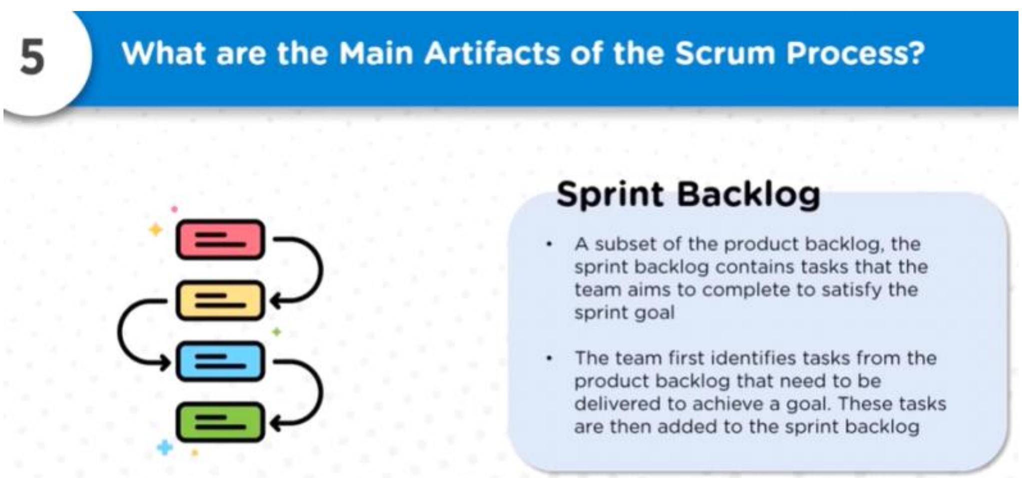

4
Differentiate Between Agile and Scrum.
Roles
Agile
The Agile Manifesto mentions collaboration and interactions to become agile
Scrum
Roles are defied, the scrum master, the product owner, and cross-functional, self-organizing teams


4
Differentiate Between Agile and Scrum.
Flexibility
Agile
The manifesto mentions the focus required on working software and change
Scrum
It enables teams to react to changes quickly


4
Differentiate Between Agile and Scrum.
Delivery
Agile
The methodology requires frequent delivery to the end user
Scrum
With sprints, builds are delivered to clients for feedback


4
Differentiate Between Agile and Scrum.
Collaboration
Agile
Stresses on individual and customer interactions
Scrum
Daily stand-up meetings help and other scrum events


5
What are the Main Artifacts of the Scrum Process?
Product Backlog
The product backlog consists of a list of
• new features
• changes made to existing features
• bug fixes,
• changes to the infrastructure and several other activities to ensure a particular outcome


5
What are the Main Artifacts of the Scrum Process?
Sprint Backlog
• A subset of the product backlog, the sprint backlog contains tasks that the team aims to complete to satisfy the sprint goal
• The team first identifies tasks from the product backlog that need to be delivered to achieve a goal. These tasks are then added to the sprint backlog

5 What are the Main Artifacts of the Scrum Process?

Product Increment

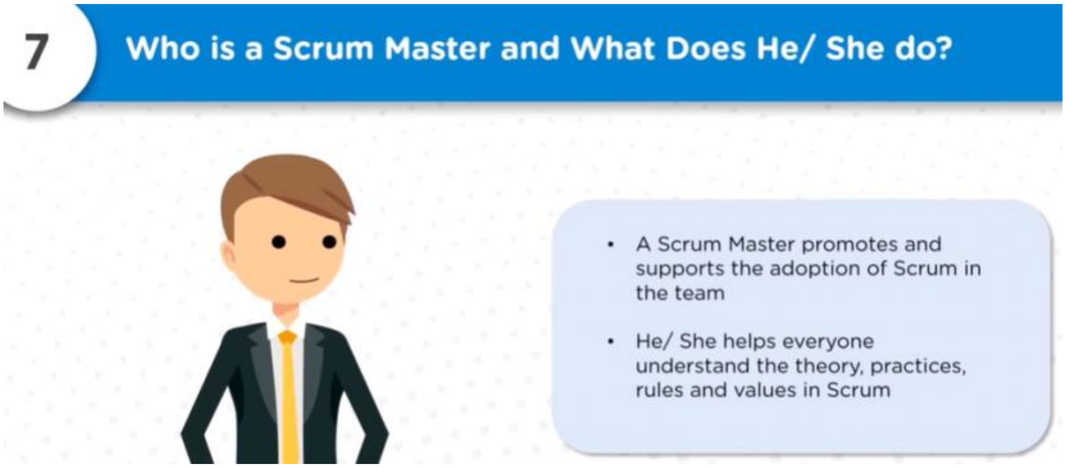
- The product increment is the combination of all product backlog tasks completed in a sprint and the value of increments of previous sprints
- The outcome should be in usable condition, even if the product owner doesn't decide to release it

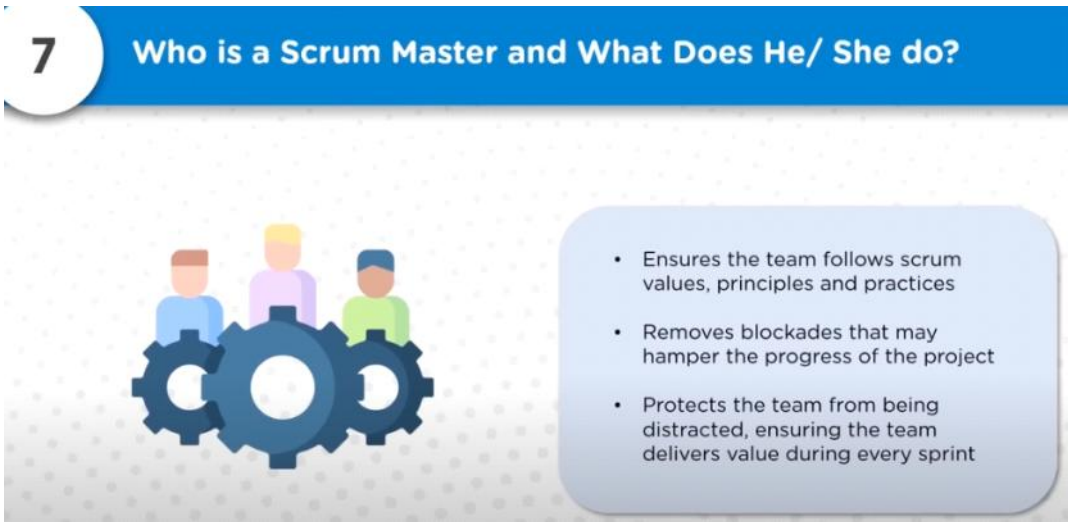
6 How are the Product and Sprint Backlog Different from One Another?

Product Backlog	Sprint Backlog
It is a list of items that need to be completed for developing the product	It is a list of items to be completed during each sprint
The backlog is collected from the customer by the product owner and assigned to the team	The team collects the backlog from the product owner and sets up the time frame for the sprint
It is specific to the end goal	It is specific to a sprint

6 How are the Product and Sprint Backlog Different from One Another?

Product Backlog	Sprint Backlog
Based on customer vision	Will vary based on the product vision defined by the product owner
Independent of the sprint backlog	Dependent on the product backlog
Until the project is complete, the product owner maintains the backlog	Each new sprint has backlogs added by the team


7
Who is a Scrum Master and What Does He/ She do?
• A Scrum Master promotes and supports the adoption of Scrum in the team
• He/ She helps everyone understand the theory, practices, rules and values in Scrum


7
Who is a Scrum Master and What Does He/ She do?
• Ensures the team follows scrum values, principles and practices
• Removes blockades that may hamper the progress of the project
• Protects the team from being distracted, ensuring the team delivers value during every sprint


8
What Happens in Daily Stand-up Sessions?
• Stand-up sessions are daily discussions that take place and are usually 15 minutes long
• The objective is to understand
• What went well
• What tasks were completed from the last stand-up session
• What tasks were pending, and
• The obstacles the team was facing

8 What Happens in Daily Stand-up Sessions?

- The objective of the meeting is to understand the overall scope and status of the project
- Further individual discussion can be taken up after the daily stand-up sessions

9 What Is Scrum-ban?

- Scrum-ban is a development methodology that's a combination of Scrum and Kanban
- Scrum-ban can be used to meet the needs of a team that aims to minimize the batching of work and to adopt a pull-based system

10 What is Sprint 0 and Spike?

- Sprint 0 refers to the short amount of effort put for the creation of vision, a rough skeleton of the product backlog
- It also includes insights towards estimating the release of products

10 What is Sprint 0 and Spike?

- The Spike is a set of activities used in extreme programming (XP) involving research, design, investigation, making POCs, etc.
- The objective of the spike is to reduce the risk of the technical approach, by gaining knowledge which helps understand requirements and to improve reliability

11 What is the 'Scrum of Scrums'?

- The 'Scrum of Scrums' is a terminology used for scaled agile techniques which is required to control and collaborate with multiple scrum teams
- It is best used in situations where teams are working together on complex assignments

11 What is the 'Scrum of Scrums'?

It is used to establish the required transparency, collaboration, adaption and adoption to the required scale, to ensure the products can be developed and delivered

12 What is User-Story Mapping?

User story mapping represents and arranges user stories that help with understanding the system's functionalities, the system's backlog, planning releases and providing value to customers

14 What is Empirical Process Control in Scrum?

- Empiricism refers to work based on facts, experiences, evidences, observation and experimentation
- Empirical process control is established and followed in Scrum, ensuring project progress and project interpretation is based on facts of observations

14 What is Empirical Process Control in Scrum?

- These rely on transparency, observation and adaption.
- The mind-set of the team and the shift required in the thought process and culture are very essential to accomplish the agility required by the organization

15 What are Some Drawbacks of Using Scrum?

- Scrum requires individuals who have experience, without whom the project risks failure
- Scrum requires teams to be collaborative and committed to ensure the required results

15 What are Some Drawbacks of Using Scrum?

- A less experienced Scrum master can cause the failure of the project
- If tasks are not well defined, the project can have many inaccuracies
- Scrum works better for smaller projects, and would be difficult to scale to large complex projects

16 What are the Skills of a Scrum Master?

A strong understanding of Scrum and Agile concepts

Fine tuned organizational skills

16 What are the Skills of a Scrum Master?

Being familiar with the technology that the team uses

The ability to coach and teach the team to follow Scrum


16
What are the Skills of a Scrum Master?
The ability to handle conflicts and resolve them quickly
To be a servant leader


17
How can Discord be Dealt With In the Scrum Team?
• The root cause of the issue needs to be identified and addressed
• Complete ownership needs to be established
• Try to diffuse the disagreement
• Emphasize on focus areas that complements the project


17
How can Discord be Dealt With In the Scrum Team?
• The root cause of the issue needs to be identified and addressed
• Complete ownership needs to be established
• Try to diffuse the disagreement
• Emphasize on focus areas that complements the project

18 What Is a User Story?

- A user story doesn't go into detail. It just mentions how a certain type of work will bring value to the end-user
- The end-user, in this case, could mean external end-users or even internal customer or colleagues within the organization

18 What Is a User Story?

- User stories also form the building block of agile frameworks like epics and initiatives
- These help ensure that the teams work towards the goals of the organization through epics and initiatives

19 How are User Stories, Epics and Tasks Different ?

User Stories

User Stories provides the team with simple explanations of the business' requirements created from the end-user's perspective

Epics

An epic is a collection of related user stories. They are usually large and complex

Tasks

Tasks are used to further break down user stories. They're the smallest unit in Scrum used to track work. A person or a team of two people usually work on a task

20 What Is a Sprint?

- Sprint is the terminology used in Scrum, which refers to a time-boxed iteration
- The creation of the specific module or feature, which is part of a product, is done during a Sprint
- The duration of Sprint varies between a week or two

21 What Is Velocity?

- Velocity is a metric used in Scrum that is a measurement of the amount of work that is completed by team during a sprint
- It refers to the number of user stories completed in a single sprint

22 What are the Responsibilities of a Product Owner?

- They define the vision of the project
- To anticipate the needs of the customers and to create appropriate user stories
- To evaluate the progress of the project
- To be a liaison for all product related questions

23 What Is a Burnup and Burndown Chart?

A burnup chart is a tool that's used to track the amount of work that's been completed, and to show the total amount of work the needs to be done for a sprint/ project

23 What Is a Burnup and Burndown Chart?

- A burndown chart is a representation of how fast the team is working through user stories
- It shows total effort against the amount of work for each iteration

24 How Is Estimation In a Scrum Project Done?

- The estimation of user stories is done based on its difficulty
- A particular scale is used to assess the difficulty of the user stories. Some type of scales are:
  - Numeric Sizing (1 - 10)
  - T-shirt Sizes (S, M, L, XL...)
  - Fibonacci Series (1, 2, 3, 5, 8...)
  - Dog breeds (Great Dane, Chihuahua...)

25 What are Some Risks in Scrum and How are they Handled?

Some types of risks in Scrum are:

- Budget: The risk of exceeding budgets
- People (team): Team members need to be of appropriate skill and capability
- Sprint (duration and deliverables): Exceeding the duration, addition of scope
- Product (user stories, epics): Having ill defined user stories and epics
- Knowledge and capability: Not having the appropriate resources

25 What are Some Risks in Scrum and How are they Handled?

- Risk management involves identification, assessment, analysis, defining & implementation of risk responses, monitoring and management of risks
- These are done on a continual basis right from the start of the project till completion. It is essential to understand that the impact of the risk is based on the proximity of the actual occurrence of the risk

26 What are Some Popular Tools Used in Scrum?

GitHub

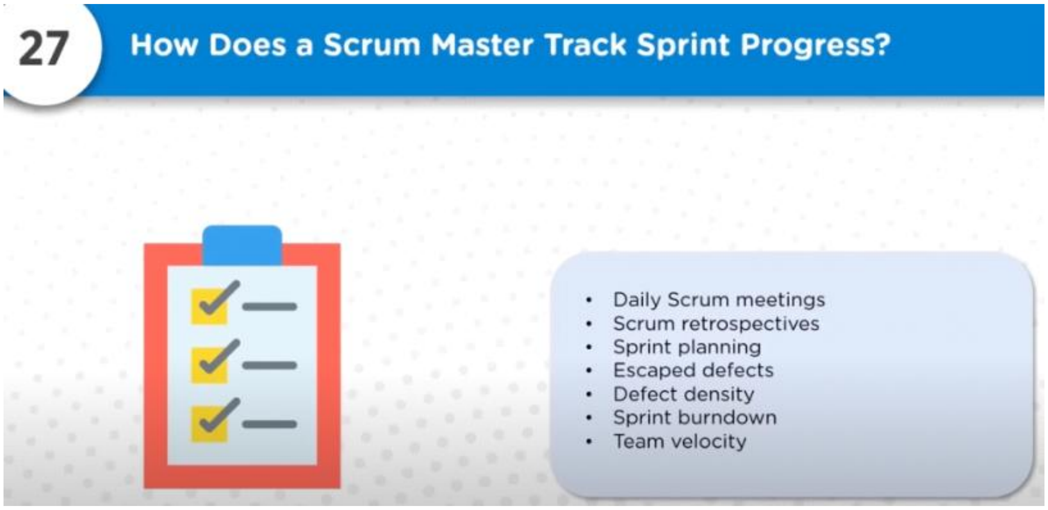
ZOHO

Trello

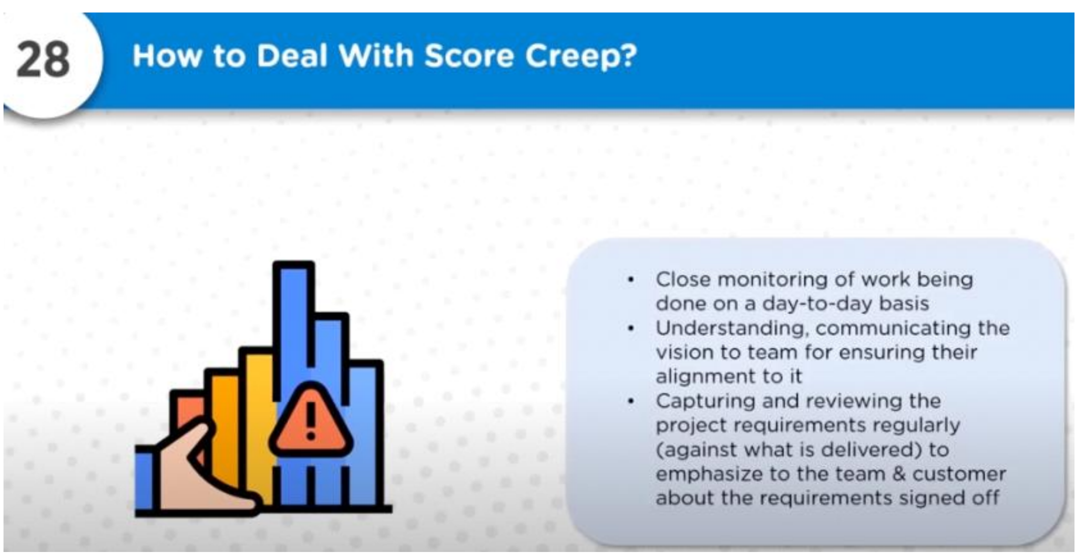
Jira Software

Yodiz

VERSIONONE
Agile Made Easier


27
How Does a Scrum Master Track Sprint Progress?
• Daily Scrum meetings
• Scrum retrospectives
• Sprint planning
• Escaped defects
• Defect density
• Sprint burndown
• Team velocity


28
How to Deal With Score Creep?
The term scope creep refers to the change which is uncontrolled and added without checking the impact on the constraints like scope, time, cost, etc.


28
How to Deal With Score Creep?
• Close monitoring of work being done on a day-to-day basis
• Understanding, communicating the vision to team for ensuring their alignment to it
• Capturing and reviewing the project requirements regularly (against what is delivered) to emphasize to the team & customer about the requirements signed off

28 How to Deal With Score Creep?

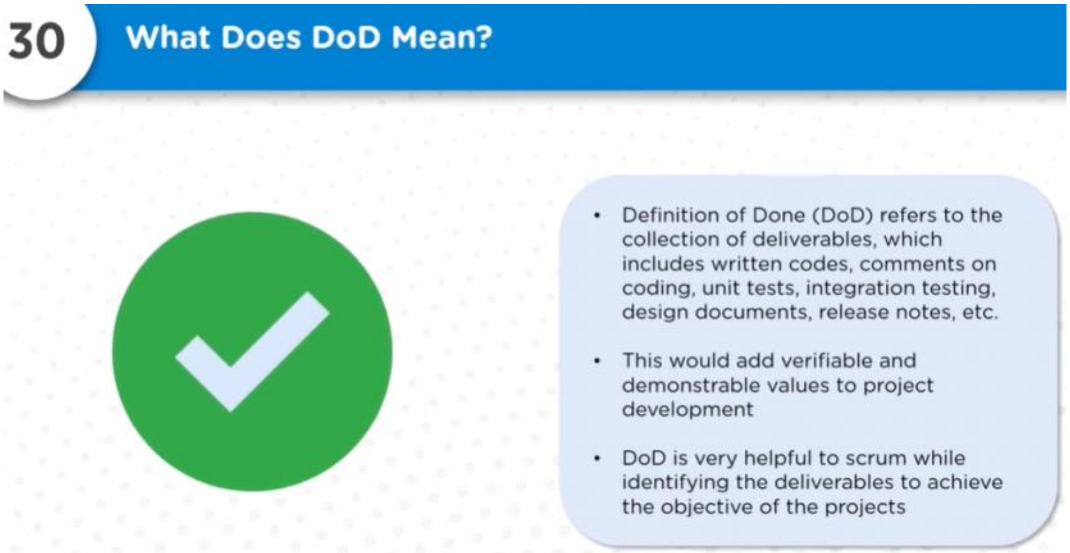
- Ensuring any changes introduced to go through change control and are implemented based on the approval for change request
- Avoid gold plating

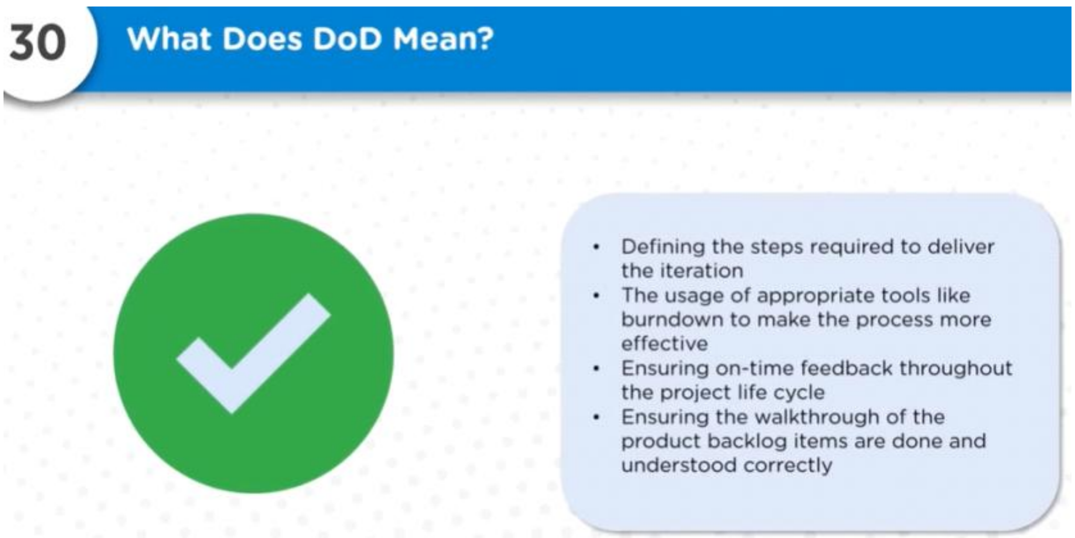
29 What Is MVP and MMP?

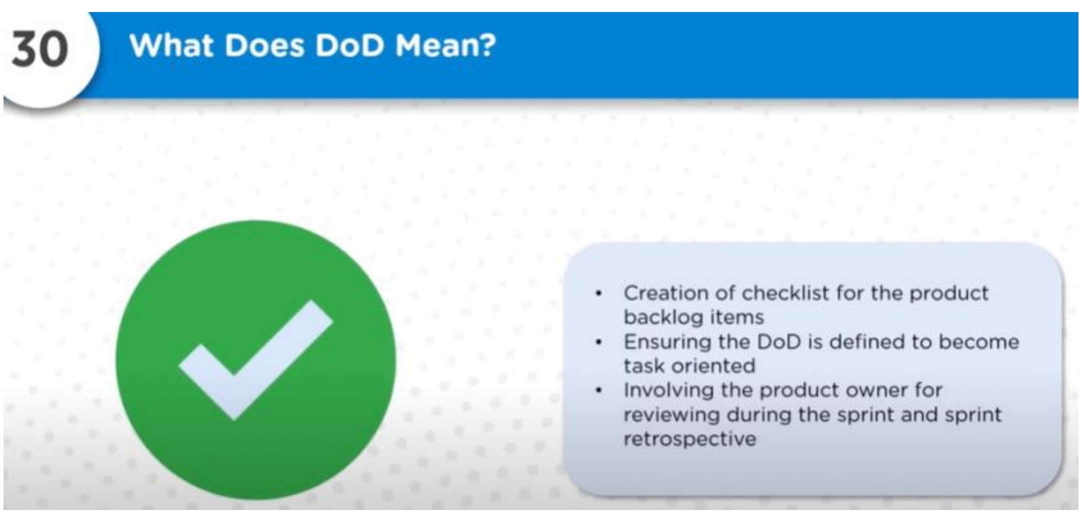
- Minimum viable product (MVP) is a concept of Lean Startup which stresses upon the impact of learning while performing product development
- This allows one to test and understand the idea by getting exposed to the initial version of the product for target customers & users
- To accomplish this, one has to collect all the relevant data and learn from that collected data

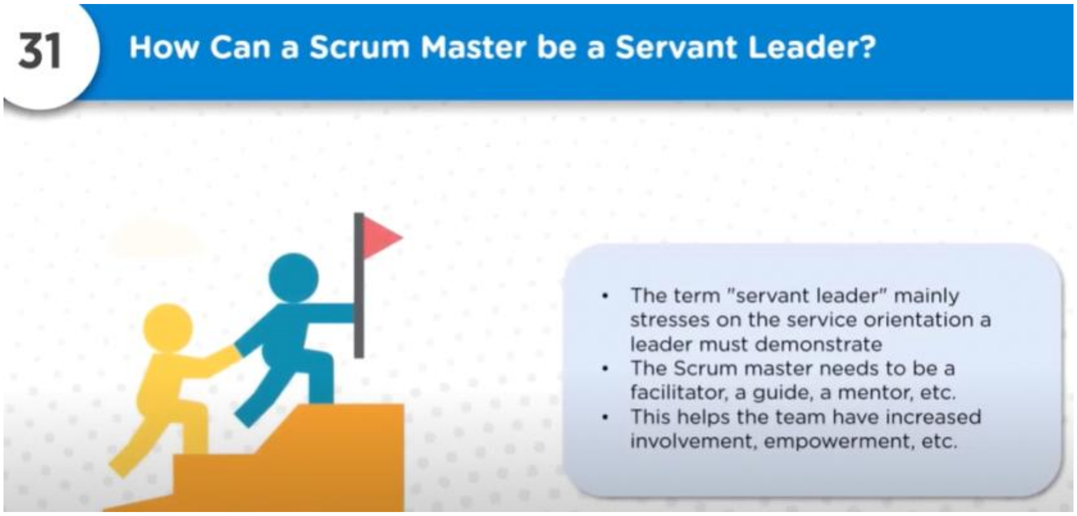
29 What Is MVP and MMP?

- The MMP (Minimal Marketable Product), refers to the description of the product, which will have a minimal number of features, that address the requirement of the users
- The MMP would help also help the organization reduce the time to market

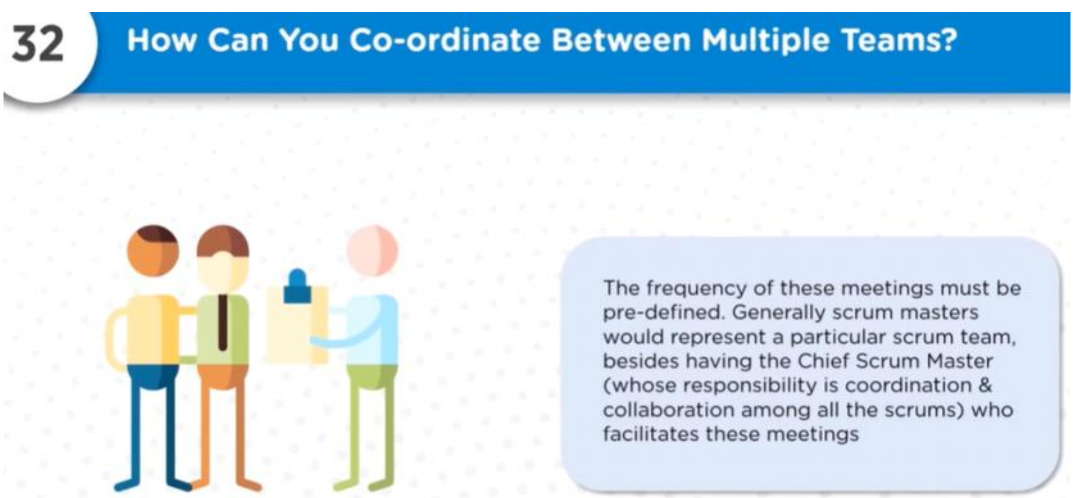

30
What Does DoD Mean?
• Definition of Done (DoD) refers to the collection of deliverables, which includes written codes, comments on coding, unit tests, integration testing, design documents, release notes, etc.
• This would add verifiable and demonstrable values to project development
• DoD is very helpful to scrum while identifying the deliverables to achieve the objective of the projects


30
What Does DoD Mean?
• Defining the steps required to deliver the iteration
• The usage of appropriate tools like burndown to make the process more effective
• Ensuring on-time feedback throughout the project life cycle
• Ensuring the walkthrough of the product backlog items are done and understood correctly


30
What Does DoD Mean?
• Creation of checklist for the product backlog items
• Ensuring the DoD is defined to become task oriented
• Involving the product owner for reviewing during the sprint and sprint retrospective


31
How Can a Scrum Master be a Servant Leader?
• The term "servant leader" mainly stresses on the service orientation a leader must demonstrate
• The Scrum master needs to be a facilitator, a guide, a mentor, etc.
• This helps the team have increased involvement, empowerment, etc.


32
How Can You Co-ordinate Between Multiple Teams?
One of the most common approaches for this is Scrum of Scrums (SoS) meeting, where members representing each scrum team discuss the progress, performance, issues, risks etc. together.


32
How Can You Co-ordinate Between Multiple Teams?
The frequency of these meetings must be pre-defined. Generally scrum masters would represent a particular scrum team, besides having the Chief Scrum Master (whose responsibility is coordination & collaboration among all the scrums) who facilitates these meetings

Create a Public Cloud Account & Approval Processes

Tuesday, October 1, 2024 7:08 PM

To set-up a non-prod account: [1 - Set up a Non-Production Cloud Account - Public Cloud | HCP Docs (uhg.com)]

To create a public cloud account, you need to be a contributor or owner of an HCP resource group.
That resource group must be in a GL-funded HCP account.
Using that *resource group*, an HCP account owner can organize HCP resources and assign role permissions.

- An HCP account owner does the following:
  - Obtains an ASK ID / AIDE ID if one is needed.
  - Goes to HCP Account Manager to create an HCP resource group with the ASK ID / AIDE ID.
  - Assigns owner or contributor privileges to application team members.
- Engineers can view the HCP resource groups they have access to in Account Manager. Depending on their access level, they can procure resources to resource groups as well as add other resource group members
- With a **non-production** public cloud account, an app team can provision and configure cloud resources, set up CI/CD pipelines, and find security solutions; test scenarios and fix issues. They ***cannot*** **use PII or PHI data** with a non-production public cloud account.
- Create a non-production cloud account
  1. Create a public cloud account request with Public Cloud Account Management in HCP Console. (First make sure the HCP resource group you have access to is selected in the drop-down to the top left of HCP Console.)
     a. Select **Create Public Cloud Account**. Then select **Nonprod** for the environment type.

     b. Fill out the required information in the **Create Public Cloud Account** dialog box – cloud service provider, account type, cloud account name, cloud account owner, etc. Select **Submit**. You can check the Public Cloud Account Management Dashboard in HCP Console for the status of your new public cloud account (initially *pending*).
  2. You will receive a Secure notification email with the **Read** role (using your MSID) for your public cloud account. And a second email with details on specific global groups created for your cloud account.
  3. In Secure, request access (for an Azure request) to either the **xxxx_Contributor_JIT** or xxxx_Owner_JIT global group to use Privileged Identity Management (PIM).
  4. Access your public cloud non-production account using your Cloud Native ID to log in. (See Access your cloud subscription for further details.)
  5. Your application team members to make a request in Secure for the global group(s) appropriate for them to work in your public cloud account.
     Focus on essential features that optimize performance. Use the infrastructure, networking, and storage that scale to your needs.
  6. Leverage Tech Landscape for cloud technologies that are either Preferred or Acceptable (PADU) - filter by Azure Preferred
  7. When working with data, use HITRUST eligible services.
  8. Use UHG/Optum platform services when possible
  9. Data transmission between the public cloud and an Optum data center must be designed optimally

  10. Before obtaining a production account, all resources need to be deployed with Infrastructure as Code (IaC), such as Terraform, stored in Github; CI/CD tooling (Github Actions, Jenkins, etc.) must be used to provision resources in non-prod
  11. You are responsible for the end-to-end support of your application (DevSecOps)
  12. Before seeking approval for your app in your non-production cloud account, **Network Perimeter** must be configured for supported CSPs (currently only Azure), and your non-prod account must be fully compliant with NetGuru.

Complete architecture and security reviews https://docs.hcp.uhg.com/public-cloud/complete-security-architecture-review

Deployed workload must meet cloud standards and security rule:
- Go to the cloud intake team for review and recommendations
- Some business units request a review by a line-of-business architect
- Go to Enterprise Information Security (EIS) for review
- After approvals, your production account will be provisioned

Complete your cloud intake request

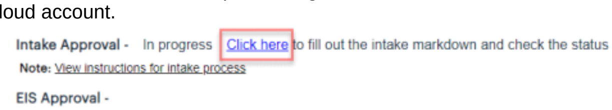
You need GitHub access to complete the Cloud Intake Request form

1. After receiving a public cloud account in HCP Console, you will need an "intake.md" markdown file to provide the details on your application.
   a. Go to the Public Cloud Account Management Dashboard in HCP Console and select **Request Intake** to the right on the row for your public cloud account.

	State	Type	Date Submitted	Approvals	Actions
-5dc63f759afe	Active	NonProd	08/20/2023	Request Intake	⋮

   b. At the bottom of the information page for your non-prod public cloud account, select the **Start Intake Approval** button - triggers the creation of an "intake.md" markdown file for your account under its own branch in the hcc-intake repo and in its own directory. This branch and directory have the same name as your account name.

   c. When it's available after processing, select **Click here** to fill out the Intake markdown document. This brings you to the GitHub repo branch named after your public cloud account.



   d. In GitHub, select the link toward the top to **add [account name]/intake.md file**.
   e. You can edit the intake.md in the GitHub web UI or use an IDE/editor of your choice. Place all supporting files in the directory with the intake.md file.

2. Provide the following information on your application in the Intake markdown document. Make sure your documentation – including diagrams – matches the infrastructure that you automate through your deployment pipeline (CI/CD).

Sections of Intake document	
General Information	Describe your **application's use, architecture and infrastructure**; users of your application, etc.
Data Governance	Describe details about **data you will be hosting or collecting, upstream sources, targets** where you are returning data to, etc.
Code Delivery & Management	Provide **links to IaC/Terraform code** you are using to build cloud components
Architecture Summary	Provide an architecture diagram with components you are using
Network Summary	Provide a network diagram with network flow details for components

3. After you finish your edits to the intake.md document and add supporting images, open a pull request to merge it into the *main* branch of the hcc-intake repo. Add images as files to the repo, rather than links. The **cloud intake team will review your pull request** and provide feedback on any issues that need to be resolved. The most **common issues are lack of IaC, lack of CI/CD, or Netguru findings** that must be resolved.

4. Resolve issues found in the PR review

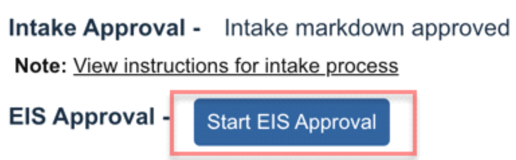
Line-of-business architect review

Certain business units (OI, ORx, etc.) have their own architect approval step. This happens between the cloud intake team's approval and the EIS endorsement request. If you are from one of these business units, the **Start EIS Approval** button described below will remain unavailable until this happens.

Request EIS endorsement

EIS endorsement is led by the Enterprise Information Security (EIS) team. It is **mandatory for all production accounts** and is a thorough assessment of cloud architecture and security.

1. Soon after the cloud intake team has reviewed and approved the intake (merges your pull request), the button in HCP Console Public Cloud Account Management (PCAM) will be available. You can then request EIS endorsement.
   a. Go to the Public Cloud Account Management Dashboard in HCP Console and select your non-production cloud account.
   b. Select **Start EIS Approval**.



   After your EIS approval request, an RITM number shows in place of the button. It links to ServiceNow where you can track your EIS request.
   EIS endorsement requirements:
   - Cloud architecture diagram
   - Network diagram
   - Security assessment
   - Review of security groups: All workload ingress and egress access must be implemented with as least-privileged access as possible; all access must be documented with business functionality justification
   - All Critical and High cloud vulnerabilities and cloud configuration violations need to be remediated
   - All Critical and High violations identified in required security scans (SAST/IAST, DAST etc.) must be addressed
   - Security Platform is used for viewing violations, risk records, scan requirements, etc.
     HITRUST certified services and modules: All Cloud Service Provider (CSP) services and modules used within a subscription or account must be certified by the CSP and included in the BAA agreement to be used. Any services or modules not HITRUST Certified must be reviewed and approved by EIS
2. After the EIS review, update the intake.md and other files as requested
3. If necessary, create a new branch/fork to make your changes, commit them, and then create a new pull request against *main*

Resourcing

Tuesday, August 20, 2024 1:40 PM

Objective: Ramp up Dave on getting ready to lead this effort. The following needs to be done:

- Request required resources
- Plan for the effort by building a workplan (See attached from MRIS)
- Prepare to manage to budget by preparing burn report

Contractor Resource Request Link:
https://app.smartsheet.com/b/form/4ec1d016250646a7a5f806efbf662992

- NOTE: for offshore contractors, the entry in the link above is just to give an FYI to OAS team managing resource procurement (Maureen Shurson, MagenPuerzer, & Team). The actual request to get contractors we work with LTIM and Persistent so will also have to request from them.
- To Request actual resources:
    - Define the resource requirement outlining “Key require skillsets” and providing context for the primary responsibilities of the requested resources.
    - Send email to LTIM
        - Sandy Jones Sandy.Jones@ltimindtree.com
        - Ajayakumar Charuvil Damodaran Ajayakumar.CharuvilDamodaran@ltimindtree.com
    - Send email to Persistent
        - Sushil Chokhandre sushil_chokhandre@persistent.com
        - Sujata Ranawade sujata_ranawade@persistent.com
    - OGA Leads to CC when sending request email to LTIM or Persistent
        - Bangale, Gaurav gaurav.bangale@optum.com
        - Sagi, Vipin vipin_sagi@optum.com

From: Shin, John S <john.shin@optum.com>

Date: Thursday, August 15, 2024 at 5:42PM

To: Cheema, Dave <dave.cheema@optum.com>

Cc: Malhotra, Dinesh <dinesh.malhotra@optum.com>

Subject: Re: RAD Phase 1 Story for Alex to Communicate Up

Dave – FYI. This opportunity is going to be moving forward. I will setup a call with you tomorrow so we can request the offshore resources. I will walk you through the process so you can take point on this effort.

John Shin (he/him)
Sr. Director Cloud Data Engineering & Solutions | Optum Advisory | Optum

O 1-952-833-7266
M 1-612-321-6400
john.shin@optum.com

12125 Technology Drive
Eden Prairie, MN 55344

Optum

Integrity • Compassion • Inclusion • Relationships • Innovation • Performance

From: Alex Schmidt <alex_schmidt@uhg.com>

Date: Thursday, August 15, 2024 at 4:56PM

To: Malhotra, Dinesh <dinesh.malhotra@optum.com>

Cc: Shin, John S <john.shin@optum.com>, Whitmer, Patrick <patrick_whitmer@uhc.com>

Subject: RE: RAD Phase 1 Story for Alex to Communicate Up

Yes, please send through (note only phase 1 at this time); I have a meeting with Ken shortly to let him know its coming too.

Best,
Alex

From: Malhotra, Dinesh <dinesh.malhotra@optum.com>
Sent: Thursday, August 15, 2024 2:26 PM
To: Alex Schmidt <alex_schmidt@uhg.com>
Cc: Shin, John S <john.shin@optum.com>; Whitmer, Patrick <patrick_whitmer@uhc.com>
Subject: RE: RAD Phase 1 Story for Alex to Communicate Up

Alex – Congratulations for securing the approvals

We will send you and the Finance Approver a formal SoW for approval – just like last time.

Would Ken Baxter be the Finance Approver for this work as well?

Best,
Dinesh Malhotra
M: +1 773-398-7713

From: Alex Schmidt <alex_schmidt@uhg.com>
Sent: Thursday, August 15, 2024 4:23 PM
To: Malhotra, Dinesh <dinesh.malhotra@optum.com>
Cc: Shin, John S <john.shin@optum.com>; Whitmer, Patrick <patrick_whitmer@uhc.com>
Subject: RE: RAD Phase 1 Story for Alex to Communicate Up

Hi Dinesh,

Let's roll on Phase 1 😌

I will have to find a solve for the other phases, but time for action is now!

We are pumped up.

Best,
Alex

From: Malhotra, Dinesh <dinesh.malhotra@optum.com>
Sent: Wednesday, July 24, 2024 3:42 PM
To: Alex Schmidt <alex_schmidt@uhg.com>; Whitmer, Patrick <patrick_whitmer@uhc.com>
Cc: Shin, John S <john.shin@optum.com>
Subject: RE: RAD Phase 1 Story for Alex to Communicate Up

Hi Alex / Patrick –

Here are couple of high quality links that I had shared with our team members in the past for Azure Fundamentals Certification training on YouTube

https://www.youtube.com/watch?v=NPEsD6n9A_I&list=PLGjZwEtPN7j-Q59JYso3L4_yoCjj2syrM

https://www.youtube.com/watch?v=5abffC-K40c

Best,
Dinesh Malhotra
M: +1 773-398-7713

From: Shin, John S <john.shin@optum.com>
Sent: Wednesday, July 24, 2024 9:30 AM
To: Malhotra, Dinesh <dinesh.malhotra@optum.com>; Alex Schmidt <alex_schmidt@uhg.com>; Whitmer, Patrick <patrick_whitmer@uhc.com>
Subject: Re: RAD Phase 1 Story for Alex to Communicate Up

Hi Alex,

Please take a look at the slide I added (slide 3) for RAD Team training guidance and let me know if you need anything further.

Hope this helps.

John Shin (he/him)
Sr. Director Cloud Data Engineering & Solutions | Optum Advisory | Optum

O 1-952-833-7266
M 1-612-321-6400
john.shin@optum.com

12125 Technology Drive
Eden Prairie, MN 55344

Optum

Integrity • Compassion • Inclusion • Relationships • Innovation • Performance

From: john.shin@optum.com
When: 2:00 **PM - 2:30 PM July 22, 2024**
Subject: RAD Phase 1 Story for Alex to Communicate Up
Location: Microsoft Teams Meeting

Objective: Review executive summary (to defined by John) why starting the training of RAD Team on cloud tech stack and, at the same time, engaging OAS to conduct a Solution Definition & Design effort would help setup the RAD team to effectively execute on the RAD Data Modernization program.

Microsoft Teams Need help?

Join the meeting now

Meeting ID: 277 018 917 882
Passcode: S9cmYk

Join on a video conferencing device
Tenant key: 425899727@t.plcm.vc
Video ID: 112 074 241 85
More info
For organizers: Meeting options
Org help

Roles

Saturday, August 17, 2024 7:49 AM

Scrum Master

Looking for an experienced Scrum Master to lead a critical on-premise data migration project to Azure Snowflake. The ideal candidate will be a seasoned Agile practitioner with a deep understanding of data migration projects, cloud technologies, and the ability to drive high-performing teams. This role requires a collaborative, results-oriented individual who can effectively facilitate Scrum ceremonies, remove impediments, and ensure the team delivers value incrementally.

Responsibilities:
1. **Agile Leadership**: Serve as the Agile leader and coach, guiding the project team and stakeholders in adopting and implementing Agile principles, values, and practices.
2. **Scrum Process Management**: Facilitate all Scrum events, including daily stand-ups, sprint planning, sprint reviews, and retrospectives.
3. **Collaboration and Communication**: Create effective collaboration and communication within the team and with stakeholders.
4. **Project Planning and Execution**: Work closely with the Product Owner and development team to define and prioritize the product backlog; define sprint planning, backlog refinement, and release planning activities.
5. **Reporting**: Track and report on team progress and metrics.
6. **Risk and Issue Management**: Identify and mitigate risks and issues that could impact project successful completion. Proactively address any obstacles or dependencies that arise during the migration process.
7. **Stakeholder Management**: Collaborate with stakeholders, such as client SMEs and business users to understand their requirements and objectives. Provide regular project updates.
8. Remove impediments and create a conducive environment for team self-organization and problem-solving.
9. Foster a collaborative and positive team culture.

Requirements:
1. Bachelor's degree in Computer Science, Information Systems, or a related field.
2. Certified Scrum Master (CSM) or equivalent certification is highly desired.
3. Proven experience as a Scrum Master in data migration projects, preferably to Snowflake.
4. Deep understanding of Agile principles, frameworks, and practices, with hands-on experience in implementing Agile methodologies, e.g., Scrum, Kanban, etc.
5. Experience with Azure cloud platform and Snowflake data warehousing is preferred.
6. Knowledge of data migration tools and processes.
7. Strong communication, facilitation, and problem-solving skills.
8. Ability to manage multiple priorities, work under pressure, and adapt to changing project needs.
9. Ability to adapt to changing priorities and environments.
10. Excellent interpersonal skills, with the ability to build collaborative relationships with diverse stakeholders

Desired Skills:
1. Experience with data warehousing and ETL processes.
2. Understanding of cloud migration challenges and best practices.
3. Experience with DevOps practices and tools.
4. Strong analytical and problem-solving skills.

++

Business Analyst

Looking for a skilled and detail-oriented analyst to support the on-premise data migration project to Snowflake. The analyst will play a critical role in gathering and analyzing business requirements, documenting processes, and ensuring successful data migration to the Snowflake platform.

Key Responsibilities

- **Business Requirements Gathering and Analysis:**
  - Identify and document the business needs and objectives driving the migration.
  - Understand current business processes, data usage patterns, and dependencies.
  - Assess the impact of data migration on existing systems and processes.
- **Data Mapping and Profiling:**
  - Collaborate with data owners and stakeholders to identify critical data elements.
  - Define data mapping rules and transformations between on-premise and Snowflake.
  - Conduct data profiling to assess data quality, consistency, and completeness.
- **Data Governance and Compliance:**
  - Ensure data governance policies and standards are defined and implemented.
  - Identify data privacy and security requirements and ensure compliance.
  - Develop data retention and archival strategies.
- **Change Management:**
  - Develop change management plans, including communication and training strategies.
  - Manage stakeholder expectations and address concerns.
- **Testing and Validation:**
  - Define test cases and scenarios to verify data accuracy and integrity.
  - Collaborate with the testing team to execute test plans and validate results.
  - Ensure data quality and consistency after migration.

Skills and Knowledge
- **Business Analysis:** Strong analytical skills to understand business requirements and translate them into technical specifications.
- **Data Analysis:** Proficiency in data analysis techniques to assess data quality and identify data transformation needs.
- **Data Governance:** Knowledge of data governance principles and best practices.
- **Change Management:** Ability to manage change effectively and communicate with stakeholders at all levels.
- **Communication:** Excellent written and verbal communication skills to interact with both technical and business teams.
- **Technical Understanding:** Basic understanding of database concepts, ETL processes, and cloud technologies.

Challenges and Opportunities
- **Data Quality Issues:** Address data inconsistencies and inaccuracies and opportunities to improve data quality.
- **Data Security and Privacy:** Identify data protection and compliance with regulations.
- **Change Management:** Overcome resistance to change and manage stakeholder expectations for a successful migration.

- **Performance Optimization:** Identify performance bottlenecks to improve the value of the migrated data.

Colors

Thursday, October 10, 2024 7:17 PM

Green - #00b050
Yellow - #ffff96
Red - #ff0909
Royal Blue - #348dcd
Navy Blue - #00226b

Team PTOs

Friday, October 18, 2024 4:02 PM

Patrick: 11/11 - 11/15, 11/25 - 11/29 (1 week) = 15 hrs.

Preeti: 10/31 - 11/04 = 13.5 hrs.
Suman: 11/01 = 8.5 hrs.
Mahesh: 11/01 = 8.5 hrs.
Asit: 11/01 - 11/04 = 8.5 hrs.
Muquaddar: 10/31 - 11/01 = 17 hrs.

John S.: 11/18 - 11/22, 11/25 - 11/27 = 8 hrs.
Dave C.: 11/28 - 11/29 = 16 hrs.

Total: 13.5 + 8.5+ 8.5 + 9 + 17 + 8 + 15 + 16 = 96 hours

Hours/day: 1.5+4.25+8.5+8.5+4.25+8.5+1+8=44.5
96/44.5=2.1573+2 (Thanksgiving holidays) = 4.1573 days
Data Product approval: 3 days
Data Product Subscription approval: 4 days
Global Review: 10 days

Final delivery will be delayed by 1 week

Roles

Wednesday, September 4, 2024 9:03 AM

Dinesh Malhotra - Executive Leadership
John Shin - Delivery Oversight
Dave Cheema - Delivery Lead
Preeti Khurana - Scrum Master
Asit Sahoo - Cloud Administrator
Muqaddar Baig - Lead Cloud Data Engineer
Mahesh Kumar - Cloud Data Engineer
TBD (Suman Kopulla) - Cloud Data Engineer

Tracker

Thursday, May 8, 2025 3:20 PM

Issue	Start Date	Resolution Status	Resolution Date	Assigned To	Comments
Verify if all needed Gold views are completed	5/8/2025	assigned		Patrick Whitmer	
Create Gold view definitions for the PeopleSoft tables	5/8/2025	Assigned		Patrick Whitmer	Patrick will discuss it with Grant
Do we have the needed clarifications for the 20 ESG data sources?	5/8/2025	WIP		Patrick Whitmer	Patrick has shared the initial tables Bronze and Silver tables. OA will study and compile questions
					\