ANALYTICS 2.0

Analytic Digital Foundation (ADF) Blueprint

June 2019

KAISER PERMANENTE®

Agenda Slide

- Introduction
- A20 Program and Systems of Intelligence (SOI) Overview
- Foundation Services
- Break
- Foundation Data
- Tenant Services
- Organizational Readiness

Purpose

- The goal of the A20 Program is to deliver high value, high quality, and timely insights to KP's decision makers
- The ADF Blueprint covers a wide depth and breadth of vital information that lays the foundation for the A20 Program, and is central to understanding the ADF and how it is being designed and built
- This presentation was developed to support ADF blueprint deep dive sessions, and is critical for anyone needing a deeper understanding of the A20 Program foundation
- As a key stakeholder in the A20 Program community, you are essential to the successful implementation of the ADF and A20 Program. We appreciate your engagement!

Key Considerations

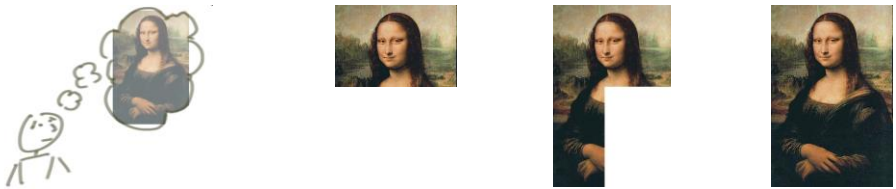
- Baseline level of knowledge about A20 Program is assumed
  - Information and resources available on A20 Program site
- The ADF blueprint will continue to be updated as program development and execution matures
  - Feedback is essential to the development of this work and will be gathered during this session, as well as after the session via survey and A20-Updates mailbox
- This is the first in a wider communications and training strategy to promote understanding of the ADF
  - Stay tuned for more information on the ADF Blueprint in the future, and the recording location

Iterative vs. Incremental

What do we mean by iterative and incremental?

Building iteratively as we refine the product – Iteratively will have a rough draft to start with, and refine further with an increase in details

Building incrementally one piece at a time – Incrementally may call for the need to have the full idea formulated



Source: Jeff Patton

We Want Your Feedback!

- Please use mural to provide questions and suggestions
- Use chrome browser to open the link in the chat box and log in as “Anonymous”
- Use the outline to navigate the view (this is located on the right side of the screen or button on the top right, if not visible)
- Use the sticky notes (categorized by topic/color) to document and classify your question/suggestion. You can move the sticky note to the appropriate slide, using the Outline navigation if helpful
- Use the Parking Lot if you aren’t sure which slide/category applies
- Share questions/suggestions with Val, Hector or Maday if you have issues with mural
- Email A20-Updates@kp.org with questions/suggestions after the session

Program Development/Roles & Responsibility	Platform/Technology	Change Management/Training	Future State/Production Support

ADF Blueprint Scope

To create the ADF blueprint, we organized five tracks of work:

Foundation Services

- **Tenant Data Collection**: Engage with tenants to understand and document **analytic personas, analytic asset's technical and functional capabilities, data related needs, and compliance** requirements
- Develop ADF architecture related artifacts
- Develop artifacts to support **tenant onboarding & engagement**

Foundation Data

- Develop an **integrated set of architectural artifacts** that describe the data dimension of the Analytic Digital Foundation (ADF) including:
  - Data zones definition
  - Design patterns of interaction with other key data disciplines
  - **Information & Analytic Domains**

Reference Implementation

- Document the features, **components and user stories for Revenue Management tenant** implementation
- Document reference **implementation solution approach** options, pros vs cons, guidelines and assumptions for **Change Data Capture (CDC), Clarity*N and Data Validation** use cases

Standards & Best Practices

- Develop 10-12 standards and best practices in 2 page format (additional scope defined for tranche 4)
- Design **framework and process to develop, review and approve ADF standards** and best practices
- Design and content recommendations for **ADF standards and best practices website**

Organizational Readiness

- Develop a **comprehensive strategy** to guide the execution of change management and communication activities supporting the A20 Program
- Provide a **framework** and resources for **developing trainings** by tracks and tenants
- Define three **operating model design options** to consider as a framework for how the A20 Program ADF ecosystem will work together in the future-state

Objectives: ADF Blueprint In-Depth Review

Section	Objectives
A20 Program and SOI Overview	✓ Inform on A20 Program background, scope, structure expected benefits ✓ Share how the ADF aligns with Systems of Intelligence architecture
Foundation Services	✓ Describe the A20 Program Architecture guiding principles, technology choices ✓ Provide an overview of A20 Program Solution Patterns, Toolkits that enables tenants ✓ Describe A20 Program Non-functional Requirements and Foundation Service/DR plans
Foundation Data	✓ Describe components of Foundation Data blueprint and guiding principles that are driving design ✓ Define the ADF data zones and Foundation Data Management (FDM) disciplines and describe how they support analytic needs ✓ Share Data Domain taxonomy and how it will enable and help to manage analytic solutions ✓ Provide an overview of key features – Change Data Capture, Clarity * N, Data Validation – applied to Revenue Management as a Reference Implementation ✓ Understand where to locate data ingest status
Tenant Services	✓ Describe the Tenant Data Collection Approach and key learnings ✓ Describe the Tenant Engagement Process and toolkits that have been developed to facilitate Tenant onboarding, development, delivery and benefit realization ✓ Discuss the importance of and share the content and process for establishing standards and practices for the ADF
Organizational Readiness	✓ Provide information about the approach the A20 Program organizational readiness team is using to ensure stakeholders are ready to adopt the changes that will result from the A20 Program ✓ Provide details on in-progress and upcoming organizational readiness activities

A20 Program and Systems of Intelligence (SOI) Overview

Key Objectives

- Inform on A20 Program background, scope, structure expected benefits
- Share how the ADF aligns with Systems of Intelligence architecture

A20 Program Vision

Deliver high value, high quality, and timely insights to KP's decision makers

- Create a cloud-based analytics platform
- Reduce time and effort to generate insights
- Enable new kinds of sophisticated analytics
- Enable both decision makers and analysts through easy and intuitive self-service capability
- Create the source of trusted, standardized, KP-wide data for decision makers and analytics purposes

A20 Program Guiding Principles

Self Service Enablement

When possible, decision makers and their analytics support teams should be able to easily fetch and analyze the needed information themselves. This will require diligent metadata management, intuitive user interfaces, pre-made joins, and data models that fit the inquiries.

Flexibility

There should be a wide spectrum of tools and ways to fetch and interact with the data, including visual data exploration, drill through, various ways of slicing the data and simple ways to lay analytic functions on top of data.

Standardization

Establish and use same definitions across the program, both for raw data and for enrichments, when appropriate. Default to industry standard first and to KP-wide standards second, but enable multiple versions of truth when necessary.

Speed to Insight

Information is of value when data analyses lead to action. While action is often beyond the reach of analysts and supporting systems, the goal should be to generate insights (processed, analyzed and decision relevant information) in as little time as possible.

User Experience

The system should be designed such that it will be easy and delightful to use. Accessibility during the workflow, aesthetics, and intuitive interfaces are all important components of well-designed system.

Agility

When possible, we should design the systems such that we do not lock ourselves to specific vendors, tools or processes. Moreover, changes to data structures, tools, and processes should be easy and quick to implement.

Re-use

Enable sharing and re-use of code. Establish ways to capture analytic process insights (on data and analytic algorithms) to make it easier for others to conduct the same or similar analyses in the future.

Quality Enablement

Quality of analytics and the resulting insights begins with data quality: accuracy, uniqueness, consistency, completeness, timeliness and validity. Quality enablement makes it easy for users to identify and conduct the right data, analytic and quality control processes to ensure output that the analyst can proudly stand behind.

KAISER PERMANENTE®

A20 Program Current and Future

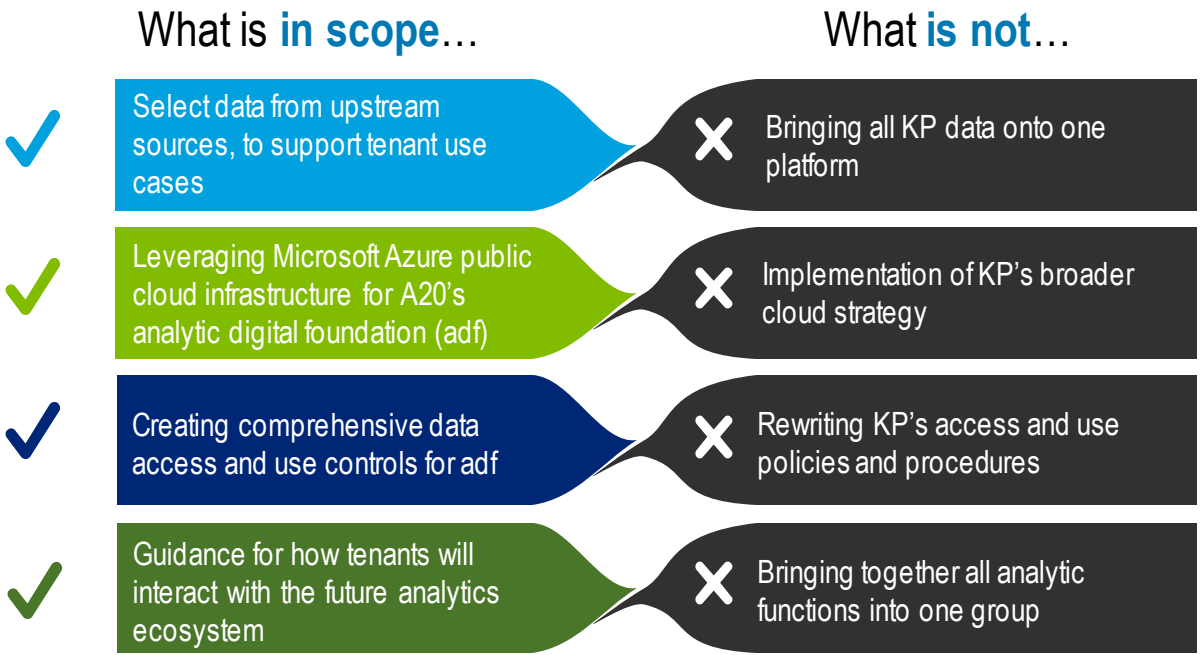
		Current	Future
Operating Model: Describes the way KP conducts its data and analytics business	ECONOMICS	**Project Based** • 20% utilization, 7x duplication	**Pay per Use; Cloud** • 75% utilization, common data stage
	PROCESS	**Mode 1** • Annual / Semi-Annual Releases • Analyst spending 75% of time in data preparation	**Mode 2** • Continuous / Daily / Weekly Release • Analyst spend 10% in data preparation
	PEOPLE	**Traditional Data warehouse** • ETL* Developers, Database Administrators, Report Developers	**Next Gen Data Engineering** • Data Engineers, Data Scientists, Platform Engineers
Data & Analytics Capabilities: Describes how data and analytics capabilities come together to enable decision makers	DATA	**Data Silos** • Redundant, unmanaged data provisioning • Functional and regional data warehouses	**Managed Data Platform** • Common secure data provisioning • Organized by data domains
	CAPABILITIES	**Dis-Jointed, Batch** • Batch, Duplicated data • Silo-data warehouse	**Right time, Right Data** • Realtime, common provisioning • Managed Secure Data Platform
	TOOLS	**Closed Legacy technologies** • SAS, Informatica, Oracle, Cognos, Business Objects	**Open Data Platform** • Kafka, Spark, Hadoop, SQL, NoSQL, R, Machine learning, Artificial Intelligence
IT Infrastructure	INFRASTRUCTURE	**Build to Order** • On-premise; vertical scaling	**Elastic, Scalable Compute** • Cloud; horizontal scaling
	PROVISIONING SPEED	• Weeks, Months	• Minutes, Hours

* ETL = Extract-Transform-Load

KAISER PERMANENTE®

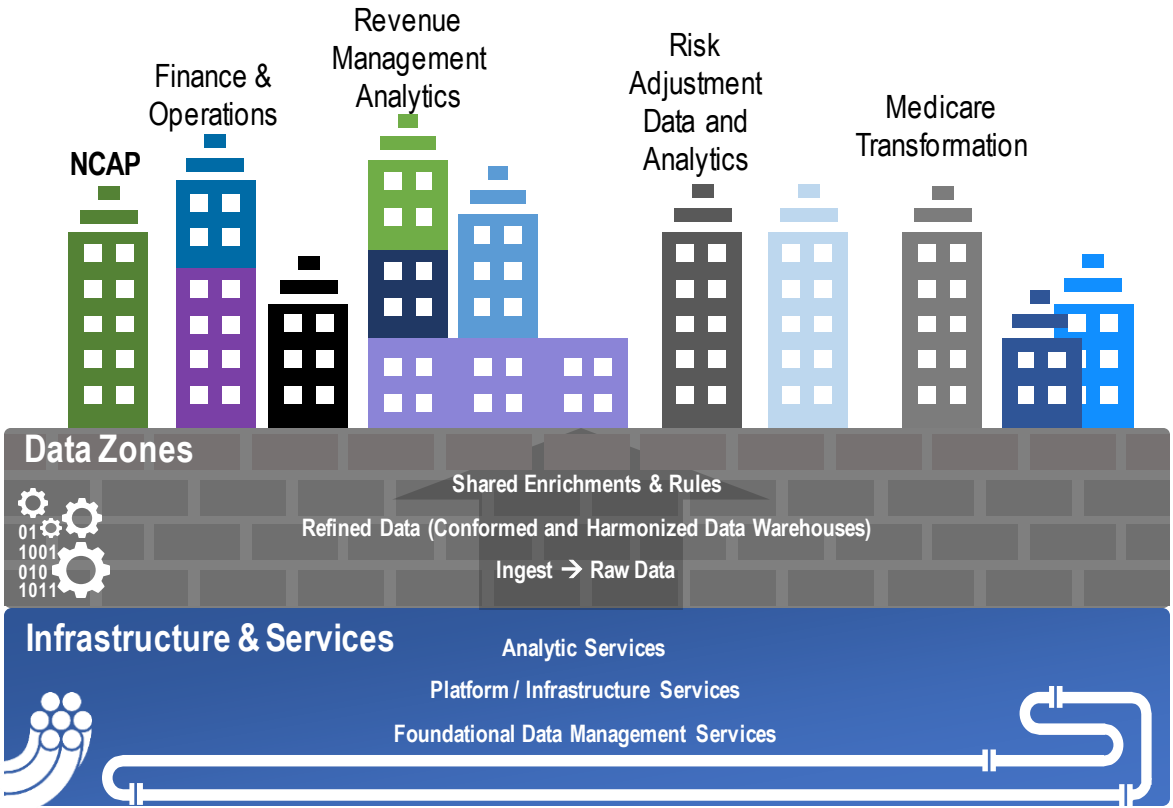
A20 Program Scope Myths

Scope clarity and focus is important



KAISER PERMANENTE®

A20 Analytic Digital Foundation and Initial Tenants



A20 Program Benefits

Benefits

For Decision Makers:

- Improves **quality, consistency and timeliness of information** for decision-makers.
- **Streamlines access** to business-critical information.
- Establishes **single source of truth** that targeted decision makers can rely upon.
- Enables **broader self-service capabilities** that allow decision makers to answer their own questions.
- **Supports KP's strategic priorities**, particularly affordability, by enabling expedited identification and evaluation of opportunities.

For Analytic Organizations:

- **Reduces time spent on data preparation** and cleansing, increases insights and analytic yield.
- Enables development of **innovative, near-time analytic models**.

For Information Technology:

- Consolidates regional and functional data onto a **common technical platform**.
- Reduces the size of the analytic infrastructure by **rationalizing and integrating data**.

For Future Analytic Projects:

- **Reduces time to value** by leveraging ready to consume capabilities.
- **Speeds up development** of new capabilities by reusing technology infrastructure.
- **Reduces costs** by enabling reuse of the Analytic Digital Foundation (Foundation Services and Foundation Data).

A20 Program Efficiencies and Capabilities

Efficiencies

- Creating a shared, re-usable Analytic Digital Foundation that includes **program wide (all regions), health plan, finance and operational data** and shared data management and analytic services **on a single platform**.
- The Analytic Digital Foundation can be re-used, **reducing the expense of future analytic projects**, and **avoiding duplicative cost of data management, storage and compute**. A20 will develop tools, training and support to migrate new and existing analytic systems to the foundation.
- Integrating onto the Analytic Digital Foundation, and **transitioning 25 analytic systems,** supported by Finance and Decision Support (includes former Management Information Analytics, Decision Support Services and Northwest Data and Information Management Enhancement).
- Conforming and rationalizing duplicative data, business rules and definitions, **bringing consistency and comparability to program wide metrics** and enabling benchmarking. Delivers a shared dimensional data model, allowing KP to **easily, and consistently evolve analytics and metrics** to new areas.

Capabilities

- Creates new Analytic Tenants on the Analytic Digital Foundation, with new dashboards and program wide metrics that support **apples-to-apples comparison of key performance indicators** for Group Presidents, National and Regional Senior Executives and Regional and Functional Leaders.
    - Presidents Dashboard, Highly Reliable Operating Model, Operating Plan Dashboard
    - 50+ program wide metrics across 10 domains (quality, service, access) with drill to detail
- Develops new and extended analytic and data management services, including a **common portal, integrated master and reference data, a shared data catalog and integrated data quality and validation** rules.
- Supports open source software and creation of standard data schemas, allowing KP to **leverage industry predictive models and artificial intelligence libraries**.

A20 Program and Governance Structure (as of 3/27/19)

Membership

Executive Sponsors
Dick Daniels
Pat Courneya, MD
Nancy Gin, MD
Edward Lee, MD
Kathy Lancaster
Paul Minardi, MD

Steering Committee
Debbie Szoke – Tri-Chair
Vivian Tan – Tri-Chair*
Maggie Vashel – Tri-Chair*
Hovannes Daniels*
George Di Salvo
Margo Gordon*
Eric Liederman, MD
Jeff Johnson*
Steve Stock
Long Thai**

Program Realization Team (PRT)
Vivian Tan – Chair
Hovannes Daniels
Erin Dirks
Chris Guy
Mark Kano
Leland Louie
Steve Stock
Debbie Szoke
Ganesh Thondikulam
Allison Winkler

Governing & Operational Bodies

High-level

- Executive Sponsors ★
- Steering Committee ★

Implementation

- Program Realization Team (PRT)
- Lean Portfolio Management (LPM)
- Integrated Program Management Office (IPMO)

Program Tenants and Tracks

- NCAP: Profitability, Utilization & Cost
- Finance & Operations
- Medicare Transformation
- Risk Adjustment Reporting
- Service, Utilization & Quality
- Pricing & Underwriting
- Revenue Management
- Financial Planning & Reporting
- Employer Group Reporting
- Membership & Market Analytics
- Finance & Supply Chain

- Foundational Data
- Foundational Services
- Foundational Data Management
- Tenant Services

Advisors & Stakeholders

- Analytic Infrastructure Council
- Chief Technology Office
- Community of Practice ★
- Change Advisory Network
- Business Owners/Stakeholders

Lean Portfolio Management (LPM)
Leland Louie – LPM Chair
Chris Guy
Ganesh Thondikulam
Mark Kano
Doug McMartin
Jae Chong***
Lucille Villanueva***

Key
★ PMG representation

*also AIC member
**AIC and NCAP Co Ex member
***LPM project management support (non-member)

KAISER PERMANENTE®

A20 Program Leads and Roles

A20 Realization Leads	Chris Guy	KP Insight
	Ganesh Thondikulam	Corporate Services IT
Foundation Data	Claudine Martin	KP Insight
	Prasad Shenoy	Corporate Services IT
Foundation Data Management	Kate Kalikstein	Chief Data Office
Foundation Services	Lisa Gros	KP Insight
	Subha Parthasarathy	Corporate Services IT
Integrated PMO	Leland Louie	Finance Management Services
Organizational Readiness	Erin Dirks	KP Insight
Tenant Services	Kathryn Clift	KP Insight

Analytic Digital Foundation (ADF)

SOI Architecture Evolution

Enterprise Level Context

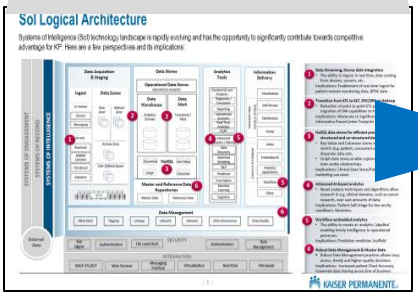
- SOI Architecture
  - SOI Architecture provides enterprise level context and capabilities required for transforming data into meaningful and actionable information

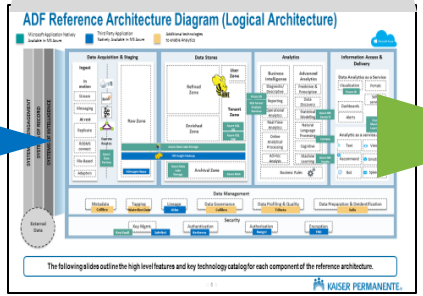
Program Level Context

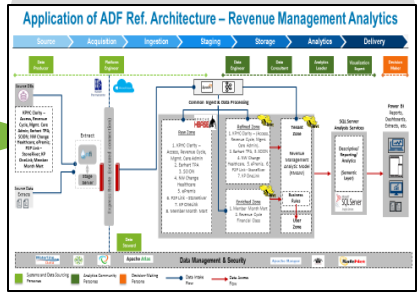
- ADF Reference Architecture
  - Analytics 2.0 program aims at development of a **shared, re-usable data and analytics foundational platform** consisting of cost-effective and efficient foundation data zones – raw, refined, and enriched and foundation services that meet current and future needs of the analytic tenants

Project/Capability Level Context

- Solution Architecture(s), at a specific Tenant level (e.g., RMA, NCAP, etc.)



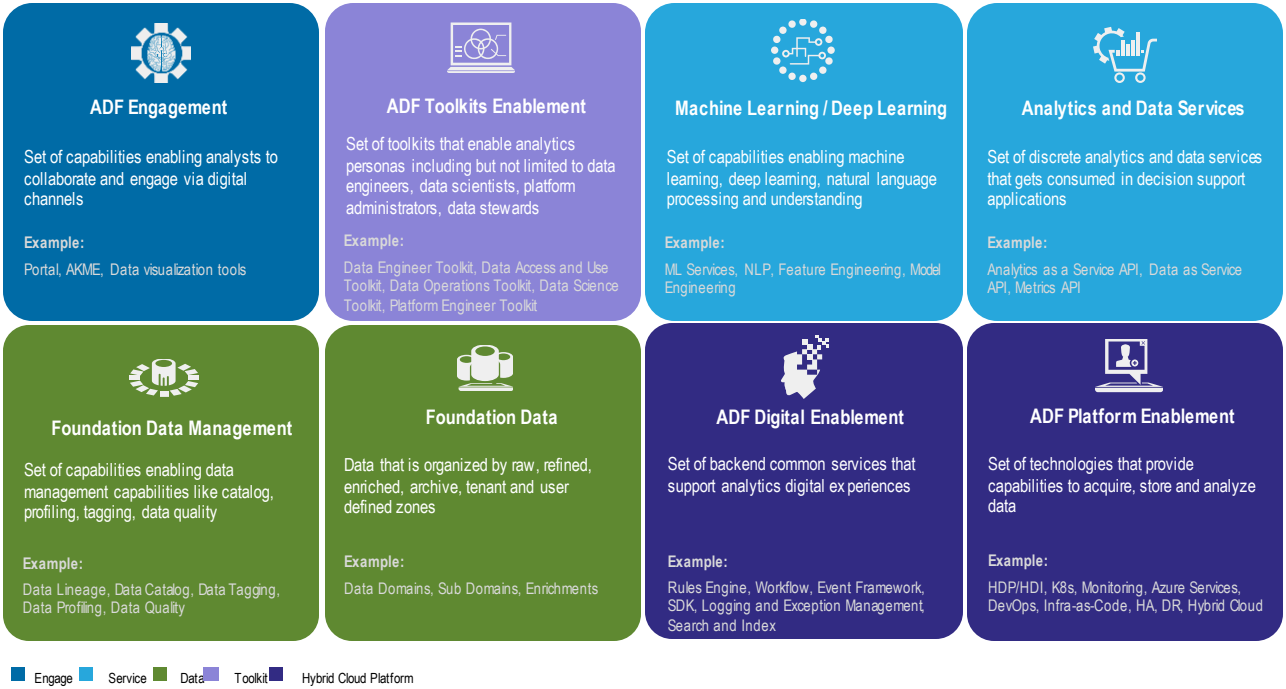




Evolution of the SOI Architecture is characterized by the ADF Reference Architecture, detailing the significant architectural components and capabilities within the ADF (Azure) Platform, which becomes the baseline for development of Solution Architecture by Tenants

KAISER PERMANENTE

Analytic Digital Foundation Pillars



KAISER PERMANENTE®

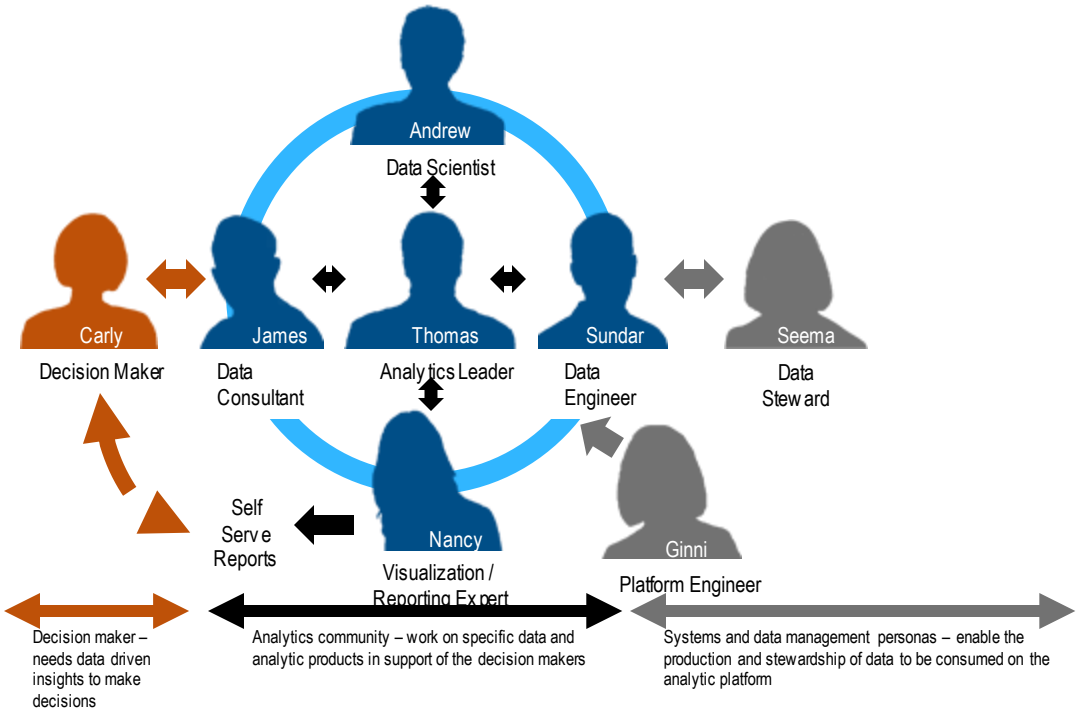
Analytic Capability Framework and Personas

The analytics capability framework, as noted below, shows the end to end process from planning and managing analytics to enabling and delivering insights that will be made possible through the development and use of the Analytic Digital Foundation (ADF).

KAISER PERMANENTE®

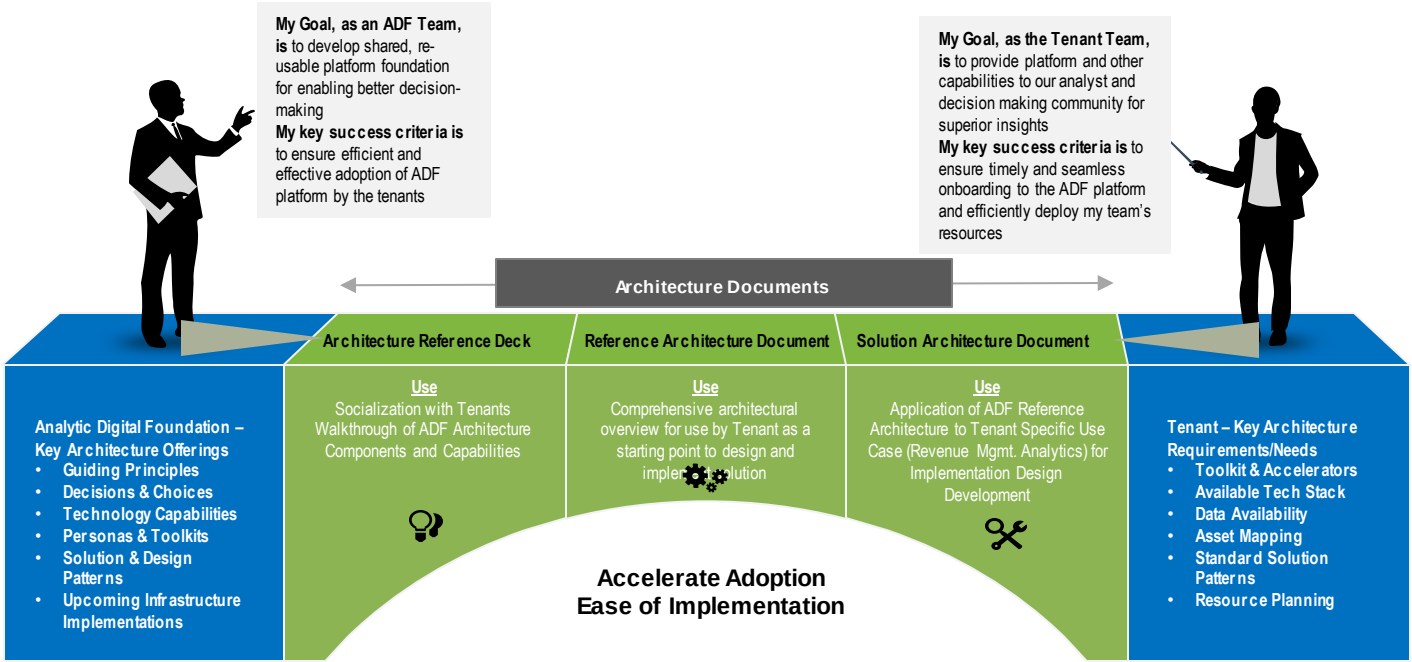
Persona Interaction Patterns

The interactions among personas are complex and vary depending on the activity being executed. The illustration below captures common interaction patterns among analytic personas:



Activities underway by Chief Data Office to update persona titles

ADF Reference Architecture Materials – Value & Use

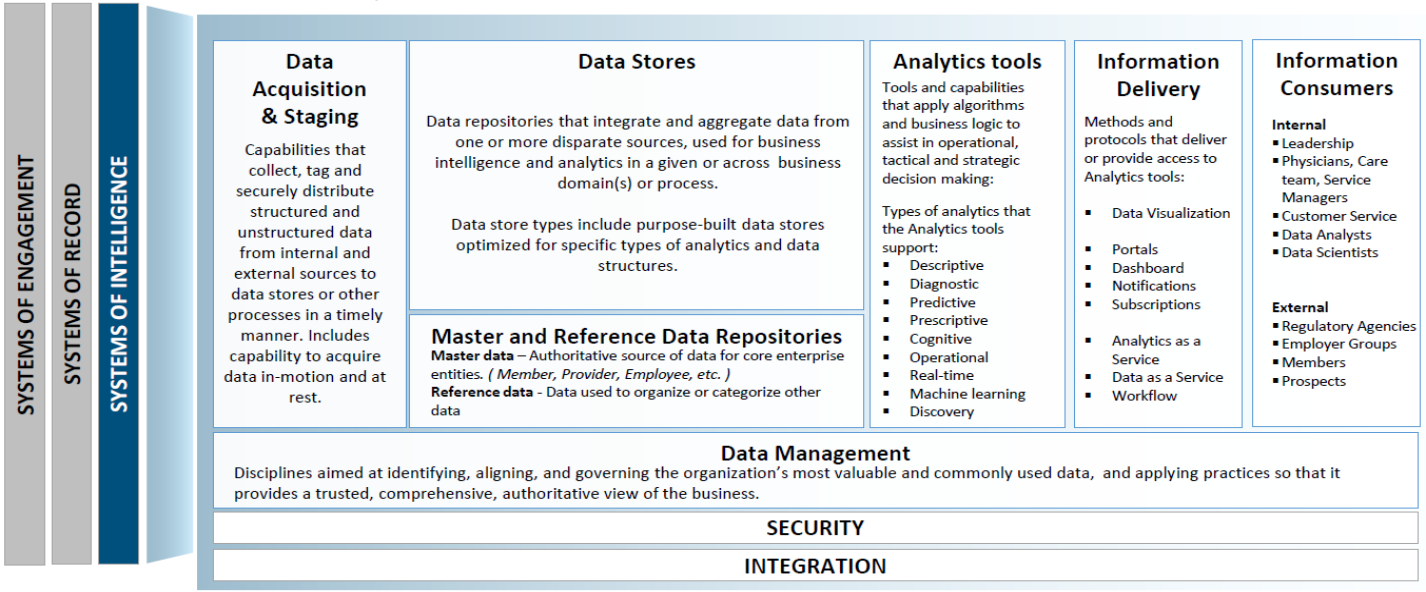


KAISER PERMANENTE®

Systems of Intelligence (SOI) Architecture

Systems of Intelligence – Level 0 Contextual Architecture

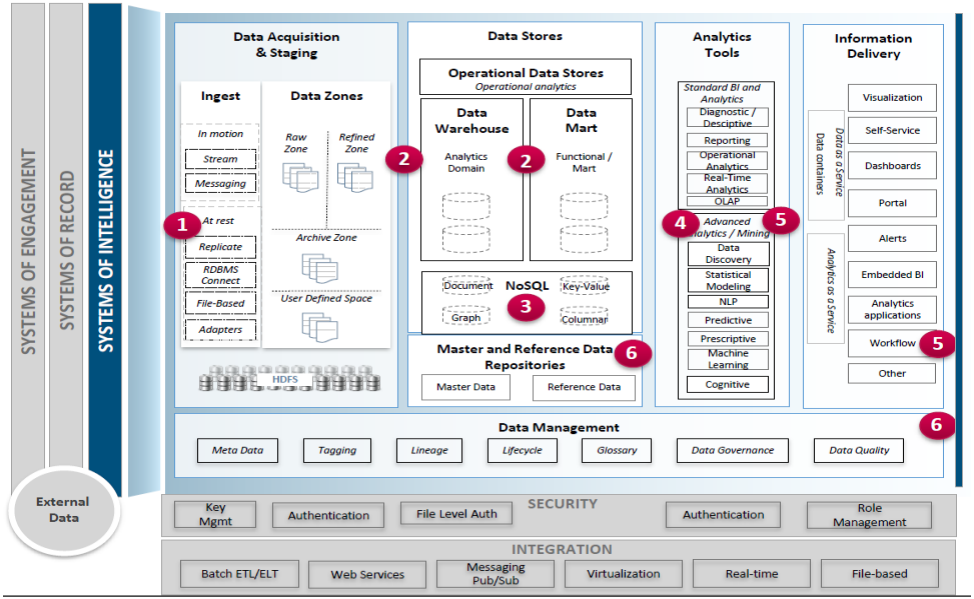
SOI represents the capability to transform data into meaningful and actionable information which aids KP in making informed business decisions and providing needed business support



KAISER PERMANENTE®

SOI Logical Architecture

SOI technology landscape is rapidly evolving and has the opportunity to significantly contribute towards competitive advantage for KP. Here are a few perspectives and its implications:



1. **Data Streaming, Device data integration**
   - The ability to ingest, in real time, data coming from devices, sensors, etc.,

   Implications: Enablement of real-time ingest for patient remote monitoring data, KPHC data
2. **Transition from ETL to ELT, DW/DM on Hadoop**
   - Reduction of point to point ETL processes and migration of DW capabilities to Hadoop

   Implications: Moderate to Significant reduction of Informatica PowerCenter Footprint
3. **NoSQL data stores for efficient processing of structured and un-structured data**
   - Key-Value and Columnar stores enable entity-centric (e.g. patient, consumer) views across disparate data sets.
   - Graph data stores enable capture and analysis of inter-entity relationships.

   Implications: Clinical Data Store/Patient 360, marketing use cases
4. **Advanced AI-based analytics**
   - Novel analysis techniques and algorithms allow research in e.g. clinical domains, such as cancer research, over vast amounts of data.

   Implications: Patient Self-triage for low acuity conditions, Genomics
5. **Workflow-embedded analytics**
   - The ability to create an analytics 'pipeline' enabling timely intelligence in operational processes.

   Implications: Prediction medicine, Scaffold
6. **Robust Data Management & Master Data**
   - Robust Data Management practices allows easy access, timely and higher quality decisions.

   Implications: Increased patient Chart Accuracy, Governed data sharing across line of business

KAISER PERMANENTE®

Foundation Services

Key Objectives

- Describe the A20 Program Architecture guiding principles, technology choices
- Provide an overview of A20 Program Solution Patterns, Toolkits that enables tenants
- Describe A20 Program Non-functional Requirements and Foundation Service/DR plans

Outline

- Guiding Principles & Architecture choices
- ADF Non-Functional Requirements & Reference Architecture
- Foundation Services & Capabilities Roadmap
- Technology Standards
- Personas & ADF Toolkits
- Solution Patterns for Architecturally Significant Use Cases
- Disaster Recovery – Current and Planned

Architecture Guiding Principles

Open Standards Based

Solution architecture is designed and based on open-standards and open-source framework for purposes of application development and use.

Elastic Platforms

Leverages the power of elastic, scalable platform capabilities provided by Microsoft Azure for cloud computing.

Modern Architecture

Built on modern architecture principles (such as cloud, loose coupling, late binding, etc.) that foster data and platform adaptability.

Data Flow Efficiency

Institute standard pipelines and norms to reduce and standardize data movement.

Intelligent Data Access

Facilitate access to the data where it is located (moving users/processing to the data, not vice versa).

Reduce Data Latency

Employ standards and structure (data zones, staging etc.) to reduce latency between data acquisition, retrieval and consumption.

Avoid Data Duplication

Leverage existing governed platforms and data assets where applicable to limit new data creation and duplication.

Data Zone Modularity

Facilitate data optimization through segregation of the data zones for its specialized purpose. This will result in enhancing the value of the existing data by effectively utilizing the value of the data across each zone vertically and laterally across each of those zones.

Self-Service Enablement

Self service capabilities, enabled through the ADF platform, will permit users to access data easily for use themselves.

Data Lineage and Tagging

Data lineage and tagging processes enable the enterprise to comprehensively monitor the data moving through the ADF platform for accuracy and quality.

Data Security

Include data security components to ensure secured access, protection, tagging, and classification of data.

Platform Security

Secured by a unified security architecture, consisting of an integrated set of security software, systems and processes.

General | Data Acquisition and Staging | Data Stores | Analytics | Data Disciplines

KAISER PERMANENTE®

Architectural Choices

The following decisions have been made by Kaiser Permanente in the course of developing this architecture. They are captured here to provide the necessary background to those interested in deeper rationale for the architecture.

Platform Choices

Use of Azure Platform-as-a-Service

Microsoft Azure Platform-as-a-Service option was selected to accommodate for the need for quicker speed to execution, elasticity and scalability of the platform, on-demand service from Microsoft certified technology teams, etc.

Hive First (but not Hive only)

The **Apache Hive™ will be used first** for access of data across data zones. A "Hive First" guideline will be followed on ADF platform for data access and use. Use of the **SQL databases** will be **limited only to the Tenant Zone.**

Use of Azure Native Platform Tools

Use of **natively assessed (by KP TRO, etc.) tools** on the Microsoft Azure platform will be **prioritized** across all functions of the ADF platform. Next priority is placed on open source third party tools that are also natively assessed before using other third party tools.

Data Choices

Modern Data Architecture

A data storage architecture that is built and based on **independent data zones** is a key architectural component of the ADF platform. This will replace the legacy architecture which depended on data warehouses and data marts.

Use Tenant Zone for Analytics

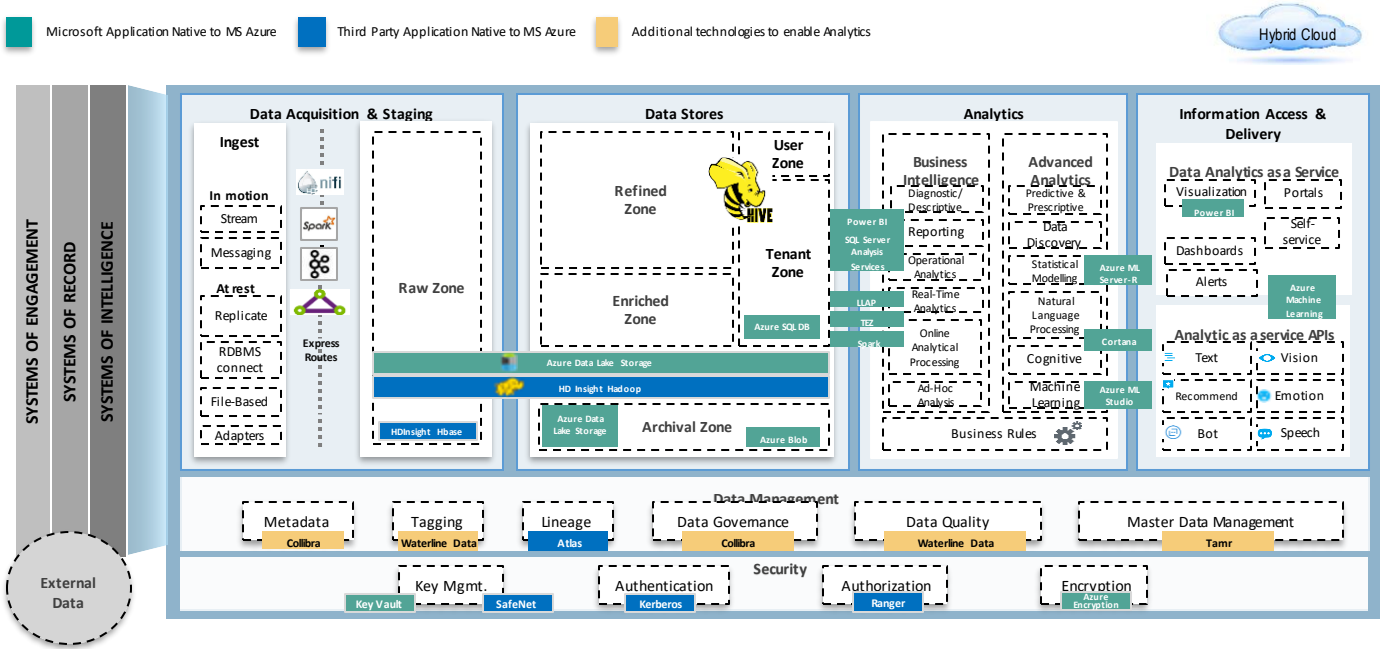
Dominant **analytic capabilities**, available on the ADF platform, will utilize primarily the **Tenant Zone** for purposes of data provision and data analysis.

Limit Data Movement to Tenant Zone

Data required from other zones, (e.g., Refined Zone and Enriched Zone) will be pointed to for use within the Tenant Zone itself through the **logical architecture schemas** (i.e. data will not be physically pulled into the Tenant Zone).

Foundation Services – Overview

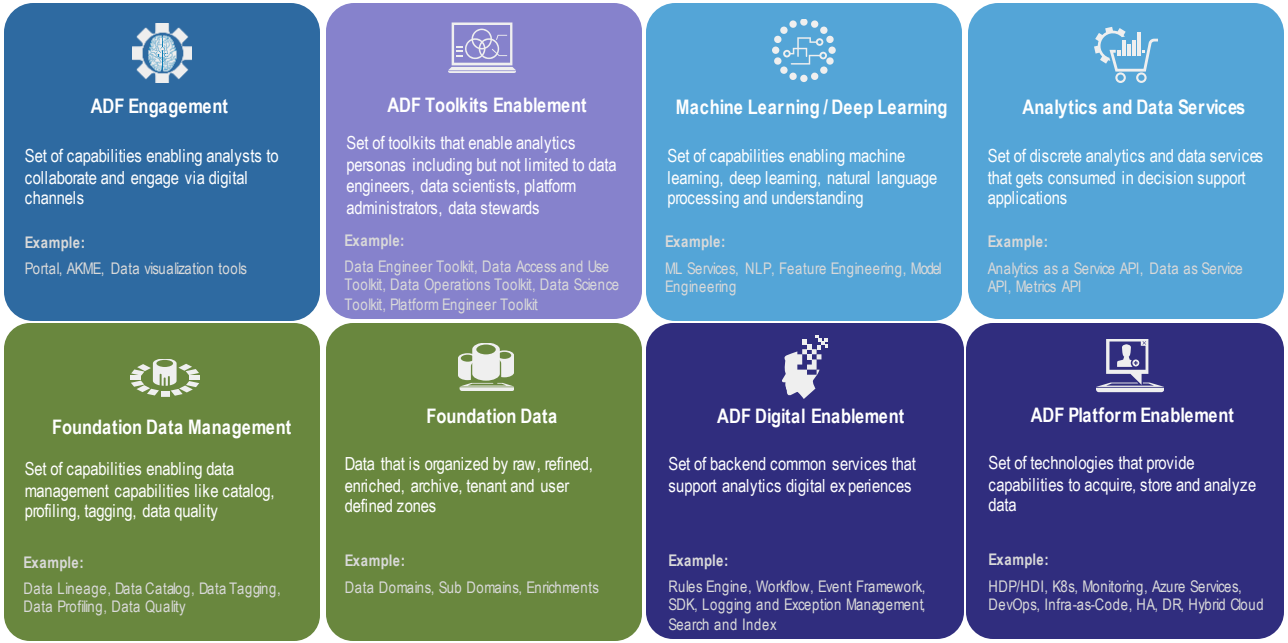
ADF Reference Architecture Diagram (As of April 2019)



The following slides outline the high level features and key technology catalog for each component of the reference architecture.

KAISER PERMANENTE®

Analytic Digital Foundation Pillars



KAISER PERMANENTE®

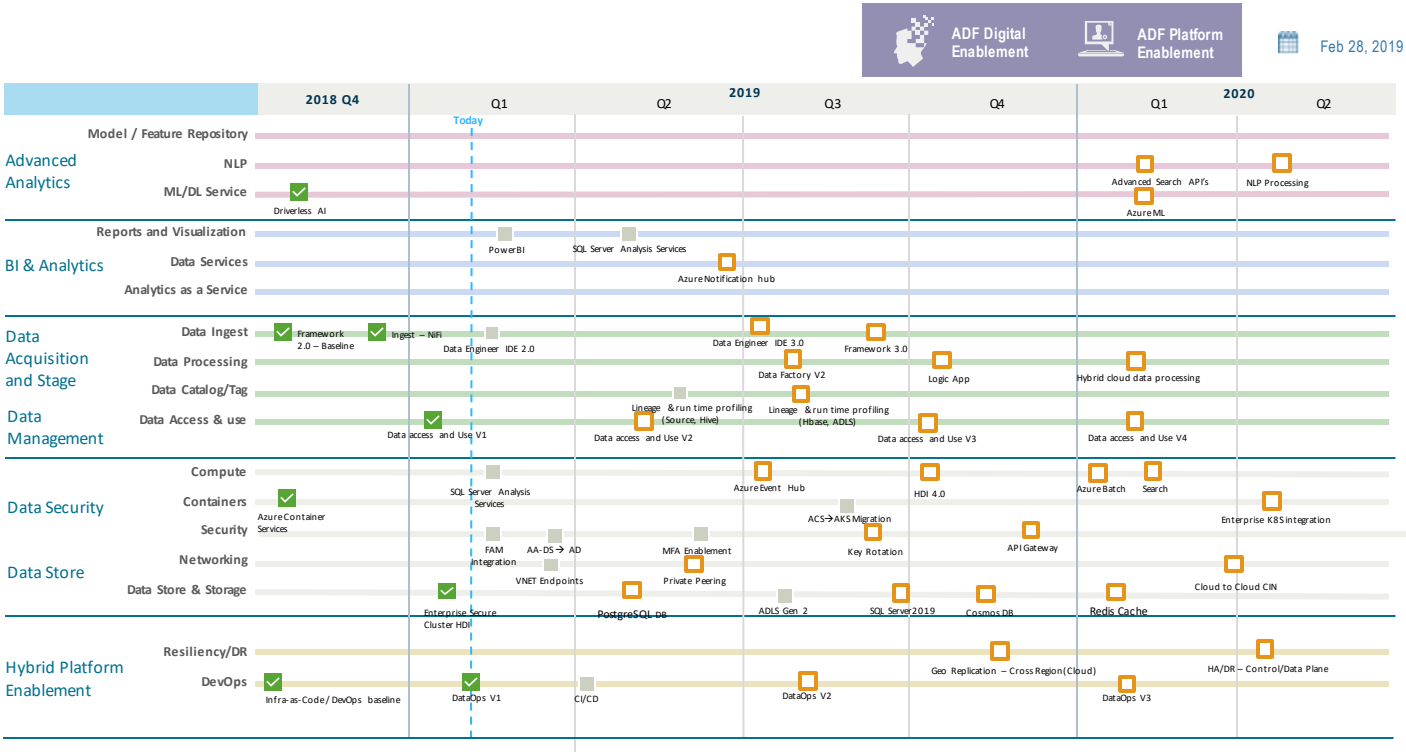
Analytic Digital Foundation Capabilities

ADF Engagement	**E1: Omni-Channel Portal** Provide decision Makers to consume insights via various channels including portals	**E2: ADF Intake Self-Service** Provide ability to intake requests for ADF and service requests	**E3: Analytics Knowledge Management (AKME)** Provide ability for analysts to collaborate and manage analysis workflow			
ADF Toolkits Enablement	**T1: Data Engineer Toolkit** Capabilities and tools for Data Engineers to onboard data to ADF	**T2: Data Scientist Toolkit** Capabilities and tools for Data Scientist to build and deploy predictive models	**T3: Data Operations Toolkit** Capabilities and tools to operate data on the ADF	**T4: Data Access and Use Toolkit** Capabilities and services to manage data access and use on ADF	**T5: Data Analyst Toolkit** Capabilities and tools for Analysts to manage data, set thresholds on the ADF	**T6: Platform Engineer Toolkit** Capabilities and tools for platform engineers to operate ADF platform
Machine Learning/ Deep Learning	**M1: Feature Repository** Provide ability to engineer and store features	**M2: Model Repository** Provide ability to develop and store predictive models	**M3: Machine Learning Pipeline** Provide ability to develop machine learning pipelines	**M4: NLP Processing, NLU** provide capability to develop and enable Natural Language Processing and Natural Language Understanding		
Analytics and Data Services	**S1: BI and OLAP Services** provide ability to deliver on Business Intelligence and OLAP Services	**S2: Data and Analytics API Services** Provide ability to deliver data and analytics as a service for consumption	**S3: Metrics Services** Provide ability for analysts to manage and publish business metrics			
Foundation Data Management	**D1: Data Catalog** Provide capabilities to manage and provide two way interface to enterprise catalog solution	**D2: Data Lineage** Provide ability to capture data lineage that is integrated into the data pipeline	**D3: Data Profiling** Provide automated capability to profile data on ADF	**D4: Data Tagging** Provide automated capability to profile data on ADF	**D5: Data Quality Mgm.** Provide ability to manage quality of data on ADF	
Foundation Data	**F1: Data Domains** Data that is organized, managed by information domains available for analytical processing	**F2: Analytical Insights** Data that is organized by logical groupings of information for purpose of decision making	**F3: Shared Enrichments** Data that is derived shared data, generated by manipulating and blending data from the Raw and Refined Zones	**F4: Master Data** Master data for core entities that permeate the enterprise	**F5: Reference Data** Internal and External reference data	
ADF Digital Enablement	**N1: Event Framework** Enable actions based on events and provide ability to manage dependencies amongst events and provide multi-channel notification	**N2: Analytics Workflow Mgmt.** Enable workflows to analysts and data stewards to manage the information delivery workflows	**N3: ADF Logging Framework** Provide ability to manage logging and exception management	**N4: ADF Search and Index Mgmt.** Provide ability to search on structured and unstructured data on ADF	**N5: Software Development Toolkit (SDK)** Provide ability for software developer to extend and develop new data nodes and contribute towards the ADF toolkit	
ADF Platform Enablement	**P1: DevOps** enhance speed through tooling, automation of software deployment	**P2: Infrastructure as Code** Enhance speed through tooling, automation of infrastructure deployment	**P3: Resiliency/DR** Enabling service resiliency and Disaster Recovery of ADF	**P4: MS Azure Platform Service Enablement** MS Azure Services to enable data processing, data store and analytics	**P5: Hybrid Cloud Enablement** enabling capabilities to leverage on premise and cloud infrastructure that provides the flexibility to optimize for both speed and cost	

Engage Service Data Toolkit Hybrid Cloud Platform

KAISER PERMANENTE®

Foundation Services Capability Roadmap*



Note:* Planned services are in process of re-prioritization as a result of 2019 budget reduction

Planned Scheduled Completed

KAISER PERMANENTE®

Technology Standards Employed by ADF (1 of 3)

Catergory*	Tool Name	Purpose	Rationale
Ingest	HDInsight Spark	HDInsight Spark is built on top of Spark core SQL engine, it provides scalable and fault-tolerant stream processing capability	Spark has now been accepted as an industry standard engine for processing distributed data sets. HDInsight Spark is Hortonworks's version of Spark available on Microsoft Azure, part of PaaS service HDInsight. Since ADF has standardized on Azure, HDInsight Spark is best choice for processing distributed datasets
	HDF Nifi	HDF Nifi is an open-source software to automate the flow of data between software systems	HDF Nifi automates flow of data between systems. Nifi is a scalable platform, i.e. more compute nodes can be added if the need grows. It allows creation of reusable components that enables standardized implementation of extract jobs. The reusable components includes, error detection, logging, and notifications
	HDInsight Kafka	HDInsight Kafka is a distributed streaming platform, it provides low-latency, high-throughput, fault-tolerant publish and subscribe pipelines and is able to process streams of events	Kafka is HDInsight PaaS offering and as a result, represents a natural choice for stream processing needs on ADF
	Azure Databricks**	Azure Databricks is a Spark platform designed by Microsoft and Databricks company. The platform provides advanced capabilities for cluster management and scaling	Azure Databricks Spark platform provides auto scale up/down capabilities and Microsoft is making it available as PaaS component. Databricks Spark provides additional capabilities/tools not available in HDInsight Spark, such as, Runtime, Delta, Workspace, Production jobs and workflows, etc. Additionally, ADF may migrate to Azure Databricks spark version after it is approved by KP TRO.

- *Please reference the category definitions available in the Appendix section of the ADF Reference Architecture Word Document*

** *Planned*

Technology Standards Employed by ADF (2 of 3)

Catergory*	Tool Name	Purpose	Rationale
Storage	Azure Data Lake Storage (ADLS)	An enterprise-wide hyper-scale repository for big data analytic workloads. Can capture data of any size, type, and ingestion speed in one single place for operational and exploratory analytics	ADLS is the default distributed storage offering from Microsoft, thereby making it an ideal choice for use on the ADF platform
	Azure Blob Storage	Azure Blob storage is Microsoft's object storage solution for the cloud. Blob storage is optimized for storing massive amounts of unstructured data.	
Staging	HDInsight Hadoop	HBase is a column-oriented database management system that runs on top of Hadoop Distributed File System (HDFS)	Hadoop framework and tools on Microsoft Azure is from Hortonworks HDP platform and are branded as HDInsight Hadoop. This product is offered as a PaaS component on Azure that aligns with ADF architectural principles of using Azure native techs and use of PaaS components
	HDInsight Hive	HDInsight Hive data warehouse software facilitates reading, writing, and managing large datasets residing in distributed storage using SQL. Structure can be projected onto data already in storage.	Hive supports relational views of the data is SQL compliant, these are essential requirements for ADF data zones. In addition to this, there are no limits as far as data size is considered, it can also support hundreds of concurrent users with the help LLAP (for performance) and scaling up the compute nodes. SQL DW was considered as an option for Refined and Enriched zones but it currently has limitations as to the number of concurrent queries and cross database it can join
Data Warehouse and Data Marts	Azure SQL DB	SQL Azure is Microsoft's cloud database service. Based on SQL Server database technology and built on Microsoft's Windows Azure cloud computing platform, SQL Azure enables organizations to store relational data in the cloud and quickly scale the size of their databases up or down as business needs change.	SQL DB is used for enabling tenant specific marts (currently there restrictions on what type of data can be stored on SQL DB).

** Please reference the category definitions available in the Appendix section of the ADF Reference Architecture Word Document*

KAISER PERMANENTE®

Technology Standards Employed by ADF (3 of 3)

Catergory*	Tool Name	Purpose	Rationale
Caching	Azure Redis Cache	Azure Redis is a fully managed, open source, in-memory data storing service that powers fast, high-performing applications	Azure Redis is offered as a PaaS component on Microsoft Azure and provides all desired capabilities required for data caching in web or mobile application
Standard BI and Analytics	Power BI	Power BI adds data visualizations, interactive reports and dashboards to applications seamlessly. Analysis Services allows use of advanced mashup and modeling features to combine data from multiple data sources, define metrics, and secure your data in a single, trusted tabular semantic data model.	Power BI is PaaS/SaaS offering from Microsoft for data visualization and analytical reporting. In terms of capabilities, it is comparable with Tableau.
	SQL Server Analysis Services	Analysis Services is an analytical data engine used in decision support and business analytics. It provides enterprise-grade semantic data models for business reports and client applications such as Power BI, Excel, Reporting Services reports, and other data visualization tools.	SQL Server Analysis Services is being used temporarily until security related issues are resolved for Azure Analysis Services
Metadata	Collibra	Collibra is used for data governance. It's a cross-organizational data governance platform that delivers openness and freedom. It breaks down traditional data silos and opens up organizational data so all users can find the data they need, collaborate on it, and easily understand its meaning.	Collibra is a KP Standard tool utilized for data governance purposes
Tagging	Waterline Data	Its platform discovers sensitive data, intermediate files, and data lineage, as well as enables data stewards to manage tags; and builds an inventory of Hadoop data.	Waterline Data provides necessary capabilities for tagging data in different zones It also works well with Apache Atlas, Ranger and Collibra that is essential for implementing data access security policies
Lineage/ Governance	Apache Atlas	Atlas is a scalable and extensible set of core foundational governance services – enabling enterprises to effectively and efficiently meet their compliance requirements within Hadoop and allows integration with the whole enterprise data ecosystem	Apache Ranger utilizes Atlas to implement data classification based security policies

** Please reference the category definitions available in the Appendix section of the ADF Reference Architecture Word Document*

KAISER PERMANENTE®

Foundation Services – Toolkits

Personas

Various non-functional requirements have been considered for development of the Analytic Digital Foundation platform, as listed below:

Decision Maker

Use data-driven insights to make high quality decisions. The quality of the decision will depend on the integrity of the data, quality of its analysis and skill of interpretation

Data Consultant

Work with business partners to define and scope the problem, understand the workflows that affect data context, and explain and interpret analytic outputs to the business partners

Analytics Leader

Coordinating analytics tasks for their organization, including creating effective strategies to collect data, analyze information, conduct research, and implement analytics solutions for their products or services

Visualization Expert / Report Builder

Create actionable business insights through the development of data visualizations, reports, and dashboards

Data Engineer

Responsible for the development and maintenance of data pipelines and analytics infrastructure that enables almost every other function in the data world

Data Scientist

Extract meaning from and interpret data, which requires both tools and methods from statistics and machine learning

Platform Engineer

Develop and maintain systems and performance

Data Steward

Manage data for a specific information domain, including improving data maturity over time, applying data standards that are relevant to the domain, and helping create remediation plans for data quality issues

Activities underway by Chief Data Office to update persona titles

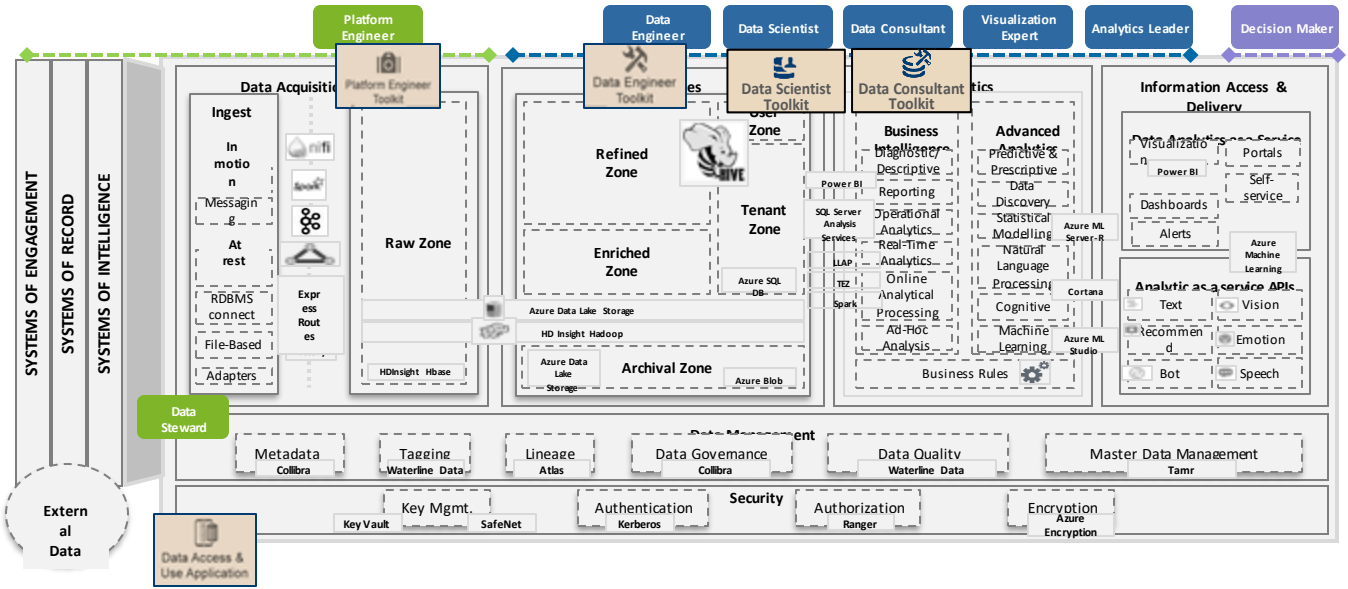
ADF Reference Architecture – Personas & Toolkits (As of March 2019)

Systems and Data Sourcing Personas

Analytics Community Personas

Decision-Making Persona

Note: *Data Science Toolkit, while not represented in the architecture below, is under development and will be added upon completion*



KAISER PERMANENTE®

Toolkit – Data Engineer (1 of 5)

Category	Description
Purpose	The primary purpose of the data engineer toolkit, deployed on the ADF platform, is to support the data engineer and enterprise data team design and develop a stable data analytics infrastrcuture and environment. This toolkit enables the data engineer to maintain the various databases, datasets and other data assets in an optimum level so as to facilitate other users (such as data scientists, consultants, etc.) to focus on advanced data modeling and statistical analysis to support superior insights generation by end users.
Features	Key components enabled by the use of this toolkit include the following: 1. Data Transformation and processing (e.g., ability to ingest data files from network, distributed files systems, SFTP, ADLS, Blob, ability to combine data from multiple sources and enable parallel processing, apply business rules, etc.) 2. Data pipeline and scheduling (e.g., ability to create complex workflows – multi-step data processing, ability to logically organize workflows by user groups and projects/target, etc.) 3. Alerts and notification (e.g., ability to logically organize workflows by user groups and projects/target) 4. Adoption and ease of use (e.g., ease of use by non-technical and business users, browser based user interface, etc.) 5. DevOps, CI/CD support (e.g., easy to deploy on Microsoft Azure, ability to utilize resources effectively, ability to scale up or down, integration with Kaiser source code management system and DevOps pipeline, etc.) 6. Integration with advanced analytics and machine learning tools
Personas	Prominent personas utilizing this toolkits, include: 1. Data Engineer 2. Data Steward
Underlying Technologies	Java, Scala, AngularJS, Spark, Kafka, SQL, etc.

Toolkit – Platform Engineer (2 of 5)

Category	Description
Purpose	The primary objectives of the platform engineer toolkit are i) to help manage detailed standards and operating procedures related to the creation, storage, processing and consumption of various data assets on the ADF platform. This toolkit enables the operations team to effectively monitor and manage the data stores assets, data transformation processes and related technologies. ii) platform technology team to maintain and monitor the platform infrastructure and performance for the ADF users. A stable ADF platform environment, powered by the capabilities of this toolkit (Infrastructure-as-a-Code, Platform-as-a-Service tools), not only ensures expansion of the platform (scalability) but also helps extracts maximum analytics value through integration and execution of various advanced technologies
Features	Key performance metrics enabled by this toolkit include the following: 1. Sustaining system performance and reliability 2. Executing improvements on the platform 3. Monitoring and maintenance of a stable and secure platform environment 4. Log Analytics Metrics (e.g., cluster count, collect usage metrics at environment level, collect Kafka metrics, foundational data metrics for enriched zone in hive/SQL DB, foundational data metrics for Raw Zone in HBase) 5. Platform Operations Metrics (e.g., HDInsight components metrics 6. Hadoop services level metrics, Kafka disk usage, Ambari metrics, application metrics, yam queue metrics, availability, edge node application metrics, uptime/downtime clusters/nodes) 7. Storage Metrics (Blob, ADLS) (e.g., request latency, authorization error, access report, error rate, data volume at zone level, file count, network error, egress metrics, ingress metrics) 8. Databases (SQL DB, SQL DW, Cosmos DB) (e.g., database alerts, performance, audit report, security alerts, vulnerability detection, usage reports) 9. Jenkins Pipeline Metrics (e.g., build metrics, code quality metrics, security scan metrics, exception/error metrics) 10. Tools Metrics (e.g., NIFI metrics, Waterline metrics, Atlas service metrics) 11. On-Prem Cluster Metrics (e.g., On-Premise K8s Cluster Metrics, NIFI, ATLS cluster infra metrics) 12. Application Metrics (e.g., request and response metrics, data pipelines metrics (system/custom), ADF framework component metrics, memory utilization, data engineer metrics, application logs metrics, error/exception metrics, data flow errors/exceptions, application performance metrics) 13. Tenant Support Metrics (e.g., data pipelines metrics (system/custom) , ADF framework component metrics, application logs metrics,# of users, access privileges, data exposure, audit metrics, top used tables, users)
Personas	Prominent personas utilizing this toolkits, include: 1. Data Steward 2. Data Consultant 3. Platform Engineer
Underlying Technologies	Java, AngularJS, Kafka, SQL, etc.

Toolkit – Data Access and Use (3 of 5)

Category	Description
Purpose	The purpose of the data access & use toolkit is to empower the data steward teams to properly control and monitor access, provision and use of the data to ensure its continuous reliability and transparency for use. This toolkit enables the critical process of safeguarding the quality and trust of the data ingested and consumed on the ADF platform as the analytics generated depends on its data authenticity.
Features	Various metrics and processes are instituted to ensure proper custody and reliability of the data. Certain such key components include: 1. DAU ATLAS Integration (e.g., integrate DAU application and atlas to get hive metadata and associated tags) 2. Tenant Management (e.g. tenant onboarding, tenant admin provisioning, create tenant,add tenant tables (data custody based), map tenant admin with tenant) 3. Tag Based Policy (Hive) (e.g. create access policy, edit access policy,create mask policy,edit mask policy,revoke access policy,revoke mask policy) 4. KPIM, CARMA & TNS Integration (e.g. user access provisioning through KPIM. CARMA and TNS onboarding) 5. Resource Based & Tag Based Policy for SQL Server & Analysis Services (e.g. create access policy, edit access policy,create mask policy,edit mask policy,revoke access policy,revoke mask policy) 6. Advanced Schema Management (e.g. schema copy , new tables creation alerts, email notifications etc.) 7. Reports (e.g. view policy at schema level,view available tenants,view available tenant admins, audit reports,generate user entitlement)
Personas	Prominent personas utilizing this toolkits, include: 1. Data Steward 2. Data Consultant 3. Others (ADF Admin, Tenant Admin, etc.)
Underlying Technologies	Java, AngularJS, Ranger, Atlas, etc.

Toolkit – Data Consultant (4 of 5)

Category	Description
Purpose	The purpose of the data consultant toolkit is to enable data analysts and other business users to set up end to end analytics workflows that may involve human interaction and data processing on the ADF platform
Features	Certain such key features & components include: 1. Easy to use users interface to setup and monitor end to end analytics workflow pipelines 2. Allow data consultants to setup end to end analytics workflow pipelines 3. Allow data consultants to add & track human interaction nodes, such as, approvals, denials in the analytics pipeline 4. Allow data consutants to add data processing nodes/workflows in the pipeline 5. Allow data consultants to add notification objects, for ex. notify a user when a certain task is completed or failed 6. Allow data consultants to add notification objects, for ex. notify a user when a certain task is completed or failed
Personas	Prominent personas utilizing this toolkits, include: 1. Data Steward 2. Data Consultant 3. Others (ADF Admin, Tenant Admin, etc.)
Underlying Technologies	Java, AngularJS, Ranger, Atlas, Airflow, etc.

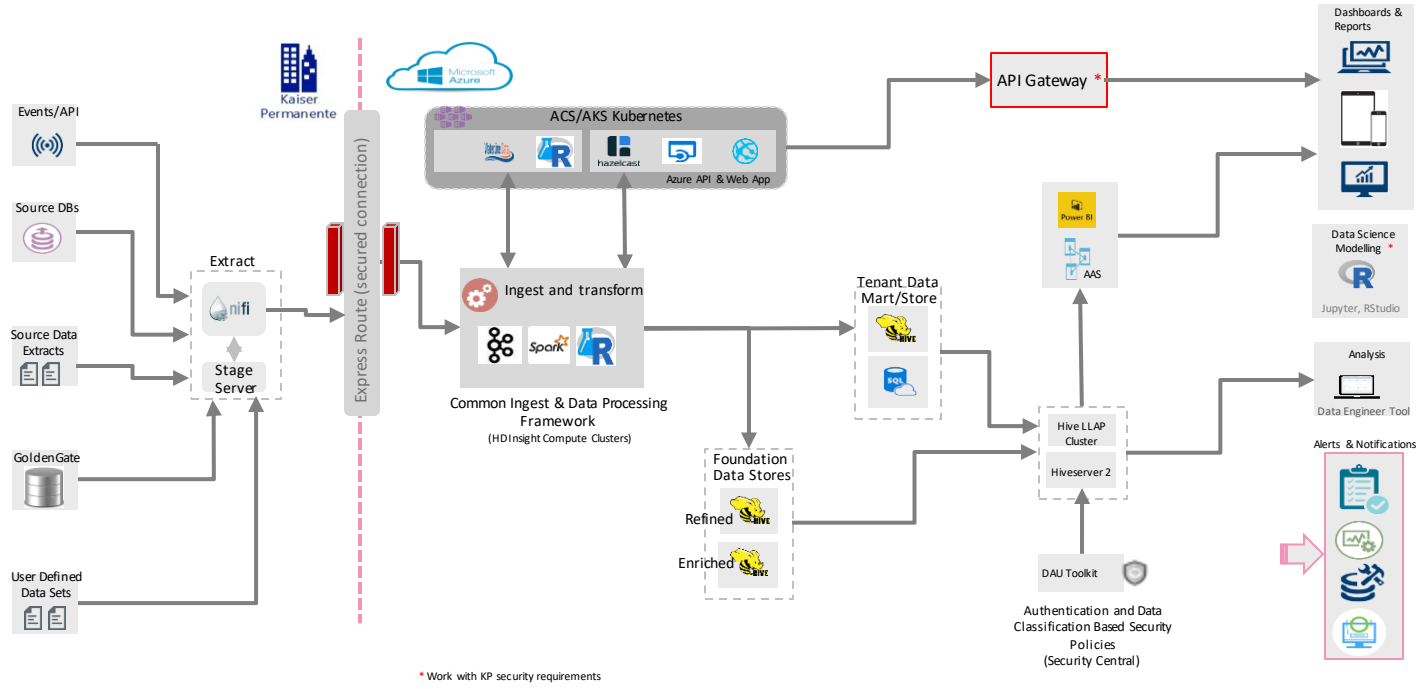
Toolkit – Data Scientist (5 of 5)

Category	Description
Purpose	The purpose of the data scientist toolkit is to enable data scientists to build, test, manage, and deploy advanced analytics models on ADF platform
Features	Certain key features & components include: 1. Provide tools required to build, test, and manage models 2. Integrtation with DevOps pipelines to deploy models 3. Provide secure environment to perform development activities 4. Provide ability to acquire required compute to deploy and execute models
Personas	Prominent personas utilizing this toolkits, include: 1. Data Scientist 2. Data Steward 3. Data Consultant 4. Data Engineer
Underlying Technologies	Jupyter Notebook, Java, AngularJS, Ranger, Atlas, other DS specific tools, etc.

Foundation Services – Solution Patterns

Solution Pattern – Overall View (1 of 10)

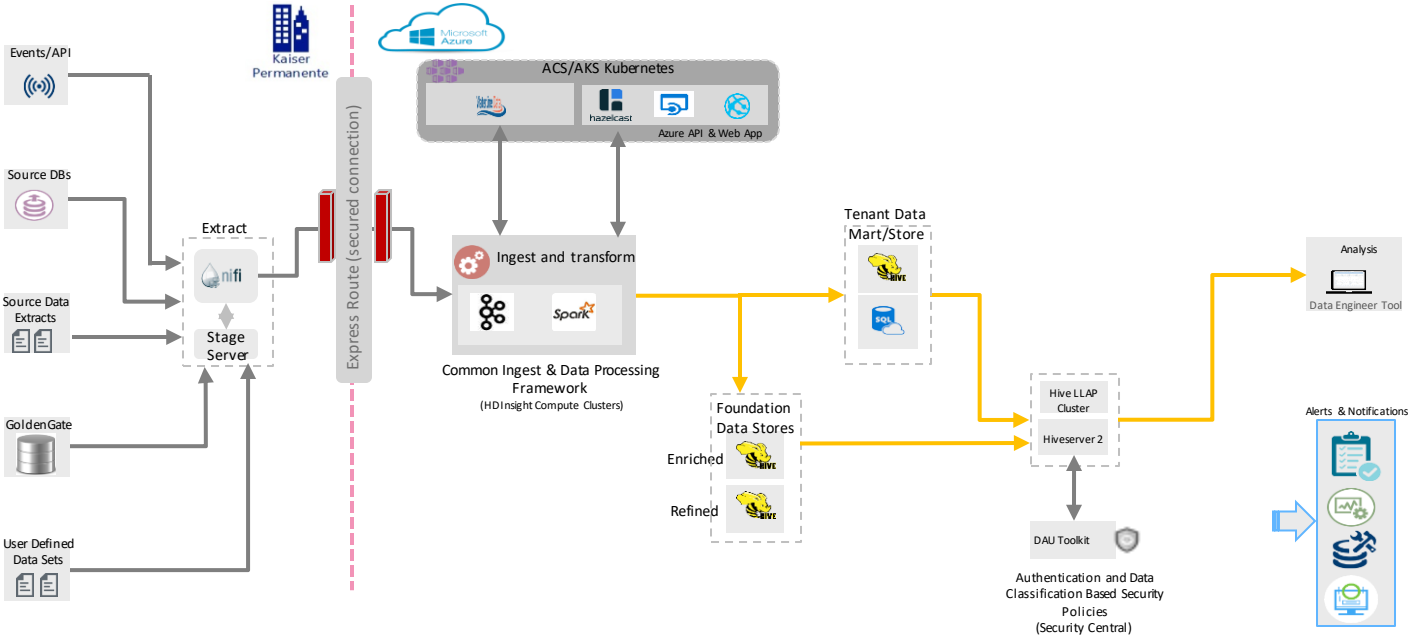
Standard solution and design patterns have been developed for the architecturally significant use cases for assisting tenant with their specific user cases



KAISER PERMANENTE®

Solution Pattern – Tenant Data Mart/Store (2 of 10)

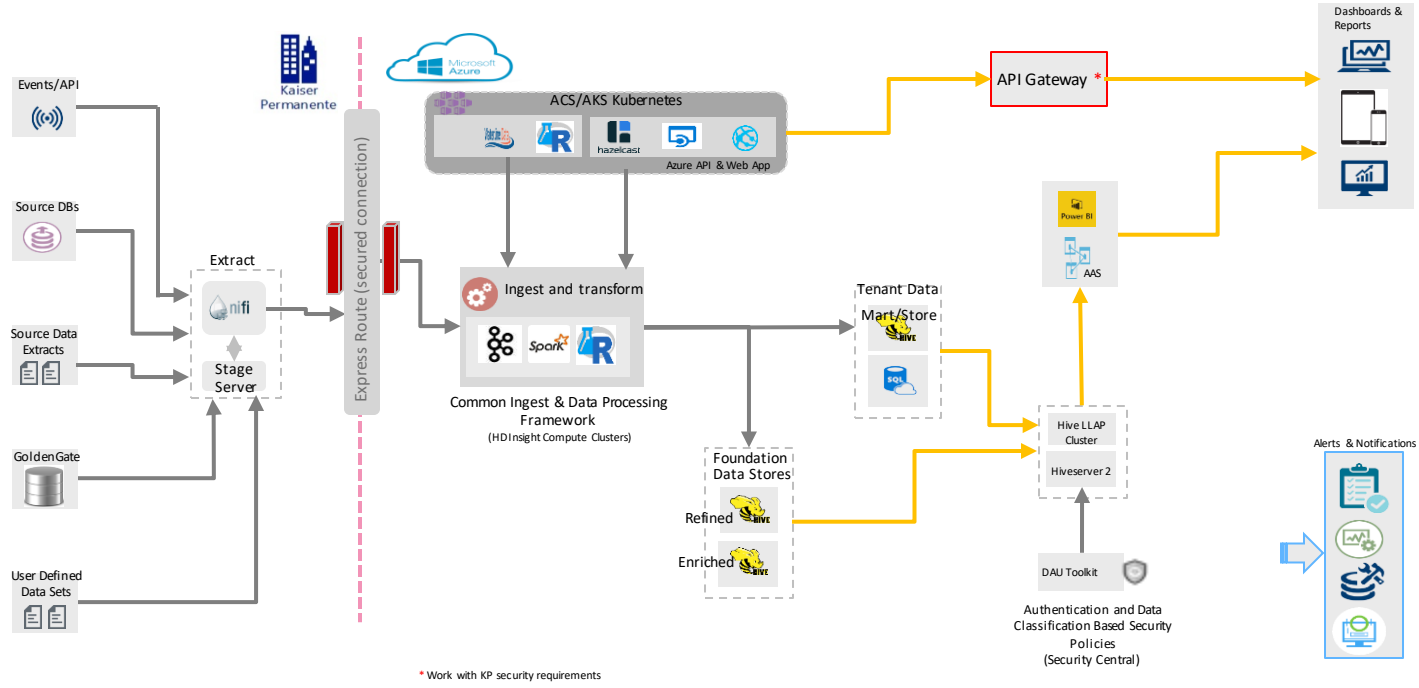
Solution pattern meets tenants need for a separate data mart/store. The data is typically sourced from Refined and Enriched data zones.



KAISER PERMANENTE®

Solution Pattern – Analytical Applications – Streaming & Batch (3 of 10)

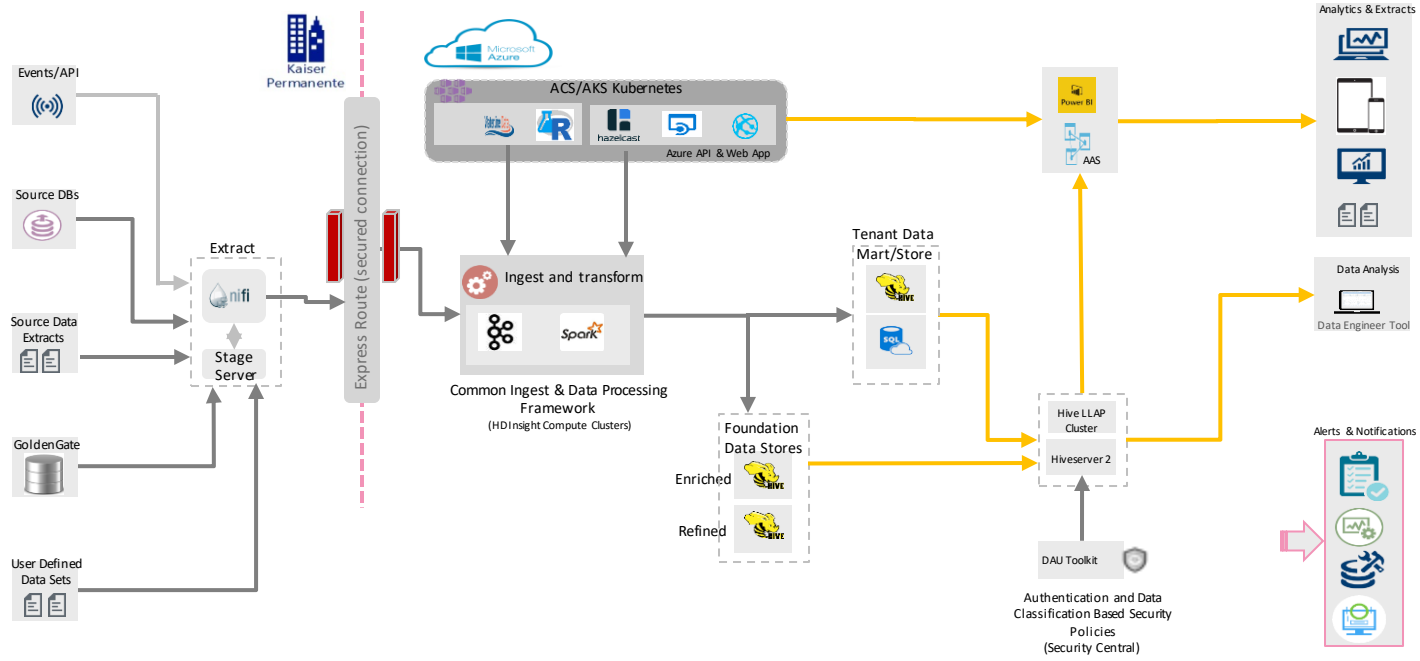
Solution pattern enables near-real time dashboards through mobile, web, desktops



KAISER PERMANENTE®

Solution Pattern – BI and Reporting (4 of 10)

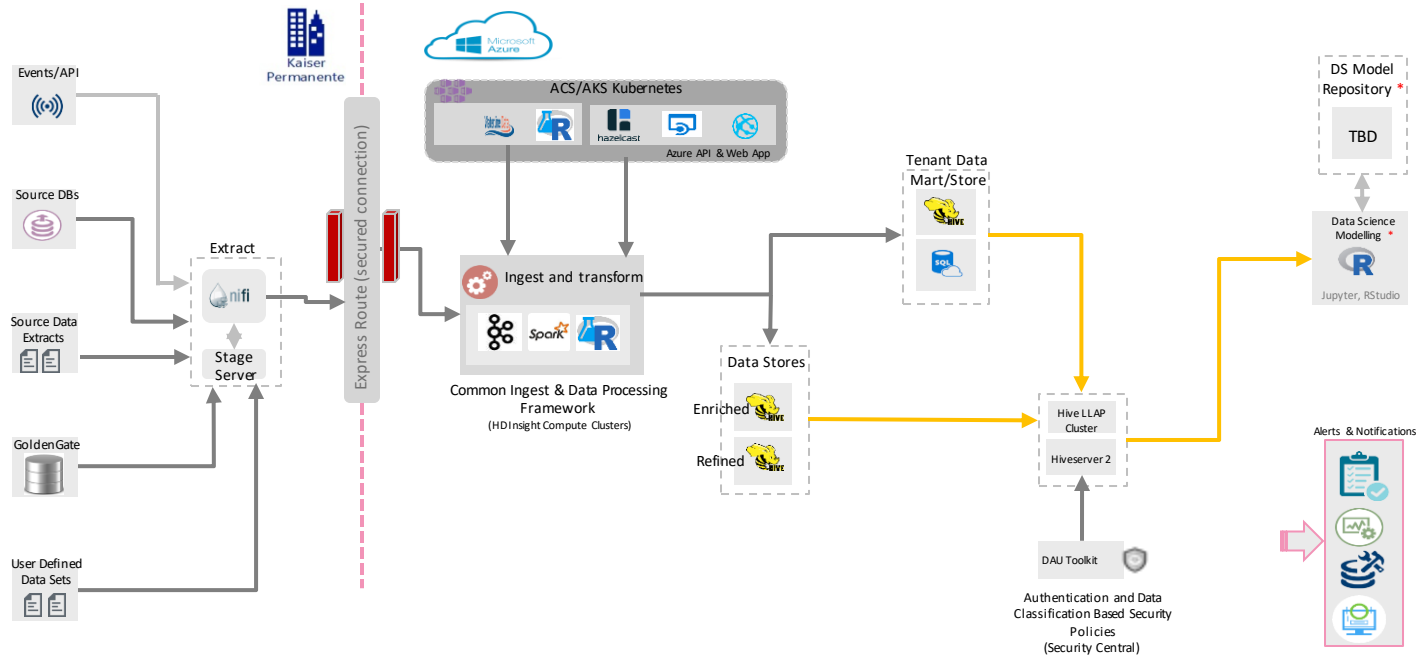
Solution pattern enables standard analytics and reporting capability with tools, such as, Power BI and Azure Analysis Services



KAISER PERMANENTE®

Solution Pattern – Machine Learning and Data Science Model (5 of 10)

Solution pattern provides for creating and training data science and deep learning models

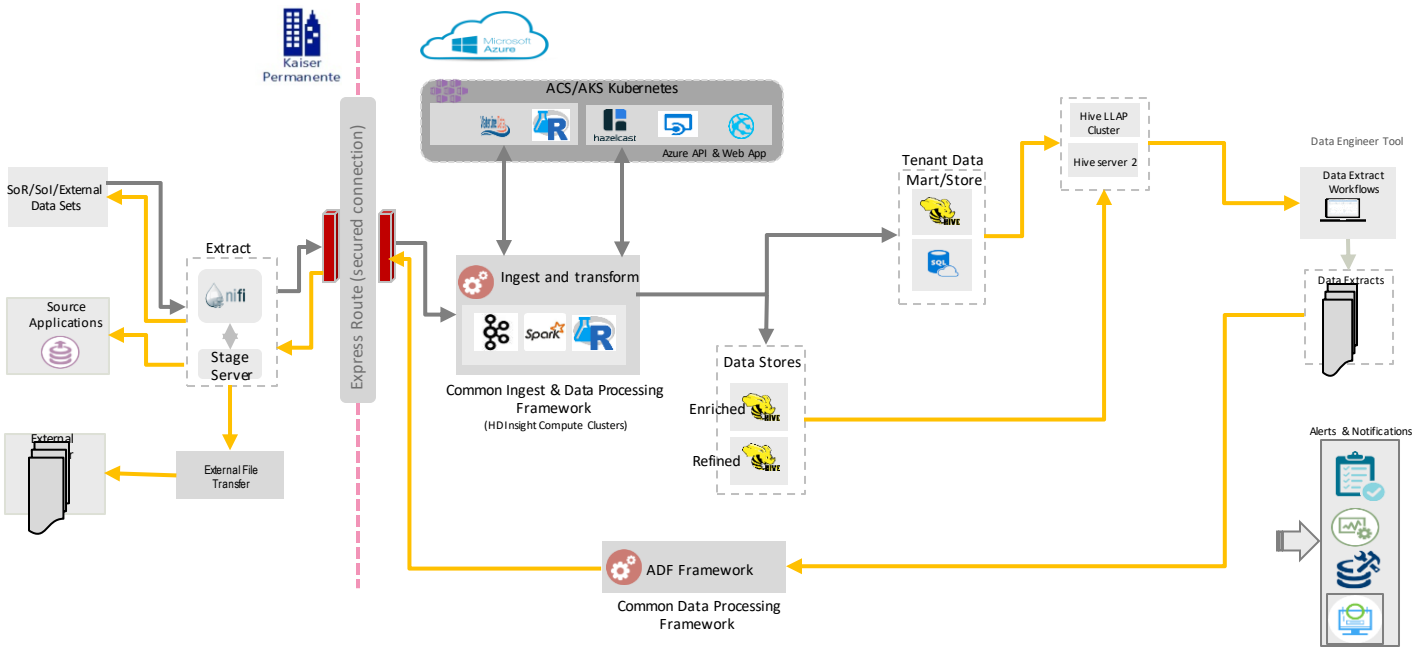


* Tools & Capabilities standards for analytics tool committee is working on defining standards

KAISER PERMANENTE®

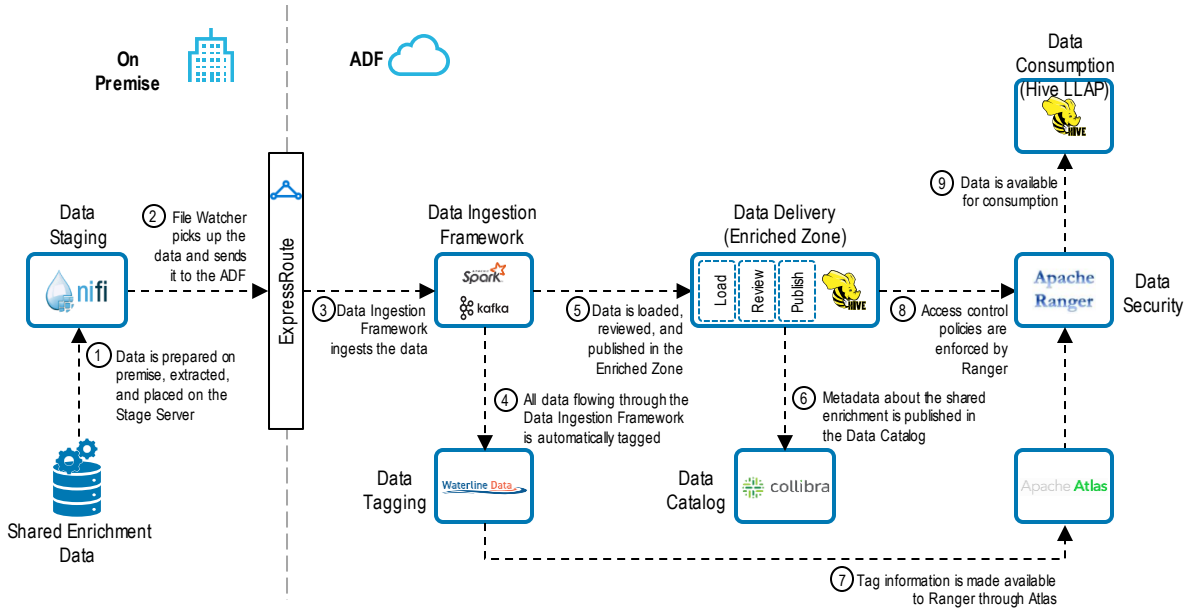
Solution Pattern: Third Party and/or SoR Data Integration (6 of 10)

Solution pattern provides for data integration between ADF Data Zones and Third Party Applications and/or Systems of Record (SoR)



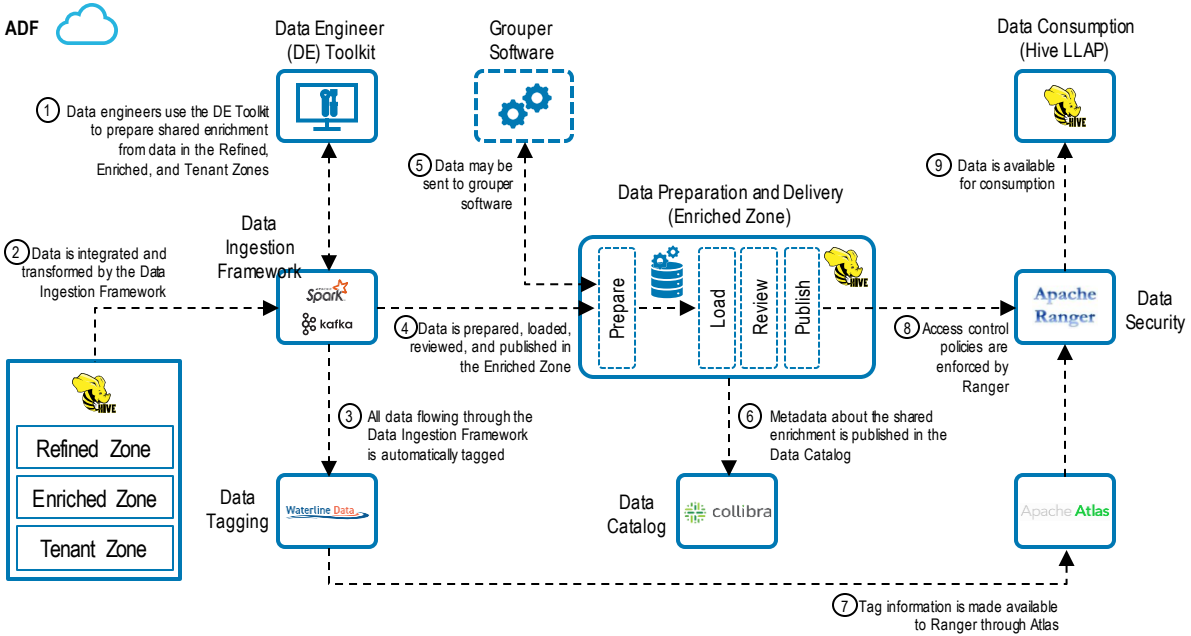
KAISER PERMANENTE®

Solution Pattern (Shared Enrichment Prepared On Premise) (7 of 10)



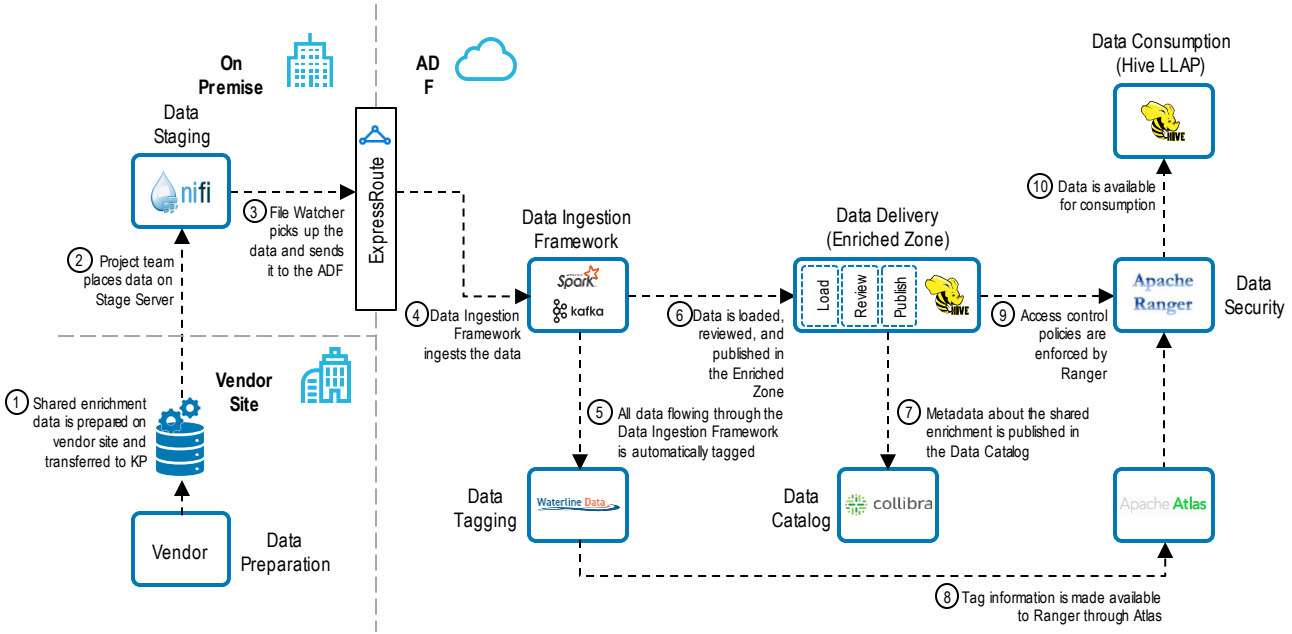
KAISER PERMANENTE®

Solution Pattern (Shared Enrichment Prepared On ADF) (8 of 10)



KAISER PERMANENTE®

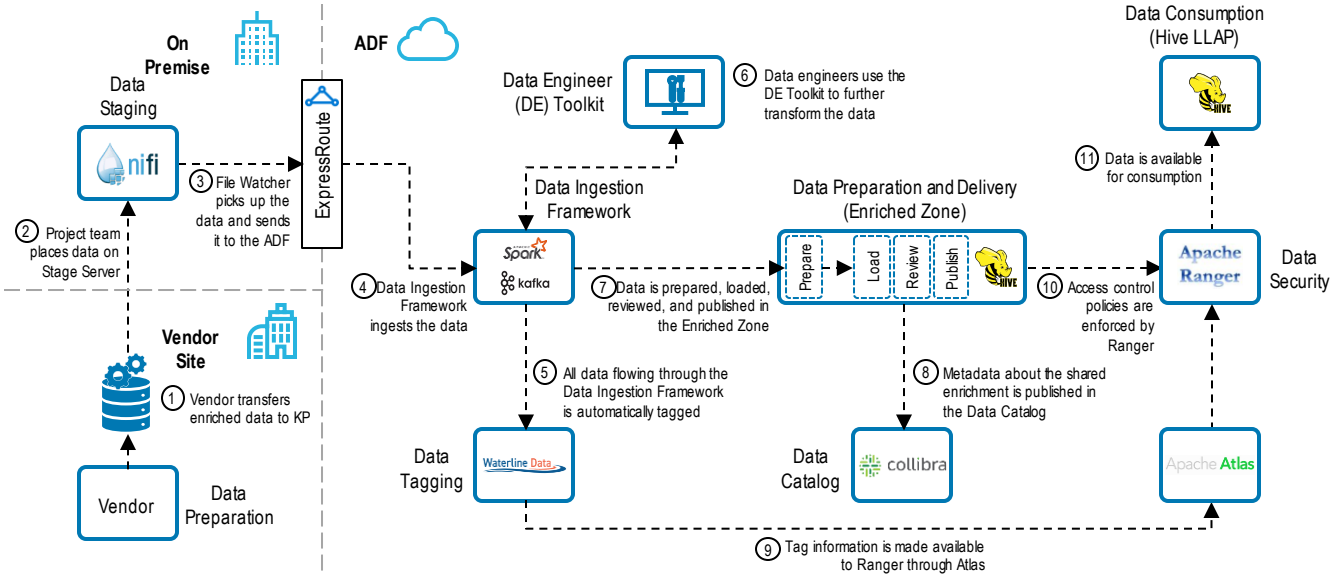
Solution Pattern (Shared Enrichment Prepared On Vendor Site) (9 of 10)



KAISER PERMANENTE

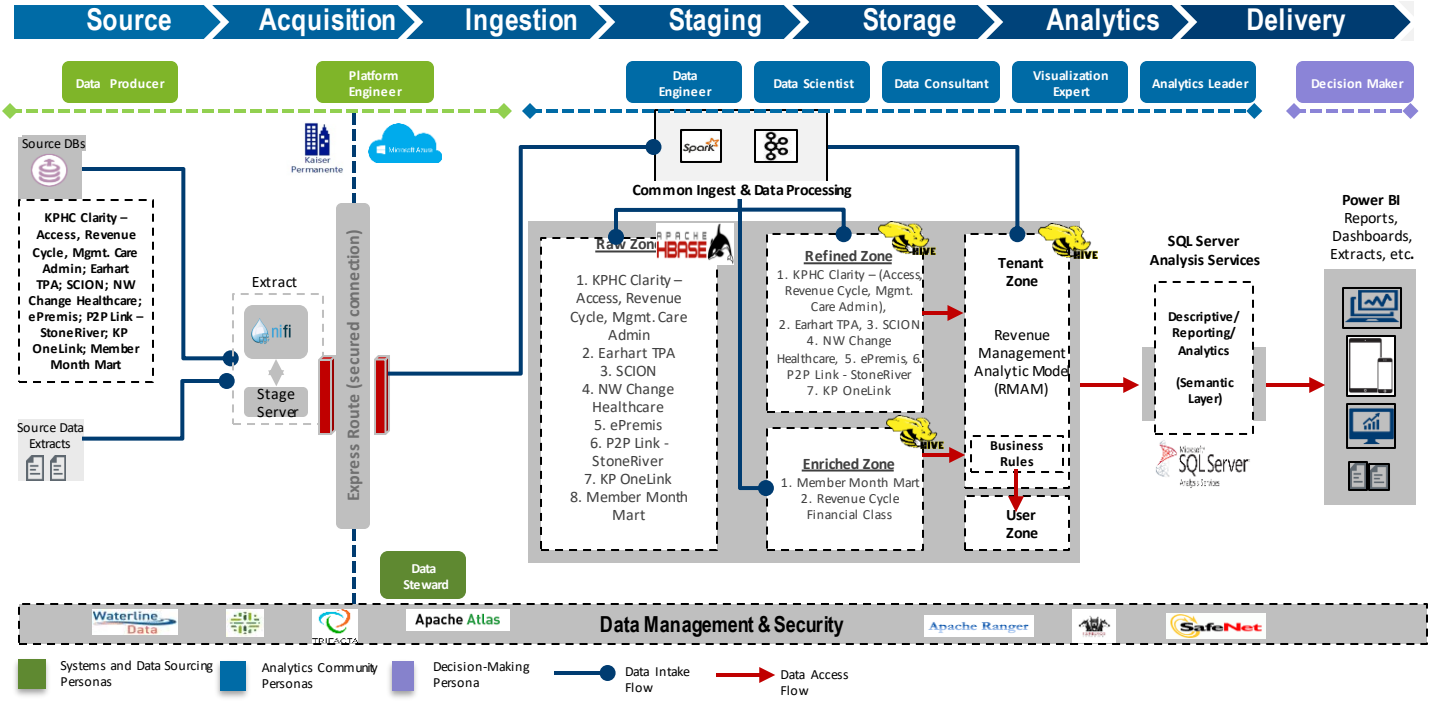
Solution Pattern (Hybrid) (10 of 10)

- Solutions may combine the use of several solution patterns, e.g., enriched data from a vendor that is further transformed on the ADF.



KAISER PERMANENTE®

Application of ADF Ref. Architecture – Revenue Management Analytics



KAISER PERMANENTE®

Foundation Services – Disaster Recovery and High-Availability

ADF Disaster Recovery Class of Service (CoS) Definitions

Recovery Designation	DR CoS	RTO	RPO	Disaster Recovery Exercise Requirements *			
				Operations (OPS) Based	Discussion Based **	DR Plan Review	Tape Restore***
Continuous Availability	0	≤1 Hr		Semi-Annually One-Operations Based DR Exercise must be performed annually			N/A
Advanced Recovery	1	≤4 Hr	≤2 Hr	Annually	Optional		
	2	≤24 Hr	≤8 Hr	Bi-annually	Annually		
	3	≤72 Hr	≤48 Hr				
Standard Recovery	4	≤1 Wk	≤72 Hr	Every 4 Years	Bi-annually	Annually	Bi-annually
	5	≤2 Wk	Last Offsite Back-up				
	6	≤1 Mo		Optional			
Custom Recovery	7	Best Effort					
	8	No Recovery					

ADF CoS

Class of Service (CoS) represents the business requirement for recovery in the event of a catastrophic incident requiring the recovery of an application at an alternate location.

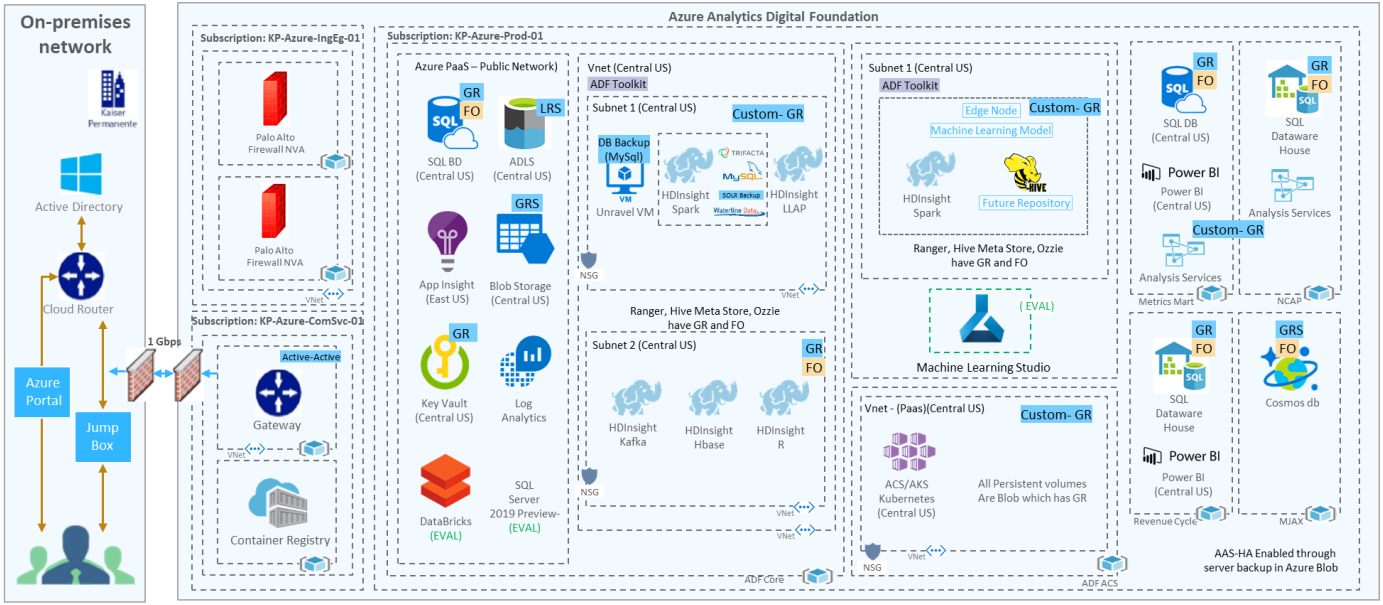
Recovery Point Objective (RPO) is the point in time to which data will be restored after recovery of an IT Service. This may involve loss of data.

Recovery Time Objective (RTO) is the maximum time allowed for recovery of an IT Service following an interruption.

ADF: Azure HA/DR Configuration Options

Type	LRS	ZRS	GRS	RA-GRS	Geo-Replication	Geo Redundancy	Custom	Azure Site Recovery
Descriptions	Locally redundant storage (LRS) will take three Replicas within single Data Center • Copies : 3 • Availability SLA : 99.9% Read/Write • Custom Solution to Copy to different region's storage accounts	Zone-redundant storage (ZRS) replicates data synchronously. • Three Replicas may span across two or three Datacenters on one or two regions • Copies : 3 • Availability SLA : 99.9% Read/Write	Geo-redundant storage (GRS) , replicates data synchronously across two regions • Copies : 6 • Availability SLA : 99.9% Read , 99.9% Write • If Primary region is not unrecoverable ,MSFT will automatically trigger Geo Failover	Read Access - Geo Redundant Storage (RA-GRS) – enables higher availability by reading the storage account's data in the secondary region • Copies : 6 • Availability SLA : 99.9% Read , 99.9% Write • If Primary region is not unrecoverable ,MSFT will automatically trigger Geo Failover	Geo-replication feature allows to create readable replicas of database in Data Centers within same or across regions • Copies : User Defined • Replication: Asynchronous • Any restore point within 35 days • Auto Failover configuration • Standard /Active Geo-Replication • ERT < 30s • RPO < 5s	• Enable global distribution of account by pairing two Azure Regions • Data will be replicated across all regions consistently • Availability SLA: 99.999% read and Write for multiregional DBs • Provides Automatic failover during regional outage	• Custom Built solution for the Azure Components in which native replication/failover features are not available • All Metastore will be Geo replicated • HDInsight & ACS Clusters will be created in West US and Central US on Demand basis • Necessary components will be redeployed	• The Azure Site Recovery service contributes to managing and orchestrating replication, failover, and failback of on-premises machines, and Azure virtual machines (VMs)
Storage	• ADLS Gen 1 • ADLS Gen 2 • Blobs Storage	• ADLS Gen 2 • Azure Blobs Storage	• ADLS Gen 2 • Azure Blobs Storage	• ADLS Gen 2 • Azure blob Storage	NA	NA	NA	NA
Database	NA	NA	NA	NA	• SQL Database • SQL DW	• Cosmos DB	NA	NA
Security	NA	NA	NA	NA	• Key Vault	NA	NA	NA
Compute-PaaS	NA	NA	NA	NA	NA	NA	• HDInsight Clusters • ACS Clusters • Analysis Services	NA
Compute - IaaS	NA	NA	NA	NA	NA	NA	NA	• Azure VMs

ADF – HA and DR Configuration



LEGEND: **LRS** – Local Redundant Storage | **GRS** – Geo Redundant Storage | **RA-GRS** – Read Access Geo Redundant Storage | **FO** – Fail Over | **GR** –Geo Redundant

KAISER PERMANENTE®

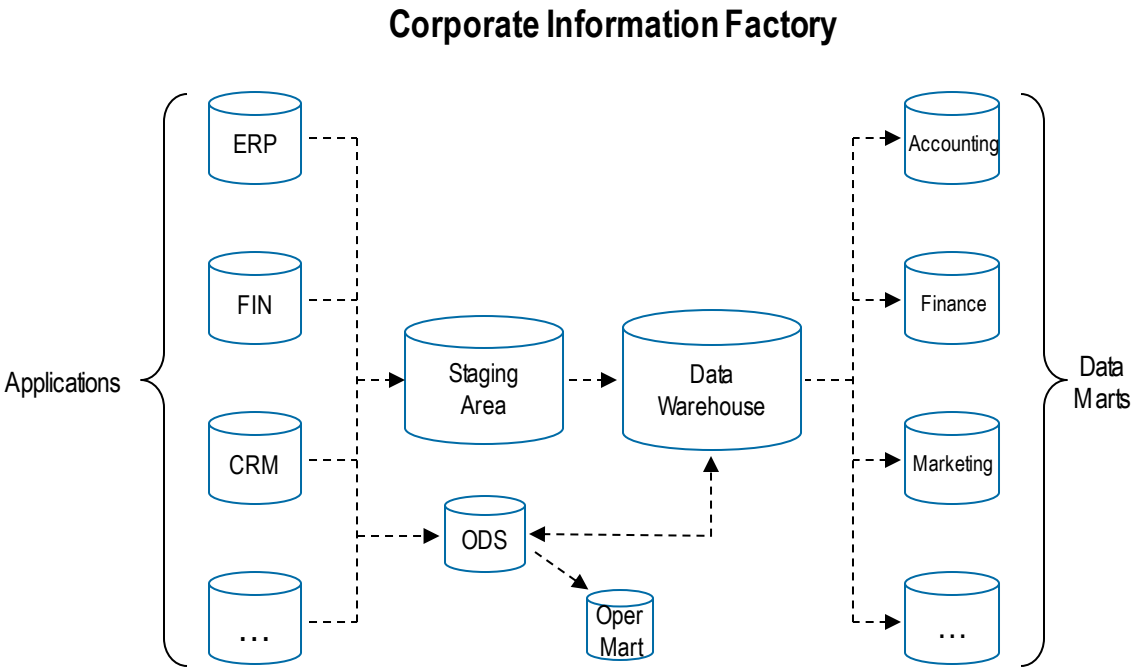
Foundation Data

Key Objectives

- Describe components of Foundation Data blueprint and guiding principles that are driving design
- Define the ADF data zones and Foundation Data Management (FDM) disciplines and describe how they support analytic needs
- Share Data Domain taxonomy and how it will enable and help to manage analytic solutions
- Provide an overview of key features – Change Data Capture, Clarity * N, Data Validation – applied to Revenue Management as a Reference Implementation
- Understand where to locate data ingest status

Foundation Data - Data Zone Architecture

90's Architecture for Decision Support Systems

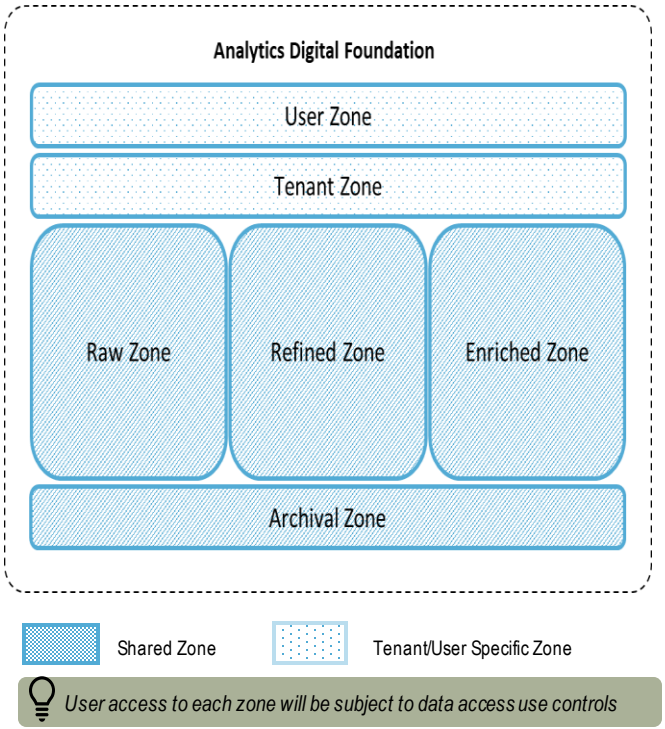


A need to service different user classes requires different types of data repositories:

- Data warehouses
- Data marts
- Operational data stores
- Oper marts

KAISER PERMANENTE®

Data Zone Architecture



Zones within the Analytic Digital Foundation (ADF) replace data repositories and embody separation of utilities, support data pipelines, data lifecycle stages, data security, and other data requirements.

Data Zone Architecture Guiding Principles

Reusability

Data in the raw, refined, and enriched zones are collected once and used many times to support tenants.

Modularity

Separation of utilities embodied in the zoned design allows each zone to be optimized for its specialized purpose independent of other zones.

Reliability

Data management is a key focus to ensure the discoverability, quality, consistency, and trustworthiness of the data.

Agility

Ready access to cross-functional program wide data in one place speeds up analytic throughput.

Persona Enablement

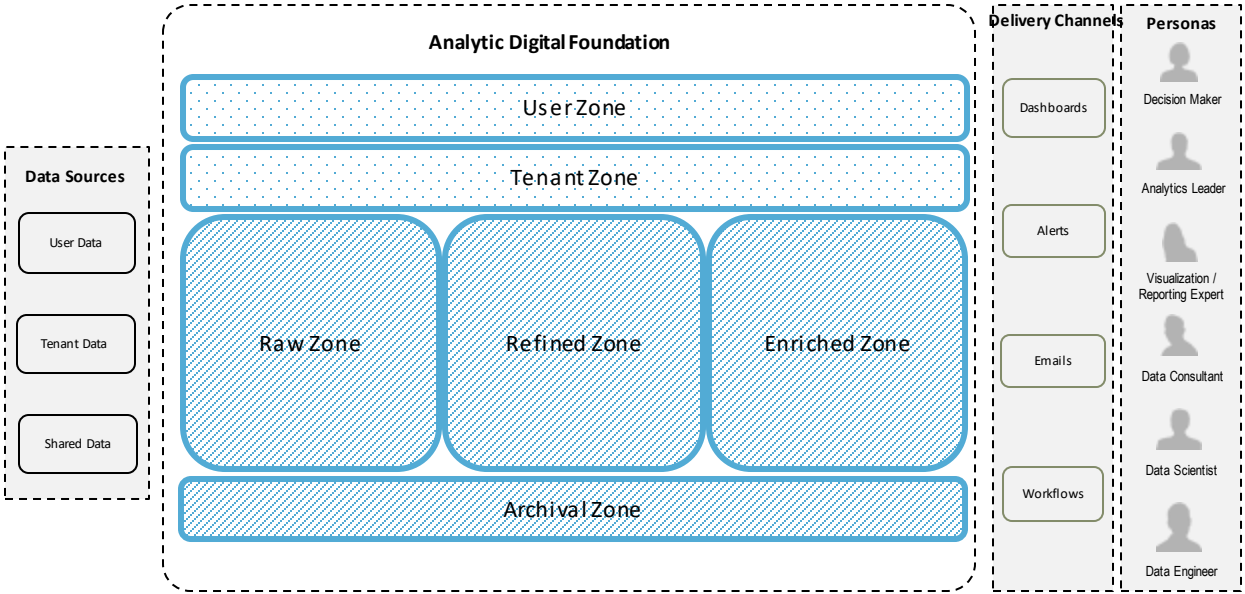
Zones storing different kinds of data (structured, unstructured) at various levels of curation cater to multiple personas and capabilities.

Security

Access and use of data are controlled by a common, shared, foundational service, and limit access to the minimum necessary to perform a user's role(s).

Analytic Digital Foundation: Data Zone Architecture

Data zones will be established to meet varied analytical needs of the tenants and enable reusability of data assets.



Shared data is brought into the Raw Zone and propagated to Refined and Enriched Zones
Tenants can bring in data to Tenant Zone to build **solutions** for their use cases
Individuals can bring in their own data into the User Zone for **prototyping/exploratory analysis**

Shared Zone

Tenant/User Specific Zone

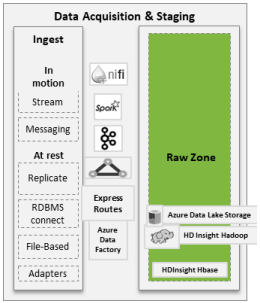
KAISER PERMANENTE®

Raw Zone

Definition and Purpose

The Raw Zone is a structure where **shared data** from data sources first land in the Analytic Digital Foundation for eventual processing into the Refined and Enriched Zones.

The Raw Zone provides a structure where data from enterprise-level data sources can be acquired and stored in its **original Raw form** once, and used many times to support the data needs of tenants.



Characteristics

Data Characteristics

Volume, variety, and velocity are the dominant characteristics of data stored in the Raw Zone. This zone focuses on **rapid ingestion** (batch, near real time) of large amounts of a variety of data (structured, semi-structured and unstructured data).

Data Organization

Data is organized by the **data source**, e.g., KPHC Clarity, OneLink.

Data Refresh

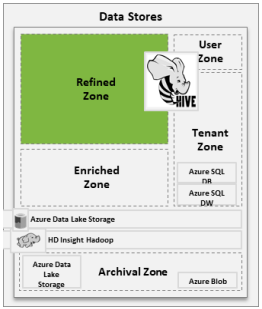
Supports varying latencies of data integration, from daily batches to continuous real-time integration, and are scalable to support more aggressive data freshness based on the needs of the maturing organization.

Refined Zone

Definition and Purpose

The Refined Zone is a structure designed to store and deliver **shared data** that has been **prepared** for the use of all tenants.

The Refined Zone provides a structure where data preparations needed by all tenants can be performed and **stored once, and used many times**.



Characteristics

Data Characteristics

Volume is the dominant characteristic of data in the Refined Zone as all historical records are maintained. Velocity varies based on business goals to support the multiple data freshness service levels.

Data Organization

Data is organized by **information data domains** which logically group data based on functional categories of the business, e.g., claims, membership, care delivery.

Data Refresh

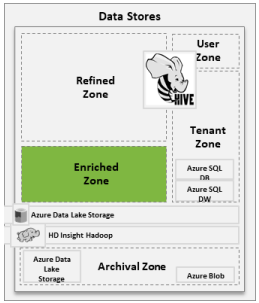
Supports different service levels for data freshness driven by business goals. Time stamping and view constructs are used to support the multiple data freshness service levels.

Enriched Zone

Definition and Purpose

The Enriched Zone is a structure designed to store and deliver newly derived **shared data**, generated by **manipulating** and blending data from the Raw and Refined Zones, for the use of all tenants.

The Enriched Zone provides a structure where **common enrichments** can be developed and stored once, and used many times. Sharing common enrichments improves the consistency of information produced **across the tenants**.



Characteristics

Data Characteristics

Veracity is the dominant characteristic of data in the Enriched Zone as it is needed to store **high quality data** for strategic decision making.

Data Organization

Data is organized by the **category of enrichment**, e.g., costed utilization (includes care delivery, health plan, medical group, IBNR, etc.), member month (includes product line, line of business, service area primary, member lookup, etc.).

Data Refresh

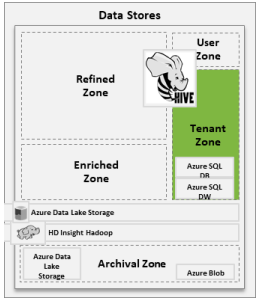
Supports mostly **long-term decision making** where the stability of the data needs to match the grain of analysis. Calendar rollups (day, week, month, quarter, and year) are used instead of servicing each grain with its own set of tables.

Tenant Zone

Definition and Purpose

The Tenant Zone is a structure designed to store and deliver data **produced by tenants** for their users to consume.

The Tenant Zone provides a structure that is optimized for the **delivery of analytics solutions** while the Raw and Refined Zones are optimized for the acquisition and preparation of data for reuse (separation of concerns).



Characteristics

Data Characteristics

Generally veracity is the dominant characteristic of data in the Tenant Zone as different types of data that is required for **use cases of specific tenant**. However, volume, variety, and velocity are dependent on Tenant use cases. In order to make data accessible to users, data is predominantly stored using pre-defined format / schemas (schema-on-write approach). However, tenant specific data from external sources can also be stored using schema-on-read approach.

Data Organization

Data is organized by the **tenant domain**, e.g., profitability reporting and cost, revenue management analytics, sales and marketing analytics.

Data Refresh

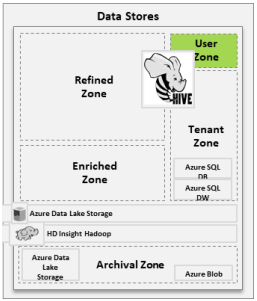
Data freshness is driven by the tenant goals.

User Zone

Definition and Purpose

The User Zone is a structure where **advanced users** (e.g., data scientists) can introduce their own data needed to perform **exploratory** data science and analysis.

Successful data science efforts can be operationalized to use the Raw, Refined, and Enriched Zones for broader consumption by tenants.



Characteristics

Data Characteristics

Variety is the dominant characteristic of the User Zone as it allows qualified users to bring in different types of data, either from Refined, Enriched, Tenant, or user specific data, for the purpose of exploratory analysis.

Data Organization

Data organization is **highly personalized** and based on an individual's use case requirements.

Data Refresh

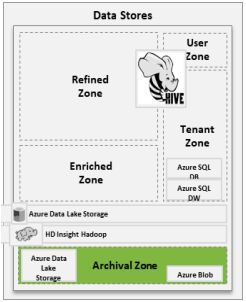
Data freshness is driven by the user goals.

Archival Zone

Definition and Purpose

The Archival zone is a structure designed to store Tenant required historical and **inactive data** from the Refined, Enriched, and Tenant Zones for long-term Archival.

The Archival Zone provides a structure for cost-effective long-term storage of data that is no longer actively needed for everyday operations but remains important for future reference or **regulatory compliance.**



Characteristics

Data Characteristics

Volume is the dominant characteristic of data in the Archival Zone as it stores structured data from Refined, Enriched, and Tenant Zone for an extended period of time for the purpose of audit and compliance. **Retired data assets** that are part of the initial migration to the platform are also archived, if needed.

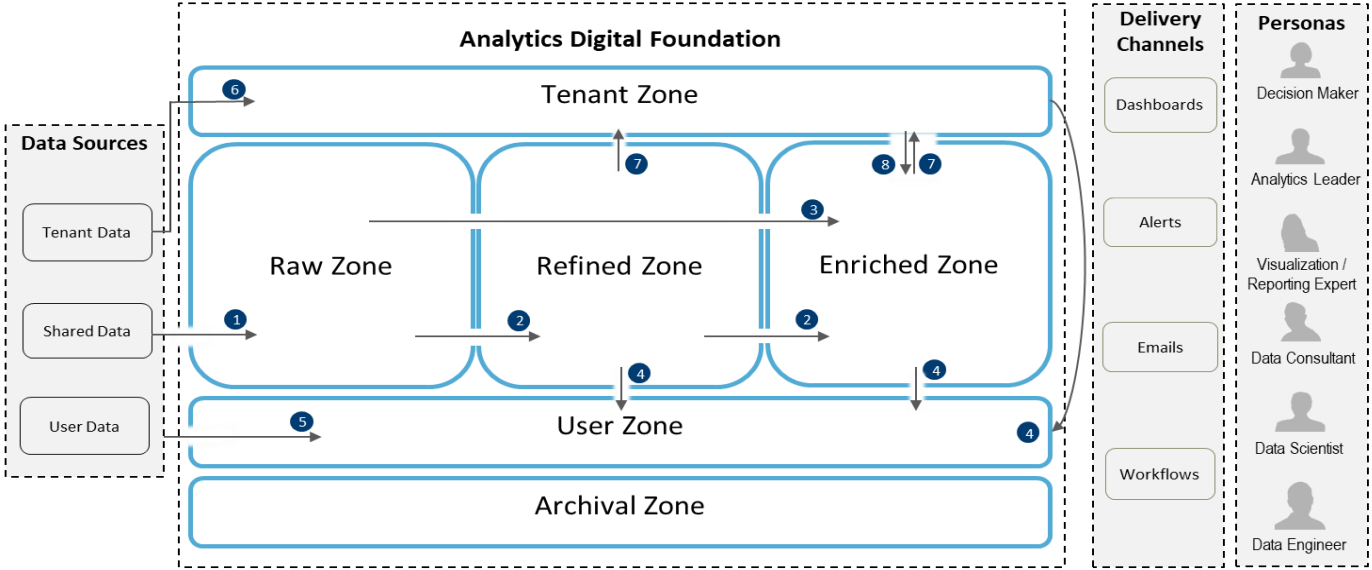
Data Organization

Data is organized based on the **origination data zone**, e.g., Refined Zone, Enriched Zone, Tenant Zone.

Data Refresh

Data will be loaded into the Archival Zone based on specific **data retention polices** established by the tenants and the foundation data teams.

Analytic Digital Foundation: Data Flow Patterns



1. Shared data coming into the ADF platform will be persisted in the Raw zone to provide complete traceability for audit and compliance.
2. As data moves from Raw zone to Refined to Enriched it will be transformed. Unstructured data will only reside in Raw Zone; information derived from unstructured data will be stored in Refined/ Enriched Zones as structured data.
3. Data can also move from Raw to Enriched Zone (e.g. internal or external application providing enriched data sets like ETGs to ADF).
4. Users can access and move data, if transformations are required, from Refined, Enriched and Tenant Zone into the User Zone for exploratory analysis and rapid prototyping.
5. Users can bring in their own data into the User Zone to build prototypes.
6. Tenants can also bring in internal or external data sources into the Tenant Zone.
7. Tenants will have access to data from Refined and Enriched Zones.
8. Tenant enrichments that need to be shared across multiple tenants will need to be brought into the Enriched Zone.

Foundation Data - Data Management

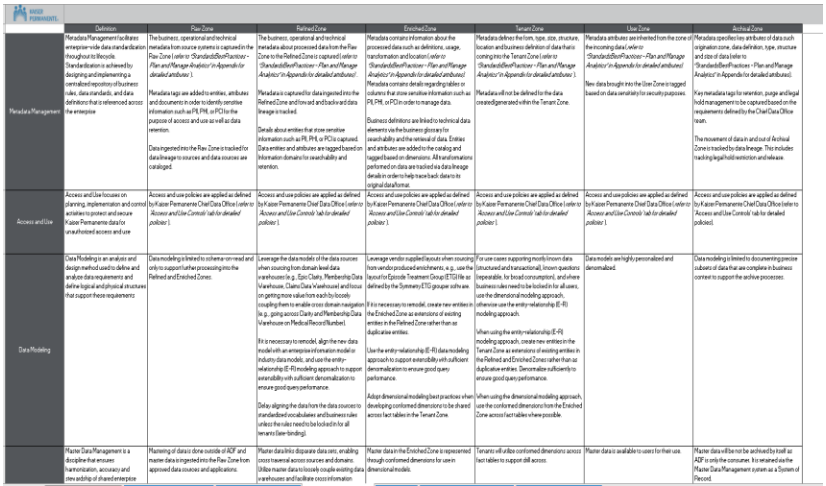
Data Management Disciplines

Data Management comprises all disciplines related to managing data as a valuable resource.

Data Governance	Data Governance focuses on establishing organizational constituencies and a framework of policies, processes, and enabling technologies to ensure that enterprise data is owned and stewarded accurately and consistently to meet business goals
Metadata Management	Metadata Management facilitates enterprise-wide data standardization throughout its lifecycle. Standardization is achieved by designing and implementing a centralized repository of business rules, data standards, and data definitions that is referenced across the enterprise
Access and Use	Access and Use focuses on planning, implementation and control activities to protect and secure Kaiser Permanente data for unauthorized access and use
Data Modeling	Data Modeling is an analysis and design method used to define and analyze data requirements and define logical and physical structures that support these requirements
Master Data Management	Master Data Management is a discipline that ensures harmonization, accuracy and stewardship of shared enterprise data assets to create a single version of truth
Reference Data Management	Reference Data Management is a discipline for recognizing, standardizing and sharing relatively static data and their permissible values across the enterprise
Data Quality	Data Quality is the capability to provide reliable data that satisfies the business functions and technical requirements of the ADF platform. Data quality is typically measured in terms of accuracy, consistency, completeness, accessibility, timeliness, and validity
Data Lifecycle and Archive Management	Data lifecycle management is an approach to data and storage management that recognizes that the value of information changes over time and that it must be managed accordingly. It also recognizes that in a heavily regulated environment, data through its lifecycle must be managed in accordance with Data Retention Policies

Data Management Blueprint

Analytic Digital Foundation Data Management practices have been developed in partnership with the **Chief Data Office**. Guardrails have been defined for each **Data Discipline** by Data Zone and align with the enterprise vision.



	Definition	Raw Zone	Refined Zone	Enriched Zone	Tenant Zone	User Zone	Archival Zone
Metadata Management	Metadata Management facilitates enterprise-wide data standardization throughout its lifecycle. Standardization is achieved by designing and implementing a centralized repository of business rules, data standards, and data definitions that is referenced across the enterprise	The business, operational and technical metadata from source systems is captured in the Raw Zone (refer to 'Standards/BestPractices' - 'Plan and Manage Analytics' in Appendix for detailed attributes). Metadata tags are added to entities, attributes and documents in order to identify sensitive information such as PII, PHI, or PCI to the purpose of access and use as well as data retention. Data ingested into the Raw Zone is tracked for data lineage to sources and data sources are cataloged.	The business, operational and technical metadata about processed data from the Raw Zone to the Refined Zone is captured (refer to 'Standards/BestPractices' - 'Plan and Manage Analytics' in Appendix for detailed attributes)'. Metadata is captured for data ingested into the Refined Zone and for lead and backward data lineage is tracked. Details about entities that store sensitive information such as PII, PHI, or PCI is captured. Data entities and attributes are tagged based on information domains for searchability and retention.	Metadata contains information about the processed data such as definitions, usage, transformation and location (refer to 'Standards/BestPractices' - 'Plan and Manage Analytics' in Appendix for detailed attributes)'. Metadata contains details regarding tables or columns that store sensitive information such as PII, PHI, or PCI in order to manage data. Business definitions are linked to technical data elements via the business glossary for searchability and the retrieval of data. Entries and attributes are added to the catalog and tagged based on dimensions. All transformations performed on data are tracked via data lineage details in order to help trace back data to its original data format.	Metadata defines the form, type, size, structure, location and business definition of data that is coming into the Tenant Zone (refer to 'Standards/BestPractices' - 'Plan and Manage Analytics' in Appendix for detailed attributes). Metadata will not be defined for the data created/generated within the Tenant Zone.	Metadata attributes are inherited from the zone of the incoming data (refer to 'Standards/BestPractices' - 'Plan and Manage Analytics' in Appendix for detailed attributes) New data brought into the User Zone is tagged based on data sensitivity for security purposes.	Metadata specifies key attributes of data such as origination zone, data definition, type, structure and size of data (refer to 'Standards/BestPractices' - 'Plan and Manage Analytics' in Appendix for detailed attributes). Key metadata tags for retention, purge and legal hold management to be captured based on the requirements defined by the Chief Data Office team. The movement of data in and out of Archival Zone is tracked by data lineage. This includes tracking legal hold restriction and release.
Access and Use	Access and Use focuses on planning, implementation and control activities to protect and secure Kaiser Permanente data for unauthorized access and use	Access and use policies are applied as defined by Kaiser Permanente Chief Data Office (refer to 'Access and Use Controls' tab for detailed policies).	Access and use policies are applied as defined by Kaiser Permanente Chief Data Office (refer to 'Access and Use Controls' tab for detailed policies).	Access and use policies are applied as defined by Kaiser Permanente Chief Data Office (refer to 'Access and Use Controls' tab for detailed policies).	Access and use policies are applied as defined by Kaiser Permanente Chief Data Office (refer to 'Access and Use Controls' tab for detailed policies).	Access and use policies are applied as defined by Kaiser Permanente Chief Data Office (refer to 'Access and Use Controls' tab for detailed policies).	Access and use policies are applied as defined by Kaiser Permanente Chief Data Office (refer to 'Access and Use Controls' tab for detailed policies).
Data Modeling	Data Modeling is an analysis and design method used to define and analyze data requirements and define logical and physical structures that support these requirements	Data modeling is limited to schema-on-read and only to support further processing into the Refined and Enriched Zones.	Leverage the data models of the data sources when sourcing from domain level data warehouses (e.g., Epic Clarity, Membership Data Warehouse, Claims Data Warehouse) and focus on getting more value from each by loosely coupling them to enable cross domain navigation (e.g., going across Clarity and Membership Data Warehouse on Medical Record Number). If it is necessary to remodel, align the new data model with an enterprise information model or industry data models, and use the entity-relationship (E-R) modeling approach to support extensibility with sufficient denormalization to ensure good query performance. Delay aligning the data from the data sources to standardized vocabularies and business rules unless the rules need to be locked in for all tenants (late-binding).	Leverage vendor supplied layouts when sourcing from vendor produced enrichments; e.g., use the layout for Episode Treatment Group (ETG) file as defined by the Symmetry ETG groupe software. If it is necessary to remodel, create new entities in the Enriched Zone as extensions of existing entities in the Refined Zone rather than as duplicative entities. Use the entity-relationship (E-R) data modeling approach to support extensibility with sufficient denormalization to ensure good query performance. Adopt dimensional modeling best practices when developing conformed dimensions to be shared across fact tables in the Tenant Zone.	For use cases supporting month/known data (structured and transactional), known questions (repeatable, for broad consumption), and where business rules need to be locked in for all users, use the dimensional modeling approach, otherwise use the entity-relationship (E-R) modeling approach. When using the entity-relationship (E-R) modeling approach, create new entities in the Tenant Zone as extensions of existing entities in the Refined and Enriched Zones rather than as duplicative entities. Denormalize sufficiently to ensure good query performance. When using the dimensional modeling approach, use the conformed dimensions from the Enriched Zone across fact tables where possible.	Data models are highly personalized and denormalized.	Data modeling is limited to documenting precise subsets of data that are complete in business context to support the archive processes.
	Master Data Management is a discipline that ensures harmonization, accuracy and stewardship of shared enterprise	Mastering of data is done outside of ADF and master data is ingested into the Raw Zone from approved data sources and applications.	Master data links disparate data sets, enabling cross traversal across sources and domains. Unlike master data it loosely couple existing data warehouses and facilitate cross information	Master data in the Enriched Zone is represented through conformed dimensions for use in dimensional models.	Tenants will utilize conformed dimensions across fact tables to support drill across.	Master data is available to users for their use.	Master data will be not be archived by itself as ADF is only the consumer. It is retained as the Master Data Management system as a System of Record.

A detail of each principle by data zone can be found A20 Program Standards and Practices Confluence site.

KAISER PERMANENTE®

Foundation Data - Data Domain Taxonomy

Data Domain Taxonomy – Definition, Principles and Business Value

A domain is a group of things, organized through a distinct lens, to serve a specific purpose. Data domains are groupings of data for the purpose of classifying data from a business perspective.

Guiding Principles

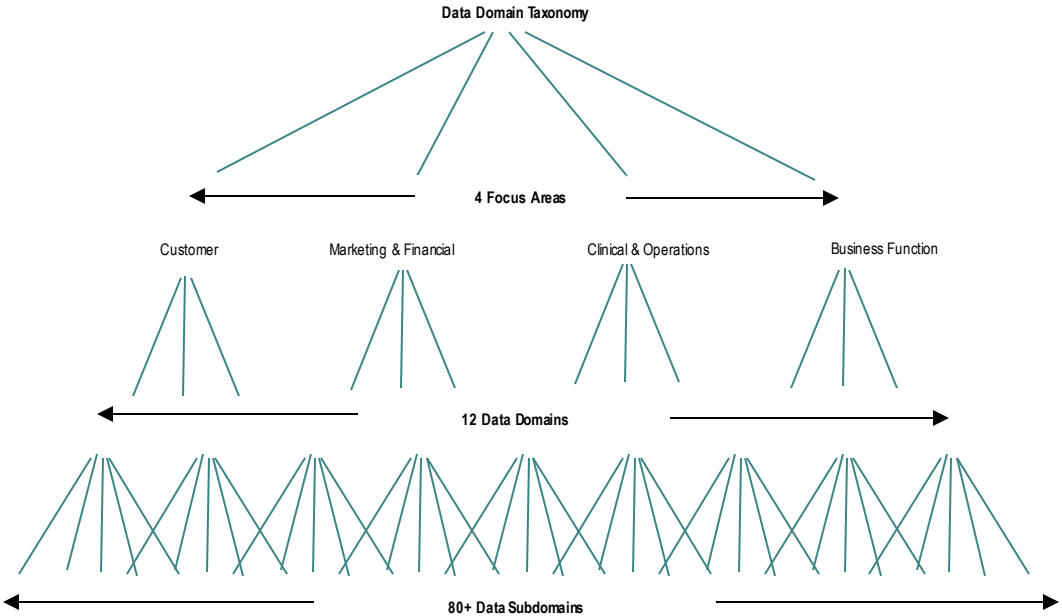
- Data domains are used to group related data aligned for business uses and processes
- Data domains support requirements and use cases for data governance and replace the legacy data governance domains
- Data domains support the entire enterprise and, as such, are managed by the Chief Data Office in KP-IT Corporate Services
- Data domains are not intended to convey "ownership" of data but enable data stewardship identification
- Data domains span the enterprise and are not region specific

Key Business Value in the adoption of Data Domains

- Data domains describe and manage KP's data to enable maturation of data as an asset at KP
- Data domains enable KP to identify and organize roles (such as data stewards) for data management and governance purposes
- Data domains define scope for appropriate / inappropriate use
- Data domains enable data to be organized and located effectively

Data Domain Taxonomy

The Data Domain Taxonomy is centered around four business focus areas, and broken out into 12 data domains and 80+ data subdomains.

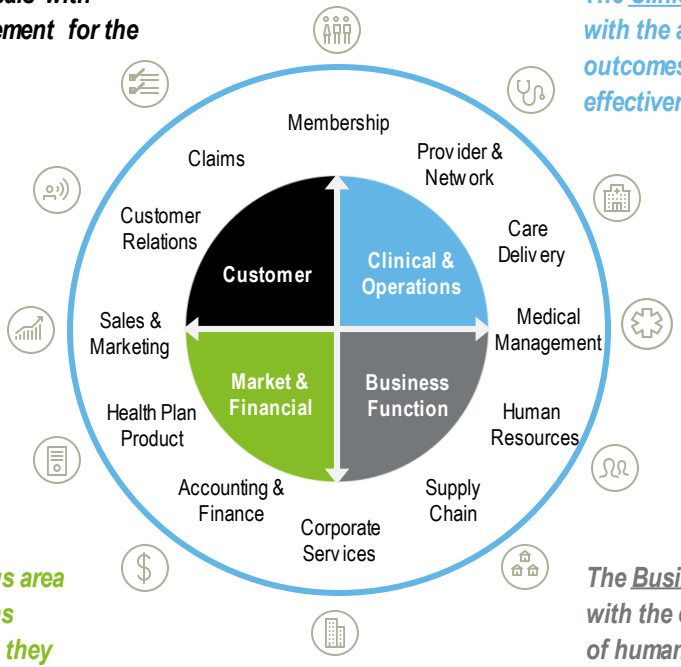


The following document showcases the organization and definitions of focus areas, data domains, and data subdomains.

KAISER PERMANENTE®

Data Domain Taxonomy

The Customer focus area deals with services and claims management for the communities they serve.



The Clinical & Operations focus area deals with the acuity, safety, satisfaction and outcomes necessary to improve the effectiveness of care.

The Market & Financial focus area deals with how organizations understand the populations they serve, the products they provide and the associated revenue and costs.

The Business Function focus area deals with the organization's efficiency in terms of human resources, supply chain and other resources used in the delivery of care.

KAISER PERMANENTE®

Data Domains and DRAFT Subdomains

The following data subdomains are a comprehensive and extensible, but not exhaustive, grouping of data managed actively at KP.

Membership

- Customer Billing
- Eligibility
- Employer Groups
- Enrollments
- Fulfillment
- Members

Claims

- Submissions
- Payment and Outputs
- Claims Correspondence

Customer Relations

- Appeals
- Call Center Data
- Cases
- Grievances
- Surveys

Sales and Marketing

- Agents / Brokers
- Campaigns
- Leads
- Market Research
- Offers
- Prospects
- Quotes
- Responses
- Segments

Health Plan Products

- Benefit Management
- Contract Management
- Fee Schedules
- Plan Accreditation
- Pricing Factors
- Product Design
- Rate Models

Accounting and Finance

- Budgets
- Expenditures
- Financial Accounts
- Forecasts
- Invoices
- Payables
- Receivables
- Revenues
- Patient Billing (Revenue Cycle)

Provider and Network

- Affiliations
- Contracts
- Credentials
- Networks
- Organizations
- Providers
- Provider Services
- Reimbursements

Care Delivery

- Care Plans
- Care Teams
- Consents
- Diagnoses
- Encounters
- Health Assessments
- Images / Radiology
- Labs
- Medical Procedures
- Patients
- Patient Movements (ADTs)
- Referrals
- Schedules
- Pharmacy Services

Medical Management

- Authorizations
- Benchmarks
- Care / Disease Management
- Care Pathways
- Medical Policies
- Medical Records
- Programs
- Quality Management
- Social Determinants
- Utilization Management
- Research

Human Resources

- Benefits &Compensation
- Learning & Development
- Employees
- Performance
- Payroll
- Time & Labor
- Recruitment

Supply Chain

- Distribution
- Inventory
- Orders
- Procurement
- Shipping / Routing
- Suppliers
- Vendors

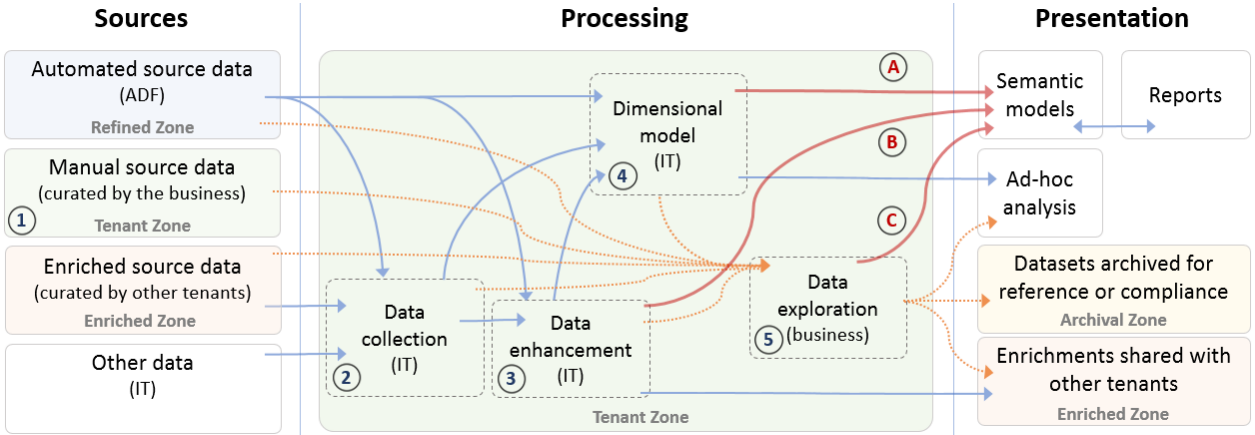
Corporate Services

- Corporate Communications
- Community Benefits
- Locations & Facilities
- Labor Relations
- Legal
- Regulatory / Compliance
- Information Technology Assets

Foundation Data - Reference

Reference Implementation Overview

A full stack deployment that demonstrates multiple capabilities of the ADF to deliver information and enrichments through a fully developed tenancy.



Sources

- **Clarity*N, fully validated and complete**
- Non-programmatic data sources (goals, for example)
- Tenant specific data from programmatic sources

Five **Tenant Schemas**: (1) Curation, (2) Collection, (3) Enhancement, (4) Dimensional Model, (5) Exploration

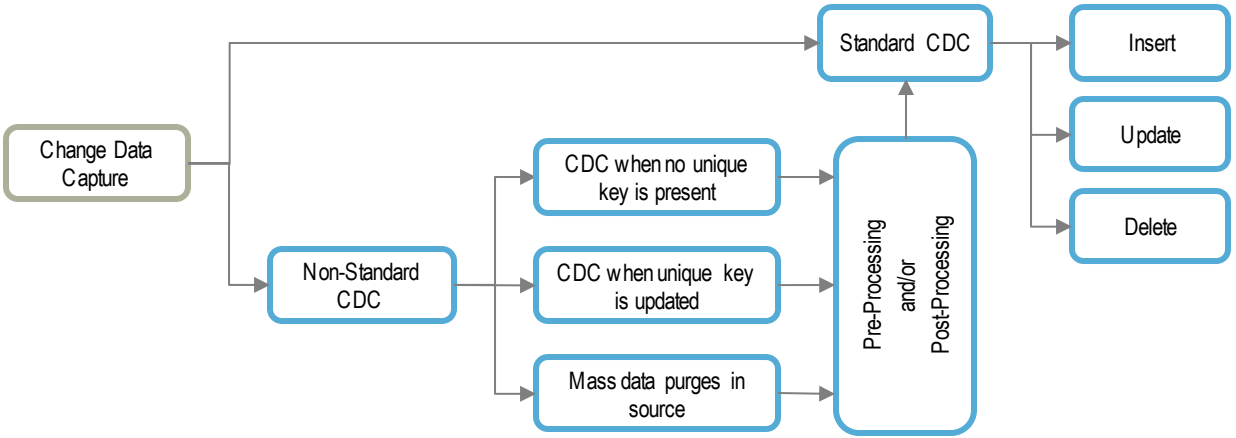
Tenancy

- Data collection – Staging schema for data not sourced from the refined zone, or requiring pre-processing for use
- Data enhancement – IT supported schema containing complex custom routines that produce 'source of truth' reference data or metrics for consumption by the tenant or publication to the Enriched Zone
- Dimensional model – Star schema underpins the OLAP semantic model for self serve reporting and analysis
- Data exploration – Tenant analysts develop ad-hocs and perform research for business partners

KAISER PERMANENTE®

Change Data Capture – Feature Overview

Change Data Capture (CDC) is the process of detecting changes in source data and ensuring that the target tables are appropriately updated, added to, or modified to preserve an accurate reflection of that source.



Standard Change Data Capture

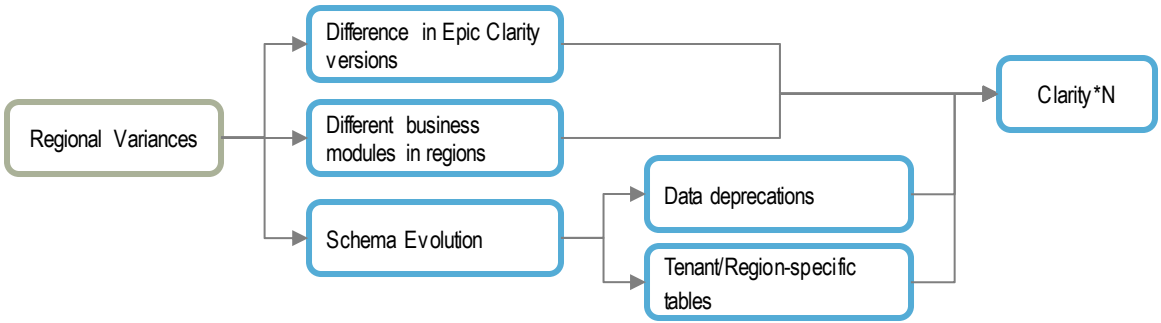
- Identify net new data from source and insert in ADF
- Detect updates from source and apply to existing ADF data
- Identify deletes in source and logically delete in ADF

Non-Standard Change Data Capture

- Handle changes to keys in source and implicit deletes
- Ingest entities with no unique keys
- Detect mass data purges/updates that occur in the source

Clarity*N – Feature Overview

Clarity*N is a single view of Clarity that sources data from regional Clarity instances and addresses version skew. This combined view of regional data can be consumed by regional and program level tenants for reporting and analytics.



Schema Design

- Option1: Physical Schema
- Option 2: Logical Schema

Data Model Design

- Option 1: Standard Epic Clarity Data Model Definition with Data Type Transformations
- Option 2: Superset of Regional Clarity Structures

Address Regional Variances

- Difference in Tables Across Regions
- Difference in Columns Across Regions
- Difference in Data Types Across Regions

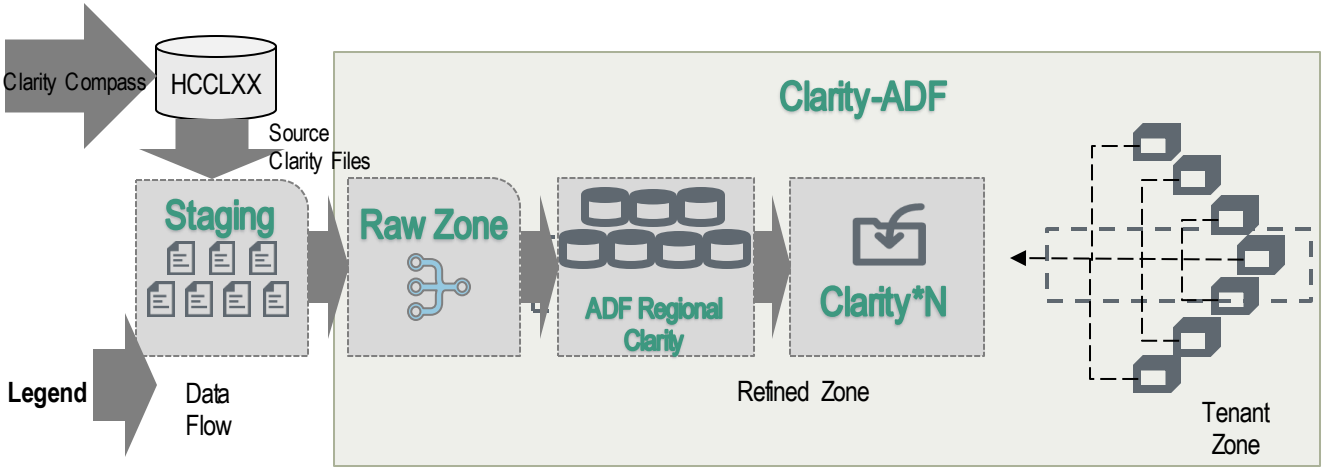
Address Schema Evolution

- Schema Evolution for Physical or Logical Schema of Clarity*N
- Impact of schema evolution on Clarity*N

KAISER PERMANENTE®

Clarity*N – Data Flow Design

Clarity*N receives data from source HCCLXX schemas. This data flows into ADF via the Raw Zone before being loaded into the Refined Zone of ADF. Data from regional instances is present in ADF Regional Clarity in their native form. Clarity*N is created by combining the regional data from ADF Regional Clarity.



Foundation manages the technical variation; Tenants manage content harmonization.

KAISER PERMANENTE®

Clarity*N – Technical Variation Metrics

Technical variations exist across regional instances of Clarity due to differences in Epic versions, different workflow needs, and regional customizations to meet region specific or program needs.

Clarity *N – Regional Clarity Variation Analysis metrics

S.No	Analysis Detail		Affected Table Count	Total Table Count	Percentage
1	Clarity_Tables_Not_Released_In_All_7_Region		1378	21849	6.306924802
			Affected Column Count	**Total Column Count**	**Percentage**
2	Clarity_Columns_Not_Released_In_All_7_Region		892	195511	0.456240314
3	Clarity columns with different presicion_Scale		920	195511	0.470561759
4	Clarity columns with different presicion_Scale_And_Defined_As_PK		119	920	12.93478261
5	Clarity columns with different datatype	Compatible Datatype	106	195511	0.054216898
		InCompatible Datatype	71	195511	0.036315092
6	Clarity columns with different datatype_And_Defined_As_PK	Compatible Datatype	21	106	19.81132075
		InCompatible Datatype	10	71	14.08450704
			Affected Table Count	**Total Table Count**	**Percentage**
7	Tables do not have a primary key defined		3	21849	0.013730606

Clarity*N – Solution Design

Recommendation

Logical Schema (View)

- A logical schema with views that collate and transform the data from the ADF Regional Clarity instances in ADF.
- Views are defined to logically combine and transform the data from ADF Regional Clarity instances. There is no redundancy of data in Clarity*N.

Standard Epic Clarity Data Model

- Adopt standard Epic Clarity data definitions from the latest Clarity implemented across all regional source systems.
  - All standard Epic entities will be present in Clarity*N.
  - Tenant or region-specific entities (e.g. X_* tables) will not be in Clarity*N.

Advantages

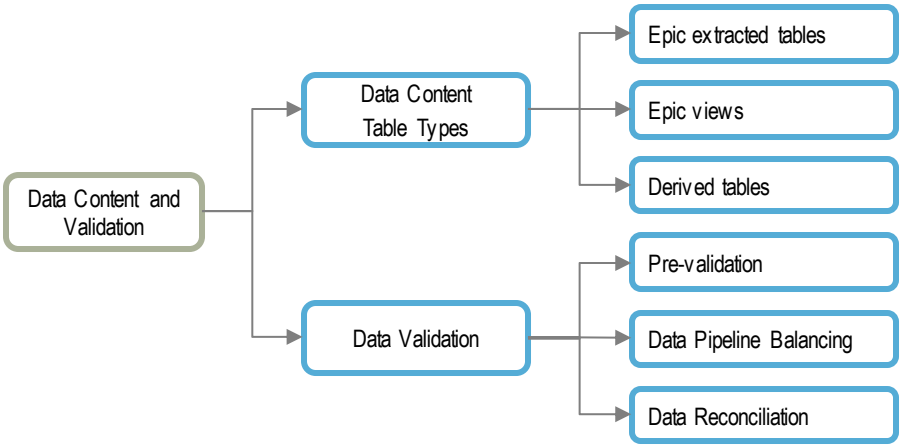
- No data replication and redundancy between ADF Regional Clarity and Clarity*N.
- No data model management is required to manage table and attribute definitions, datatypes, constraints, and referential integrity.
- When source schema changes, only the views need to be maintained. There is one less hop to be maintained in the ADF pipeline.
- High availability of data - as soon as data is available in ADF Regional Clarity, the views will reflect the latest data automatically. There is no ETL required.

- Accelerated adoption of Clarity*N data model due availability of the standard Epic data definition.
- Maintenance of Clarity*N is simple as standard Epic data definitions are adopted.
- When source schema changes, Clarity*N will only address data compatibility^ to ensure there is no data truncation.
- Low latency when source schema changes, as only data type transformations are addressed.
- If the user requires data from multiple regions, then the standard epic schema is easier to consume.

Future Decision: Upon completion of the Proof of Concept to confirm feasibility and operationalization of the Solution Design recommendation, results will be presented and approval will be requested of the Steering Committee for adoption of design.

Data Content and Validation – Feature Overview

Data Content addresses the types of Clarity data structures that must be stored in Clarity*N. Data Validation is the processes to perform balancing of data during data load as well as reconciliation post data load to ensure parity between the regional Clarity source data and Clarity*N data.



Data Content

Clarity*N must support the following data structures from HCCLXX schemas which is dependent on Clarity*N design options:

- Epic extracted tables
- Epic views
- Derived tables
- Tenant or region-specific tables

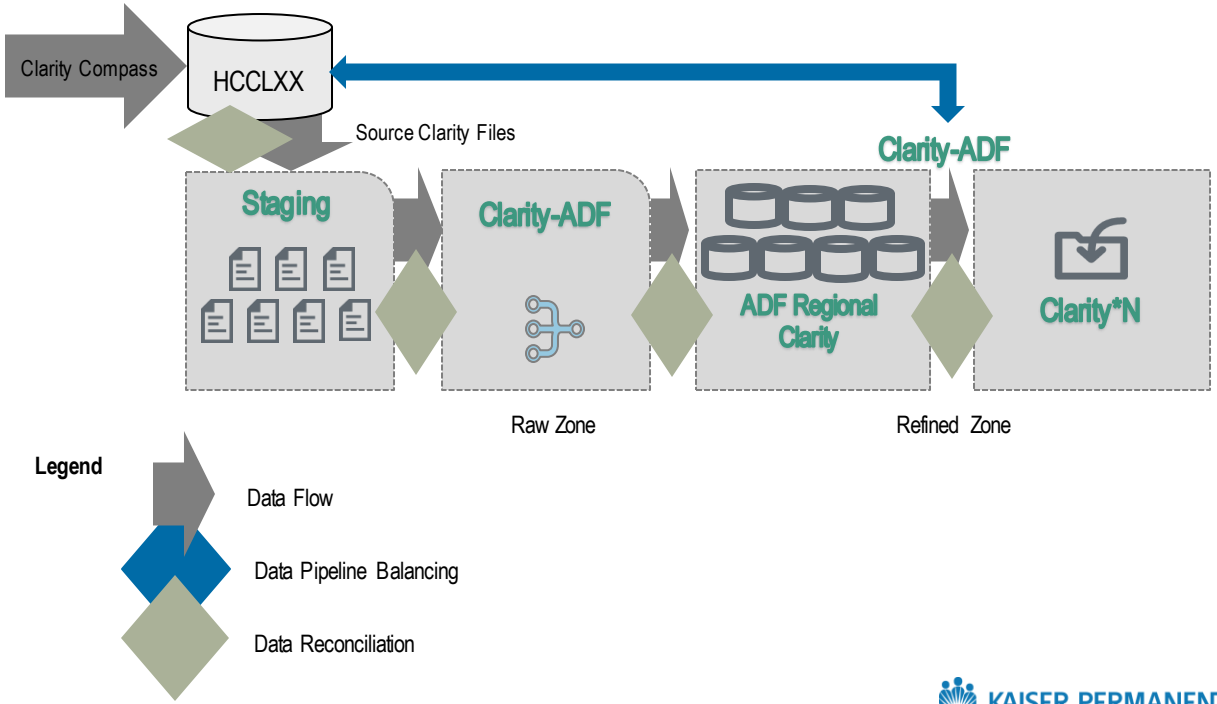
Data Validation

- Clarity*N must have validation processes to check for accuracy and completeness of data.
- Data validation occurs in three stages:
  - Pre-validation
  - Data Pipeline Balancing
  - Data Reconciliation

Data Validation – Approach Overview

Data Validation in Clarity-ADF is divided into the below 3 categories:

- Pre-validation of data when data is loaded from the source Clarity files into the staging layer of ADF.
- Data pipeline balancing that occurs across the Clarity-ADF pipeline.
- Data reconciliation to verify parity of data between source and Clarity-ADF.



KAISER PERMANENTE®

Data Validation – Proposed Approach

Purpose

The purpose of data validation is to ensure:

- Completeness of data load.
- Accuracy of data in Clarity-ADF
- Parity between source Clarity data and Clarity-ADF

Approach

Pre-validation

- Pre-validation of data occurs on the source Clarity files that are received from the regional source Clarity systems - HCCLXX schemas.
- In this stage, the following checks are performed:
  - Source file eligibility check
  - Total counts
  - Identification of full/delta loads
  - Threshold calculation and checks

Data Pipeline Balancing

- Error logging and balancing will be performed across the data pipeline from HCCLXX to Clarity*N.
- Errors or warning identified at every stage will be logged in a centralized repository.
- These logs will be made available via subscription-based reports.
- Remediation of exceptions and errors must be performed to account for all the records in the final parity check.

Data Reconciliation

- Data reconciliation will be performed between ADF Regional Clarity and source Clarity HCCLXX to ensure accuracy of data.
- Data validation will be performed against the source for Clarity*N data accounting for schema changes.
- Some methods to perform data reconciliation include:
  - Counts
  - Checksums
  - Minus Queries
  - Distribution Queries

Future Decision: Upon completion of the Proof of Concept to confirm feasibility and operationalization of the Solution Design recommendation, results will be presented and approval will be requested of the Steering Committee for adoption of design.

Foundation Data - Data Ingest

Foundation Data (FD) Solution Plan Overview

The Foundation Data Solution Plan provides the sequence and timeline to deliver the data sources, reference/master data, and enrichments on the Analytic Digital Foundation (ADF)

Timeline view with milestones for Foundation Data on the ADF*

11 TENANTS

77 DATA SOURCES *(Domain and Sub-Domain)*

39 MASTER & REFERENCE DATA

40 DATA ENRICHMENTS

Approach to Developing FD Solution Plan

Inputs

- **Primary Input:** The Tenant Data Collection effort identified a Tenant's data source, master and reference data, and enrichment needs by analytic asset
- **Other Inputs:** Analytic asset migration dates (A20 Integrated Roadmap), scope of the migration, available capacity/funding, and data producer's capability/ readiness to support the ingestion of data to the ADF

Create and Update

- **Create:** An initial version of the FD Solution Plan was developed based on the Tenant Data Collection input and established plan for 2019 with some high-level assumptions
- **Update:** The "Go Live" dates for data sources, master and reference data, and enrichments will need to be constantly managed as inputs change to strike the balance between tenant needs and data producer's capability while taking into account constraints like capacity and funding

Outputs

- **Timeline**: Produce a detailed timeline (at the domain and sub-domain level) with "Go-Live" milestones*, gaps, and dependencies for data sources, master and reference data, and enrichments.
- **Alignment:** Summarized high-level view of the FD Solution Plan milestones will be incorporated into the A20 Integrated Roadmap to facilitate alignment across the various A20 workstreams

***Note:** The milestones represent the availability of data on ADF in a zone(s) accessible to the tenants

KAISER PERMANENTE®

Foundation Data Solution Plan – 2019 Targeted Delivery Scope

KPHC Clarity	• Clarity One-time load for all regions#~800 tables (NCAP, RMA, F&O) • Tables required by RMA incremental data in Production • Clarity*N capability
OneLink	• OneLink GL domain - KFH and KFHP one-time and incremental load for all regions in Production
Claims Data Warehouse	• Tables required by NCAP one-time load for all regions • Professional, Institutional, Pharmacy claims, Claims Finance incremental data in QA
Pharmacy Data Warehouse	• Bridge Feed, Sub Ledger and Point of Sale (PoS) data for all regions one-time and incremental load for all regions in QA
Member Month	• Analytic view and supporting tables one-time and incremental load for all regions in Production

Data Ingest Status Reporting

ADF – Data Ingest Status 2019

As of 5/24/19, 12:00 PM

Data Custody has been completed for all Data Domains

	SOR	Region	QA	Product Support	Production	Target Dates (Based on Earliest Tenant Request)
Resolute Inpatient and Professional Billing, OpTime, Structure Files	KPHC Clarity	NW	4-Apr	24-Apr	12-Jul	Target Prod Support Date: 2/15/2019 Target Production Date: 2/25/2019
		NCAL	10-Apr	26-Apr	12-Jul	
		MAS	29-Apr	20-May	12-Jul	
		SCAL	3-Jun	17-Jun	12-Jul	
		CO	1-Jul	3-Jun	12-Jul	
		GA	17-Jun	1-Jul	12-Jul	
		HI	1-Jul	1-Jul	12-Jul	
General Ledger • KFH and KFHP entities only	OneLink	All Regions	28-Mar	29-Mar	12-Jul	Target Prod Support Date:3/01/2019 Target Production Date:
Pharmacy Utilization	PDW		19-Apr	3-May	22-Sep	Target Prod Support Date:7/14/2019 Target Production Date:
Pharmacy Claims	CDW		5-Apr	3-May	22-Sep	Target Prod Support Date:9/13/2019 Target Production Date:
Professional and Institutional Claims	CDW		5-Apr	3-May	22-Sep	Target Prod Support Date:9/13/2019 Target Production Date:
Member Month • Analytic View	NPSA		14-Jun	20-Sep	22-Sep	Target Prod Support Date:9/13/2019 Target Production Date:

- **KPHC Clarity**
  - GA - QA : 130 tables do not exist in data delivered from KPHC, confirming with KPHC team this matches the WITS environment
  - MA, CO – PS : Data content validation is complete, ADF-Data Delivery team not enabled to apply/ validate PII tags to the data, ADF Services team working on an interim solution

- CDW Intake process completed, working with SOR team to fast track onetime load of 40 tables as needed by NCAP
- Pharmacy requirements and scope for NCAP complete, working with SOR team to plan for onetime load as needed by NCAP.

Complete | On Track | At Risk | Off Track | Planned but not Committed Date | Committed and yet to Start

KAISER PERMANENTE®

Status updated weekly and available on A20 Site; Link: https://sp-cloud.kp.org/sites/A20Program/SitePages/A20-Tracks.aspx

KAISER PERMANENTE®

Tenant Services

Key Objectives

- Describe the Tenant Data Collection Approach and key learnings
- Describe the Tenant Engagement Process and toolkits that have been developed to facilitate Tenant onboarding, development, delivery and benefit realization
- Discuss the importance of and share the content and process for establishing standards and practices for the ADF

Tenant Services – Tenant Data Collection

Tenant Data Collection Approach

Tenant Personas

Collect key characteristics of the tenant personas and map to the analytics assets

- **1.1** Tenant Persona Validation
- **1.2** Persona to Analytic Asset Mapping

Analytic Asset - Deep Dive

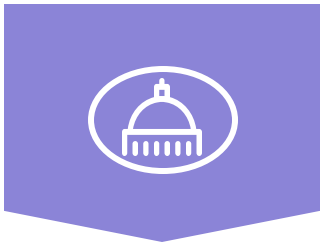
Capture the roadmap, business cases, technical details, and non-functional requirements

- **2.1** Analytic Asset -Overview
- **2.2** Analytic Asset -Technical Overview
- **2.3** Analytic Asset - Non Functional Requirements

Analytic Asset - Sources & Downstream

Identify information domains relevant to the analytic assets and document generated outputs

- **3.1** Analytic Asset - Data Sources
- **3.2** Analytic Asset - Master & Reference
- **3.3** Analytic Asset - Enrichments
- **3.4** Analytic Asset - Output



Regulatory & Compliance

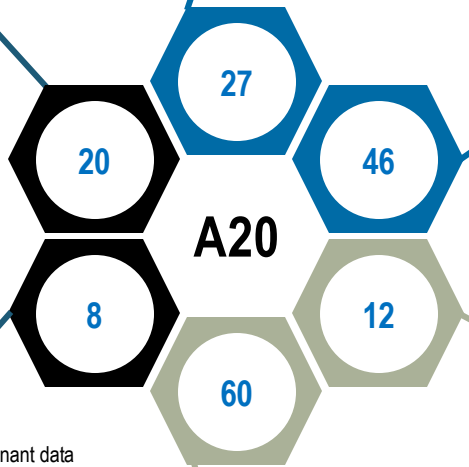
Capture applicable regulatory and compliance areas for applicable analytic assets

- **4** Regulatory & Compliance

Data Collection Template was pre-filled prior to the kick-off sessions based on available documentation

KAISER PERMANENTE®

Tenant Data Collection – By the Numbers



Touchpoints with Tenants

- **27** kick-off meetings and working sessions were hosted by the core A20 Program team with Tenant stakeholders to address questions and gather feedback on the data collection template

Analytics Assets[1]

- **20 of 26** targeted assets (partial scope - NVT and IDR) were covered in the tenant data collection responses (as of Feb 13)
- Responses for the 20 analytic assets covered **53%** of questions in the Tenant Data Collection template

Tenant Stakeholders

- **46** stakeholders (excluding the core A20 Program team) participated in the kick-off sessions with the ten tenants
- Additional tenant team members were engaged by their respective leads to complete the data collection template

Data Collection Activity Duration[1]

- **12 weeks,** was the duration of the tenant data collection activities
- First kick-off session was held on 11/29/18 and final set of gap-closure/validation activities were completed on 02/13/19

Tenant Responses[1]

- **8 out of 10** targeted tenants responded to the tenant data collection request
- Service, Utilization, and Quality, Membership and Market Analysis did not complete the data collection request by Feb 13
- F&O wasn't requested to complete the tenant data collection template

Supporting Documents

- **43** documents were utilized to pre-populate the data collection template before the kick-off meetings with tenants
- **17** additional documents were received from tenants as a part of their response to the data collection request
- All of the supporting documents have been uploaded to the A20 Program site

Perspective Inputs Results

[1] Refer to the Appendix for details on the assets covered, and the timeline of data collection activities

KAISER PERMANENTE®

Key Takeaways from Tenant Responses

Tenant Personas

No new Personas *were identified and all activities were found to be applicable. All of the analytic assets (except for MARC) currently* ***satisfy the needs*** *of the associated Personas in the Tenant organization*

Varying Needs

Analytic assets have ***varying data needs*** *(e.g., latency, history, purging and archival policies), which need to be synchronized as they are migrated to ADF and* ***may require code remediation***

Data Needs

Approximately ***one-third of the data sources satisfy nearly 50% of all source data needs*** *identified by the tenants. However, the input provided by Tenants will need to be analyzed (e.g., NVT requested the most data sources, master/reference data, and enrichment needs)*

Migration Risks

Ability of the Analytic Digital Foundation (ADF) to ***support Tenant needs*** *and* ***disruption from the migration*** *(including remediating downstream dependencies) were the common themes across a majority of the risks raised by Tenants*

Slightly Off-Target

Non-functional requirements ***didn't provide the required information*** *to set-up and configure the ADF. Additional targeted working sessions will be required to gather the required information*

Data Management

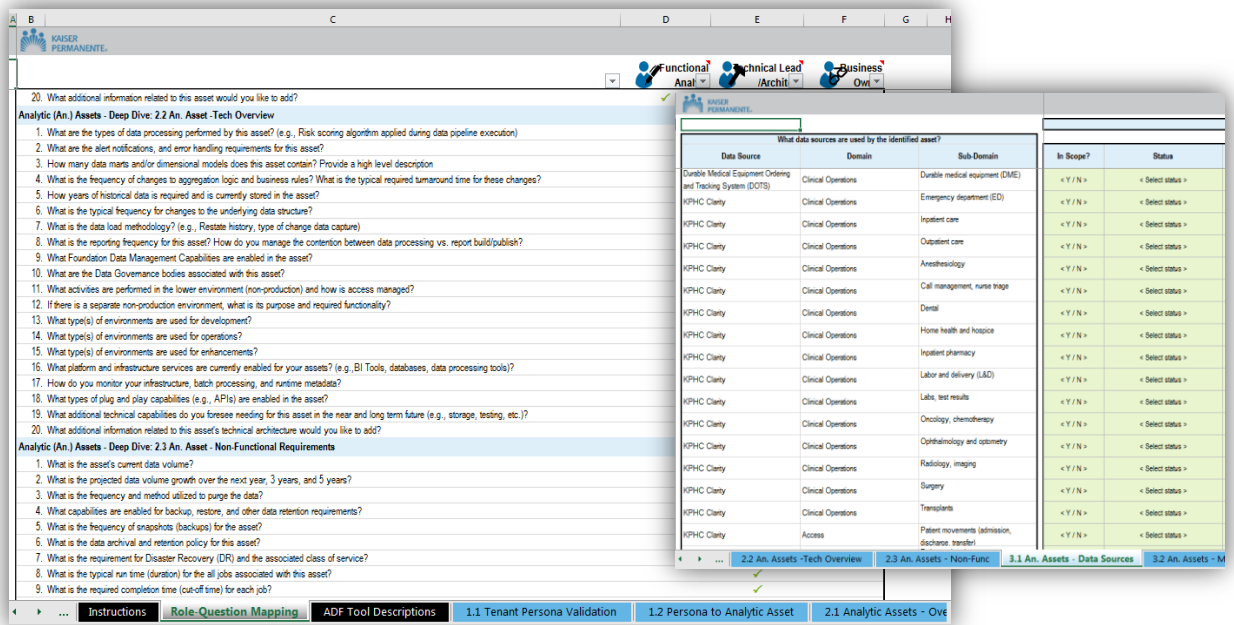
All in-scope analytic assets either requested or currently have ***one or more Data Management capabilities****, underscoring the importance for CDO's engagement in developing the ADF*

Org. Readiness | Foundation Services | Foundation Data

KAISER PERMANENTE®

Tenant Data Collection Template (link)

In-scope tenants completed an extensive and structured questionnaire to inform the A20 Program blueprint efforts. This questionnaire will be reviewed as tenants on-board to the analytic digital foundation.

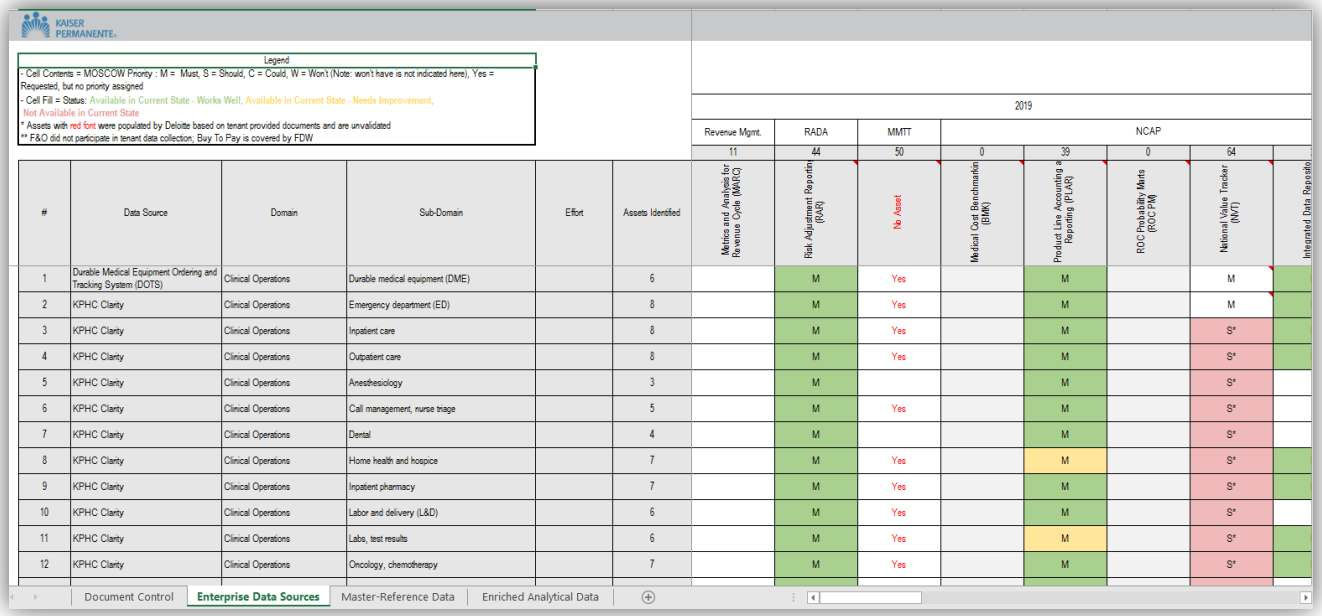


Future Decision: Data Access & Use Workgroup will advance the Tenant Data Collection Template as the intake form for Approved Control#4.

KAISER PERMANENTE

Tenant Data Needs Cross-Analysis (link)

Tenant data needs were consolidated into a cross-analysis by asset and data type: Sources, Master-Reference, and Enrichments.



KAISER PERMANENTE

Legend

- Cell Contents = MOSCOW Priority : M = Must, S = Should, C = Could, W = Won't (Note: won't have is not indicated here), Yes = Requested, but no priority assigned
- Cell Fill = Status: Available in Current State - Works Well, Available in Current State - Needs Improvement, Not Available in Current State
* Assets with red font were populated by Deloite based on tenant provided documents and are unvalidated
** F&O did not participate in tenant data collection; Buy To Pay is covered by FDW

						2019							
						Revenue Mgmt.	RADA	MMTT	NCAP				
						11	44	50	0	39	0	64	
#	Data Source	Domain	Sub-Domain	Effort	Assets Identified	Metrics and Analysis for Revenue Cycle (MARC)	Risk Adjustment Reporting (RAR)	No Asset	Medical Cost Benchmarking (BMK)	Product Line Accounting & Reporting (PLAR)	ROC Probability Metrics (ROC-PM)	National Value Tracker (NVT)	Integrated Data Repository
1	Durable Medical Equipment Ordering and Tracking System (DOTS)	Clinical Operations	Durable medical equipment (DME)		6		M	Yes		M		M	
2	KPHC Clarity	Clinical Operations	Emergency department (ED)		8		M	Yes		M		M	
3	KPHC Clarity	Clinical Operations	Inpatient care		8		M	Yes		M		S*	
4	KPHC Clarity	Clinical Operations	Outpatient care		8		M	Yes		M		S*	
5	KPHC Clarity	Clinical Operations	Anesthesiology		3		M			M		S*	
6	KPHC Clarity	Clinical Operations	Call management, nurse triage		5		M	Yes		M		S*	
7	KPHC Clarity	Clinical Operations	Dental		4		M			M		S*	
8	KPHC Clarity	Clinical Operations	Home health and hospice		7		M	Yes		M		S*	
9	KPHC Clarity	Clinical Operations	Inpatient pharmacy		7		M	Yes		M		S*	
10	KPHC Clarity	Clinical Operations	Labor and delivery (L&D)		6		M	Yes		M		S*	
11	KPHC Clarity	Clinical Operations	Labs, test results		6		M	Yes		M		S*	
12	KPHC Clarity	Clinical Operations	Oncology, chemotherapy		7		M	Yes		M		S*	

Document Control | Enterprise Data Sources | Master-Reference Data | Enriched Analytical Data

KAISER PERMANENTE

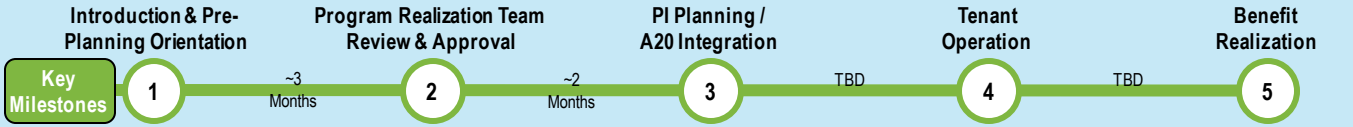
Tenant Engagement

A20 Program Tenant Journey Overview

Tenants are built upon data and services in the analytic digital foundation (ADF) to **deliver analytic and reporting capabilities** that address business objectives and produce insights for **distinct decision makers**.

A20 Program leaders will partner with tenant leadership to facilitate the tenant journey from initial onboarding, through product development and delivery, to ongoing operations and final benefit realization

A20 Tenant Journey



- ~4-5 hour pre-planning meeting with A20 Program leaders to knowledge exchange and begin onboarding process
- Additional follow-ups with key A20 Program leaders (i.e., ADF)

- Present charter to Program Realization Team (PRT) for formal A20 Program approval
- Resolve any open items or requests (if needed, at the recommendation of the PRT)

- Participate in first Program Increment (PI) Planning event
- Work with ADF and other tenants to review and accept dependencies

- Work with IPMO support functions (roadmap team, finance, SAFe coaches) to complete work over required 10-week PI cycles
- Demo final product for business stakeholders

- Live in production on the A20 Program platform
- Work with Benefit Tracking Team (BTT) to document realized benefit (e.g., asset retirement or transition, efficiency gains, cost savings)

KAISER PERMANENTE®

A20 Program Tenant Engagement Model Checklist

The tenant journey consists of five key milestones with critical activities and integration with key A20 Program partners

	1 Introduction & Pre-Planning Orientation	2 Program Realization Team Review & Approval	3 PI Planning / A20 Integration	4 Tenant Operation	5 Benefit Realization
Key Milestones					
Tenant Activities	❑ Complete New **Tenant Onboarding Charter** ❑ Complete (or update) **Tenant Data Collection Template** ❑ Confirm funding source / business case ❑ Confirm Tenant resource requirements ❑ Confirm A20 resource support (e.g., IPMO, SAFe) ❑ Pre-planning discussion with IPMO resource management team ❑ Pre-planning discussion with IPMO finance management to develop initial forecast with confidence level / funding	❑ Assign SAFe team roles ❑ Attend SAFe training ❑ Prep for first PI planning session ❑ Set up Confluence / JIRA ❑ Develop status reporting capabilities (integrate with IPMO) ❑ Integrate with IPMO risk and issues escalation process ❑ Initial integration with A20 integrated program roadmap ❑ Engage A20 org readiness for change management and communications toolkit ❑ Ongoing Finance Mgmt activities	❑ Establish "scrum of scrums" meetings ❑ Develop long-term PI planning roadmap and MVP sequencing ❑ Add long-term planning and dependencies to integrated program roadmap ❑ Manage story, feature, and product backlogs ❑ Confirm ongoing "operate" resource plan ❑ Ongoing Finance Mgmt activities ❑ Define benefits with benefit tracking team & align on benefit cadence	❑ Ongoing Finance Mgmt activities ❑ Review direct and indirect expenses for benefit tracking (with Tenant Finance and Benefits Tracking Teams)* ❑ Inform benefit tracking team and confirm completion of retirement checklist* ❑ Determine level of tenant support for A20 prod	
Analytic Digital Foundation (ADF) Integration	❑ Conduct pre-planning discovery meeting ❑ Tenant overview (including business objectives discussion) ❑ Foundation services ❑ Foundation data ❑ A20 support structure ❑ Lessons learned ❑ Next steps ❑ Conduct follow-up discussions with ADF leaders, as needed ❑ ADF / IPMO team to complete **Estimation Model & Tenant Assessment**	❑ Provide initial ADF needs and dependencies ❑ Confirm data access and use standards, e.g., training and toolkits ❑ Coordinate data custody agreements ❑ Review prioritized list of standards for ADF approved by the Community of Practice (CoP) ❑ Address compliance, HIPAA, and/or SOX requirements	❑ Confirm services and roles from IT / ADF (dev, post-production support) ❑ Provide ongoing data needs and deliver on dependencies ❑ Adhere to S&Ps approved by CoP for ADF (e.g., Data Security, HIPAA, SOX Compliance)	❑ Confirm services and roles from IT / ADF (enhancements) ❑ Continue, monitor adherence to S&Ps, and review any new ones approved by CoP ❑ Conduct due diligence on asset dependencies and retirement feasibility, including user impact* ❑ Transition production business processes to A20 Prod* ❑ Support tenant business processes go-live on A20 Prod ❑ Transition on-prem asset to read-only state and/or set as backup* ❑ Turn off on-prem asset instance*	

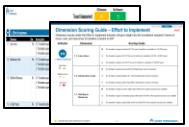
Data collection and effort estimation must be completed prior to conducting the Tenant Assessment

❑ Black check box = Self-Service
❑ Orange check box = A20 / IPMO integration
* Applicable for tenants with asset refirements in scope

KAISER PERMANENTE®

Overview of the Tenant Assessment Process

The tenant assessment is executed primarily by the A20 Program team through a three step process by utilizing existing A20 Program Tenant onboarding/ ADF artifacts and a scoring guide to score the dimensions and report the summary findings to the PRT

	Collect	Analyze	Report
	Assemble all of the data points required to enable the analysis of a Tenant's complexity and readiness for transition to ADF in the Analyze phase; Distribution of Effort: 20%	*Utilize the scoring guide and the data points collected to score the applicable dimensions in the Tenant Assessment Template; Distribution of Effort: 70%*	*Present an assessment of a Tenant's complexity and transition readiness to the PRT through by summarizing the scores developed in the Analyze phase; Distribution of Effort: 10%*
Inputs	• Specific data points from A20 and ADF work products for each dimension of an indicator	• Complete set of inputs from the Collect step • The Dimension Scoring Guide • The Indicator and Global Assessment Templates	• Assessment Scores from the Analyze step • The PRT Indicator and Global Assessment Templates
Actions	• Access the appropriate sources to collect relevant inputs • Verify inputs with responsible stakeholders	• Derive a transition maturity score for each specific dimension • Average the dimension scores for each indicator and derive indicator scores • Average all dimension scores for and derive a global scores	• Record indicator and global scores • Provide and present the Indicator and Global Assessment Templates to the PRT
Outputs	• A complete set of inputs that are aligned to each dimension  Input Collection Guide	• A quantitative assessment of a Tenant's complexity and readiness  Scoring Guide and Excel Template	• An executive-facing summary of a Tenant's complexity and readiness to transition to ADF  Tenant Assessment PRT Summary Templates

KAISER PERMANENTE®

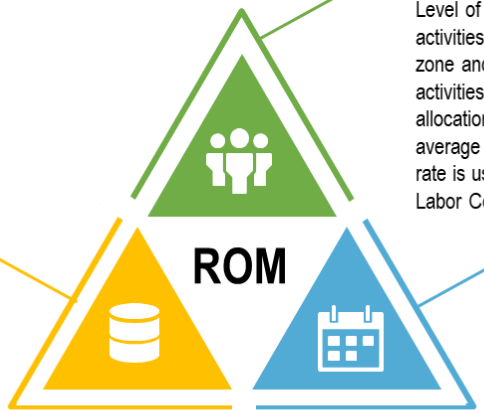
Estimation Model Template Overview

The purpose of the Estimation Model Template is to develop Rough Order of Magnitude (ROM) estimates for labor cost/hours and infrastructure cost. The scope and ROM estimate is typically completed in Phase 1 of a tenant's onboarding journey.

Key Features of the Estimation Model Template

- **Parameterized Input**
- **Complexity-Based Estimation**
- **Highly Configurable**

Key Outputs Generated by the Estimation Model Template



LABOR COST & HOURS

Level of effort is generated utilizing activities considered in scope for a zone and the complexity of those activities. Onshore and offshore allocation is then used to calculate an average rate per hour. This blended rate is used to calculate the Total Labor Cost.

INFRASTRUCTURE COST

Infrastructure Costs are generated based on the software/hardware components and the pricing tiers selected across the various environments (Production, Development, Testing, and Production Support).

RESOURCE ALLOCATION

Role, Resource Type, Employee Type, Start Date, Duration, and Skillsets are identified to satisfy the LoE hours estimated for each zone. Using this information, a Resource Calendar is formulated to identify allocation of resources on a weekly basis.

KAISER PERMANENTE®

Standards & Practices

Advisory Groups Focus on Key Themes for Analytics

- Make recommendations on Standards & Practices for ADF to Steering Committee
- Work began with A20 Program kick off (Basecamp) in 2018
- Informal advisory groups have been established to support ADF blueprint, meet weekly

Metrics Delivery

- Develop recommendations to establish common dimensions
- Develop a standard approach to metric definition, code development, validation, and user acceptance testing to enable accurate and comparable operational definitions
- Develop recommendations about technical implementation
- Basecamp slides 38, 39, 40

Foundation Data Consortium

- Establish uses, behaviors, and characteristics of data flow on the analytic digital foundation
- Define and prioritize activities and interactions necessary to ensure reliability of the data to promote transparency, quality, and integrity by zone of information
- Address data modeling needs and challenges related to delivery of analytics solutions
- Basecamp slides 41, 42

Capabilities & Tools

- Establish and document decision criteria (business & IT) for analytics tool selection
- Focal use cases include data management, data visualization & self-service, statistical analysis, and advanced analytics
- Basecamp slides 47, 48

Data Access & Use

- Establish standards and implement processes to support the Data Access & Use proposal
- Basecamp slides 45, 52, 53, 54, 55

Steering Committee Guidelines

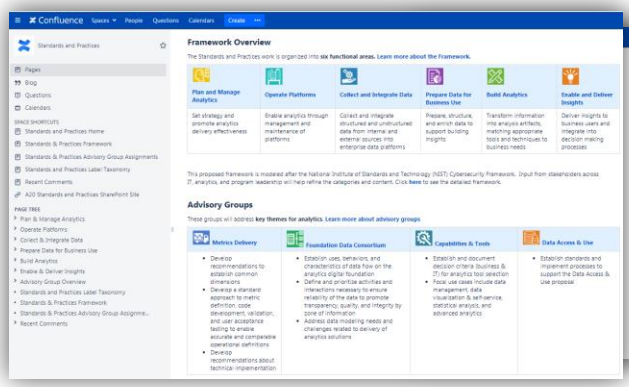
1 **Approve** potentially controversial standards

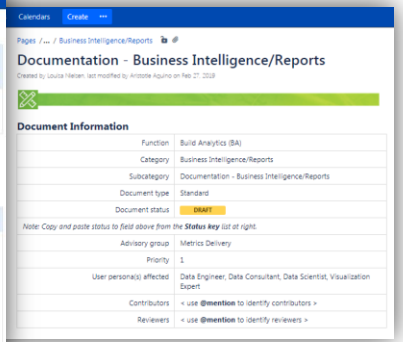
A20 Program Advisory Groups will advance to the A20 Program Steering Committee **sensitive and/or impactful architecture, solution designs, standards and controls and topics**. The A20 Program Steering Committee can also request that Advisory Group topics be advanced to the Steering Committee.

2 **Decide** when advisory groups cannot agree

A20 Program Advisory Groups will advance architecture, solution designs, standards and controls where they cannot reach agreement. The A20 Program Steering Committee will **provide direction and be the arbitrator for decisions** where Advisory Groups **cannot reach agreement**.

3 **Informed of** advisory group activities, **recommend** participants





- Website - https://confluence.kp.org/display/SP

KAISER PERMANENTE®

Framework for Standards and Practices

The Analytics Capability Framework* depicts 6 foundational areas necessary to support operational analytics. These align with architectural and persona-based perspectives.

PERSONAS


Analytics Leader


IT Support


Data Engineer


Data Consultant


Visualization / Reporting Expert


Data Scientist


Decision Maker

Plan and Manage Analytics	Operate Platform	Collect and Integrate Data	Prepare Data for Business	Build Analytics	Enable and Deliver
Set strategy and promote analytics delivery effectiveness	Enable analytics through management and maintenance of platforms	Collect and integrate structured and unstructured data from internal and external sources into enterprise data platforms	Prepare, structure, and enrich data to support building insights	Transform information into analysis artifacts, matching appropriate tools and techniques to business needs	Deliver insights to business users and integrate into decision making processes

ARCHITECTURE

Tenancy

Foundation Services

Foundation Data

KAISER PERMANENTE®

Framework for Standards and Practices

The proposed framework is modeled after the National Institute of Standards and Technology (NIST) Cybersecurity Framework*.

Function	Category	Category Description
Plan & Manage Analytics	Infrastructure	Reusable infrastructure services are discoverable & understood.
	Platform	Reusable platform services are discoverable & understood.
	Foundational Data Management	Reusable data assets are discoverable, understood and governed transparently.
	Governance	Oversight to data domains, business rules, enrichments, tools and services on the Analytics Digital Foundation (ADF) helps ensure they are complete, non-duplicative, communicated and evolved.
	Analytics Services	Flexible services support innovation and speed to insight.
Operate Platforms	Security Framework	Access to platforms and infrastructure is limited to authorized users, processes, and devices, and is managed consistent with the assessed risk of unauthorized access to authorized activities and transactions. (NIST)
	Data Security	Information and records (data) are managed consistent with KP's risk strategy to protect the confidentiality, integrity, and availability of information. (NIST)
	Production	Procedures are maintained and used to manage information systems and assets.
Collect & Integrate Data	Data Ingest (Raw Data)	Raw data is stored in its native or near-native format, typically sourced from a System of Record (SoR). Though not intended to meet the needs of wider audiences who require curated data, it may address the needs of a limited data science audience.
	Data Refinement (Refined Data)	Refined data is prepared for analyst use by removing inconsistencies and errors, improving quality and semantic consistency, and simplifying access. The goal is to provide good-quality, trusted data in a way that is understood and accessible.
	Shared Enrichments	Enriched
Prepare Data for Business Use	Dimensional Models	Data prep
	Tenant Improvements	Tenants a
	Extracts	Data extra
	Analytic Data Layer	Data prep
Build Analytics	Question/problem definition	Analysis q
	Business Intelligence / Reports	Descriptiv
	Model Development	Methods,
	Results Communication	Results ar
Enable & Deliver Insights	Decision Making	Decisions
	Management & Operations Reporting	...
	Embedded Insights	...
	External Customer Delivery	...

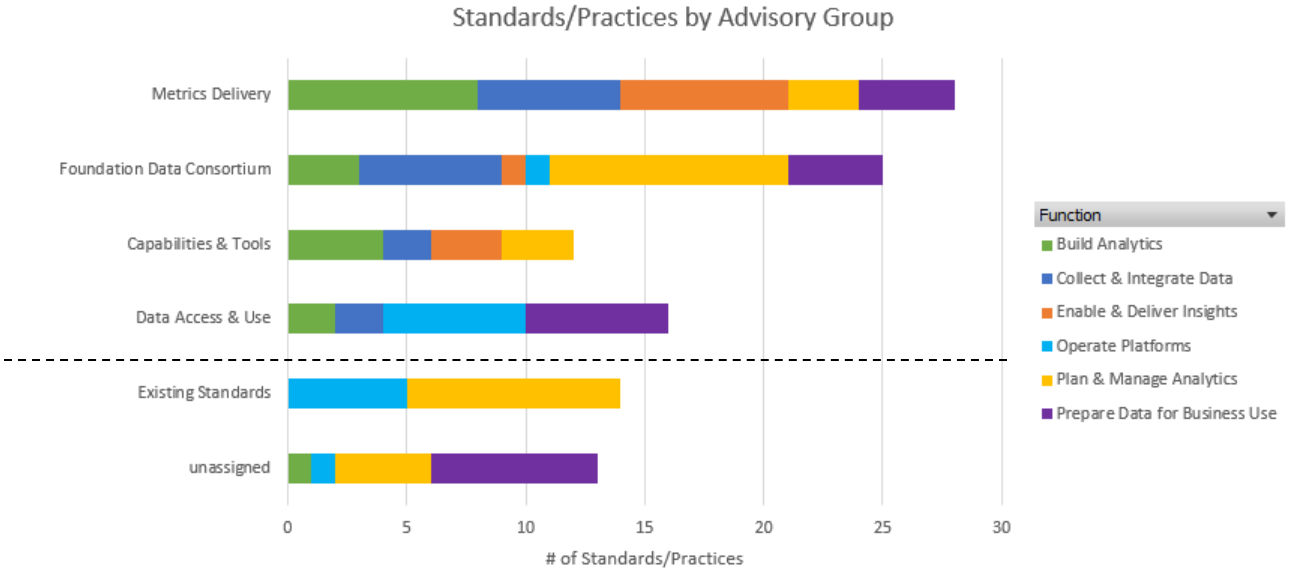
Function	Category	Category Description	Sub-Category
Build Analytics	Question/problem definition	Analysis question, benefits, stakeholders and data inputs are identified. Prior analysis results and production reports are searchable to prevent duplication of effort.	Search
			Requirements gathering
			Data use
			Stakeholder analysis
			Expected benefit
	Business Intelligence / Reports	Descriptive reports are accessible on-demand, discoverable, and documented to facilitate re-use and adoption.	Access control (National/Regional, Provider, Member)
			Publish to portal
			Documentation
			Change management
	Model Development	Methods, workflows, data sources, and tools are established and evaluated based on the ability to address identified questions.	Methods selection
			Workflow development
			Data relationships
			Test/Train/Prod dataset development
			Documentation
			Tenant-specific tool integration
	Results Communication	Results are communicated as broadly as appropriate after considering sensitivities related to information sharing.	Data visualization
			Publish to portal
			Documentation

* https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.04162018.pdf#page=30

KAISER PERMANENTE®

Advisory Groups – Identified Relevant Topics

- The proposed framework includes **~100 topics** across 6 functional areas.
- Standards & Practices have been "assigned" to advisory groups and are **under review for relevance and priority by the advisory groups**.
- Participants at the A20 kick off (Basecamp) **identified 14 "Existing Standards,"** most of which are platform-related.
- Some standards are currently "unassigned," and may reflect gaps related to Production Support and Governance.



KAISER PERMANENTE®

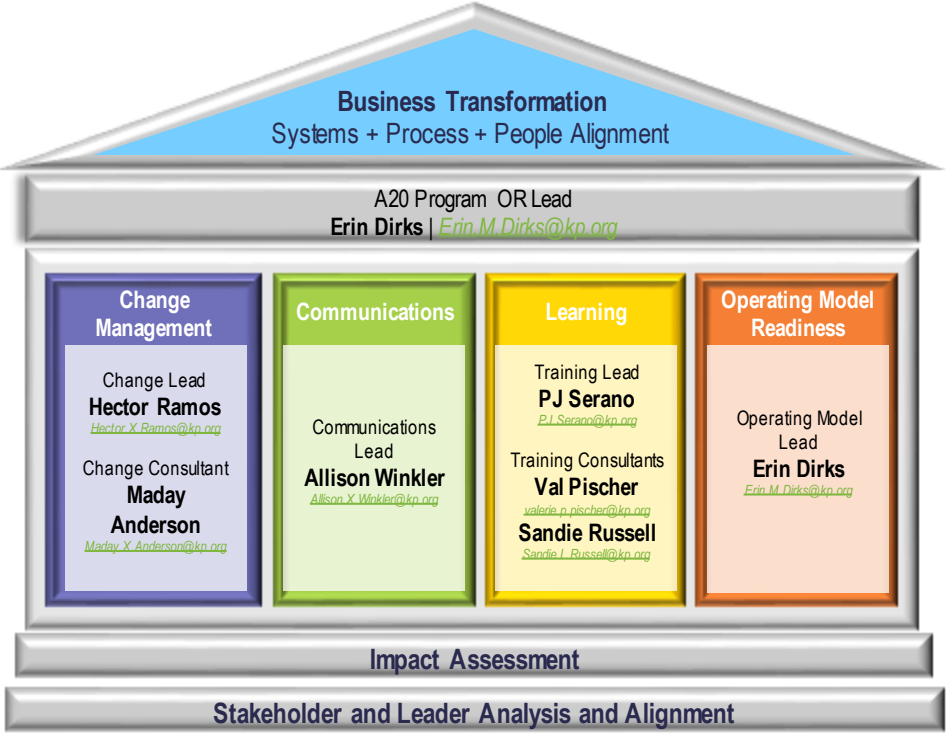
Organizational Readiness

Key Objectives

- Provide information about the approach the A20 Program organizational readiness team is using to ensure stakeholders are ready to adopt the changes that will result from the A20 Program
- Provide details on in-progress and upcoming organizational readiness activities

Organizational Readiness Overview

Organizational Readiness Overview and Team



Goal: Improved Readiness

We are here to support and consult with you on preparing you and your stakeholders for change

Change Management

- Determine **effects of change** on sponsors and employees
- **Develop and execute** change management strategy
- Align **stakeholders and leaders** to new reality

Communications

- Plan and Execute **strategic communications**
- Create **key messages** to support on-going program objectives
- Support Change Management Strategy
- **Measure effectiveness**

Learning

- **Assess learning needs** across program
- Develop and deliver training to **meet program objectives**

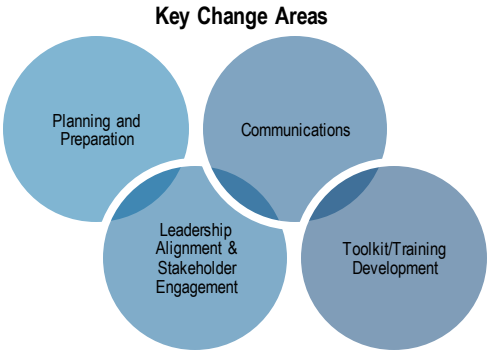
Operating Model Readiness

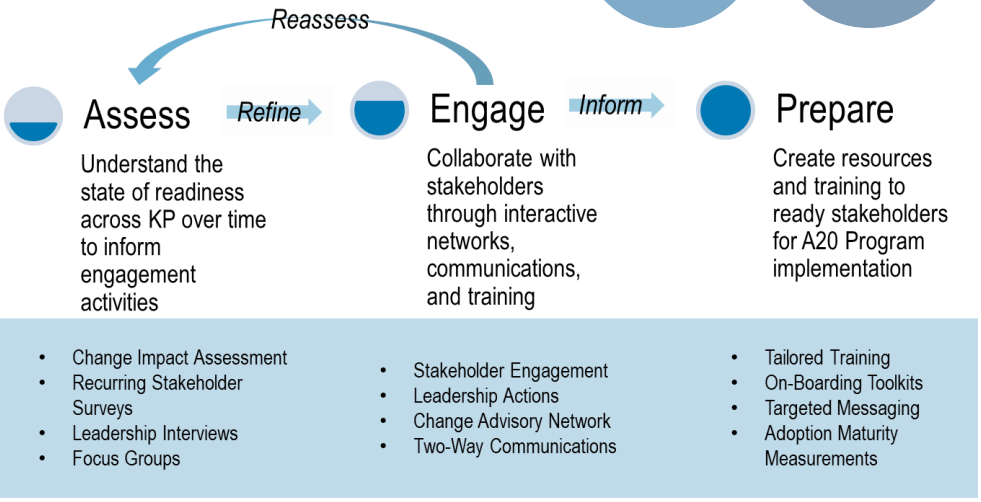
- **Assess readiness** to select operating model design
- **Refine operating model design** based off stakeholder feedback
- **Develop change management strategy and plan** to transition into the operating model
- **Conduct** change management activities

Organizational Readiness Approach

Organizational Readiness Approach Summary

Due to the use of SAFe methodology and tenant onboarding cycles, the Organizational Readiness (OR) Strategy requires an **iterative and integrated approach** for executing activities and developing resources around key change areas





Expected Outcomes:

- ✓ Prepare stakeholders for **changes**
- ✓ Establish a baseline for organizational and stakeholder **change readiness**
- ✓ Present a **compelling vision** to broad audiences using targeted campaigns
- ✓ Promote transparency and **manage expectations** among key stakeholders
- ✓ **Accelerate** the realization of the A20 Program's business **value**
- ✓ **Alignment with the A20 Program's strategy,** guiding principles, timing, and expectations

KAISER PERMANENTE®

Stakeholder Communication Approach

The OR team utilizes a top-down, middle-out, and bottom-up communication to drive a comprehensive approach to communicating with KP stakeholders

Leaders, Program Sponsors & Team Members

drive top-down adoption by communicating the program vision and objectives across the organization to align all key stakeholders

- Executive Sponsors
- Steering Committee/AIC
- Program Realization Team
- Community of Practice
- The A20 Program/NCAP Program teams

Key Stakeholders

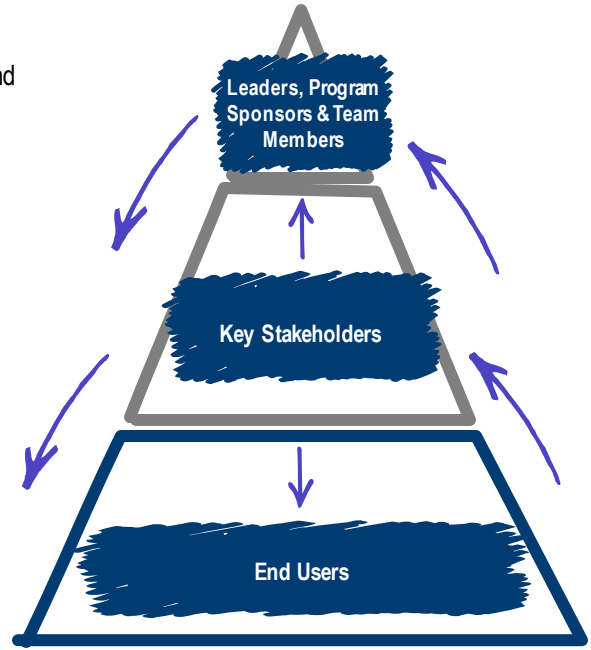
drive middle-out adoption by cascading program information, supporting end-user adoption and funneling questions/concerns back through to the sponsors and change team

- KP Insight leaders
- KPIT leaders
- PMG/Federation leaders
- Regional finance leaders
- Tenant leads
- Other management

End Users

drive bottom-up adoption by identifying and communicating issues and concern

- KP Insight teams
- KPIT teams
- Tenant teams
- Other end users



KAISER PERMANENTE®

Communication Vehicles

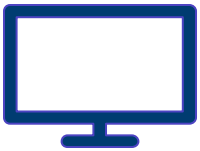
- Utilizing a variety of communication vehicles will maximize the effectiveness of messages and increase the number of key stakeholders engaged

A20 Program Website and Workplace

Central location for program information; program updates, FAQs, and collaboration

Emails

Monthly newsletters; Targeted, action-oriented messaging to specific stakeholder groups



Videos

Key announcements, program update, and features and functionality

Info Sessions & Roadshows

In-person and virtual sessions to receive program information and live Q&A sessions

Face-to Face and Virtual Meetings

Provide context around changes and gain buy-in with key stakeholders

Engagement and Collaboration = High Impact

Change Impact Approach

Change Impacts are differences in the way employees will do their jobs and tasks as the A20 Program is implemented. Identifying change impacts and preparing stakeholders accordingly is critical to the success of A20 Program

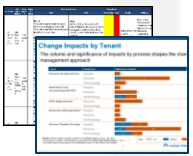
Change Impact Objectives

- ***Identify*** People, Process and Technology Changes
- ***Rate*** degree of change impacts on stakeholders
- ***Gather*** information that will feed into the change management activities for the A20 Program

Inputs

- ✓ Tenant Charters
- ✓ Completed Tenant Data Collection Templates
- ✓ Output from PI Planning Sessions
- ✓ Output from Standards and Best Practices Focus Groups
- ▯ Persona Analysis

Change Impact Assessment and Scorecard



Outputs

Change and Communication Activities:

- Targeted communications for stakeholders most impacted
- Focused messaging for program website, toolkits, newsletters
- Leadership and stakeholder alignment and engagement plans
- Messaging for Change Advisory Network

Process and Technology Training Activities:

- Specific examples for training materials

✓ What we have ▯ In progress/being completed

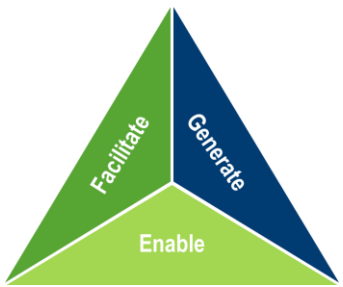
KAISER PERMANENTE®

Change Advisory Network Approach

The A20 Program Change Advisory Network (the A20 Program CAN) members act as change agents by providing information to other stakeholders, and also helping the program team capture feedback and adjust accordingly

Aligning Activities

- CAN members **facilitate** and lead knowledge transfer on A20 Program value, goals, and status to mature organizational readiness across KP.
- CAN members collaborate across the organization to **generate** and communicate unique opportunities, challenges, and risks for the A20 Program.
- CAN members utilize their unique positions to inform the direction and **enable** the success of A20 Program activities.

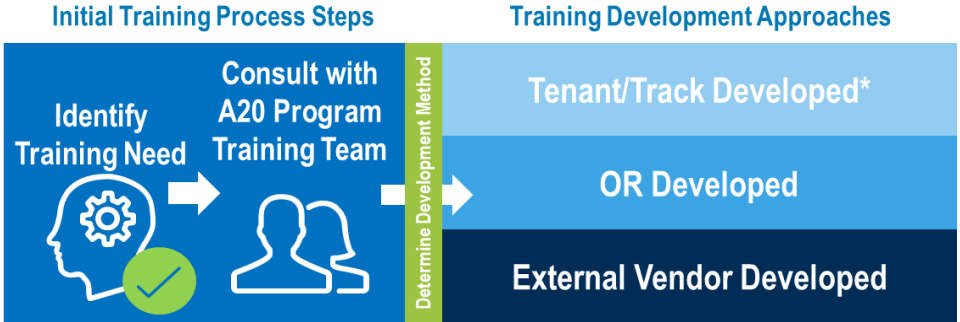


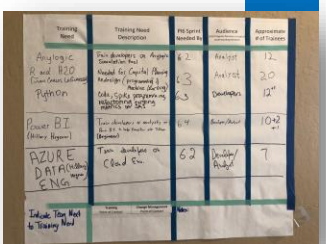
Change Advisory Network	
Responsibilities	• Provide input into the training strategy • Review and provide input into the Change Impacts • Provide input to toolkit development • Provide input to stakeholder engagement approach • Provide input to relevant measurement approach
Expectations	• Engage with A20 Program Team, CAN members, and relevant leadership • Represent and communicate A20 Program values across KP • Collaborate and solicit information from relevant stakeholders
Potential Member Sources	• KP Tenant Teams • KP Insight Staff • A20 Program Team • KP IT • Representation from A20 Persona Leaders

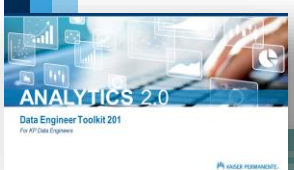
KAISER PERMANENTE®

Training Process Pathway

The A20 Program training team has **developed a comprehensive process** to gather and evaluate training needs, provide training development support, and create training content





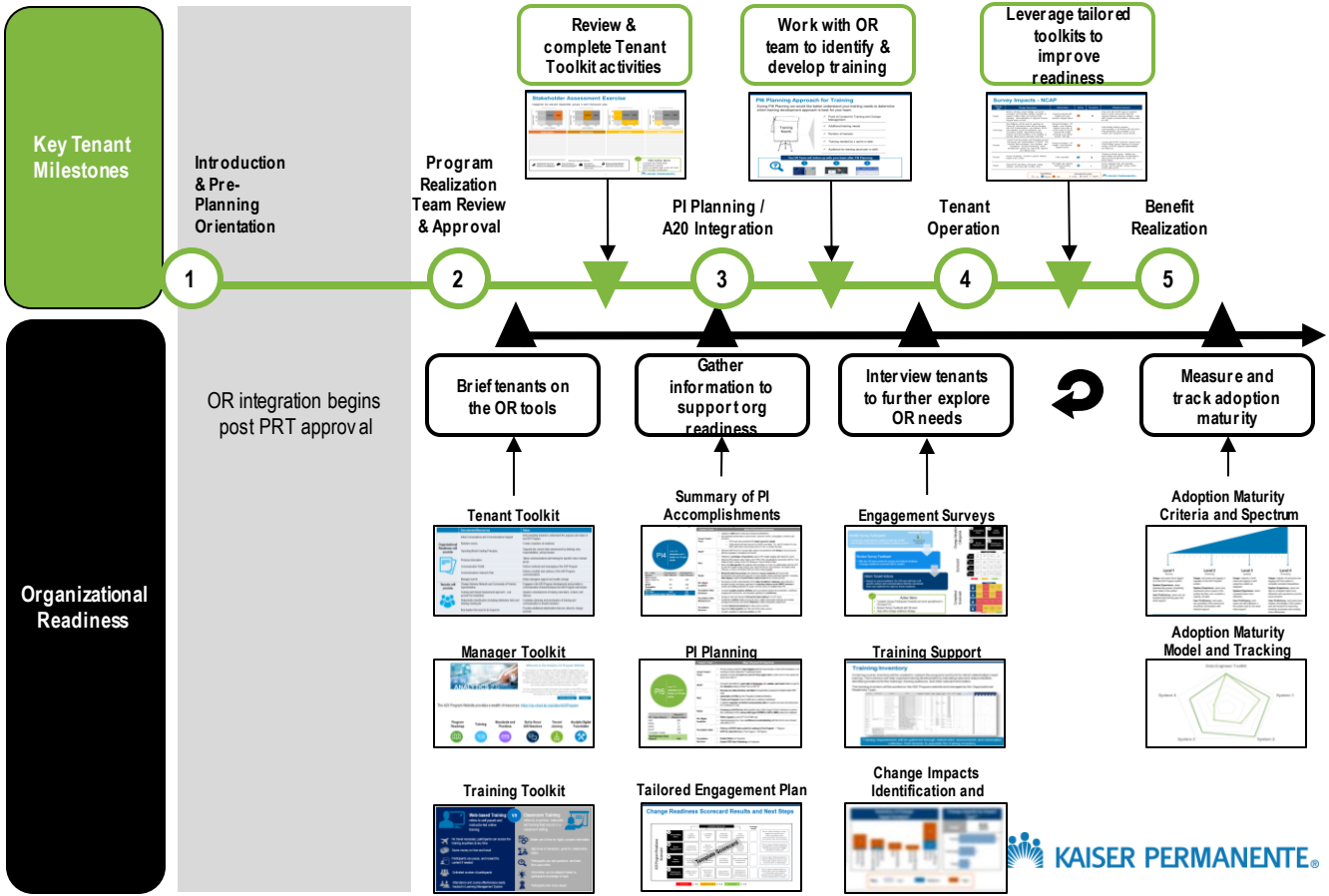


KAISER PERMANENTE®

* Developed with support of A20 Program Track and Tenant Training Toolkit and A20 Program training team

Org Readiness and Tenant Engagement

The Organizational Readiness team will integrate with tenant engagement key milestones and provide a series of resources at each integration point



KAISER PERMANENTE®

OR Integration In Action Overview

Stakeholder Engagement Assessment
(Stakeholder Survey, Focus Groups, and Leadership Interviews)

- The A20 Program roadmap is unclear
- There is a lack of two-way feedback channels

Change Impacts Collection
(Tenant Charters, Change Impacts Survey)

- People will use new tools and processes
- Roles and responsibilities will change

Tenant Needs Discussions
(PI Planning Sessions)

- Tenants have immediate training needs
- Tenants want Change Management tools/templates

- Conducting Get to Know A20 Sessions
- Publishing information to the Program Website
- Developing A20 Overview Training
- Standing up a Change Advisory Network
- Developing Persona Training
- Developing a Training Development Toolkit
- Developing OR Toolkits for Tenants

The OR team operates in an **integrated** fashion using stakeholder outreach to inform Change Management, Communications and Training activities

Note: This slide represents key findings and OR activities and is not intended to be the full list.

Change Management | Communications | Learning

KAISER PERMANENTE®

Examples of OR Resources

Manager's Toolkit

A guide for supporting and leading you and your team through the A20 Program changes

Leading Through Change

Understand the stages of change and how you can support your team members

- Learn about the five goals and outcomes of successful change (Prosci ADKAR® Model) and how the A20 Program is addressing these areas
- Assess your own change readiness so you can be better prepared to support and lead your team through change
- Familiarize yourself with KP resources on leading through change
- Direct your team to KP's EAP website to schedule manager consultations, access wellness resources and more

All About A20

Get to know A20 and stay up to date on Program activities

- Visit the A20 Program Website for an Overview of the A20 Program
- Watch past Get to Know A20 videos and read about Program FAQs
- Get to know the Change Advisory Network and their role in supporting change
- Learn about the A20 Personas and identify the one(s) most closely related to the work that you and your team do
- Stay up to date with what's happening in the A20 Program on Workplace
- Learn about SAFe Agile methodologies by registering for SAFe training

Sharing With Your Team

Leverage the Organizational Readiness Team and resources to support you through change

- Visit the A20 Program Website Training page to find relevant Tenant or Persona training
- Contact the A20 Program Organizational Readiness team for support assessing your team's change readiness and conducting change impacts analysis
- Download and use A20 Program Key Messages in conversations with your team

Have Questions? Email: A20-Updates@kp.org

KAISER PERMANENTE®

Organizational Readiness Toolkit

The goal of Organizational Readiness (OR) is to prepare A20 Program stakeholders for change through training, change management and communications activities. Tenant leaders and teams are key partners in this work given their subject matter expertise and stakeholder relationships. The OR team will provide resources and support to enable the tenant leads and teams to successfully lead their stakeholders through the changes.

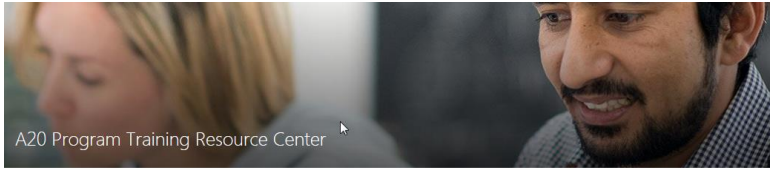
Resources to support and prepare Tenants for change

OR is a shared responsibility

	Resources	Resource Description	OR Activities	Tenant Activities
Change Management	Readiness Survey	Survey that can be administered by the OR team to understand the areas impacting end user readiness	• Administer survey and analyze results	• Provide stakeholder list and participate in survey
	Stakeholder Assessment	Framework to help identify stakeholders, considering their degree of influence and level of engagement to determine the appropriate engagement approach for each key stakeholder group	• Provide resources and stakeholder assessment exercise	• Complete stakeholder assessment exercise
	Leadership Action Plan	Template for documenting key leaders and steps to move each leader to the desired level of engagement	• Provide resources and leadership action plan exercise	• Complete leadership action plan exercise
	Change Impact Assessment	How to assess and the value of understanding what's changing and the level of impact of each change. Assessment is on-going in Agile environment.	• Conduct change impact assessment • Provide Tailored OR Plan	• Provide Tenant documentation (e.g., Charter) • Participate in change impact validation meetings
	Change Advisory Network (CAN)	Network of change agents who gather information and feedback from stakeholders to help the A20 Program team adjust accordingly	• Host monthly CAN meetings • Document outcomes and participation	• Review CAN materials and nominate a CAN representative
Communications	Communication Resources	Series of resources describing the A20 Program's approach for communications	• Provide resources	• Review communication objectives, approach, and vehicles
	Manager's Toolkit	Series of resources tailored for Managers to help lead their teams through change	• Provide toolkit to support Managers in leading through change	• Utilize Manager's Toolkit
	A20 Program Website and Workplace	Information about how to access the Program Website and Workplace	• Provide Program updates and resources via the A20 Program Website and Workplace	• Visit the Website and join Workplace
Learning	Training	Introduction to the A20 Program Training Development Approaches (3 pathways) and contact information for the A20 Program Training Team	• Consult with Tenants on training needs • Maintain A20 Program training inventory and support training development • Provide training toolkit to support Tenant training development	• Work with OR team to identify training needs and development approach • Use training toolkit as needed • Provide map of staff to the A20 Program Personas

Training Resource Center Now Available

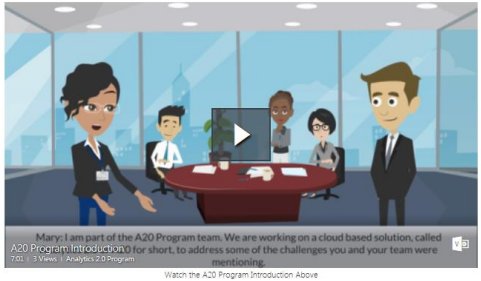
- **A20 Program** Training Resource Center
  - Self-service hub for A20 Program training news
    - A20 Program Training Classes
    - Toolkit Overview
    - Training Inventory
    - Learn more about SAFe
  - A20 Training contact portal



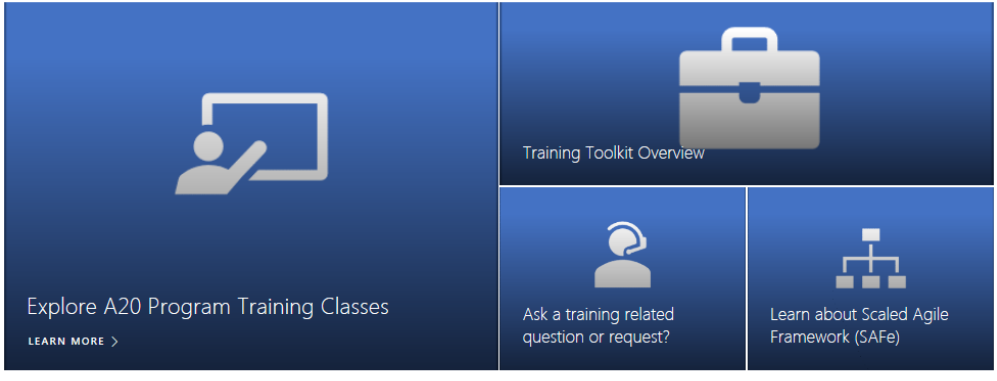
One important piece of the A20 Program Organizational Readiness is to provide our program stakeholders with knowledge advancement and education. This page is designed to assist in providing helpful information on upcoming trainings, as well as key resources.

To get started with training, take a look at our short A20 Program video to the right, and then explore the training inventory below. If you still have further questions, please fill out the form to request more information.

Beyond what you see here, the Organizational Readiness team has ready to use training toolkits, guides and other training-related materials available upon request. So if you don't see something here that meets your needs, let us know by emailing A20-Training@kp.org.



Watch the A20 Program Introduction Above



https://sp-cloud.kp.org/sites/A20Program/SitePages/A20-Program-Training.aspx

Contact A20 Program Training: A20-Training@kp.org

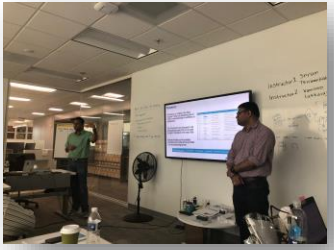
KAISER PERMANENTE®

Data Engineer Trainings

Data Engineer training courses and an initial training pathway have been developed. Introductory courses are available in KP Learn and 2 instructor led trainings (ILT) have been deployed with 5 qualified trainers to over 45 people trained as of March 27th

Data Engineer Persona 101

Data Engineer Toolkit Overview



Data Engineer Toolkit 101 and 201 ILTs

Available Publicly Through KP Learn

Available Through Active Tenant Manager's Approval

Courses available in KP Learn

Catalog Search

New Catalog Item | Print | Export | Modify Table

Title	Id	Type	Available From	Discontinued From	Created By	Created On	Domain	Actions
A20 DE Data Engineer Basic 101	BCS:A20 DE 101.PI6.01	Course	28-FEB-2019	31-DEC-2021	K117201	28-FEB-2019	Kaiser Permanente	Edit Advanced Edit New Class Assign Course View Classes
A20 DE Data Engineer Intermediate 201	BCS:A20 DE 201.PI6.01	Course	28-FEB-2019	31-DEC-2021	K117201	28-FEB-2019	Kaiser Permanente	Edit Advanced Edit New Class Assign Course View Classes
A20 DE Data Engineer Persona Video	BCS:A20 DE 001	Course	28-FEB-2019	31-DEC-2021	K117201	28-FEB-2019	Kaiser Permanente	Edit Advanced Edit New Class Assign Course View Classes

If interested, search for current courses in KP Learn:

Course ID Keyword: BCS:A20

KAISER PERMANENTE

Stay Updated through the A20 Program Website

A20 Program

Home | A20 Program Overview | Get to Know A20 | A20 Tenants | A20 Tracks | ... | Edit

★ Following | Share site | Search this site

New | Page details | Published 3/20/2019 | Edit

Welcome to the Analytics 2.0 Program Website

The Analytics 2.0 (A20) program will bring together programwide, cross-functional data onto a re-usable, cloud-based analytics digital foundation. It will include re-platforming and transitioning 25 analytic systems within KP Insight. The outcome will be a platform that includes consistent and comparable data that can be easily accessed by a suite of analytics services, enabling analysts to deliver critical, high quality, and timely insights to Kaiser Permanente decision makers.

This site is designed to provide information and updates on the A20 Program. As the program progresses the site will be updated with new information. If you have questions on the A20 Program, contact A20-Updates@kp.org

The A20 Program Website provides a wealth of resources: https://sp-cloud.kp.org/sites/A20Program

Program Overview	Get to Know A20 Sessions	A20 Tenants	ADF	Standards and Practices	Training Resources

The A20 Program Workplace site is an enterprise tool that facilitates collaboration and encourages dialogue between A20 Program stakeholders

Erin Dirks, A20 Program Organizational Readiness Lead
Erin.M.Dirks@kp.org

Hector Ramos and Maday Anderson, A20 Program Change Management Leads
Hector.X.Ramos@kp.org; Maday.X.Anderson@kp.org

Allison Winkler, A20 Program Communications Lead
Allison.X.Winkler@kp.org

PJ Serano, A20 Program Training Lead
PJ.Serano@kp.org

Presentation Change Log

Change Log

Version	Update date and time	Owner	Description of changes, including slide numbers impacted if applicable
1.0	May 21, 2019, 3 PM	Erin D	Added change log slide, removed operating model slides
1.1	May 29, 2019, 2:30 PM	Katie C	Reviewed Tenant Services content with minor revisions. Moved governance & lead roles to intro, deleted advisory group names, deleted duplicate slides, changed A20 tenant journey slide
1.2	May 30, 2019, 9:45 AM	Erin D	Updated organizational readiness slides with more recent content.
1.3	5/31/2019 12:51 PM	Claudine M	Updated slides for Foundation Data including Data Zone Architecture, Data Disciplines, Domain Taxonomy, and Data Ingest. Moved Data Management capability back to Foundation Services and removed Tenant Data Collection Needs from Foundation Data. Ref Impl updated to recent content. Kate provided CDO input.
1.4	5/31/2019 5:20 PM	Erin D	Updated FS section with slides from Sumit; incorporated requests from Judy; updated section headers
1.5	6/3/2019, 11 AM	Erin D and Allison	Updated master slide, added introductory slides, updated objectives, updated formatting throughout presentation, moved change log to end of presentation
1.6	6/4/2019, 7:30 AM	Erin D	Updated agenda, section title, objectives, formatting on slide 38
1.7	6/10/2019, 4:45 PM	Erin D	Added intro slides on incremental vs iterative, and use of mural
1.8	6/11/2019, 6:30 AM	Erin D	Added slide on 90s architecture (from Ernest)
1.9	6/14/2019, 10:30 AM	Erin D	Updated formatting in tenant services and org readiness sections, added additional instruction to mural instructions slide; added break to agenda
1.10	6/19/2019, 8:30 PM	Sumit	Updated ADF, SOI and FS sections
1.11	6/26/2019, 6:30 AM	Erin D and Allison	Updated formatting