



Whitepaper

Innovaccer's Unified Data Model (UDM)



Table of Contents

Executive Summary	3
Next Generation Healthcare Data Lake	4
Innovaccer's UDM Design Principles	5
Leveraging Multiple Data Levels to Support a Wide Range of Use Cases	6
Extensive Support for Business and Analytics Use Cases	7
• Population Health Management	
• Interoperability Engine	
• Real-time Analytics	
• Data Warehousing and Marts	
• Archival Storage with Random Access	
About Innovaccer	10

Executive Summary

Innovaccer's Unified Data Model (UDM) is the foundational component of Innovaccer's FHIR®-enabled Data Activation Platform. With a gamut of optimized business and technical data models, the UDM supports a wide range of healthcare use cases, from CMS compliance needs to core clinical operations and analytics applications at scale. UDM has matured over years of instrumentation across the healthcare value chain, including providers, payers, government organizations, and life sciences companies. Through purpose-built healthcare enterprise data lakes and patient-centered business enablement, Innovaccer UDM has significantly accelerated business outcomes for many healthcare entities.

The digital transformation of US healthcare accelerated with legislative reforms such as the Hitech Act and the Patient Protection and Affordable Care Act. By striving to improve care delivery, resource allocation & prioritization, and funding & technology decisions, the reforms facilitated easier access to reliable and readily available healthcare data.

However, there remains immense challenges to integrating healthcare data from disparate sources and empowering business applications to make decisions and improve performance. Furthermore, the volume of the data, evolving regulatory landscape, and requirements around real-time, population-level analytics further adds complexity to instrument and design a comprehensive foundation for business intelligence and advanced analytics.

Innovaccer approached the data challenges with a unique and hybrid approach of leveraging modern distributed data technologies with data model foundations based on data architecture best practice principles. This approach has provided a robust set of business and technical data models inspired by HL7 principles, but expanded to modern analytics use cases that are extensible and scalable. This paper discusses how the aforementioned approach better fits a modern healthcare organization's unique environment and offers significant competitive advantages and business benefits.

Next Generation Healthcare Data Lake

Healthcare data lakes centralize crucial information that can be useful in driving improved care delivery and health outcomes. This data-driven approach is not only mandated at healthcare providers, but also in the other healthcare organizations (e.g., CMS 9115-F rule mandating FHIR APIs and payer-to-payer data exchange by the start of 2022)¹.

However, there is a need for a concrete Data infrastructure plan to organize the data, make it usable, and establish interoperability across the healthcare value chain. Innovaccer envisioned and embraced the opportunity with the next generation Hybrid Data Lake (Figure 1). This provides a competitive advantage to Innovaccer, where it doesn't have a specific system for storage (e.g., Hadoop, MPP, OLTP and/or NoSQL), but offers customized services depending upon solution and business requirements.

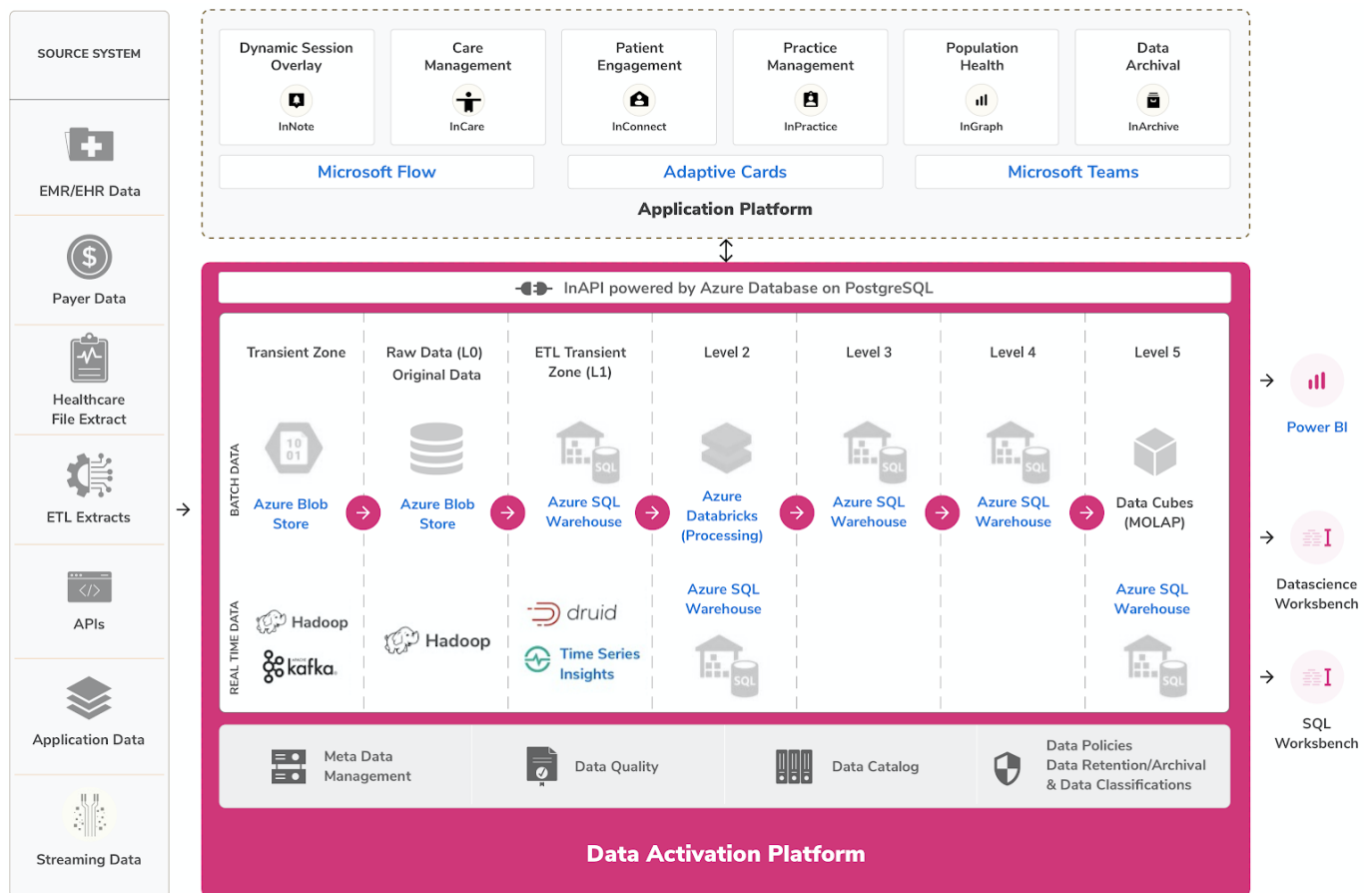


Fig 1. Hybrid data lake infrastructure diagram.

Innovaccer's UDM

UDMs play an important role in developing applications based on business requirements such as during the design phase of a data lake solution. As mentioned in the previous section, Innovaccer leverages and instruments one or more modern distributed data systems for any of its solutions depending on data volume and use case-driven requirements.

Innovaccer's UDM is built with a foundation of strong flexibility in the data systems by not assuming any pre-defined design principles such as normalization or denormalization, but by supporting the full spectrum of possibilities within the UDM. This approach was made possible by following a set of key design principles:

Sl. No.	Principle	Description
1	Patient at the center	The data model is centered around the patient / member, but also supports market-level aggregations (e.g., provider, payer) as secondary support to power up dashboards and applications.
2	Optimization across all the supported storage in the data lake	The UDM is optimized for every storage across the data lake (including distributed file system, ROLAP, MOLAP, and OLTPs) through normalization, de-normalization and further distribution, indexes, and real-time data synchronization, and reconciliation.
3	Healthcare industry standards support	Broad and expanding standards support including FHIR, HL7, HIPAA, CCD, CCPA, GDPR 1, OMOP, CDISC-BRIDG, and SAMHSA.
4	Data level security	Data classification, archival, and retention policies, with access policies at each storage of the data lake ensuring data security across protocol access.
5	Extended genomics coverage	Support for multiple genotype and phenotype data sets.
6	Interoperability and FHIR (batch and real-time interoperability support)	UDM is optimized for applications through FHIR APIs with a normalized data model which also supports all the guarantees on FHIR such as versioning and source meta information. Furthermore, UDM data models at the FHIR level are compatible with interoperability tools to support two-way interoperability.
7	Extensible and backward-compatible	UDM is backward-compatible irrespective of HL7 versions, thus ensuring standard version-agnostic data model without any modifications needed to the upstream or downstream data pipelines.

Leveraging Multiple Data Levels to Support a Wide Range of Use Cases

Recognizing the importance of no-one-size-fits-all, the UDM structures and levels the data to overcome various management and performance limitations (Figure 2). There are multiple levels where the data is stored physically and virtually from the raw storage to summarized, aggregated data to optimize cost, data quality, and downstream SLA use cases. Furthermore, storage, security, data quality, and data modeling are optimized and enriched at each level of the data lake for a true data model for the full spectrum of potential use cases.







Level 0	Level 1	Level 2	Level 3	Level 4	Level 5
Raw data from source system as-is Cloud (Azure Data lake Storage)	Source data parsed and normalised Cloud (Azure Data lake Storage)	Source data parsed and normalised Cloud (Azure Data lake Storage)	Innovaccer Analytical Views Cloud (Distributed Relational Storage)	FHIR Data Model Cloud (Distributed Relational Application Storage)	Analytics Model for Business Intelligence Cloud (MOLAP & ROLAP)
					

Fig 2. Innovaccer UDM data levels overview.

Extensive Support for Business and Analytics Use Cases

UDM supports an extensive range of business and analytics use cases for payer, provider, and life science markets. Key examples are summarized below.

Market-agnostic Use Cases

Real-time Analytics

Use cases such as Web Interface Reporting and global EMPI require real-time AI/ML analytics at the patient / member-level, and are supported with sub-second SLAs. This applies more broadly to use cases where information must be presented in real-time, especially where underlying analytics and business logic define for whom and what is presented. UDM supports real-time analytics use cases with very high concurrency.

Data Warehousing and Marts

UDM contains comprehensive data designs for creating atomic and dimensional enterprise data warehouses with a central repository. Data organized around business entities with multiple levels with members, providers, organizations, ACOs, and geographical regions are more useful for analysis than data committed to applications that support vertical functions of the business.

Payer Use Cases

CMS Compliance

The UDM is foundational to supporting the payer regulatory API requirements per the CMS interoperability and patient access final rule (CMS 9115-F). The UDM enables sourcing of necessary data elements, mapping to APIs to be exposed, and management of user-level access to the data. Innovaccer's UDM supports compliance with patient access API, provider directory API, payer-to-payer data exchange, and federal-state data exchange requirements.

True Interoperability Engine

UDM supports a canonical data model and HL7 semantics out of the box and simplifies interoperability across systems, enabling direct EMR integration and customized workflows for payers. Innovaccer's essentially 'built-for-healthcare' UDM has integrated FHIR capabilities and support for other healthcare standards like HL7, CDA, X12, and others. This enables seamless cross-system connectivity (e.g., care management system and automated member outreach), as well as unified member-level records (e.g., claims, EMR, labs, pharmacy data). The Innovaccer UDM enables both a high-performing payer healthcare data platform and a robust application suite to support true interoperability with absolute ease.

Provider Use Cases

Population Health Management

A foundational use case of the UDM is to drive Population Health Management (PHM) outcomes with data aggregated from disparate data sources (e.g., electronic medical records, claims, pharmacies, labs). This leverages data cleaning, standardization, and enterprise master patient index (EMPI) capabilities, and enables clinicians to drive cost and quality outcomes at the individual patient level, from physicians and medical assistants at the point of care to care managers' care protocols. The aggregated data also enables actionable population-level analytics, with insights on specific patient cohorts directly linked to interventions that can be initiated for those sub-populations (e.g., targeted outreach campaign, new care protocol with automated referrals).

Archival Storage with Random Access

UDM supports SQL access on top of a distributed file system to provide a cost-effective option for archival storage. This is effective for real-time patient access and supports FHIR APIs for applications, such as Innovaccer's Patient-360 for full patient-level data access.

Life Sciences Use Cases

Clinical Data & Analytics

The UDM unlocks real world data (RWD)-driven needs for life sciences organizations. Specifically, de-identified, patient-level, longitudinal data covering the full spectrum of care can be utilized for therapeutic research and development. In most if not all cases, data sets with this level of breadth and depth requires aggregation across multiple sources (e.g., claims, EMR clinical data) and enrichment, as well as ongoing data feeds for iterative updates. Paired real world evidence (RWE) needs are likewise addressed with accompanying advanced statistical analyses, dynamic dashboards, and analytics customizations, again leveraging the UDM.

Clinical Trials Management

Streamlining clinical trials performance is a core focus for many pharmaceutical companies and contract research organizations (CROs) alike. By leveraging UDM-unlocked, aggregated patient-level clinical and trial operations data, key business needs can be addressed:

- **Protocol feasibility & optimization:** analyze potential patient enrollment counts at target sites, and identify patient population cohorts
- **Recruitment:** support site recruitment of patients who meet clinical trial criteria so study is sufficiently powered
- **Tracking:** track trial progress vs. plan, and support operations to improve performance over time

About Innovaccer

Innovaccer, Inc. is a leading San Francisco-based healthcare technology company committed to making a powerful and enduring difference in the way care is delivered. The company leverages artificial intelligence and analytics to automate routine workflows and reduce manual overhead to facilitate more person-centered care. Its KLAS-recognized products have been deployed all over the U.S. across more than 1,000 locations, enabling more than 25,000 providers to transform care delivery and work collaboratively with payers. Innovaccer's FHIR®-enabled Data Activation Platform has been successfully implemented with healthcare institutions, private health plans, and government organizations. By using the connected care framework, Innovaccer has unified more than 3.8 million patient records and generated more than \$400M in savings.

For more information, please visit innovaccer.com.



535 Mission Street Floor 18th, Suite 1829
San Francisco, CA 94105
innovaccer.com

☎ +1 415 504 3851

✉ team@innovaccer.com

Copyright © May 2020
Innovaccer Inc. All Rights Reserved