MarkLogic

RESPONSE TO WEST VIRGINIA BUREAU FOR MEDICAL SERVICES MES STRATEGIC PLANNING RFI BMS2200000001

GARY KATZ • SR. ACCOUNT EXECUTIVE, STATE AND LOCAL HHS • JANUARY 2022

EMAIL: GARY.KATZ@MARKLOGIC.COM
PHONE: 215-840-8371

To meet the increasing demands of Human Services delivery, Medicaid Enterprise Systems must quickly adapt to the volume, variety, velocity, and complexity of health data while also supporting modularity requirements. A Data Hub approach helps State Medicaid Agencies accomplish this quickly via a simplified architecture and without sacrificing enterprise features such as security, scalability, and searchability.

MarkLogic

MarkLogic Corporation
999 Skyway Road, Suite 200
San Carlos, CA 94070

January 10, 2022

State of West Virginia
Bureau of Medical Services
Attention: Crystal Hustead (crystal.g.hustead@wv.gov)
Fax: 304-558-3970
2019 Washington Street East
Charleston, WV 25305

Re: MarkLogic's Response to Request for Information #105180

Dear Ms. Hustead,

MarkLogic Corporation is pleased to provide this response to the State of West Virginia Bureau of Medical Services (BMS) Medicaid Enterprise Systems (MES) Strategic Planning Request for Information (RFI). This response details how MarkLogic's Medicaid Data Hub can address the BMS' requirements, allowing you to modernize and securely deploy new applications and modules with reduced cost and shortened time-to-value. Our team has fully considered the BMS requirements outlined in the RFI and has developed a response that describes not only how today's requirements can be met, but how the power of a flexible approach can help address inevitable future requirement changes. MarkLogic's role is as a data management technology partner to our state customers. Thus, our answers to RFI questions focus primarily on those pertaining to that aspect of MES environments.

MarkLogic is a powerful data platform that is designed to rapidly ingest data in any form and from any source, help de-duplicate and improve the quality of the data, and make that data available to users and systems via modern web services. MarkLogic offers a Medicaid Data Hub that combines an Operational Data Store (ODS) with built-in master indices that will enable BMS to get a common view of data coming from multiple sources. You will be able to provide a Medicaid Data Hub platform with rich enterprise features and a technology roadmap to reduce the infrastructure, non-recurring engineering, and operations and maintenance costs across the entire portfolio of systems – all while satisfying MITA tenants. More and more enterprises are leveraging the power of our data platform to address their Health and Humans Services (HHS) requirements, including providing immediate value, getting results faster and future-proofing their MES environments from unforeseen additions or changes in data.

We are enthusiastically prepared to discuss our solution and demonstrate MarkLogic capabilities that can further BMS' mission. I am the account executive dedicated to the State of West Virginia, please feel free to contact me at 215-840-8371 or at Gary.Katz@marklogic.com with any questions or communications that will assist BMS in the evaluation of our response.

We look forward to your review and the opportunity to discuss with BMS in more detail.

Sincerely,

Gary Katz
Sr. Account Executive, US State and Local HHS, MarkLogic Corporation

Contents

SECTION 4.2 QUESTIONS

ID	Response
4.2.1	Please describe any elements BMS should consider incorporating into its vision, planning, and implementation for a modernized, modular MES.

The road from moving from a legacy Medicaid Management Information System (MMIS) to one that embraces CMS' modular architecture is paved with the hopes of providing more agile and flexible capabilities, but it is also paved with challenges and risks that must be properly mitigated.

One of the major challenges of modularity stems from data and its management, which is the true lifeblood of MMIS systems. Specifically, as new modules are being developed, how can the State ensure that the data from existing silos can be quickly integrated into the new modules, and be confident that the data is of high quality? Similarly, once the data exists in its new module, how can the State ensure key parts of the data can be easily shared among the different modules while providing a 360° view of the key entities to ensure effective service delivery and oversight?

Traditional solutions to solve these problems have utilized legacy technologies. However, these approaches now impose severe limitations to MMIS modernizations efforts because they were simply not designed to easily integrate data from many different sources, nor do they have the agility to adjust to evolving data requirements.

One of the primary issues with the legacy approaches is the upfront data modeling required in order to acquire new data sources. A precise definition and unanimous agreement on every field and relationship – before users get access to any of it – must be established before the data can be used. This process is an expensive, time-consuming task requiring an intimate knowledge of all data sources and data usage patterns or analytics coupled with detailed ETL mapping, coding, and testing for precision. This imposes significant limitations when tuning mastering algorithms by prolonging data discovery, and instead refocusing most of the effort on how to persist the data rather than how to exploit it. Furthermore, the query patterns need to be known in advance in order to create appropriate indexes to deliver acceptable performance. This approach, depicted below, requires a lengthy waterfall-type development process, during which time, the original requirements or formats often change. Legacy approaches therefore assume a static environment where data sources and access patterns do not change over time. This process is severely flawed and too difficult for any complex data, including healthcare data.
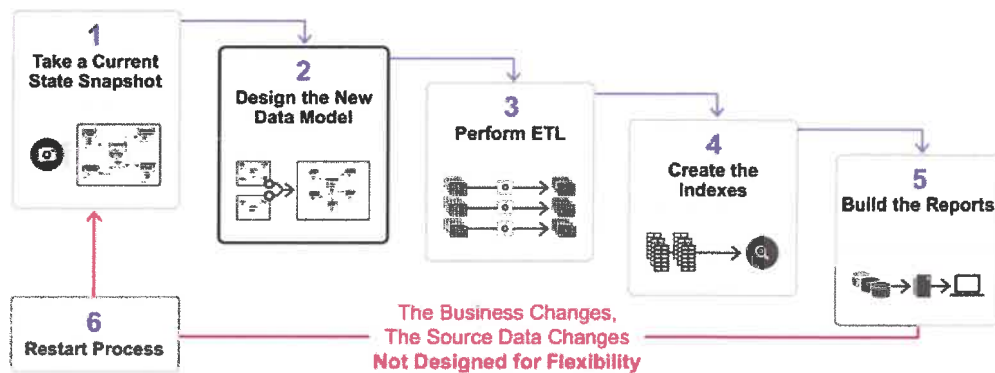
To make matters worse, most of today's data sources in healthcare involve semi-structured, poly-structured, or unstructured data (i.e. text, PDFs, images, and multimedia). Using these other data sources requires adding newer approaches and technologies, (e.g., data lakes), in order to hold or file the data for eventual use, plus various other data technologies and supporting databases in order to make the data useful or to allow access for comparison to other data stores. This results in significant architectural complexity.

To overcome these challenges, MarkLogic recommends leveraging an architectural pattern that enables a consistent and flexible data management strategy and that employs an underlying technology with a proven track record of adapting to changes in data and query requirements—**a true Data-First approach**. By leveraging next-gen technology to quickly keep pace with changes to data, many scheduling and programmatic challenges can be quickly overcome.

MarkLogic is the provider of the MarkLogic Data Hub for Medicaid which enables this architectural and data management flexibility. The MarkLogic Data Hub for Medicaid combines an Operational Data Store, Master Patient/Provider Index, and analytical capabilities all in one consolidated platform. It offers a unique, fast, incremental, and secure approach to get complete visibility |

across your data silos and perform "Smart Mastering" of data to improve data quality. Smart Mastering is MarkLogic's flexible approach to master data management and disambiguation of duplicate or conflicting data. The proliferation of function- specific systems supporting Health and Human Services (HHS) over the years has made it increasingly difficult to get a consistent view of often conflicting or duplicate data across applications. MarkLogic enables HHS agencies to use all of their data to deliver more comprehensive, accurate, secure, and reliable data to both internal and external stakeholders with a low risk, modular approach.

By ingesting data as is – without upfront ETL – data comes in faster and goes out faster. The "load as is" pattern supported by incremental data quality and enrichment to the data lets you feed operational transactional applications in real time, support analysis directly, and deliver data through bulk output to downstream systems. This enhanced, enriched, and higher quality data is then available for use in the mastering algorithms.

MarkLogic addresses data integration directly inside an enterprise database – with security, governance, indexing and fast, transactional access – centralizing it without demanding a complex data model to be built up-front. At times, this can seem counterintuitive -- why put the data in a database in order to integrate it? The answer to this is a natural evolution of how MarkLogic customers frequently use our platform: they have found that it is far easier and faster to integrate data by leveraging a technology that allows it to be loaded as-is, and then queried as part of an iterative discovery process using enhanced search to determine how best to leverage and exploit it. In this regard, MarkLogic is as much the integration platform for fusing the data, just as it can be the database that is queried for business purposes.
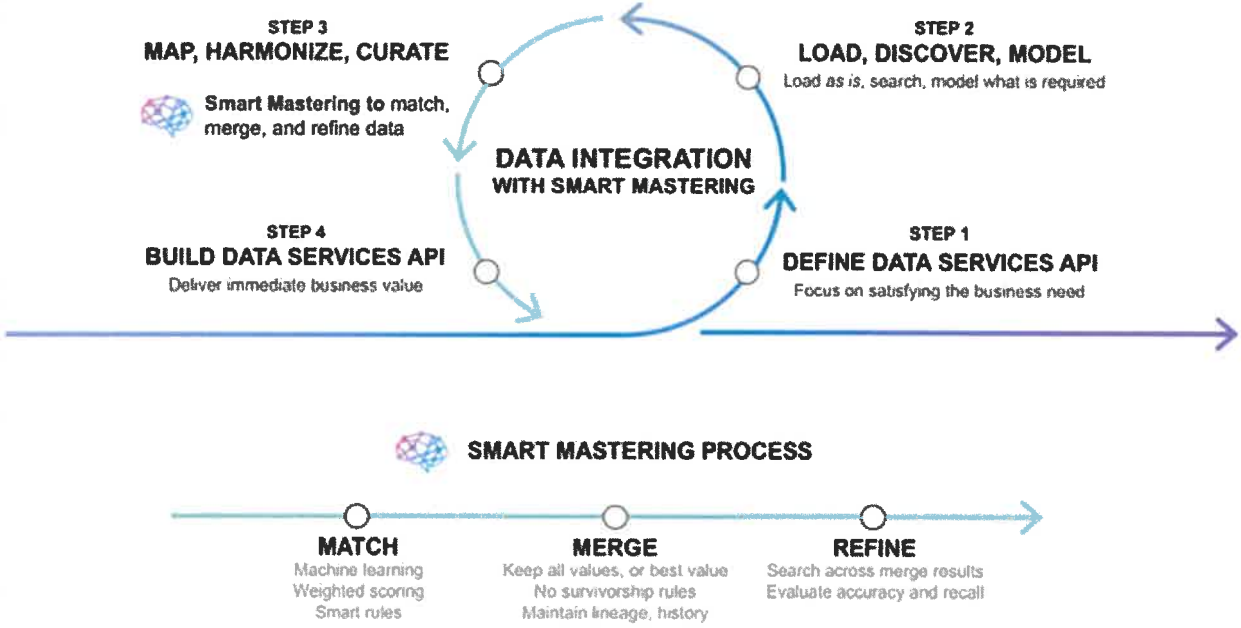

| 4.2.2	In the projects you have been on, what was the optimal configuration of MES modules specific to functionality, integration of other solutions, and management of data? MarkLogic recommends taking a Data Hub approach to configuring MES systems. Focusing on getting the data "right" at the outset enables optimal module configuration by allowing modules to better leverage consistent, shared data and thus reducing redundancy within module-specific implementations. This is accomplished by leveraging an Operational Data Store (ODS)-first implementation, often performed during the Systems Integration effort, which lays the foundation of the data management activities into one centralized repository that serves out data to the modules. An example of this architecture and how modules would interface with it can be seen below.
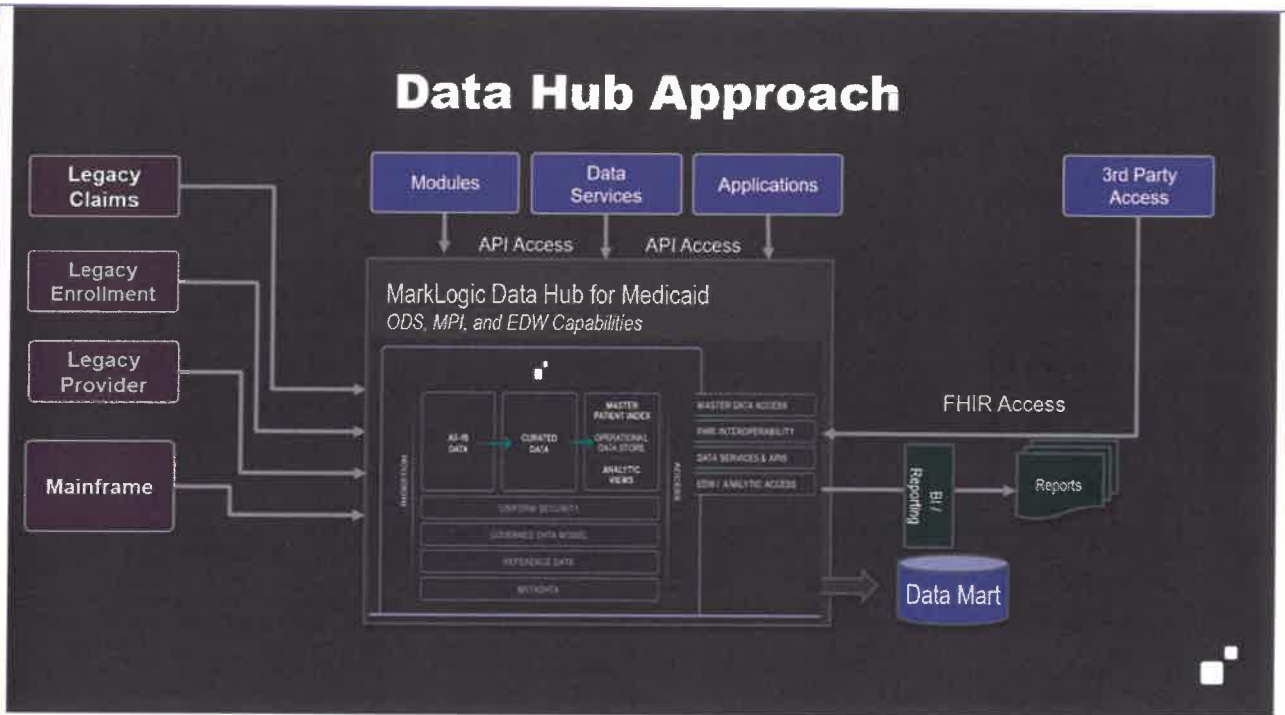
Without leveraging a Data Hub approach, the architecture and systems integration required to support efficient data management for module consumption in MES environments is considerably more complex, brittle, and risk-prone, as seen in the image below:
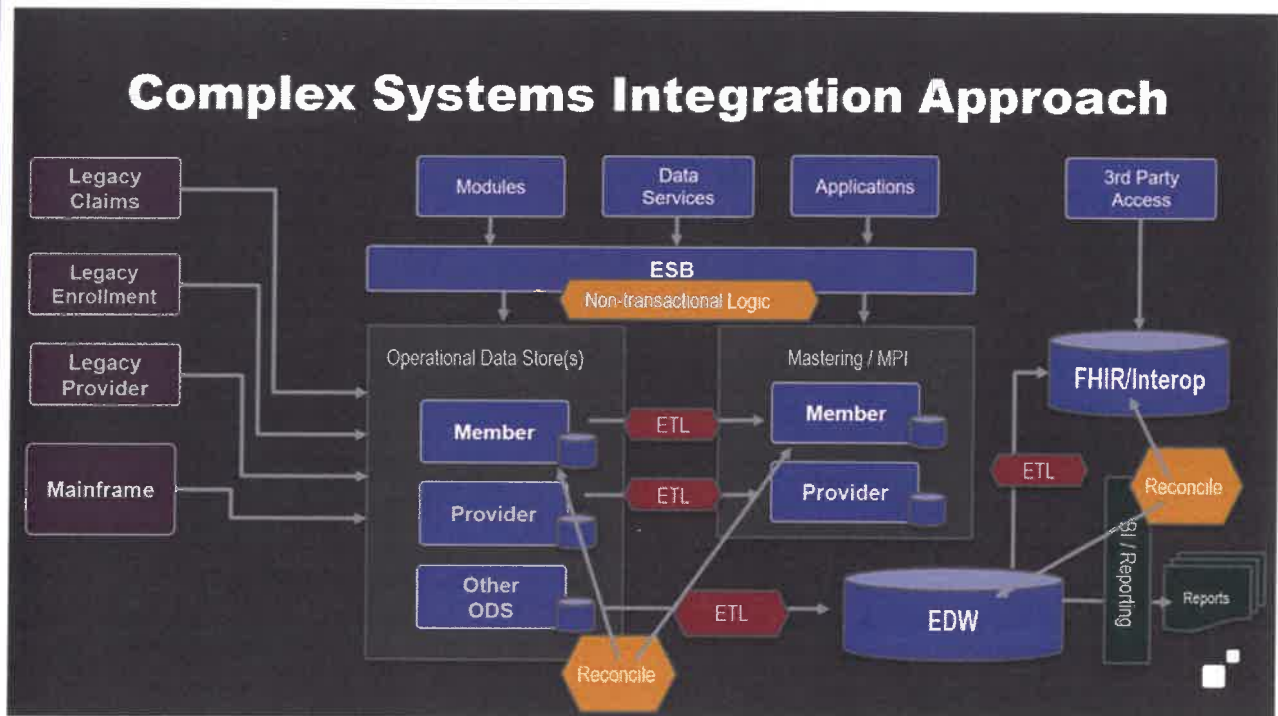
4.2.3	Describe Medicaid Enterprise solutions your organization provides or is developing that BMS should consider during its roadmap planning. BMS is interested in learning about the following: 1. The Medicaid Enterprise business processes or discrete functionalities targeted by the Medicaid Enterprise solution. 2. How the Medicaid Enterprise solution is packaged (i.e., commercial-off-theshelf (COTS) or proprietary; modular or tightly integrated; cloud or local). 3. How the Medicaid Enterprise solution is priced (please include methodology only, e.g., Per Member per Month, fixed price per year, data usage—please do not

provide actual purchase prices). 4. In how many states is your Medicaid Enterprise solution currently deployed, or expected to be deployed, and how long has it been in use. 5. Configurations and customizations typically requested to adapt the product for use in a State Medicaid Program. 6. Technical architecture and processing capacity/scalability. 7. User-facing and self-service capabilities. 8. Interface support, flexibility, and extensibility to other stakeholders and State agencies.

MarkLogic Corporation provides a Medicaid Data Hub that enables States to gain a 360 degree view of members, providers, and claims across their Medicaid Enterprise System (MES). MarkLogic's Medicaid Data Hub combines and Operational Data Store (ODS) with built in Master Indices that consolidates and simplifies MES infrastructure, thus lowering risk to and enabling modules to be more quickly developed. The MarkLogic Medicaid Data Hub is COTS technology that supports modularity and can be supported in local, cloud, or hybrid environments. The solution is priced by volume of cores (measured in 8-core packs per unit) necessary to support ingestion, data volume, and query-speed performance.
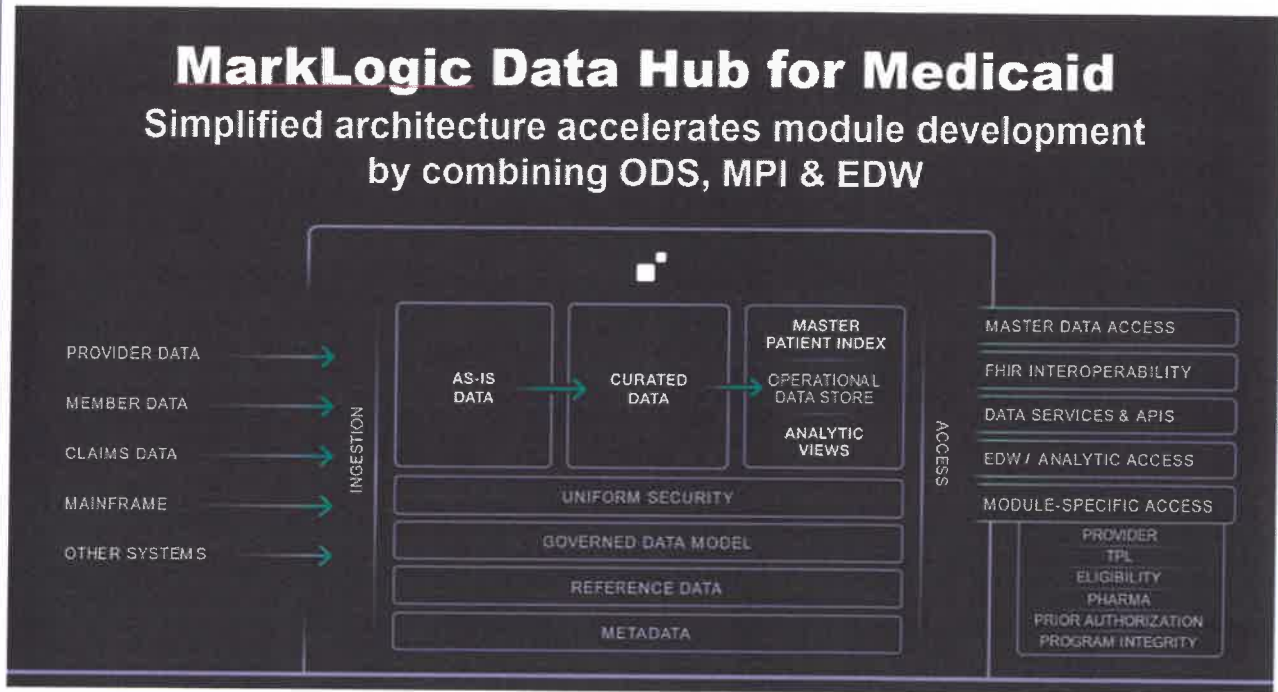

The MarkLogic Medicaid Data Hub is built atop our industry-leading Data Hub platform. The Data Hub provides more capabilities than any traditional ODS – including built-in master data indices, flexible query, text search and security. It integrates data from any legacy system or new module with ease and reduces the risks and costs of MMIS modernizations. As a proven ODS approach in use in multiple States, we help ensure that State Medicaid agencies are HIPAA, MITA and MECT-compliant, while providing a ODS platform that accommodates change and positions States for success in the out years, rather than being locked into a particular Module or technology stack. Commercially, MarkLogic's Data Hub platform is used to support some of the most mission critical applications in highly regulated industries, such Healthcare, Insurance, Intelligence and Defense, Financial Services, as well as Entertainment, Publishing, Energy and Manufacturing. Most visibly, MarkLogic has been the back-end of the Data Services Hub for HealthCare.gov, having supported our 8th successful Open Enrollment since the program's inception. The system scales to tremendous data and query load volumes, handling 70,000 active users at peak times and greater than 6500 complex transactions per second.

Proven MMIS Modernization Success

Around the US, states are successfully transforming Medicaid services with MarkLogic.

- **MMIS replacement to improve health services for 800,000 beneficiaries**: A southwestern US state Human Services Department (HSD) selected MarkLogic to drive transformation in the data processing core of their MES systems.
- **Improved Medicaid eligibility and claims processing:** In just 120-days, this southeastern state's Department of Health and Human Services (DHHS) configured the core of their Medicaid Enterprise System (MES) to streamline processing and serve as the foundation for a series of new modules that have since been built out and integrated, including administrative services (AS), claims processing, third party liability and others.
- **Modernizing MMIS in the cloud:** A state department of community health is building their future MMIS using MarkLogic as the ODS and Master Data Management solution on the AWS Cloud. MarkLogic was chosen due to its proven success with other state health departments, robust product features and strong solutions engineering

capabilities. The first phase involved data consolidation and analysis, and future plans include extending the Medicaid data to support case management and a variety of human services.

The amount of configuration/customization required to support State Medicaid Programs is different for each State, as they each have their own architectural considerations and specific use cases to support. MarkLogic provides complete APIs to facilitate ease of integration into existing MES environments, which is often done via REST.
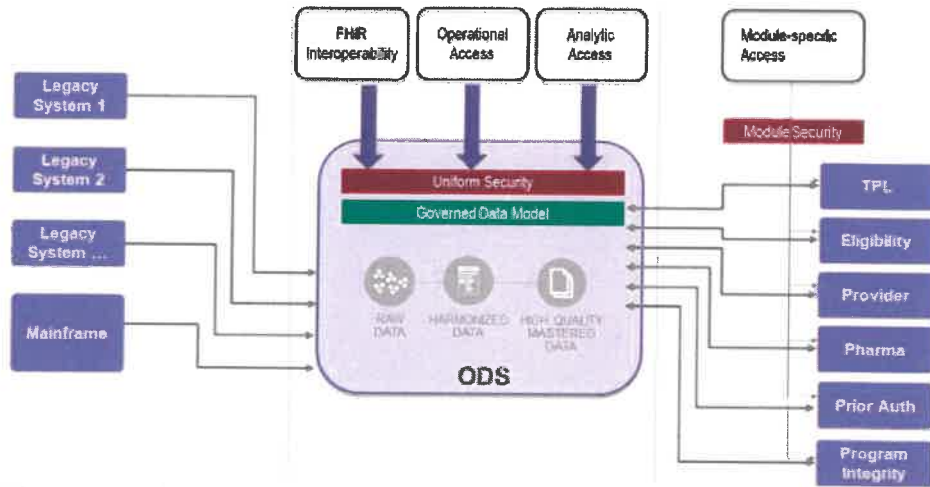

MMIS Modernization Reference ODS Architecture

MarkLogic's Medicaid Data Hub can sit behind the scenes serving as the "engine" for data processing. This can be managed through Data Hub Central interface, seen below.
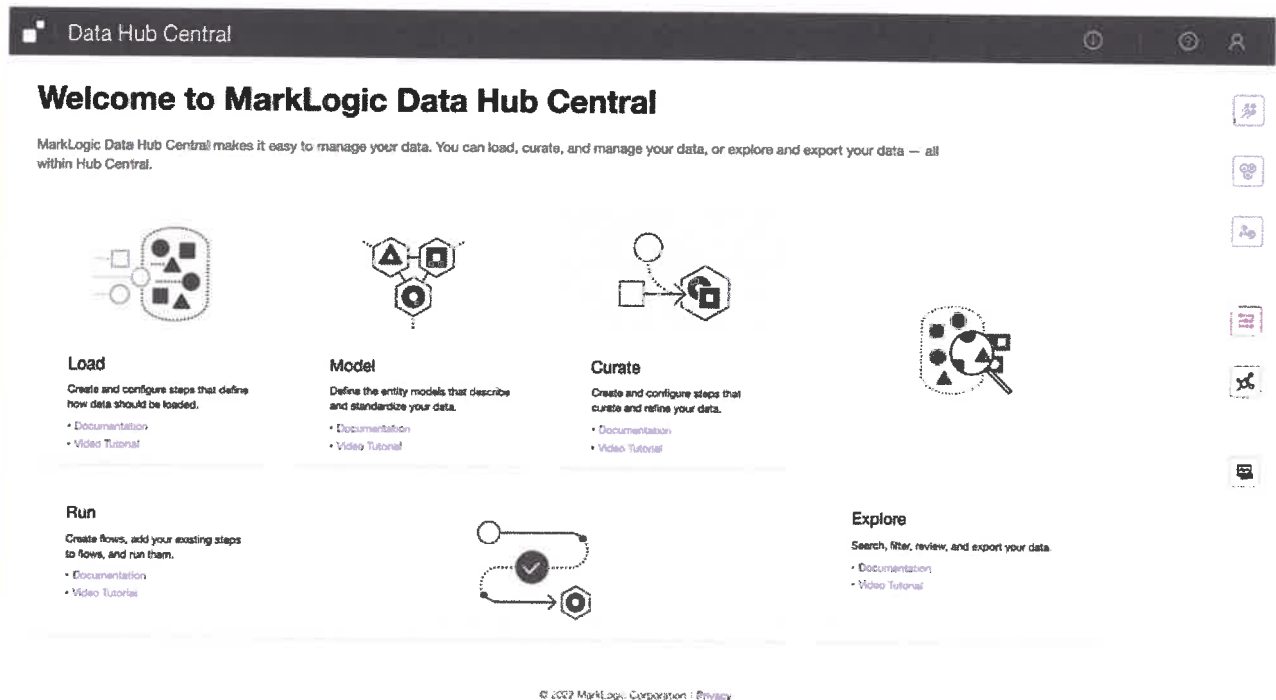
Once data is brought into the Medicaid Data Hub, it can then be made available to third party applications through standard interfaces such as analytic and BI reporting tools:
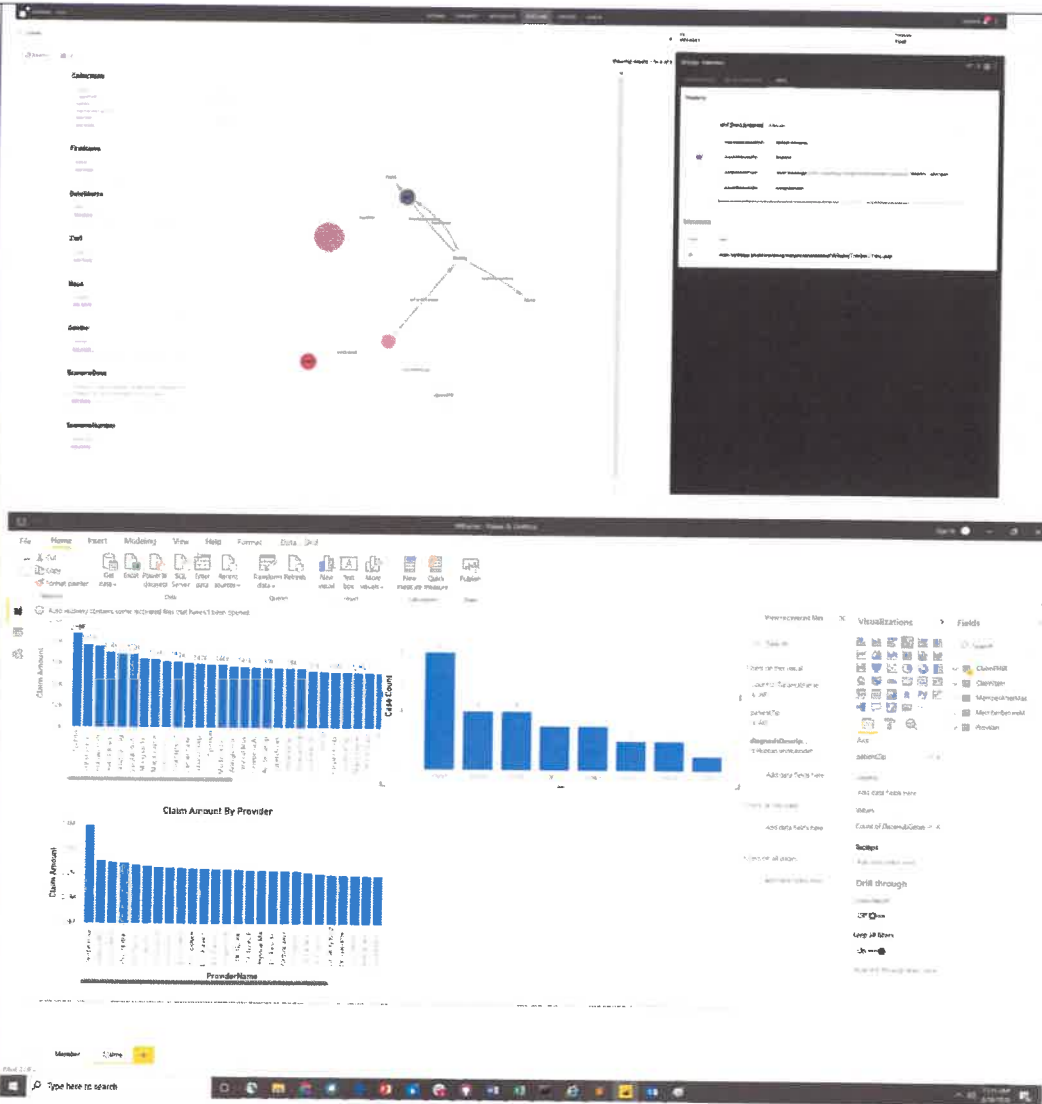
## Technical Architecture for Processing/Scalability:

MarkLogic leverages a shared-nothing distributed architecture for its data hub. To achieve speed and scale beyond the capabilities of one server, MarkLogic clusters across commodity hardware connected on a LAN. Every host in the cluster runs the same MarkLogic process, but there are two specialized roles. Some hosts (Data Managers, or D-nodes) manage a subset of data. D-nodes contain the transaction layer for node, as well as journals, caches and indices. Other hosts (Evaluators, or E-nodes) handle incoming user queries and internally federate across the D-nodes to access the data. They contain their own caches, mechanisms to broadcast and aggregate query results, evaluators specific to the type of query being processed, as well as communication interfaces.

A load balancer spreads queries across E-nodes. As you load more data, you add more D-nodes. As your user load increases, you add more E-nodes. In some cluster architecture designs, the same host may act as both a D-node and an E-node. In a single-host environment that's always the case. Clustering enables high availability. In the event an E-node should fail, there's no host specific state to lose, just the in-process requests (that can be retried), and the load balancer can route traffic to the remaining E-nodes. Should a D-node fail, that subset of the data needs to be brought online by another D-node. You can do this by using either a clustered filesystem (allowing another D-node to directly

access the failed D-node's storage and replay its journals) or intra-cluster data replication (replicating updates across multiple D-node disks, providing, in essence, a live backup). Offering maximum fault tolerance and flexibility in our architecture.

A MarkLogic cluster does not differ from a software perspective based upon its deployment into an on-prem environment or a cloud. In fact, MarkLogic is both cloud-neutral and platform-agnostic, meaning the software can be deployed in any cloud environment. This will help BMS avoid cloud vendor lock-in and future-proof itself from changes to cloud strategy.
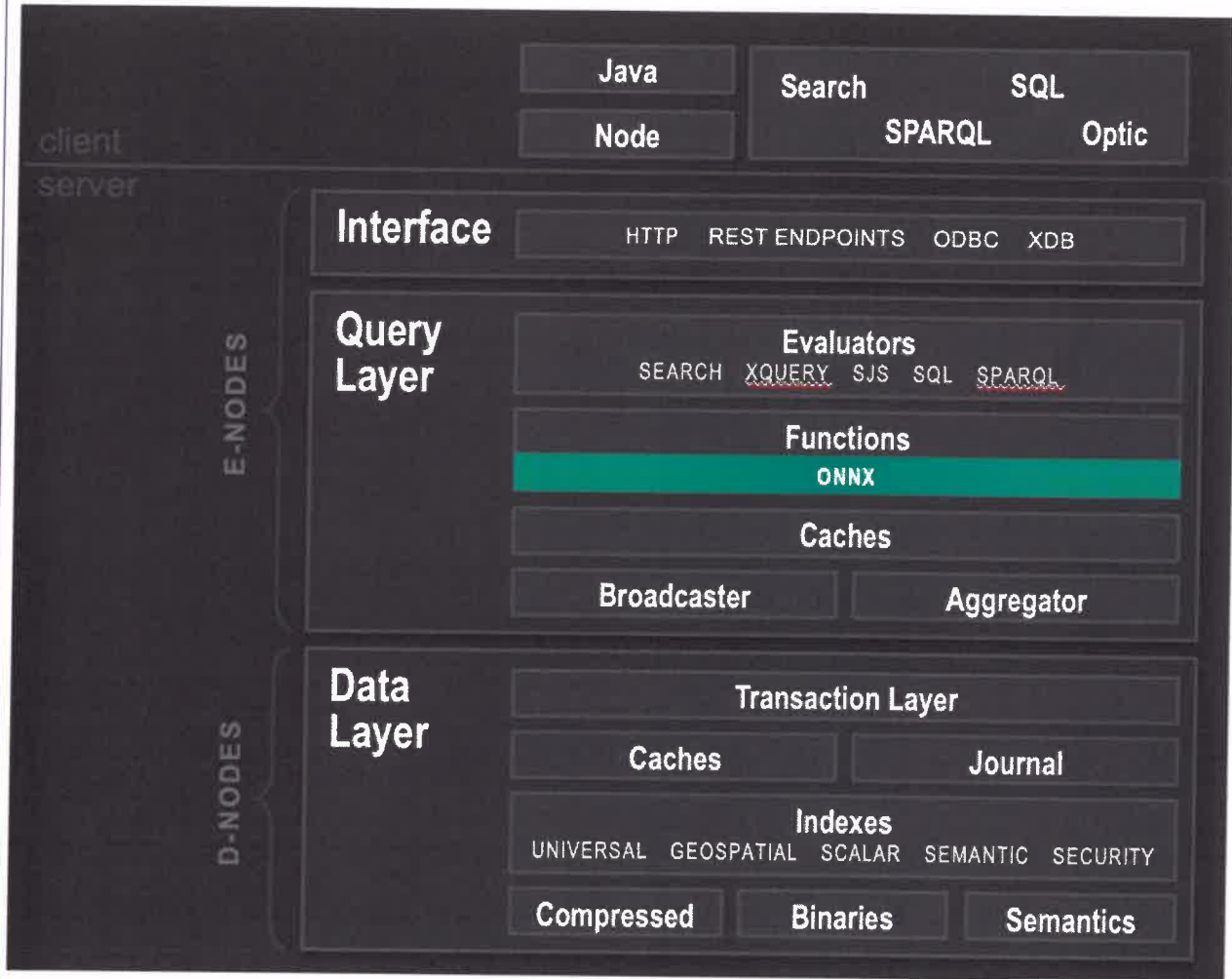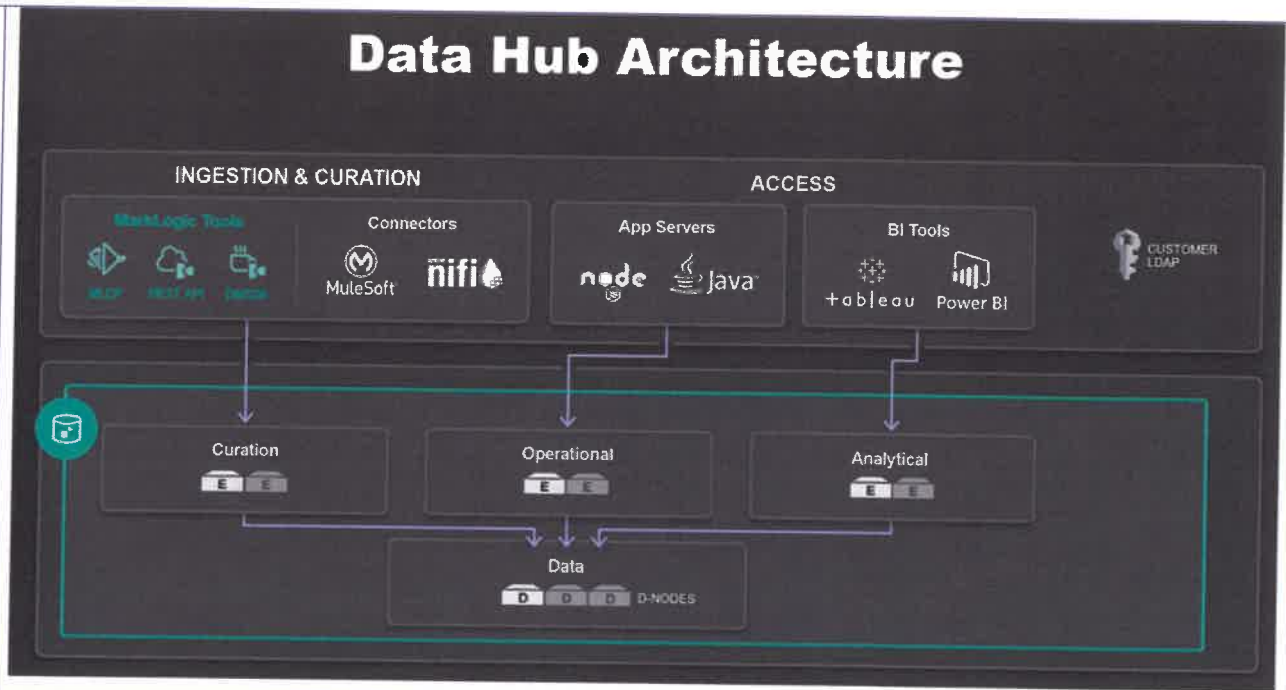
| 4.2.7 | Please provide your recommended strategy for ongoing compliance with the CMS Interoperability and Patient Access final rule (CMS-9115-F). The rule can be found at the following location: https://www.cms.gov/files/document/cms-9115-f.pdf. |

Using MarkLogic as the Data Hub platform storage mechanism is recommended in part because it simplifies implementing the Interoperability rule. MarkLogic has a "FHIR Mapping Medicaid Accelerator" component which allows users to convert data in a MarkLogic Medicaid Data Hub to compliant FHIR messages. This innovative approach means that the Data Hub can directly serve FHIR-based, compliant data without copying or transforming into another data store, reducing both cost and complexity.

This "Accelerator" includes a range of key resources including Member, Provider and Claim with sample mappings, and these are integrated into the HAPI FHIR server. The popular HAPI FHIR server is a FOSS capability that implements the Interoperability Rule, but requires an underlying data store. This innovative approach allows the Data Hub in a Medicaid environment to directly provide data to HAPI FHIR. Notably, the MarkLogic FHIR Mapping Medicaid Accelerator uses HAPI as it's default FHIR server. However, the Accelerator is available as open source software and can be updated to work with your preferred FHIR server, as well.

A mapping GUI is provided to enable a business analyst or data expert (vs a developer) to map all non-included FHIR resource types as needed quickly and easily. Interoperability requires certain resources are mapped if the State has data for them, and each State will want to customize the exact resource set. This "MarkLogic FHIR Mapping Medicaid Accelerator" is a free, open-source tool for interoperability, but is built specifically to work with the MarkLogic Medicaid Data Hub technology stack.

MarkLogic is also unique is addressing the broad range of search capabilities required or supported by FHIR specification. This allows extension to capabilities that are supported within FHIR, but are not required by the Interoperability Rule (at least not required at this time). They include:

- Text search is included in MarkLogic, including: exact, starts-with, wildcarded, and match-anywhere. Case-sensitivity, diacritic markings and even sounds-alike or nickname-based search is also supported. Name search in particular requires these search capabilities.
- Structured record query is supported in MarkLogic since MarkLogic is a NoSQL database that natively queries JSON and XML data. Most FHIR query capabilities are structured queries of this type.
- Ontology-based lookup and search is supported, where any condition may be queried based on super- or sub-categories of a known ontology like SNOMED. In this way, a condition or medical situation can be queried and related categories or sub-categories will automatically be included. This is optional in the FHIR spec. MarkLogic supports SPARQL 1.1 W3C standard for querying ontology or RDF data.
- Versioned and historic data is supported in MarkLogic using MarkLogic's temporal features. This is an optional feature in the FHIR spec, but very useful.
- Geospatial proximity searches are supported in MarkLogic, and are particularly useful for Provider data, which is required to be shared by the CMS Interoperability rule. MarkLogic supports proximity searches using distance to a

point, inside a bounding box, or even within a complex polygon such as a county or zip code. These may be useful for finding providers in a service area or close to a member.

- All above searches can be combined seamlessly in MarkLogic, making it an ideal fit for this requirement.

The FHIR data standard is often referenced with regards to claims processing: We believe it is now time to start to use FHIR for other purposes, including claims adjudication. The CMS Interoperability rule serves multiple purposes, not least of which is to drive maturity and widespread adoption of FHIR as a data standard. CMS is putting substantial resources behind maturing and driving adoption, so we believe that being compatible will yield more and more benefits in the future as FHIR becomes ubiquitous.

Currently, FHIR may not support all data required for payer processing, and may require extensions or changes in some areas. In the future, FHIR may evolve to be more payer-centric (vs patient and health information centric). Even if changes are needed, the core structures and idioms need not change (codable concepts to adapt to multiple code systems, common address structures, duration handling etc..). Thus investments in FHIR to support interoperability will avoid wholesale rework of the core enterprise data model in the future.

In any case FHIR v4 is the current standard and should be embraced. As noted above, this suggests a NoSQL JSON-based data strategy, and a NoSQL data storage approach. Note that MarkLogic data lenses (via the Template Drive Extraction feature that creates SQL views over JSON) allow projections from FHIR data (or any JSON data) to relational views, and in turn this allows both JSON-based API access and SQL-based analytic or operational reporting access from the same Data Hub database. These mappings may be standardized or shared across States if both MarkLogic and FHIR-based processing are widely adopted as this RFI response recommends.

Put another way, MarkLogic is betting on FHIR adoption, and building enabling technology that makes many things easier when a FHIR data model is used. Everything from lists of PII fields to data interchange APIs to reports and queries can be better standardized as States and others adopt MarkLogic as a storage technology together with FHIR as a standard model.

4.2.8	Provide your strategy for compliance with the Health Insurance Portability and Accountability Act (HIPAA) and Federal Risk and Authorization Management Program (FedRAMP) Requirements. Information about HIPAA compliance can be found at the following location: https://www.hhs.gov/hipaa/for-professionals/privacy/index.html. Information about FedRAMP can be found on www.fedramp.gov. MarkLogic is cloud agnostic and can be deployed onto most any FedRAMP certified cloud service provider. Our most common deployments are into Azure and AWS. MarkLogic includes the security features to support HIPAA and FedRamp requirements such as element-level security, redaction, and database encryption at rest with external key management. Security is provided through role-based access control (RBAC) along with a "compartment security" option. MarkLogic's government-grade security has been hardened through many DoD and Intelligence Community projects. The product has completed multiple vulnerability assessments and been validated by the National Information Assurance Partnership (NIAP) as complying with the Common Criteria DBMS profile at Evaluation Assurance Level 3 (EAL3) augmented with ALC_FLR.3 (highest level of flaw remediation). It has also been certified and accredited to protection level 3 (PL/3) by numerous Intelligence Community and Dept. of Defense customers, with authority to operate (ATO) in place on NIPR, SPIR, JWICS, and several other classified networks. Many MarkLogic applications have received ATO through DITSCAP and DCID 6/3 C&A process. DISA operates MarkLogic systems that have passed Mission Assurance Category (MAC) II and III. The product has received Common Criteria certification in its most recent release through the National Information Assurance Partnership (NIAP) – one more proof-point that information will be secure when stored in the MarkLogic Data Hub for Medicaid and you can ensure HIPAA compliance. It is also worth nothing that MarkLogic serves as the backend of the HealthCare.gov Data Services Hub and Marketplace, which also now sit in AWS infrastructure and must comply with HIPAA, FedRAMP,and Federal Tax Information (FTI) security requirements.

The Most Secure Data Hub. Period.

- Secure by default
- Advanced encryption
- Fine-grained, role-based security
- Data anonymization and redaction
- Best-in-class data security across the integration lifecycle

MORE SHARING WITH LESS RISK

Sample Clients

HealthCare.gov   aetna   OPTUM   J.P.Morgan   CREDIT SUISSE

MarkLogic

HealthCare.gov

Providing health insurance for millions of citizens

- Marketplace and Data Services Hub with changing requirements that Oracle and ETL failed to handle
- 280,000 concurrent users, 6,500 transactions/second
- 10+ million citizens enrolled, met congressional deadline

"The largest personal data integration government project in the history of the Republic."

Stephen Parente
PHD, MPH, MS

| 4.2.10 | What approaches do you suggest for Disaster Recovery processes in a modular MES that accounts for integration and communication across multiple partners? |

From a data management standpoint, which is where the MarkLogic Data Hub for Medicaid is addressing, having full High Availability and Disaster Recovery capabilities is essential to ensuring continuity of operations in the event of catastrophic failure in a MES environment. If the data is preserved and available, it will be much easier to coordinate integration across multiple partners and modules, as data is the lifeblood of MMIS environments.

MarkLogic has enterprise-class high availability and disaster recovery (HA/DR) so that you can have confidence that your data is always available and scheduled downtime is minimized, reducing risk and avoiding interruptions. MarkLogic achieves HA/DR using a shared-nothing architecture that provides redundancy for failover and high-performance scaling.

- o Shared-nothing architecture has no master-slave relationships, eliminating any single point of failure
- o MarkLogic has point-in-time recovery and ACID transactions to ensure full redundancy and consistency
- o Changes do not require a server restart (re-indexing, adding nodes, or changing configurations)
- o Database replication between sites is secured with SSL out-of-the-box
- o Incremental backups consume less storage and can be completed quickly

High availability entails continuity within a cluster, protecting against component failure with computers, networking equipment, and power supplies. MarkLogic clusters are protected against these sorts of failures by providing fault tolerance in both evaluator nodes (e-nodes) and data nodes (D-nodes). if either an e-node or D-node fails, other nodes automatically pick up the workload so that the data stored in forests is always available. high availability works with either local disk failover on Das or ssD, or shared disk storage such as SAN, Amazon S3, or HDFS.

Disaster recovery protects against failures of an entire data center (power outages, natural disasters, etc.). With disaster recovery, you can backup selected components or the entire database using ssl out-of-the-box. You can also do incremental backups, coupled with journal archiving, in order to restore the database to a point-in-time that minimizes the recovery point objective (RPO) and uses less storage.

Avoiding scheduled downtime is important to ensure a highly available system, which is why Marklogic is designed so that administrators can keep systems online and operational while performing changes and maintenance:
- — **Online Database Backup Operations** - Full and consistent database backup operations run while the system is up, avoiding downtime to backup data
- — **Hot Configuration Changes** - Most configuration changes, including adding nodes, do not require a server restart
- — **Automatic Index Optimization** - On-disk data structures are designed for fast indexing without shutting down the system, and queries can even be continued while re-indexing

Flexible replication allows you to replicate a specific subset of your database to support information distribution. You can select specific portions of documents, or even perform data transformations during delivery to remote sites. This asynchronous, document-level replication is ideal when replicating data to thousands of edge nodes in remote areas of the world where internet access is poor, and mobility and security are at a premium.

4.2.14	What roles and responsibilities have you seen for a system integrator (SI) in a modular systems environment? Was this role fulfilled by a separate vendor, incorporated with other services, or performed by the state Medicaid agency itself? What are the key success factors and risks to success related to using a SI? MarkLogic has worked with both contracted Systems Integrators, as well as SMA's who serve this function themselves. If SMA's have depth of knowledge and technical skills in-house with dedicated development and PM teams, then surely SMAs can be successful serving as their own integrator. Where SMA's have been successful as their own integrators is where investments in providing technical staff can be made. By serving as its own SI, an SMA can have more control over technology choices and direction. Contracted Systems Integrators generally bring greater volumes of technical depth and staffing resources. Systems Integrators have generally been utilized to stand up the core infrastructure of Medicaid Modernization environments that can be leveraged by module developers. Having a Systems Integrator can be an advantage to SMA's by outsourcing the development tasks and holding these integrators accountable to schedule, cost, and delivery. However, SMA's should be careful when selecting SI's to ensure that the solutions they bring to bear are aligned with SMA objectives and provide requisite innovation and automation to help mitigate cost concerns. The use MarkLogic's Medicaid Data Hub helps to offset risks to SI contracts by providing a platform that contains the necessary application and enterprise capabilities that minimize risks associated with systems integration tasks. This allows for SI's to instead focus on data integration and supporting module development.
4.2.15	Describe your depth, breadth, and frequency recommendations for performing periodic vulnerability scans of production and development environments? MarkLogic does not perform vulnerability scans of production and development environments. However, our software has participated in such scans by other processes and these techniques can readily be applied.
4.2.16	What processes, techniques, and solutions does your organization consider critical for delivering optimal data sharing throughout the MES? It is critical for development teams within the Medicaid Data solution and other downstream systems such as modules to have access to data feeds that are as close to production data as possible. Only data that is required by the recipient should be provided. Techniques such as redaction, de-identification, and element-level security should be used to facilitate data sharing.

These techniques have allowed our customers to be able to provide data that has been cleansed of Protected Health Information (PHI) and Personally Identifiable Information (PII). The data can be replaced with valid values allowing for development teams to test with "production like" data without having to deal with PII and PHI.

When sharing actual production data it's important to ensure its security and to track what data was shared and to whom it was shared.

We have seen a number of approaches and techniques that facilitate data sharing:

- Use Data APIs (typically RESTful with JSON data payloads) to standardize data interchange, and decouple components. APIs are easily documented (e.g. using the OpenAPI spec), can be standardized across the enterprise, and can be versioned, discovered and managed by API gateways and other technology. As a standard technology they are familiar to developers and can usually be consumed by various components. Exposing data primarily via APIs is consistent with a services-based approach.
  - o Note that some systems will use only business APIs, with direct database access and tight coupling to underlying data formats and stored. This is not at all what we mean here – it is critical to have *data APIs* served directly from the Data Hub to avoid tight coupling to data formats and expensive relational-JSON mapping layers.
- "Data First" development: define data and implement the Data Hub aspects early. To standardize, it is important to define a good model that will work well for all data consumers. This is much easier if done early on, and not as part of (or even concurrent with) the first MMIS modules or systems that are integrated. Starting early allows the State to stand up a team to think through and standardize APIs and interchange formats while considering the big picture for Medicaid and State future needs. As noted elsewhere in this response, we support basing the enterprise standard model on FHIR. Data dictionaries, documentation and industry expertise standardized in FHIR can be leveraged even if the persistent model is slightly different from fully-compliant FHIR. Conversion to true FHIR for interoperability is then easy even as the persistent layer is customized for other needs.
  - o Early focus on data processing also provides data access and APIs early on to drive module development, and even to drive module evaluation during procurement.
- FHIR supports JSON, XML and RDF interchange, but we recommend using JSON, due to the ubiquity and simplicity of that format, to drive greater and easier sharing via RESTful APIs.
- We recommend direct NoSQL (non-relational) storage of the JSON messages in the Data Hub data store. MarkLogic is an ideal technology to use here, as it natively augments JSON data with RDF and SQL data for storage. RDF is useful because JSON data is tree-structured, and it must be augmented with "graph" linkages such as RDF provides – this helps with "graph-y" data such as provider networks, family structures and other relationships that are not purely tree-based. RDF is a W3C standard for rich graph relationship data. MarkLogic adds SQL projections to provide a convenient interface for typical reporting and BI tools, as well as some data interchange tools while still enjoying benefits of a NoSQL storage approach.
- Fresh, timely data improves data sharing by making data more actionable. Therefore, move processing toward real-time to the extent possible. MarkLogic's open-source "State Machine" processing framework is ideal for this, using data-oriented events and conditions to run data through required transforms, validations and modifications as it flows into and through a Data Hub.
- Combine operational and analytic systems into one system, to avoid complexity. Generally speaking, start by using ODS as the source to feed a specific analytic need (as opposed to wholesale deferring data management to EDW needs). MarkLogic has built-in technology to support this (Template Driven Extraction creates transactionally up-to-date SQL views of all JSON data such as FHIR records), and we find this greatly facilitates data access and sharing because:
  - o Combined analytic and operational data storage allows real-time reporting to be the default, rather than an add-on or proprietary approach.
  - o Combining them reduces barriers to access and delays by eliminating complex and slow re-modeling of the operational data into a dimensional EDW model for analytics. It also avoids the issues and delays inherent in data movement jobs, monitoring, DevOps provisioning. Also combining them avoids additional technology components, additional vendor management, additional monitoring, and additional security compliance overhead of a separate EDW component.
- Paradoxically, strong security enhances and enables data sharing. Therefore, build in security at the Data Hub and database level. The reason security enhances sharing is that a person cannot access a data store at all if that data store (database) contains un-secured data they should not have access to. (E.g. adding a bank account number to a database without RBAC or similar controls immediately means nobody without financial data access can use that database directly). To facilitate this, the Data Hub database should support
  - o Role-based security (RBAC)
  - o Attribute-level (field-level) security
  - o Query-based security (QBAC, or query-based access control, allowing complex conditions to enforce security)

	o Universal security enforcement regardless of access pattern. API-based access, file exports, and SQL views should all respect the same, universal, declarative security policy.
	
	• Build a capability to produce and manage De-identified data
	     o Much as how good security actually facilitates data sharing (see above), data sharing also sometimes requires de-identification of the data, per the HIPAA rule and guided by NIST 800-122.
	     o Use Limited Data Sets for testing and procurement. True "de-identification" per the HIPAA rule is often difficult for highly-dimensional data (because the combination of various data values can technically identify a person), and therefore the production of Limited Data Sets per the HIPAA rule is often needed. Limited Data Sets require removal or masking (or randomization or other obfuscation) of a set of clear identifiable information such as name, address, and ID numbers, but allow certain data such as dates and diagnoses to be left in place.
	     o Produce de-identified or Limited data sets with a particular "purpose" in mind – think "fit for purpose" when sharing data. How much, and how fully de-identified is required? One purpose – statistical analyses for population health – requires diagnoses and population demographics to remain intact, but MMIS module software testing requires completely different data to remain the same to exercise module business logic.
	     o Declarative configuration for de-identification and Limited Data Set production. We have found that declarative de-identification configurations are more governable because they are understandable by the data and security experts, who are often not coders, and declarative artifacts can be tracked and versioned more easily than scripts or routines written in python or Java code that implement de-identification via coding.
	
	• Robust, multi-model data storage and query. Consider all types of data access, including future access. Initial uses are often simple access using simple queries, but ultimately full-text search, graph analyses and binary handling are all likely to be needed. E.g. processing a claim may only require simple, structured data queries for enrollment status and third-party liability checks. However, finding a provider may require text name search together with geospatial proximity query and graph analysis to determine the participation of the providers via their networks. A "no wrong door" social services support capability requires a graph-based query of the family and non-family relationship structure. Handling all types of data in this way requires "multi-model" storage and processing, or requires that the systems integrator cobble together a collection of multiple databases of different types (document store, relational, graph database, search engine, binary data store, and geospatial query tool). Choosing a multi-model storage technology up front reduces cost in the out years and avoids building a brittle, non-portable architecture.

4.2.17	What standards and practices would you recommend with regards to key data governance, master data management, data stewardship, and data-sharing concerns? What approaches do you recommend for engaging business data owners separately from technical data system managers? Data Stewards need to have access to the data. Data stewards should be involved in the mapping of the data from the source systems to the canonical model. As well as the mastering of the data to create a 360 view of the data. Its recommended to have technology that gives the Data stewards the ability to own the modeling, mapping, and mastering of the data. Some mappings and mastering may require complex implementation beyond the skills set of the Data stewards. Data stewards should interact with the technical development team to handle these outlier cases. Data Stewards should not be required to ensure the systems are running. The systems should generally handle operations themselves. The technical data system managers should handle any case of manual intervention. The Medicaid Data solution has to deal with many different data types and processes. There will inevitably be different Data Stewards of the different data types and processes. It is recommended to allow the Data Stewards to only have access to the data that they are stewards over.
4.2.21	What factors (technologies, development methodologies, frameworks, etc.) would you recommend BMS require in an RFP in order to accelerate the DDI of MES modules? It is recommended to use Agile Methodology in the designing, developing, and implementing of the solution. It allows for adaptive designing, flexible development, and early delivery. A phased approach for the Medicaid Enterprise Systems (MES) is recommended which aligns well with Agile Methodology this will reduce risk by allowing key stakeholders to see the solution in action as its being developed. Technologies should be able to handle the high volume, velocity, and variety of the data. The data will be coming in many different formats and will need to be processed promptly. The Technologies should be able to handle the security requirements of the Medicaid Data solution. Such as encryption at rest and in transit. The Technologies should be able to implement the principle of least privilege and be able to only give out data that is required for the user or system that is requesting it. Technologies should be able to run on-prem and on the cloud and should be able to be integrated with continuous integration

and a continuous delivery pipeline. Technologies should be enterprise-ready and have features such as High availability and disaster recovery.

MarkLogic's core product provides the architecture to build the data hub and framework called MarkLogic Data Hub Framework (DHF) to accelerate the development of integration layer including the Smart Mastering (Match/Merge) tool and architecture for BMS solution. MarkLogic Enterprise NoSQL database provides Schema-less architecture to ingest the structured/ unstructured data AS-IS without any need of creating the schemas to ingest the data. This ingest As-Is feature in MarkLogic simplifies and accelerates the integration project to great extent (10-12 times faster) compared to traditional relational/ ETL based solution. MarkLogic DHF provides the framework architecture and tool to create the Entity Model for the target entities and ability to connect the entities in the context of business relationship between entities. And the entire process of Entity model is iterative and flexible. You don't need to model everything at once and allows flexibility to add/ changes to the model as the business needs changes. MarkLogic DHF provides the APIs (Java API, REST API, Optic API, SQL and SQL-Views) to consume the curated, mastered and harmonized data for BMS for Analytics and reporting and to provide data to any downstream applications and/or external systems as needed with built in role-based access control data security at element/ attribute level granularity and redaction (masking) capability for PII/PIH data. These APIs and built-in government grade data security, accelerates and simplifies the data consumption layer.

4.2.23	What recommendations do you have for establishing procurement and implementation timelines that help deliver value sooner, reduce risk, maximize Federal Financial Participation (FFP), and achieve Outcomes-Based Certification or Streamlined Modular Certification?

When establishing procurement there are many ways from relying on one system integrator (SI) to having different SIs or RFPs for each module. We have found States to be more successful with having different RFPs for each module and sometimes the SIs win multiple RFPs. When having different RFPs it's important to have a vehicle to source development teams that can implement connectors between the different modules. As well as setting up a center of excellence for the MES Project to reduce risk and ensure the technology is being implemented correctly.

Using Agile Methodology is recommended to lower defect rates, and decrease development times. The Agile Methodology allows for a phased approach in creating Medicaid Enterprise Systems (MES). The Medicaid Management Information System (MMIS) has many different processes within it. It is recommended to create a plan where implementation schedules and milestones match to processes within MMIS so that these processes can be evaluated and then retired.

The Medicaid Data Hub and Data Integration layer allow States to own their data and schedules. Once the data for a process is in the Medicaid Data Hub the States can feed that data back to the source system ensuring it's ready for a module. |
| 4.2.24 | Describe the major trends in your Medicaid Enterprise solution category that you believe BMS should be aware of, including any product or approach changes that you believe will come to market within the next 12 – 24 months. How do your Medicaid Enterprise solution roadmaps stay current with such trends? If possible, please be specific regarding how these trends affect Medicaid, WVCHIP, or healthcare IT in West Virginia.

The West Virginia (WV) Managed Care Quality Strategy is a good example of the importance of our Medicaid Enterprise solution category. The WV goal to increase the quality of care and identify areas for performance management and improvement among the managed care organizations (MCOs) in the Medicaid and WVCHIP populations will increase demands for complex data management and expose areas of data challenges. However, enabling analytics with deeper insights into the data is the right path for improving outcomes for your members, and especially as Medicaid enrollments are increasing at a healthy rate.

Enabling this level of data insights to inform these important business decisions can be supported by the following;

- **The technical ability to access and exchange members healthcare data with various sources** - In August 2021 during the MESC conference in Boston, MarkLogic announced the first two Medicaid Accelerators released as part of its Solution Accelerator Program. The first of these solution accelerators is the FHIR Mapper (intended for existing MarkLogic customers) to make their data accessible via FHIR based queries. The second accelerator is a Starter Kit for the Medicaid Data Hub. Intended for new projects, the Starter Kit includes FHIR-friendly persistent data models, security configurations, member and provider mastering configurations, redaction rules, and FHIR-based data access layer. The purpose of this is to enable customers to have a sample project from which to begin their new Medicaid project that can be customized and tailored to the SMA's specific needs and the accelerates their time to value.
- **Ensure data quality, privacy and security mandates are being managed in information exchange and access** - MarkLogic is currently working on its next accelerators, to include Data Anonymization and Redaction rules specific to PII and PHI seen in Medicaid environments, as well as to expedite the creation of CMS reports. More broadly, MarkLogic has a larger roadmap including enhanced Data Quality, Governance, Metadata Management, 3rd Party Connectors, and other tooling to enhance the capabilities and user experiences of the Data Hub. |

- ***Evolving your data architecture to provide deeper insights -*** MarkLogic recently acquired SmartLogic, the provider of leading AI and Knowledge Management capabilities which we anticipate will be incorporated into the MarkLogic Data Hub in 2022. SmartLogic's capabilities have significant applications in the healthcare IT in the areas of predictive analytics, patient service, drug and treatment knowledge surfacing, etc. SmartLogic's Semantic AI platform, Semaphore, reveals qualified contextual data using a centralized integrated platform that identifies enterprise information with capabilities to create and manage semantic metadata, active metadata, and information extraction.

There is a trend towards serving the 'Whole Person', i.e. addressing the Social Determinants of Health - Social, economic and environmental factors are the primary drivers for health outcomes, care costs and care quality. In 2021, CMS issued guidance to the states to drive the adoption of strategies that address the social determinants of health (SDOH) in Medicaid and the Children's Health Insurance Program (CHIP) so states can further improve member health outcomes, reduce health disparities, and lower costs in Medicaid and CHIP.

From a data perspective, this means that states must have a complete 360° view of the member across various program silos that manage the members various benefits programs, e.g. Medicaid, SNAP, TANF, Healthcare, etc. These benefits programs typically exist in disparate systems, and connecting these silos with traditional ETL approaches is a costly, long term and risky proposition. MarkLogic is meeting this challenge head on by providing its Data Hub's capability to quickly (months not years) ingest structured and unstructured data from the program silos, and then harmonize end curate the data, creating a 'gold record' of the member enabling the 360° views. This member master data can be consumed by the SDoH Care Manager using a choice of tools, for example either their department standard analytics tools or perhaps a Care Management purpose-built application. MarkLogic Data Hub is the data technology foundation to enable management of the SDoH.

MarkLogic invests to ensure our product management plans and roadmaps are current and forward looking. In December 2021, Gartner listed MarkLogic Corporation as a Visionary in its Gartner Magic Quadrant for Cloud Database Management Systems.

4.2.25	Identify any innovations in your Medicaid Enterprise solution for addressing Medicaid Business Priorities (cost savings, performance efficiencies, improved care outcomes, etc.).
	The MarkLogic Medicaid Data Hub provides tremendous innovation that addresses many business priorities including cost, performance, and improved care outcomes.
	The MarkLogic Medicaid Data Hub consolidates many different capabilities traditionally seen in different software packages into one consolidated platform. It is worth noting that this is not multiple different products and/or different codebases all "under the MarkLogic umbrella," but rather one tightly integrated software product that provides the necessary capabilities to build and scale your Medicaid infrastructure. Ultimately, this means development cycles can be spent focusing on *data integration* tasks and *module development* instead of *system integration* tasks for your core Medicaid infrastructure. The MarkLogic Medicaid Data Hub provides integration, search, government-grade security, an operational data store, master indices for Master Provider and/or Patient Indices, and redaction/masking for de-identification, flexible query, and much more as configurable out-of-the-box capabilities. This is not seen elsewhere in the market, and enables MarkLogic customers to compete project up to 4 times faster than with traditional legacy approaches. Further, it is because of this innovation that MarkLogic customers have seen reduced risk to Systems Integration bids – allowing SI partners to spend the majority of their time on application development in Full Operating Capability (FOC) modes rather than rote activities the prolong achieving even Initial Operating Capability (IOC). This faster time to mission lowers cost, improves performances, and shorten time to improved care outcomes as a result of system utilization.

| 4.2.26 | Identify any innovations in your Medicaid Enterprise solution for addressing technical risk management. |

Technical risk on a large, system-of-systems program comes from overall complexity, and high coupling (technical dependencies) among components or modules that should be separate. MarkLogic recommends a number of strategies to reduce technical risk, some described in more detail elsewhere in this RFI response. They include API-based de-coupling of modules, centralized data processing in a single Data Hub or ODS, using a NoSQL data store that stores data aligned to message APIs, and tools that allow declarative or GUI-based configuration of many key functions. Combining operational and analytic data access is also a good way to remove a lot of technical complexity, and therefore reduce overall system complexity, and in turn reduce technical risk.

- API-based de-coupling: A system of systems lives or dies based on the quality of the system-to-system communications. One module or component can always be replaced, but if the core data model and handling is tightly coupled to the rest of the system everything becomes difficult to fix, upgrade or replace.
- Centralized data processing in a single Data Hub or ODS: Curated, high-quality data is extremely valuable, and had been processed in various ways to be fit-for-purpose. This high-value data should be stored (not merely queried and sent around, and transformed or cleansed while in transit). Storing it allows remediation processes, de-duplication, validation, and indexing for easy query and access. Then the central data repository becomes a simpler and well-governed place to implement or review all data processing, including validation, security policy, reprocessing, data lineage tracking, transforms, data version management and other tasks.
- NoSQL Databases, with a multi-model capability: To avoid building many different data models it is less risk and effort to have the API data formats (often FHIR) align to the storage models. This often means JSON formats for both persistence and API messages. Storing data as rows and columns means building a completely new RDBMS data model

and constantly "shredding" JSON API messages into these relational formats – only to join the rows back together with more technically complex code to serve every API call.

- o Multi-model capabilities allow many formats such as JSON, XML, RDF, relational, text, binary and others to be handled in one data store, without complex integration among many data stores (such as a NoSQL key/value store, graph database, text search engine and triple-store).
- Tools for declarative configuration: Code increases technical risk. Only the developer who writes it is guaranteed to understand it, and the code for at least some developers invariably becomes complex and hide bugs. Declarations are understandable by business users and also regularize the system since declarations are more constrained to a particular purpose. E.g. a declarative security file can directly enforce access for various records and data properties, vs writing code to check roles against data access calls. Similarly, a declarative data mapping file can simplify the transform among data formats vs writing long scripts in a programming language.
- GUIs to further simplify configuration and allow more types of people to understand the system: GUIs are the next level up the understandability chain from declarative configuration. A GUI takes a declarative configuration and formats it graphically, and ideally also makes it interactive. E.g. the mapping configuration noted above can have a left and right side in a GUI and a "test" button that shows the effect in real time of a mapping change. This further avoids obscure work that is overly-complex and not easily monitored or understood by various (non-programmer) stakeholders in a project.
- Combined operational and analytic access: Much as it simplifies a system to consolidate the message data model and persistent model using NoSQL, it is also a great simplification to use the main data store for analytic reports wherever possible. The goal should be to reduce or eliminate the data synchronized to a separate, downstream enterprise data warehouse (EDW) by reporting directly on the operational Data Hub or ODS store, or a read replica of it. This removes yet another data model (the dimensional model), a data transform layer, and data synch process and various monitoring and reconciliation processes. It increases overall agility as well.

4.2.27	Describe 1 to 3 use cases where innovations in your Medicaid Enterprise solution would apply and the value your Medicaid Enterprise solution would add when applied to them. **Agile Data Integration** – The MarkLogic Medicaid Data Hub platform enables ease of large scale integration of multiple, siloed data sources with a fast, iterative and less risky approach – regardless of the form of the data (structured, sem-structured, or unstructured). The Data Hub provides centralized repository that provides a full data discovery and access layer that allows for data to be interrogated/queried in a multitude of ways, and all with built-in data governance and security. **Member/Provider 360** – The MarkLogic Medicaid Data Hub leverages its built-in Master Indices and multi-model capabilities to provide a single view of member or provider information. This is helpful from a Medicaid standpoint, but also from a holistic approach in sharing information across state agency boundaries, thus supporting "No Wrong Door" initiatives and enabling state HHS organizations to expedite time to benefit and ensure needed services are delivered. **Operational and Analytical Reporting** – The MarkLogic Medicaid Data Hub can support the transactional loads required of an Operational Data Store, but also provide curated data for analytical reporting typically seen in Enterprise Data Warehouse environments. MarkLogic can help to consolidate this by providing one centralized data store capable of handling the transactional integrity demands of day-to-day data, while also enabling downstream applications to glean insights from this data via standard interfaces and BI reporting.
4.2.28	In the states where you have implemented, what have been some of the higher value outcomes? What performance metrics were you able to provide to substantiate this success? MarkLogic's outcomes have been expedited time to mission and results using the MarkLogic Data Hub. By way of example, in a southeastern state the initial project stand-up (including all project iterations, data flows, etc) for a use case was stood up in 4 months with one (1) data source. Over the next 30 days thereafter, an additional eight (8) data sources were able to be added to the system.
4.2.29	Discuss any experiences you have had integrating your Medicaid Enterprise solution with legacy system management and lessons you have learned for implementing new Medicaid Enterprise solutions. Do you recommend any specific approach for modifying, interfacing with, and managing the legacy system while implementing a new Medicaid Enterprise solution? Every time MarkLogic has implemented our solution, we have successfully integrated with legacy systems. Typically, legacy systems are Mainframe systems. The best way we have interfaced with these Mainframe systems is to use the existing exports (typically a fixed-length flat file) with their data definition (typically a copy book). We have developed tooling that uses the data definitions to read the export and insert the data into the MarkLogic Medicaid Data Hub solution. New exports can be made but

|	
	that requires work from the team that is maintaining the legacy system. We have also seen other legacy systems such as Relational Database Management Systems (RDBMS) and have the tooling to extract data from these systems and insert this into the MarkLogic solution, as well. Our solution supports an agile iterative phased approach that takes parts of the legacy systems and management in phases. These parts are typically business process and data-driven such as enrollment and member data. Typically, the legacy systems still need all the data to be kept up to date once the business process is moved out of the legacy system. For example, claims processing needs member enrollment data. Our solution handles this challenge by interfacing with legacy systems to keep the data up to date for the duration of the migration from the legacy systems to the MES. Once the migration has been verified and the legacy system is retired then the export to the legacy system is no longer needed.
4.2.31	Describe the System Development Lifecycle (SDLC) approach that you use for implementing your Medicaid Enterprise solution. Can your SDLC approach be incorporated into an environment that uses a traditional "waterfall" SDLC approach? What about "agile" methodologies to support the implementation of your Medicaid Enterprise solution? If so, how can this be accomplished? MarkLogic can support both Waterfall and Agile Methodologies for development, and is particularly well suited for the latter. MarkLogic facilitates the Agile delivery model by taking a different approach to data and recommends a "Data Services First" approach. The fundamentals of project delivery using Data Services First, e.g. the skills needed, are not different from those in legacy approaches. But the recommended steps and their order reflect a better alignment with the goals of Agile: progressive, iterative, and fast delivery. MarkLogic was built for modern software delivery practices. And the good skills and foundations that your existing teams have acquired through their enterprise implementations are fully transferable for smooth and fundamentally sound implementation. Experience has taught that focusing on delivering the most high-value functionality early in a project increases the chances of a successful delivery. Delivering a minimum viable product (MVP) into production quickly is the goal of most modern software projects. An MVP that is scoped to provide high value with low risk, and with progressive enhancements is the exemplar of Agile methodologies. Using MarkLogic while adopting a Data Services Approach allows project teams successfully deliver on the promise of the MVP. Below are the basic steps for implementing a MarkLogic MVP using a Data Services First approach. **Define the data service** - Define what the output requirements are based upon what the service will be used for. A data service is a fixed interface to the data managed in MarkLogic expressed in terms of how the data will be consumed by the application tier. For example, data services can run queries ("Find eligible insurance plans for an applicant"), updates ("Flag this claim as fraudulent"), or both ("Adjust the rates of plans that haven't made claims in the last year"). The data service allows the service developer to obscure the physical layout of the data and constrain or enhance queries and updates with business logic. This allows application developers to contribute fully to the project without needing expertise in MarkLogic. **Ingest data** - Ingest your source data as is into MarkLogic. The harmonization steps defined by your data service model(s) will transform those pieces of the data needed for the defined data service(s). Any unharmonized data can be processed later if and when it is needed by a newly defined data service. This is a truly progressive approach to data modeling. **Model only what you need for each data service** -Create a model with the minimum set of output properties required. Modeling only what you need, when you need it, provides the opportunity to ensure that the model is created collaboratively, that it serves the business requirement and works well for developers to interact with. **Allow application developers to access the data service(s)** - Now that one or more data services have been created they can be exposed so that your application developers can begin writing code against the real data. Code is tested against real data. This close alignment between data modeling and development significantly reduces the risk that code will need to be refactored later. No more surprises in the middle of your project! **Iterate** - While your application developers are busy writing and releasing features with their existing set of data services, the next set of data services can be defined and built. Data Services facilitates a truly iterative methodology. Adopting the Data Services First approach reduces the complexity of your project. Dispensing with the need for separate data and development streams reduces risk and administrative oversight requirements.
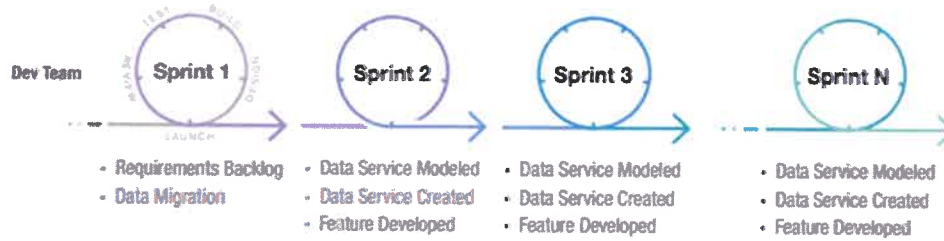
| 4.2.32 | What is the typical duration of a project to implement your Medicaid Enterprise solution? How does this timeline break down across the planning and DDI phases?

It is recommended to have the Medicaid Data Hub be an iterative and phased solution. The Medicaid Data Hub should provide the modules the data they need and would be dependent on the timelines of the modules. Data can be brought into the Medicaid Data Hub before the modules need them. However, development time should be allowed for the specific needs of the modules. Such as connections and data feeds. |
|---|---|
| 4.2.33 | What do you see as the key cost drivers for implementing your Medicaid Enterprise solution? What recommendations do you have for managing MES costs and demonstrating outcomes that mitigate any unnecessary costs of a Medicaid Enterprise solution?

The primary cost drivers for implementing a MarkLogic Medicaid Data Hub are the following:

- **Accelerate Modernization** - MarkLogic's Medicaid Data Hub enables data to be integrated and made discoverable up to 4 times faster than with legacy approaches. We recommend building the ODS before individual modules are implemented, which jump-starts data modeling, sharing and governance immediately.
- **Simplified Architecture** – MarkLogic's Medicaid Data Hub contains a multitude of capabilities traditionally seen several different software technologies. By leveraging MarkLogic, states can realize the cost benefit of not having to perform prolonged systems integration tasks, and instead focus on application and module development activities. Additionally, MarkLogic's Medicaid Data Hub contains the necessary enterprise features – such as high availability, disaster recovery, government-grade security, ACID transactions – to ensure that states can rely upon to run mission critical systems. The MarkLogic ODS platform eliminates or reduces the need for other systems, such as an enterprise data warehouse or data lakes. The same real-time, service-based architecture in the ODS supports business intelligence and reporting as well, saving hours and complexity otherwise needed to model, copy, transform and manage multiple data stores. MarkLogic natively handles JSON, XML, relational, text and graph data in one platform.
- **Share data safely:** MarkLogic secures data *at the data layer*. Every query, record and field can be secured using declarative configurations. Every "data lens" from REST query calls to SQL access or even knowledge graph access via RDF respect the unified security configuration in MarkLogic. Test or analytic data exports can be run through de-identification and redaction, to enable data sharing instead of inhibiting it.
- **Enable unified 360° views needed for use outside of Medicaid**: A complete view of members, providers, hospitals and any other important entity supports integrated case management and integrated eligibility and verification for benefits programs like TANF, SNAP, energy subsidies, and even services like behavioral health programs and child welfare provided by other agencies. With MarkLogic Smart Mastering, getting a complete, real-time, person-centric view from disparate data silos – a true golden record – is faster, easier and cheaper.

Future-proof your MES - APIs, REST services, and schema flexibility ensure that your organization can swap out modules or take in new data sources without disrupting the entire enterprise. If the APIs are stable and tested, internal data changes will not break downstream data consumers. |
| 4.2.35 | What do you believe would be the optimum duration and the minimum duration for DDI of your Medicaid Enterprise solution?

It is recommended to have the Medicaid Data Hub be an iterative and phased solution. The Medicaid Data Hub should provide the modules the data they need and would be dependent on the timelines of the modules. Data can be brought into the Medicaid Data Hub before the modules need them. However, development time should be allowed for the specific needs of the modules. Such as connections and data feeds |
| 4.2.36 | List and describe the documentation developed by your company and/or the state Medicaid agency that is essential to DDI and operations of your Medicaid Enterprise solution. |

|	
	The background and documentation needed would be around the data itself and the source systems that currently retain the data as well as the downstream systems that require the data such as modules. Healthcare data traditionally is transferred as large batch files although recently there has been movement in the market for real-time data. The background and documentation on the requirements of how the data can be transferred from the source system and to the downstream systems would be needed. The different formats of the data and the size of the data would also be required. The required frequency at which data would need to be refreshed to the downstream systems and from the source systems would need to be provided. The capabilities of the source and downstream systems would need to be known such as change data capture, interfaces, and authentication methods. Other pertinent software that the State currently owns such as an Enterprise Job Scheduler and Enterprise ETL tool, and preference of the state in using these tools.

4.2.37	Detail how your Medicaid Enterprise solution could support BMS in improving data analytics and reporting capabilities, data sharing initiatives, and overall confidence in health data.
	
	It is critical for development teams within the Medicaid Data solution and other downstream systems such as modules to have access to data feeds that are as close to production data as possible. Only data that is required by the recipient should be provided. Techniques such as redaction, de-identification, and element-level security should be used to facilitate data sharing.
	
	These techniques have allowed our customers to be able to provide data that has been cleansed of Protected Health Information (PHI) and Personally Identifiable Information (PII). The data can be replaced with valid values allowing for development teams to test with "production like" data without having to deal with PII and PHI.
	
	When sharing actual production data it's important to ensure its security and to track what data was shared and to whom it was shared.
	
	We have seen a number of approaches and techniques that facilitate data sharing generally.
	
	• Use Data APIs (typically RESTful with JSON data payloads) to standardize data interchange, and decouple components. APIs are easily documented (e.g. using the OpenAPI spec), can be standardized across the enterprise, and can be versioned, discovered and managed by API gateways and other technology. As a standard technology they are familiar to developers and can usually be consumed by various components. Exposing data primarily via APIs is consistent with a services-based approach.
	    o Note that some systems will use only business APIs, with direct database access and tight coupling to underlying data formats and stored. This is not at all what we mean here – it is critical to have *data APIs* served directly from the Data Hub to avoid tight coupling to data formats and expensive relational-JSON mapping layers.
	• "Data First" development: define data and implement the Data Hub aspects early. To standardize, it is important to define a good model that will work well for all data consumers. This is much easier if done early on, and not as part of (or even concurrent with) the first MMIS modules or systems that are integrated. Starting early allows the State to stand up a team to think through and standardize APIs and interchange formats while considering the big picture for Medicaid and State future needs. As noted elsewhere in this response, we support basing the enterprise standard model on FHIR. Data dictionaries, documentation and industry expertise standardized in FHIR can be leveraged even if the persistent model is slightly different from fully-compliant FHIR. Conversion to true FHIR for interoperability is then easy even as the persistent layer is customized for other needs.
	    o Early focus on data processing also provides data access and APIs early on to drive module development, and even to drive module evaluation during procurement.
	• FHIR supports JSON, XML and RDF interchange, but we recommend using JSON, due to the ubiquity and simplicity of that format, to drive greater and easier sharing via RESTful APIs.
	• We recommend direct NoSQL (non-relational) storage of the JSON messages in the Data Hub data store. MarkLogic is an ideal technology to use here, as it natively augments JSON data with RDF and SQL data for storage. RDF is useful because JSON data is tree-structured, and it must be augmented with "graph" linkages such as RDF provides – this helps with "graph-y" data such as provider networks, family structures and other relationships that are not purely tree-based. RDF is a W3C standard for rich graph relationship data. MarkLogic adds SQL projections to provide a convenient interface for typical reporting and BI tools, as well as some data interchange tools while still enjoying benefits of a NoSQL storage approach.
	• Fresh, timely data improves data sharing by making data more actionable. Therefore, move processing toward real-time to the extent possible. MarkLogic's open-source "State Machine" processing framework is ideal for this, using data-oriented events and conditions to run data through required transforms, validations and modifications as it flows into and through a Data Hub.

- **Combine operational and analytic systems into one system**, to avoid complexity. MarkLogic has built-in technology to support this (Template Driven Extraction creates transactionally up-to-date SQL views of all JSON data such as FHIR records), and we find this greatly facilitates data access and sharing because:
  - o **Combined analytic and operational data storage allows real-time reporting to be the default, rather than an add-on or proprietary approach.**
  - o Combining them reduces barriers to access and delays by eliminating complex and slow re-modeling of the operational data into a dimensional EDW model for analytics. It also avoids the issues and delays inherent in data movement jobs, monitoring, DevOps provisioning. Also combining them avoids additional technology components, additional vendor management, additional monitoring, and additional security compliance overhead of a separate EDW component.
- Paradoxically, strong security enhances and enables data sharing. Therefore, build in security at the Data Hub and database level. The reason security enhances sharing is that a person cannot access a data store at all if that data store (database) contains un-secured data they should not have access to. (E.g. adding a bank account number to a database without RBAC or similar controls immediately means nobody without financial data access can use that database directly). To facilitate this, the Data Hub database should support
  - o Role-based security (RBAC)
  - o Attribute-level (field-level) security
  - o Query-based security (QBAC, or query-based access control, allowing complex conditions to enforce security)
  - o Universal security enforcement regardless of access pattern. API-based access, file exports, and SQL views should all respect the same, universal, declarative security policy.
- Build a capability to produce and manage De-identified data
  - o Much as how good security actually facilitates data sharing (see above), data sharing also sometimes requires de-identification of the data, per the HIPAA rule and guided by NIST 800-122.
  - o Use Limited Data Sets for testing and procurement. True "de-identification" per the HIPAA rule is often difficult for highly-dimensional data (because the combination of various data values can technically identify a person), and therefore the production of Limited Data Sets per the HIPAA rule is often needed. Limited Data Sets require removal or masking (or randomization or other obfuscation) of a set of clear identifiable information such as name, address, and ID numbers, but allow certain data such as dates and diagnoses to be left in place.
  - o Produce de-identified or Limited data sets with a particular "purpose" in mind – think "fit for purpose" when sharing data. How much, and how fully de-identified is required? One purpose – statistical analyses for population health – requires diagnoses and population demographics to remain intact, but MMIS module software testing requires completely different data to remain the same to exercise module business logic.
  - o Declarative configuration for de-identification and Limited Data Set production. We have found that declarative de-identification configurations are more governable because they are understandable by the data and security experts, who are often not coders, and declarative artifacts can be tracked and versioned more easily than scripts or routines written in python or Java code that implement de-identification via coding.

Robust, multi-model data storage and query. Consider all types of data access, including future access. Initial uses are often simple access using simple queries, but ultimately full-text search, graph analyses and binary handling are all likely to be needed. E.g. processing a claim may only require simple, structured data queries for enrollment status and third-party liability checks. However, finding a provider may require text name search together with geospatial proximity query and graph analysis to determine the participation of the providers via their networks. A "no wrong door" social services support capability requires a graph-based query of the family and non-family relationship structure. Handling all types of data in this way requires "multi-model" storage and processing, or requires that the systems integrator cobble together a collection of multiple databases of different types (document store, relational, graph database, search engine, binary data store, and geospatial query tool). Choosing a multi-model storage technology up front reduces cost in the out years and avoids building a brittle, non-portable MES architecture.

4.2.38	Describe or illustrate your data visualization capabilities.
	The MarkLogic Data Hub for Medicaid is not a data visualization product. However, it does prepare data for visualization, and often serves as the back-end data repository that feeds BI/Reporting/Analysis tools. The MarkLogic Data Hub for Medicaid can also enhance visualizations by refining data sets based on user input, allowing for sub-second response times on queries that include keywords, geospatial, temporal, and relationship-based data. Below are several illustrations of MarkLogic-supported visualizations in different tools.
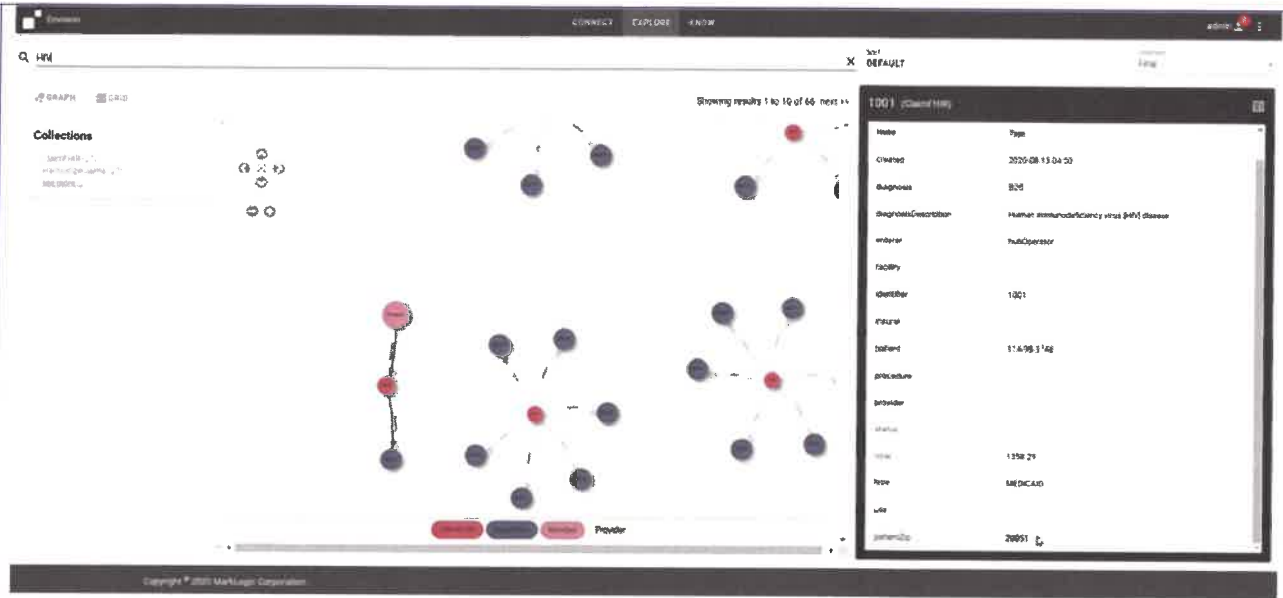
| 4.2.39 | How does your Medicaid Enterprise solution improve the coordination of care, detect and prevent fraud, waste, and abuse to support Medicaid program integrity, and improve stakeholder access to state Medicaid Enterprise data?

The MarkLogic Data Hub for Medicaid serves as a centralized repository for all MES data – claims, members, providers, reference data, etc. This data can be used, as appropriate, for application or module-specific functions to improve coordination of care or to support Program Integrity objectives. For instance, MarkLogic allows for machine learning to be run directly on the data, which can be beneficial to surfacing potential incidents of fraud/waste/abuse. While possible to do this, these types of applications are often purpose-driven with their own custom user interfaces to accomplish these tasks. MarkLogic has partners who have built these types of applications on our platform. However, it should be noted that the MarkLogic Data Hub for Medicaid does not support these functions out of the box, but can readily be used as the underlying platform upon which these applications are built.

However, where there MarkLogic Data Hub for Medicaid does provide tremendous capability out of the box is the ability to improve stakeholder access to state Medicaid data. One of the key tenants of MarkLogic is the ability to access data in a multitude of ways, and to do so securely. MarkLogic provides a robust "Ask Anything Universal Index" that indexes all data in the Data Hub and provides sub-second responses on queries. MarkLogic also allows for the data to be queried in a multitude of ways – from "Google-esque" search with full Boolean grammar, to using graph or relationship-based queries, to relational views of data. By leveraging the MarkLogic Data Hub for Medicaid, you enable your stakeholders to ask questions of the data and retrieve data in a plethora of ways, and to do so safely knowing that the data is secured with government-grade security built directly into the Data Hub. |

Integrating the Medicaid Silos

An *Operational and Analytical* Data Hub

EASY TO GET DATA IN
Flexible Data Model

- Data ingested *as is* (no ETL)
- Structured and unstructured data
- Data and metadata together
- Adapts to changing data and changing data structures

EASY TO GET DATA OUT
Ask Anything Universal Index

- Index once and query endlessly
- Real-time and lightning fast
- Open standards, open APIs
- Multi-model database unifies query across JSON, XML, text, geospatial, and semantic triples in one database

TRUSTED TO RUN YOUR BUSINESS
Enterprise Ready

- Flexible cloud deployment
- Enterprise-grade data security
- Reliable data and transactions (100% ACID compliant)
- Seamless integration with your existing environment

MarkLogic Data Hub for Medicaid
Simplified architecture accelerates module development by combining ODS, MPI & EDW
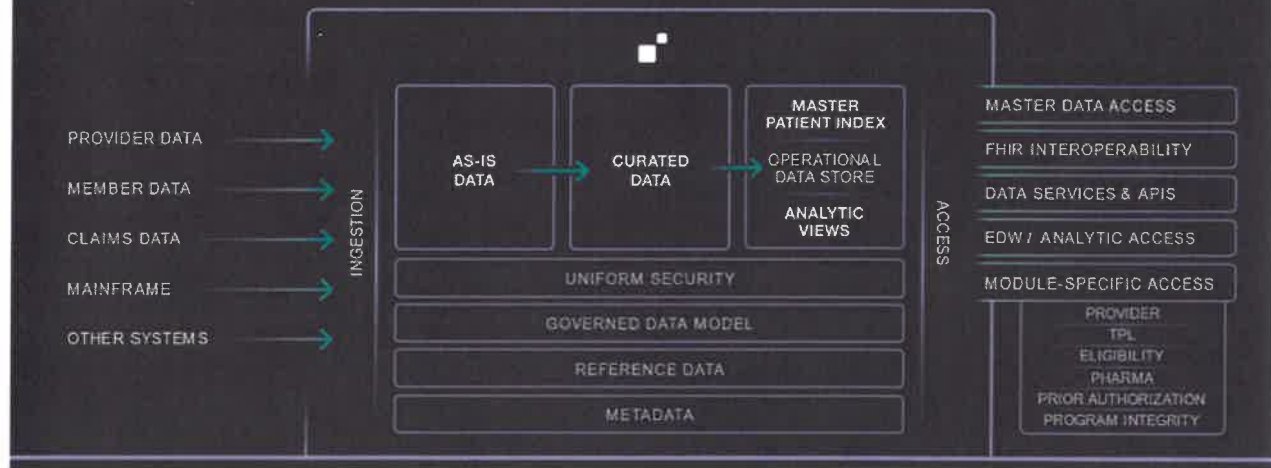
4.2.40	Describe how your Medicaid Enterprise solution increases access and shared use of data with both the State and other vendors, improves healthcare quality management, and increases automation capabilities.

The MarkLogic Data Hub for Medicaid provides tremendous capabilities to query and access shared data in a multitude of ways. This is addressed in more detail in 4.2.39 regarding the "Ask Anything Universal Index" capabilities.

However, the key to increasing shared used of data within the State and with other vendors is having robust security at the data level – and this is exactly what MarkLogic does. Security is often thought of as the inhibitor to sharing, largely because of the different "fiefdoms" that stove-piped decentralization demands. In the MarkLogic Data Hub for Medicaid, security is a forethought, not an afterthought, and SMA's can ensure that data is shared appropriately because the data is secured by default in MarkLogic, and only allowed to be shared with those authorized to see it. By having a centralized security model at the data-level in the MarkLogic Data Hub for Medicaid, SMA's inherently promote sharing by ensuring compliance with security policies at the outset that is built into the Universal Index access layer.

4.2.41	If applicable, how does your Medicaid Enterprise solution improve access to endusers, such as a user's data or access to additional services?
	Please see answers for section 4.2.39 and 4.2.30 which expands on data access and security in more detail. The MarkLogic Data Hub for Medicaid provides a multitude of ways to access data leveraging it's "Ask Anything Universal Index" that is tied directly to the MarkLogic security model. This allows for improved and secured access to end user data, and many ways for it to be rendered out in applications.
4.2.44	Do you have a short demonstration of your approach and/or Medicaid Enterprise solution that you would like to present to BMS? If so, please describe the method of presentation for the demonstration and suggestions for who should attend. If BMS wishes to take part in a demonstration, they will reach out to the Respondent for further information.
	Yes. The MarkLogic team would welcome the opportunity to present a demonstration of the MarkLogic Medicaid Data Hub to the State of West Virginia. We are happy to accommodate in this in person or via remote conferencing (i.e., Microsoft Teams, Zoom, Webex, or similar). We recommend that the session be attended by Medicaid technical and business leadership, and their supporting advisors who are involved in the evaluation and decision making process (i.e., Medicaid Directors, CIO/CTO/CDO, Enterprise Architects, etc).
4.2.45	Is there additional information you would like to share with BMS related to the topics addressed in this RFI?
	This RFI has been comprehensive and there is no additional information to share at this time.

ATTACHMENT 1: SECTION 5.2.4 TWO-PAGE OVERVIEW

(Attachment on subsequent pages, this page intentionally left blank)

MarkLogic·

MMIS Modernization

Operational Data Store and MPI for Medicaid Transformation

To improve the delivery of healthcare services and meet requirements for health data interoperability, technology leaders of state health and human service departments need to modernize their aging Medicaid Enterprise Systems (MES). The *MarkLogic Data Hub for Medicaid* combines an Operational Data Store (ODS) and Master Patient/Provider Index (MPI) that allows them to confidently accelerate migration of legacy data to a more modern platform that simplifies data integration while future-proofing IT architecture.

MMIS Modernization Challenges

State government Medicaid Management Information Systems replacement (MMIS-R) projects are monumental undertakings that traditionally suffer from lengthy development cycles – resulting in costly systems integration and deployment delays.

Legacy ODS and separate Master Data Management (MDM/MPI) technology are simply not up to the task of quickly and reliably delivering a complete, real-time, person-centric 360° view from disparate data silos. To get to a true golden record of your members and providers – and ensure the data security you need – you need a new approach for integrating data from legacy and newer systems to modern MES modules. MarkLogic further simplifies modernization by handling much of the EDW load directly from the ODS.

Why MarkLogic Data Hub for Medicaid?

In a word: Simplification. By combining many components, MarkLogic reduces risk, complexity, and time-to-value.

MarkLogic provides an enterprise-ready MMIS ODS with MPI & EDW capabilities. Our ODS platform provides more capabilities than any traditional ODS – including built-in master data indices, search, security, and analytics. It integrates data from any legacy system or new module with ease and reduces the risks and costs of MMIS modernizations. As a proven ODS approach, we ensure that State Medicaid Agencies (SMAs) are HIPAA, MITA, MECT and FHIR-compliant, accommodate change, and position states for success in the out years – rather than being locked into a particular module or technology stack.

Having all member, provider and program data in one place with easy and secure accessibility is also a first step towards improved operational execution and better service delivery, such as cross-department benefits coordination and population health management.
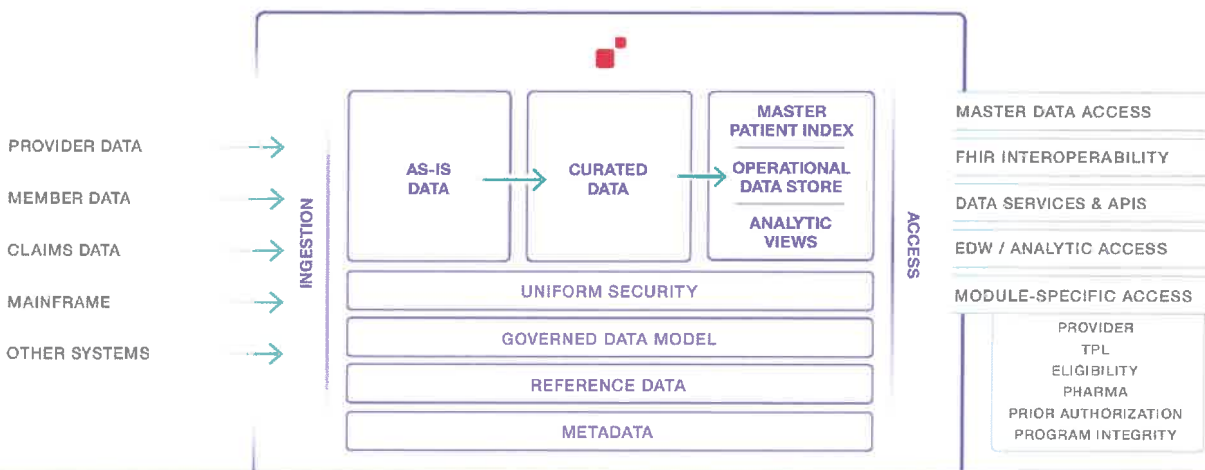
MarkLogic Data Hub for Medicaid accelerates module development by combining ODS, MPI & EDW capabilities

MarkLogic·

Benefits of Modernizing with MarkLogic

- **Accelerate development and modernization:** A MarkLogic ODS can be deployed before individual modules are implemented, allowing a State to get started on data modeling, sharing and governance immediately. And with intuitive tooling, data owners are less reliant on software engineering for all data integration tasks.

- **Streamline your CMS-compliant architecture:** Using the MarkLogic Data Hub for Medicaid, your organization now has an enterprise-ready MMIS ODS with comprehensive capabilities that can eliminate or reduce the need for other storage systems, such as enterprise data warehouses and lakes. This simplified architecture for data integration is HIPAA-compliant and consistent with MITA and MECT guidelines

- **Enable unified 360° views needed for service development:** By curating, deduplicating and linking all data, systems can access a complete view of members, providers, hospitals and any other important entity. This is critical for Medicaid processing, and may also support integrated case management and eligibility for other programs such as TANF, SNAP, LTC, asylee or energy subsidies.

- **Directly provide reporting and analytics:** Analytic outputs and reports can be served directly from "data lenses" within the MarkLogic ODS, which provide customized analytic views. This direct reporting and analytic capability can reduce or eliminate the need for custom analytic components such as a full-scale EDW.

- **Share data safely:** Upfront, built-in, certified security–including support for secure de-identification and redaction – enables data sharing instead of inhibiting it.

- **Future-proof your MES:** Schema flexibility ensures that your organization can integrate data from any legacy system or new module with ease, and adapt to future changes.

Proven MMIS Modernization Success

Around the US, states are successfully transforming Medicaid services with MarkLogic.

MMIS replacement to improve health services for 800,000 beneficiaries
A southwestern US state Human Services Department selected MarkLogic to drive transformation improvements in constituent services by delivering a secure, complete, and adaptable 360° view of citizens, providers, and hospitals for new health analytics and integrated case management, eligibility, and verification.

Improved Medicaid eligibility and claims processing
Completed in 120-days, this southeastern state's Department of Health and Human Services now has a modern MES that streamlines claims processing and reduces data management costs.

Modernizing MMIS in the cloud
A state department of community health is building their future MMIS using MarkLogic as the ODS and Master Data Management solution on the AWS Cloud. MarkLogic was chosen due to its proven success with other state health departments, robust product features and strong solutions engineering capabilities.

About MarkLogic

MarkLogic helps Medicaid agencies achieve a highly secure and actionable 360° view of their data, updated in real-time –reducing the risk and cost of modernization efforts, and positioning each agency to easily adapt to future changes.
For more information visit: www.marklogic.com/solutions/industry/state-local-government

ATTACHMENT 2: RFI ACKNOWLEDGEMENT FORM

(Attachment on subsequent page, this page intentionally left blank)

Request for Information
CRFI BMS2200000001
Medicaid Enterprise System (MES)

By signing below, I certify that I have reviewed this Request for Information in its entirety; understand the requirements, terms and conditions, and other information contained herein; that I am submitting this response for review and consideration on behalf of my organization.

MarkLogic Corporation
(Company)

Gary Katz, Sr. Account Executive
(Representative Name, Title)

M: 215-840-8571
(Contact Phone/Fax Number)

1/10/2022
(Date)

ATTACHMENT 3: ADDENDUM ACKNOWLEDGEMENT FORM

(Attachment on subsequent page, this page intentionally left blank)

ADDENDUM ACKNOWLEDGEMENT FORM
SOLICITATION NO.: BMS2200000001

Instructions: Please acknowledge receipt of all addenda issued with this solicitation by completing this addendum acknowledgment form. Check the box next to each addendum received and sign below. Failure to acknowledge addenda may result in bid disqualification.

Acknowledgment: I hereby acknowledge receipt of the following addenda and have made the necessary revisions to my proposal, plans and/or specification, etc.

Addendum Numbers Received:
(Check the box next to each addendum received)

[✓] Addendum No. 1   [ ] Addendum No. 6

[✓] Addendum No. 2   [ ] Addendum No. 7

[✓] Addendum No. 3   [ ] Addendum No. 8

[ ] Addendum No. 4   [ ] Addendum No. 9

[ ] Addendum No. 5   [ ] Addendum No. 10

I understand that failure to confirm the receipt of addenda may be cause for rejection of this bid. I further understand that any verbal representation made or assumed to be made during any oral discussion held between Vendor's representatives and any state personnel is not binding. Only the information issued in writing and added to the specifications by an official addendum is binding.

MarkLogic Corporation
Company

[signature]
Authorized Signature

_1/10/2022_
Date

NOTE: This addendum acknowledgement should be submitted with the bid to expedite document processing.
Revised 6/8/2012

About MarkLogic

For almost two decades, organizations around the world have come to rely on MarkLogic to power their innovative information applications. As the world's experts at integrating data from silos, MarkLogic's operational and transactional Enterprise NoSQL database platform empowers our customers to build next generation applications on a unified, 360-degree view of their data. Headquartered in Silicon Valley, MarkLogic has offices throughout the U.S., Europe, Asia, and Australia.

MarkLogic

999 SKYWAY ROAD, SUITE 200, SAN CARLOS, CA 94070
+1 650 655 2300 · +1 877 992 8885 · www.marklogic.com · sales@marklogic.com