STATE OF ALASKA
REQUEST FOR INFORMATION (RFI)

Department of Administration
Shared Services of Alaska
333 Willoughby Avenue, 8th Floor
PO Box 110210
Juneau, Alaska 99811-0210

State of Alaska Division of Office of Information Technology Enterprise Data Hub and Service Delivery Platform

General Information

This Request for Information (RFI) is meant to identify potential sources and approximate costs for the microservices of the data hub and service delivery platform. The information in this document is not meant to be a comprehensive list of requirements but should give enough high-level information on the conceptual model to assist the Department of Administration (DOA) in developing intermediate cost estimates for the enterprise data hub and service delivery platform.

Responses should be geared toward describing the functionality of existing products or services and should include a rough idea of the costs based on past implementations of similar projects. Responses may include multiple cost projections based on varying scales of implementation. Cost estimates will not become part of any later Request for Proposal (RFP) process.

The data hub and service delivery platform will use service-oriented and microservice architecture. Microservices within this system should be designed to perform specific business functions and operate independently of each other yet provide robust interoperability through application programming interfaces (APIs). When providing cost and capability information, please provide a separate estimate for each proposed microservice. Components of the system may be a combination of open source, commercial off-the-shelf (COTS), and custom development.

Responses should support the business functionality described in this RFI; however, responses do not need to include all functionality described under Business Functions. Responses to this RFI may include features not listed within descriptions, and features may cross over between microservices; however, microservices must maintain a narrow scope and support interchangeability, cost effective maintenance and long-term sustainability.

Summary

DOA is seeking information on an enterprise data hub and service delivery platform to support data exchange across disparate information systems, and the separation of concerns across business, information, technology and application layers within the systems development lifecycle. Microservices[1] within this system should include support for the following business functions:

- Semantic Layer (SL)

---

[1] Individual software services that provide a very narrow scope of functionality to support interchangeability, cost effective maintenance and long-term sustainability. Services are made interoperable through robust application programming interfaces (APIs).

- Data Hub (DH)
- Data Dictionary (DD)
- Data Classification and Configuration Management (DCCM)
- Data Catalog (DC)
- Master Data Management (MDM)
- Data Warehousing (DW)
- Business Rules, Process, Decision and Case Management (BRPDCM)

Background

The State of Alaska continues to face fiscal challenges that require it to improve service delivery to Alaskans while using fewer resources. This means the State must focus on changing how it does business. The Office of Management and Budget has identified an overarching need for agencies to re-evaluate their scope and responsibility to ensure programs align with core services and to identify potential areas of opportunity for revenue generation, improved efficiencies, cost savings, and economic development. To be effective in implementing sustainable change, the State must first put in place the proper support mechanisms. Two such mechanisms are a data hub and service delivery platform.

The State envisions a hybrid data hub[2] to connect data across its many disparate siloed systems to enable informed decision making through analytics and business intelligence, and the use of standardized and master data to improve service delivery and reduce costs associated with developing and maintaining information systems.

Furthermore, a hybrid data hub will allow a smoother transition to enterprise cloud-based infrastructure. Cloud deployments are inherently challenging, and a hybrid data hub will greatly increase the State's chances of success by overcoming data silos, ensuring security through data classification and configuration management, and helping to overcome latency issues through virtualization.

A service delivery platform will further extend the State's ability to use its data and digital assets for the greatest return on investment for constituents. A service delivery platform will provide functionality through microservices. As envisioned, microservices will be consumed through the application layer of information systems resulting in shorter software development times and a significant reduction in costs related to the development and maintenance of information systems across State entities.

Business Functions

The business functions in this section are not meant to be a comprehensive list of requirements. Responses to this RFI may include features not listed within descriptions, and features may cross over between microservices; however, microservices must maintain a narrow scope and support interchangeability, cost effective maintenance and long-term sustainability.

Semantic Layer (SL)

There are many competing definitions for the SL, so for this RFI, SL is defined as the ontologies[3], taxonomies[4] and resulting common vocabularies[5] and terms that can be applied to data elements to give them context and extend their usability. Common vocabulary terms allow data to be easily discovered, accessed and understood. In addition, terms should allow data to be translated into machine readable code that convey context to computer programs or software applications. This machine-readable code will help to create interoperability between disparate information systems.

---

[2] A hybrid data hub allows real time data exchange across a virtually unlimited number of data sources. Data is aggregated into the hub using application programming interfaces. Additional hub features include data indexing; harmonization; contextualizing with metadata; governing over component configuration, classification, security and privacy; master data management, virtualization and storage; and discoverability and search features.

[3] An ontology is the description of concepts including entities and their properties, and the relationships between entities. An ontology defines a common vocabulary of terms that can be applied to data elements.

[4] Taxonomies are classifications of entities. For example, books are classified within the Dewey Decimal system by division and subdivision.

[5] Common vocabularies are a collection of terms that are defined within an ontology.

Common vocabularies will provide the agility to define data domains that will greatly extend the use and the value of State data through a data hub. Among other knowledge areas, domains can represent data schemas, definitions, configurations for configuration management, business process groups, workflows, reports, information systems, even entire organizational structures and legal statutes.

Semantic Layer (SL) functionality should provide the ability to visually model and manage ontologies, taxonomies and common vocabularies (OTCV). The system should allow the definition of entities, properties and the relationships between entities, and should provide the ability to publish unique namespaces[6] using discoverable Universal Resource Identifiers (URIs). The namespaces will document common vocabularies and provide access to terms for inclusion in Resource Description Framework (RDF) documents, and other machine-readable formats. For example, the URI for the DCAT or Data Catalog Vocabulary is http://www.w3.org/ns/dcat#. The ontology is described in RDF/XML here: https://www.w3.org/ns/dcat.rdf. And finally, the human readable version of the documentation is located at: https://www.w3.org/TR/vocab-dcat/.

Once common vocabularies are developed and accessible through a namespace, we can map data elements from microservices and disparate information systems to vocabulary terms through the data hub. The semantic layer must be kept separate from other microservices to avoid locking data to one domain, and for security of the data itself. An effective SL allows data to be translated into a virtually unlimited number of domains and uses through a data hub. Other capabilities of this function are the ability to visually model and manage versions of OTCVs.

Data Hub (DH)
The data hub is where data elements from disparate systems (including microservices) are mapped to common vocabulary terms to transform the data for use within domains. Domains are any area of knowledge described using a common vocabulary. For example, we can create a report domain or definition that helps transform data from multiple databases by mapping the data elements to terms. Metadata related to the terms are used within algorithms to aid transformation of the data to the standardized definition of the domain. The hub also implements virtualization of data through APIs, governance through the enforcement of business rules, security through data classification and configuration management, and performance management through intelligent caching.

Data virtualization (DV) should provide the ability to retrieve data across disparate systems without the data user having any knowledge of the originating information system or data schema. DV should provide an additional layer of security between the data users and the originating systems. Within this system DV should be delivered through robust application programming interfaces, that deliver data according to domains defined within the SL. The semantic layer, in combination with data classification and the management of data as a configuration item, prevents the combination of data that could contribute to reidentification of deidentified fields across datasets.

The DH should also provide the ability to define access by profiles. Profiles should be defined for users and information systems through data classification and configuration management rules. Respondents to this RFI should review NIST SP 800-60, which addresses the Federal Information Security Management Act (FISMA) direction to develop guidelines recommending the types of information and information systems to be included in each category of potential security impact.

As the State moves toward cloud-based infrastructure, and we have less direct control over our network, it will be important to have mechanisms in place that help improve performance. The hub should provide intelligent caching that helps reduce latency across many different information systems within the cloud and on-premise data centers. The performance of data connections should be monitored, and caching implemented for services with slow response times.

---

[6] Documentation of the common vocabulary and ontology represented through a Universal Resource Identifier (URI). The URI is a form of Uniform Resource Locator (URL) that provides an accessible unique identifier for entities and attributes in a common vocabulary. The namespace represents the exact location of a data definition. URIs are also used for RESTful Web services.

Data Dictionary (DD)

The State maintains many types of database management systems and data warehousing solutions; however, data schemas are often not documented. This creates difficulty understanding and extracting data, and often increases the cost of data use. DD functionality should provide the ability to document existing data schemas in a standardized way across the enterprise to contribute to better understanding of data assets, and the ability to use data across disparate systems for improved decision making, reporting, and service delivery. DD functionality will also create the basic data inventory needed for data classification and configuration management.

Data Classification and Configuration Management (DCCM)

Data classification and configuration management is necessary for service-oriented and cloud infrastructure security, privacy, and audit logging to verify that data governance rules are being followed and are effectively securing data.

At a minimum, A DCCM microservice should allow the documentation of information system component inventories, configuration items, configurations and baseline configurations. It must also provide the ability for information system stewards to manage change over configurations, perform security impact analysis, monitor configurations and set roles that are relevant to managing configuration processes.[7] The system must offer the ability to index data elements as configuration items; set configurations and baseline configurations according to data classifications; and to link component inventories, configuration items, configurations and baseline configurations to common vocabulary terms though the data hub.

Configuration management functionality should allow the verification and security of configurations, and the logging of use in the system to substantiate that configuration rules are being upheld. Verification should flag discrepancies with any new item being added to the system, or with changes to the system. This could include verifying that software versions and patches are up to date, or that unauthorized software is not installed. Configurations allow vulnerabilities to be easily identified and remedied. Audit logs substantiate that configuration rules are being upheld and provide the documentation of traffic through the system according to those rules.

In response to this RFI, please list all areas in which your solution is federally compliant. Examples could include but are not limited to: NIST-SP 800-60 *Guide for Mapping Types of Information and Information Systems to Security Categories*; NIST-SP 800-128 *Guide for Security-Focused Configuration Management of Information Systems*; or *Centers for Medicare & Medicaid Services Information Security and Privacy Group, Risk Management Handbook.*

Data Catalog (DC)

Data catalog functionality will make it easier to browse and search for data using metadata. A semantically enabled data catalog should provide domain-level documentation only. This means that the schema and location of the data is not exposed, only the definitions of the data domains. This allows all data sources to be documented, so users know what could potentially be available, to which user profiles, and in which formats, and under which classifications and access rules. The data catalog provides documentation of the States data portfolio. The data catalog should also provide a standardized API interface that allows secure and private access to data based on data classification and related profiles.

Master Data Management (MDM)

MDM is at the core of the State's data and digital asset management strategy. MDM should provide the ability to cleans and transform data according to semantic data domains; store standardized data; store master data IDs in association with the originating IDs to allow information systems to pull back the master record. MDM should also provide robust API functionality for all push and pull operations.

Data Warehousing (DW)

Among other transformations, the semantic layer will help transform data for loading into an enterprise data warehouse. Common vocabularies will allow data to be mapped to dimensions, attributes and measures, and will provide automated documentation of the data warehouse structure for use in analytics, visualizations, reporting and

---

[7] NIST SP 800-128, *Guide for Security-Focused Configuration Management of Information Systems*, October 2007.

performance monitoring. DW functionality must allow loading from an API. Data will be warehoused only when required for specific uses such as longitudinal studies and for master data management.

Business Rules, Process, Decision and Case Management (BRPDCM)

Business rules and processes are ever changing and represent a significant cost center within information system design and maintenance. Government business rules and processes are subject to changes brought about by statute, federal regulation, administrative code, policy and procedure, and lawsuits. Given the ever-changing nature of BRPDCM, it makes sense to decouple these concerns from the application layer. BRPDCM should provide the ability for business analysts to record business rules and process groups, then define the processes and decision points necessary to automate workflows and provide case management. APIs should provide access to all elements for use in the application layer.

<u>Important Notice</u>

DOA does not intend to award a contract from this RFI, nor will DOA be financially responsible for the preparation, or administrative costs incurred in responses to this RFI. All costs associated with responding will be solely at the interested party's expense. Not responding to this RFI does not preclude participation in any future request for proposal (RFP), if any is issued. DOA may or may not choose to meet with potential offerors to get further clarification of potential capability to meet requirements.

RESPONSE INFORMATION

Interested parties must submit a response via email to the following email address: doa.dgs.info@alaska.gov by **Thursday, November 21, 2019, at 4:00 p.m. ADT** directed to the Procurement Officer, Jennifer Krick.

All questions must be directed to the Procurement Officer via email.

Department of Administration
Shared Services of Alaska
Attention: Jennifer Krick
P.O. Box 110210
Juneau, Alaska 99811-0210
Phone: 907-465-1066
Questions Email: Jennifer.Krick@alaska.gov

Please note the State does not accept responsibility for failed e-mailed response deliveries. It is the responsibility of the interested party to follow up with the individual listed above to ensure your response was received prior to the deadline specified above.

Questions about this RFI can be directed to Jennifer Krick at Jennifer.Krick@alaska.gov.