|  |  |  |
| --- | --- | --- |
| ETF_logo_large | STATE OF WISCONSIN **Department of Employee Trust Funds**  **Robert J. Conlin**  SECRETARY | 4822 Madison Yards Way  Madison, WI 53705-9100  P. O. Box 7931  Madison, WI 53707-7931  http://etf.wi.gov |

Date: May 20, 2020

To: All Potential Responders to ETF RFI ETJ0058

RE: **Request for Information (RFI) ETJ0058 – Master Data Management (Part 2)**

Vendor Questions and Department Answers

|  |  |  |
| --- | --- | --- |
| Q # | Question | Department Answer |
| Q1 | Where does the State plan on hosting this solution? On the cloud or on premise? | We will consider both options. We encourage vendors to provide a response that includes a recommendation that deploys their solution in the most strategic and effective manner. |
| Q2 | If on the cloud, which cloud provider? | We have not chosen a cloud provider. We encourage vendors to recommend a provider as part of their response. |
| Q3 | Do you have a containers strategy in place? A container is basically a little boxed off area where developers can write code and deploy applications without the interruption of other parts of the IT environment. Containers run on kubernetes, leveraging an open source platform. | No, ETF does not have a container strategy in place at this time. The development of one is on the roadmap for next calendar year. |
| Q4 | If you are hosting this on the cloud, who is going to be responsible for hosting? The State or the vendor? | We will consider both scenarios and encourage vendors to provide a recommendation based on what is best for their solution in their response. |
| Q5 | FC57 – Please list the arithmetic and logic functions supported by the application. Please also list any known unsupported functions.  Are you referring to rules applied to data on imports and such? | It could be rules applied during the import process, during the matching/mastering process (less likely but possible), or when data is put into the space where other systems will consume it. |
| Q6 | FC71- Please describe your data dictionary capabilities.  What are you trying to achieve with the data dictionary. | We would like to maintain a standard data dictionary within the tool with a list of the key data elements, a definition of them, and a link to the lineage of the data. |
| Q7 | What is the deployment preference? (On-Premise, Cloud, or Hybrid) | We will consider all options. We encourage vendors to provide a response that includes a recommendation that deploys their solution in the most strategic and effective manner. |
| Q8 | If Cloud, is there a preferred vendor? (AWS, Google, Azure) | See Department Answer for Q2 |
| Q9 | What is the license model of preference? (Subscription or Perpetual) | We have not chosen a licensing model. We encourage vendors to recommend a model as part of their response. |
| Q10 | The RFI indicates an Individual/Member Domain and an Employer/Organization Domain. Will other Domains like Brokers be required? | Other domains may be managed within the system as well including reference data. The initial focus will be on person and organization related data. |
| Q11 | What volumes will be required to support the Individual/Member Domain? Assuming 622,000 unique employees for year 1, what is your expected growth rate in unique members per year? | Approximately 750,000+ person records as ETF needs to maintain records past a person’s lifespan. The expected growth rate is relatively low in the 1-2% range. In addition, we will have beneficiary and dependent ‘people’ records as well but they will be attached to and mastered as part of the member record. |
| Q12 | What volumes will be required to support the Employer/Organization Domain? Can we assume a not to exceed 1,500 unique entities? | We have approximately 1,500 employers that ETF interacts with. ETF also interacts with around 20 Third Party Administrators and data associated with them. Other organization data (trustees, charities, etc.) will eventually pass through the MDM solution as well. |
| Q13 | Is there a preference for Cloud or on-premise deployment? | See Department Answer for Q7 |
| Q14 | If Cloud deployment is preferred, is there a preferred Cloud vendor? | See Department Answer for Q2 |
| Q15 | How many years should be included for recurring costs? | Please include 5 years of recurring costs. |
| Q16 | What is the current state of the ETF Data center to support HA and DR? What is the desired end state? Are there any RTO and RPO targets? | ETF’s data center services are provided by the Department of Administration – Division of Enterprise Technology. The data centers (primary and secondary) can support HA and DR functionality. The desired end state is to have a reliable, scalable solution and we will consider both on-premise and hosted options to accomplish that. The RTO is 8 hours and the RPO is 24 hours. |
| Q17 | How many source systems will provide member and provider data for mastering? | Initially there will be relatively few (less than a dozen) source systems providing data but over time the number will grow and approach 100. |
| Q18 | How many systems will consume the mastered data? | This will change over time and will range from a couple when initially implemented to potentially dozens over time. |
| Q19 | What are the master data domain(s) that will be supported by this new solution? | Person and Organization will be the first two domains that will be mastered within the solution. Other domains may follow in the future. |
| Q20 | If “customer” is one of the domains, is ETF interested in creating a customer 360 view so that attributes of any individual can be tracked across all programs they participate in? | Yes, this is something we may be interested in doing at some point. |
| Q21 | What are specific requirements driving the need for this solution, at this time? | ETF needs to improve the quality of the data it uses to make business decisions. Having a system in place to govern and manage data that allows the creation of a ‘golden record’ or single source of the truth will build trust in the data and allow ETF staff to make decisions more effectively. |
| Q22 | What process, analytic, or visibility improvements is ETF hoping to recognize with this solution? Can you offer an example? | ETF is expecting to improve the quality of the data it uses to make business decisions. It also plans to create a ‘golden record’ or single source of the truth for mastered data domains. The data from the MDM solution will be used to feed a data warehouse which will be used for ETF analytical purposes. We expect issues with data will be made more visible and transparent with this solution so they can be found and resolved more quickly leading to better service to ETF’s customers. One example would be a person’s address. It may be accurate in the last system that was updated with that information, but other systems may not have the same address. |
| Q23 | Will this solution be replacing any existing systems or processes? | There is an existing internally built data quality management solution it will be replacing. |
| Q24 | Please provide an overview of your current technology landscape and relevant information, i.e. number of data sources (internal and external), level of duplication, consumption requirements, etc. | The current technology landscape is primarily centered around the use custom developed Java applications that utilize DB2 as the back-end database. However, that will be changing over the next few years as those systems are replaced by COTS solutions. |
| Q25 | What is the approximate total volume of master data (at least for the initial domain)? What is the expected volume for on-going or daily changes? | In terms of records there will be around 675,000 person records and 1,500 employer records that will be part of the initial implementation. The size in bytes will be dependent on the amount of data we end up storing and how the system stores the information. There is not a significant amount of new data coming in on a daily/weekly basis but at certain times of the year significant volumes of data come through where most person and employer records are updated. |
| Q26 | Does ETF have a tool in use currently for data cleansing and standardization? | There is an internal tool that monitors data quality but it does not do any cleansing or standardization. |
| Q27 | Can you please share the scoring method for this RFI so that we can understand which aspects of the solution WI ETF feels are most important? | ETF will be scoring the RFI based on criteria it has created. All aspects of the RFI are important to ETF. |
| Q28 | Does ETF intend to follow this RFI with a formal RFP? If so, will that RFP be posted on Wisconsin’s eSupplier portal, or will it be distributed only to companies that respond to this RFI? | ETF will determine next steps after reviewing the RFI. Those next steps could include purchasing a solution off of contracts ETF has available, performing an RFP, or another path. |
| Q29 | With regard to pricing, is ETF looking for a general range estimate; perhaps based on projects of similar size and scope? For more specific pricing, please provide information about each of the implementation domains, number of data sources (internal and external), data volumes, and current data quality level. | A relatively narrow price range is an acceptable response based on the vendors experience with other organizations and projects of similar size and scope. The initial implementation will be focused on the person and organization data domains but other domains may follow. While the initial number of data sources will most likely be less than a dozen there may be up to a hundred data sources (both internal and external) that are included over time. The current data quality is reasonable but there are issues with consistency between systems and a lack of a single source of the truth. There is an existing system that monitors data quality and provides information for people to utilize to improve data but it has limitations. |
| Q30 | Who is the primary audience of the questionnaire - so we know how best to phrase our answers? | Data managers, business analysts, data stewards, security staff, and technology staff. |
| Q31 | FC20 and more  How does the solution support initial matching and de-duplication of records in bulk?  How detailed of an explanation are you looking for? | As detailed as you are willing to provide. Please clarify how the solution does this. Using examples related to how it works with people data would provide the best context for the answer in evaluating how it will work for ETF. |
| Q32 | Consistency: Data is the same when comparing two or more representations of a thing against a definition.  We are wondering if you are talking about Matching rules and/or profiling. | This is primarily a method to check the quality of data coming from multiple systems and verifying that data is the same across them. For example, we want to be able to check the consistency of a member’s birth date based on data received from multiple systems. |
| Q33 | Reconciliation: Data needs to align to verify components equal the whole. Typically, this verifies that total row amounts must add up to stored amount for all rows  We understand that you have 622,000 members. What we are looking for is what we call base objects, which is the term we use to define the total number of unique records? We are looking for the kinds of volumes are you are looking to Master? | We will have a unique record for each member and each organization. In addition, we will have beneficiary and dependent ‘people’ records as well but they will be attached to and mastered as part of the member record. |
| Q34 | Data Monitoring: Please clarify the level of issue tracking you are looking for. We have workflows that assign actions on data to specific personnel. | We want to be able to track data issues from origination to resolution. We want to be able to see who the issue is assigned to and what its status is throughout its lifecycle. |

This Addendum is available on ETF’s Website at <https://etf.wi.gov/node/16036> .