University Information Technology Services

Request for Information
For
Data Catalog Solution to Support Data Governance

ADDENDUM # 1
RFI L202002

**Please mark all proposal submission
Envelopes with the following information
Sealed RFI # L202002
Due on August 26, 2019 no later than 2:00pm MST**

Request for Clarification

1) What type of data does UA plan to store in the Data catalog?
   A: UA is not storing data in the data catalog. The catalog will be connected to a data platform and will provide insight into the data in the platform.

2) Are there going to be any security requirements around PII data (if applicable?)
   A: Yes, data security requirements will be applied in accordance with UA data security policies, as well as other relevant data security policies.

3) Is UA looking for a SaaS based Data catalog solution? Is UA Open for a custom developed solution as well?
   A: The UA is open to which ever type of solution provides the best value for the collection of features sought.

4) Is there any specific layout / design that the UA has in mind regarding display of various data sets based upon a user's profile accessing the catalog? If so, pls share.
   A: No.

5) Does UA wish to have a cloud-based solution for the data storage or want to use existing infrastructure (cloud or physical)?
   A: We are open to both options.

6) Is UA looking for LDAP sort of integration with the data catalog to control access of the data catalog limited to the users defined in the UA LDAP (or similar)?
   A: Ideally, yes. Either via LDAP or Shibboleth/SAML groups.

7) As Data Steward works on data quality issues with data owners, is UA expecting to have an option for data steward to be able to correct the data at source or will that be a responsibility of the data owners?
   A: Correcting the data will not be a role for a data steward under the data governance process. Responsibility for correcting data will fall on the existing party with that responsibility.

8) What is the number of data tables and fields that will need to be synced from Salesforce?
   A: We can't answer that with certainty, at this point. The connected data will expand over time but we're uncertain at what rate and breadth that expansion will occur.

9) What is the number of records that will be available through search?
   A: >50

10) Is UA open to an onsite/offshore model of development? If yes, are they any limits to the type of roles that can be offshore.
    A: The UA prefers onsite.

11) What is the desired system for the search interface?
A: There is no preferred system, at this time. This RFI process will elucidate the systems available and the features these systems will offer.

12) Please clarify if the contractor will be responsible for creating training videos or documentation for the data catalog solution.
A: Training will be a collaborative effort between the UA team and the vendor. It is expected that the vendor should be able to provide some training on use of the data catalog.

13) Do you have any existing data governance (processes, data dictionary/business glossary, organizational structures) or data quality processes in place?
A: Yes. And other processes are being developed in parallel.

14) How many data source systems would be providing data/metadata to the solution and consuming data/metadata from it? Are those only Salesforce, PeopleSoft, AWS S3 and Oracle DW as listed in the non-functional requirements?
A: Initially Salesforce and AWS with other system to potentially follow.

15) What is the expected number/volume of records that will be processed by the solution?
A: We can't answer that with certainty, at this point. The connected data will expand over time but we're uncertain at what rate and breadth that expansion will occur.

16) What is the expected number of Catalog items?
A: We can't answer that with certainty, at this point. The connected data will expand over time but we're uncertain at what rate and breadth that expansion will occur.

17) Will connectivity to Big Data sources be required? Which technologies? Processing within the cluster or pulling data outside?
A: Potentially, yes. However, specific technologies are not known at this time.

18) How many users do you expect to:
- *Work with Business Glossary and Catalog*
- *View DQ profiling/monitoring results*
- *Define/implement business rules*
- *Participate in data remediation (improvement)*

A: We do not have an estimated number of users at this time. However, we would like to see how pricing scales as users increase over time.

19) Will there be any sources of reference data? If yes, should this Reference Data be managed in Ataccama RDM (Reference Data Manager)?
A: Potentially, yes. We are open to the solution that best meets our requirements.

20) What is your preferred deployment scenario?
- *On-premise / Private cloud managed by the University of Arizona*
- *Platform as a Service*
- *Managed Solution*

A: We have no preference between these options.

21) We see the requirement for 1 hardcopy form. Our firm has strong environmentally friendly or green initiatives. We usually email a PDF copy. Would it be acceptable to submit a PDF response by email instead of hardcopy form?
A: The University of Arizona does not accept formal submissions via email at this time.

22) We also see the request for an electronic copy on flash drive. If it's an electronic copy, would an email PDF copy suffice?
A: The University of Arizona does not accept formal submissions via email at this time.

23) What are the current or legacy University of Arizona solutions for MDM, Data prep, ETL and Data Quality? What other data management tools do you have?
A: We are contracting with SSB for MDM. DataStage is our predominant ETL tool, although StreamSets is part of our toolset as well.

24) How many datasets across all the data sources?
A: We can't answer that with certainty, at this point. The connected data will expand over time but we're uncertain at what rate and breadth that expansion will occur.

25) How many Analysts?
A: We do not have an estimated number of users at this time. However, we would like to see how pricing scales as users increase over time.

26) How many Developers?
A: We do not have an estimated number of users at this time. However, we would like to see how pricing scales as users increase over time.

27) How many Data Stewards?
A: We do not have an estimated number of users at this time. However, we would like to see how pricing scales as users increase over time.

28) Total number of people using Data Catalog?
A: We do not have an estimated number of users at this time. However, we would like to see how pricing scales as users increase over time.

29) Data catalog by itself does not include any ETL, data quality or data prep. Do you require any of this functionality or do you currently have a solution for this? If so, what are the systems or tools in place today?
A: We do not require any of these functionalities to be a part of your response. We are contracting with SSB for MDM. DataStage is our predominant ETL tool, although StreamSets is part of our toolset as well.

END OF ADDENDUM