

Response to Bidders Queries - Power Information Technology Company (PITC)

Pre- Bid meeting held on 11-11-2022 at 11:00 A:M

Tender No. PITC/G-224(74)/10-2022 for "Supply of Software & Services (Data Analytics Product)"

Sr no.	Queries by bidders		Clarifications by PITC / (Responses by PITC)
1	General	Is there a Data Analytics platform in place? Please share the details of the current data analytics platform if it exists.	Multiple small scale dimensional models in place in the form of single tables containing aggregated data to fulfill a/one dashboard's
2		Hardware for this implementation should also be proposed or it is not part of this	Hardware sizing to be proposed, but price quote not required. PITC will arrange for the
3		Is there any BI tool currently being used? If yes then please share details along with	No BI tool currently being used, however there are 9 dashboards (number of simpler reports
4		As part of hardware, is it expected to include all the required infrastructure	Yes
5		Is there any requirement of migration of data or reports from existing to new	Data migration from source systems is in scope. No existing DWH data to be brought in.
6		Is CDC required as part of this engagement?	No change data captures solution is required, data to be sourced from Oracle, SQL Server,
7		Please provide insights on the current pain points with respect to the existing platform implementation/features.	Architecture is not scalable anymore as it is developed in bits and pieces over years with new requirements coming in. Please share
8		Please provide any timeline expectations	
9	Sources & Ingestion	What is the total number of source systems to be covered under the scope of this project? Please provide the list of all sources that will be part of this engagement.	Type of databases are Oracle, SQL Server, My SQL, COBOL flat files and SAP HANA database. However it would be great to know how many different systems and how many different databases are we receiving the data from. Pleaes confirm how many Oracle DBs, how
10		Do you currently have any source system	No
11		What are the expected load windows for the data sizes of each sources?	Data will be loaded on at most once a day basis. Some data comes once a month and has
12		What is the total number of tables within each source system? Getting to know at least an upper limit number is important	
13		What will be the mode of availability of data from each source system, i.e. 1) pull or push mechanism and 2) through files or	
14		Can you please confirm the expected data	Daily at most or lesser.
15	Volume	Can you please confirm the expected data volume for each data source?	20 TB overall data
16		What is the estimated year-on-year organic growth of data for each source?	8-10% increase in data volume
17		What is the expected incremental data volume from each source i.e., Daily,	

18		Can you please share the daily transaction count that is expected to be processed	
19	Format	Please mention all the formats in which data is available?	structured data only being sourced from RDBMSes, flat files and columnar database of
20		Are there any other source system data types other than database, text files, and csv files? If yes please share.	No
21	Real Time	What are the specific use cases for real time data ingestion and downstream	No case for real time data.
22	Data Retention	What is the data retention volume and duration for each of the source?	NA
23	& Repository	If no data retention policy has been applied already, then what data retention is required for each source system?	5-10 years
24	Environment	Is there a requirement for multiple environments (Prod., Dev, Test, DR, etc.)?	Yes
25		Is High Availability required for production environment only?	Yes
26		Does the disaster recovery site need to have the same architecture, size and high Availability as the production site?	No
27	Reporting	What is the total number of reports expected to be created in this	150
28		Can you please help us understand the current inventory of operational and analytical reports?	The reports comprise of multiple aggregation at multiple hirearcy levels; however, successful bidder will be provided with detail structure of
29		What is the expected complexity breakdown of the total no of reports ? (Example: Complexity: Low 60 % , Medium 30 % , High 10%)	High
30		Are there any Ad-hoc reporting requirements to be catered as well?	Yes
31		Any tool of choice for Data Visualizations?	NA
32		What is the total possible number of KPIs to be created?	Will be discussed with successful bidder
33		What are the high level categories of the KPIs to be created?	Will be discussed with successful bidder
34		how many Admins, developers and business users would be required to access BI application? Is it possible to identify the numbers of users accessing	Initially during pilot project total 10 comprising of Admin, developer and business user accounts would be required; however, additional users in each categories will be
35		How many dashboards are required as an outcome of this engagement?	Dashboards will be discussed with the successful bidder to be designed, developed
36		What type of customization is required in dashboards/reports	High

37		What is the expected complexity breakdown of the total no of dashboard? (Example: Complexity: Low 60 % , Medium 30 % , High 10%)	High
38		Is there any requirement for self-service BI enablement?	Yes
39	Security	Is there any current security standard followed by the IT team which needs to	NA
40		Would basic security be required (Kerberos authentication, perimeter security, LDAP integration, InfoSec policy	NA
41		Would advanced security be required (encryption at rest and/or in motion, data masking, anonymization etc.)?	NA
42	AI/ML	IS there a requirement for AI/ML? If yes, please share the details of AI/ML use	No; however proposed solution must have capability for AI/ML
43	Access-Management	Is there a requirement for role-based access management?If yes, then for what purpose?	Yes
44	documentati	Can you provide clarity on the expected documentation?	Technical documentaion, operational manuals
45	Should the Solution be open source or licensed?		
46	IBS Data Source is structured or Not?		structured
47	No. Of Data sources?		Multiple
48	Frequency and Size of Data being loaded into the data		Will be shared with successful bidder
49	What is the main purpose of Using AI or ML in the Pro		Predictive capabilities
50	No. of Users?		Initially during pilot project total 10 comprising of Admin, developer and business user accounts would be required; however, additional users in each categories will be
51	Do you have an LDAP System Deployed? Are your		NA