

F. No. 14001/1/2019-Stats.
Government of India
Ministry of Corporate Affairs
(Statistics Division)

1st Floor, Jeevan Vihar Building,
No. 3, Parliament Street, New Delhi-110001,
Dated 23.05.2019

OFFICE MEMORANDUM

The undersigned is directed to circulate the summary record of Pre-Bid discussions on Request for Proposal for selection of DW/BI vendor for Corporate Data Management Project held on 17.05.2019 and 22.05.2019, at Conference Hall, 1st Floor, Jeevan Vihar Building, Sansad Marg.

2. This issues with the approval of the competent authority.

(Manish Joon)
Assistant Director

To

All participants.

Copy to :

1. Sr. PPS to Secretary/Additional Secretary, MCA.
2. PS to Joint Secretary (e-Gov.), MCA.
3. PS to DDG, MCA.

Enclosures : As stated (5 Pages).

Summary record of Pre-Bid discussion on Request for Proposal for Corporate Data Management Project.

A Pre-Bid Meeting on the Request for Proposal (RFP) for DW/BI Component of the Corporate Data Management Project was held on 17.05.2019 and 22.05.2019.

2. List of Participants is enclosed.
3. A brief of the CDM Project was given by DD(EN) highlighting the CDM Data Portal, Enforcement Module, Monitoring Module, CSR Module and Data Dissemination Module. It was informed to the participants that the DW/BI vendor would be responsible for maintaining and further enhancement of the pre-existing DW/BI setup. CDM now handles Corporate filing Data from F.Y. 2005-06 onwards, which has to be kept up to date with the latest filings. The DW/BI setup is currently based on SAP Software.
4. With reference to the major contents of the requirements stated in the RFP document, the following major points were discussed at length:
 - a. The DW/BI service providers would be responsible for maintaining the Corporate Data in Warehouse mode, by processing the e-forms (XBRL and Non-XBRL) filed by companies from time to time.
 - b. The DW/BI vendor has to have capacity to process unstructured data (like scanned documents) received as attachments to Annual filings of the companies and other sources such as the PMO.

- c. The DW/BI Service provider should have capacity to use Machine Learning and Artificial Intelligence algorithms for analyzing corporate data from the backend/Database level itself and use machine driven codes to flag non-compliance by companies. This feature will e required to be rolled out from 01.09.2019 after stabilization of the process involved.
 - d. The DW/BI Service Provider is expected to have worked on Datalakes, as going in future, the service provider should have the capacity of transitioning current Data Base (SAP/Sybase) to Data lake architecture without interruption/dislocation of the current working model.
 - e. The DW/BI Service Provider should be able to provision seamless transmission of data through APIs to various stakeholders, as may be decided by the competent authority.
5. After the walk through of RFP, queries/questions from bidders were taken up and answered. These have been listed at Annexure-I.
6. Whiles going through the RFP, in the meeting with the prospective bidders a typographical error (in the specimen/format) was noticed. Although, it is not material to the technical and/or administrative requirements stated in the RFP, by way of abundant caution and in the interest of fairness, transparency and competitiveness, it was decided to extend the period of submitting bids by a further period of 7 days.

7. It was decided that a copy of these minutes be also sent by e-mail to the various participants besides being hosted on MCA website with a 'What is new' tag.

8. The meeting ended with a vote of thanks to the Chair.

List of Participants

1. Shri E. Nagachandran, Deputy Director, MCA.
2. Shri Manish Joon, Assistant Director, MCA.
3. Shri Rahul Tiwari, Chief Consultant, CDM.
4. Shri Vikas Singh, Senior Consultant, CDM.
5. Ms. Snita Verma, Consultant, CDM.

List of Participants from Vendor:

Prospective Bidder: Object Technology Solutions India Pvt. Ltd. (17.05.2019)

1. Shri Rakesh Dubey.
2. Ms. Anupma Yaddav.

Prospective Bidder: Logicsoft International Pvt. Ltd. (22.05.2019)

1. Shri Rishi Bhaskar.

Annexure-I

S. No	Page No.	Clause	Query	CDM-Response
1	8	'Project Scope migrate current data-warehouse architecture to versatile data-lake architecture'	<p>What is the existing data size?</p> <p>Who would provide the infra?</p> <p>What is the data growth rate?</p> <p>How is the cluster monitoring?</p>	<p>1. Raw Data is available in multiple formats running in more than 100 TB. Processed Data during the entire project cycle would be within 5 TB.</p> <p>2. Infrastructure would be provisioned by MCA.</p> <p>3. Project is in evolution stage hence incremental data is variable but approximately 1 TB a month.</p> <p>4. DW/BI service provider in-sync with Ministry Officials for the entire project duration.</p>
2	6	'MCA - CDM must be capable of processing unstructured data'	<p>What is the processing expected over these files?</p> <p>Is it insert and retrieval associated with a tag?</p>	<p>1. Would depend on filing to filing basis and associated requirement.</p> <p>2. May also include tag based as well.</p>
3	8	'Technical Requirements Extraction of data from scanned pdf (text content) files, whether as attachment to e-form or otherwise, and to have them stored in an appropriate table in RDBMS'	<p>Extraction of scanned tables cannot be done as it is into RDBMS.</p> <p>However, we can extract the content and put it in a table as a single row or text.</p>	<p>Extraction needs to be done into RDBMS.</p>
4	9	'knowledge and understanding of data science technologies'	<p>Any technology preference?</p> <p>What is the expected outcome?</p>	<p>No Specific Technology Preference, to be finalized for enhancements into the existing system and as detailed in the RFP.</p>

5	9	'understanding and worked on block-chain technologies'	Exact use case/ need for block chain?	It will be used for enhancing security features as the data will be used in a statutory framework and has legal ramifications. The exact mode of roll-out will be finalized as per directions of the competent authority. Likely date of roll-out is 01.09.2019.
6	9	'Developing Reports from SAP IQ (Sybase Database) using BODS as the extraction tool Information Designer Tool and SAP Business Objects.'	If there is data lake then why IQ and what is the expected outcome from Data lake	The work has to be in lines of the existing system on Sybase IQ. The migration to Datalakes architecture is expected to take place along with MCA21 V3, i.e., around the time of end of the current tenure of this project. The migration, if and when required to be done, has to be seamless without affecting or disrupting the working system.