



**REQUEST FOR EXPRESSION OF INTEREST (EOI)
FOR PROCUREMENT OF
INTEGRATED SOLUTION FOR DATA LAKE AND DATA WAREHOUSE**

Ref: SBI/GITC/Data Warehouse/2023/2024/48

dated 08.12.2023

Data Warehouse Department

Global IT Centre

Sector -11, CBD Belapur

Navi Mumbai – 400614

(Maharashtra)

Schedule of Events

SI No	Particulars	Remarks
1	Contact details of issuing department (Name, Designation, Mobile No., Email and office address for sending any kind of correspondence regarding this EOI)	Name: Shailendra Kumar Uniyal Designation: Dy, General Manager Email ID: rpfordwh@sbi.co.in Contact Address: Data Warehouse Department, Tower no. 7, 4th floor, Railway Station Building, State Bank of India, Global IT Centre, Sector 11, C.B.D. Belapur, Navi Mumbai-400614 Contact Number: Sanjay Nandanwar - 9167397872
2	Bid Document Availability including changes/amendments, if any to be issued	EOI may be downloaded from Bank's website https://www.sbi.co.in procurement news from <u>08.12.2023</u> to <u>10.1.2024</u>
3	Last date for requesting clarification	22.12.2023 (5:00 PM) All communications regarding points / queries requiring clarifications shall be given in writing or by e-mail.



4	Pre - bid Meeting at (venue will be advised in due course)	29.12.2023						
5	Clarifications to queries raised at pre-bid meeting will be provided by the Bank.	02.01.2024						
6	Last date and time for Bid submission	10.01.2024 (5:00 PM)						
7	Address for submission of Bids	To be submitted online through the e-tendering Website as under: https://etender.sbi/SBI/						
8	Date and Time of opening of Technical Bids	10.01.2024 (5:30 PM) Authorized representatives of Bidders may be present online during opening of the Technical Bids. However, Technical Bids would be opened even in the absence of any or all of Bidders representatives.						
11	Tender Fee	NIL						
12	Contact details of e-Procurement agency appointed for e-procurement	M/s E-Procurement Technologies Ltd, Ahmedabad Website: https://etender.sbi/SBI/						
		<table border="1"> <thead> <tr> <th>Name</th> <th>Phone Number</th> <th>Email Id</th> </tr> </thead> <tbody> <tr> <td>Fahad Khan</td> <td>6352631766</td> <td>fahad@eptl.in</td> </tr> </tbody> </table>	Name	Phone Number	Email Id	Fahad Khan	6352631766	fahad@eptl.in
Name	Phone Number	Email Id						
Fahad Khan	6352631766	fahad@eptl.in						



		Shaikh Nasruddin	6352632098	shaikh@eptl.in
		Hiral Purohit	6352631968	hiral.purohit@eptl.in
		Mubassera Mansuri	7859800621	mubassera@eptl.in
		Jay Vyas	9265562819	jay.v@eptl.in



INDEX

PART-I6

1. INVITATION TO BID:..... 7

2. BACKGROUND:..... 8

3. DISCLAIMER:..... 9

4. DEFINITIONS:..... 10

5. ELIGIBILITY AND TECHNICAL CRITERIA/ SCOPE OF WORK: 10

6. COST OF BID DOCUMENT:..... 11

7. CLARIFICATION AND AMENDMENTS ON EOI/PRE-BID MEETING: 11

8. CONTENTS OF EOI DOCUMENT: 12

9. BID PREPARATION AND SUBMISSION:..... 13

10. DEADLINE FOR SUBMISSION OF BIDS: 15

11. MODIFICATION AND WITHDRAWAL OF BIDS: 15

12. BIDDING PROCESS/OPENING OF TECHNICAL BIDS: 15

13. CONFLICT OF INTEREST: 16

14. CODE OF INTEGRITY AND DEBARMENT/BANNING: 18

15. GOVERNING LANGUAGE:..... 22

16. APPLICABLE LAW:..... 22

17. TENDER FEE:..... 23

18. NOTICES:..... 24

19. OTHER TERMS & CONDITIONS..... 24

PART-II26

Appendix-A: BID FORM (TECHNICAL BID) 27

Appendix-B: Bidder’s Eligibility Criteria 30

Appendix-C: Technical Criteria/Scope of Work..... 33

Appendix-D: Bidder Details..... 100

Appendix-E: Pre-Bid Query Format 102

Appendix-F: Format for Submission of Client References 103



PART-I

1. INVITATION TO BID:

- i. **State Bank of India** (herein after referred to as '**SBI/the Bank**'), having its Corporate Centre at Mumbai, various other offices (LHOs/ Head Offices /Zonal Offices/Global Link Services, Global IT Centre, foreign offices etc.) of State Bank of India, branches/other offices, Subsidiaries and Joint Ventures available at various locations and managed by the Bank (collectively referred to as **State Bank Group or 'SBG'** hereinafter). This Request for EOI is issued by **the Bank** on behalf of **SBG** for procurement of Integrated Solution for Data Lake and Data Warehouse (ISDLDW).
- ii. In order to meet the IT requirements, the Bank proposes to invite online EOI from eligible Bidders wishing to be considered for short listing for participating in bidding process of the proposed procurement.
- iii. Bidder shall mean any entity (i.e. juristic person) who meets the eligibility criteria given in **Appendix-B** of this EOI and willing to provide the product and services required by the Bank through RFP process in pursuant to this EOI. The interested Bidders who agree to all the terms and conditions contained in this EOI may submit their Bids with the information desired in this EOI. Consortium bidding is not permitted under this EOI.
- iv. Address for submission of online Bids, contact details including email address for sending communications are given in Schedule of Events of this EOI.
- v. The objective of this EOI is to identify eligible Bidders who are interested in providing the required product and services and have suitable capacity, capability and experience.
- vii. This EOI document shall not be transferred, reproduced or otherwise used for purpose other than for which it is specifically issued.
- viii. Interested Bidders are advised to go through the entire EOI before submission of online Bids to avoid any chance of elimination. The criteria and the actual process of evaluation of the responses to this EOI and

subsequent selection of the successful Bidder will be entirely at Bank's discretion.

2. BACKGROUND:

Bank is keen on embarking on a Data Analytics journey and gain an advantage over its competitors in terms of innovation, time to market, user delight, cost of services despite huge data size and scale.

There is an increasing trend globally in large Banks to use customer data in real time / near real time from their mobile click-stream / touch-stream data, purchase patterns to make customized offers / improve service experience / support for advanced AI-ML models.

With data at SBI growing at an astronomical rate, a robust, scalable and cost-effective infrastructure for data storage and processing is the need of the hour. The ability to handle complex data i.e. semi-structured and unstructured data in addition to the conventional structured data processing, maximizes value potential and quality of services in the financial sector. Operational efficiency in the digital age is one of the factors that determine a Bank's growth. By managing petabytes of data efficiently, SBI has the potential to be a trend setter for smart data management globally. Empowering representatives of Bank at every level with analytics, thus standardizing and speeding up the decision-making process, will be a natural outcome of smart data management.

Data Warehouse has migrated on new infrastructure on 10th May 2023 as per five-year Agreement between Bank and M/s IBM India Private Limited in February 2020 for implementation and maintenance of the Data Warehouse. With the implementation of project, DWH has been shifted from PDOA (Pure data for Operation Analytics) technology to IIAS (IBM Integrated Analytical System) technology, along with a small-scale Data Lake and Data Virtualization capabilities. Data Lake is implemented with a view to enrich the capabilities of sourcing unstructured and semi-structured data for advance analytics with AI / ML tools.

Technical details including Scope of Work, project mandates, etc are furnished in Appendix - C

3. DISCLAIMER:

- i. The information contained in this EOI or information provided subsequently to Bidder(s) whether verbally or in documentary form/email by or on behalf of SBI, is subject to the terms and conditions set out in this EOI.
- ii. This EOI is not an offer by SBI, but an invitation to receive responses from the eligible Bidders.
- iii. The purpose of this EOI is to provide the Bidder(s) with information to assist preparation of their Bid proposals. This EOI does not claim to contain all the information each Bidder may require. Each Bidder should conduct its own investigations and analysis and should check the accuracy, reliability and completeness of the information contained in this EOI and where necessary obtain independent advices/clarifications. Bank may in its absolute discretion, but without being under any obligation to do so, update, amend or supplement the information in this EOI.
- iv. The Bank, its employees and advisors make no representation or warranty and shall have no liability to any person, including any Bidder under any law, statute, rules or regulations or tort, principles of restitution or unjust enrichment or otherwise for any loss, damages, cost or expense which may arise from or be incurred or suffered on account of anything contained in this EOI or otherwise, including the accuracy, adequacy, correctness, completeness or reliability of the EOI and any assessment, assumption, statement or information contained therein or deemed to form or arising in any way for participation in this EOI.
- v. The Bank also accepts no liability of any nature whether resulting from negligence or otherwise, howsoever caused arising from reliance of any Bidder upon the statements contained in this EOI.
- vi. The Bidder is expected to examine all instructions, forms, terms and specifications in this EOI. Failure to furnish all information required under

this EOI or to submit a Bid not substantially responsive to this EOI in all respect will be at the Bidder's risk and may result in rejection of the Bid.

- vii. The issue of this EOI does not imply that the Bank is bound to select a Bidder or subsequently to award the contract to the shortlisted Bidder, as the case may be, for the project and the Bank reserves the right to reject all or any of the Bids or Bidders without assigning any reason.

4. DEFINITIONS:

In this connection, the following terms shall be interpreted as indicated below:

- i. **"The Bank"** 'means the State Bank of India (including domestic branches and foreign offices), Subsidiaries and Joint Ventures, where the Bank has ownership of more than 50% of voting securities or the power to direct the management and policies of such Subsidiaries and Joint Ventures.
- ii. **"Bidder/Channel Partner"** means an eligible entity/firm submitting the Bid in response to this EOI.
- iii. **"Bid"** means the written reply or submission of response to this EOI.

5. ELIGIBILITY AND TECHNICAL CRITERIA/ SCOPE OF WORK:

- i. Bid is open to all Bidders who meet the eligibility and technical criteria/scope of work as given in **Appendix-B & Appendix-C** of this EOI. The Bidder has to submit the documents substantiating eligibility criteria as mentioned in this EOI document.
 - (a) If any Bidder submits Bid on behalf of Principal/OEM, the same Bidder shall not submit a Bid on behalf of another Principal/OEM under the EOI. Bid submitted with options of multiple OEMs shall also be considered Bid submitted on behalf of multiple OEMs.

(b) Either the Bidder on behalf of Principal/OEM or Principal/OEM itself is allowed to Bid, however both cannot Bid simultaneously.

6. COST OF BID DOCUMENT:

The participating Bidders shall bear all the costs associated with or relating to the preparation and submission of their Bids including but not limited to preparation, copying, postage, delivery fees, expenses associated with any demonstration or presentations which may be required by the Bank or any other costs incurred in connection with or relating to their Bid. The Bank shall not be liable in any manner whatsoever for the same or for any other costs or other expenses incurred by a Bidder regardless of the conduct or outcome of the bidding process.

7. CLARIFICATION AND AMENDMENTS ON EOI/PRE-BID MEETING:

- i. Bidder requiring any clarification on EOI may notify the Bank in writing strictly as per the format given in **Appendix-E** at the address/by e-mail within the date/time mentioned in the Schedule of Events.
- ii. All queries to be raised in the pre-bid meeting will relate to the EOI alone and no queries related to detailed analysis of scope of work, payment terms and mode of selection will be entertained. These issues will be amply clarified at the RFP stage.
- iii. A pre-Bid meeting will be held in person or online on the date and time specified in the Schedule of Events which may be attended by the authorized representatives of the Bidders interested to respond to this EOI.
- iv. The queries received (without identifying source of query) and response of the Bank thereof will be posted on the Bank's website or conveyed to the Bidders.
- v. The Bank reserves the right to amend, rescind or reissue the EOI, at any time prior to the deadline for submission of Bids. The Bank, for any reason,

whether, on its own initiative or in response to a clarification requested by a prospective Bidder, may modify the EOI, by amendment which will be made available to the Bidders by way of corrigendum/addendum. The interested parties/Bidders are advised to check the Bank's website regularly till the date of submission of Bid document specified in the Schedule of Events/email and ensure that clarifications/ amendments issued by the Bank, if any, have been taken into consideration before submitting the Bid. Such amendments/clarifications, if any, issued by the Bank will be binding on the participating Bidders. Bank will not take any responsibility for any such omissions by the Bidder. The Bank, at its own discretion, may extend the deadline for submission of Bids in order to allow prospective Bidders a reasonable time to prepare the Bid, for taking the amendment into account. Nothing in this EOI or any addenda/corrigenda or clarifications issued in connection thereto is intended to relieve Bidders from forming their own opinions and conclusions in respect of the matters addresses in this EOI or any addenda/corrigenda or clarifications issued in connection thereto.

- vi. Queries received after the scheduled date and time will not be responded/acted upon.

8. CONTENTS OF EOI DOCUMENT:

- i. The Bidder must thoroughly study/analyse and properly understand the contents of this EOI, its meaning and impact of the information contained therein.
- ii. Failure to furnish all information required in this EOI or submission of Bid not responsive to this EOI in any respect will be at the Bidder's risk and responsibility and the same may finally result in rejection of its Bid. The Bank has made considerable effort to ensure that accurate information is contained in this EOI and is supplied solely as guidelines for Bidders.
- iii. The Bid prepared by the Bidder, as well as all correspondences and documents relating to the Bid exchanged by the Bidder and the Bank and supporting documents and printed literature shall be submitted in English.

- iv. The information provided by the Bidders in response to this EOI will become the property of the Bank and will not be returned. Incomplete information in Bid document may lead to non-consideration of the proposal.

9. BID PREPARATION AND SUBMISSION:

- i. The Bid is to be submitted on portal of e-Procurement agency for EOI of **Integrated Solution for Data Lake and Data Warehouse** in response to the EOI No. **SBI/GITC/Data Warehouse/2023/2024/48** dated **08.12.2023**. Documents mentioned below are to be uploaded on portal of e-Procurement agency with digital signature of authorised signatory:
- (a) Index of all the documents, letters, bid forms etc. submitted in response to EOI along with page numbers.
 - (b) Bid covering letter/Bid form on the lines of **Appendix-A** on Bidder's letter head.
 - (c) Proof of remittance of Tender Fee as specified in this document.
 - (d) Specific response with supporting documents in respect of Eligibility Criteria as mentioned in **Appendix-B** and technical criteria/scope of work on the lines of **Appendix-C**.
 - (e) Bidder's details as per **Appendix-D** on Bidder's letter head.
 - (f) Audited financial statement and profit and loss account statement as mentioned in Part-II.
 - (g) A copy of board resolution along with copy of power of attorney (POA wherever applicable) showing that the signatory has been duly authorized to sign the Bid document.
 - (h) If applicable, copy of registration certificate issued by competent authority as mentioned in **SI No 2 of Eligibility Criteria** under **Appendix-B**.

ii. Bidders may please note:

- (a) While submitting the Technical Bid, literature on the proposed solution/services should be segregated and kept together in one section.
- (b) The Bid document shall be complete in accordance with various clauses of the EOI document or any addenda/corrigenda or clarifications issued in connection thereto, duly signed by the authorized representative of the Bidder. Board resolution authorizing representative to Bid and make commitments on behalf of the Bidder is to be attached.
- (c) It is mandatory for all the Bidders to have class-III Digital Signature Certificate (DSC) (in the name of person who will sign the Bid) from any of the licensed certifying agency to participate in this EOI. DSC should be in the name of the authorized signatory. It should be in corporate capacity (that is in Bidder capacity).
- (d) If deemed necessary, the Bank may seek clarifications on any aspect from the Bidder. However, that would not entitle the Bidder to change or cause any change in the substances of the Bid already submitted.
- (e) The Bidder may also be asked to give presentation for the purpose of clarification of the Bid.
- (f) The Bidder must provide specific and factual replies to the points raised in the EOI.
- (g) The Bid shall be typed or written and shall be digitally signed by the Bidder or a person or persons duly authorized to bind the Bidder to the Contract.
- (h) All the enclosures (Bid submission) shall be serially numbered.
- (i) The Bank reserves the right to reject Bids not conforming to above.
- (j) The Bid document shall be complete in accordance with various clauses of the RFP document or any addenda/corrigenda or clarifications issued in connection thereto, duly signed by the authorized representative of the Bidder. Board resolution authorizing representative to Bid and make commitments on behalf of the Bidder is to be attached.
- (k) Bidder(s) should prepare and submit their online Bids well in advance before the prescribed date and time to avoid any delay or problem during the bid

submission process. The Bank shall not be held responsible for any sort of delay or the difficulties faced by the Bidder(s) during the submission of online Bids.

- (1) Bidder(s) should ensure that the Bid documents submitted should be free from virus and if the documents could not be opened, due to virus or otherwise, during Bid opening, the Bid is liable to be rejected.

10. DEADLINE FOR SUBMISSION OF BIDS:

- i. Bids must be submitted online on portal of e-Procurement agency by the date and time mentioned in the “Schedule of Events”.
- ii. In the event of the specified date for submission of Bids being declared a holiday for the Bank, the Bids will be received upto the appointed time on the next working day.
- iii. In case the Bank extends the scheduled date of submission of Bid document, the Bids shall be submitted by the time and date rescheduled. All rights and obligations of the Bank and Bidders will remain the same.

11. MODIFICATION AND WITHDRAWAL OF BIDS:

- i. The Bidder may modify or withdraw its Bid after the Bid’s submission, provided that modification, including substitution or withdrawal of the Bids, is received on e-procurement portal, prior to the deadline prescribed for submission of Bids.
- ii. No modification in the Bid shall be allowed, after the deadline for submission of Bids.

12. BIDDING PROCESS/OPENING OF TECHNICAL BIDS:

- i. All Bids received up to the specified time and date will be opened for initial evaluation on the time and date mentioned in the schedule of events. The Bids will be opened in the presence of representatives of the Bidders who choose to attend the same on portal of e-Procurement agency. However,

Bids may be opened even in the absence of representatives of one or more of the Bidders.

- ii. Prior to the detailed evaluation, the Bank will determine the responsiveness of each Bid to the EOI. For purposes of these Clauses, a responsive Bid is one, which conforms to all the terms and conditions of the EOI in toto, without any deviation.
- iii. If a Bid is not responsive, it will be rejected by the Bank and will not subsequently be made responsive by the Bidder by correction of the non-conformity.
- iv. After examining the EOI, some or all of the Bidders may be asked to make presentation of the solution and demonstrate proof of concept (POC). The Bidder shall bear all costs associated with submission of EOI, presentation/POC desired by the Bank. The Bank shall not be responsible or liable for any cost thereof, regardless of the conduct or outcome of the process.
- v. SBI may shortlist the Bidders who fulfill the eligibility and technical criteria specified in this EOI, have solution/services as per the requirement of the Bank and are agreeing to abide by the terms and conditions of the Bank. Bank's judgment in this regard will be final.
- vi. Bank may issue a Request for Proposal (RFP) to shortlisted Bidder for next process of procurement. However, please note that short listing of Bidders should not be treated as a contract for the proposed work.
- vii. Nothing contained in this EOI shall impair the Bank's Right to issue 'Open RFP' on the proposed solution/services.
- viii. Bidders will be advised about shortlisting of their EOIs or otherwise. However, Bidders will not be provided with information about comparative position of their EOIs with that of others.

13. CONFLICT OF INTEREST:

- i. Bidder shall not have a conflict of interest (the "Conflict of Interest") that affects the bidding Process. Any Bidder found to have a Conflict of Interest shall be disqualified without prejudice to any other right or remedy that may

be available to the Bank under the EOI and/ or the subsequent RFP or otherwise.

ii. Without limiting the generality of the above, a Bidder shall be deemed to have a Conflict of Interest affecting the bidding Process, if:

(a) the Bidder, its Member or Associate (or any constituent thereof) and any other Bidder, its Member or any Associate thereof (or any constituent thereof) have common controlling shareholders or other ownership interest; provided that this disqualification shall not apply in cases where the direct or indirect shareholding of a Bidder, its Member or an Associate thereof (or any shareholder thereof having a shareholding of more than 5% (five per cent) of the paid up and subscribed share capital of such Bidder, Member or Associate, as the case may be) in the other Bidder, its Member or Associate, has less than 5% (five per cent) of the subscribed and paid up equity share capital thereof; provided further that this disqualification shall not apply to any ownership by a bank, insurance company, pension fund or a public financial institution referred to in section 2(72) of the Companies Act, 2013. For the purposes of this Clause, indirect shareholding held through one or more intermediate persons shall be computed as follows: (aa) where any intermediary is controlled by a person through management control or otherwise, the entire shareholding held by such controlled intermediary in any other person (the "Subject Person") shall be taken into account for computing the shareholding of such controlling person in the Subject Person; and (bb) subject always to sub-clause (aa) above, where a person does not exercise control over an intermediary, which has shareholding in the Subject Person, the computation of indirect shareholding of such person in the Subject Person shall be undertaken on a proportionate basis; provided, however, that no such shareholding shall be reckoned under this sub-clause (bb) if the shareholding of such person in the intermediary is less than 26% of the subscribed and paid up equity shareholding of such intermediary; or



- (b) a constituent of such Bidder is also a constituent of another Bidder; or
 - (c) such Bidder, its Member or any Associate thereof receives or has received any direct or indirect subsidy, grant, concessional loan or subordinated debt from any other Bidder, its Member or Associate, or has provided any such subsidy, grant, concessional loan or subordinated debt to any other Bidder, its Member or any Associate thereof; or
 - (d) such Bidder has the same legal representative for purposes of this Bid as any other Bidder; or
 - (e) such Bidder, or any Associate thereof, has a relationship with another Bidder, or any Associate thereof, directly or through common third party/parties, that puts either or both of them in a position to have access to each other's information about, or to influence the Bid of either or each other; or
 - (f) such Bidder or any of its affiliates thereof has participated as a consultant to the Bank in the preparation of any documents, design or technical specifications of the EOI.
- iii. For the purposes of this EOI, Associate means, in relation to the Bidder, a person who controls, is controlled by, or is under the common control with such Bidder (the "Associate"). As used in this definition, the expression "control" means, with respect to a person which is a company or corporation, the ownership, directly or indirectly, of more than 50% (fifty per cent) of the voting shares of such person, and with respect to a person which is not a company or corporation, the power to direct the management and policies of such person by operation of law or by contract.

14. CODE OF INTEGRITY AND DEBARMENT/BANNING:

- i. The Bidder and their respective officers, employees, agents and advisers shall observe the highest standard of ethics during the bidding Process. Notwithstanding anything to the contrary contained herein, the Bank shall reject Bid without being liable in any manner whatsoever to the Bidder if it

determines that the Bidder has, directly or indirectly or through an agent, engaged in corrupt/fraudulent/coercive/undesirable or restrictive practices in the bidding Process.

- ii. Bidders are obliged under code of integrity to Suo-moto proactively declare any conflicts of interest (pre-existing or as and as soon as these arise at any stage) in EOI process or execution of contract. Failure to do so would amount to violation of this code of integrity.
- iii. Any Bidder needs to declare any previous transgressions of such a code of integrity with any entity in any country during the last three years or of being debarred by any other procuring entity. Failure to do so would amount to violation of this code of integrity.
- iv. For the purposes of this clause, the following terms shall have the meaning hereinafter, respectively assigned to them:
 - (a) “**corrupt practice**” means making offers, solicitation or acceptance of bribe, rewards or gifts or any material benefit, in exchange for an unfair advantage in the procurement process or to otherwise influence the procurement process or contract execution;
 - (b) “**Fraudulent practice**” means any omission or misrepresentation that may mislead or attempt to mislead so that financial or other benefits may be obtained or an obligation avoided. This includes making false declaration or providing false information for participation in a EOI process or to secure a contract or in execution of the contract;
 - (c) “**Coercive practice**” means harming or threatening to harm, persons or their property to influence their participation in the procurement process or affect the execution of a contract;
 - (d) “**Anti-competitive practice**” means any collusion, bid rigging or anti-competitive arrangement, or any other practice coming under the purview of the Competition Act, 2002, between two or more bidders, with or without the knowledge of the Bank, that may impair the transparency, fairness and the progress of the procurement process or to establish bid prices at artificial, non-competitive levels;

(e) **“Obstructive practice”** means materially impede the Bank’s or Government agencies investigation into allegations of one or more of the above mentioned prohibited practices either by deliberately destroying, falsifying, altering; or by concealing of evidence material to the investigation; or by making false statements to investigators and/or by threatening, harassing or intimidating any party to prevent it from disclosing its knowledge of matters relevant to the investigation or from pursuing the investigation; or by impeding the Bank’s rights of audit or access to information;

v. **Debarment/Banning**

Empanelment/participation of Bidders and their eligibility to participate in the Bank’s procurements is subject to compliance with code of integrity and performance in contracts as per terms and conditions of contracts. Following grades of debarment from empanelment/participation in the Bank’s procurement process shall be considered against delinquent Vendors/Bidders:

(a) **Holiday Listing (Temporary Debarment - suspension):**

Whenever a Bidder is found lacking in performance, in case of less frequent and less serious misdemeanors, the Bidders may be put on a holiday listing (temporary debarment) for a period upto 12 (twelve) months. When a Bidder is on the holiday listing, he is neither invited to bid nor are his bids considered for evaluation during the period of the holiday. The Bidder is, however, not removed from the list of empaneled vendors, if any. Performance issues which may justify holiday listing of the Vendor are:

- Vendors who have not responded to requests for quotation/tenders consecutively three times without furnishing valid reasons, if mandated in the empanelment contract (if applicable);
- Repeated non-performance or performance below specified standards (including after sales services and maintenance services etc.);

- Bidders undergoing process for removal from empanelment/participation in procurement process or banning/debarment may also be put on a holiday listing during such proceedings.

(b) Debarment from participation including removal from empanelled list

Debarment of a delinquent Bidder (including their related entities) for a period (one to two years) from the Bank's procurements including removal from empanelment, wherever such Bidder is empaneled, due to severe deficiencies in performance or other serious transgressions. Reasons which may justify debarment and/or removal of the Bidder from the list of empaneled vendors are:

- Without prejudice to the rights of the Bank under Clause 14(i) hereinabove, if a Bidder is found by the Bank to have directly or indirectly or through an agent, engaged or indulged in any corrupt/fraudulent/coercive/undesirable or restrictive practices during the bidding Process, such Bidder shall not be eligible to participate in any EOI/RFP issued by the Bank during a period of 2 (two) years from the date of debarment.
- The Bidder fails to abide by the terms and conditions or to maintain the required technical/operational staff/equipment or there is change in its production/service line affecting its performance adversely, or fails to cooperate or qualify in the review for empanelment;
- If Bidder ceases to exist or ceases to operate in the category of requirements for which it is empaneled;
- Bankruptcy or insolvency on the part of the Bidder as declared by a court of law; or
- Banning by Ministry/Department or any other Government agency;
- Other than in situations of force majeure, technically qualified Bidder withdraws from the procurement process or after being declared as successful bidder: (i) withdraws from the process; (ii) fails to enter into a

Contract; or (iii) fails to provide performance guarantee or any other document or security required in terms of the RFP documents;

- If the Central Bureau of Investigation/CVC/C&AG or Vigilance Department of the Bank or any other investigating agency recommends such a course in respect of a case under investigation;
- Employs a Government servant or the Bank's Officer within two years of his retirement, who has had business dealings with him in an official capacity before retirement; or
- Any other ground, based on which the Bank considers, that continuation of Contract is not in public interest.
- If there is strong justification for believing that the partners/directors/proprietor/agents of the firm/company has been guilty of violation of the code of integrity or Integrity Pact (wherever applicable), evasion or habitual default in payment of any tax levied by law; etc.

(c) Banning from Ministry/Country-wide procurements

For serious transgression of code of integrity, a delinquent Bidder (including their related entities) may be banned/debarred from participation in a procurement process of the Bank including procurement process of any procuring entity of Government of India for a period not exceeding three years commencing from the date of debarment.

15. GOVERNING LANGUAGE:

The governing language shall be English.

16. APPLICABLE LAW:

The Contract shall be interpreted in accordance with the laws of the Union of India and shall be subjected to the exclusive jurisdiction of courts at Mumbai.

Bidder shall abide by the provisions of the DPDP Act, 2023 - 11th August, 2023; CG-DL-E-12082023-248045 as and when the relevant rules and guidelines come into force.

17. TENDER FEE:

Non-refundable Tender Fee should be directly credited to the designated account as mentioned in Schedule of Events. Proof of remittance of Tender Fee in the designated account should be enclosed with the technical bid. The Bids without tender fee will not be considered valid.

18. EXEMPTION OF EMD AND TENDER FEE:

Micro & Small Enterprises (MSE) units and Start-ups* are exempted from payment of EMD and tender fee provided the products and/or services they are offering, are manufactured and/or services rendered by them. Exemption as stated above is not applicable for selling products and/or services, manufactured/ rendered by other companies.

Bidder should submit supporting documents issued by competent Govt. bodies to become eligible for the above exemption.

Bidders may please note:

i. NSIC certificate/ Udyog Aadhar Memorandum/Udyam Registration Certificate should cover the items tendered to get EMD/tender fee exemptions. Certificate/ Memorandum should be valid as on due date / extended due date for Bid submission.

ii. "Start-up" company should enclose the valid Certificate of Recognition issued by Department for Promotion of Industry and Internal Trade (DPIIT), (erstwhile Department of Industrial Policy and Promotion), Ministry of Commerce & Industry, Govt. of India with the technical bid.

iii. *Start-ups which are not under the category of MSE shall not be eligible for exemption of tender fee.

iv. Bidder who solely on its own, fulfils each eligibility criteria condition as per the RFP terms and conditions and who are having MSE or Start-up company status, can claim exemption for EMD/ tender fee.

v. If all these conditions are not fulfilled or supporting documents are not submitted with the technical Bid, then all those Bids without tender fees /EMD will be summarily rejected and no queries will be entertained.

19. NOTICES:

Any notice given by one party to the other pursuant to this EOI shall be sent to other party in writing and confirmed in writing to other Party's address. The notice shall be effective when delivered or on the notice's effective date whichever is later.

19. OTHER TERMS & CONDITIONS

- i. Lodgement of an EOI is evidence of a Bidder's consent to comply with the terms and condition of Request for EOI process and subsequent bidding process. If a Bidder fails to comply with any of the terms, its EOI may be summarily rejected.
- ii. Willful misrepresentation of any fact within the Bid will lead to the disqualification of the Bidder without prejudice to other actions that the Bank may take. All the submissions, including any accompanying documents, will become property of the Bank. The Bidders shall be deemed to license, and grant all rights to the Bank, to reproduce the whole or any portion of their

Bid document for the purpose of evaluation and to disclose the contents of submission for regulatory and legal requirements.

- iii. Bidders must advise the Bank immediately in writing of any material change to the information contained in the EOI application, including any substantial change in their ownership or their financial or technical capacity. Copies of relevant documents must be submitted with their advices. For shortlisted Bidders, this requirement applies until a contract is awarded as a result of subsequent bidding process.
- iv. Shortlisted Bidders must not advertise or publish the same in any form without the prior written consent of SBI.
- v. Brief overview of the proposed procurement/scope of work given in this document may be further elaborated, viz., more details may be included in the Request for Proposal (RFP) document to be issued as a result of evaluation process of EOIs.
- vi. The Bank reserves the right to formulate any terms & conditions while framing the RFP, even if these are in variance with the terms provided in this EOI. Further, the Bidders shall have no claim in this regard.
- vii. The Bank shall have the right to cancel the EOI process itself at any time, without thereby incurring any liabilities to the affected Bidders. Reasons for cancellation, as determined by the Bank in its sole discretion include but are not limited to, the following:
 - a. Services contemplated are no longer required.
 - b. Scope of work not adequately or clearly defined due to unforeseen circumstance and/or factors and/or new developments.
 - c. The project is not in the best interest of the Bank.
 - d. Any other reason.



PART-II



Appendix-A: BID FORM (TECHNICAL BID)

[On Bidder's letter head]

Date: _____

To:

< Address of tendering office >

Dear Sir,

Ref: EOI No. SBI:xx:xxdated dd/mm/yyyy

~~~~~

We have examined the above EOI, the receipt of which is hereby duly acknowledged and subsequent pre-bid clarifications / modifications / revisions, if any, furnished by the Bank. We submit our bid and shall abide by the terms and conditions spelt out in the EOI.

**i. While submitting this Bid, we certify that:**

- The undersigned is authorized to sign on behalf of the Bidder and the necessary support document delegating this authority is enclosed to this letter.
- We declare that we are not in contravention of conflict of interest obligation mentioned in this EOI.
- We have not induced or attempted to induce any other Bidder to submit or not to submit a Bid for restricting competition.

- ii. We undertake that, in competing for the above EOI, we will strictly observe the laws against fraud and corruption in force in India namely "Prevention of Corruption Act 1988".
- iii. We undertake that we will not offer, directly or through intermediaries, any bribe, gift, consideration, reward, favour, any material or immaterial benefit or other advantage, commission, fees, brokerage or inducement to any official of the Bank, connected directly or indirectly with this EOI process, or to any person, organization or third party related to this EOI in exchange for any advantage in the EOI, evaluation and shortlisting.
- iv. We undertake that we will not resort to canvassing with any official of the Bank, connected directly or indirectly with the EOI process to derive any undue advantage. We also understand that any violation in this regard, will result in disqualification of bidder from further bidding process.
- v. It is further certified that the contents of our Bid are factually correct. We have not sought any deviation to the terms and conditions of this EOI. We also accept that in the event of any information / data / particulars proving to be incorrect, the Bank will have right to disqualify us from the EOI or any subsequent bidding process without prejudice to any other rights available to the Bank.
- vi. We certify that while submitting our Bid document, we have not made any changes in the contents of the EOI, read with its amendments/clarifications provided by the Bank.
- vii. We understand that you are not bound to accept any Bid you may receive and you may reject all or any Bid without assigning any reason or giving any explanation whatsoever.
- viii. We hereby certify that our name does not appear in any "Caution" list of RBI / IBA or any other regulatory body for outsourcing activity.
- ix. We hereby certify that on the date of submission of Bid for this EOI, we are not under any debarment/blacklist period for breach of contract/fraud/corrupt practices by any Scheduled Commercial Bank/ Public Sector Undertaking/ State or Central Government or their agencies/departments.



- x. We hereby certify that on the date of submission of Bid, we do not have any Service Level Agreement pending to be signed with the Bank for more than 6 months from the date of issue of purchase order.
- xi. We hereby certify that we have read the clauses contained in O.M. No. 6/18/2019-PPD, dated 23.07.2020 order (Public Procurement No. 1), order (Public Procurement No. 2) dated 23.07.2020 and order (Public Procurement No. 3) dated 24.07.2020 regarding restrictions on procurement from a bidder of a country which shares a land border with India. We further certify that we and our OEM are not from such a country or if from a country, has been registered with competent authority. We certify that we and our OEM fulfil all the requirements in this regard and are eligible to participate in this EOI.
- xii. We, further, hereby undertake and agree to abide by all the terms and conditions stipulated by the Bank in the EOI document.

Dated this ..... day of ..... 202

\_\_\_\_\_

*(Signature)*

*(Name)*

*(In the capacity of)*

Duly authorised to sign Bid for and on behalf of

\_\_\_\_\_

**Seal of the company.**

**Appendix-B: Bidder's Eligibility Criteria**

Bidders meeting the following criteria are eligible to submit their Bids along with supporting documents. If the Bid is not accompanied by all the required documents supporting eligibility criteria, the same would be rejected:

| Sl. No. | Eligibility Criteria                                                                                                                                                                                                                                                     | Compliance (Yes/No) | Documents to be submitted                                                                                                                                                  |
|---------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1.      | The Bidder must be an Indian Company/ LLP /Partnership firm registered under applicable Act in India.                                                                                                                                                                    |                     | Certificate of Incorporation issued by Registrar of Companies and full address of the registered office along with Memorandum & Articles of Association/ Partnership Deed. |
| 2.      | The Bidder (including its OEM, if any) must comply with the requirements contained in O.M. No. 6/18/2019-PPD, dated 23.07.2020 order (Public Procurement No. 1), order (Public Procurement No. 2) dated 23.07.2020 and order (Public Procurement No. 3) dated 24.07.2020 |                     | Bidder should specifically certify in <b>Appendix A</b> in this regard and provide copy of registration certificate issued by competent authority wherever applicable.     |
| 3.      | The Bidder must have an average turnover of minimum Rs. 1275 crore during last 03 (three) financial year(s) (2022-23, 2021-22, 2020-21)                                                                                                                                  |                     | Copy of the audited financial statement for required financial years. (Certificate from statutory auditor for preceding/current Financial Year may be submitted.)          |

|    |                                                                                                                                                                                                                                                               |  |                                                                                                                                                     |
|----|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|-----------------------------------------------------------------------------------------------------------------------------------------------------|
| 4. | The Bidder should be profitable organization on the basis of profit before tax (PBT) for at least 02 (two) out of last 03 (three) financial years mentioned in para 2 above.                                                                                  |  | Copy of the audited financial statement along with profit and loss statement for corresponding years and / or Certificate of the statutory auditor. |
| 5  | Bidder should have experience of minimum 3 years in implementation and managing large-scale (~petabytes) enterprise data warehouse / data lake preferably in BFSI sector.                                                                                     |  | Copy of the order and / or Certificate of completion of the work. The Bidder should also furnish user acceptance report from client(s).             |
| 6. | Client references and contact details (email/ landline/ mobile) of customers for whom the Bidder has executed similar projects in India. (Start and End Date of the Project to be mentioned) in the past <b>(At least two client references are required)</b> |  | Bidder should specifically confirm on their letter head in this regard as per <b>Appendix-F</b>                                                     |
| 7. | Past/present litigations, disputes, if any (Adverse litigations could result in disqualification, at the sole discretion of the Bank)                                                                                                                         |  | Brief details of litigations, disputes, if any are to be given on Company's letter head.                                                            |
| 8. | Bidders should not be under debarment/blacklist period for breach of contract/fraud/corrupt practices by any Scheduled Commercial Bank/ Public Sector Undertaking / State or Central Government or their agencies/                                            |  | Bidder should specifically certify in <b>Appendix A</b> in this regard.                                                                             |



|    |                                                                                                                                                            |  |                                                                         |
|----|------------------------------------------------------------------------------------------------------------------------------------------------------------|--|-------------------------------------------------------------------------|
|    | departments on the date of submission of bid for this EOI.                                                                                                 |  |                                                                         |
| 9. | The Bidder should not have any Service Level Agreement pending to be signed with the Bank for more than 6 months from the date of issue of purchase order. |  | Bidder should specifically certify in <b>Appendix A</b> in this regard. |

Documentary evidence must be furnished against each of the above criteria along with an index. All documents must be signed by the authorized signatory of the Bidder. Relevant portions, in the documents submitted in pursuance of eligibility criteria, should be highlighted.

**Name & Signature of authorised signatory**

**Seal of Company**



## Appendix-C: Technical Criteria/Scope of Work

### **C1. Background**

State Bank of India (SBI) is keen on embarking on a Data Analytics journey and gain an advantage over its competitors in terms of innovation, time to market, user delight, cost of service, despite its behemoth size and scale.

There is an increasing trend globally in large Banks to use customer data in near real time from their mobile click-stream / touch-stream data, purchase patterns to make customized offers / improve service experience / support for advanced AI-ML models.

With data at SBI growing at an astronomical rate, a robust, horizontally scalable and cost-effective infrastructure for data storage and processing is the need of the hour. The ability to handle complex data from semi-structured and unstructured data sources, in addition to the conventional structural data processing, maximizes value potential and quality of services in the financial sector. Operational efficiency in the digital age is one of the factors that determine a Bank's growth. By managing petabytes of data efficiently, SBI has the potential to be a trend setter for smart data management globally. Empowering representatives of Bank at every level with analytics, thus standardizing and speeding up the decision-making process, will be a natural outcome of smart data management.

Data Warehouse has migrated on new infrastructure on 10<sup>th</sup> May 2023 as per five-year Agreement between Bank and M/s IBM India Private Limited in February 2020 for implementation and maintenance of the Data Warehouse. With the implementation of project, DWH has been shifted from PDOA (Pure data for Operation Analytics) technology to IIAS (IBM Integrated Analytical System) technology, along with a small-scale Data Lake and Data Virtualization capabilities. Data Lake is implemented with a view to enrich the capabilities of sourcing unstructured and semi-structured data for advance analytics with AI / ML tools.

## **C2. Project Mandate**

### **C2.1 Goals and Objectives**

- a) Integrated solution with real time / near real time sourcing of structured and unstructured data with real time analytical capabilities, object storage, NFS storage, columnar data storage supporting in-memory analysis/ analytics. Right architecture with right components should be proposed.
- b) Real time integration with upstream and downstream users with latest technologies like streaming and APIs for data movement / utilization.
- c) Capabilities of data ingestion from all source system available with the Bank or any future standard source systems that the Bank will decide to use with high throughput and low latency of the Bank in automated manner.
- d) Optimization of solution / tool (s) for Data Federation/Virtualization to ensure seamless integration of data in real time when stored in multiple sources without physical movement of data sets for the purpose of reporting / analytics. Data federation should be enhanced with more intelligent real-time query optimization, caching, in-memory and hybrid strategies that are automatically (or manually) chosen based on source constraints, application need and network awareness.
- e) Since our present Data Warehouse system is on-prem, it has limited scope of scalability and it requires complex forecasting to determine how to scale the data warehouse for future needs. With a view to overcome these limitations, the proposed solution should be implementable in Physical hardware / Bank's own cloud (Meghdoot) and scalable to public / hybrid cloud in future. This solution should also support cloud bursting for heavy workload on demand.
- f) Implementable solution for availability of DR.
- g) Low cost integrated archival solution with support for easy movement of data between active and archive data sets.

- h) Support for effective data governance for data discovery, data modelling, data quality, data catalogue, data profiling, data reconciliation and data lineage.
- i) Data reorganization to enhance usability and eliminate redundancy.
- j) Migration of existing ETL pipelines / jobs and analytical models.
- k) Dev, UAT, Production and DR setup to be done along with sandbox for analytics.
- l) Self-service BI reporting by end users.
- m) Seamless real time / near real time integration with various source systems including but not limited to support for oracle 19c and higher versions with TDE.

### C3. Current Components

#### C3.1 Present Architecture of Data Warehouse

Present Architecture of existing Data Warehouse and technology stack will be shared with the intended bidders who participate in this EOI proposal on execution of Non-Disclosure Agreement (NDA). Copy of NDA will be shared in due course.

#### C3.2 Volumetric of Existing Data Warehouse:

| SI. No. | Area               | Volumetric |
|---------|--------------------|------------|
| 1       | Production DB Size | ~1 PB      |
| 2       | Archival DB Size   | ~1 PB      |
| 3       | Data Lake          | ~730 TB    |
| 4       | Total ETL Jobs     | ~27000     |

|    |                                   |        |
|----|-----------------------------------|--------|
| 5  | Total Tables                      | ~50000 |
| 6  | Data Source Systems               | ~100   |
| 7  | Data Marts                        | ~30    |
| 8  | Monthly Data Ingestion            | ~20 TB |
| 9  | Scheduled Workload (jobs per day) | ~5000  |
| 10 | Ad-hoc Workload (jobs per day)    | ~200   |
| 11 | Number of Cognos Reports          | ~100   |

#### C4. Functional Requirements

Bank is planning to upgrade existing Data Warehouse and Data Lake solution to a modern and state of art integrated solution. In addition to project mandates mentioned in the section C3, the solution should provide following additional capabilities:

- 1) Support of Artificial Intelligence, Machine Learning, Business Intelligence, Data Engineering, Data Management features together on a single platform.
- 2) Enables different teams to use unified platform across all the enterprise data for a range of products including data science, machine learning and business intelligence.
- 3) Support to Improved data reliability, reduction in cost, data deduplication, improved data management and governance.
- 4) ACID transaction support (Atomicity, Consistency, Isolation, and Durability)
- 5) Unstructured, semi-structured and Streaming data support
- 6) Application Programming Interface (API) support.
- 7) Decoupled Storage and Compute resources: to support independent scalability of storage and compute without dependency on other.
- 8) Integrated Business Intelligence solution.

- 9) Data Management capabilities with multilayered architecture with different layer of data processing and data cleansing.
- 10) Supporting multi-programming languages for data lake and data warehouse functioning.
- 11) Data optimization capabilities to optimize data for faster analytical performance through measures such as clustering, distributing, caching, indexing etc.
- 12) Ability to store diverse set of data and use in variety of BI, analytics and data science applications, including both batch and streaming workloads.
- 13) Ability to relate data across diverse type of data environment.
- 14) Capability of In-memory and real-time analytics.
- 15) Analytics: The product's ability to perform advanced analytic operations within the dbPaaS. It is evaluated on the basis of what functionality is offered in the current version of the product, and what functionality is being used by customers.
- 16) Analytics on Streaming Data: Analytical capabilities specifically tied to stream data, including windowing analytics, analytics on in-flight data, joins to static data and support for streaming standards, among others.
- 17) Stream Optimization: The ability to optimize the capture, ingestion and analysis of stream data.
- 18) Compatible with Oracle CDC (Goldengate) at near real-time
- 19) Operational Intelligence: The ability to handle large numbers of concurrent users running fixed short analytical queries while still providing appropriate performance SLAs for operational workloads.
- 20) Data Science: The ability to build, train and deploy models in the DBMS, the inclusion of a data science workbench, the ability to implement a smooth flow of model development and deployment, support for feature stores and other capabilities relevant to the exercise of data science workloads.
- 21) Intercloud: The ability to deploy and operate analytical activities across nodes on multiple private/public cloud environments and on-premises. Inter-cloud means the ability to use data and execute operations across multiple clouds as a single logical entity.
- 22) Distributed Capabilities: The ability to access data outside the internal storage of a DBMS and optimize distributed access by a variety of

methods, such as push-down, extended metadata, statistics collection and a distributed catalog, among others. : This category also covers the ability to carry out distributed transactions.

- 23) Governance: The ability to forecast, budget usage, monitor and control costs by throttling, workload or user prioritization or other means.
- 24) Application Development Support: The ability to support multiple application languages (Python mandatory) and their APIs, the ability to support stored procedures and UDFs, and the ability to implement constraints, among other features.
- 25) Performance Features: This capability includes optimization, statistics collection, the ability to use static and dynamic plans, partitioning, partition elimination and storage tiering for performance and materialized views, among other features.
- 26) Relational Attributes: The ability to support complex relational operations involving many tables.
- 27) Multi model Support: The ability to support different storage and logical models within the DBMS, such as JSON data and external tables, efficiently and with performance optimization, as well as additional capabilities such as temporal, time series, geospatial and graph, among others.
- 28) Management and Administration: The ability to manage instances and resources, monitor operations, track and implement security, high availability and disaster recovery, and to do these and other tasks at enterprise scale.
- 29) Resource Usage: The ability to automatically handle different types and sizes of workload simultaneously while enforcing or dynamically extending policy-based resource limits, handle varying and conflicting workloads while optimizing response times and prioritize workloads to meet policy-defined service levels.
- 30) ML platform which supports a common collaborative environment capable of running all analytic processes in one place. This platform should enable us to ingest data from many sources, build data pipelines, train ML models, automate ML models, monitor and retrain ML models. Pipelines to orchestrate the entire machine learning process from data preparation to deployment.
- 31) Platform should support running models on both CPU and GPU. GPU to be provisioned accordingly.

- 32) Unified model monitoring platform for all the deployed models through dashboard.
- 33) Platform should support the visibility of models and features for model governance.
- 34) Capability to deploy models over microservice. Deploy models/workflows as industry standard web service
- 35) Data governance, data security at rest and in transit, unified authorization, access tracking, data privacy and governance standards across the AI/ML journey
- 36) Catalog of prepackaged ML use cases, models, and algorithms on wide ranging topics which can be directly used with Bank's data to get insights
- 37) Integrated data visualization tool which enables to build dashboards of the deployed models. Support to build /host dashboards on the data and deploy to end users
- 38) Object storage support; Support to run analytical queries against object stores
- 39) Support Spark based processing
- 40) Segregated Development and Production environment
- 41) Integration with R, Python, Keras, TensorFlow, Theano, scikit-learn etc and other frameworks / languages.
- 42) Support ML, Deep Learning, AI, Generative AI Models
- 43) Support automated experiments
- 44) Support Explainable AI

### **Architectural Requirements**

- 1) Cloud Native (Container Clusters, with on-demand instances across on-premises/private cloud/public cloud)
- 2) Commodity H/w with ability to integrate Private Cloud (VMWare Tanzu /ESXI) nodes into cluster for compute or storage.
- 3) Open File Format based: Should use Parquet/ORC or similar open standard file format.
- 4) Open DBMS layer:
  - a. Hudi or similar DBMS Layer

- b. Direct JDBC connectivity to DBMS
- c. Should be able to perform:
  - i. Schema Evolution
  - ii. Incremental Writes
  - iii. Update records
  - iv. Streaming data
- 5) Open ETL Layer: Should use open standard ETL platforms like Apache NiFi
- 6) Note: All components must be based on open standard/open source software, customized for enterprise use (Open Core)

EOI Proposal should include following items for each proposed solution:

- a) Following Specifications for each of PROD, DEV, UAT and DR environments:
  - i. Architecture Diagram
  - ii. Network Diagram
  - iii. Process/Data Flow Diagram
  - iv. Proposed software specification along with OEM and features of each software product
  - v. Proposed hardware specification
  - vi. Proposed internal and external network specification.
  - vii. Rackspace, Power, Cooling, Network connectivity & Bandwidths
  - viii. Any other relevant artefacts
- b) Tentative project implementation timelines including major milestones e.g setup, security review, migration, validation plan, etc.
- c) Cost model (e.g. how licensing will be done) - Actual commercials are not required to be mentioned at this stage.





- d) Team structure
- e) Performance benchmark for proposed solution with comparable solution available in the market.

**Present Architecture of existing Data Warehouse and technology stack will be shared with the intended bidders who participate in this EOI proposal on execution of Non-Disclosure Agreement (NDA). Copy of NDA will be shared in due course.**

The applicant should describe how their solution will meet the required parameters and provide details thereof in their EOI proposal on the following lines by answering all the points given in below table as capability (Yes/No).

| Sr. No                | Parameters                                                                                                                                                                                                                                                                                                                                                           | Capability (Yes / No) |
|-----------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| <b>Data Ingestion</b> |                                                                                                                                                                                                                                                                                                                                                                      |                       |
| 1                     | Capable of ingesting data from any source system in automated manner currently implemented in the Bank, or any future standard source systems that the Bank will decide to use with high throughput and low latency. Vendor to propose performance benchmarking for the same.                                                                                        |                       |
| 2                     | Data may be structured, semi-structured, and unstructured. It may come from internal or external sources. It may come in batches, incremental additions or real-time feeds. There should be no limitation on the type, format and size of data ingested. Data may include log, feeds, audio, video, image, NOSQL, RDBMS, unstructured text, through ERP systems, etc |                       |

|   |                                                                                                                                                                                                                                                                                                                                                                  |  |
|---|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| 3 | Graphical User Interface (GUI) based framework to configure sources to ISDLDW                                                                                                                                                                                                                                                                                    |  |
| 4 | Data ingestion system should be capable to configure data ingestion from various sources sharing all type of data through various methods e.g streaming, batch, files, archive logs, etc.                                                                                                                                                                        |  |
| 5 | Solution proposed for Data Ingestion should be compatible to ingest and replicate data simultaneously on all available regions.                                                                                                                                                                                                                                  |  |
| 6 | End objective for the data ingestion is to publish the dashboards for end users or any job related to reporting and analytics maximum within 3 hours from day end for batch data loads on all days including month-end / quarter-end / year-end. Vendor may propose suitable solution to achieve the objective. Refer Annexure – C for existing data flow in DWH |  |
| 7 | Data sanity checks, automated reject processing, validations and reconciliation of data should be available as part of data ingestion solution to ensure the integrity of data.                                                                                                                                                                                  |  |
| 8 | Functionality of existing jobs to be finetuned. Re-runnability checkpoints should be present in ETL jobs. New ETL jobs should be able to parallel read and write data.                                                                                                                                                                                           |  |
| 9 | Auto Resume – Aborted ETL jobs must resume from check point without any manual clean up in associated table or file being written                                                                                                                                                                                                                                |  |



|    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |  |
|----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| 10 | Ingestion subsystem should generate audit and diagnostic logs                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |  |
| 11 | An alerting report and monitoring utility about the ingest pipelines should be available as part of the solution.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |  |
| 12 | Trigger mechanisms in identifying any structural changes at source                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |  |
| 13 | The vendor should be able to design solutions to handle data volumes and complexity in source data with decompression logic wherever required. An example of such workload is provided below (Workloads of this type are in scope of this engagement); A source system table with ~ 300 columns in an Oracle database generates data of ~ 800 GB on a daily basis with ~ 400 columns in DWH. 100 columns in DWH are compressed in a single column at source in Oracle. The ingestion tools should be able to perform a change data capture on source systems of this nature with run time decompression functionality. |  |
| 14 | Vendor should list out all types of risks they expect from the ingestion subsystem (e.g., dropping of data packets during ingestion, security loopholes, unprotected personally identifiable information, etc.) along with mechanisms and processes they would implement for mitigating such risks.                                                                                                                                                                                                                                                                                                                    |  |
| 15 | Proposed solution should be able to scrap encrypted log, capture Metadata changes at source level completely, scrapping 4000-5000 logs daily having log size of ~ 32 GB each scalable up to 20000 logs.                                                                                                                                                                                                                                                                                                                                                                                                                |  |



|                     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |  |
|---------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
|                     | Proposed solution should be capable of scrapping logs generated by any type of Database. E.g. Oracle Database, IBM DB2 Database etc.                                                                                                                                                                                                                                                                                                                                                    |  |
| 16                  | Solution should be able to handle DDL change without manual reorg/runstat.                                                                                                                                                                                                                                                                                                                                                                                                              |  |
| 17                  | A job scheduler, along with process management controls that provide things like runtime monitoring and error alerting, handling, and logging.                                                                                                                                                                                                                                                                                                                                          |  |
| 18                  | When an ETL package runs, it is very important to be able to log how it executed i.e. we need to know how long it takes, what time it started/ended, who triggered it, was it successful or failed, what was the error message, etc. Apart from logging, when an ETL tool runs, it also needs to be able to run at scheduled times, re-run when failed, and limit the execution duration. ETL meta data repository should be query-able by designated user to extract this information. |  |
| 19                  | Tool should have in-built capability to optimize/prioritize the ingestion pipeline system as a whole. For example if multiple pipelines are batched in the system tool should determine the cost of the pipeline and optimize & prioritize accordingly.                                                                                                                                                                                                                                 |  |
| 20                  | Tool should have the capability to notify any transmission loss through connectivity, network, system, hardware or any other failures. It should automatically resume from the checkpoints with minimal or no manual intervention.                                                                                                                                                                                                                                                      |  |
| <b>Data Storage</b> |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |  |



|   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |  |
|---|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| 1 | Vendor should propose effective number of data storage layers (landing, staging, history, data mart, etc) in ISDLDW between data ingestion and data consumption.                                                                                                                                                                                                                                                                                                                                                                                             |  |
| 2 | Vendor should propose the type of storage (object storage etc) to opt for in ISDLDW (SQL, NO-SQL etc) and provide details of the hardware requirements, supported components for storage management and any other relevant artifacts for each storage subsystem                                                                                                                                                                                                                                                                                              |  |
| 3 | A multi-temperature data management solution to be proposed by vendor where data that is frequently accessed on fast storage—hot data—compared to less-frequently accessed data stored on slightly slower storage—warm data—and rarely accessed data stored on the slowest storage —cold data. System should also be capable automated storage tiering and seamless data transfer between hot, warm and cold storage. Data residing in any of these storage areas must be seamlessly mixed / merged according to requirements without impacting performance. |  |
| 4 | Storage replication (e.g. RAID) should be automatically managed by the platform.                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |  |
| 5 | Tool should have capability to store/swap data in memory, disk and distributed storage areas depending on the age of the data determined by its usages through user queries.                                                                                                                                                                                                                                                                                                                                                                                 |  |



|    |                                                                                                                                                                                                                                                                                           |  |
|----|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| 6  | Provision of periodic backup with facility to selective restoration of specific data set without hampering database access to any schedule / adhoc workloads.                                                                                                                             |  |
| 7  | Processing engines should have capabilities to process all kinds of Data Storage formats used by the Bank.                                                                                                                                                                                |  |
| 8  | Storage should support data compression without degrading the system performance.                                                                                                                                                                                                         |  |
| 9  | The storage should be horizontally and vertically scalable. Redistribution of data across the ISDLDW should be possible automatically and seamlessly.                                                                                                                                     |  |
| 10 | Ensuring real time health checks, monitoring and alerting about data storage / utilization of storage / failure handling of storage components. Actionable dashboard must be available to designated users to monitor health checks and tool should automatically issues alerts to users. |  |
| 11 | The storage system should be robust to handle at least 1,50,000 concurrent queries (Select/DML) by processing engines / ETL jobs / end users scalable up to 6,00,000 concurrent queries in next 5 years (assuming parallelism of 100 degree).                                             |  |
| 12 | The ISDLDW will have storage intensive, compute intensive and balanced workload scenarios. The storage mechanism should serve as a single entity that can be utilized by multiple processing engines or                                                                                   |  |

|                                  |                                                                                                                                                                                                                                                                                                                                                                                                                       |  |
|----------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
|                                  | compute nodes designated for a specific purpose in the ISDLDW.                                                                                                                                                                                                                                                                                                                                                        |  |
| 13                               | Vendor should list out comprehensively the types of risks they expect from the storage subsystem (e.g., loss of data blocks due to insufficient replication, data corruption, problems of insufficient storage during operations, inefficient retrieval of blocks or files, etc). For each such risk, they should describe the processes and mechanisms that would be put in place to avoid and/or mitigate the risk. |  |
| <b>Data Processing Framework</b> |                                                                                                                                                                                                                                                                                                                                                                                                                       |  |
| 1                                | For the data to be accessible and consumable by businesses / downstream applications, the ISDLDW should have robust, highly efficient and parallel execution of data transformation jobs.                                                                                                                                                                                                                             |  |
| 2                                | ISDLDW ecosystem should have state of the art data processing engines that can perform in-memory processing to reduce the time for data transformations and query in case of real time requirements.                                                                                                                                                                                                                  |  |
| 3                                | Framework should allow joining multiple sources/tables/inputs etc.                                                                                                                                                                                                                                                                                                                                                    |  |
| 4                                | Framework should be able to put the processed data in a predefined area separate from the raw data, while minimizing data replication.                                                                                                                                                                                                                                                                                |  |
| 5                                | Framework should be capable of performing validation checks pre-and post-processing.                                                                                                                                                                                                                                                                                                                                  |  |



|    |                                                                                                                                                                                                                                                                           |  |
|----|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| 6  | <p>Have a workflow management and scheduling solution to schedule data transformation, data acquisition or data delivery jobs.</p> <p>Allocations of separate workload channel to designated queries</p>                                                                  |  |
| 7  | <p>Should have failure, retry, alert and escalation logic defined. Successful file transfer should also be automatically acknowledged</p>                                                                                                                                 |  |
| 8  | <p>Should have audit and error logs for auditing and troubleshooting</p>                                                                                                                                                                                                  |  |
| 9  | <p>Automatic recovery of data after failure/rejection of record needs to happen without any manual intervention</p>                                                                                                                                                       |  |
| 10 | <p>Framework should have mechanism to encrypt data at rest, data at motion and data in use.</p>                                                                                                                                                                           |  |
| 11 | <p>ETL/ELT tool for data extraction should be AI/ML features for suggesting / improving Query / ETL / ELT Stages</p>                                                                                                                                                      |  |
| 12 | <p>Existing reports and extracts generation jobs on DWH should be analyzed and transformed to the ISDLWD. The vendor should use preferably off-the-shelf tools and not resort to building from scratch.</p>                                                               |  |
| 13 | <p>Data transformations should be triggered in parallel. The ISDLWD should be capable to run multiple transformation jobs in parallel. The ISDLWD should be able to run at-least 500 jobs in parallel (concurrently), scalable up to 1500 in next 5 years, of varying</p> |  |





|    |                                                                                                                                                                                                                                                                                                                                                                                        |  |
|----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
|    | <p>complexity - simple, medium, complex, in batch or near real time mode every day.</p>                                                                                                                                                                                                                                                                                                |  |
| 14 | <p>There should be provision to monitor, log and assess the progress of each transformation job.</p> <p>Job design (s) should be re-usable.</p> <p>There should be a provision to organize individual transformations as user defined operators. Provision should be there in framework to tie these operators to data structures and organize the entire process as a single job.</p> |  |
| 15 | <p>Organizing jobs, scheduling jobs and triggering jobs should be possible from an easy to use interface. Reports on job status and success / failure / retrigger should be sent to concerned stakeholders on a continuous basis. The tool(s) used for data transformation should be aligned to data sources and types / formats of data residing on ISDLDW.</p>                       |  |
| 16 | <p>The processing pipelines for ETL/ELT jobs also include real time, daily, weekly, monthly, quarterly and annual reports, feeding data structures for downstream consumption. These activities are in-scope for this engagement.</p>                                                                                                                                                  |  |
| 17 | <p>The workflows should work with standard schedulers. Monitoring and management of workflows should be possible from an easy to use interface. Workflow management tool(s) should have connectors / pluggable interfaces to already existing / in-use proprietary software available with the Bank. These could be (and not restricted to) data repositories,</p>                     |  |



|                                                           |                                                                                                                                                                                                                                                |  |
|-----------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
|                                                           | reporting tools, data analysis tools and generic interfaces for data transfer. Scheduled jobs status should be made available to the Bank in Monitoring dashboard on real time basis.                                                          |  |
| 18                                                        | Tool should be flexible enough to implement the data privacy and security policies of the Bank while transforming and storing the data.                                                                                                        |  |
| <b>Data Federation/Virtualization</b>                     |                                                                                                                                                                                                                                                |  |
| 1                                                         | Vendor to propose a solution / tool (s) for Data Federation/Virtualization to ensure seamless integration of data in real time when stored in multiple sources without physical movement of data sets for the purpose of reporting / analytics |  |
| 2                                                         | Data federation support intelligent real-time query optimization, caching, in-memory and hybrid strategies that are automatically (or manually) chosen based on source constraints, application need and network awareness.                    |  |
| 3                                                         | Semantic integration of structured & unstructured Data.                                                                                                                                                                                        |  |
| 4                                                         | Data virtualization should support the use of APIs.                                                                                                                                                                                            |  |
| 5                                                         | All data should within eco-system and across other system whenever possible be made discoverable and integrable easily through a single virtual layer which will expose redundancy and quality issues faster.                                  |  |
| <b>Migration from Existing Setup to Proposed Solution</b> |                                                                                                                                                                                                                                                |  |



|    |                                                                                                                                                                                                                                                                              |  |
|----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| 1  | Vendor should propose a detailed seamless automated migration plan from existing setup to proposed solution. Plan should focus on less manual intervention, data reconciliation between the systems and minimum parallel run of existing and proposed solution.              |  |
| 2  | Data migration from Staging and Data Marts, user tables and any other schemas identified by Bank.                                                                                                                                                                            |  |
| 3  | Data migration from existing archival solution to new one.                                                                                                                                                                                                                   |  |
| 4  | Migration of existing data sourcing ETL jobs.                                                                                                                                                                                                                                |  |
| 5  | Migration of existing data extraction and reporting jobs.                                                                                                                                                                                                                    |  |
| 6  | Migration of monitoring dashboard data points.                                                                                                                                                                                                                               |  |
| 7  | Migration of user details.                                                                                                                                                                                                                                                   |  |
| 8  | Migration of Data Governance, Data Lineage and Data Quality rules and policies                                                                                                                                                                                               |  |
| 9  | Migration of All the remaining components of existing ecosystem (Mentioned in Annexure - D) as and when identified by Bank like job scheduler, reports, history of version control, existing tape backup, etc.                                                               |  |
| 10 | Vendor should list out all types of risks they expect during the migration. Vendor should provide justification if any downtime is required on existing or proposed system during migration. Vendor should provide all the pre-requisites for the migration in the proposal. |  |

|                          |                                                                                                                                                                                                                                                                                    |  |
|--------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| 11                       | Vendor to review the existing architecture during migration and remove duplication of data and recommend improvements in overall setup if any                                                                                                                                      |  |
| 12                       | Vendor should provide a feasible plan for best use of existing infrastructure which is procured during last 10 years in staggered manner during the implementation of ISDLDW which will save cost to the Bank. (Annexure D gives the technology architecture of the current setup) |  |
| <b>Disaster Recovery</b> |                                                                                                                                                                                                                                                                                    |  |
| 1                        | Bank proposes to setup replication based DR capable of replication of select tables, schema and jobs.                                                                                                                                                                              |  |
| 2                        | DR will contain jobs, scripts, models, snapshots, metadata, all software requires for managing and maintaining PR cluster along with critical data defined by the Bank.                                                                                                            |  |
| 3                        | The ISDLDW – DR solution needs to be set up at a remote location at Hyderabad.                                                                                                                                                                                                     |  |
| 4                        | The DR solution should be synced with production ISDLDW. The SLA for RTO should be maximum 2Hrs as per Bank’s defined policy.                                                                                                                                                      |  |
| 5                        | The proposed solution is expected to have a monitoring engine that can determine the health of production ISDLDW and raise alerts / trigger remedial actions to bring ISDLDW – DR as the default ISDLDW                                                                            |  |
| 6                        | Switching over to ISDLDW – DR must have limited manual intervention. The handover should be soft and must not incur any data loss / data corruption. Jobs in                                                                                                                       |  |



|                                 |                                                                                                                                                                                                                                                     |  |
|---------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
|                                 | progress must be restarted with minimal impact (including data loss) with migration & automation tool.                                                                                                                                              |  |
| 7                               | User jobs will be categorized into Critical/Normal priorities as decided by the Bank so that jobs can be moved into DR based on assigned priorities.                                                                                                |  |
| 8                               | Jobs developed to be deployed in production, should be deployed both in production and DR. The same applies to any admin activity, security policy and other activities on ISDLDW                                                                   |  |
| 9                               | The Bank reserves the right to enforce mock up scenarios for failure to ensure that the ISDLDW – DR solution is working as expected. The vendor should support the Bank to demonstrate such scenarios                                               |  |
| <b>Data Archival and Backup</b> |                                                                                                                                                                                                                                                     |  |
| 1                               | Data older than specific duration as identified by Bank to be archived in low cost cold storage. Changing data archival rules should be easily configurable. Vendor to propose solution for the same with cheap and flexible storage and processing |  |
| 2                               | Data Archival solution should not be visible to end user, but Archived data should be available for all end users. For end user it should be a single view with Data Federation/Virtualization Layer                                                |  |
| 3                               | All the applications connected to the non-archived data should be available with archived as well                                                                                                                                                   |  |
| 4                               | Automated process for archiving data as per Banks policy                                                                                                                                                                                            |  |

|                                        |                                                                                                                                                                                                                            |  |
|----------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| 5                                      | Store backup of entire ecosystem on suitable cost-effective, fast recovery infrastructure (Currently tape backup is taken)                                                                                                 |  |
| 6                                      | Mixing and Merging data from Current/Live to and from Archival must not result in any significant loss of performance and response time                                                                                    |  |
| 7                                      | Archival and Backup setup must support automated Data Reconciliation whenever movement from Current/Live happens                                                                                                           |  |
| 8                                      | Capability of selective restore functionality for specific table(s).                                                                                                                                                       |  |
| <b>Cloud Integration and Migration</b> |                                                                                                                                                                                                                            |  |
| 1                                      | ISDLDW should be able to consume data from external cloud-based infrastructures.                                                                                                                                           |  |
| 2                                      | Cloud integration / data transfer to and from public/private/hybrid cloud should be available using all standard protocols. (Web requests / secure transfer channels etc)                                                  |  |
| 3                                      | The architecture of ISDLDW should enable interaction between public cloud and designated edge servers alone. ISDLDW non-edge server cannot be directly exposed to the public cloud (in case solution proposed is on-prem). |  |
| 4                                      | Transfer out of ISDLDW to public cloud should not be possible by all roles in ISDLDW. All activities with data                                                                                                             |  |



|                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                          |  |
|-----------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
|                             | transfer from Public cloud should be logged for audit and monitoring.                                                                                                                                                                                                                                                                                                                                                                                    |  |
| 5                           | In view of the intent to reduce the hardware footprint (in future), the technical architecture of ISDLDW solution should be flexible to accommodate moving of ISDLDW on cloud and also should support cloud bursting for heavy workloads. The Bank understands that there can be differences in services offered by cloud service providers. The ISDLDW solution architecture should be designed considering as-is infrastructure availability in cloud. |  |
| 6                           | Adherence to global and local standards related to cloud as well as compliance to applicable regulatory requirements.                                                                                                                                                                                                                                                                                                                                    |  |
| <b>Monitoring Dashboard</b> |                                                                                                                                                                                                                                                                                                                                                                                                                                                          |  |
| 1                           | Real time data flow in dashboard                                                                                                                                                                                                                                                                                                                                                                                                                         |  |
| 2                           | Health monitoring of ISDLDW ecosystem                                                                                                                                                                                                                                                                                                                                                                                                                    |  |
| 3                           | Drag and Drop feature to customize dashboard                                                                                                                                                                                                                                                                                                                                                                                                             |  |
| 4                           | Monitoring of all data sourcing tables/data marts on daily basis. Data Ingestion status pass / fail / in-progress for all source systems                                                                                                                                                                                                                                                                                                                 |  |
| 5                           | Operational logs analysis highlights                                                                                                                                                                                                                                                                                                                                                                                                                     |  |
| 6                           | Scalable distributed monitoring system for high-performance computing systems such as clusters and grids                                                                                                                                                                                                                                                                                                                                                 |  |

|    |                                                                                                                           |  |
|----|---------------------------------------------------------------------------------------------------------------------------|--|
| 7  | Industry best tools for performance (slowness) monitoring                                                                 |  |
| 8  | Controlling threshold breaches in pre-defined standard ways                                                               |  |
| 9  | End to end performance monitoring with automatic alerts                                                                   |  |
| 10 | Threshold control to kill the high resource consumption query                                                             |  |
| 11 | Monitoring of allocations of separate workload channel to designated queries                                              |  |
| 12 | Status of all applications / interfaces running on ISDLDW along with memory, disk, CPU usage etc                          |  |
| 13 | Automated email alerts to source systems on missing logs, files, stream etc.                                              |  |
| 14 | Email alerts from portal to downstream applications if data not loaded                                                    |  |
| 15 | Status of ISDLDW clusters, hardware status/software status etc. with need-based access                                    |  |
| 16 | Source System should be having access to portal for seeing the processing (rejection) status of the data provided by them |  |
| 17 | Monitor privileged users to monitor DBA activities & share history                                                        |  |
| 18 | Automated alerts for jobs failure, jobs successful should be available                                                    |  |





|                        |                                                                                                                                                                                                                                                                                                       |  |
|------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| 19                     | List of Job created not compiled, executed – should be available                                                                                                                                                                                                                                      |  |
| 20                     | Mechanism to ensure that file transfer failures are automatically detected                                                                                                                                                                                                                            |  |
| 21                     | DB health monitoring tool as well as monitor of N/W health                                                                                                                                                                                                                                            |  |
| 22                     | Usage statistics of tables, jobs, reports, dashboards etc.                                                                                                                                                                                                                                            |  |
| 23                     | Data Reconciliation status for every data movement on real time basis                                                                                                                                                                                                                                 |  |
| 24                     | All above parameters to be displayed in console with real time updates without any manual intervention                                                                                                                                                                                                |  |
| <b>Data Governance</b> |                                                                                                                                                                                                                                                                                                       |  |
| 1                      | Provide traceability – it should be possible to track and visualize any data transformation or any rule applied to data in the ISDLDW eco-system                                                                                                                                                      |  |
| 2                      | Provide trust – The system should be able to ensure the users that they are accessing data from the right source of information.                                                                                                                                                                      |  |
| 3                      | Provide auditability – the solution should record any access to the data to satisfy compliance audits. For example, it should be able to check on who touched the data, when did they touch it, is there a chain of custody issue, is there transparency in terms of data privacy and protection etc. |  |



|    |                                                                                                                                                                                                                              |  |
|----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| 4  | Enforce security and privacy – Data inside the ISDLDW will be accessed by only authorized users. Data at rest/in-motion should be encrypted                                                                                  |  |
| 5  | Capability to classify and store (personal identifiable information) sensitive data in encrypted /masked form and should have capability to decrypt/unmask such information in ISDLDW when required by only authorized ID's. |  |
| 6  | Capability to define clear roles and access management rules to user ID's                                                                                                                                                    |  |
| 7  | Automated propagation of changes to ISDLDW ecosystem Data Dictionary and business glossary by multiple sources as and when changes occur in source.                                                                          |  |
| 8  | Capability for purging/achieving of tables/jobs/reports which are not in use.                                                                                                                                                |  |
| 9  | Capability to review the conformance to governance policies and processes periodically (monthly/quarterly) and make necessary changes.                                                                                       |  |
| 10 | Provide detailed information about interfacing (upstream and downstream) applications.                                                                                                                                       |  |
| 11 | Data modeling capabilities to be provided by the tool                                                                                                                                                                        |  |
| 12 | Metadata Management Capability: Tool should cater to three broad categories of metadata; Business metadata, Technical metadata and Operational metadata                                                                      |  |



|                     |                                                                                                                                                                                                                                                                                                                                                                                                                            |  |
|---------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| 13                  | Architecture Review: The architecture including all the applications and policies of the ISDLDW ecosystem to be reviewed by Vendor and make necessary changes to get the maximum benefits from the setup. Vendor should provide statistics which Bank may demand in connection with this like user details who is using the services, hit rates, application runtime and downtime, hardware memory and storage usages, etc |  |
| 14                  | Vendor should provide a robust change management CI/CD framework to track versioning, change history with date and user information, track migration of code from development to production, security. Source code, IPR or customized IPR of ISDLDW will be owned by the Bank.                                                                                                                                             |  |
| 15                  | Performance benchmark of all components of ISDLDW to be given by participating Vendors                                                                                                                                                                                                                                                                                                                                     |  |
| <b>Data Quality</b> |                                                                                                                                                                                                                                                                                                                                                                                                                            |  |
| 1                   | Vendor should propose end-to-end solution for Data Quality Management starting from data origin till the data consumption. These tool (s) to be used for addressing various aspects of the data quality problem mentioned below on SBI data set during data ingestion, data processing or data consumption as advised by Bank on case by case basis                                                                        |  |
| 2                   | Parsing and standardization — Decomposition of text fields into component parts and formatting of values into consistent layouts based on industry standards, local standards                                                                                                                                                                                                                                              |  |



|    |                                                                                                                                                                                                                                                           |  |
|----|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| 3  | Generalized “cleansing” — Highlight data values which are failing to meet domain restrictions, integrity constraints or other business rules for user review                                                                                              |  |
| 4  | Recommend Matching —Identification, linking or merging related entries within or across sets of data for user review                                                                                                                                      |  |
| 5  | Profiling — Analysis of data to capture statistics (metadata) that provide insight into the quality of the data                                                                                                                                           |  |
| 6  | Monitoring — Deployment of controls to ensure ongoing conformance of data to business rules that define data quality for the organization                                                                                                                 |  |
| 7  | Recommend Enrichment — Enhancing the value of internally held data by appending related attributes from external sources                                                                                                                                  |  |
| 8  | Data quality dashboarding - A data quality dashboard can aggregate the status of continuously monitored data quality rules, as well as generate alerts to notify data stewards when they need to address an issue.                                        |  |
| 9  | Connectivity to multiple data sources - Tool (s) should be able to connect to a wide selection of data source types, both from a data management system -- e.g., RDBMS vs. NoSQL database -- and from a platform -- e.g., on premises vs. cloud -- basis. |  |
| 10 | Identity resolution - Identity resolution is the process of linking various records and is the main engine for record                                                                                                                                     |  |

|                            |                                                                                                                                                                                                                                                                                                                                                                                                                                      |  |
|----------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
|                            | de-duplication, which can enable some aspects of data cleansing.                                                                                                                                                                                                                                                                                                                                                                     |  |
| 11                         | Continuous ETL/ELT and Analytical data quality checks                                                                                                                                                                                                                                                                                                                                                                                |  |
| 12                         | Mechanism to capture feedback from end users to report Data Quality issues                                                                                                                                                                                                                                                                                                                                                           |  |
| 13                         | Dedicated Data Quality team who will thoroughly verify the Data Quality across setup                                                                                                                                                                                                                                                                                                                                                 |  |
| 14                         | Parallel processing: Data quality tool should be able to handle 500 concurrent users running any kind of job (e.g. Data Profiling on small/medium/large tables as advised by Bank in the tool)                                                                                                                                                                                                                                       |  |
| 15                         | <p>Each month 5 data quality use cases will be developed and implemented on Next Gen DW. Examples of use cases are as given below;</p> <ul style="list-style-type: none"> <li>• Profiling of Customer Master table for verifying PAN, Mobile Number, Date of Birth, Address, Pin Code, etc</li> <li>• Profiling of Branch Master for verifying branch address, contact information, branch manager/staff information, etc</li> </ul> |  |
| 16                         | System detectable Data Quality rules using business concepts defined at Data governance framework                                                                                                                                                                                                                                                                                                                                    |  |
| <b>Data Reconciliation</b> |                                                                                                                                                                                                                                                                                                                                                                                                                                      |  |
| 1                          | Developing automated reconciliation methodology suitable for each source system depending on recon                                                                                                                                                                                                                                                                                                                                   |  |



|                                |                                                                                                                                                                                                                                                                                                                                                                                                                                             |  |
|--------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
|                                | <p>implementation supported by source systems. Reconciliation should be automated such that, while loading data from source data should be reconciled. Reconciliation reports should be displayed real time for monitoring. Any discrepancies in reconciliation should be notified through alerts. Inherent database capabilities of ingestion pipeline content management DB must be leveraged for this purpose</p>                        |  |
| 2                              | <p>In the case of failure in reconciliation, tool should be able to drill down to the stage to identify the error and should be able to rectify the errors on near real time basis wherever possible. AI/ML capabilities can be leveraged for reject processing wherever applicable.</p>                                                                                                                                                    |  |
| 3                              | <p>Near real time dashboard to be published displaying records in source system and records received in DWH and action take for any discrepancies</p>                                                                                                                                                                                                                                                                                       |  |
| 4                              | <p>Data should be reconciled in every movement across table in any/all layers including data marts in automated way at the time of loading itself.</p>                                                                                                                                                                                                                                                                                      |  |
| <b>Security and Compliance</b> |                                                                                                                                                                                                                                                                                                                                                                                                                                             |  |
| 1                              | <p>Authentication and Identity Management - A comprehensive identity and access management system should be available for centralized management of users and groups. It should be possible to quickly create and revoke the identity of a user or a service by simply deleting or disabling the account in the directory. Multi-factor authentication is desired as an additional layer of security for user sign-in and transactions.</p> |  |



|   |                                                                                                                                                                                                                                                                                                                      |  |
|---|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| 2 | Authorization and access control - Role-based access controls should be provided for authorization of account-related and data-related activities.                                                                                                                                                                   |  |
| 3 | Data protection - It should be possible to protect the data in the ISDLDW throughout its lifecycle including data at rest and data in motion.                                                                                                                                                                        |  |
| 4 | Auditing - Audit or diagnostic logs should be used to log management-related activities or data-related activities. Log management and auditing of all critical activities on ISDLDW is a critical requirement. The Bank reserves right to ask the vendor to produce / analyze logs for reporting purposes.          |  |
| 5 | Data Democratization - Secure access of PROD Database to LHOs/GOCs                                                                                                                                                                                                                                                   |  |
| 6 | Data Leakage - Security CIA parameters should be achieved, and tools should be able to find and alert on Data leakage                                                                                                                                                                                                |  |
| 7 | Compliance with SBI IS Policies – some of the key areas are as under (Summary of Key policies is given in Annexure F):                                                                                                                                                                                               |  |
| 8 | <p>Compliance to Global Standards</p> <ul style="list-style-type: none"> <li>• GDPR, BCBS239, PCIDSS, DFRA and similar relevant standards</li> <li>• Office of Foreign Assets Control (OFAC)</li> <li>• Financial Crimes Enforcement Network (FinCEN)</li> <li>• Securities and Exchange Commission (SEC)</li> </ul> |  |



|                             |                                                                                                                                                                                                                                                                                                                                       |  |
|-----------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
|                             | <ul style="list-style-type: none"> <li>Office of the Comptroller of the Currency (OCC) etc.</li> </ul>                                                                                                                                                                                                                                |  |
| 9                           | Compliance of Bank's Information security policy, Data governance policy.                                                                                                                                                                                                                                                             |  |
| 10                          | Data privacy compliance                                                                                                                                                                                                                                                                                                               |  |
| <b>Regulatory Reporting</b> |                                                                                                                                                                                                                                                                                                                                       |  |
| 1                           | Vendor should follow the RBI guideline in developing the solution with which it will be easier for the Bank to migrate to the element-based data reporting envisaged by the RBI.                                                                                                                                                      |  |
| 2                           | Automation –Tool should automate analytics and reporting workflow end-to-end, including all data collection, enrichment, and management, as well as all calculations, processes to final report submission. Currently 500+ jobs are being used for Tranche 1 DCT generation along with 500 more for other regulatory reports/returns. |  |
| 3                           | Change Management –Tool for handling ongoing change in regulation or business requirements without the need for programming expertise. On and average logic for 5% of jobs being changed monthly. Data used for regulatory reporting changes on any frequency like daily / weekly / bi-weekly / monthly, etc                          |  |
| 4                           | Dashboards – Web browser dashboards can be used to present information and manage processes, such as                                                                                                                                                                                                                                  |  |





|    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |  |
|----|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
|    | data being used, data change history, job versions, current status, etc                                                                                                                                                                                                                                                                                                                                                                                                                       |  |
| 5  | Data Lineage and Transparency – Tool should retrace the journey of the source data through every single workflow processes or calculations across siloes systems all the way to disclosures.                                                                                                                                                                                                                                                                                                  |  |
| 6  | Drill Down and 100% Auditability – Every step in the data collection, enrichment and consolidation should be tracked and recorded, including all comments.                                                                                                                                                                                                                                                                                                                                    |  |
| 7  | Electronic Submission – Should support for all regulators globally in all required formats, including XBRL, XML or other file-based electronic submission.                                                                                                                                                                                                                                                                                                                                    |  |
| 8  | Flexibility to deal with any level of Complexity – Tool should handle the flow of information particular to any client’s organization, such as specific Key Risk Indicators (KRIs), syntaxes, data taxonomies and process operating models. Further, as demands like BCBS 239 require near-constant fine-tuning to address data governance, Business Rules should enable adjustments to regulatory and management changes with its out-of-the-box configurability and slice-and dice of data. |  |
| 9  | High Volume, High Performance and Reliability – Scalable and resilient architecture which will handle all volume and performance demands.                                                                                                                                                                                                                                                                                                                                                     |  |
| 10 | Leveraging Data and Processes – All data, calculations and other processes defined for one purpose are stored                                                                                                                                                                                                                                                                                                                                                                                 |  |

|                                 |                                                                                                                                                                                                        |  |
|---------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
|                                 | and should be reused wherever required to address new regulations and internal management reporting.                                                                                                   |  |
| 11                              | Pre-submission Review – Multiple report writers should allow users to review reports in various formats before submission, with the ability to drill down and make manual adjustments where necessary. |  |
| 12                              | Tool must support generation of reports in XBRL format                                                                                                                                                 |  |
| 13                              | Security and Control – Security and Compliance framework should provide a completely controlled and secure environment by providing appropriate permission to view granular data to designated users.  |  |
| <b>Audit and Log Management</b> |                                                                                                                                                                                                        |  |
| 1                               | Logging of operational activities: Support the logging of all user activities without slowing down the performance.                                                                                    |  |
| 2                               | Should have metadata enabled reporting mechanism on run time log.                                                                                                                                      |  |
| 3                               | Should have audit and error logs for auditing and troubleshooting                                                                                                                                      |  |
| 4                               | Extracting useful information from system logs to understand the efficiency of system and any fraud                                                                                                    |  |
| 5                               | User should be able to control the events/items being logged. For e. g user can opt for enabling logging of failures only.                                                                             |  |
| 6                               | Vendor to propose solution with cheap storage options for log storage of all the Bank's applications and                                                                                               |  |



|                        |                                                                                                                                                                                                                                                                                                                                      |  |
|------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
|                        | mechanisms to extract requested information from the logs as and when required                                                                                                                                                                                                                                                       |  |
| <b>Data Encryption</b> |                                                                                                                                                                                                                                                                                                                                      |  |
| 1                      | The sensitive and personal identification data on ISDLDW should be encrypted                                                                                                                                                                                                                                                         |  |
| 2                      | Data encryption applies to the following data on ISDLDW: <ul style="list-style-type: none"> <li>• Data at rest</li> <li>• Data in motion</li> <li>• Data available over an API to external sources / applications</li> <li>• Data at end users</li> <li>• Data from source system during ingestion</li> <li>• Data in use</li> </ul> |  |
| 3                      | Standard encryption/decryption techniques / policies should be implemented for this activity. Decryption should be allowed only to selective users/roles by complying the global standards like DES, SKC, PKC, 3DES, MD5 cryptographic hash, etc.                                                                                    |  |
| 4                      | The vendor team will determine tables / columns/ Cell-level encryption/Table space-level encryption that needs encryption by discussing with data owners before implementing the solution.                                                                                                                                           |  |
| 5                      | The processing time for data should include time needed for encrypting and decrypting information                                                                                                                                                                                                                                    |  |

|                        |                                                                                                                                                                                                                                                                                   |  |
|------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
|                        | needed for downstream / report generation. The overall SLA for data processing should be adhered to, keeping data encryption as an important activity.                                                                                                                            |  |
| 6                      | Proposed Solution should be capable on ingesting encrypted data from source system. It should support the encryption/decryption mechanism implemented at source system.                                                                                                           |  |
| <b>Data Masking</b>    |                                                                                                                                                                                                                                                                                   |  |
| 1                      | Capability to disable unmasking of masked data according to users' requirements                                                                                                                                                                                                   |  |
| 2                      | Tool should enable selective masking of columns in user specified tables                                                                                                                                                                                                          |  |
| 3                      | The output must be repeatable - The same source data, masked repeatedly by the same masking methodology, must yield the same output                                                                                                                                               |  |
| 4                      | Maintain referential integrity - Make sure the masked data is usable                                                                                                                                                                                                              |  |
| 5                      | Data Masking solution should support various methodologies like but not limited to below examples; <ul style="list-style-type: none"> <li>• Substitution</li> <li>• Nullifying and spacing</li> <li>• Number and date variance</li> <li>• Format-preserving encryption</li> </ul> |  |
| <b>User Management</b> |                                                                                                                                                                                                                                                                                   |  |



|                                 |                                                                                                                                                                                                                                                  |  |
|---------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| 1                               | Vendor should propose automated solution / tool (s) of User Access Management (UAM) for administration of giving access to individual users within a system access to the tools they need at the right time.                                     |  |
| 2                               | User Access Management should allow IT administrators to securely manage access to services/information/ resources for all the users.                                                                                                            |  |
| 3                               | Access rights to a user should be allocated and modified on the principle of least privilege and “Need to know” or “Need to do/have” basis.                                                                                                      |  |
| 4                               | The access privileges associated with each system product, e.g. operating system, network, database, application and system utilities, and the users to which these privileges need to be allocated should be clearly identified and documented. |  |
| 5                               | Privileges granted to the user should not conflict with Access Management -Segregation of Duties policy of the Bank.                                                                                                                             |  |
| <b>Business Continuity Plan</b> |                                                                                                                                                                                                                                                  |  |
| 1                               | Plans, measures and arrangements to ensure the continuous delivery of critical services and products, which permits the organization to recover its facility, data and assets.                                                                   |  |
| 2                               | Identification of necessary resources to support business continuity, including personnel, information, equipment, financial allocations, legal counsel, infrastructure protection and accommodations.                                           |  |

|                                    |                                                                                                                                                                                                                                                                                                                                                                                                                                        |  |
|------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| 3                                  | <p>Points to be taken care during preparation of BCP plan:</p> <ul style="list-style-type: none"> <li>• Plans must be updated and tested frequently</li> <li>• All types of threats must be considered</li> <li>• Dependencies and interdependencies should be carefully analyzed</li> <li>• Availability of key personnel</li> <li>• Network and telecommunications</li> <li>• Alternate sites</li> <li>• Employee support</li> </ul> |  |
| <b>Downstream Data Consumption</b> |                                                                                                                                                                                                                                                                                                                                                                                                                                        |  |
| 1                                  | <p>Capability to easily connect with any downstream applications currently used at the Bank and be flexible to support all end users in the Bank based on future needs.</p>                                                                                                                                                                                                                                                            |  |
| 2                                  | <p>Should support bulk data and targeted data extracts via statistical tools and APIs</p>                                                                                                                                                                                                                                                                                                                                              |  |
| 3                                  | <p>ISDLDW should support open source library / proprietary / supported open source statistical and machine learning tools configured for analysis and reporting</p>                                                                                                                                                                                                                                                                    |  |
| 4                                  | <p>Facility to generate and distribute canned / automatic bursted reports from ISDLDW to downstream end users like BID, Analytics, CRM, YONO, OFSAA, etc</p>                                                                                                                                                                                                                                                                           |  |
| 5                                  | <p>Self-service portal to extract the data on their own (Should support Data Democratization)</p>                                                                                                                                                                                                                                                                                                                                      |  |

|                                                      |                                                                                                                                                                                                                                 |  |
|------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| 6                                                    | Dedicated high-performance department wise sandboxes allocated to end users for R&D                                                                                                                                             |  |
| 7                                                    | Dedicated Data marts to be created as per BFSI industry Standard models and previous experience. All the reporting and analytics request to be sufficed with the same. Requirements for Data Marts to be reviewed periodically. |  |
| 8                                                    | Real time reporting can be done through Staging area of Data Warehouse                                                                                                                                                          |  |
| 9                                                    | The vendor as mandated by the Bank will streamline ad-hoc reports / visualizations using the data discovery capability or assist business teams build pipelines for analysis using data discovery capability.                   |  |
| <b>Data Science Platform with AI/ML Capabilities</b> |                                                                                                                                                                                                                                 |  |
| 1                                                    | Implementing end to end analytics use-cases as mandated by the Bank                                                                                                                                                             |  |
| 2                                                    | Power data / objects to existing analytics models built on proprietary tools (IBM SPSS). Migration of such models to new solution                                                                                               |  |
| 3                                                    | Availability Pre-build models which can be directly used with Bank's data to get insights                                                                                                                                       |  |
| 4                                                    | Model building capabilities on agile mode                                                                                                                                                                                       |  |
| 5                                                    | AI/ML framework that run on scalable architectures                                                                                                                                                                              |  |

|    |                                                                                                                                                                                                  |  |
|----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| 6  | Templates for common data processing / end to end applications, tag data sets and re-use operators across machine learning pipelines                                                             |  |
| 7  | In-memory computing & integration with Spark, Redis, etc                                                                                                                                         |  |
| 8  | Analytics on real-time data in real-time/near real-time                                                                                                                                          |  |
| 9  | GPUs to be incorporated in solution if possible using HDFS Hadoop like environment for better analytical results                                                                                 |  |
| 10 | Analytics notebooks which should be shared, and experiments be replicable across users / teams performing similar analysis                                                                       |  |
| 11 | The ability to automate the process of iteratively searching for the best model from a set of candidates. This feature has also been called "model factory." Platform provide this functionality |  |
| 12 | Platforms facilitate the automation of tasks such as feature engineering and hyper parameter tuning                                                                                              |  |
| 13 | Access management at the data, workflow and models                                                                                                                                               |  |
| 14 | Automate workflow execution and scaling of the servers                                                                                                                                           |  |
| 15 | Build and Publish detailed reports, insights on web portal                                                                                                                                       |  |
| 16 | Deploy models/workflows as industry standard web service                                                                                                                                         |  |
| 17 | Centralized monitoring and management                                                                                                                                                            |  |



|    |                                                                                                                                                                                                                                                                                                                                        |  |
|----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| 18 | All machine-learning platforms either support multiple models out of the box or provide an option to custom-code the same                                                                                                                                                                                                              |  |
| 19 | Integration with R, Python, Keras, Tensorflow, Theano, scikit-learn etc and other frameworks / languages                                                                                                                                                                                                                               |  |
| 20 | Stable Product                                                                                                                                                                                                                                                                                                                         |  |
| 21 | Support on demand basis.                                                                                                                                                                                                                                                                                                               |  |
| 22 | Performance and scalability                                                                                                                                                                                                                                                                                                            |  |
| 23 | Enables ease and speed with which the user can move models from a developed environment to a production environment, or embed them in a business process                                                                                                                                                                               |  |
| 24 | Annexure E gives sample use cases which are to be implemented on Next Gen Data Warehouse using structured and/or unstructured and/or semi-structured and/or any other kind of data gathered from either Data Warehouse or Data Lake or Data Virtualization or all together or any other source.                                        |  |
| 25 | <p>Vendor to provide solution / tool (s) for below scope of activities on SBI data sets;</p> <ul style="list-style-type: none"> <li>• Benchmarking</li> <li>• Predictive &amp; Prescriptive Analytics</li> <li>• Social Media Analytics</li> <li>• Web Analytics</li> <li>• Geolocation Analysis</li> <li>• Ad-Hoc Analysis</li> </ul> |  |



|                                    |                                                                                                                                                                                                                                                      |  |
|------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
|                                    | <ul style="list-style-type: none"> <li>• Trend Indicators</li> <li>• Profit Analysis</li> <li>• In-Memory Analysis</li> <li>• Statistic Analytics</li> <li>• Data Mining</li> <li>• Machine Learning</li> </ul>                                      |  |
| 26                                 | Pipelines to orchestrate the entire machine learning process from data preparation to deployment                                                                                                                                                     |  |
| 27                                 | The ability to launch and redirect training to CPU and GPU-enabled resources.                                                                                                                                                                        |  |
| 28                                 | DevOps - ability to automate build and deploy on all elements of the data science project, from the pipelines to building the models to model deployment.                                                                                            |  |
| <b>Business Intelligence Tools</b> |                                                                                                                                                                                                                                                      |  |
| 1                                  | Capability to connect to various data sources: Access to various databases and file types, such as comma-separated values files, text, Excel and XML, SQL/NoSQL databases. Bank may raise specific requirements as and when required.                |  |
| 2                                  | Data filters and drilldown: The product should enable user to filter the contents in a tabular report or visualization by data values. The product should also enable the user to drill down from summarized to more detailed data and then drill up |  |



|   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |  |
|---|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| 3 | <p>Web-based client user interface: The product's client user interface for the BI consumer role should be web-based.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |  |
| 4 | <p>Independent and interconnected mashups: When the business intelligence style enables a single-screen to display of multiple visualizations, including tabular reports, the software should allow these visualizations to be either independent of each other or interconnected. If they're interconnected, data filters and selections will affect all the visualizations; for example, if any attribute is selected, all the visualizations will share that attribute.</p>                                                                                                                                                                                                                                                                                                                                      |  |
| 5 | <p>Visualizations: BI tools must provide different types of visualizations, including but not limited to:</p> <ul style="list-style-type: none"> <li><input type="checkbox"/> Animations, Barcodes</li> <li><input type="checkbox"/> Bar, line, pie, area and radar chart types</li> <li><input type="checkbox"/> Tables, Graphs, Infographics, Filters</li> <li><input type="checkbox"/> Widgets</li> <li><input type="checkbox"/> Drag and Drop Creation, Customization</li> <li><input type="checkbox"/> Templates</li> <li><input type="checkbox"/> Freehand SQL Command</li> <li><input type="checkbox"/> Geospatial Integration</li> <li><input type="checkbox"/> Layouts</li> <li><input type="checkbox"/> Themes</li> <li><input type="checkbox"/> Ability to mix and match various combinations</li> </ul> |  |



|    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |  |
|----|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| 6  | <p>Reporting on all types of available of Data Formats;</p> <ul style="list-style-type: none"> <li><input type="checkbox"/> Structured, semi-structured, unstructured</li> <li><input type="checkbox"/> Click stream data</li> <li><input type="checkbox"/> Audit Logs</li> <li><input type="checkbox"/> Documents</li> <li><input type="checkbox"/> Multimedia data (Images/Videos/Audios)</li> <li><input type="checkbox"/> XBRL format</li> <li><input type="checkbox"/> IRIS iFILE framework</li> </ul> |  |
| 7  | <p>Security: Both user and user role-based security, designating who can create, modify, publish, use and administer the BI applications.</p>                                                                                                                                                                                                                                                                                                                                                               |  |
| 8  | <p>Print and export: The product must enable you to export print visualizations and tabular reports to PDF or other graphics. Tabular reports need to be exportable to text files at a minimum and, preferably, to spreadsheets.</p>                                                                                                                                                                                                                                                                        |  |
| 9  | <p>Select data for analysis. BI tools must enable the user to select the data used in decision-making analysis and present it as a pivot table-style interface where dimension attributes are placed in rows and columns, measures are selected, and filters are applied.</p>                                                                                                                                                                                                                               |  |
| 10 | <p>Data blending: The product must permit the user to blend data from various data sources. This includes accessing the data and mapping or creating relationships with data from multiple sources.</p>                                                                                                                                                                                                                                                                                                     |  |



|    |                                                                                                                                                                                                      |  |
|----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| 11 | Reporting framework / model – Tool must support “easy to create and maintain” reporting framework / model containing all business concepts / measures used in reporting.                             |  |
| 12 | Create measures: The product must enable the user to create and save measures or calculations for use in analysis. These are also referred to as performance measures or key performance indicators. |  |
| 13 | Create hierarchies: The product must enable the user to create dimensional hierarchies, such as by geography or product, to group and summarize data. This establishes the drill-down paths.         |  |
| 14 | Save queries and analysis: The product should enable the BI user to save the data filters, selections and drill-down paths used in decision-making process so that they can be reused.               |  |
| 15 | Create and publish by business users: The product must enable the user to save and share his or her analysis with other BI consumers.                                                                |  |
| 16 | Context-based filters: Filters will list only the choices that have values that fit the current selection of facts and dimensions.                                                                   |  |
| 17 | Context-based visualizations: Only visualizations or chart types that are relevant to the data selected will be listed as options.                                                                   |  |
| 18 | Collaboration and social interaction: BI tools enable the creation of a business community that can share and                                                                                        |  |



|    |                                                                                                                                                                                                                                                                                                                                                                            |  |
|----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
|    | <p>discuss their decision-making analysis. This would include annotating analysis to share observations and social media, which can enable discussion threads or chats.</p>                                                                                                                                                                                                |  |
| 19 | <p>Storyboarding. Business analysis often involves a process or workflow to analyze different data from different perspectives. Storyboarding enables a series of reports or visualizations to be tied together in a workflow that can be shared.</p>                                                                                                                      |  |
| 20 | <p>Microsoft Office real-time data integration. Beyond simple import and export, the product should provide real-time data integration with Microsoft Office products, which enables business people to embed analytics from the business intelligence tool into a PowerPoint or Excel presentation, for example, and refresh it automatically as the data is updated.</p> |  |
| 21 | <p>Mobile version: BI tools should be able to differentiate between viewing BI applications on a web browser on a mobile device versus a mobile BI application.</p>                                                                                                                                                                                                        |  |
| 22 | <p>In-memory analytics: The product should pull data into an in-memory or locally cached data store preferably columnar is an increasingly popular feature that enables very fast analytics once the data is loaded.</p>                                                                                                                                                   |  |
| 23 | <p>Offline updates: BI tools, when storing copies of the source data in an online analytical processing (OLAP) cube or in-memory columnar data store, should enable business users to schedule automatic data updates.</p>                                                                                                                                                 |  |



|    |                                                                                                                                                                                                                                                                                                                                                                                                                  |  |
|----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| 24 | Performance monitoring: BI applications that monitor report and data usage enable a BI group to improve analytical performance for the business, eliminating bottlenecks and enabling users to assess infrastructure needs.                                                                                                                                                                                      |  |
| 25 | Business intelligence platform administration: Although all BI software should provide code and version management, there are many application development features, such as team development and user administration, that are useful for larger BI deployments.                                                                                                                                                |  |
| 26 | Ease of analytical use: There should be different criteria defined for each type of user, such as information consumer, business analyst and IT.                                                                                                                                                                                                                                                                 |  |
| 27 | Ease of creating BI applications: There should be different criteria for each type of analytics creator, such as business analysts and IT.                                                                                                                                                                                                                                                                       |  |
| 28 | Speed of access: Query performance will vary based on the complexity of the queries and the amount of data involved. Dashboards with multiple visualizations will need to get query results from many queries. The best practice is to create several prebuilt query scenarios and compare how each product performs based on these specific examples. The worse practice is to just arbitrarily rate the speed. |  |
| 29 | The best practice is to establish a testing environment to determine scalability in terms of both the number of                                                                                                                                                                                                                                                                                                  |  |



|                                |                                                                                                                                                                                                                      |  |
|--------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
|                                | concurrent users and data metrics, such as volumes, variety and veracity.                                                                                                                                            |  |
| 30                             | There should be separate criteria for BI user versus administration training. Training may include in-person classes, online classes -- live or prerecorded -- or web recordings for specific features or processes. |  |
| 31                             | There should be separate criteria for BI user online help versus technical documentation.                                                                                                                            |  |
| 32                             | Ability to handle and summarize huge volumes of data. E.g. 30-40 million rows accessed on index and summarized over 5 to 8 metrics.                                                                                  |  |
| <b>Hardware Specifications</b> |                                                                                                                                                                                                                      |  |
| 1                              | Vendor to supply, install, test, commission, manage and maintain the required IT Hardware for Next Gen DW.                                                                                                           |  |
| 2                              | Software and Solution proposed by vendors should be compatible with all types of Hardware.                                                                                                                           |  |
| 3                              | The proposed solution envisages use of commodity hardware, and if any proprietary components are used, should be listed in the response with details and justification.                                              |  |
| 4                              | Vendor proposed hardware is expected to be enterprise class, best of breed, tested and stable release.                                                                                                               |  |
| 5                              | The proposed architecture considers vertical and horizontal scalability as one of the most important design principles.                                                                                              |  |





|    |                                                                                                                                                                                                                                                                                           |  |
|----|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| 6  | Vendor to propose hardware specifications for each component of ISDLDW ecosystem like Data Warehouse, Data Marts, Data Lake, Data Archival, Data Federation/Virtualization, Data Science Platform, Backup, Sandboxes, Functional DR, etc. for PROD, DEV and UAT environment as applicable |  |
| 7  | Vendor must ensure that the proposed servers are fault-resilient with the most comprehensive features and functionalities to ensure maximum system uptime.                                                                                                                                |  |
| 8  | The proposed hardware must not fall into 'End of Support' for at least 7 years from the date of delivery to the Bank.                                                                                                                                                                     |  |
| 9  | The hardware will be delivered in a staggered manner and Vendor to provide a plan for the same                                                                                                                                                                                            |  |
| 10 | Vendor must provide detailed configuration of the proposed Hardware, including Hosting Space Requirements, Racks, Power, Cooling and any other requirement for the fulfillment of the Vendor's obligation in this EOI.                                                                    |  |
| 11 | Vendor will supply hardware resources and related services at the desired locations (Production and DR)                                                                                                                                                                                   |  |
| 12 | The Vendor is required to supply, install, test, commission, monitor, manage and maintain the IT System along with operating system and other peripherals with one-year warranty and AMC for 4 years from the date of delivery at data centers advised by the Bank                        |  |



|    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |  |
|----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| 13 | Vendor to provide all necessary hardware and software required to make the solution work                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |  |
| 14 | <p>The Vendor shall be responsible for the following duly verified by the Bank's team</p> <ul style="list-style-type: none"> <li><input type="checkbox"/> Rack Layout along with cable layout</li> <li><input type="checkbox"/> IP Schema details and VLAN layout</li> <li><input type="checkbox"/> Ensuring proper connectivity and access to ISDLDW setup</li> <li><input type="checkbox"/> Network and Security Hardening</li> <li><input type="checkbox"/> Server Setup and configuration</li> <li><input type="checkbox"/> Storage and backup configuration</li> <li><input type="checkbox"/> Hardware Failover, Functional DR architecture and IT continuity approach (Only related to hardware)</li> <li><input type="checkbox"/> Performance and benchmarking approach and plan</li> </ul> |  |
| 15 | Architecture diagram Deployment plan - Vendor to submit architecture diagram of entire setup with network and security equipment required. Bank may change it after vetting by Information Security Dept and / or Enterprise architecture Dept. It will be binding on vendor.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |  |
| 16 | The Vendor shall ensure all Installations & Implementation to be done by OEM badged resources only which will include;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |  |

- Preparation of racks for hosting including all required cabling & all other activities required for installation of hardware
- Installation and implementation of all the servers (Production, Development, UAT, DR) including network cabling.
- BIOS-configuration configure management port and assigning IP's for managing systems remotely, RAID Configuration.
- OS (Linux) Installation, OS Configuration, OS Management Capabilities, Formatting the disks, OS and Network configuration
- OS packages Repository Configuration.
- Installation and Configuration of Network equipment
- Installation and Configuration of Storage and Backup equipment with Hot, warm and Cold data segregation
- Installation and Configuration of Security equipment
- Setting up the testing of the platform and UAT.
- All work related to cabling to utilize network infrastructure will be done by Vendor.
- All work related to patch panels will be done by Vendor.
- The Vendor shall also carry out OS Hardening, Anti-Virus installation, Create Super user for the Production, DR and UAT/Dev environment according to Bank's policy and secured configuration document

|    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |  |
|----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| 17 | <p>One Time Review Services for Production, DR, UAT and Dev environment;</p> <ul style="list-style-type: none"> <li><input type="checkbox"/> Review the HLD &amp; LLD of the hardware setup</li> <li><input type="checkbox"/> Reviewing the Platform security requirements &amp; security architecture</li> <li><input type="checkbox"/> Reviewing the implementation &amp; administration approach</li> <li><input type="checkbox"/> Reviewing Platform administration and management requirements</li> <li><input type="checkbox"/> Post Hardware Setup Review &amp; sign off based on the design document provided by the Bank</li> <li><input type="checkbox"/> Ensure the High Availability &amp; uptime requirements in line with the SLA</li> </ul> |  |
| 18 | <p>The proposed hardware is mission critical for the proposed project and support of 24 X 7 with an uptime of 99.99 % to be ensured by providing support at PR, and DR site for a period of 5 years.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |  |
| 19 | <p>Vendor will create/update process and process manuals for the management of complete environment. This will also include Installation, Housekeeping, Recovery and other operations manuals and share with the Bank</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |  |
| 20 | <p>Bank will perform a third-party validation of Hardware at its own cost to ensure all framework specifications, deployments are as per guidelines and other parameters as decided by the Bank to ensure proper functioning after commissioning of all equipments.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |  |



|    |                                                                                                                                                                                                                                                                                                                                    |  |
|----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
|    | Vendor will have to close all observations identified in third party validation. Approval from third party vendor of the Bank will be treated as completion of installation successfully.                                                                                                                                          |  |
| 21 | Vendor's team should monitor all supplied components and environment, pro actively report and resolve issues as a part of warranty and AMC obligations.                                                                                                                                                                            |  |
| 22 | The Hardware solution must be compatible to integrate with various systems in the Bank including but not limited to SOC, PIMS, NOC, Command Centre, ITAM, Service Desk, ADS, and SSO etc. at no extra cost. Vendor will have to give appropriate support to the Bank during integration with various components of IT environment. |  |
| 23 | System Administration Support- Service Provider must provide 24X7 supports for the Administration, Maintenance, Up-gradation (related to hardware) and other related activity to keep system running so that high availability can be assured.                                                                                     |  |
| 24 | Vendor needs to provide Helpdesk support 24X7 to the Bank for end to end support for hardware maintenance.                                                                                                                                                                                                                         |  |
| 25 | Vendor will have to provide support for Inventory, Asset register, License for Hardware and bundled software.                                                                                                                                                                                                                      |  |
| 26 | Vendor will have to promptly follow Bank's processes/policies for movement of HW, Manpower Resources in and out of the Data centre.                                                                                                                                                                                                |  |
| 27 | Vendor needs to give full support to the Bank related to any security upgrades, implementations, VAPT closure,                                                                                                                                                                                                                     |  |



|    |                                                                                                                                                                                                                                                                                                                                                                                |  |
|----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
|    | SCD, support to audits, patch management/TL upgrade of HW & bundled software on time to time basis as deemed by the Bank and close all the gaps identified in such activities within given timelines by the Bank as a part of warranty and AMC obligations.                                                                                                                    |  |
| 28 | Vendor will be required to comply with cyber security guidelines of the Bank which shall be revised from time to time as a part of warranty and AMC obligations.                                                                                                                                                                                                               |  |
| 29 | New patches are to be proactively identified, tested and applied by the Vendor during the period of contract. A proper record of testing, installation will be required to be maintained by the Vendor.                                                                                                                                                                        |  |
| 30 | Vendor is required to provide the minimum resources to monitor & manage the infrastructure, however it is the Vendor's responsibility to right size the resources to meet the SLA                                                                                                                                                                                              |  |
| 31 | UAT setup should be exact replica of production (all the business schemas and software / applications / services installed) and to be refreshed every month. Data to be put up in UAT will be finalized by Bank. Dummy dataset to be inserted in UAT to maximize the benefits. Either Data Anonymization on PROD data or any other suitable approach to be proposed by Vendor. |  |
| 32 | Development should be made in DEV only and testing in UAT. Framework to be implemented by Vendor to avoid improper user practices.                                                                                                                                                                                                                                             |  |
| 33 | ISDLDW should support at-least 500 concurrent users, scalable up to 1000 users in next 5 years, running                                                                                                                                                                                                                                                                        |  |



|    |                                                                                                                                                                                                                                                          |  |
|----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
|    | ETL/ELT jobs or doing ad-hoc data extraction requests on database (Not including API based access or scheduled job connections to database)                                                                                                              |  |
| 34 | Next Gen DW should support at-least 200 concurrent users, scalable up to 500 users in next 5 years, on Data Science platform performing advanced analytics and running models. GPUs can be proposed for running Machine Learning/Neural Networks models. |  |
| 35 | The web portal of Business Intelligence tool should support at-least 25000 concurrent users, scalable up to 75000 in next 5 years, accessing various reports generated                                                                                   |  |
| 36 | ISDLDW is expected to have more users and the solution should not be bound by any license model for number of users                                                                                                                                      |  |
| 37 | Ad-hoc jobs of any complexity should not hamper the scheduled jobs performance.                                                                                                                                                                          |  |
| 38 | The ISDLDW solution at SBI should be enabled for exact point-in-time recovery of files or directories in the cluster from accidental deletions or corruption due to user or application error.                                                           |  |
| 39 | While building ISDLDW, failure handling should be a design characteristic. There should be tool(s) that can notify, alert and if possible, predict node, disk or application failure.                                                                    |  |

|    |                                                                                                                                                                                                                                                                                      |  |
|----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| 40 | Necessary corrective action should be in place for each failure. Failure of node, disk or any network component should not affect the functionality of ISDLDW.                                                                                                                       |  |
| 41 | Failure tolerance is not restricted only to server and application failure. Failures of network, hardware, software or nodes are considered failure scenarios in scope.                                                                                                              |  |
| 42 | Standard practices of configuring ISDLDW for High Availability and storage namespace Federation/Virtualization (at least 99.99% - calculated monthly basis) should be implemented. ISDLDW solution should serve users even when confronted with node failures or network partitions. |  |
| 43 | New system should be available 24x7 without any failure.                                                                                                                                                                                                                             |  |
| 44 | Vendor need to propose a solution for data migration / transfer between Existing DWH (Navi Mumbai Location 1) and ISDLDW-PR (Navi Mumbai Location 2) and also between ISDLDW-PR (Navi Mumbai Location 2) and Hyderabad (DR) or any other places for PR and DR decided by the Bank.   |  |
| 45 | Database should be linearly scalable which can expand the database capacity by just adding more nodes to the existing database. If the data volume grows more hardware can be added and expand the database capacity                                                                 |  |
| 46 | DEV and UAT environments are to be set up at the DR site                                                                                                                                                                                                                             |  |





|    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |  |
|----|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| 47 | The Vendor also needs to provide the configuration for setting up of functional DR along with DEV and UAT for each and every component / application. DR, UAT and DEV should be exact replica of production.                                                                                                                                                                                                                                                                                |  |
| 48 | The vendor should provide EXACT size needed for production in the 1st year and estimated sizes for consecutive years keeping in view the growth rate predicted by Bank in this section and provide empirical evidence for the calculation of growth rate.                                                                                                                                                                                                                                   |  |
| 49 | It will be appreciated if the vendor comes with Green computing technology across infrastructure setup (which minimizes the electricity consumption of high computing computers).                                                                                                                                                                                                                                                                                                           |  |
| 50 | Annexure G gives the ISDLDW sizing forecasted for next five years. Sizing given in the Annexure G is in compressed form with compression factor of 2.5. Vendors are expected to propose storage forecast over and above given sizing in the solution to ensure fast performance of system. Vendors shall propose performance benchmark for this system as part of the solution. This is indicated projected growth, but the solution should be scalable enough to store any volume of data. |  |
| 51 | Vendor to submit all back-to-back agreement copies between Vendor and SI / OEM / Parent company etc if any and tenure of the back-to-back agreement should be same as selected Vendor's agreement with the Bank                                                                                                                                                                                                                                                                             |  |

## C5. Abbreviations and definitions

### C5.1 Abbreviations

| As used in the Document | Description                                       |
|-------------------------|---------------------------------------------------|
| ADS                     | Active Directory System                           |
| AML                     | Anti-Money Laundering                             |
| API                     | Application programming interface                 |
| AI/ML                   | Artificial Intelligence/ Machine learning         |
| ADF                     | Automated data flow                               |
| ATM                     | Automated Teller Machine                          |
| BFSI                    | Banking, financial services and insurance         |
| BCBS                    | Basel Committee on Banking Supervision's standard |
| BIOS                    | basic input/output system                         |
| BCP                     | Business continuity planning                      |
| BID                     | Business Intelligence Department                  |
| CPU                     | Central Processing Unit                           |
| CDC                     | Change Data Capture                               |
| CIA                     | Confidentiality Integrity and Availability        |
| CBS                     | Core Banking Solution                             |

|       |                                           |
|-------|-------------------------------------------|
| CRM   | Customer Relationship Management          |
| DAS   | Data Archival Solution                    |
| DC    | Data Centre                               |
| DCT   | Data Collection Template                  |
| DDL   | Data Definition language                  |
| DES   | Data Encryption Standard                  |
| DL    | Data Lake                                 |
| DLP   | Data Loss Prevention                      |
| DQ    | Data Quality                              |
| DWH   | Data Warehouse                            |
| DBA   | Database Administrator                    |
| DEV   | Development                               |
| DFRA  | Digital forensic and readiness assessment |
| DR/PR | Disaster recovery /Production             |
| DR    | Disaster Recovery Centre                  |
| EOD   | End of Day                                |
| EOI   | Expression of Interest                    |
| XBRL  | extensible Business Reporting Language    |
| XML   | Extensible Markup Language                |
| ETL   | Extract Transform Load                    |



|        |                                                      |
|--------|------------------------------------------------------|
| ELT    | Extract Load Transform                               |
| FinCEN | Financial Crimes Enforcement Network                 |
| GDPR   | General Data Protection Regulation                   |
| GUI    | Graphical user interface                             |
| GPU    | Graphics Processing Unit                             |
| HDFS   | Hadoop Distributed File System                       |
| HLD    | High Level Design                                    |
| SCAPM  | IBM Smart Cloud                                      |
| IA     | Information Analyzer                                 |
| IGC    | Information Governance Catalog                       |
| IS     | Information security                                 |
| ITAM   | Information Technology Asset Management              |
| IPR    | Intellectual property Rights                         |
| ISO    | International Organization for Standardization       |
| INB    | Internet Banking                                     |
| IDS    | Intrusion Detection System                           |
| IPS    | Intrusion Prevention System                          |
| ISDLDW | Integrated Solution for Data Lake and Data Warehouse |
| KRI    | key risk indicator                                   |
| LDAP   | Lightweight Directory Access Protocol                |

|        |                                                  |
|--------|--------------------------------------------------|
| LLD    | Low Level Design                                 |
| MDM    | Master data management                           |
| MD5    | Message-digest 5                                 |
| NOC    | Network Operations Centre                        |
| NTFS   | New Technology File System                       |
| ISDLDW | Next Generation Data warehouse                   |
| NPA    | Non-performing asset                             |
| OFAC   | Office of Foreign Assets Control                 |
| OCC    | Office of the Comptroller of the Currency        |
| OLAP   | Online analytical processing                     |
| OFSAA  | Oracle Financial Services Analytical Application |
| OEM    | Original equipment manufacturer                  |
| PCIDSS | Payment Card Industry Data Security Standard     |
| PSG    | Payment System Gateway                           |
| PIMS   | Privileged Identity Management System            |
| POC    | Proof of Concept                                 |
| PKC    | Public-key cryptography                          |
| RTC    | Rational Team Concert                            |
| RPO    | Recovery Point Objective                         |
| RTO    | Recovery Time Objective                          |



|       |                                            |
|-------|--------------------------------------------|
| RAID  | Redundant Array of Independent Disks       |
| RDBMS | Relational database management system      |
| RFP   | Request for Proposal                       |
| R&D   | Research and development                   |
| ROI   | Return on investment                       |
| SCD   | Secure Configuration Document              |
| SEC   | Securities and Exchange Commission         |
| SOC   | Security operations Centre                 |
| SLA   | Service Level Agreement                    |
| SPOC  | Single Point of Contact                    |
| SSO   | Single Sign on                             |
| SME   | Small and Medium Enterprise                |
| SBI   | State Bank of India                        |
| SPSS  | Statistical Package for the Social Science |
| SQL   | Structured Query Language                  |
| TL    | Technology Level                           |
| TWS   | Tivoli Workload Scheduler                  |
| UAT   | Unit Testing                               |
| UAM   | User Access Management                     |
| VLAN  | Virtual LAN                                |

|      |                                                  |
|------|--------------------------------------------------|
| VAPT | Vulnerability Assessment and Penetration Testing |
| YONO | You only need one                                |

### C5.2 Definitions

| As used in the Document | Description                                                                                                                                                                    |
|-------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| ADS                     | Microsoft Active Directory services of the Bank for controlling user access                                                                                                    |
| Audit Logs              | Security related chronological record of events in various systems.                                                                                                            |
| Automation              | Process or procedure performed with minimum or no human assistance                                                                                                             |
| Bank                    | State Bank of India (Domestic, Foreign branches/subsidiaries) combined                                                                                                         |
| Bank IS Policy          | Latest IS policy of the Bank                                                                                                                                                   |
| Business Glossary       | Glossary of Business concepts and terms defined and used by Bank time to time                                                                                                  |
| Business Intelligence   | Creation of architecture to create and deliver business values from the raw data of the Bank                                                                                   |
| Business Metadata       | Business attributes to help, promote the context meaning and search ability. It gives business users a way to search and understand data attributes based on business concepts |

|                 |                                                                                                                                                                         |
|-----------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Business Units  | Branches, Extension Counters, Service Units, Corporate Offices, Head Offices, Zonal Offices, Regional Offices, Other Back Offices/ Offices and any new branches/offices |
| CDC             | Change Data Capture or any process of scrapping prevailing in any of the Bank's system either existing or proposed                                                      |
| Checkpoints     | Snapshot at a specific time from which activity can be resumed without any loss of transformation/modification of data                                                  |
| Cloud           | Data centres available to many users over internet or intranet                                                                                                          |
| Columnar        | Column based storage of data either in in-memory or in secondary storage                                                                                                |
| Data Stores     | Repository of heterogeneous data                                                                                                                                        |
| Data Warehouse  | Existing Data Warehouse system of the Bank                                                                                                                              |
| Diagnostic Logs | Detailed report of activities performed / being performed by the System which will be used to diagnose any problem faced                                                |
| DLP             | Data Loss Prevention is a strategy for making sure that users do not send sensitive or critical information outside the corporate network                               |
| Edge Servers    | A device providing entry point into the network                                                                                                                         |
| Error Logs      | Detailed report of errors in the System which will be used to resolve same                                                                                              |
| Existing Setup  | Existing Data Warehouse ecosystem of the Bank along with upstream and downstream links/interfaces/systems                                                               |





|                      |                                                                                                                                                                                                                                                                  |
|----------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Functional DR        | DR with all the schemas/applications of existing setup with limited data as identified by the Bank                                                                                                                                                               |
| LHOs/GOCs            | Local Head Offices/GITC Offsite Centre of the Bank                                                                                                                                                                                                               |
| Near Real Time       | Near real time refers to processing data as and when available with an acceptable processing time of a few minutes.                                                                                                                                              |
| Operational Metadata | Lineage of data along with audit trails of success and failure of ingestion pipelines duly updated in a pre-determined frequency                                                                                                                                 |
| Processing Pipelines | ETL, ELT, any data processing / transformation / sourcing/ extraction jobs                                                                                                                                                                                       |
| Project              | The Vendors' delivery of the services, deliverables, supply, installation, testing and commissioning, integration of computer hardware / software and services with support under Warranty and annual maintenance contract, if required for the contract period. |
| Proposed Solution    | Solution to be proposed by Vendors vide this EOI                                                                                                                                                                                                                 |
| Record Rejection     | Rejection of full or part of data due to non-conformance to various rules set in system                                                                                                                                                                          |
| Regulatory Reporting | Mandatory reporting to various regulators by the Bank                                                                                                                                                                                                            |
| Reject Processing    | Successful loading of reject records in reasonable time ensuring no impact on downstream processing/reporting                                                                                                                                                    |
| Semantic Integration | Process of integrating heterogeneous sources of data and interrelating them by leveraging the semantic information embedded within these sources                                                                                                                 |



|                                   |                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|-----------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Services                          | Services ancillary to the supply of the equipment/product, such as transportation, transit insurance, installation, commissioning, customization, integration, provision of technical assistance, training, maintenance and other such obligations of the Vendor covered under the Contract.                                                                                                                                                                 |
| Solutions/ Services/ Work/ System | “Solution” or “Services” or “Work” or “System” or “IT System” means all services, scope of work and deliverable to be provided by a Vendor as described in the EOI and include services ancillary to the development of the solution, such as installation, commissioning, integration with existing systems, provision of technical assistance, training, certifications, auditing, maintenance and other obligation of the Supplier covered under the EOI. |
| Storage Subsystem                 | Includes all the components of proposed data storage like Data Warehouse, Data Marts, Data Lake, Data Stores, and any other storage proposed by Vendor                                                                                                                                                                                                                                                                                                       |
| Supported Open Source             | Open source software supported by Vendor or his partners/joint ventures/parent company                                                                                                                                                                                                                                                                                                                                                                       |
| System                            | Any existing legacy system in the Bank and proposed ISDLDW by Vendor                                                                                                                                                                                                                                                                                                                                                                                         |
| Technical Metadata                | Technical attributes including data type, length, precision, format, nullability, etc                                                                                                                                                                                                                                                                                                                                                                        |
| Tranche 1                         | Mandatory reporting to Reserve Bank of India (RBI) for Risk Based Supervision (RBS)                                                                                                                                                                                                                                                                                                                                                                          |



|            |                                                                                            |
|------------|--------------------------------------------------------------------------------------------|
| Validation | The process of checking or establishing the validity or accuracy of data/hardware/software |
|------------|--------------------------------------------------------------------------------------------|

**Name & Signature of authorised signatory**

**Seal of Company**



**Appendix-D: Bidder Details**

Details of the Bidder

| S. No. | Particulars                                                                                                                                                                             | Details |
|--------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| 1.     | Name                                                                                                                                                                                    |         |
| 2.     | Date of Incorporation and / or commencement of business                                                                                                                                 |         |
| 3.     | Certificate of incorporation                                                                                                                                                            |         |
| 4.     | Brief description of the Bidder including details of its main line of business                                                                                                          |         |
| 5.     | Company website URL                                                                                                                                                                     |         |
| 6.     | Company Pan Number                                                                                                                                                                      |         |
| 7.     | Company GSTIN Number                                                                                                                                                                    |         |
| 8.     | Particulars of the Authorized Signatory of the Bidder<br>a) Name<br>b) Designation<br>c) Address<br>d) Phone Number (Landline)<br>e) Mobile Number<br>f) Fax Number<br>g) Email Address |         |

**Name & Signature of authorised signatory**



**Seal of Company**



**Appendix-E: Pre-Bid Query Format**

**(To be provide strictly in Excel format)**

| <b>Vendor Name</b> | <b>Sl. No</b> | <b>EOI Page No</b> | <b>EOI Clause No.</b> | <b>Existing Clause</b> | <b>Query/Suggestions</b> |
|--------------------|---------------|--------------------|-----------------------|------------------------|--------------------------|
|                    |               |                    |                       |                        |                          |
|                    |               |                    |                       |                        |                          |
|                    |               |                    |                       |                        |                          |
|                    |               |                    |                       |                        |                          |



**Appendix-F: Format for Submission of Client References**

**To whosoever it may concern**

| <b>Particulars</b>                         | <b>Details</b> |
|--------------------------------------------|----------------|
|                                            |                |
| <b>Client Information</b>                  |                |
| Client Name                                |                |
| Client address                             |                |
| Name of the contact person and designation |                |
| Phone number of the contact person         |                |
| E-mail address of the contact person       |                |
| <b>Project Details</b>                     |                |
| Name of the Project                        |                |
| Start Date                                 |                |
| End Date                                   |                |
| Current Status (In Progress / Completed)   |                |
| <b>Size of Project</b>                     |                |



Value of Work Order (In Lakh) (only single work order)

Appendix-G

### Certificate of Local Content

<Certificate from the statutory auditor or cost auditor of the company (in case of companies) or from a practicing cost accountant or practicing chartered accountant (in respect of suppliers other than companies) giving the percentage of local content, on their letter head with Registration Number with seal.>

Date:

To,

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

Dear Sir,

Ref.: RFP No. : \_\_\_\_\_ Dated: \_\_\_\_\_

This is to certify that proposed \_\_\_\_\_ <details of services> is having the local content of \_\_\_\_\_ % as defined in the above mentioned RFP.





2. This certificate is submitted in reference to the Public Procurement (Preference to Make in India), Order 2017 including revision thereto.

**Signature of Statutory Auditor/Cost Auditor**

**Registration Number:**

**Seal**

**Counter-signed:**

**Bidder**

**OEM**

< Certified copy of board resolution for appointment of statutory/cost auditor should also be enclosed with the certificate of local content.>

**OR**

**Format for Self-Certification of Local Content**

**Date:**

To,



---

---

---

Dear Sir,

Ref.: RFP No. : \_\_\_\_\_ Dated: \_\_\_\_\_

This is to certify that proposed \_\_\_\_\_ < details of services > is having the local content of \_\_\_\_\_ % as defined in the above mentioned RFP.

1. The details of location(s) at which the local value addition is made are as under:

| SI No | Product details | Name of place |
|-------|-----------------|---------------|
| 1     |                 |               |
| 2     |                 |               |

3. This certificate is submitted in reference to the Public Procurement (Preference to Make in India), Order 2017 including revision thereto.

**Signature of authorised official**

**Name:**

**Company seal:**