

Request for Proposal (RFP)
For
Program DRISHTI
(Data Re-engineering for Insights and Strategic Transformation Initiative)

RFP No : SBIL/2024-25/D&I/001 Dated 26/11/2024

Corrigendum 6

Confidentiality – All information included in this RFP and contained in any subsequent communication/s is confidential and is intended only for the recipient’s knowledge. No information included in this document or any subsequent communications or disclosed in any discussions connected to it can be disclosed to any other party. Receipt and viewing of this document imply acceptance of the above confidentiality norm.

SBI LIFE INSURANCE COMPANY LTD., having its registered office at “Natraj”, M. V. Road & Western Express Highway Junction, Andheri (East), – Mumbai 400 069 and its Central Processing Centre at 8th Level Seawoods Grand Central, Tower 2, Plot No R-1, Sector 40, Seawoods, Nerul Node, Navi Mumbai- 400706.

Public

Sr. No	Query/Suggestions	Answers
1	<p>Does SBIL have an existing virtual platform E.g. Vmware, which can be leveraged by the selected partner to deploy MDM and the other necessary Data Products? If yes, would SBIL have enough capacity available or would the selected partner need to provision the licenses for this additional capacity?</p> <p>If VMware virtual platform is present, then is Site Recovery Manager used or any other recovery tools are used for DR?</p>	<p>For the analytical MDM solution built by bidder please consider provisioning infrastructure capacity appropriately</p>
2	<p>Also, from your previous responses we understand that the existing platform monitoring tools and system management toolsets (e.g. backups, logging, etc) can be leveraged for the purpose of additional data product deployment. Kindly confirm the same.</p>	<p>Yes, this is correct</p>
3	<p>It was mentioned in the RFP to use the existing storage investment, does this mean that the new MDM can leverage the existing storage platform to store all the customer 360, product 360 data? Will the existing storage have sufficient storage capacity in the range of around 80 - 100 TB of storage?</p>	<p>Based on the volumetrics provided the bidders is required to estimate if additional storage is required for the proposed solution. Existing Storage is enough for current MIS & Regulatory reporting.</p>
4	<p>As per the RFP document, we understand that there is a requirement to master Customer and Product. Please let us know if we should consider mastering Policy (it will be an association of Product and Customer), Employees, Prospects and leads also? Also, we have assumed from the volumetric analysis that the Customer count is around 3 Million and Employees are around 18K. Please provide us the records count which will help us in doing the proper sizing for hardware and software?</p>	<p>Bidder to propose the master data solution as per the requirements outlines in the SoW. Volumetric count of customers, employees etc can be assumed based on size & scale of SBI Life and publicly available information</p>
5	<p>Regarding Qlik and SAP Crystal reports, there are nearly 800 report that are being repointed and 250+ new reports are being rebuilt. We assume that existing 800 reports will be supported by existing team and only new 250+ reports would be in scope as part for this RFP. Can we please get a confirmation on this?</p>	<p>Existing reports need to ensured that are working right after any enhancements to the data model. Please consider that existing report count is not exceeding 1300 reports</p>
6	<p>Could you please clarify the operational tooling and access requirements for the support team. Specifically whether we should utilize client-provided monitoring and incident management systems or propose our own solutions. Also Please confirm that Service Desk (L0) setup is not in scope and L1 is considered for Application management and monitoring.</p>	<p>Only L1 in scope should be considered, SBIL will provide the incident management systems. Bidder to provide Monitoring tools required for the data platform</p>

7	<p>Could you please clarify if the infrastructure required for this solution needs to be provided by the SI or SBI Life existing infrastructure (including Active Directory, security services, core switches, routers, firewalls, load balancers, Servers, Switches, Racks, Network equipment, front-end backup solution) will be leveraged? We assume that existing Infrastructure and platform like Oracle Data Warehouse, Qlik & Crystal reports environments and Network changes (Routers, switches, load balancers, firewall) would be reused and hence Admin support for these environments would be out of scope as part for this RFP. Please confirm.</p>	<p>Existing infrastructure for Oracle Data Marts, Qlik will be reused while for other components proposed by the SI, infrastructure will have to be provided and setup by SI. Bidder to procure only additional servers as necessary rest of the equipment such as racks, cables switches to be provided by SBIL. Servers need to be compatible with existing infrastructure equipment</p>
8	<p>Will additional components like OEMS/OCI for Oracle need a separate EULA with Oracle or do the existing EULA also cover any other components of Oracle that can be leveraged? Regarding the timeframe for commercial proposals, the RFP mentions implementation and operate costs for a 3-year period, while the commercial template references a 5-year Total Cost of Ownership. Could you please confirm which timeframe should be used for our commercial proposal - 3 years or 5 years?</p>	<p>Additional components will require a separate EULA. The time frame for commercial proposal is 5 years and contract is for 3 years extendable to 5 years overall.</p>
9	<p>As per our understanding, setting up of this new platform (H2O.ai) on AWS and Migration of existing AI/ML models to the new AI H2O.ai platform is out of scope of this proposal. Can you please confirm this understanding?</p>	<p>Yes, this is correct</p>
10	<p>Do we need to move only the staging layer to Data Lakehouse which is basically getting data from all source systems and storing them in same table structures as source systems?</p> <p>Is there a need to get business/technical users do direct sql queries/data dump on lakehouse data</p>	<p>Please refer to the RFP and corrigendums for the architecture requirements on the staging layer. There will be a need for the analytical users and actuarial team to do a direct query on the lakehouse data</p>
11	<p>As shown below, there are two values for uptime specified in each part of the Service Uptime Penalty section, and they are not equal. So, as bidders, what uptime value should we consideration?</p>	<p>The bidder should consider the uptime differently for different services</p>
12	<ol style="list-style-type: none"> 1. Bidder expects SAST and DAST solution is in-pace and bidder has to leverage/manage the same. If Yes, Kindly share the name of OEMs. 2. Please confirm number of applications to be scanned 3. Application security solutions are integrated with CI/CD pipeline or bidder needs to integrate 	<p>As mentioned in the guardrails SBIL currently has Fortify which is used for security and vulnerability testing as well</p>

13	<p>As per the pre-bid response for Sr.104 for backup solution sizing, as per SBIL response it is not in scope of the bidder. Therefore, bidder assumes that the existing Backup solution will be leveraged for any Backup requirement. Please confirm.</p> <p>Is the bidder expected to provide BAU support for the proposed infrastructure? If yes, where should this cost be reflected in the Commercial template?</p>	<p>Backup solution sizing is not part of bidder's scope as mentioned in the pre-bid queries.</p> <p>Yes, BAU support cost to be provided by Bidder in the Commercial Sched - Support Tab.</p>
14	<p>These reports are pre-built on Qlik Sense and getting the data from Oracle using an existing data model. Bidder is not making changes any of these tools. Are these reports currently getting completed in 3 seconds or bidder has to further optimize the reports to complete them in 3 seconds? This applies to other reporting requirements as well?</p>	<p>Most of the reports are completed within 3 seconds any further optimization to be done by the bidder.</p>
15	<p>While the average turnover over the past three years exceeds ₹500 crore, the turnover for one of the individual years falls below ₹500 crore. In this context, would we still meet the eligibility criteria?</p>	<p>Yes, the eligibility criteria is met with the average turnover.</p> <p>The Bidder must be a profitable organization, having achieved profit before tax (PBT) in at least two of the last three financial years, with a minimum average standalone turnover of ₹500 Crores during the same period (2021-2022, 2022-2023, and 2023-2024)</p>
16	<p>Service Agreement Clause 3.7 This clause has reference to regulation 46 of IRDAI (Protection of Policyholders' Interests, Operations and Allied Matters of Insurers) Regulations, 2024. The said regulations pertain to opening of foreign branches by India based Insurers. This may need to be rechecked and correct regulation number should be incorporated, are we correct in assuming the same?</p>	<p>To be discussed with selected bidder and changed if needed.</p>
17	<p>Who will be responsible to Installation and management of Hardware racks .</p> <p>Who will be responsible for software Installation .</p>	<p>Bidder to procure only additional servers as necessary rest of the equipment such as racks, cables switches to be provided by SBIL. Servers need to be compatible with existing infrastructure equipment.</p> <p>The bidder will be responsible for Software installation and setup and also mentioned in the costing sheet</p>
18	<p>what should be size of delta data, expected in CDC for daily batch processing?</p> <p>What is size of each message (max size and min size of outliers with avg sizing).</p>	<p>Refer Corrigendum. 2 & 3 For CDC details</p>
19	<p>How the current architecture supports scheduling and can it be reused or new procurement is expected</p>	<p>Scheduling of ETL and transformation jobs will have to be taken care off in the New Data Platform</p>
20	<p>We have been asked to comply with 'Third Party Security Control Checklist'. We were not able to locate this document, would you please guide us to get the same, please.</p>	<p>To be provided to the selected bidder</p>

21	Request you to please provide an extension of 2 weeks for the bid submission. Due to involvement of multiple OEMs and in our pursuit of designing optimal solution, we kindly request extension of the deadline.	Please refer Corrigendum. 5
22	We have reviewed the documents and corrigendum and there was information mentioned about structured and un-structured data, however, the volumetrics for semi-structure is not provided explicitly. The same has been cross verified in the Questionnaire response document as well. Kindly share the information/details/volumetrics on semi-structured data.	Please assume 30% of the unstructured data to be considered as the semi-structured data
23	Volumetrics details provided for On-Prem Workloads in Table 1: 1.3.2.12.1 We have below observations- 1. The year-on-year growth mentioned is 100% in some cases and drastically falling to lowest in the subsequent year and again growing to 100% in the 3rd year. 2. Year on year comparison is also diminishing in some cases. Please confirm if we can go ahead with the volumetrics shared as it will have an impact on the overall sizing of the solution.	The volumetrics provided are correct
24	For eligibility, kindly consider and confirm that implementations wherein we have partnered with our overseas firm in developing solution for a client is acceptable. Due to confidentiality agreement with the client, we would be able to share limited project specifics, but when instructed and required by SBI Life we can facilitate a meeting with the client for endorsement of the work delivered by us and details over the same.	We are expecting that a Significant portion of the scope in terms of design & build of the project is done by the India firm
25	Definition of Presentation Layer – By presentation layer, do we mean it will consist of aggregated business ready data sets for consumption and analytics?	Yes, this is correct
26	Bidder requests the cloud workload will be proposed by bidder and SBIL will allocate the same since SBIL has already in contract with AWS& GCP.	Cloud workloads along with costs to be proposed by bidder and SBIL will eventually decide on allocation
	So, request to remove it from customer template since the cloud price will be variable based on consumption of the IO, Network throughput and number of API call. These variable parameters cannot be controlled.	
27	Bidder understand the "Data Catalog and Data Quality, Business Glossary" on cloud is not a mandatory. Request to confirm.	It is not mandatory but preferred from an optimized architecture perspective

28	In Prebid response, it said that "Total Contract Period is of 3 years, including build, development and support. Bidder to propose break- up of these 3 years with set-up, implementation and support timelines." whereas the commercial template asked for 5-year price. Request to confirm we have to provide for 5 years or the commercial template will be revised?	The time frame for commercial proposal is 5 years and contract is for 3 years extendable to 5 years overall.
29	there are 3000 reports to be repointed whereas inventory of 1300 is given. Additional 1700 details are not provided. Can you details of those?	Please consider that existing report count is not exceeding 1300 reports
30	There is no clarity of use cases to be developed. We request to share at least the description of use case to estimate the effort.	Bidder to propose and provide 5 use cases basis the insurance industry knowledge and expertise
31	1. How is the Data Cache Layer expected to be used ? As per the Future state diagram (2/2) data is ingested into the "Real-time App Store" by the Ingestion layer and then possibly consumed by Applications on the Consumption layer (Including but not limited to GenAI/LLM Apps). · Is the Data cache going to be used as a Vector Database? · Will it also be connected to the Lakehouse and/or data warehouse as a simple performance acceleration engine · Are Aggregates and Summaries also planned to be created directly in this Cache layer for which data may be coming from DWH, Lakehouse and other systems	Data Cache is not a vector Database but part of real - time app store. Real-time transformation for aggregate and summaries will also be needed in the cache layer with data coming in from the Lakehouse
32	1. Please help to understand the use case for Classify because Classify the data required 3rd party tool for data classification ? Please clarify is there is requirement for 3rd party data classification tool ? 2. We will tokenize the data when moving PII data to cloud and mask the data on-prem so that it can be consumed via application(on-prem/Cloud) ? Is that understanding is correct ?	There is no requirement for a 3rd party data classification tool The understanding is correct for Pt.2
33	We understand that there is a existing connectivity between DC and Public Cloud will be leverage by the successful bidder? Please confirm	Yes, this is correct
34	How many news sources are expected for feeds into data mart / ODS?	Consider 5 additional sources beyond the ones provided in the current state.
35	Cache DB- can you detail on the expectations w.r.t cache requirement?	Data Cache is not a vector Database but part of real - time app store. Real-time transformation for aggregate and summaries will also be needed in the cache layer with data coming in from the Lakehouse
36	Since only the analytical use case is expected to on-Prem, can we not leverage on-prem DG solution? Can you share the rationale behind ask for cloud-based DG solution?	Analytical data marts/Data transformation required prior to analytical modelling is required on cloud hence the DG solution is preferred on-prem. Note that it is preferred not mandatory

37	Will SBI Life continue with VAPT services of existing vendor and bidder does not need to consider the same. Please confirm. If No, does VAPT need to be considered for the entire Data Platform (Reuse and Proposed tools) within SBI Life network ? Please clarify scope of VAPT for Internet facing applications as mentioned in RFP?	For on-premises setups, SBI Life will be responsible for VAPT, whereas for cloud environments, it falls under the scope of the system integrator (SI). The SI must submit VAPT reports to SBI Life on a timely basis
38	Please confirm if SBI Life has EMS,NMS Tool to monitor the Infrastructure and network devices, same can extend to the bidder for this project.	Bidder to use existing infrastructure monitoring tools
39	Clarify for which reuse components does need to consider 5 yr support. (e.g. Oracle , Reporting Tools, Security tools , IBM ESB, Devops tools)	All the required tools will need 5-year support
40	CDC workload has been clarified in pg 3 Corrigendum 3.: "40%-50% of the daily archival volumes to be the peak log size for CDC". This works out to be approx. 600GB which is very high. Please share exact size to be considered.	Please assume this for CDC sizing
41	We understand that, the Landing zones in DC is already available for Google/AWS platform, is to be reused for this program. Please let us know if Landing zones for GCP/AWS is also available in DR that can be leveraged, or SBI Life will make it available for us to be used for DR? Please confirm.	Bidder to use existing landing infrastructure for DR