RFP for Development of Data, Reporting and Analytics Solutions

Pre-bid Session

MAY 2024

Overview of Business & Functional Requirements

Vision

LIC is embarking on a digital transformation journey to become the most valued life insurer globally

To meet this bold vision, LIC will become a technology driven life insurer supported by:
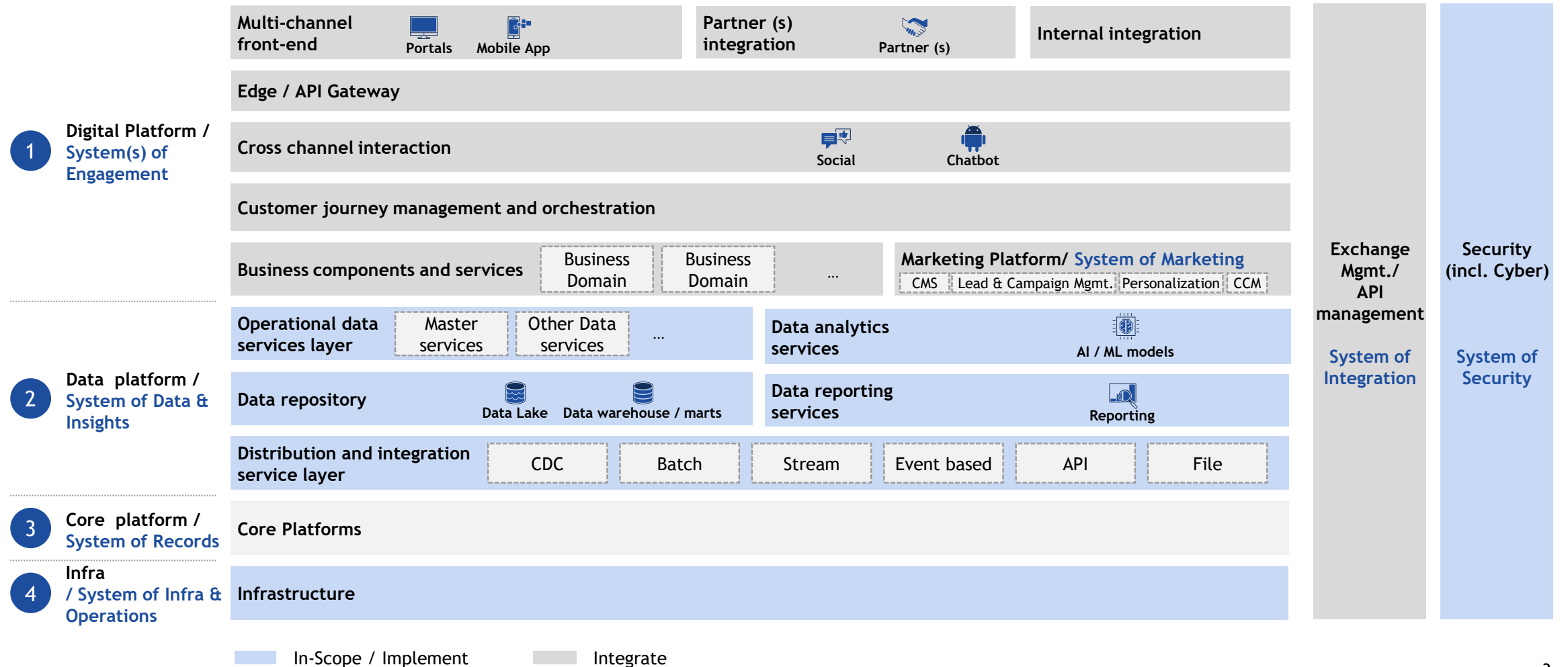
Next gen Data Lake

to efficiently manage and analyze vast amounts of data, leading to better decision-making and enhanced insights

Analytics and BI

to gain actionable insights, optimize operations, and improve risk management and customer satisfaction.

Data, Analytics and Reporting platform is an integral part of LIC's overall Digital Transformation journey

1. Digital Platform / System(s) of Engagement

Multi-channel front-end	Portals, Mobile App	Partner (s) integration	Partner (s)	Internal integration

Edge / API Gateway

Cross channel interaction — Social, Chatbot

Customer journey management and orchestration

Business components and services	Business Domain	Business Domain	...	Marketing Platform/ System of Marketing
				CMS, Lead & Campaign Mgmt., Personalization, CCM

2. Data platform / System of Data & Insights

Operational data services layer	Master services	Other Data services	...	Data analytics services	AI / ML models

Data repository	Data Lake	Data warehouse / marts	Data reporting services	Reporting

Distribution and integration service layer	CDC	Batch	Stream	Event based	API	File

3. Core platform / System of Records

Core Platforms

4. Infra / System of Infra & Operations

Infrastructure

Exchange Mgmt./ API management — System of Integration

Security (incl. Cyber) — System of Security

Legend: In-Scope / Implement | Integrate

3

Overview of project deliverables

Data Lake / Lakehouse platform

- Install suitable software and infrastructure components to build an end-to-end lakehouse platform.
- Perform data engineering tasks including ETL/ELT, and develop various data repositories as part of the overall data lake / lakehouse platform
- Create suitable data models and schemas to facilitate fast and relevant reporting and dashboarding
- Ingest structured data from multiple internal and external sources in real-time and batch mode.
- Ingest semi-structured data and unstructured data in real-time and batch mode.
- Enable end to end data governance including data lineage, metadata management, data quality, etc
- Make suitable data services available for consumption by other applications.

Advanced Analytics

- Implement MLOps to manage the overall model lifecycle development, testing, training and refinement.
- Implement and establish a data science workbench.
- Develop, test, and deploy models according to stated requirements.
- Integrate MLOps with DevSecOps to manage automated releases into production.
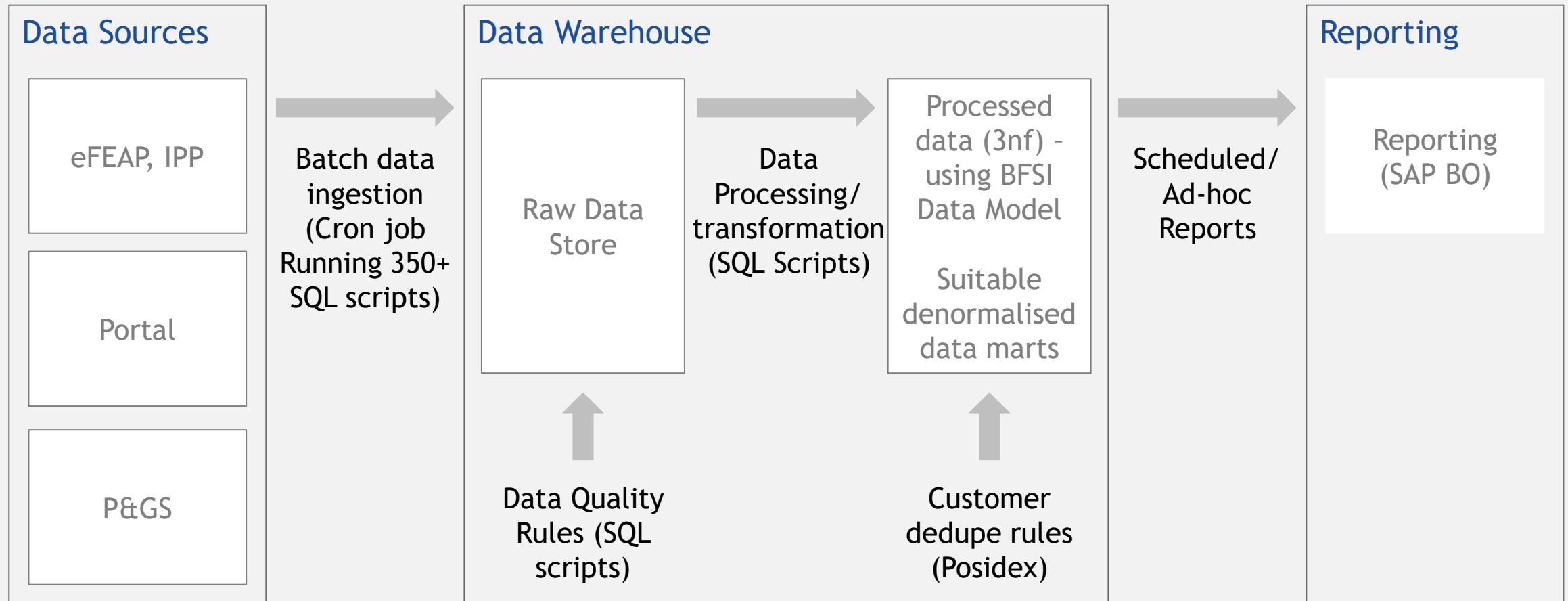
Reporting and Dashboarding

- Implement and deploy the reporting solution.
- Develop, test, and deploy reports according to stated requirements.
- Make the ad-hoc reporting solution available for end-users.
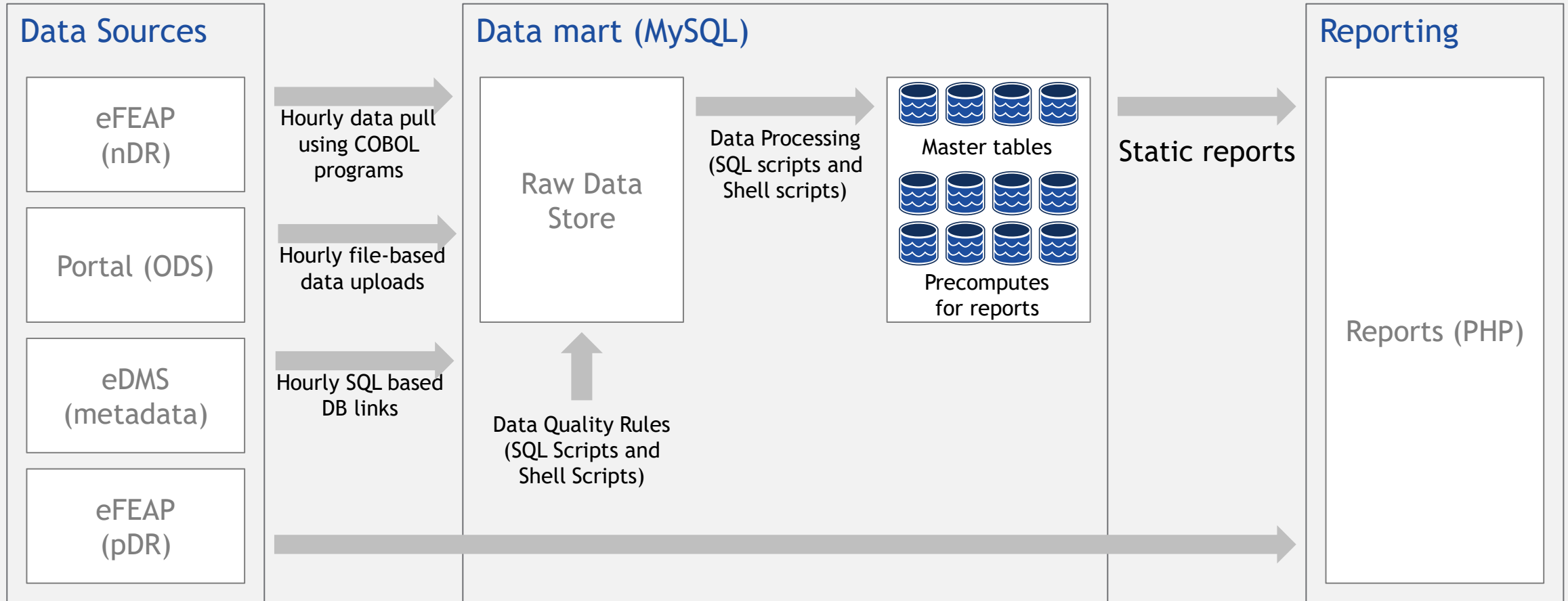
Outline of Current State & the required Target State Architecture

LIC data platform CADW | Current State

This is a centralized data warehouse platform that LIC implemented in 2004 to address MIS and dashboarding needs. The primary structure of this platform is as shown in the diagram below.

LIC's Reporting Mart COMIS| Current State

Data Sources

- eFEAP (nDR)
- Portal (ODS)
- eDMS (metadata)
- eFEAP (pDR)

Hourly data pull using COBOL programs

Hourly file-based data uploads

Hourly SQL based DB links

Data mart (MySQL)

Raw Data Store

Data Quality Rules (SQL Scripts and Shell Scripts)

Data Processing (SQL scripts and Shell scripts)

Master tables

Precomputes for reports

Static reports

Reporting

Reports (PHP)

Key Source Systems (1/4)

Core Platform: eFEAP

This is the core operational platform for the entire life business for LIC. This contains data around all customers, policies issued, premium collected, agent hierarchy, commissions and bonuses, etc.

Nature of application	Custom developed
Tech stack	Business logic: Cobol (primary) Java / STRUTS (some functionality) Web services: Java (SOAP and REST) Database: MySQL
Structure	• Overall eFeap has around 60 modules for different business functions including new business, marketing, customer service, individual pension plans (which has its independent set of databases), accounting, internal office functions, etc. • There are 125 separate instances of eFeap for divisions, ZOs, CO and other specialized areas. Hence there are 125 databases containing the data for these divisions. Each database is roughly 1.8 TB in size. • Each instance has about 3000 tables including master tables, transaction tables, control tables and interim tables.
Volume of data	Approx 200 TB.

Key Source Systems (2/4)

Core Platform: UCS
This is the core operational platform for the ULIP products. This contains data for all the ULIP policies.

📋 Nature of application	⚙️ Custom developed
Tech stack	Business logic: Java / STRUTS Web services: Java (REST) Database: MySQL
Structure	• This is a centralized application that stores policy information for ULIP products, computations around those plans, premium collections, customer service elements, etc • The application has 11 databases for the 9 zones. The total data volume is around 1 TB. The table structure is very similar to eFEAP table structure with some additional fields related to ULIP.
Volume of data	Approx. 1 TB

Core Platform: P&GS
This is the core operational platform for the pension and group business. This contains all the relevant data for that part of LIC's business.

📋 Nature of application	⚙️ Custom developed
Tech stack	Business logic: Java / STRUTS / PowerBuilder Database: Oracle
Volume of data	Approx. 1.8 TB

Key Source Systems (3/4)

Core Platform: Investment management
This is the core platform to manage the investments that LIC does. This addresses the front, mid and backoffice investment management needs.

📋 Nature of application	🔼 Off the shelf
Tech stack	Application: SAP Treasury Database: SAP
Volume of data	Approx. 1 TB

Document management: eDMS
This is used to digitally store customer related documents such as policy documents, claims, certain KYC documents, etc.

📋 Nature of application	🔼 Off the shelf
Tech stack	Application: Newgen Omnidocs and OmniXtract Web services: Java (REST) Database: Postgres
Volume of data	Approx 1.8PB

The relevant data (expected to be metadata around the documents – number and nature of documents, medical report extracts) extracted from the digital documents is expected to be ingested into the data platform.

Key Source Systems (4/4)

Data Platform: CADW

This is a centralized datawarehousing platform that LIC implemented in 2004 to address MIS and DSS needs. This is expected to be a one-time data load at the start of the program for historical data.

Data sources	• eFEAP: 600+ of the 3000+ tables within eFEAP are the sources from within eFEAP. This constitutes more than 90% of the data within CADW. • IPP: Pension plan data • Portal: Online business data • P&GS: Pension and Group business data
Data warehouse	The datawarehouse software is from Opentext Vertica (v9.1). This is a columnar database. There are two data stores within Vertica: • Raw data store: Here data comes in as per the structure in the source systems. There are data quality checks developed using SQL scripts (approx. 350) that run at this stage and identify data issues and errors. • Processed data store: The data in the raw store post cleansing and identification of unique customers are then processed to form the data store required for the reporting. This uses Teradata's logical data model for financial institutions. The processing / transformation here is done using SQL scripts. At this stage, unique customers are identified using a tool named Posidex. Customer IDs are generated and associated with the customers.
Reporting	SAP BO (v4.1) is used for BI / reporting. There are approximately 7900 ad-hoc reports that exist in the system.
Volume of data	Approx 100 TB

Target State Data Architecture | Guiding Principles (1/2)

User centricity: The solution would be centred around business or customer and their requirements and work backwards to design and deliver specific solutions.

Reusability: The use cases, data models, data marts, APIs, etc. should be designed in a way that can be extensible and reusable across different business groups and personas

Automated: The solution should automate the data pipelines end-to-end starting from data ingestion the source systems, data transformation to the data consumption in data platform. In addition, it should identify and document the meta-data, perform data quality and leakage checks and detect data anomalies and trigger corresponding alerts.

Governed: The solution must support fine-grained access control and usage framework for each data objects. It should support configuration of users and groups, classification of data according to sensitivity, protecting data in transit and in rest, etc.

Security and Resiliency: The data platform solution should support resiliency, high availability, data loss prevention, disaster recovery, and backup/restore capabilities.

Target State Data Architecture | Guiding Principles (2/2)

Ability to work seamlessly across on-premise and public cloud environments: The current requirement is to have all components on-premise. But the architecture should enable seamless movement of some / all components to public cloud as desired by LIC.

All types of data: The platform should be able to efficiently consume and store data of all popular formats – including structured, unstructured, semi-structured – logs, audio, video, clickstream, IoT, etc

Fast and real time: The data platform will need to be able to consume and process real time data streams and be able to respond in real time with analytics-based output. The platform should be able to generate reports and dashboards rapidly with minimal / no wait time.

Scalability: The platform is expected to support petabyte scale. It is expected to seamlessly scale both horizontally and vertically and support parallel processing at scale.
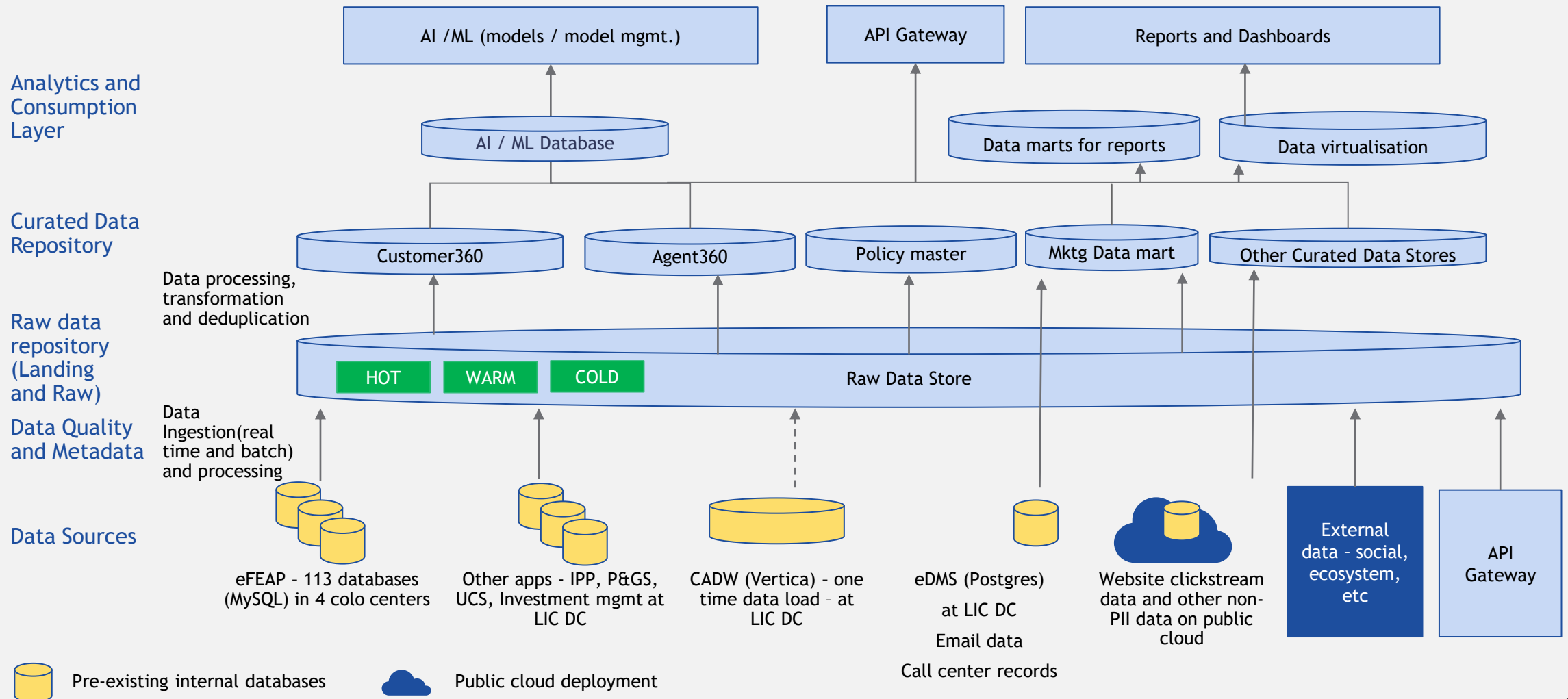
Target State Architecture for Data Lake | Component Level View

Data consumption
- BI/Visualization
- Data and Analytics APIs

Data analytics
- Business intelligence
- Nudge engine
- Recommendation engine
- Anomaly detection
- Risk analytics
- Fraud analytics

Processed data
- Other masters
- Customer 360
- Policy master
- Processed data marts
- Agent 360
- Search (Elastic)

Data cabin
- Data science workbench
- Model training
- Data preparation
- Data exploration
- Data sandbox

Data processing
- Structured data processing
- Unstructured data processing
- Stream processing

Databases
- MPP DB
- RDBMS
- No-SQL
- Cache

Data storage
- Raw data
- Processed data

Data Ingestion
- Batch
- Events/messages
- CDC
- API Gateway
- MFT

Data sources
- eFEAP
- UCS
- P&GS
- ODS
- Investment mgmt
- Website
- Mobile apps
- CADW (one time)
- eDMS
- Others

Data management
- Business glossary
- Metadata dictionary
- Data lineage
- Data quality
- Data stewardship workbench

Data security
- Identify mgmt.
- Data encryption and anonymization
- Data confidentiality tagging

Data ops
- ML orchestration
- Data orchestration
- Model mgmt.
- ETL mgmt.
- Logging/monitoring / auditing

Legend:
- Partially exists
- Does not Exist
- Exists

14

Target State Architecture for Data Lake | High Level View

Disaster Recovery and Service Availability

- The primary site will be Mumbai and the DR site will be Bangalore.
- The DR site should be maintained with High Availability (Active-Passive deployment). A DR automation solution is envisaged and the bidder is expected to procure and implement necessary tools required for enabling the same.
- All components (hardware and software) at the secondary site should be an exact replica of the primary site.

Category	Metric	Expectation
Service Availability (uptime) Uptime = Time since the system was deployed	Data Lake / Lakehouse (including data ingestion, data repositories, data processing, data governance)	99.5% (Monthly) – 3.65 hours Downtime per month
	Master data services	99.99% (Monthly) – 4m 21s Downtime per month
	Reporting and analytics solutions	99% (Monthly) – 7.31 hours Downtime per month
Disaster Recovery	RPO - Recovery Point Objective	Upto 30 mins
	RTO - Recovery Time Objective	Upto 4 hours

Target Architecture | Capabilities and Requirements (1/14)

Application use cases	Indicative list of Features and their description
	Data Lake / Lakehouse

Master data	• Master data repository and source for all key data entities including customer, prospect, policy, agent, Development Officers, etc
	• Sources for preparing the master data will include:
	  – Current core applications:
	    – Core eFEAP platform, IPP
	    – UCS for unit linked products
	    – P&GS for pension and group schemes
	  – Digital and channel applications:
	    – Digital Applications
	    – Mobile app and web app – clickstream data, cookie data, etc
	    – Call center records, emails, etc
	  – Third party data sources:
	    – Social
	    – Third party databases
	• Master tables will need to make available suitable master services (APIs) for consumption by other applications including the front end digital platforms
	• This is not expected to be an operational master data management solution.

Application use cases	Indicative list of Features and their description
	Data Lake / Lakehouse

Customer360	• Complete customer profile data captured from
	   – Customer master data, eDMS – including profile data, KYC data, endorsements, etc
	   – Customer transaction data (including policies bought, riders, premiums paid, payment mode, etc) from core platforms such as eFEAP, UCS, Portal and other digital platforms such as martech systems
	   – Customer interaction data from digital platforms, contact center, emails, customer service and grievance management systems
	   – Relevant third party data sources including social, marketing sources, etc
	• Data elements in the Customer360 will include data of the following nature:
	   – Demographic (eg: unique customer ID, age, gender, address, life stage, education, profession, family details, number of policies, customer since, etc)
	   – Medical and health details (eg: pre-existing medical conditions, output of medical tests, etc)
	   – Holdings details (eg: channel of purchase, policies owned, policy type, date of commencement, policy term, premium payment term, premium amount, premium payment frequency, premium due date, last premium paid, total premium paid so far, riders, sum assured, maturity dates, mode of payment, premium payment missed, etc)
	   – Financial data (eg: income details, PAN, Aadhaar, CIBIL, etc)
	   – Contactability details (eg: outbound contact details, inbound contact details, complaints, emails sent, opened, SMS sent, opened, SRs raised, queries raised, language preferences, interaction history on web, mobile, social, branches, etc)
	   – Digital data (eg: behavioral data on digital assets, cookie data, social data, etc)
	• Make the output available to other systems and platforms using suitable APIs
	• Similar 360 views will need to be created for the lead and the prospect as well in terms of lead360 and prospect360. As the individual progresses from prospect to lead to customer, the respective IDs will need to get mapped to the customer ID with all the relevant data

Application use cases	Indicative list of Features and their description
	Data Lake / Lakehouse

Agent360	• Complete agent profile data captured from:
	   – eFEAP and relevant KYC documents (including policies sold, customers serviced, agent profile, etc)
	 • Data elements in the Agent360 will include data of the following nature:
	   – Agent master data (eg: agent ID, age / date of birth, city, branch associated and type of branch, zone of branch, reporting manager type and ID, date of joining, gender, education, past experience, occupation, number of lives, etc)
	   – Club / MDRT data (eg: current club level, previous club levels, lives data for the past few years, lives in force data for the past few years, renewal commission data for the past few years, first year commission data for the past few years, lapsation rate, total commission, number of years in current tier, etc)
	   – Brigade data for CLIA (eg: NB premium, number of lives, number of agents supervised, active agents, qualified agents, MDRT agents, brigade related history – current tier, years in current tier, etc)
	   – New business data (eg: policy type, premium, customer ID, date bought, start date, expiry date, sum assured, riders, premium term, policy term, payment modes and frequency, etc)
	   – Renewal data (eg: policy type, premium, renewal premium due, payment history, customer ID, date bought, start date, expiry date, sum assured, riders, premium term, policy term, payment modes and frequency, etc)
	   – Customer / Policy portfolio data (eg: policy number, policy type, premium amount, premium mode and frequency, premium payment history, sum assured, riders, maturity amount, survival benefits, etc)
	   – Endorsement / cancellation data (eg: customer ID, endorsements, policies, sum assured, policy start date, policy expiry etc)
	   – Agent training data (eg: trainings completed, last training done, etc)
	   – Agent activity data – activity on digital channels, engagement with customers, DO, CLIA, etc
	 • Agent360 will incorporate various agent types including supervisors, CLIA, SBA, LICA, etc
	 • Make the output available to other systems and platforms using suitable APIs

Application use cases	Indicative list of Features and their description
	Data Lake / Lakehouse
Customer and Family ID	• Ability to uniquely identify customers and related parties (families) • The solution should have the capability to cluster similar records, identify duplicates and create golden record. • This will include using sets of deterministic parameters (such as name, date of birth, address, etc), digital data (cookie data, digital behavior data, etc), third party data (social, etc) to identify customers • Identification of related parties within customers using various parameters such as customer profile data, name, address, transactions, social data, etc and assigning family / household ID. • The solution should provide multiple data quality transformation specific to INDIAN locale which can be leveraged to identify golden record. • The solution should have the ability to identify gender of individuals using the INDIA specific vocabularies. The solution should have the ability to be customizable in terms of the vocabularies, grammars, phonetics, standardization rules, etc. • The solution should have intelligent logic for INDIA names, addresses, phone numbers, national ID, PAN No., Passport number and other identification proof documents and demographic details. • The solution should have intuitive, flexible rules for automatic merging of clustered records • Probabilistic / fuzzy matches / advanced analytics-based matches using combinations of deterministic and non-deterministic parameters will be done in cases where complete sets of deterministic parameters are not available

Application use cases	Indicative list of Features and their description
	Data Lake / Lakehouse
Prospect and Lead ID	• Ability to identify visitors on digital assets (mobile / web app). Visitors who have come onto the digital assets and where minimal data is available (cookie data, limited behavioral data, some identifier such as email ID, name, etc) will be understood as prospects. • Visitors who have expressed interest where some more data is available (such as product interested in, date of birth, mobile number, etc) will be understood as leads. • The system should be able to use data available in these cases and apply suitable rule-based or analytics-based approaches to identify these entities to the best extent possible

Application use cases	Indicative list of Features and their description
	MIS and Advanced Analytics

Customer segmentation and analytics	• Use customer master data, transaction data and interaction data across multiple sources to run segmentation models and create suitable micro segments to be used across journeys. These analytical models will be based on the data from the customer360. Indicative data elements required for the Customer360 are mentioned in the section above.
	• Use cases to be serviced include:
	  – Next best action to predict what would be the best action to be taken / best product to position for a specific customer basis real time events and the customer micro-segment
	  – Predict the probability of a specific customer to buy a specific policy based on propensity to buy models that would use customer segmentation and suitable ML models
	  – Provide personalized input to the agent on the likelihood of a specific customer to pay premium
	  – Provide personalized input to the agent on the likelihood of a specific customer to renew / revive a policy term

Target Architecture | Capabilities and Requirements (7/14)

Application use cases	Indicative list of Features and their description
	MIS and Advanced Analytics
Personalisation	• Identify best campaign / nudge for a customer basis customer360 data and specific events – either life stage related or pre-defined trigger event on digital asset (eg: customer expresses interest on a specific plan, etc) • Develop optimized, customer-specific pricing based on analytical models • Develop optimized offers specific to customers using suitable AI / ML models • Hyper-Personalized recommendation engine of product/plan offers (including ability to recommend cross-sell/upsell offers) • Use AI/ML models to identify right channels to use to engage with a specific customer / customer segment • Use AI / ML models to identify right time slot for communication with a specific customer / customer segment
Churn prediction use cases	• Build suitable AI / ML models to identify customers that are likely to churn • Identify suitable actions / campaigns / nudges and their timing and channel for such customers depending on likelihood of response

Application use cases	Indicative list of Features and their description
	MIS and Advanced Analytics
Other ML use cases	• Agent segmentation: Use agent master data, transaction data and interaction data across multiple sources to run segmentation models and create suitable agent micro segments to be used across journeys. These analytical models will be based on the data from the agent360. Indicative data elements required for the agent360 are mentioned in the section on Agent360 above. • Customer nudges: Develop propensity to pay models to identify likelihood for customers to renew. Identify high and low propensity cases to help drive optimized campaigns. • Lead Scoring and Qualification: Use AI / ML based models to score and prioritize leads for follow up by channel • Agent performance nudges: Identify suitable nudges to drive high sales performance for agents based on performance details and other details as per Agent360 data and Customer360 data (across new business and renewal business) • Supervisor performance nudges: Identify suitable nudges to drive high sales performance for DO/CLIA/SBA/LICA based on performance details and other details as per Agent360 data and Customer360 data. • Upsell opportunities: Identify up-sell opportunities based on customer360 data and specific policy being sold or for existing customers based on policies owned. • Cross sell and upsell opportunities: Identify cross sell / up-sell opportunities based on lead data for new customers • Cross-sell opportunities: Identify cross-sell opportunities based on customer360 data (including elements such as life stage of customer) and specific policy being sold or for existing customers based on policies owned.

Application use cases	Indicative list of Features and their description
	MIS and Advanced Analytics

Other ML use cases	• Identify opportunities to nudge customer at maturity / survival benefit payout
	  – Reinvestment, with suggestion for next best offer
	  – Send nudges to customers to finish required steps (bank account verification, PAN/ Aadhar verification) to receive maturity payout on time
	• Identify nudges customers to be served with nudges for auto debit registration
	• Identify customers to whom nudges could be sent to revive lapsed policy based on propensity to revive
	• Agent eligibility related nudges: Send nudges to agents around eligibility to clubs, loans and other facilities and probability of agents to meet the eligibility.
	• Agent churn prediction: Run models to identify potential agent churn
	• Customer experience related nudges: Send behavior related nudges to agents basis analysis of elements such as customer grievances, service TATs, freelook cancellations, etc
	• Activity and engagement related nudges: Send specific activity related nudges to agents basis number of customer visits done, digital activity, trainings done, etc
	• Fraud detection: Use AI / ML models to use internal and external third party data to identify potential fraudulent claims
	• Sentiment analysis: Use techniques such as NLP and suitable AI / ML models to understand the sentiment of the customer as the customer communicates using contact mechanisms such as Whatsapp, call center, etc and either frame suitable responses or guide the live agent

Application use cases	Indicative list of Features and their description
	MIS and Advanced Analytics
Other ML use cases	• Customer engagement: Use clickstream data to understand behavioral parameters and drive suitable communication or campaigns. • Customer drop-offs: Use suitable advanced analytics models to assess drop offs and trigger suitable action basis customer profile / segment and behavioral parameters. • Medical underwriting: Image analytics to identify health characteristics, smoker / non-smoker, etc • "People like you" analytics to compare people with others in the same segment / cohort • Use "People like you" analytics to identify what other people in the same cohort is buying and use that to drive campaigns / nudges • Personalized nudges basis micro-market driven prospect sourcing methodology to depict high opportunity areas/low sales penetrated areas (e.g., heatmaps to depict region-wise LIC penetration) • Use NLP and suitable AI / ML modelling to categorise customer comments/ incidents into grievances/ service requests/ queries. Categorise incidents into topics (e.g., renewals related, claims related etc.). • Integrated GenAI based tools to provide sales intermediaries with customized multi-lingual (e.g.., Hindi, English, Gujarati, etc.) sales pitch basis product benefits and prospect/customer/policyholder demographics and other profile details (e.g., last purchase, etc.)

Application use cases	Indicative list of Features and their description
	MIS and Advanced Analytics
MIS and Dashboard	**Sales Related:** • Daily activity metrics by LIC (e.g., daily active users, concurrent users, service request raised) • Near Real-time Sales performance dashboards on multiple parameters including distribution channel, product type and customer type and by branch, Division, Zone and CO – with drill down features • Option to view number of policies due for renewal in next 7 days/month **Business Performance Related:** • Summarized view of business productivity metrics (FTD, MTD and YTD metrics for policy premium (e.g., FYP, NOP, etc.) • Personalized 'Tip of the day' targeted to improve business performance (e.g., low renewal ratio compared to agents in similar cohort). • Performance Dashboard: Individual performance reports of the sales intermediary on key metrics like number of policies sold, Current tier of club, Contest Leaderboards, etc.) • Comprehensive Business Dashboards: Feature to see business performance overview as well as options to view detailed views of metrics like FYP, Renewals etc. and trends, qualification for club tiers, competitions/contests • Availability of variety of filters and aggregation options like month-wise records, business-wise records, etc. • Ability to depict visually (e.g., RAG color code) Actual vs Target business achievement for monthly/quarterly/annual performance metrics • 'Top Performer' dashboard in zone/city and benchmark with top performing branches

Target Architecture | Capabilities and Requirements (12/14)

Application use cases	Indicative list of Features and their description
	MIS and Advanced Analytics

MIS and Dashboard	• Ability to visually depict actual vs target achievement using color schemes, option to filter basis use case (e.g. sales intermediaries with less than 50% target achievement)
	• Mapped Agent Cohort Performance Dashboard: Cohort performance reports on key metrics like number of policies sold, total rewards eligible/earned, pay-outs and incentives earned with regards to different business activities and performance parameters
	• Business operational reports including:
	  – New business details and trends across branches / divisions / zones and LIC as a whole
	  – Payments and surrender details and trends
	  – New policies generated and trends
	  – Loans and advances details
	  – Agent performance details
	  – Agent eligibility on clubs and contests
	  – Online business performance
	  – Persistency related reports
	Service Related:
	• Real-time MIS and dashboard for different service requests, customer segments, etc.
	• Detailed reports providing strong insights into areas such as customer service and sales

Application use cases	Indicative list of Features and their description
	MIS and Advanced Analytics

MIS and Dashboard
	• Rewards Related:
	• Ability to track status of rewards, incentives and payouts associated with different business activities and performance parameters.
	• Pay-outs and Incentive dashboard customized basis supervisory role to view segments by channel (e.g. Agency, Bancassurance, etc.), geography, agent cohorts, etc.
	• Ability to track agent wise status of rewards, incentives and pay-outs associated with different business Milestones
	Others:
	• Status for ongoing/mandatory LIC training, Learning and development programs and completion.
	Regulatory:
	• Regulatory reports as required by IRDAI (Further details in Appendix D) including (but not limited to):
	  – New business related reports
	  – Collection related reports
	  – Persistency
	  – Claims related reports
	  – BAP reports
	  – Regulatory return related reports
	  – Commissions, rewards and remuneration
	  – Agency and other channel related reports
	  – Financial statements and trial balance related
	  – Income and collections related
	  – Expenses related and IRDAI reports around finance and accounting

Application use cases	Indicative list of Features and their description
	MIS and Advanced Analytics
MIS and Dashboard	Business Development Related: • Business operations reports (interactive - channel, geo, target audience, campaign...) • Detailed campaign reporting & monitoring – Against digital marketing KPIs (e.g., spends, impressions, clicks, Cost/click, ROAS, etc.) • In-flight optimization dashboard – Optimizations made (bid change, key word addition, audience definition change, creative update...) – Impact on key campaign metrics • Custom reports based on defined KPIs – Self-service: Ability to query the data; for user to filter out the data, Run ad hoc reporting/queries, Manageable by a non-tech, marketing person Total Reports and Dashboards • We expect a total of 3500 reports and 50 dashboards to be developed on the new platform • The complexity definition of the reports and dashboards as below: – Simple Reports / Dashboards are 50% of the total numbers of reports and dashboards respectively. The reports are tabular report, and underlying queries depend on up to 4 fact tables and 2 dimensions. – Moderately Complex Reports / Dashboards are 30% of the total numbers of reports and dashboards. The reports are tabular reports, and underlying queries depend on up to 7 fact tables and 4 dimensions. – Complex Reports / Dashboards are 20% of the total numbers of reports and dashboards. The reports are tabular reports and underlying queries depend on up to more than 7 fact tables with required dimensions. • This will need to be accompanied with strong self-service ability where end users can customize their own views and cuts of data.

Walkthrough of the Technical Evaluation Criteria
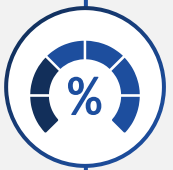
RFP | Techno-Commercial Evaluation Overview

Selection method will be Quality Cost Based Selection (QCBS)

Relative weightages for technical and commercial score will be 70:30

Technical Bids will be evaluated and assigned a technical score out of a maximum of 100 marks.

Only the bidders with an aggregate technical score of 70% or more will qualify for the commercial bid evaluation

Technical Evaluation Criteria

1 **Understanding of Life Insurance Business and LIC context**
- Life insurance business understanding and key challenges faced by customers, agents today from a data / reporting / analytics perspective

5 Marks

2 **Bidder's Experience**
- Bidder should submit two case studies showing in detail the scope of work, architecture implemented, size & complexity

20 Marks

3 **Quality of Proposed Technical Solution Architecture**
- Data Ingestion
- Data Storage
- Data Processing
- Data Quality and Governance
- Data security & Access control
- Data consumption
- Monitoring
- Proposed partnerships and OEMs

30 Marks

4 **Implementation Approach**
- Overall detailed project plan
- DevSecOps and MLOps
- Tech Documentation
- User Training
- Platform Run Operations

15 Marks

5 **Quality of Team**
- Bidder should deploy best of class professionals to ensure successful execution of the project – evaluation based on CV & interviews

20 Marks

6 **References**
- Bidder shall provide 2 references that will be verified by LIC via visit/call.

5 Marks

Walkthrough of the Commercial Evaluation Criteria

Commercial Bid Structure | Summary

A Total Cost of Ownership – 5 years

#	Description of cost items	Cost	5-year TCO
1	One Time Implementation[1]		
2	Hardware[2]		
3	Software[3]		
4	ATS / AMC /AMS[4]		
5	Change Request Fees[5] (to be utilized on a need basis)		

1. One-time Implementation cost to be quoted separately for each of the 3 modules – Data Lake, Advanced Analytics, Reporting and Dashboarding – includes design, implementation, testing, deployment into production and roll-out support for the solution
2. Hardware – the vendor to procure and install the hardware on behalf of LIC; LIC to provide the physical space for hosting the hardware. Vendor to submit a detailed Bill of Materials Table specifying the hardware needed for DC, DR
3. OEM Software Licenses – the vendor to procure and install the OEM software on behalf of LIC. Vendor to submit a detailed Bill of Materials Table specifying the nature of quantity and rate of the licenses needed
4. Annual maintenance and support services cost to be specified as a lumpsum for all three modules as well as the annual maintenance of hardware and OEM software licenses required to enable the same.
5. Vendor is required to submit a detailed rate card as shown in the next page. The cost of will be calculated at a Blended Rate as per Rate card.

Commercial Bid Structure | One Time Implementation

One-off Implementation cost to be quoted separately for each of the 3 modules –

- Data Lake / Lakehouse
- Advanced Analytics
- Reporting and Dashboarding

It includes design, implementation, testing, deployment into production and roll-out support for the solution

1 Table#1: Bill of Material (One Time Implementation)

Sr. No.	Application	One Time Implementation *(in INR lakhs)*
1	Data Lake / Lakehouse	
2	Advanced Analytics	
3	Reporting and Dashboarding	
	Total	

Commercial Bid Structure | Hardware

The bidder is expected to size and install adequate hardware for the scope mentioned in the RFP in the following manner -

- Vile Parle (Primary DC) – Dev, SIT, UAT, Prod, any other environment
- Bangalore (Far DR) – Prod
- An interim development environment on Public Cloud/on-premise as considered fit by the bidder until the hardware is commissioned. This development environment will be decommissioned and migrated to the Primary DC once the hardware is installed and ready

2 Table#2: Bill of Material (Hardware)

Sr.No.	HW (Hardware)	Description	Primary DC / DR	Environment (Interim Dev, Dev, SIT, UAT, Prod, Any Other)	License Type Perpetual / Subscription	Qty	Rate Per Unit	5-year TCO (in INR lakhs)
						A	B	C = A*B
1								
2								
...								
							Total	<transfer to TCO table>

Quoted 5-year TCO is to be discounted to the net present value (NPV) (Discounting rate to be used: 10%)

Commercial Bid Structure | Software

The bidder is expected to size and install adequate OEM Software for the scope mentioned in the RFP in the following manner -

- Vile Parle (Primary DC) – Dev, SIT, UAT, Prod, any other environment
- Bangalore (Far DR) – Prod
- An interim development environment on Public Cloud/on-premise as considered fit by the bidder until the hardware is commissioned. This development environment will be decommissioned and migrated to the Primary DC once the hardware is installed and ready

③ Table#3: Bill of Material (Software)

Sr.No.	HW (Software)	Description	Primary DC / DR	Environment (Interim Dev, Dev, SIT, UAT, Prod, Any Other)	License Type Perpetual / Subscription	Qty	Rate Per Unit	5-year TCO (in INR lakhs)
						A	B	C = A*B
1								
2								
...								
						Total		<transfer to TCO table>

Quoted 5-year TCO is to be discounted to the net present value (NPV) (Discounting rate to be used: 10%)

Commercial Bid Structure | AMC/ATS/AMS

- Bidder should provide details of the annual maintenance and support services for own developed custom applications, third-party (OEM) software and hardware in a segregated manner as shown below.

- Includes application maintenance of the data lake/lakehouse, data pipelines, reporting and advanced analytics solutions as well as the annual maintenance of hardware and OEM software licenses required to enable the same.

- During the AMC period, the bidder is liable to provide onsite support resources to LIC.

4 Table#4: Bill of Material (AMC)

S. No.	Nature of support	Y1 (in INR lakhs)	Y2 (in INR lakhs)	Y3 (in INR lakhs)	Y4 (in INR lakhs)	Y5 (in INR lakhs)	TCO for 5 years (in INR lakhs)
1	ATS (Hardware)	x					
2	AMS (custom developed applications)	x	x				
3	AMC (OEM software)	x					
						Total	\<transfer to TCO table\>

Commercial Bid Structure | Change Request Fees

5 Table#5: Change request costs: Team deployed for 12 months

Sr. No	Role	Estimated weightage for a resource type for a CR	Person Month Price *(in INR lakhs)*	12 Months Price *(in INR lakhs)*	One Time Cost/ Year 1 Cost *(in INR lakhs)*
		A	B	C = (A * B)*12	
1	Project Manager	1			
2	Insurance data expert	0.25			
3	Dev/Pod lead	1			
4	Sr. Data Engineer	2			
5	Data Engineer	2			
6	Data security and privacy expert	0.25			
7	Cloud and Infra Architect	0.25			
8	Principal Data Scientist	0.5			
9	Data Scientist	2			
10	Reporting and Visualization Experts	5			
11	Devops Engineer	0.5			
12	Test / QA Engineers	4			
	Total				<transfer to TCO table>

Commercial Bid Structure | Summary (2/2)

B Rate Card

Professional Figure	Price/Month *(in INR lakhs)*
Project Director	
Project manager	
Team Lead	
Insurance Data Expert	
Data Architect	
Infra Architect	
Sr. Big Data Engineers	
Big Data Engineers	
Data Security and Data Privacy Experts	
Data Scientists	
Principal Data Scientist	
Reporting and Visualization Experts	
MLOps engineer	
Test Architect and Lead	
Test Engineers	
Release manager	

Bidder should provide a monthly rate card for the mentioned resources that will be utilized when calculating change requests costs if the actual cost exceeds the total estimated cost mentioned in Section A

Walkthrough of the Payment Terms and Timelines

Payment terms | 1. Implementation (1/2)

1 Payment Terms : Implementation

T_0: From the date of issuance of Letter of Intent (LOI)

Sr.No	Deliverables	Milestone	Payment terms as per Commercial Bid (T-6)
1	Architecture, High level Design (HLD) signed off for data lake / lakehouse	T_0 + 2 month	1%
2	Initial setup and installation of key data platform solution components in the interim dev environment	T_0 + 2 months	2%
3.	Setup and installation of key data platform solution components in the actual dev environment	T_0 + 4 months	2%
4.	Wave 1: Go-Live - Full launch of first set of data and analytics services for the digital program No P1 (Critical) and P2 (High) bugs open. Data completeness – 100% Customer data uniqueness false positives– 98% Model Accuracy greater than 75%. F1 Score greater than 0.7 Report accuracy – 100%	T_0 + 9 months	30%

Payment terms | 1. Implementation (2/2)

T_0: From the date of issuance of Letter of Intent (LOI)

Sr.No	Deliverables	Milestone	Payment terms as per Commercial Bid (T-6)
5.	Wave 2: Go-Live - Full launch of second set of data and analytics services for the digital program No P1 (Critical) and P2 (High) bugs open. Model Accuracy greater than 75%. F1 Score greater than 0.7 Report accuracy – 100%	T_0 + 12 months	20%
6.	Wave 3: Go-Live - Full launch of third set of data and analytics services for the digital program No P1 (Critical) and P2 (High) bugs open. Model Accuracy greater than 75%. F1 Score greater than 0.7 Report accuracy – 100%	T_0 + 15 months	15%
7.	Wave 4: Go-Live - Satisfactory Delivery of all features as per the scope of RFP No P1 (Critical) and P2 (High) bugs open. Model Accuracy greater than 75%. F1 Score greater than 0.7 Report accuracy – 100%	T_0 + 18 months	30%

Note: Prioritization of features/ functionalities going live in each wave may be modified during the contract based on LIC's business requirements/ exigencies and as mutually agreed by LIC and the successful bidder

Deliverable and payment plan | 2. Hardware

T_0: From the date of issuance of Letter of Intent (LOI)

Sr. No.	Deliverables	Timeline	Payment (%)
Hardware Deployment at LIC			
1	Hardware delivery and installation for Non-Prod (Dev, UAT) at LIC DC (Vile Parle) or LIC approved co-lo	T_0 + 4 months	25%
2	Env set-up on new Hardware for Non-prod (Dev, UAT) LIC data center	T_0 + 5 months	20%
3	Hardware delivery and installation for Prod at LIC DC (Vile Parle), DR (Bangalore) or LIC approved co-lo	T_0 + 4 months	25%
4	Production environment set up and signed off by LIC at LIC DC (Vile Parle), DR (Bangalore) or LIC approved co-lo	T_0 + 5 months	20%
5	Final Payment	T_0 + 12 months	10%
Total			**100%**
Additional hardware (if required) for subsequent years			
6	Delivery & installation of the Hardware / Appliances and their satisfactory deployment on all applicable environments (Dev, SIT, UAT, Prod, any other environment) duly signed off by LIC.	As Applicable	100%

Deliverable and payment plan | 3. Software

Sr. No.	Deliverables	Timeline	Payment (%)
Software licenses (perpetual and subscription) for first year			
1	Delivery of Software licenses. The required documents to be provided are original invoice along with Original Delivery Challans dully stamped and signed by the LIC Official & Selected Bidder representative.	T_0 + 1 months	50%
2	Interim payment	T_0 + 3 months	25%
3	Interim payment	T_0 + 6 months	20%
4	Final payment	T_0 + 12 months	5%
	Total		100%
Software licenses (perpetual and subscription) for subsequent 4 years			
5	**Delivery of Software Licenses and their installation on all applicable environments.** The required documents to be provided are original invoice along with Original Delivery Challans dully stamped and signed by the LIC Official & vendor representative. LIC official to sign off that new licenses have been satisfactorily installed.	T_0 + 22 months T_0 + 34 months T_0 + 46 months T_0 + 58 months	100% 100% 100% 100%

Deliverable and payment plan | 4. Annual Maintenance and Support Costs (1/2)

Sr. No	Deliverables	Milestone	Payment terms as per Commercial Bid (T-6)
AMS (custom developed applications)			
1	Year 3	T_0 + 24 months	25%
		T_0 + 27 months	25%
		T_0 + 30 months	25%
		T_0 + 33 months	25%
		Total	100%
2	Year 4	T_0 + 36 months	25%
		T_0 + 39 months	25%
		T_0 + 42 months	25%
		T_0 + 45 months	25%
		Total	100%
3	Year 5	T_0 + 48 months	25%
		T_0 + 51 months	25%
		T_0 + 54 months	25%
		T_0 + 57 months	25%
		Total	100%

Deliverable and payment plan | 4. Annual Maintenance and Support Costs (2/2)

T$_0$: From the date of issuance of Letter of Intent (LOI)

Sr. No	Deliverables	Milestone	Payment terms as per Commercial Bid (T-6)
ATS (Hardware) and AMC (OEM software)			
1	Year 2	T$_0$ + 13 months	100%
2	Year 3	T$_0$ + 25 months	100%
3	Year 4	T$_0$ + 37 months	100%
4	Year 5	T$_0$ + 49 months	100%

भारतीय जीवन बीमा निगम
LIFE INSURANCE CORPORATION OF INDIA