

### **Scope for the POC forming part of Technical evaluation of RFP for IDP– Scope A**

- 1) Demonstrate using your Demo application the following features and the solution capabilities using a set of documents provided by SBI Life.
- 2) Document Ingestion
  - a. to accept two documents as input
  - b. to accept a compound document (PDF with more than one document contained in same or different pages)
- 3) Image Pre-processing & classification of the following documents

<b>Document Type</b>
SBI Life Proposal Form
SBI Life Claim Form
PAN
Voter ID
Driving License
Passport
Aadhaar
Bank Statement/passbook
Cancelled Cheque
ITR
Death Certificate
If any document other than the above then classify as “Other Documents”



- 4) Compound Document handling –
  - a. List the documents found in a single input file
  - b. Create a new pdf file as output with a distinct file name for each of the above documents as found in the single input file
  - c. If more than one document is found on the same page, then the relevant document has to be cropped and given as a new pdf file
- 5) Extraction, Validation, Capture and tagging of data from documents

Document Type	Data Points to Extract
Proposal Form	Name of Life Assured Proposal Number
SBI Life Claim Form	Policy Number Name of the Deceased Name of Claimant Relationship to Deceased Date of Death Cause of Death
PAN	PAN No Name as per PAN
Voter ID	Voter ID Number Name Father's Name DOB or Year of Birth (calculated as age less as on date) Address

Driving License	DL Number Name Father's Name Address
Passport	Passport Number Name Father's Name Address Expiry date
Aadhaar	Aadhaar Number Extract following from the QR <ul style="list-style-type: none"> <li>○ Last 4 digits of Aadhaar</li> <li>○ Name</li> <li>○ Father's Name</li> <li>○ Address</li> <li>○ Mobile Number</li> </ul>
Bank Statement or Cheque	Account Number IFSC Account holder Name
Death Certificate	Name of Deceased Date of Death Data contained in QR code

6) Identify human face and match human faces appearing in documents read with photo in Proposal form

7) Present output of (5) and (6) as under

DOC No	Document Type	Labelled Data Extracted with confidence score for each field	Face Extracted & Match Score	Name Match score
1	<<doc type>>	<<data extracted as mentioned at 5>>  Present as <<field tag>>-<<value>>-<<confidence score>>	Photo :   Face Match Score (vs Doc 2)- <<score>>	Name Match Score (Vs Doc 2)- <<score>>
2	<<doc type>>	<<data extracted as mentioned at 5>>  Present as <<field tag>>-<<value>>-<<confidence score>>	Photo :   Face Match Score (vs Doc 1)- <<score>>	Name Match Score (Vs Doc 1)- <<score>>

<b>Parameters to Showcase through POC</b>	
1	Accuracy of classification of medical & KYC documents
2	Accuracy of data extracted – Medical & KYC - Printed Text - Handwritten Text
3	Accuracy of identification & tagging of named entities
4	Creation of summary sheet & grouping of tests as per SBIL grouping
5	Human Face detection & extraction
6	Facial & Name match accuracy