Sl. No.	Page No.	Para No. (Sub para, if any)	Original text	query raised	Reply
1	12	Annexure B	Automated/straight-through periodic processing of documents (PDF/scanned or system generated images of all formats) from a predefined location should be possible.	Please share the list of document types required to be processed	Bank statements,KYC documents, account opening/loan application forms etc. in either PDF, or images format
2	12	Annexure B	Automated/straight-through periodic processing of documents (PDF/scanned or system generated images of all formats) from a predefined location should be possible.	Please share sample documents for each type of document	Bank statements,KYC documents, account opening/loan application forms etc. formats as prevalent right now to be done. These are generaly available with Bank customers.
3	12	Annexure B	The application should also have provisions for manual verification of the extracted data in case of any rule-based errors, insufficient accuracy or any other requirements. Support for NLP and Machine Learning is an added advantage	Since support for NLP and Machine Learning is an added advantage, please indicate the use cases / document types for which this capability will be required.	Bank statements,KYC documents, account opening/loan application forms etc. in either PDF, or images format
4	3	Background	Current volume of data and projected volume of Data	Please share daily document volumes, peak document volumes and time available to process peak and daily volumes	Average volume estimated is 55000 per day Peak volume estimated is 300000 pages per day. Time available for processing will be 24 hours. It is expected that any document uploaded /provided for conversion should be done in 12 hours maximum.
5	12	1	Scanned/system generated images/PDF file	What are different types of input file formats apart from PDF, Scanned images?. Is there any hand written documents also needs to be covered in this scope of work?	Different type of input files format may be PDF/all image type etc. Hand writen documents(forms filled in capital english letters if available will be an added advantage)

Sl. No.	Page No.	Para No. (Sub para, if any)	Original text	query raised	Reply
6	12	1		What will be the different input sources of the files?. Is it from the shared folder, as an attachment in the email, FTP site, any information system etc?	The input source will be different input folders for different projects and tool is expected to run at predefined time interval to fetch the doucments availabe in input folder , convert it and then transfer it to designated output folder with same name convention/pre defined rule as used for input folder.
7	12	1		The input documents are in same standard structure/format or unstructured documents?	It may be of same standard structure/format or unstructured documents?
8	12	1.1	Automated/straight-through periodic processing of documents	Do you expect fully automated OCR tool without manual intervention?	Yes.
9	12	1.2	The application should be able to automatically identify documents	What do you mean by this statement? Does it means that application should be able to auto identify whether it's a PAN card, driving licence, invoice etc?. Need more clarification on this	Document definition input facility must be in tool. After proper document definition each type of document will be kept in input folder with respective definition.
10				Kindly share the document types that needs to be extracted	Bank statements,KYC documents, account opening/loan application forms etc. formats as prevalent right now to be done. These are generaly available with Bank customers.
11	12	1.2	There should also be ability to define document structures	Please provide more clarification on this	Document definition input facility must be in tool (for each type of document)

Sl. No.	Page No.	Para No. (Sub para, if any)	Original text	query raised	Reply
12	12	1.4	Application should have the ability to save the extracted data in multiple standard file formats (such as excel/word etc.)	Please provide all the formats that the application should save the extracted data (apart from excel/word).	The application should save the extracted data in excel/word.
13	12	1.4	There should be an archival and classification mechanism to store the extracted data along with the electronic copies of the documents, in a database or other information systems	In how many systems data needs to be stored? Whether they are web/desktop based? Which database (SQL, Oracle etc.) do you expect the application to store the extracted data?	Data needs to be stored in Bank's serverin Oracle Database.
14				Please let us know the integration touch points for this project	No integration with any other application as such needed
15				How many users from Bank will be accessing the extraction platform	Appx 5000 concurret user at a time will acess the input folder on the server through front end application.
16				From how many branches / locations users will login	The input source will be different input folders for each project and tool is expected to run at predefined time interval to fetch the doucments availabe in input folder and to convert it and then transfer it to designated output folder

Sl. No.	Page No.	Para No. (Sub para, if any)	Original text	query raised	Reply
17	12	1	The document imaging and data extraction software is required for automation of extraction of desired data from scanned/ system generated images/PDF files.	In scanned images, required data is printed or handwritten and which language used ?	Required data is expected to be printed primarily in english language. Writen documents(forms filled in capital english letters) if available will be and added advantage.