NBO –GAR/GAM   
Text Analytics– Lot 1 [ENGLISH]

|  |
| --- |
|  |
| **Requirement Specification Document** |

|  |  |  |
| --- | --- | --- |
|  |  |  |

Table of Contents

[1. Document Purpose 3](#_Toc477879007)

[2. Activity Objective and Scope 3](#_Toc477879008)

[3. Text Analytics tool access 4](#_Toc477879009)

[4. Functional Design Requirements 5](#_Toc477879010)

[4.1. Use Case: Extracting Key Data Elements from Prospectus 5](#_Toc477879011)

[4.2. Criteria of sourcing the KDEs [To be concluded] 7](#_Toc477879012)

[4.3. Business Rules of sourcing the KDEs 10](#_Toc477879013)

[5. Exception Handling 14](#_Toc477879014)

[6. Acceptance Criteria 15](#_Toc477879015)

[7. Business Process Continuity Steps 16](#_Toc477879016)

[8. Appendix and Attachments 17](#_Toc477879017)

[9. Definitions 17](#_Toc477879018)

[10. Document Control Sheet 17](#_Toc477879019)

[11. Appendix 18](#_Toc477879020)

[11.1 Appendix 1: Samples of prospectus 18](#_Toc477879021)

[11.2 Appendix 2: KPI –Audit report [To be concluded] 18](#_Toc477879022)

[11.3 Appendix 3: Attribute analysis 18](#_Toc477879023)

# Document Purpose

The purpose of this document delineates the functional and technical design requirements in capturing the 11 critical KDEs from prospectus using Text Analytics technology. The document outlines the activity scope, high level design, design considerations and requirements.

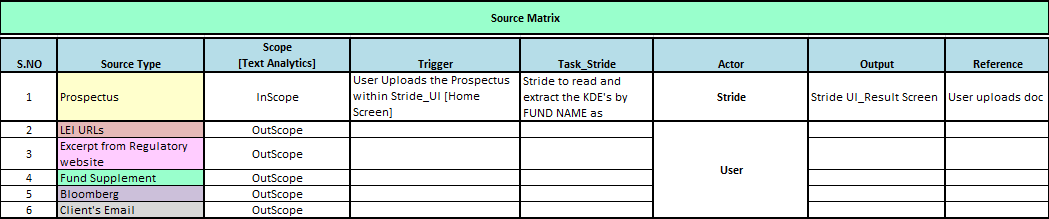
# Activity Objective and Scope

The scope is to extract the Key Data Elements [KDEs] from the prospectus through Text Analytics technology solution aiming to gain efficacy in obtaining critical KDEs of GAR/GAM Onboarding process. Further user can export the results into excel and strike the derived values into relevant business applications.

The scope of the KDE’s are represented in the below table.  
***Table 2.1:*** *Attributes required*

|  |  |
| --- | --- |
| S.No | Attribute Name |
| 1 | Asset Manager |
| 2 | Custodian |
| 3 | Delegated Asset Manager |
| 4 | Registered Address |
| 5 | Country of Domicile |
| 6 | ISIN |
| 7 | LEI |
| 8 | FBL Exemption |
| 9 | Leverage Ratio / Commitment Approach / Value at Risk Approach |
| 10 | Hedge Fund |

The scope of sources are as represented in the below table.  
***Table 2.2:*** *Attributes required*



The below are the list of languages in scope.

***Table 2.3: Languages required***

* English
* German
* Spanish
* French
* Italian

The whole project has been broken down into 5-Lots to cover the prospectus covering all the above mentioned 5-distinct languages.

Requirements are standard across the languages along with the file extension and file type [pdf and prospectus respectively]  
  
***Table 2.4:*** *Lot planning against Modules*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Prospectus Lingua Module** | **Lot 1** | **Lot 2** | **Lot 3** | **Lot 4** | **Lot 5** |
| English Extraction Module |  |  |  |  |  |
| German Extraction Module |  |  |  |  |  |
| Spanish Extraction Module |  |  |  |  |  |
| French Extraction Module |  |  |  |  |  |
| Italian Extraction Module |  |  |  |  |  |

# Text Analytics tool access

***Table 3:*** *Application/Access required*

|  |  |
| --- | --- |
| S.No | Application /Access required |
|  | Internet Explorer |
|  | Google Chrome |

# Functional Design Requirements

This section lists out the various tasks that are to be automated. It includes a name and description of each task and specifies the functional requirements for each task.

## Use Case: Extracting Key Data Elements from Prospectus

Detailed process steps have been outlined below:

*Table 4.1:* *Extracting Key Data Elements from Prospectus*

|  |  |  |
| --- | --- | --- |
| **Step** | **Process** | **Comments/Design considerations** |
| 4.1.1 | User launches Text Analytics tool by using URL shared by Stride | User should be provided with a URL to launch text analytics tool |
| 4.1.2 | User logs in Text Analytics tool by using credentials | 1. User should be provided with unique credentials to access text analytics tool    1. User name   Username will be in the following syntax;  <Firstname>.<lastname>  e.g. pavan.kaki  If more matches with same name [may be rare scenario] found, use below syntax  <Firstname1>.<lastname>  e.g. pavan1.kaki   * 1. Password  1. Should have a ‘Submit’ button to enter into Stride UI |
| 4.1.3 | User uploads the document | 1. User clicks on Browse button within Stride UI 2. User chooses the file from system and uploads the selected file/prospectus to extract KDEs 3. User inputs the Entity Name for which information needs to be sourced [Refer **BR007 from** *Table 4.3:* *Extracting Key Data Elements from Prospectus* |
| 4.1.4 | User clicks on Submit button | The moment user clicks on Submit button, text analytics tool read/process the document uploaded and display a message whether the operation is successful/failed with a message through a pop up window |
| 4.1.5 | Text Analytics Results UI | 1. On successful processing, results should sync with the scope of KDE’s asked   Please refer below mentioned table from  ***[Table 2.1:*** *Attributes required]*  ***Please Note:*** *Attributes with synonyms/keywords are provided in Appendix 3 to ensure completeness and accuracy in extracting even when the same attribute is called by different string name [or synonyms/key words]*   1. UI should display in two frames,    1. Left Frame to display results       1. Values obtained should be represented in a tabular format against the attribute name       2. All values extracted from prospectus/doc should be hyperlinked    2. Right Frame to locate the text in Yellow colour highlight from where the resulting value has been derived. [on a click of value in hyperlink]   This helps user to make sure the results are good to go |
| 4.1.6 | Download Excel | User should be allowed to download standardize input extracted resultsinto MS Excel. Further to leverage data into federated systems |
| 4.1.7 | KPI-Log Files [Audit Report] | User should be able to extract the KPI-Log files as and when needed to view the KDEs sourced against the number of documents/instances run by user along with the required details.  This report is required to store the data in stride database to accommodate user to call the data following periodicity 1.Daily 2.Weekly 3. Monthly [TBD]  This is to record the actions with results against the no.o f prospectus ran by user and results generated by Stride and further used for data quality.  Please refer **Appendix 2** |

## Criteria of sourcing the KDEs [To be concluded]

*Table 4.2:* *Extracting Key Data Elements from Prospectus*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **S.No** | **Attribute Name** | **Field Type** | **Synonyms/Key Words** | **Source Priority#1** | **Data Type** | **Business Rule** |
| 1 | Asset Manager | **Mandatory** | Asset Manager | Prospectus | Char |  |
| Management Company | Prospectus |  |
| Investment Manager | Prospectus |  |
| Investment Advisor | Prospectus |  |
| Asset Management | Prospectus |  |
| ACD - Authorized Corporate Director | Prospectus |  |
| 2 | Custodian | **Mandatory** | Custodian | Prospectus | Char |  |
| Depository | Prospectus |  |
| Administrator (for few countries) | Prospectus |  |
| Trustee | Prospectus |  |
| 3 | Delegated Asset Manager | **Mandatory** | Delegated Asset Manager | Prospectus | Char | **BR006(a)** |
| Sub Investment Manager | Prospectus | **BR006(a)** |
| Sub Investment Advisor | Prospectus | **BR006(a)** |
| Sub-Advisor | Prospectus | **BR006(a)** |
| Delegate Asset Manager | Prospectus | **BR006(a)** |
| Delegate Asset Management | Prospectus | **BR006(a)** |
| Delegated Asset Management | Prospectus | **BR006(a)** |
| must be read in conjunction with the Supplement | Prospectus | **BR003 (a)** |
| Supplement to the Prospectus | Prospectus | **BR003 (a)** |
| Supplement for | Prospectus | **BR003 (a)** |
| 4 | Registered Address [Fund] | **Mandatory** | Registered Address [Fund] | Prospectus | Char & Alphanumeric | **BR004** |
| Head Office [Fund] | Prospectus |
| Registered Office [Fund] | Prospectus |
| 5 | Country of Domicile | **Mandatory** | Country of Domicile | Prospectus | Char |  |
| Country of Residence | Prospectus |  |
| Place of Registration | Prospectus |  |
| Place of incorporation | Prospectus |  |
| incorporated | Prospectus |  |
| incorporated in | Prospectus |  |
| incorporated with | Prospectus |  |
| resident in |  |  |
| registered in | Prospectus |  |
| under the laws of | Prospectus |  |
| incorporated under | Prospectus |  |
| 6 | ISIN | **Mandatory** | ISIN | Prospectus | Alphanumeric | **BR006(b)** |
| 7 | LEI | **Mandatory** | Legal Entity Identifier | Prospectus | Alphanumeric |  |
| LEI | Prospectus | Alphanumeric |  |
| 8 | FBL Exemption | **Mandatory** | Pension Funds | Prospectus | Char | **BR005** |
| SICAV's | Prospectus | **BR005** |
| Private Equity | Prospectus | **BR005** |
| UCITS | Prospectus | **BR005** |
| 9 | Leverage Ratio &Commitment Approach / Value at Risk Approach | **Mandatory** | Commitment Approach / Value at Risk Approach | Prospectus | Char & Alphanumeric | **BR008** |
| Leverage Ratio | Prospectus |
| NAV | Prospectus |
| Commitment Approach | Prospectus |
| Value at Risk Approach | Prospectus |
| Relative Value-at-Risk | Prospectus |
| Value at Risk (VaR) approach | Prospectus |
| Absolute VaR | Prospectus |
| Var | Prospectus |
| Expected Return | Prospectus |
| Global Exposure | Prospectus |
| Commitment Method | Prospectus |
| Var approach | Prospectus |
| Commitment Leverage | Prospectus |
| Commitment | Prospectus |
| Var approach | Prospectus |
| 10 | Hedge Fund | **Mandatory** | Hedge Funds | Prospectus | Boolean [Yes/No] | **BR009** |
| Minimum initial investment | Prospectus |
| Eligible/qualified Investors | Prospectus |
| Valuation frequency -  Not daily | Prospectus |
| Redemption frequency - Not within the month | Prospectus |
| Redemption gate | Prospectus |
| Strategies - Hedge fund, Short selling, speculative strategy | Prospectus |
| Net asset value calculation - within a month | Prospectus |
| Fund of funds | Prospectus |
| investment in/of hedge fund | Prospectus |
| Master fund | Prospectus |
| Feeder fund | Prospectus |
| Performance based fees | Prospectus |
| created for specific purposes | Prospectus |
| invested/investment in vehicles | Prospectus |
| close end/ended fund | Prospectus |
| Special Investment funds | Prospectus |
| société d'investissement en capital à risque | Prospectus |
| SPV - special purpose vehicle | Prospectus |

## Business Rules of sourcing the KDEs

*Table 4.3: Rules for* *Extracting Key Data Elements from Prospectus*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Business Rules** | | | | | |
|
|  |  |  |  |  |  |
| **Code** | **Case** | **Attribute Name** | **Exception** | **Description** | **Output** |
| BR001 | When Search Result = 0 | 1. Asset Manager 2. Custodian 5. Country Of Domicile 6. ISIN 7. LEI 9. Leverage Ratio / Commitment Approach / Value at Risk Approach | 3. DAM [Delegated Asset Manager] 4. Registered Address 8. FBL Exemption  10. Hedge Fund | In case if Stride returns **no value** from prospectus for all KDEs except the following. 1. DAM [Delegated Asset Manager] 2. FBL Exemption 3. Registered Address  10. Hedge Fund | "NA" [Not Available] |
| BR002 | When Search Result = 1 | All 10 KDE's | - | When Stride returns only ONE VALUE for queried KDE | Update its <VALUE> as it is sourced from prospectus |
| BR003 | When Search Result = 1  OR  When Search Result = 0 | 3. Delegated Asset Manager | - | In case if Stride returns no value/Only one value/more than one value from prospectus for **DAM [Delegated Asset Manager]** | should never be populated as "NA" [Not Available] |
| BR003(a) | When Search Result = 1 | - | Check synonyms/keywords within prospectus. If found- | "Refer Fund Supplement" |
| BR003(b) | When Search Result = 0 | - | Check synonyms/keywords within prospectus. If not found- "Asset manager" value should be populated in results | "Asset Manager" |
| BR004 | When Search Result = 0 | 4. Registered Address [Fund] | - | If "Registered address" is "Not Available" even with synonym/key word within Prospectus, populate the address of "Asset Manager" | "Refer BDR Address\_AM" |
| BR005 | When Search Result = 1  OR  When Search Result = 0 | 8. FBL Exemption | - | In case if Stride returns no value even after following order of priority for **FBL Exemption** | "FBL Exemption Value" is derived based on the below rules |
| BR005(a) | When Search Result = 1 | - | If Leverage Ratio >300%, FBL Exemption = | N |
| BR005(b) | When Search Result = 1 | - | If Leverage Ratio <300%, FBL Exemption = | Y |
| BR005(c) | When Search Result = 0 | - | If Leverage Ratio is "NOT AVAILABLE", FBL Exemption = | "NA" [Not Available] |
| BR006 | When Search Result >1 | 3. Delegated Asset Manager 6. ISIN | - | When more than ONE Delegated Asset Manger / ISIN found in a prospectus, please capture all the available information from prospectus | Capture ALL |
| BR006 (a) | When Search Result >1 | 3. Delegated Asset Manager | - | When more than ONE Delegated Asset Manger is found in a prospectus, please capture all the DAM names available in prospectus |
| BR006 (b) | When Search Result >1 | 6. ISIN | - | When more than ONE ISIN is found in a prospectus, please capture all the ISINs available in prospectus |
| BR007 | When Search Result = 0 OR When Search Result = 1 OR When Search Result >1 | Fund Name [User Input in Stride UI] | - | KDE's should be extracted basis on the **FUND NAME provided by the USER** | Fund Name as provided by the USER |
| BR008 | When Search Result = 0  OR  When Search Result = 1 | 9.Leverage Ratio &Commitment Approach / Value at Risk Approach | - | Handling Leverage Ratio based on the search result outcome | Should update as mentioned below |
| BR008(a) | When Search Result = 0 | - | When Stride returns NO VALUE for queried KDE | "NA" [Not Available] |
| BR008(b) | When Search Result = 1 | - | When Stride returns only ONE VALUE for queried KDE | "Numeric Value" and "Text" |
| BR009 | When Search Result = 0  OR  When Search Result = 1  OR  When Search Result > 1 | 10. Hedge Fund | - | Handling **Hedge Fund** on the search result outcome | Should update as mentioned below |
| BR008(a) | When Search Result = 0 | - | When Stride returns NO VALUE for queried KDE with synonyms/Keywords | "NO" |
| BR008(b) | When Search Result = 1  OR  When Search Result > 1 | - | When Stride returns only ONE/MORETHAN ONE VALUE for queried KDE with synonyms/Keywords | "YES" |

# Exception Handling

***Table 5:*** *Exception handling methodology*

|  |  |
| --- | --- |
| **Exception Type** | **Exception handling** |
| * 1. KDEs extracted when details are missing within prospectus | Update Value obtained as ‘Not Available’ when value is missing within the document uploaded by user  ***Refer Business Rules 4.3*** |
| * 1. Error in uploading document | Populate with right string in message pop-up window  **User Controls:**   1. “**Incorrect File**” in case user uploads irrelevant file (may be not a prospectus!) 2. **“Incorrect File format”** may be not in pdf which text analytics tool doesn’t support 3. **“Scanned/Image prospectus” are** not supported by STRIDE.   *when user uplods scanned pdf /pdf containing images.* an *information message* will be displayed on UI instructing user to upload [**Read-only**] prospectus  This helps user to re-process correct file and format to avoid inconsistencies. |
| * 1. Error in downloading excel document from results page of text analytics | Text analytics tool is expected to support ‘*export to excel’* functionality irrelevant KDE’s are Available or Not.  In case of any issues associated to disruption of server/internet/others, should be populated with right string in message pop-up window for user’s understanding |

# Acceptance Criteria

***Table 6:*** *Acceptance criteria recommended by Business*

|  |  |  |
| --- | --- | --- |
| **Criteria** | **Acceptance level** | **Testing comments** |
| Accuracy of the data extracted from Prospectus | 100% | To maintain 100% accuracy level for all the extracted data points |
| Completeness of extraction of Key data points | 100% | * Asset Manager * Custodian * Delegated Asset Manager * Country of Domicile * Leverage / Commitment Approach/Value of risk Approach * FBL exemptions * Hedge Fund |
| Completeness of extraction of Key data points | 80% | * Register Address * ISIN * LEI |
| The automation should extend to all five languages | 100% | Text analytics should be performed on the following languages:   * English * French * German * Spanish * Italian |
| The data extraction should be in a excel format | 100% | All the Key data points identified through text analytics should be extracted into excel |
| User interface should have hyperlinks | 100% | All the data points that are captured should have a hyper link for the analyst to validate. |
| Peak capacity (no. of users) | 100 |  |
| UI | As in functional requirements please refer above | |
|  |  | |

# Business Process Continuity Steps

The Business Continuity Plan (BCP) provides a reference for resources to remediate a situation where an automated process has been disrupted. The BCP will break down the activities being performed by the automation and establish what will occur if an automation is disrupted. *For example, will a process owner coordinate and establish manual workaround due to the impact to customer or upstream process.*

|  |  |  |  |
| --- | --- | --- | --- |
| **Automation Task No** | **Automation Task** | **Implications of process failing**  (What are will happen if the bots stop working? i.e. how will it impact customers? Are there any upstream impacts to consider?) | **Continuation Strategy**  (If automation is no longer executing will a manual process be actioned? Who will action and own the process?) |
| 7.1 | Text analytics to extract the KDEs from prospectus. | Automation solution is intended to extract the KDEs for faster processing of Onboarding requests in GAR/GAM process, no direct impact to customer or upstream is identified in the event of down time. | The disruption of automated process will not prevent the analyst to perform the task manually if required. |

***Table 7:*** *Business Continuity Plan*

# Appendix and Attachments

|  |  |
| --- | --- |
| Appendix number | Title |
| Appendix 1 | Samples of Prospectus |
| Appendix 2 | KPI-Log Files template [Audit Report] |
| Appendix 3 | Attribute names referred with different names , business rules along with current and desired state |

# Definitions

| Term | Definition |
| --- | --- |
| KDE | Key Data Element |
| KPI | **Key Performance Indicator** |
|  |  |
|  |  |

# Document Control Sheet

Who to reach out to if you have any queries, questions, changes or concerns.

|  |  |
| --- | --- |
| **Name** | Pavankumar KAKI |
| **Position** | Business Analyst |
| **Phone** | +91 806731 **3546** |
| **Email** | [pavankumar.kaki@socgen.com](mailto:Nipu.mohan@socgen.com) |

If you have a suggestion for improving this document, please contact the person listed above.

|  |  |  |
| --- | --- | --- |
| Version number | Version date | Details of the change |
| 1.0 | 03/21/2017 | First Draft |
|  |  |  |
|  |  |  |

# Appendix

### Appendix 1: Samples of prospectus

Please refer spreadsheet with data extracts against the prospectus.

Note: **5** samples are covered with prospectus.



### Appendix 2: KPI –Audit report [To be concluded]

Definitions and Business Rules are covered in the below attached spreadsheet



### Appendix 3: Attribute analysis

The below attachment helps to identify the attribute name referred with different names, business rules along with current and desired state

