Data Quality Assurance & Governance Framework

Bondlinc Private Limited
Version 1.0.3
28 September 2020

Overview of Document Contents

1. Document Information and History

1.1. Version History

Version No.	Change Date	Approval Date	Changes and Description	Reviewed and Approved by
1.0.0	15 Sep 2020	16 Sep 2020	Base document formation	CTO
1.0.1	22 Sep 2020	23 Sep 2020	Inclusion of technology framework	CTO
1.0.2	25 Sep 2020	27 Sep 2020	Inclusion of business framework	VP, BD
1.0.3	28 Sep 2020	30 Sep 2020	Holistic review and additionals	COO

1.2. Distribution

New approved versions of this document must be distributed to the functions listed below. It is the responsibility of the document owner to initiate (re)approval processes and thereafter the responsibility of the approvers to approve the Information Security Policy.

Approvers:

Company & Function
Bondlinc Internal Management Team

Specification of third parties to whom distribution of approved versions of this document has been authorized:

Company & Function
Bondlinc Employees
Private Wealth Management Association (PWMA)

2. General

2.1. Purpose and Goals

The objective of this document is to outline the governance framework in place, for an end-to-end data management process and ensure high data quality is adhered to.

The data management and quality process will apply end-to-end, from data sourcing to processing, storage and eventual consumption by the users.

2.2. Scope and Target Audience

The data management and data quality framework apply to all staff and employees of the organization, contractual third parties and agents of the organization who have access to the organization's information systems or information.

The data specified as critical (as outlined in Appendix 1) must be subjected to the requirements in the data quality framework.

3. Data Management Framework

The information architecture for Bondlinc Private Limited ("Bondlinc") covers the following areas:

i. Data Governance
ii. Master Data Management
iii. Metadata Management
iv. Data Architecture
v. Data Access
vi. Data Storage

3.1. Data Governance

Refers to a set of capabilities for the lifecycle management of debt instruments, data stewardship and data quality.

There are separate processes in place to manage the above capabilities, as defined in sections 4 through 8 in this document.

3.2. Master Data Management

This refers to managing master data consistent with a single version of truth. Bondlinc maintains a reference data system which acts as a single source of truth to manage incoming data.

All data received from the various data sources is captured and updated in this system for post processing and published downstream to financial institutions. All data standardizations from multiple sources are handled at this layer.

3.3. Metadata Management

This refers to establishing lineage, common data dictionary and business glossary. Bondlinc receives data from multiple sources that are mapped to a common business glossary. Rules are further defined on top of these data elements to manage data quality.

All data quality checks and transformations are handled at a single layer, hence there is a clear lineage from the source to the downstream systems.

3.4. Data Architecture

This refers to the definition, structure, and design of enterprise data models' entities and attributes. Bondlinc has its own data model for bonds, to which the source data is mapped.

The data model ensures data integrity, enforces reference data standardization and stores data in a standard, consistent and predictable manner.

3.5. Data Access

This refers to the platforms and tools provisioned for users to access data. Bondlinc provides flexibility for the users to customize and view data elements on their respective platforms.

Users are authorized to view data points specific to their firm (e.g. Loan-To-Value, Product Risk Rating) on their dedicated platforms via two-fold checks:

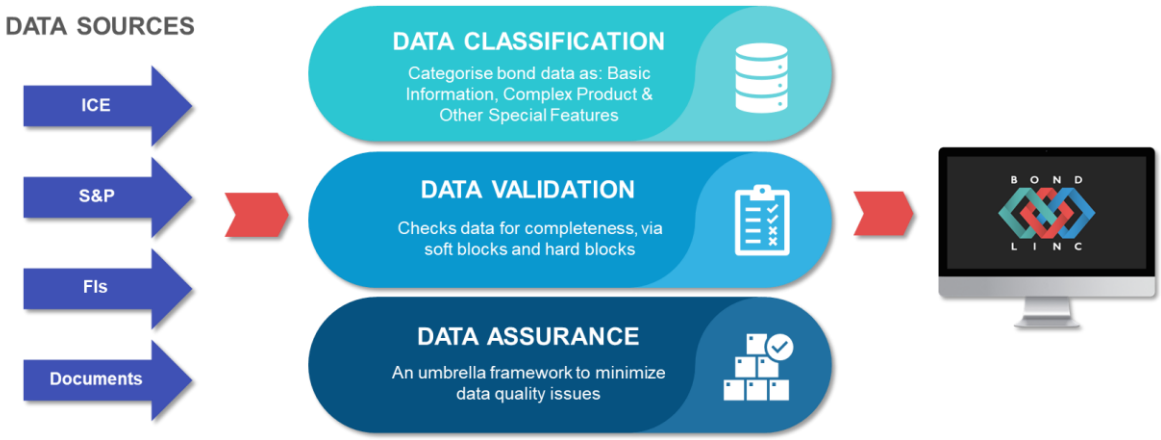
i. Their user profiles are configured to grant only the necessary visibility; and
ii. The financial institution ("FI") they belong to, is validated against, to ensure the users are only allowed to view content specific to their FI.

3.6. Data Storage

Data is stored on highly available storage devices which maintains confidentiality, integrity and availability.

4. Debt Instruments Life Cycle Management

Bondlinc provides preliminary, primary and secondary bond information on its platform. The diagram below provides an overview of the life cycle of data that comes in from Bondlinc's partners, to the processing and publishing of data on the platform for clients to access.



4.1. Data Sources

Bondlinc receives data from over 45 different sources in the form of documents, reference data, market data as well as FI specific data (e.g. Loan-To-Value, Product Risk Ratings, Solicitation flags). This puts us in a unique position, allowing us to cross-reference data to obtain a greater degree of completeness and accuracy. Having this holistic view allows us to challenge our data vendors and their rationale on behalf of our clients.

In addition to this, Bondlinc computes data from basic data points such as Yield-to-Maturity, Yield-to-Call, Yield-to-Worst, to more advanced data points such as Bond Complexity amongst others.

The types of data have been further elaborated in section 5.

4.2. Data Processing

The data received is then processed in three stages as outlined below, and has been described in detail in sections 6 through 8.

4.2.1. Data Classification

The data points are classified into three broad categories depending on their use cases:

i. Basic Bond Information
ii. Complex Product and Special Features Information
iii. Other Miscellaneous Features

4.2.2. Data Validation

Data is checked at this stage for completeness via built-in rules, to flag missing data components, and either allow or block the data from being loaded onto the platform, depending on its criticality.

The actions undertaken on the missing or incomplete data are categorized as:

i. **Soft Blocks:**
Missing or incomplete data elements are documented in Bondlinc's internal logs for further action, however the bond information as a whole is allowed to be published onto the platform as the missing elements are not deemed critical.

ii. **Hard Blocks:**
Missing or incomplete data elements are logged, and the bond information is prohibited from being published onto the platform. The Data Quality & Governance Team ("DQGT") will review the logs on a daily basis and source or rectify the data, where possible, in order to allow the data to proceed.

4.3. Data Assurance

It is never possible to completely eradicate data issues, hence the goal is to understand and realise that monitoring data is a ongoing and not a one-off process. At Bondlinc, data is closely scrutinised to attain a higher degree of accuracy, to provide more compelling and assuring datasets.

4.4. Data Output on the Bondlinc Platform

Once the data has been processed, it is made available on the Bondlinc platform for clients to access.

The environment may be dedicated to individual clients if there are customizations involved in accordance with the client's specific needs, or a shared environment (e.g. the Bondlinc

PWMA environment) where all members with the relevant subscription have access to the same set of information or data rules.

4.5. Data Quality & Governance Team

The Data Quality & Governance function of Bondlinc manages and oversees data handling, with different team members responsible for carrying out differing roles:

i. **Data Stewards**
Responsible for understanding and aligning with new regulations and business requirements (e.g. SFC regulations), sourcing required data fields from different vendors or through data partnerships, and capturing the requisite data into the system.

ii. **Data Quality Analyst**
Responsible for data validation, exceptions handling and data quality inspection.

iii. **Data Governance Council**
Responsible for liaising with the Private Wealth Management Association of Hong Kong (PWMA) and clients for data governance, and hosting board reviews where necessary.

5. Types of Data

The types of data available on the Bondlinc Platform have been outlined in the segments below. All data types go through the data classification, validation and assurance process which is described in sections 6 through 8.

5.1. Offering Documents and Prospectuses

All documents in the system are automatically categorised based on predefined rules (main and sub labels) into one of the following classifications:

i. **Prospectus - MTN Base**
This label is used on documents that are the base Medium-Term Note Programme. These documents do not contain any bond but the generic debt issuance programme.

ii. **Prospectus - MTN Bond**
Bonds that are issued as part of an offering programme are classified as MTN bonds, these are typically accompanied by an MTN Base, however there may be instances where an MTN Bond is issued with only the Base and a Pricing Sheet, skipping the Prospectus - MTN Bond document entirely.

iii. **Prospectus - Non-MTN Bond**
Bonds that are not issued as part of an offering programme are classified as non-MTN bonds.

iv. **Supplement - MTN Bond**
This categorisation is used to label supplement documents related to a Medium-Term Note Programme.

v. **Supplement - Non-MTN Bond**
This categorisation is used for prospectuses that are issued subsequent to a Prospectus – Non-MTN Bond, typically for updates to the offering risks.

vi. **Pricing Sheet**
Final term sheets or pricing sheets are categorised under this label.

vii. **Supplement - Pricing Sheet**
Typically used for future tranches of a pricing sheet, or a pricing supplement.

viii. **Others**
A document is defaulted to this categorisation if it does not fall under any of the previous categories.

ix. **Unclassified**
Documents are tagged to this category if the automatic classifier is unable to determine the document type and requires human intervention. This typically occurs for Marketing documents or Reports which are not prospectus related and immaterial.

These classifications are reviewed by the DQGT at the point of receipt before they are sent out to the clients and periodically, as part of a two-pronged approach, to ensure accuracy and remediation on a going basis.

An end-of-day batch file is generated by the system, with a list of all available documents on the platform and their respective classification. This file serves as an actionable insight for reconciliation by the DQGT.

5.2. Reference Data

Bondlinc sources its reference data from several data partners. These include the Fortune 500, New York Stock Exchange (NYSE) listed Intercontinental Exchange (ICE) and the renowned Standard & Poor (S&P) CUSIP Global Services, to information derived from bond prospectuses and other data contributors.

All reference data processed by the system goes through a set of mandatory Data Quality (“DQ”) checks, as outlined in sections 6 through 8.

These DQ checks include any mandatory basic information and complex product data points as required by the business. Exceptions flagged out are not loaded into the system but are diverted to the attention of Bondlinc's DQGT for investigation and further action.

In addition to this, data points that are not available from the secondary bond data providers are sourced from the bond prospectus or other proprietary mediums; for example, structural subordination and Chapter 37 classification. This ensures a greater degree of completeness and more encompassing data on the Bondlinc platform.

5.3. Market Data

Similar to the previous segment, market data is sourced from data partners and various other contributors. All market data that is processed by the system goes through a set of mandatory DQ checks as well.

The checks on market data are not hard blocking (i.e. the data will be loaded on the platform), but provide insights on missing prices, yield and modified duration, etc. Likewise, information of the missing data points are made available to the DQGT for review and further action.

5.4. Derived Data

Derived data is composed of information that has been computed or formulated based on multiple base data elements, these are made up of 2 or more data points. Derived data points include the likes of Yield-to-Worst (YTW), Yield-to-Maturity (YTM), Yield-to-Call (YTC), Loss Absorption, Special Features and Complex Product data points, amongst several others.

Computation of all derived data is documented by Bondlinc's Business Team based on their industry knowledge and experience as practitioners, in consultation and collaboration with fixed income specialists from the client institutions. The data elements go through validation and user acceptance testing before being implemented in production for general access.

Furthermore, all derived data are subjected to the data quality assurance process as described in section 8.

5.5. Financial Institution Specific Data

This includes data points received from FIs like Product Risk Rating (PRR), Loan To Value (LTV) and Bond Taggings (e.g. Buy / Sell lists, Solicitation lists, Sanctions Lists) that are specific to the individual FI.

The data received from the FIs are auto-forwarded to the system for processing, and involves little to no manual intervention. Data Validation checks are performed on this data, including checks for incorrect formatting and/or missing data. If the Data Validation checks fail, the FI is immediately notified for rectification and is required to resubmit the data for processing. The files sent through are documented for data audit purposes in ensuring its accuracy and mitigating any risks of malicious attempts to manipulate the data, thereby avoiding any reputational risk to the client.

All data processed is segregated and logically stored, and is only accessible from the specific FI's platform.

6. Data Classification

Data classification, in the context of debt instruments, is the categorization of data based on its use case(s) into mandatory information and the level of complexity of the debt instrument. The classification is based on the bare information required to trade the bond as well as for pre-trade checks.

As mentioned earlier in Section 4., Bondlinc classifies all data elements into three categories:

i. Basic Bond Information
ii. Complex Product and Special Features Information
iii. Other Miscellaneous Features

6.1. Basic Bond Information

This categorization applies to trade related bond information and comprises fundamental information required to perform a trade; such as a bond's identifier, issuer details, maturity, coupon, yield, ratings, industry classification, amongst others.

Appended below is a sample screenshot of the Bondlinc Platform outlining the basic bond details (please note: data fields may be subject to change depending on further enhancements).

Bond Details

ISIN	XS1515239942	Coupon Frequency	Semi Annual
Issuer Name	Huarong Finance II Co., Ltd.	Coupon Type	Fixed
Guarantor	China Huarong International Holdings Ltd	Reset Coupon	-
Currency	USD	Reset Date	-
Issuance Size	1,350,000,000	Country	China
Amount Outstanding	1,350,000,000	Industry Group	Investment Companies
Min Size / Incr Size	200,000 / 1,000	NAICS Industry Classification	Finance & Insurance
Issue Date	22 Nov 2016	Payment Rank	Sr Unsecured
Maturity Date	22 Nov 2021	S&P / Moody's / Fitch (Rating Status)	BBB+ / - / A (Rating Category : IG)
Callable	Yes	Product Risk Rating	3
Call Date	-	Documents	
Call Price	-	Factsheet (PDD)	
Coupon Rate	3.625	One-pager Document	

Note: (-) refers to datapoint being not available

6.2. Complex Product and Special Features Information

The Bondlinc complex product classification is implemented using a dedicated rule engine which is configurable by the Business Team. The rule engine supports customization to the rules implemented, as changes or updates may be required by the different FIs to classify complex products based on requirements mandated by their regional or even global teams.

The rule engine evaluates all logic in a granular and systematic manner to provide the following output variants: *YES*, *NO* and *INSUFFICIENT_INFORMATION*. All data tagged as *INSUFFICIENT_INFORMATION* is specifically validated and cross examined by the DQGT to false negatives.

The Bondlinc system screenshot sample below outlines the complex product data points currently in use (please note: data fields may be subject to change depending on further enhancements).

Complex Product Features

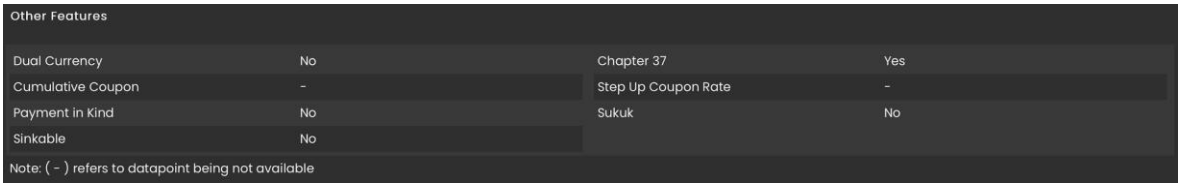
Complex Product	Yes	Special Features	Yes
Loss Absorption	-	Loss Absorption Mechanism	-
Defaulted	No	Non Viability Trigger	No
Basel III Designation	Not Applicable	Bail-in Feature	No
Subordinated	No	Puttable	No
Perpetual	No	Variable Coupon	No
Extendible	No	Deferred Interest	No
Convertible	No	Change of Control Trigger	Yes
Exchangeable	No	Keep Well Deed	Yes

Note: (-) refers to datapoint being not available

6.3. Other Miscellaneous Features

In addition to the complex product data points described above, there may be additional data fields that certain FIs wish to reference and/or disclose to their clients in order to provide additional depth of information - these are amalgamated under this section. Appended below

is a sample screenshot of the Bondlinc system which outlines some of the data points currently available (please note: data fields may be subject to change depending on further enhancements).



Other Features			
Dual Currency	No	Chapter 37	Yes
Cumulative Coupon	-	Step Up Coupon Rate	-
Payment in Kind	No	Sukuk	No
Sinkable	No		

Note: (-) refers to datapoint being not available

7. Data Validation

When data is received from the data partners, Bondlinc validates it to ensure its completeness. The appropriate checks are applied depending on whether the data points are deemed as critical elements or otherwise.

Critical elements are identified as data points that are essential for any bond to be made available on the Bondlinc platform, without which the platform would not host those bonds until the requisite data points have been sourced. The list for Critical Data Elements can be found under Appendix 1.

The data validation checks are divided into two categories - Soft Blocks and Hard Blocks.

7.1. Soft Blocks

Soft blocks apply to data points whose absence will not impact the bond information access or usage as they are not deemed as critical elements.

For instance, if a bond's reference data is received without a price or first coupon date, it is treated as a soft block. The soft blocks are recorded in internal logs and the bond would still be pushed through onto the platform, even without these data points. The exception records are available to the DQGT for review and further action(s).

Soft blocks are typically enforced on data that do not affect the fundamental nature of trading the asset.

7.2. Hard Blocks

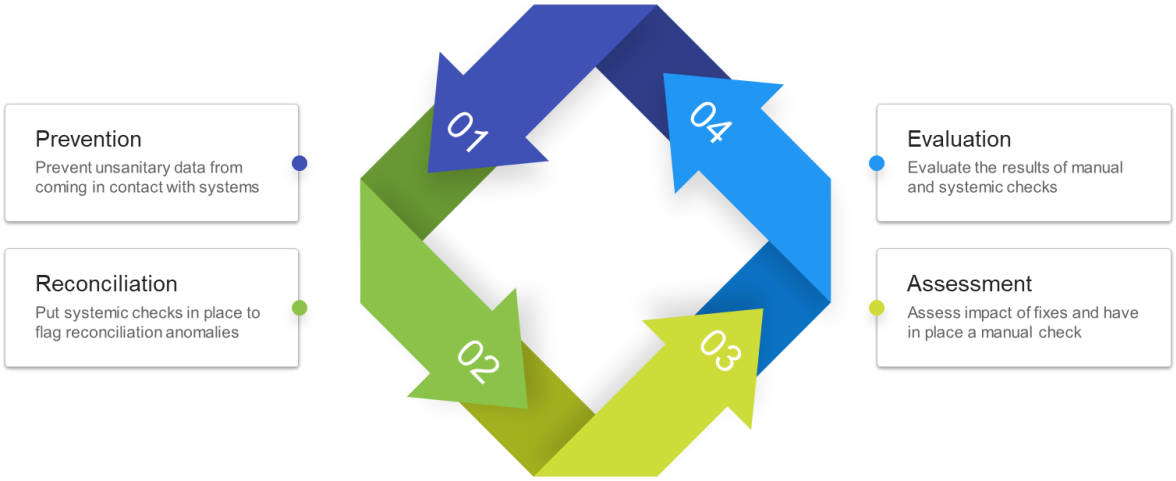
Hard blocks are applied to data points that have been identified as critical elements. A hard block would be triggered if information for these data points are missing when the bond reference data is received from sources. The bond as a whole will be prohibited from being loaded onto the platform, and the information would be recorded in internal logs for follow up action(s).

For instance, if bond data is received without an ISIN or coupon, it would give rise to a hard block and the data for this bond would not be loaded to the Bondlinc platform. The DQGT would then review the Hard Blocks Log and take the necessary action by either sourcing the missing information from the data partners or filing data challenges to the data vendors if necessary.

8. Data Assurance

As mentioned in the earlier segments of the document, data is constantly monitored for accuracy, with the goal of minimizing data quality issues using an umbrella framework.

This is achieved through a multi-fold process we dub the P.R.A.E methodology:



i. **Prevention**
The sourced data is scrutinised for various possible data quality issues before ingestion into the Bondlinc system. This facilitates in ensuring only clean data is introduced to the platform.

ii. **Reconciliation**
Daily reconciliation is performed on all data loaded to the reference data and pricing systems, to ensure data integrity is strictly adhered to.

iii. **Assessment & Evaluation**
Quarterly data quality checks are performed via sampling, and a Data Quality Scorecard is generated for internal management's review.

8.1. Source Data

Data for all preliminary, new and secondary bonds are checked for the following types of issues before access is granted for the data to be put through to the system.

i. **Missing Data:**
These errors are handled during the ingestion process with mandatory data quality checks, as outlined in Section 7, via soft blocks and hard blocks.

ii. **Duplicate Records:**
Bondlinc has a well-defined data model, and as a result, if there are any data violations, these are raised during the ingestion process.

We check for specific duplicates that help the system identify if a bond is a false positive or negative. These include global identifiers such as CUSIP and ISIN numbers where a duplicate should never occur.

iii. **Inconsistent Data Across Multiple Sources:**
These are highlighted as errors during cross-referencing secondary data against the primary data harvested.

iv. **Data Capturing Errors**:
Data capturing errors are generally detected through mandatory DQ checks (missing data) or through client feedback and data challenges. For such cases the origin of the data and the data in its default state is cross-referenced against all available sources. If it were found that the error arose externally, a data challenge would be raised with the source to rectify the issue.

Sources for performing checks differ across the different statuses of a bond as the degree of information available varies accordingly.

Bond Status	Syndicate Emails	Offering Documents[1]	Term Sheets / Prospectus[1]	Data Vendors
Preliminary Bonds	●	●	○	○
New Bonds	●	●	●	●
Secondary Bonds	●	●	●	●
Inactive Bonds	○	○	○	○

1 – availability of these vary across the bond universe

i. **Preliminary Bonds:**
These are bonds which have yet to be priced and are offered as initial public offerings. Information for these are limited to the syndicate desks, and they may in some circumstances be dropped off or cancelled depending on the level of interest garnered from the market.

ii. **New Bonds:**
These are bonds which have recently been priced and allocated within what we define as a 90-day window period.

iii. **Secondary Bonds**:
Secondary bonds are issuances which have traded for well over 90 days.

iv. **Inactive Bonds**:
Inactive bonds are not supported on a going basis as they would have either matured, defaulted or redeemed. A redemption event may arise as a call or put depending on the terms associated with the note.

8.2. System Errors & Reconciliation

All data loaded to the Reference Data and Pricing System are reconciled end-to-end daily against the source to confirm integrity and ensure every record is accounted for. All exceptions are handled on a case by case basis, depending on the system issues or the mandatory data check failures.

8.3. Data Quality Inspection

Bondlinc's system flags out anomalies as and when they are detected. We supplement this by performing additional manual checks on a quarterly basis to validate the data quality using a mix of stratified sampling and random sampling methods across the bond universe. The goal of such two-pronged approach paves the way for a more focused and unbiased inspection.

i. **Stratified Sampling:**

Stratified sampling involves dividing the bond universe into sub-universes that may differ in category definitions. This allows us to draw more precise conclusions by ensuring that every sub-universe is properly represented in the sample.

We start by dividing the universe into stratas based on specific characteristics we are looking to test (e.g. Change of Control, Dual Currency, Keep Well Deed, etc).

Based on the overall size of the universe, we proceed to calculate how many ISINs should be sampled from each strata and further apply a random sampling method to select a sample from each strata.

This method is typically adopted when performing checks on specific data points we are looking to test.

ii. **Random Sampling:**

We use random sampling when looking at the population of ISINs as a whole to pick out bonds without bias, and offering each ISIN an equal chance of selection.

Random sampling allows us to perform analysis on the data that is selected with a lower margin of error. The sampling occurs within specific boundaries that dictate the sampling process, and because the whole process is randomized, the random sample more accurately reflects the entire bond universe, allowing the results to be more representative and insightful.

Appendix

1. Classification of Data Elements

Basic Data Points		
Issuer Name	Incremental Size	Payment Rank
Guarantor	Issue Date	Credit Ratings
Issuance Size	Country	Product Risk Rating
Amount Outstanding	Industry Group	Loan-to-Value (LTV)
Minimum Size	NAICS Industry Classification	Reset Coupon
Reset Coupon	Reset Date	Price
Yield	Duration	Callable
Call Date	Call Type	

Critical Data Points		
ISIN	Coupon Rate	Maturity Date
Ticker	Coupon Frequency	
Currency	Coupon Type	

Complex Product & Special Features		
Complexity	Subordinated	Bail-in Feature
Loss Absorption	Perpetual	Puttable
Loss Absoprtion Mechanism	Convertible	Variable Coupon
Default Status	Exchangeable	Deferred Interest
Basel III Designation	Non-viability Ttrigger	Change of Control Trigger
Keep Well Deed	Structurally Subordinated	Multiple Credit Support

Other Miscellaneous Features		
Sinkable	Dual Currency	Cumulative Coupon
Sukuk	Payment In Kind	Chapter 37