

Solving the FRTB Data Management Challenges through an AI approach

June 2021

AGENDA



Story

Context

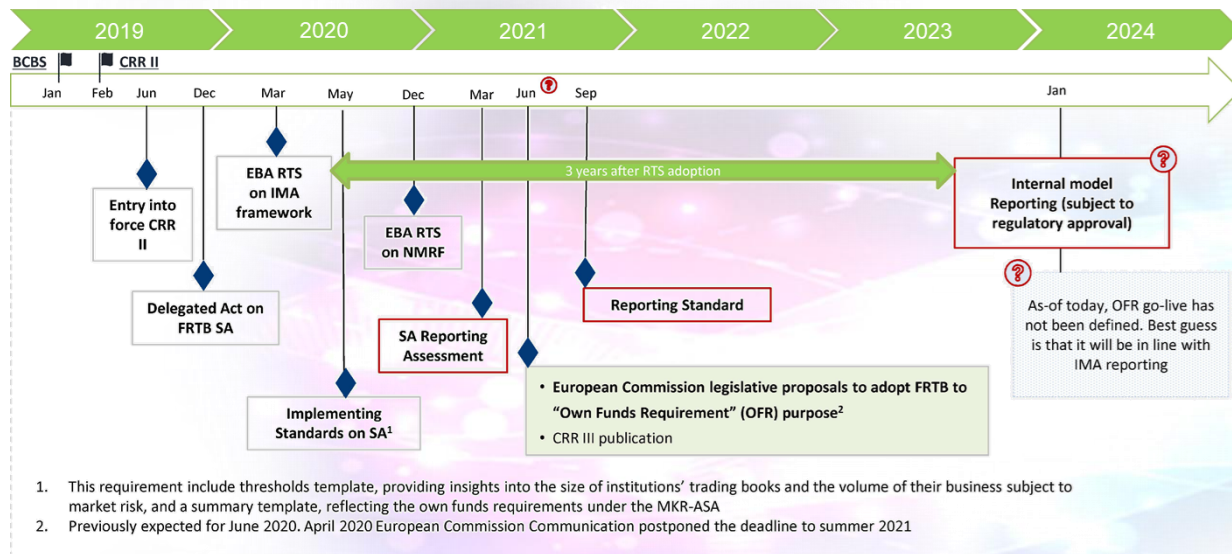
Challenge

FRTB Challenges

Hero

AI Approach





✓ **September 2021 Go-live Standard reporting**

- EBA launched a series of consultation **RTS** for a set of topics related to both IMA and SA implementation.

• IMA

- IMCC:** clarify that the data used to compute the ES risk measure should be accurate, appropriate, frequently updated, and complete and consistent
- DRC:** clarify the requirements to be met for the estimation of PDs and LGDs

• SA

- RRAO:** clarify the scope of the RRAO; specifying a non-exhaustive list of instruments bearing residual risks, and a list of risks that, in themselves, do not constitute residual risks (*due date June 2021*)
- JTD:** specify the key inputs needed for computing the Default Risk Charge (*due date June 2021*)
- EQUITY:** provide a list of countries with an advanced economy

THE OPTIONS WE HAVE

**Sensitivity Based Method
(SBM)**

Delta Risk, Vega Risk, Curvature Risk, calculated for each **risk class**, weighted according to **specific risk weight** and aggregated using **intra/cross bucket regulatory correlations**

**Residual Risk Add-On
(RRAO)**

Capital charge for **residual risk** in positions that are not in the scope of Delta, Vega e Curvature. It is calculated as the sum of the gross notional of the instruments multiplied by specific risk weight

**DEFAULT RISK CHARGE
(DRC)**

Capture default risk of credit and equity trading book exposure with no diversification with other market risks

Internal model approach**GLOBAL EXPECTE SHORTFALL
(IMCC)**

Modellable risk factor must be included in the bank's internal, firm-wide, expected shortfall model

**STRESS CAPITAL ADD-ON
(SES)**

Non-Modellable risk factor has to be capitalised using a stress internal model

**DEFAULT RISK CHARGE
(DRC)**

Capture default risk of credit and equity trading book exposure with no diversification with other market risks

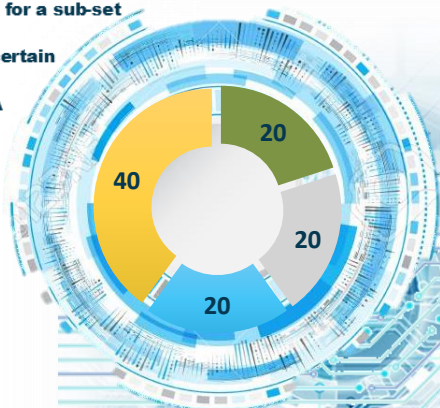
Standardised Approach

■ IMA for all desk

■ IMA for a sub-set

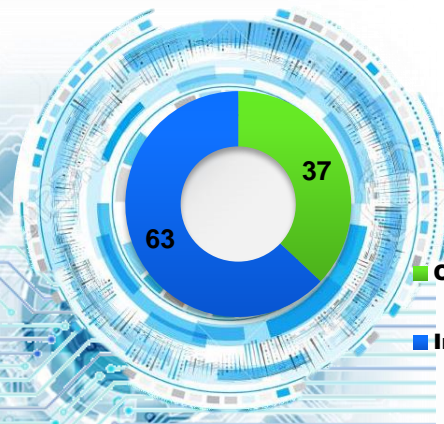
■ Uncertain

■ SBA



Main metrics

Sensitivities for SA
 IMA Non-Default
 metrics (IMCC)
 IMA NMRF metrics
 IMA Default metrics
 (IDRC)
 Validations test
 metrics



■ Outsourcing

■ In-House



2 Challenge WHERE WE ARE STRUGGLING

The regulatory framework sets methodological challenges both on SA and IMA point of view

SBA

As expected the FRTB SA is more compliant to Regulation – due to a more close deadline. However, the computation of all sensitivities based on **buckets** and all risk indicators has determined an increase of IT systems performance for the huge amount of data to be collected and aggregated.

IMA

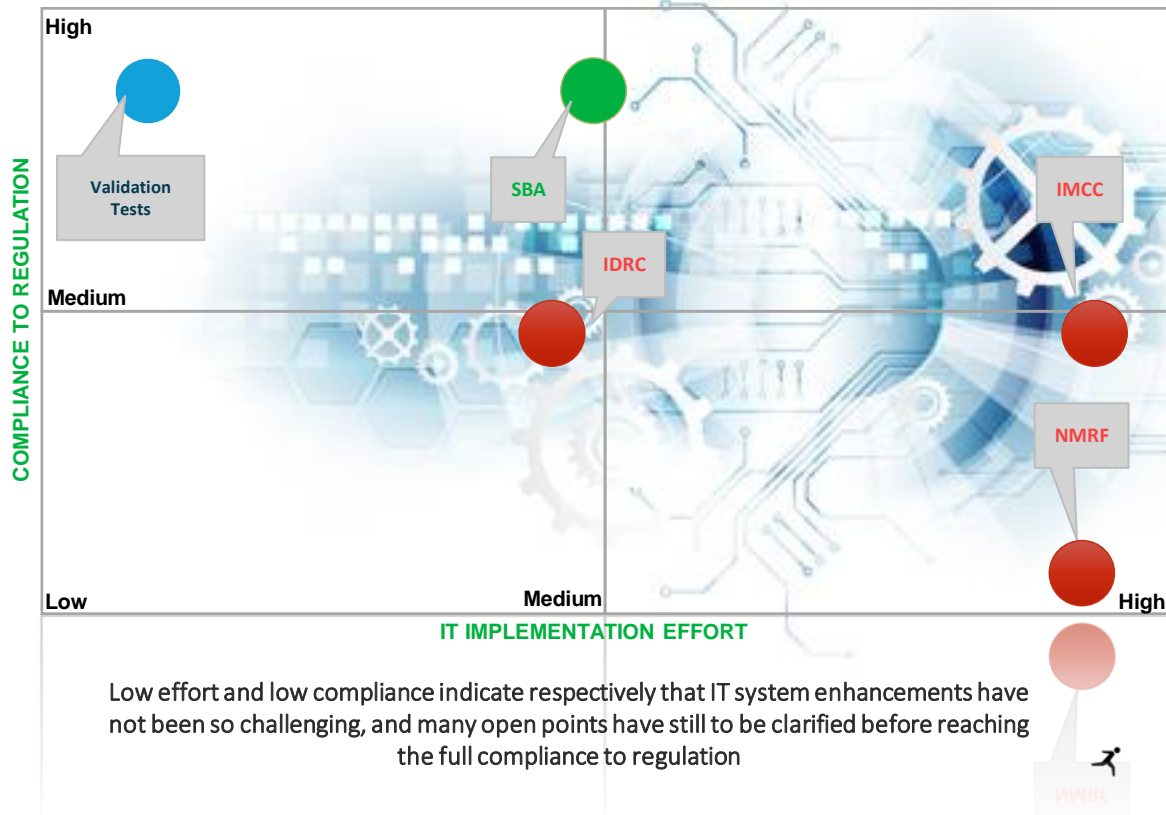
There are lot of open points still debated in the Industry. Some of the most impacting areas in terms of both business model organization and IT system solutions and implementation can be classified as follows:

- ✓ Backtesting and P&L validation test
- ✓ Modelling – Non-Modelling Risk Factors
- ✓ PD floor for Default Risk Charge

Validation Test

The implementation of P&L for Backtesting and PLA test metrics requires that IT systems have a high data storage capacity at disposal in order to store at least one-year observations (250) for all positions of trading desk in scope, and also a consistent computational power to compute, aggregate and visualize all the metrics in reasonable times. **For banks that decide to adopt internal model, one of the most difficult challenges could be to pass validation test for all trading desks.**

Main drivers: hardware enhancements, methodological computation and aggregation data logic, perimeter of revaluation, and data storage capacity



2

Challenge

AND WHY...

Data was the number one challenge identified for FRTB implementation, with 59% of 131 respondents to a Bloomberg poll citing it as their main challenge

As the SA is the default fallback for the IMA, the data and calculation engine required should be in place for all of the trading desks. A bank now has to compute at least **79 different calculation inputs for each sensitivity** class for risk computation under SA

The SA requires market data to correctly identify and compute sensitivities, and then aggregate exposures in accordance with FRTB rules; and then apply the necessary **weightings** before calculating their market risk capital

Look Through Approach for index and fund requires their **decomposition into constituent and underlying**

Curvature risk, it is similar to stress-testing and this poses significant challenges in terms of data availability

NMRFs present further complexity and are linked to the RFET. NMRFs require banks to collect real price observations for executed trades or committed quotes on a regular basis

The bank must identify for the risk factor at least **24 real price** observations per year

The bank must identify for the risk factor at least **100 “real” price** observations over the previous 12

FRTB mandates banks conduct Expected Shortfall (ES) and Default Risk Charge (DRC) calculations based on **data sets that go back to 2007 and 10 years** respectively

Identification of the **term structure** on which to map the risk factors for **each sensitivity**

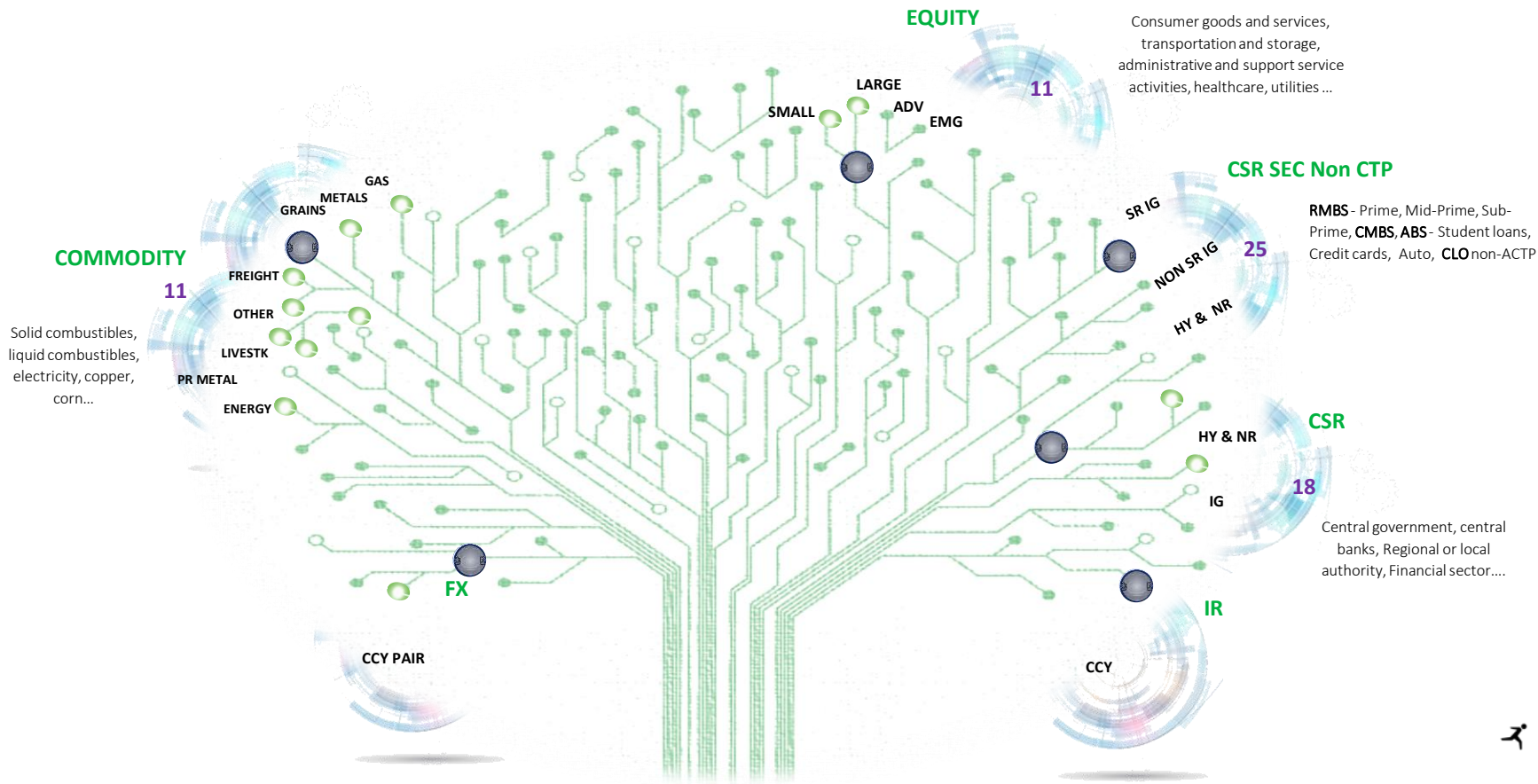
Differentiated **liquidity horizons** by risk class to compute Expected Shortfall



2

Challenge

... THE TREE OF LIFE



COMPLIANCE WITH THE REVISED MARKET RISK CAPITAL CHARGE IS A COMPLEX
JOURNEY

DATA WILL BE A CRUCIAL COMPONENT

POOR DATA WILL LEAD TO HIGHER
CAPITAL REQUIREMENTS

3

AI Approach

THE HERO IN ACTION

ASSISTED
BUCKETING

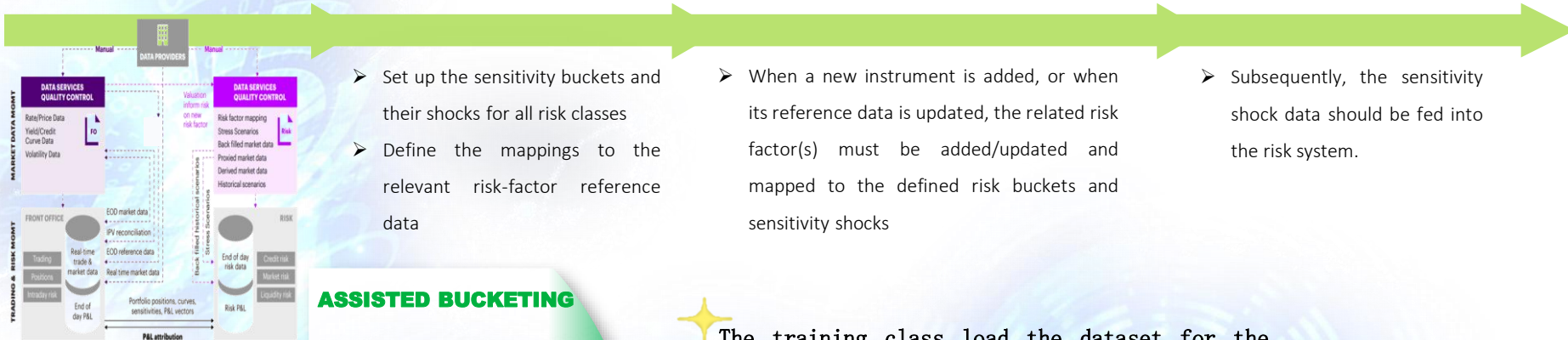
NMRF
MEASUREMENT

AUTOMATION

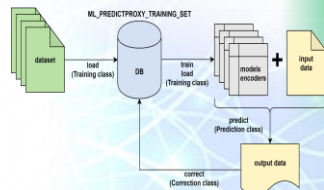


3 ASSISTED BUCKETING

CURRENT BUCKETING PROCESS



ASSISTED BUCKETING



A scheme is defined by targeting each bucket characteristic (ML_PREDICTBUCKET_TRAINING SET)

The training class load the dataset for the bucketing process aligning the encoders and the reference models (e.g. Decision Tree Classifier Bond & Equity)

Given the input portfolio, encoders and models are iteratively loaded in order to predict, for each asset, the corresponding bucket

It is possible to correct model predictions by modifying the corresponding scheme

3

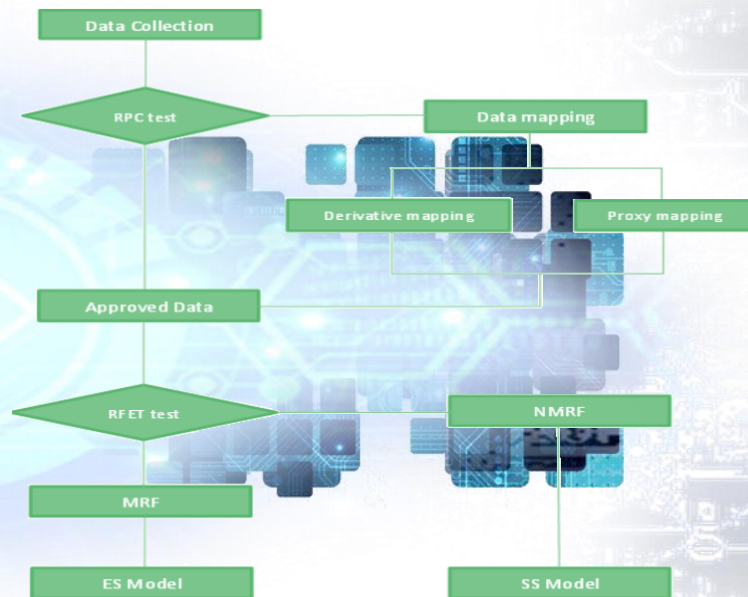
AI Approach

NMRF MEASUREMENT

The impact of NMRF remains significant, leading banks to look for other ways of reducing the number of non-modellable risk factors
One area worth exploring further is the use of proxies where modellability data for a risk factor is unavailable

Statistical modelling

A statistical model is used to discover the best proxy by looking for the most correlated risk factors and using data from that MRF as a proxy for the NMRF



Linear Gaussian Discriminant Analysis
Quadratic Discriminant Analysis
Naive Bayes Classifiers
k-th Nearest Neighbour
Logistic Regression
Decision Trees
Support Vector Machines (SVM)
Neural Networks

	NN-Tangent-ActFunc	NN-ElliptSig-ActFunc	SVM-Poly-kernel	SVM-Gauss-kernel	Bagged Tree-LC30	NN-Linear-ActFunc	QDA-FullCov
Classifiers							
Accuracy	99.3%	98.5%	96.8%	96.6%	96.0%	95.5%	95.2%
Ranking	1	2	3	4	5	6	7
Stdev	0.6%	1.6%	1.5%	1.9%	2.2%	0.6%	1.6%
Ranking	1	5	3	6	8	2	4

	KNN-city-dist	KNN-eucl-dist	SVM-Linear-kernel	KNN-maha-dist
Classifiers				
Accuracy	95.2%	95.1%	93.4%	92.6%
Ranking	8	9	10	11
Stdev	2.3%	1.9%	2.5%	2.6%
Ranking	9	7	11	14

Re-Run automation

The use of history sets stretching back to 2007 and the Liquidity Horizon buckets in the IM will result in an increase of computations, with one publication suggesting about 30 times more. Rather than the results be found by an analyst staying up late (or even by end users the next day)

Machine Learning algorithms could be applied to identify spurious/incorrect results and immediately commence a re-run for that portfolio.

Elliptic Envelop considers the distribution of multi-dimensional events to identify when an event has different characteristics in terms of correlation between the variables that form it.

Report publishing optimisation

With an increased amount of data that will be produced under FRTB, network bandwidth and storage capacity is likely to come under pressure.

We use Machine Learning algorithms such as **transfer learning and generative adversarial networks (GAN)** to determine when numbers are produced and push the optimal amount of reports and data out (or cache on a local server) to end users

Carmine Lombardi
Manager

Mobile: +39 3406393134
E-mail: ca.lombardi@reply.it

Via Castellanza, 11
20151 Milano



Thank you