



Irving Fisher Committee on
Central Bank Statistics

BANK FOR INTERNATIONAL SETTLEMENTS

IFC – Bank Indonesia International Workshop and Seminar on *“Big Data for Central Bank Policies / Building Pathways for Policy Making with Big Data”*

Bali, Indonesia, 23-26 July 2018

Exploring big data to sharpen financial sector risk assessment¹

David Roi Hardoon,
Monetary Authority of Singapore

¹ This presentation was prepared for the meeting. The views expressed are those of the author and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the meeting.

EXPLORING BIG DATA TO SHARPEN FINANCIAL SECTOR RISK ASSESSMENT

“Building Pathways for Policy Making with Big Data”

BI-IFC / BIS International Seminar on Big Data, 26 July 2018

David R. Hardoon
Monetary Authority of Singapore



1 | THE VISION

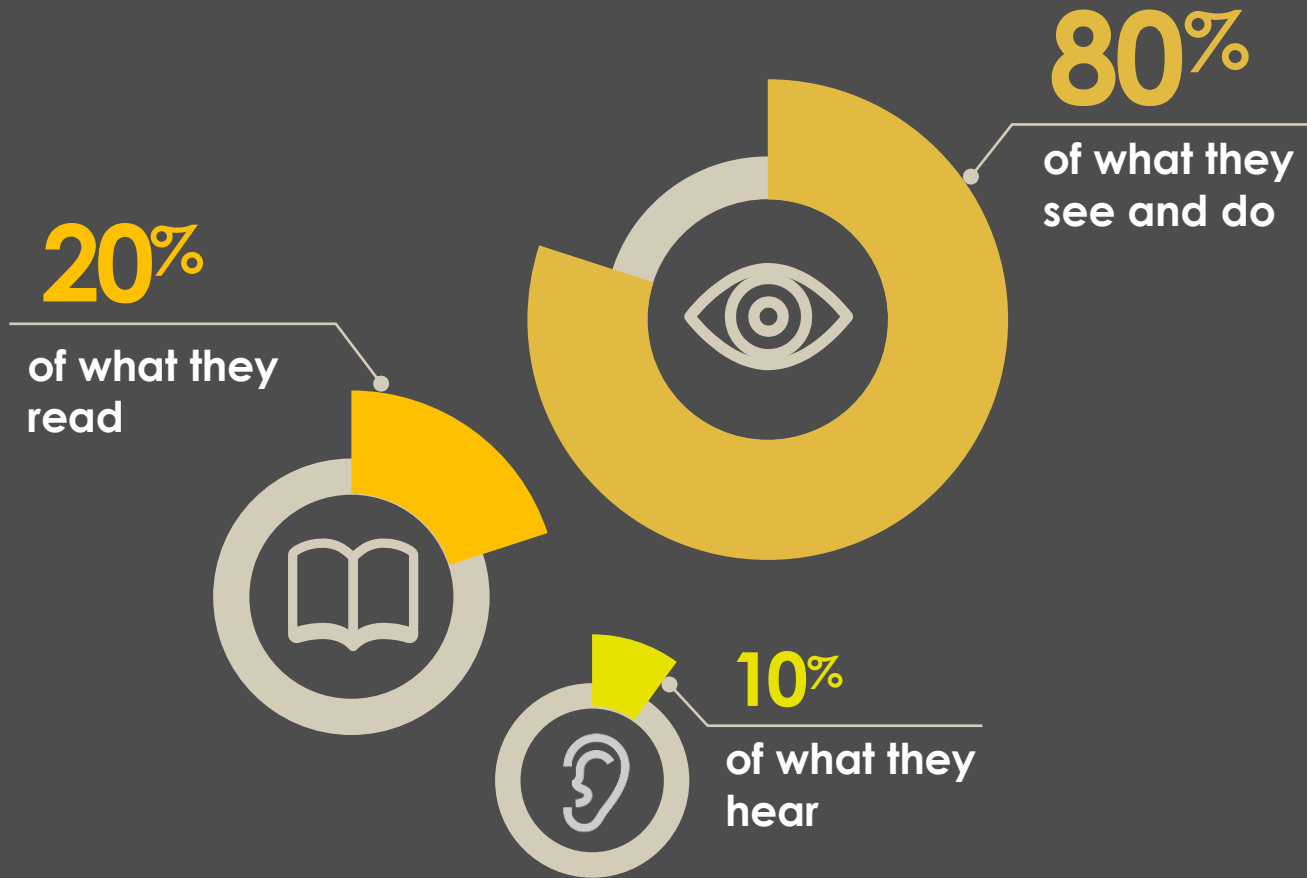


**DOING
BUSINESS
AS
UNUSUAL**



2 | VISUALISATION

People remember....



3 | TRAINING PLANS

Interactive Data Analytics Course Catalogue

- In-house
- External - Online
- External - Classroom

Advanced				
<ul style="list-style-type: none"> ● Data Science Specialisation ● Reproducible Data Science 			<ul style="list-style-type: none"> ● Deep Learning Specialisation ● Advanced Data Science: Machine Learning 	<ul style="list-style-type: none"> ● Building A Data Science Platform with Kubernetes, Anaconda & Jupyterhub
Intermediate				
<ul style="list-style-type: none"> ● Intermediate Data Science (Python) ● Introduction to Python for Data Science ● Introduction to Python Programming ● Getting Started with Data Science (Python) ● Foundational Python for Data Science ● Master Class on Advanced Analytics ● Master Class in Decision Making with Data Analytics 	<ul style="list-style-type: none"> ● Intermediate Data Management & Analysis ● Data Preparation & Exploration Using Python ● Data Analytics and Techniques with Excel and VBA ● Intermediate Data Collection & Analysis 	<ul style="list-style-type: none"> ● Data Analyst - Visualisation II ● Introduction to Coded Visualisation ● Tableau Test Drive (for Power Users) ● Tableau Townhall (for Normal Users) ● Data Analyst - Visualisation I 	<ul style="list-style-type: none"> ● Introduction to Machine Learning 	<ul style="list-style-type: none"> ● Stata for Public Policy Workshop ● Statistics for Bureaucrats - Stats 101 for Policymakers ● Workshop on Randomised Controlled Trials in Public Policy
Basic				
<ul style="list-style-type: none"> ● Data Analytics: Core Concepts & Techniques ● Data Analytics for Managers ● Data Science for Managers: Crash Course in Data Science ● Starting a Data Science Project 	<ul style="list-style-type: none"> ● Customer Research through Surveys and Analytics ● Introduction to Data Management & Analysis ● Survey Questionnaire - Design & Implementation ● Introduction to Data Collection & Analysis 	<ul style="list-style-type: none"> ● The Art and Science of Data Visualisation ● Data Analytics & Visualisation for Managers ● Business Intelligence & Dashboarding ● Introduction to Data Visualisation and Communication ● Data Visualisation for Effective Communication 	<ul style="list-style-type: none"> ● Introduction to Predictive Analytics - Tools & Techniques ● Data Science & Artificial Intelligence for Senior Executives 	<ul style="list-style-type: none"> ● Stats Alive - A Foundation for Analytics ● Master Class on Building An Analytics Culture ● Basic Analytics for HR Professionals ● Driving Communications with a Data Analytics Strategy
Data Science	Collection, Preparation and Exploration	Visualisation	Machine Learning	Others

4 | CURRENT TECHNOLOGY & PROJECTS



MACHINE LEARNING



**TEXT ANALYTICS &
NATURAL LANGUAGE
PROCESSING**



VISUALISATION



AUTOMATION

5

TEXT ANALYTICS & NATURAL LANGUAGE PROCESSING

Examples of Text Analytics and NLP



SENTIMENT ANALYSIS



TOPIC MODELLING

1. Scrape Chinese news websites
2. Clean data

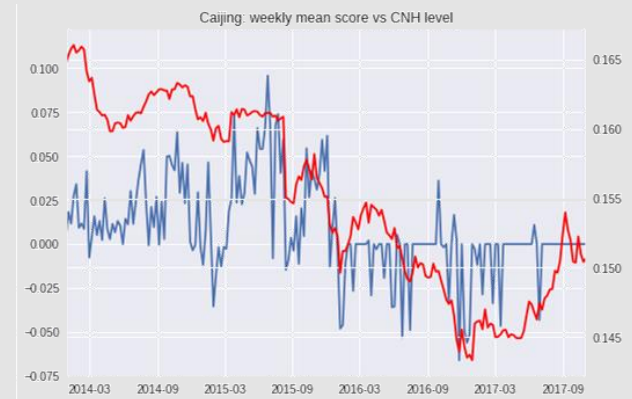


3. Calculate sentiment scores

```
In [8]: text = '''10月15日，中国人民银行行长周小川在华盛顿出席国际货币基金组织/世界银行年会期间，在G20国际银行业讨论会上就中国经济前景发表演讲。周小川指出，今年以来，中国经济增长动能有所回升，下半年增速有望实现7%，整体杠杆率开始出现下降，并有望保持这一趋势。未来，将进一步深化改革，逐步推动经济去杠杆，加强金融监管协调，推动金融市场平稳健康发展，维护金融稳定，经济增长动能回升。去杠杆去产能取得初步效果过去几年来，中国经济增速持续放缓，自此前高于10%降至2012年的8%左右以后，继续降至2016年的6.7%。'''

In [9]: print(", ".join(jieba.cut(text)))
```

4. Visualise results and compute correlation



Correlation coeff: 0.56

6

TEXT ANALYTICS & NATURAL LANGUAGE PROCESSING

Examples of Text Analytics and NLP




SENTIMENT ANALYSIS



TOPIC MODELLING

DATA

Name	Feedback
A	Encourage better email practices, for example reducing the number of emails circulated and expectations of instant replies. This can allow staff to focus on their work instead of checking emails all the time.
B	Reduce the amount of processes for greater efficiency and effectiveness. There are already a lot of rules in government etc. Having in additional checks and balances is already tight process increases significant work but the additional benefit. Additional safeguards should not be introduced unless it is a significant risk. For example, the procurement cycle has been revised several times just this year alone. There may be good reasons for doing so, but perhaps it is more important for us to reduce the cost, then to tighten the system every time there is an 'infringement'. The other's way to do and implement, but it comes at a great cost to the organization in terms of productivity.
C	Increased reliance on technology should be tempered. Relying on technology as an engine of growth is a good idea, forward and drive forward thinking. However, pushing out new technology advances have to be done more carefully, not all it offers is suitable for the differing needs of different departments, and it is important that the software be thoroughly tested and suitably flexible for all departments.
D	As departments' annual budgets could be cut despite having accurately budgeted for their needs, they would need to look for funds to be used to make up for the shortfall. Departments would be questioned on the reasons for the shortfall, even though their original budget would have been accurate if it was not cut.

TOPIC MODELLING (e.g. NMF, LDA)



TOPICS

Topic 1	work processes, streamlining, prioritisation
Topic 2	workload, headcount, resources
Topic 3	productivity, efficiency, systems and technology

7

VISUALISATION

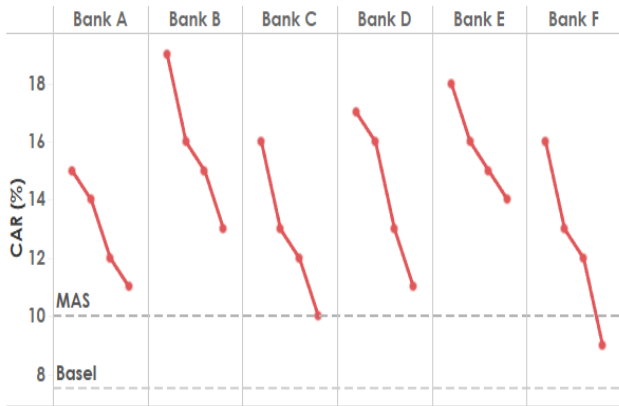
Industry Wide Stress Test (IWST) Dashboard



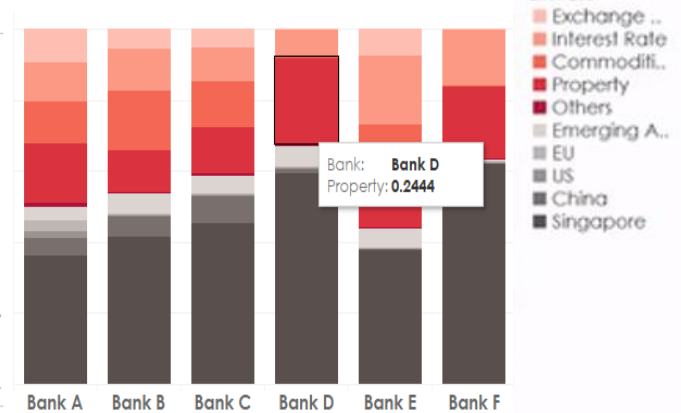
IWST Stress Scenario



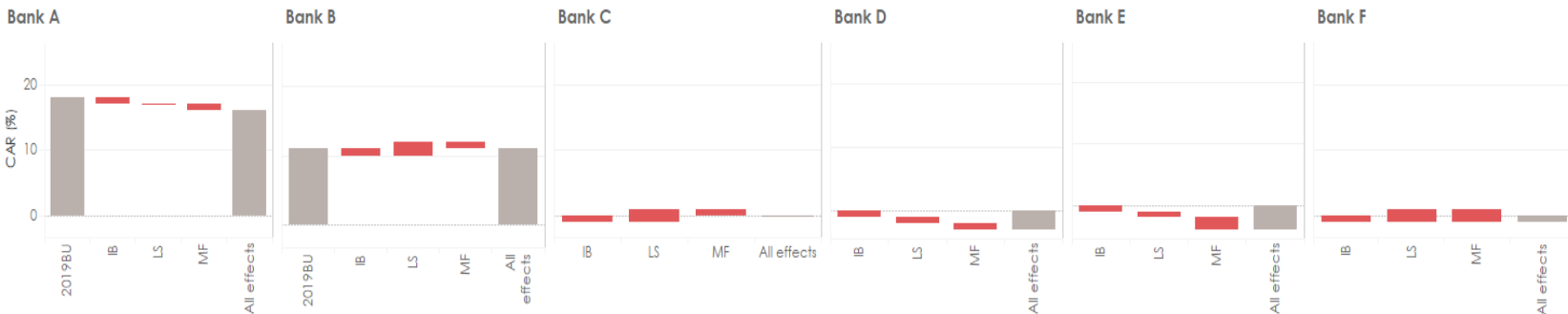
CAR (2016-2019)



CAR Impact Attribution



Impact of Second Order Effects



Legend

- BU : Bottom Up
- LS : Liquidity Solvency
- MF : Macro-Financial
- ALL : All effects

8

AUTOMATION

Filtering News Articles

Automating news surveillance

PROBLEM

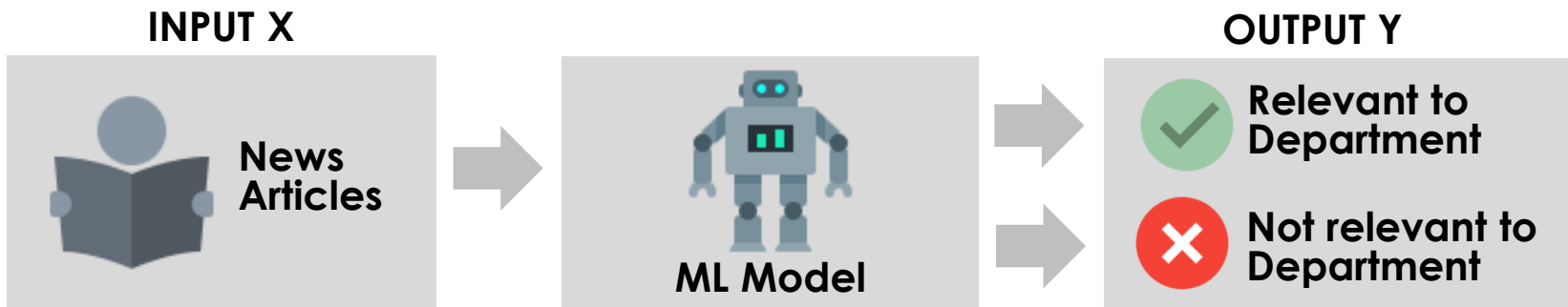
Many departments in MAS currently filter news alerts by keywords, and rely on daily checks by support officers to assess if there are specific events that require attention. This process is time-consuming.

AUTOMATION

To help automate the process, we can train a model based on labelled articles.

PRODUCTIVITY GAINS

The model predicts whether a new article is relevant to a department. A similar model is currently used by Comms in filtering articles relevant to MAS, saving them a lot of time from the manual filtering.



9 | FUTURE OF MAS DIGITAL SUPERVISION

Example of 'Bank A' Portal View

Latest news directly relevant to Bank A

Ongoing communication btw MAS & Bank A

Suspicious transactions reported by Bank A

MAS' assessment of Bank A's risk over time

Supervisor's rating of Bank A on different indicators

Inspection assessment of Bank A

Supervision Of Banks

Event Impact Analysis

Supervision Conditions

BANK A	Q1	Q2	Q3	Q4	Line
ACF, (\$)	80	70	80	90	
CAR, (%)	50	30	60	60	
LDR, (%)	70	80	60	30	
MCR, (%)	80	90	60	70	
MLA, (%)	60	60	30	30	
TPUG, (%)	50	30	70	30	

CRAFT Dashboard

Outstanding Case

- Case No. 100001
- Case No. 100002

STR Analysis

STR No.	Review	Escalate
100001		
100002		
100003		
100004		

Inspection

Bank	Assessment
Bank B	Pass
Bank C	Substandard
Bank D	Pass
Bank E	Pass

10 | ALERT PRIORITIZATION

The Pain & The Cure



130 Alerts per week

11.6% are important



4 Analysts

2 days per week

Pain



Cure



Machine to learn and
classify alerts by
importance

75%



To capture as many
important alerts as possible



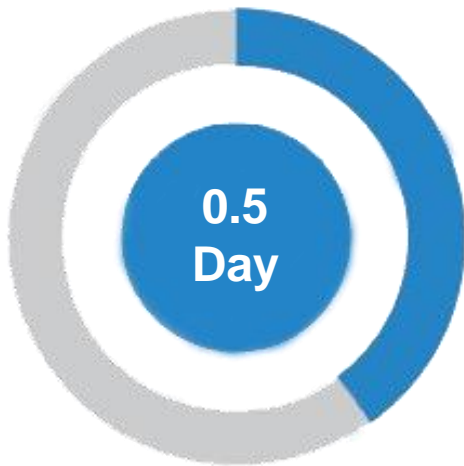
25%



To reduce workload

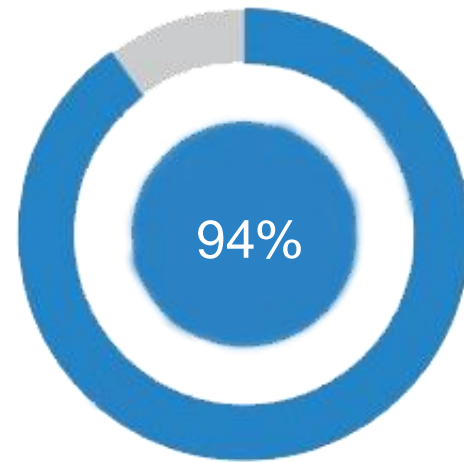


11 | ALERT PRIORITIZATION - THE RESULTS



Workload Efficiency Gain

2 days per week reduced to 1.5 days. Monthly gain of 4 days



Alert Capture Recall

Objective is to capture as many positive alerts as possible

Note: The results are based on the optimal model and 10-fold cross validation

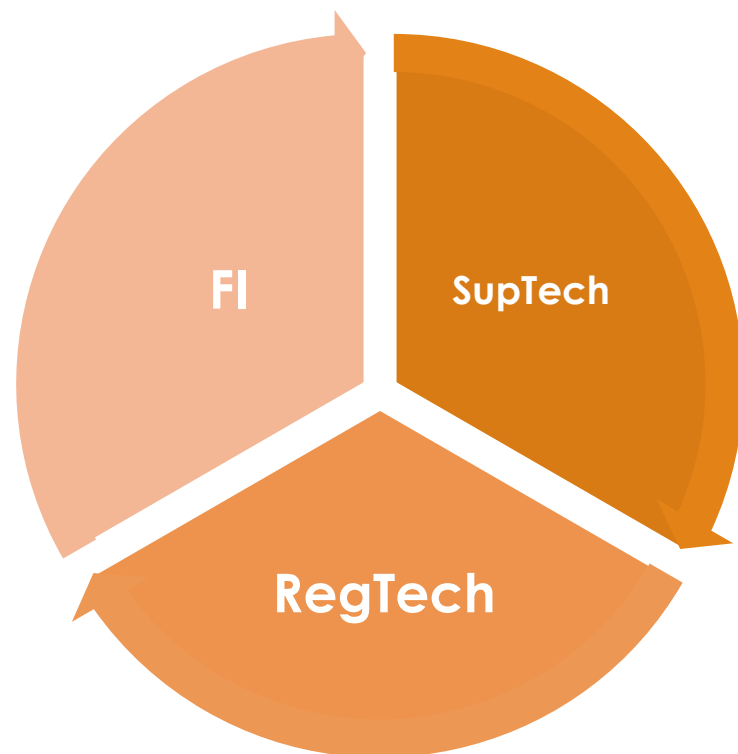
12 | FI-REGTECH DIALOGUE

Enhance regulatory compliance with RegTech and latest technology

1 RegTechs to better understand the problem statements from FIs

2 FIs to better understand solutions from RegTechs

3 Encourage RegTech/FI interaction and potential POCs



13 | ENABLING CROSS BORDER INSIGHTS



THANK YOU!

