IFC-Bank of Italy Workshop on "Data Science in Central Banking: Applications and tools"

14-17 February 2022

Integrating natural language processing technologies to central bank operations at Bank of Thailand[1]

Jiradett Kerdsri and Pucktada Treeratpituk,
Bank of Thailand

[1] This contribution was prepared for the workshop. The views expressed are those of the authors and do not necessarily reflect the views of the Bank of Italy, the BIS, the IFC or the other central banks and institutions represented at the event.

ธนาคารแห่งประเทศไทย
BANK OF THAILAND

Integrating Natural Language Processing To Central Bank Operations at Bank of Thailand

Pucktada Treeratpituk, Jiradett Kerdsri

Department of Data Management and Data Analytics
Bank of Thailand

15 Feb 2022

ธนาคารแห่งประเทศไทย
BANK OF THAILAND

NLP for BOT’s Operations

Why NLP?

- Lots of textual data, ranging from unstructured economic data to financial documents to social media information that can be used to enhance efficiency & effectiveness of Central Bank operations.
- Additionally, due to the cheer volume, it is desirable to use ML to process these data.
- Data-driven Organization - data in many departments are mainly textual data.

Challenges

- Unstructured & Data Quality (eg. PDF scanned, social media, free-text data fields)
- Thai Language (characteristics of the language + lack of readily available tools).

ที่ผ่านมา ธปท. ได้ติดตามและเร่งบริษัทปรับระบบให้มีกระบวนการแสดงตนและพิสูจน์ตัวตนลูกค้าบุคคลธรรมดา (KYC) สอดคล้องตามหลักเกณฑ์ของกฎหมายป้องกันและปราบปรามการฟอกเงินและการสนับสนุนทางการเงินแก่การก่อการร้าย ซึ่งเป็นกระบวนการสำคัญที่จะช่วยดูแลให้การทำธุรกรรมทางการเงิน มีความปลอดภัย ไม่เป็นช่องทางให้เกิดการทุจริตหรือฟอกเงิน ที่เป็นภัยต่อประชาชนและระบบการเงินของประเทศ

example

Projects

1. Behavioral & Culture Supervision though Board of Directors’ meeting minutes.
2. Automate Statistics Compilation of Government Expenditure.
3. Assessing and Managing BOT’s Reputational Risk through social media analysis.

ธนาคารแห่งประเทศไทย
BANK OF THAILAND

1. Behavioral & Culture Supervision

Assessing organizational behavior and culture (B&C) of financial institutions to promote good corporate governance

B&C Toolbox

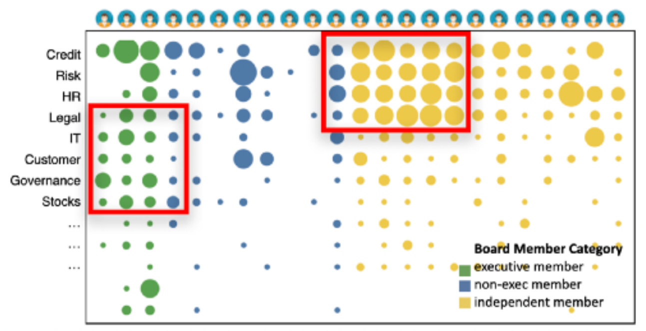
1. ***Self-Assessment***
2. ***Board Observation***
3. ***Interview & Dialogue***
4. ***Desk Research***
5. ***Expectation Framework***
6. ***Enforcement***

Board effectiveness & Risk Culture

Because BOT cannot directly observe board meetings, BOT evaluates the Board's Effectiveness through board survey and analyzing board meeting minutes.

board of directors meeting minutes

PDF → ***Thai OCR*** → *Document Structure Parsing (heuristics,* ***HMM****)* → ***Thai Word Segmentation****, Topic Modeling, Name-Entity Extraction*



Topics of Discussion

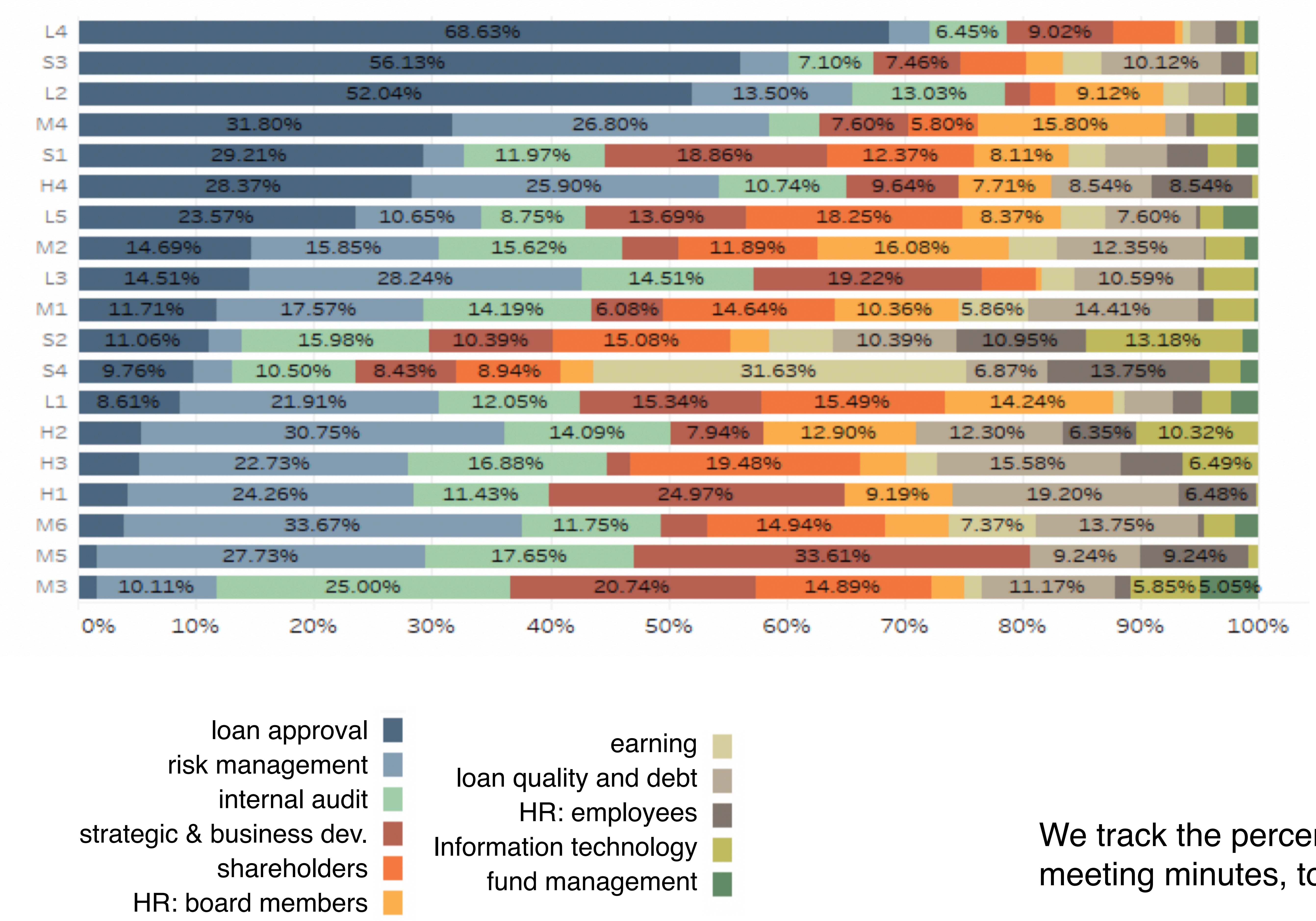
Participation of Board Members

Suitability of Time & Discussion

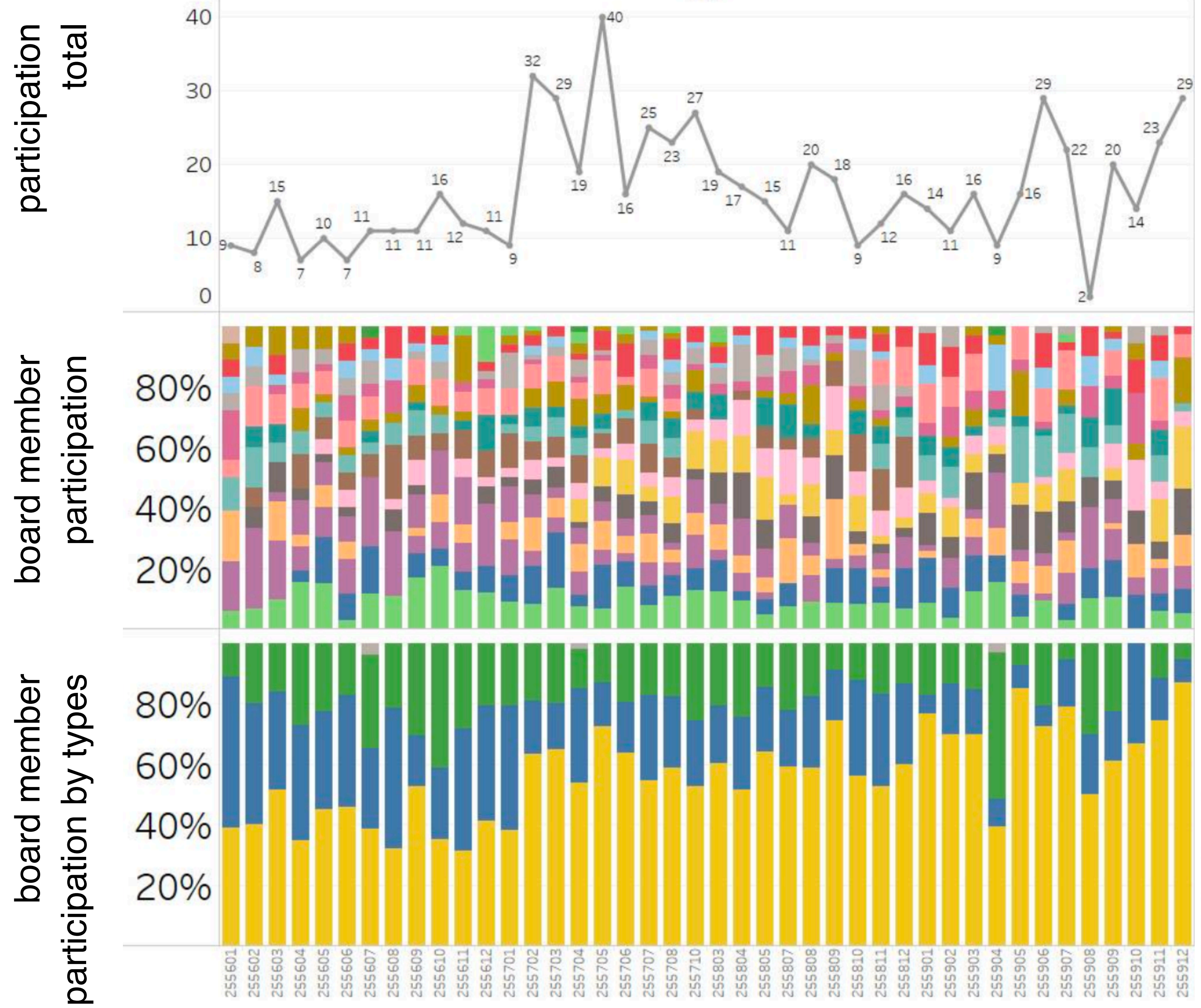
ธนาคารแห่งประเทศไทย
BANK OF THAILAND

1. Behavioral & Culture Supervision

% of topics discussed by different FIs

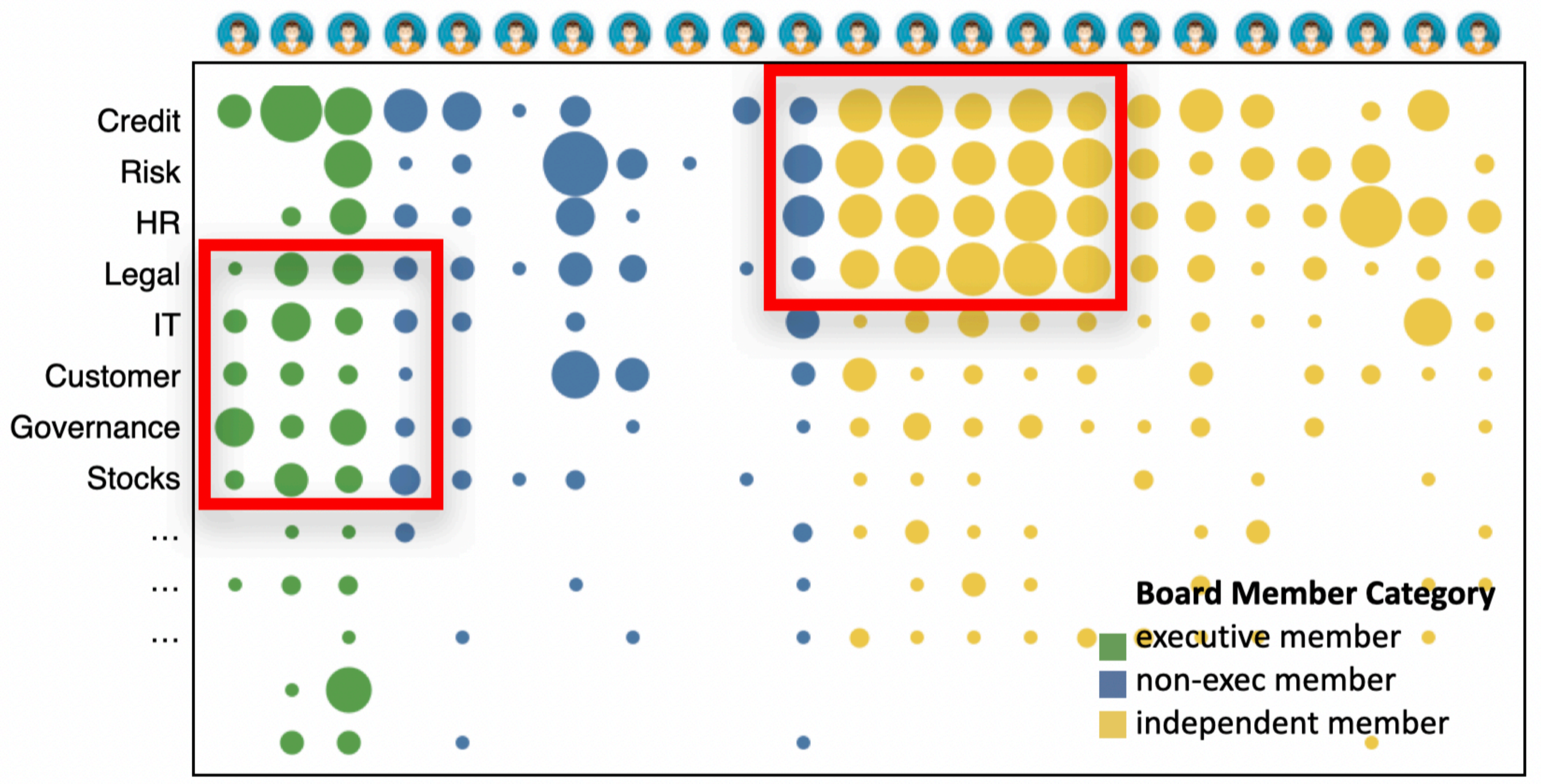


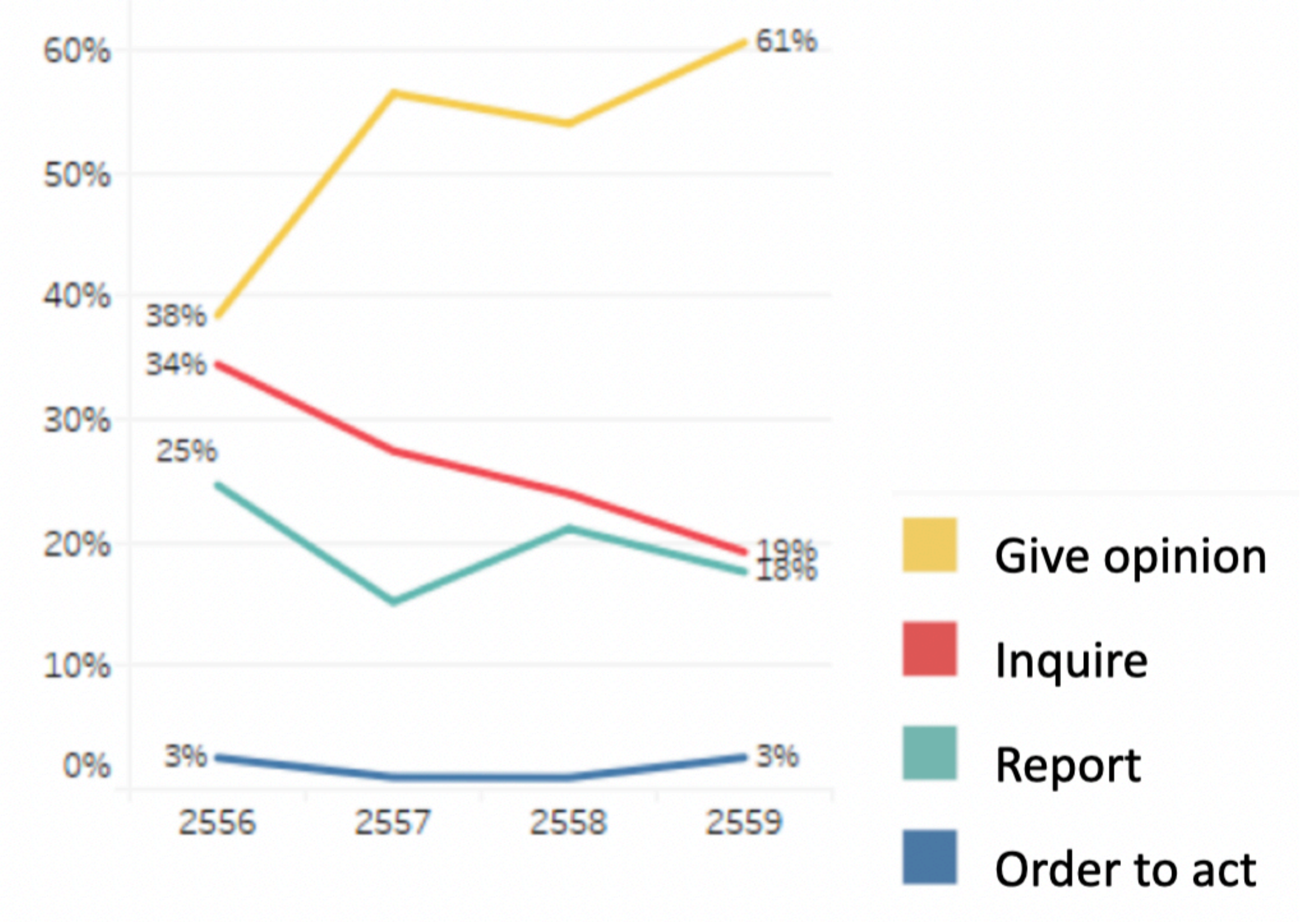
board member participation according to meeting minutes



We track the percentage of topics discussed by different FIs in their board of director meeting minutes, together with the participation rate of each board members over time.

1. Behavioral & Culture Supervision





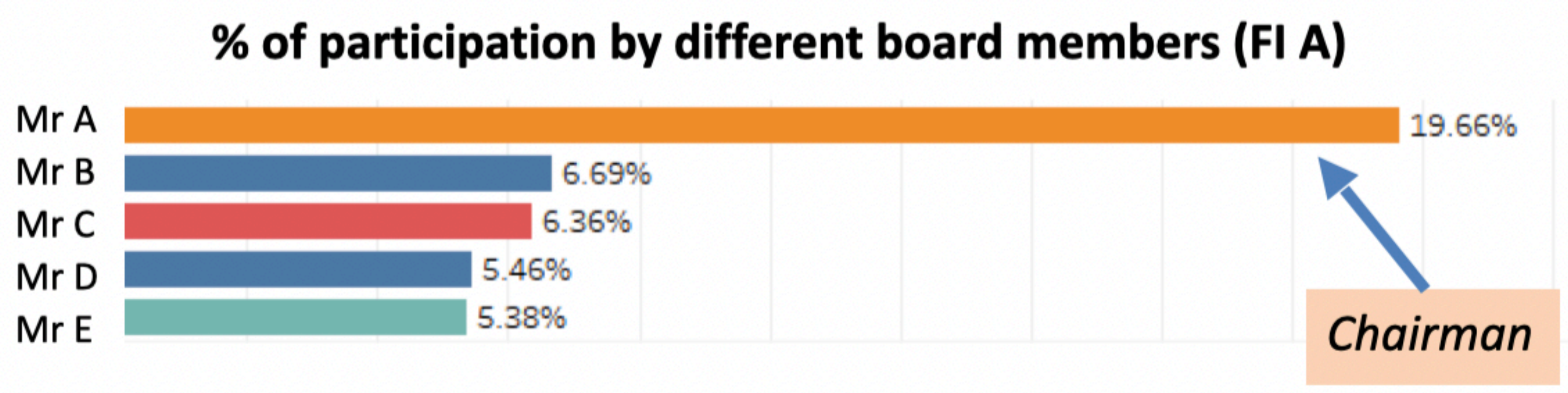
Additionally, We track the participation of each board member of each FI, in different topic of discussion. This enables us to study and to monitor different culture of each FI board. For instance, in one FI, we see that for certain specialize topics such as IT and legal issues, only some specialize board members participate in the discussion (sparse circles). While in some topics such as credit and risk, there are more open discussion (dense circles in the box).

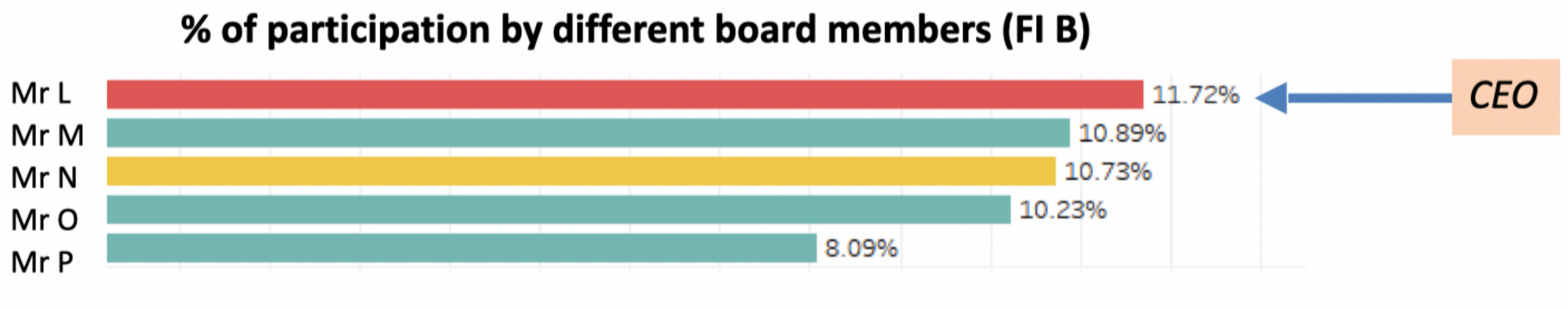
We also track "type" of participation by each board member whether they simply inquire, report, give opinion or giving order. This allow us to study the change in behavior of each board members. For instance, some board members become more opinionated the longer they stay on the board.

ธนาคารแห่งประเทศไทย
BANK OF THAILAND

1. Behavioral & Culture Supervision

Examples of different pattern of board member participation between different FIs



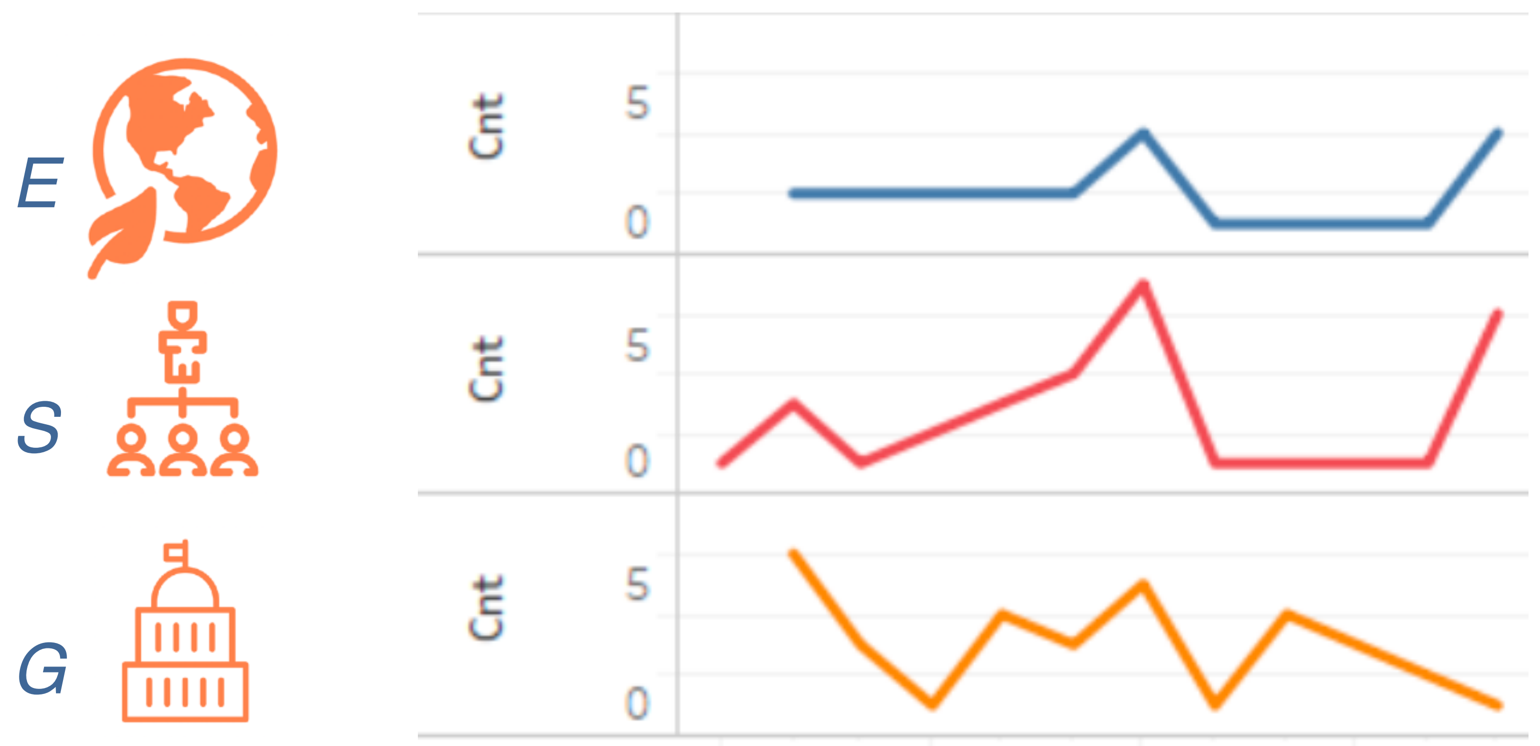


We found that each FI displays different behaviors of board member participations. Some FIs' board discussions are highly dominated by either the chairman or the CEO, while some FIs has high level of participation by all board members.

Results from the analysis of board meeting minutes were incorporated into B&C final report, which were sent to FIs.

Various components, eg. OCR, hmm document parser, word segmentation, were reused in other pilot efforts such as ESG dashboard, textual analysis on historical archives.

ESG dashboard (pilot)



ธนาคารแห่งประเทศไทย
BANK OF THAILAND

2. Government Expenditure Statistics Compilation

BOT routinely compile various statistics of data from other public or private agencies and share with other government and international agencies

Data from the Government Fiscal Management Information System (GFMIS), maintained by the Comptroller General's Department, Ministry of Finance, are among the most challenging ones.

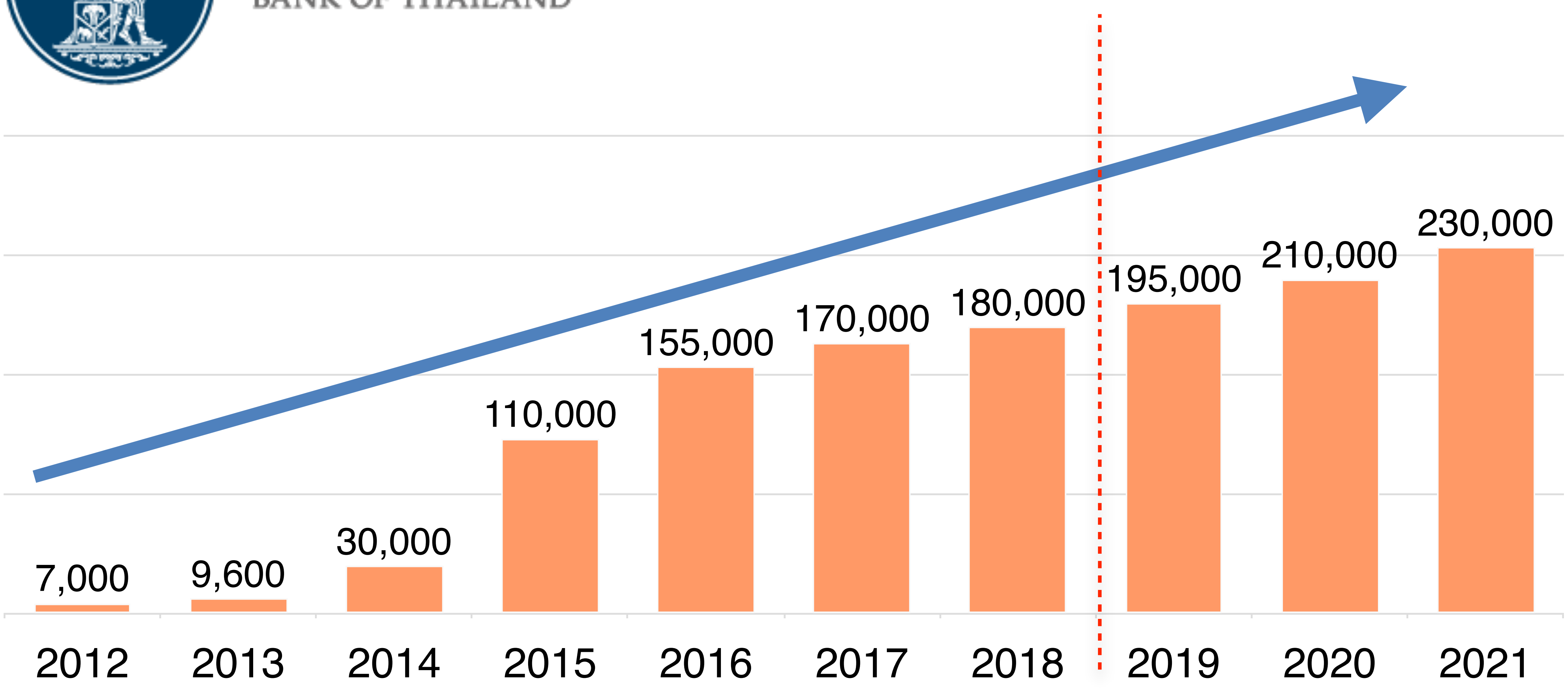
	Govt. Expenditure Item	Agency	Amount (Baht)
1	ค่าตอบแทนสมาชิกกองอาสารักษาดินแดน	กรมการปกครอง	25,000,000
2	ค่าใช้จ่ายการพัฒนาขีดความสามารถด้านข่าวกรองทางทะเล	กองทัพเรือ	215,000,000
3	ค่าปรับปรุงอาคารสำนักงาน กศน. ตำบลในเมือง อำเภอเมืองเพชรบูรณ์ จ. เพชรบูรณ์	สำนักงานปลัดกระทรวงศึกษาธิการ	42,000,000
4	เตียงนอนพร้อมที่นอนผู้รับบริการขนาด 3.5 ฟุต บ้านพักเด็กและครอบครัวจังหวัดอำนาจเจริญ ตำบลโนนหนามแท่ง อำเภอเมืองอำนาจเจริญ จังหวัดอำนาจเจริญ	สำนักงานส่งเสริมสวัสดิภาพ เด็ก ผู้ด้วยโอกาสฯ	500,000

Code GFS	Functional Code
01	General public services
02	Defense
⌄	
09	Education
10	Social Protection

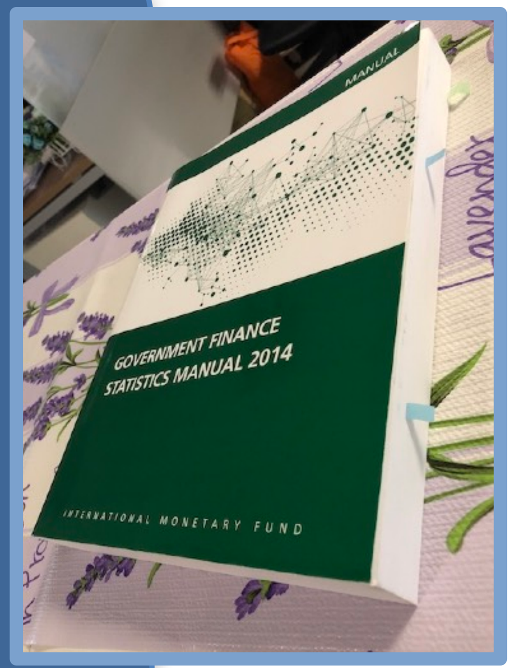
Code GFS	Economic Code
21	Compensation of employees
2111	Wages and salaries in cash
22	Used of goods and services
⌄	
32	Acquisition of financial assets
33	Principal Repayments

ธนาคารแห่งประเทศไทย
BANK OF THAILAND

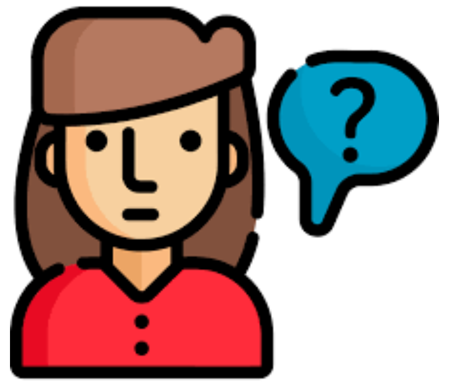
2. Government Expenditure Statistics Compilation



- The number of government expenditure transactions in the system increase significantly over the last 7 years to ~ 230,000 items in the 2021 budgeting year.
- The assignment of Economic Code & Functional Code require domain experts to manually classify each items (also time-consuming).
- The data are not well formatted and do not have proper categorical information, and are in Thai.
- The Government expenditure statistics is also needed to be compiled monthly in the timely-manner to be shared with other govt. agencies.



deployed since 2019



Manual Compilation → **Automate Compilation (ML)**

3 man-days → 30 mins

Accuracy

Economic code 99.5%

Functional code 96.5%

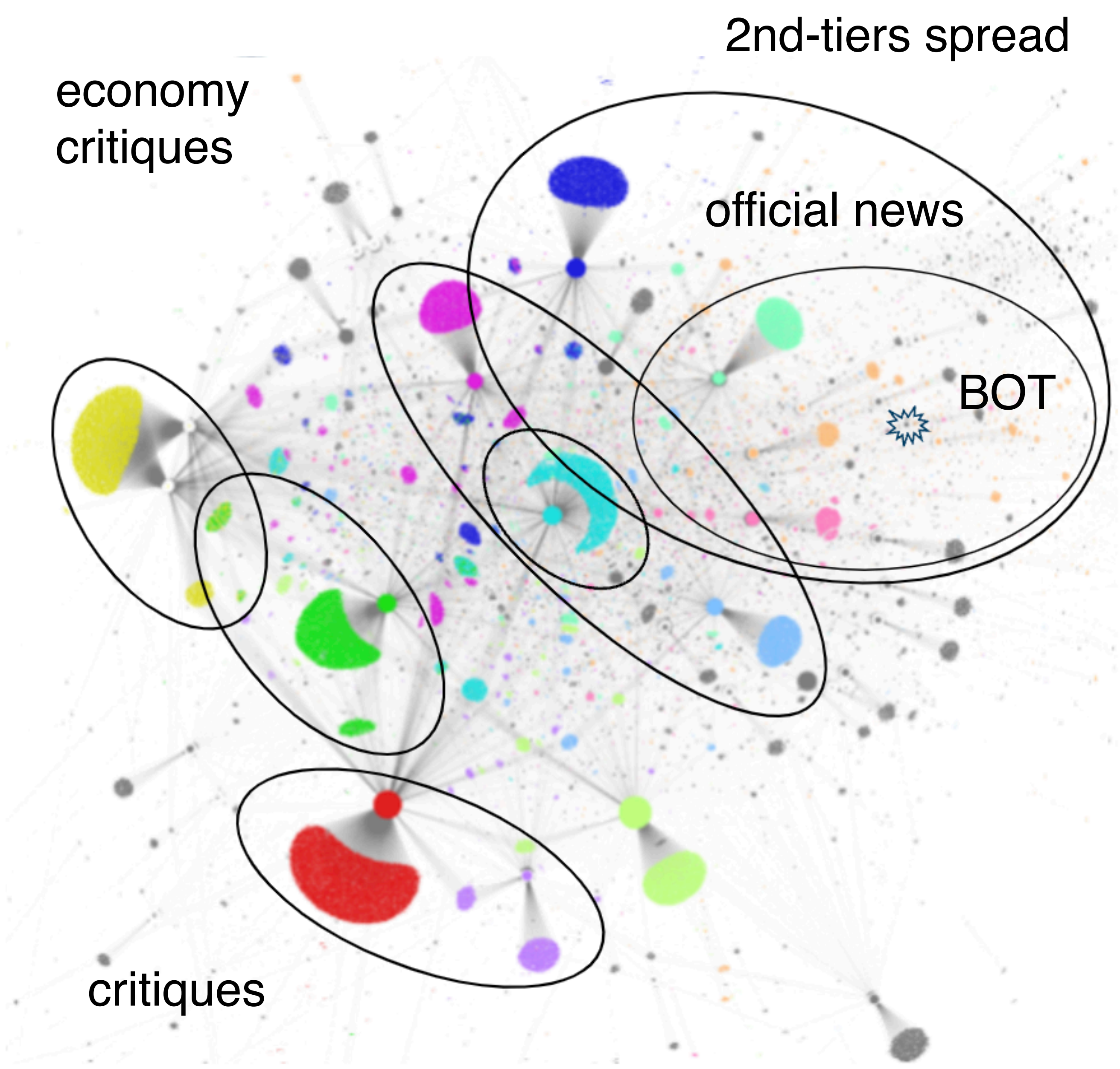
Automation enable more time for further analysis of the government expenditure data (which was not possible before due to the time-constraint) such as analysis of government funding allocation in the local government level.

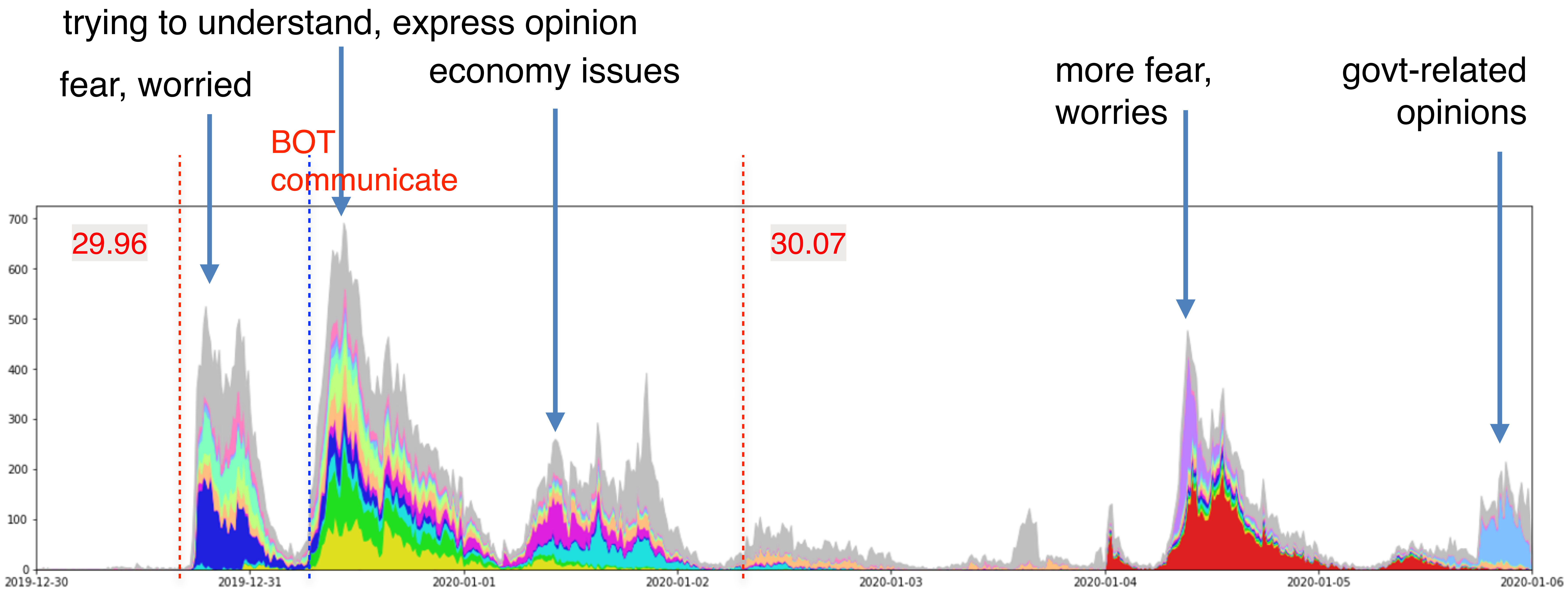
ธนาคารแห่งประเทศไทย
BANK OF THAILAND

3. Social Media Data for Reputational Risk

Monitoring engagement and public opinions on related issues and central bank policy in social media (FX, soft loan policy, counterfeit banknote, etc).

- clustering topics for public opinions monitoring
- monitoring virality of issues
- network analysis for community discovery (interested in neural voices and channel reach)





ธนาคารแห่งประเทศไทย
BANK OF THAILAND

Thank You for you attention

Take Away

- Lots of Potential Use Cases for NLP in Central Bank Operations.
- Use Cases need to be explored collaboratively between Data Analytics Team and Operational Team.
- Long-Term Adoption can be a challenge.
- However, tools developed can still be re-repurposed as building-block for future projects