More Than Words

A Textual Analysis of Central Bank Communication

Reviewer
Wirote Aroonmanakun

Sep 15, 2016
BOT Symposium 2016
Outline

• Data Revolution
• Comments on the paper
• Possibility to do that on Thai texts?
DEFINING THE DATA REVOLUTION

Since the phrase was coined in May 2013 in the report of the High-Level Panel of Eminent Persons on the post-2015 Development Agenda, the “data revolution” has come to mean many things to many people. Here, we take it to mean the following:

The data revolution is:

➤ An explosion in the volume of data, the speed with which data are produced, the number of producers of data, the dissemination of data, and the range of things on which there is data, coming from new technologies such as mobile phones and the “internet of things”, and from other sources, such as qualitative data, citizen-generated data and perceptions data;

➤ A growing demand for data from all parts of society.

The data revolution for sustainable development is:

➤ The integration of these new data with traditional data to produce high-quality information that is more detailed, timely and relevant for many purposes and users, especially to foster and monitor sustainable development;

➤ The increase in the usefulness of data through a much greater degree of openness and transparency, avoiding invasion of privacy and abuse of human rights from misuse of data on individuals and groups, and minimising inequality in production, access to and use of data;

➤ Ultimately, more empowered people, better policies, better decisions and greater participation and accountability, leading to better outcomes for people and the planet.
Data revolution

(Big)Data

machine learning

classification
clustering
association
sequence prediction

Knowledge

Information
Data revolution

(Big)Data

machine learning

classification
clustering
association
sequence prediction

• positive/negative comments
• handwriting recognition
• service/ sale/ product docs

Information

Knowledge
Data revolution

(Big)Data → machine learning → classification

Information

Knowledge

clustering

association

sequence prediction
Data revolution

(Big)Data → Information → Knowledge

machine learning →

classification
clustering
association
sequence
prediction

• discover interesting relations between variables
• beer - diaper
Data revolution

(Big)Data

Information

Knowledge

machine learning

classification

clustering

association

sequence prediction

• shopping history
  =>
  what is next
Data revolution

(Big)Data → Information

Information → Knowledge

Information → number

Information → texts

texts → hidden info.
Text mining

- character sequence
- word sequence
- part of speech
- syntactic structure
- semantics
- meaning in situation
- discourse structure

https://ischool.syr.edu/infospace/2013/04/23/what-is-text-mining/
Comments on paper

• Readability:
  • Flesch-Kincaid (FK) grade level
  • ETS TextEvaluator

• Topics:
  • Topic modeling using LDA

• Tones: positive / negative outlook
  • Dictionary method
ETS TextEvaluator

<table>
<thead>
<tr>
<th>Dimension of Variation/Component Score</th>
<th>Effect on Complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sentence Structure</strong></td>
<td></td>
</tr>
<tr>
<td>Syntactic Complexity</td>
<td>↑</td>
</tr>
<tr>
<td><em>(Increases Complexity)</em></td>
<td></td>
</tr>
<tr>
<td><strong>Vocabulary Difficulty</strong></td>
<td></td>
</tr>
<tr>
<td>Academic Vocabulary</td>
<td>↑</td>
</tr>
<tr>
<td><em>(Increases Complexity)</em></td>
<td></td>
</tr>
<tr>
<td>Word Unfamiliarity</td>
<td>↑</td>
</tr>
<tr>
<td><em>(Increases Complexity)</em></td>
<td></td>
</tr>
<tr>
<td>Concreteness</td>
<td>↓</td>
</tr>
<tr>
<td><em>(Decreases Complexity)</em></td>
<td></td>
</tr>
<tr>
<td><strong>Connections Across Ideas</strong></td>
<td></td>
</tr>
<tr>
<td>Lexical Cohesion</td>
<td>↓</td>
</tr>
<tr>
<td><em>(Decreases Complexity)</em></td>
<td></td>
</tr>
<tr>
<td>Interactive/Conversational Style</td>
<td>↓</td>
</tr>
<tr>
<td><em>(Decreases Complexity)</em></td>
<td></td>
</tr>
<tr>
<td>Level of Argumentation</td>
<td>↑</td>
</tr>
<tr>
<td><em>(Increases Complexity)</em></td>
<td></td>
</tr>
<tr>
<td><strong>Organization</strong></td>
<td></td>
</tr>
<tr>
<td>Degree of Narrativity</td>
<td>↓</td>
</tr>
<tr>
<td><em>(Decreases Complexity)</em></td>
<td></td>
</tr>
</tbody>
</table>
changes in policy statements (in terms of language)
Figure 3.3: Flesch-Kincaid Grade Level Prediction (Avg. 2008-2015)

- Romania: 17.5
- FED: 16.8
- S. Korea: 15.8
- Armenia: 14.6
- Indonesia: 14.5
- ECB: 14.5
- BOJ: 14.2
- Philippines: 14.1
- Thailand: 13.9
- Chile: 13.5
- Canada: 13.3
- Peru: 13.0
- Serbia: 12.8
- Turkey: 12.5
- Hungary: 12.4
- Colombia: 12.3
- Australia: 12.1
- New Zealand: 11.8
- Poland: 11.7
- Czech: 11.1
- Iceland: 10.7
- Norway: 9.8

The median is indicated by the blue arrow.
อย่างไรก็ตาม ค่าเงินบาทในช่วงที่ผ่านมา ในนั้นมีค่ารั้นขึ้นเมื่อเทียบกับเงินสกุลหลักต่างประเทศ และอาจไม่เป็นผลดีต่อการฟื้นตัวของเศรษฐกิจไทย

นอกจากนี้ คณะกรรมการฯ เห็นว่าการรักษาขีดความสามารถในการดำเนินนโยบาย (policy space) ยังมีความจำเป็น หากสภาวะการเงินในระยะต่อไปส่งผลกระทบต่อเศรษฐกิจไทย ความเสี่ยงต่อการดำเนินนโยบายได้เพิ่มขึ้น โดยเฉพาะการเพิ่มขึ้นของสภาพแวดล้อมทางการเงินที่ต้องเผชิญกับความไม่แน่นอนทั้งการเงิน และความมั่นคง นโยบายทางการดำเนินนโยบายการเงินของประเทศ อุตสาหกรรมหลักที่ส่งผลให้เงินลงทุนเคลื่อนย้ายและอัตราแลกเปลี่ยนผันผวนมากขึ้น

ในระยะต่อไป คณะกรรมการฯ เห็นว่านโยบายการเงินควรอยู่ในระดับผ่อนปรนต่อเนื่อง และพร้อมที่จะใช้เครื่องมือการเงินที่มีอยู่อย่างเหมาะสม เพื่อให้การเงินโดยรวมเอื้อต่อการฟื้นตัวของเศรษฐกิจ ควบคู่กับการรักษาเสถียรภาพการเงินของประเทศ

 nonprofit, the baht appreciated against some major currencies over the recent period, which might not be beneficial to the ongoing economic recovery.

In addition, financial stability risks from the prolonged low interest rate environment, including search-for-yield behavior, continued to warrant close monitoring.

The Committee saw the need to preserve policy space given that the Thai economy would still be facing more uncertainties going forward, particularly the fragile global economic recovery, and uncertainties in the monetary policy directions of major advanced economies that might induce greater capital flow and exchange rate volatility.

Going forward, the Committee viewed that monetary policy should remain accommodative, and stands ready to utilize an appropriate mix of available policy tools in order to ensure that monetary conditions are conducive to the economic recovery, while ensuring financial stability.
Figure 3.3: Flesch-Kincaid Grade Level Prediction (Avg. 2008-2015)

- Romania: 17.5
- FED: 16.8
- S. Korea: 15.8
- Armenia: 14.6
- Indonesia: 14.5
- ECB: 14.5
- BOJ: 14.2
- Philippines: 14.1
- Thailand: 13.9
- Chile: 13.5
- Canada: 13.3
- Peru: 13.0
- Serbia: 12.8
- Turkey: 12.5
- Hungary: 12.4
- Colombia: 12.3
- Australia: 12.1
- New Zealand: 11.8
- Poland: 11.7
- Czech: 11.1
- Iceland: 10.7
- Norway: 9.8

Median: 13.9

Legend:
- Dark blue: Flesch-Kincaid Grade Level
- Orange: # Words Per Sentence
<table>
<thead>
<tr>
<th><strong>FOMC</strong></th>
<th><strong>BOT</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>The Federal Reserve Board on Thursday released a policy statement detailing the framework the Board will follow in setting the Countercyclical Capital Buffer (CCyB) for private-sector credit exposures located in the United States.</td>
<td>Mr. Jaturong Jantarangs, Secretary of the Monetary Policy Committee (MPC), announced the outcome of the meeting on 22 June 2016 as follows.</td>
</tr>
<tr>
<td>The CCyB is a macroprudential tool that can be used to increase the resilience of the financial system by raising capital requirements on internationally active banking organizations when the risk of above-normal losses is elevated.</td>
<td>The Committee voted unanimously to maintain the policy rate at 1.50 percent. One MPC member was unable to attend this meeting.</td>
</tr>
<tr>
<td>The CCyB would then be available to help banking organizations absorb shocks associated with declining credit conditions.</td>
<td>In deliberating monetary policy, the Committee judged that the economy would continue to recover, and inflation would return to the target band within the latter half of the year as previously expected.</td>
</tr>
<tr>
<td>Implementation of the buffer could also help moderate fluctuations in the supply of credit.</td>
<td>Meanwhile, monetary conditions remained accommodative and conducive to the economic recovery.</td>
</tr>
<tr>
<td>The policy statement provides background on the range of financial-system vulnerabilities and other factors the Board may take into account as it evaluates settings for the buffer, including but not limited to, leverage in the nonfinancial sector, leverage in the financial sector, maturity and liquidity transformation in the financial sector, and asset valuation pressures.</td>
<td>Hence, the policy rate should be kept on hold at this meeting.</td>
</tr>
<tr>
<td>FOMC</td>
<td>BOT</td>
</tr>
<tr>
<td>--------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>Because economic and financial risks are constantly evolving, the</td>
<td>The Thai economy continued to recover on the back of public</td>
</tr>
<tr>
<td>range of indicators and models that the Board may consider is likely</td>
<td>expenditure and broad-based improvement in the tourism sector, while</td>
</tr>
<tr>
<td>to change over time.</td>
<td>private consumption expanded as anticipated.</td>
</tr>
<tr>
<td></td>
<td>**The Thai economy continued to recover on the back of public</td>
</tr>
<tr>
<td></td>
<td>expenditure and broad-based improvement in the tourism sector, while</td>
</tr>
<tr>
<td></td>
<td>private consumption expanded as anticipated.</td>
</tr>
<tr>
<td>In response to comments, the final policy statement clarifies</td>
<td>However, expansion in private investment remained low, and</td>
</tr>
<tr>
<td>that the Board expects that the CCyB will be activated when</td>
<td>merchandise exports still contracted in tandem with other Asian</td>
</tr>
<tr>
<td>systemic vulnerabilities are meaningfully above normal and that</td>
<td>economies which were weaker than expected.</td>
</tr>
<tr>
<td>the Board generally intends to increase the CCyB gradually.</td>
<td>Overall, the economic momentum from domestic demand and the tourism</td>
</tr>
<tr>
<td></td>
<td>sector offset the impact from contraction in merchandise exports.</td>
</tr>
<tr>
<td>The final policy statement also emphasizes that the Board</td>
<td>As such, the economy was expected to expand at the same rate as</td>
</tr>
<tr>
<td>expects to remove or reduce the CCyB when the conditions that led</td>
<td>previously assessed.</td>
</tr>
<tr>
<td>to its activation abate or lessen and when the release of CCyB</td>
<td>Nevertheless, there are remaining downside risks from slower-than-</td>
</tr>
<tr>
<td>capital would promote financial stability.</td>
<td>expected growth of trading partners, and private sector’s confidence</td>
</tr>
<tr>
<td></td>
<td>which is still fragile.</td>
</tr>
<tr>
<td>The Board generally would expect to provide notice to the public</td>
<td>Nevertheless, there are remaining downside risks from slower-than-</td>
</tr>
<tr>
<td>and seek comment on the proposed level of the CCyB as part of</td>
<td>expected growth of trading partners, and private sector’s confidence</td>
</tr>
<tr>
<td>making any final determination to change the CCyB.</td>
<td>which is still fragile.</td>
</tr>
<tr>
<td></td>
<td>However, concerns pertaining to the impact of the drought have</td>
</tr>
<tr>
<td></td>
<td>subsided, and prices of some agricultural commodities have shown</td>
</tr>
<tr>
<td></td>
<td>sign of recovery.</td>
</tr>
<tr>
<td>The CCyB applies to banking organizations that are subject to the</td>
<td>Nevertheless, there are remaining downside risks from slower-than-</td>
</tr>
<tr>
<td>advanced approaches capital rules, generally those with more than</td>
<td>expected growth of trading partners, and private sector’s confidence</td>
</tr>
<tr>
<td>$250 billion in assets or $10 billion in on-balance-sheet foreign</td>
<td>which is still fragile.</td>
</tr>
<tr>
<td>exposures, and to any depository institution subsidiary of such</td>
<td>However, concerns pertaining to the impact of the drought have</td>
</tr>
<tr>
<td>banking organizations.</td>
<td>subsided, and prices of some agricultural commodities have shown</td>
</tr>
<tr>
<td></td>
<td>sign of recovery.</td>
</tr>
</tbody>
</table>
Result of Readability

• Readability tends to fall when central banks lower their policy interest rates

• What is the implication?
Topic modeling
Topic modeling

- Preprocessing: remove stop words, stemming, remove punctuations, remove numbers, etc.
- Use Latent Dirichlet Allocation (LDA)
Topic modeling

<table>
<thead>
<tr>
<th>Topic assignment</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inflation</td>
<td>topic0</td>
<td>remain</td>
<td>pressure</td>
<td>low</td>
<td>recent</td>
<td>moder</td>
<td>data</td>
<td>continu</td>
<td>growth</td>
<td>inflationari</td>
</tr>
<tr>
<td></td>
<td>0.083</td>
<td>0.06</td>
<td>0.057</td>
<td>0.054</td>
<td>0.048</td>
<td>0.033</td>
<td>0.032</td>
<td>0.031</td>
<td>0.027</td>
<td>0.024</td>
</tr>
<tr>
<td>Labor market</td>
<td>topic1</td>
<td>increas</td>
<td>wage</td>
<td>product</td>
<td>growth</td>
<td>labour</td>
<td>employ</td>
<td>cost</td>
<td>result</td>
<td>adjust</td>
</tr>
<tr>
<td></td>
<td>0.091</td>
<td>0.041</td>
<td>0.037</td>
<td>0.034</td>
<td>0.033</td>
<td>0.031</td>
<td>0.028</td>
<td>0.023</td>
<td>0.021</td>
<td>0.02</td>
</tr>
<tr>
<td>Domestic demand</td>
<td>topic2</td>
<td>growth</td>
<td>demand</td>
<td>domest</td>
<td>sector</td>
<td>household</td>
<td>continu</td>
<td>credit</td>
<td>invest</td>
<td>loan</td>
</tr>
<tr>
<td></td>
<td>0.088</td>
<td>0.065</td>
<td>0.056</td>
<td>0.048</td>
<td>0.04</td>
<td>0.039</td>
<td>0.036</td>
<td>0.034</td>
<td>0.032</td>
<td>0.028</td>
</tr>
<tr>
<td>Economic condition</td>
<td>topic3</td>
<td>econom</td>
<td>committe</td>
<td>condit</td>
<td>inform</td>
<td>support</td>
<td>time</td>
<td>improv</td>
<td>current</td>
<td>assess</td>
</tr>
<tr>
<td></td>
<td>0.091</td>
<td>0.055</td>
<td>0.046</td>
<td>0.032</td>
<td>0.031</td>
<td>0.03</td>
<td>0.028</td>
<td>0.027</td>
<td>0.023</td>
<td>0.021</td>
</tr>
<tr>
<td>Oil prices</td>
<td>topic4</td>
<td>price</td>
<td>increas</td>
<td>oil</td>
<td>consum</td>
<td>rise</td>
<td>food</td>
<td>commod</td>
<td>mainli</td>
<td>energi</td>
</tr>
<tr>
<td></td>
<td>0.274</td>
<td>0.076</td>
<td>0.046</td>
<td>0.039</td>
<td>0.036</td>
<td>0.03</td>
<td>0.025</td>
<td>0.023</td>
<td>0.02</td>
<td>0.018</td>
</tr>
<tr>
<td>Risk outlook</td>
<td>topic5</td>
<td>risk</td>
<td>outlook</td>
<td>factor</td>
<td>current</td>
<td>effect</td>
<td>develop</td>
<td>balanc</td>
<td>account</td>
<td>chang</td>
</tr>
<tr>
<td></td>
<td>0.097</td>
<td>0.071</td>
<td>0.038</td>
<td>0.032</td>
<td>0.03</td>
<td>0.029</td>
<td>0.028</td>
<td>0.028</td>
<td>0.027</td>
<td>0.024</td>
</tr>
<tr>
<td>Financial market</td>
<td>topic6</td>
<td>market</td>
<td>financi</td>
<td>exchang</td>
<td>intern</td>
<td>us</td>
<td>includ</td>
<td>dollar</td>
<td>rand</td>
<td>volatil</td>
</tr>
<tr>
<td></td>
<td>0.128</td>
<td>0.087</td>
<td>0.042</td>
<td>0.036</td>
<td>0.034</td>
<td>0.025</td>
<td>0.018</td>
<td>0.017</td>
<td>0.017</td>
<td>0.017</td>
</tr>
</tbody>
</table>

Topic 1: remain, pressure, low, recent, modern, data, continue, growth, inflationary, confirm  
= inflation
Topic modeling

- Need specialist to interpret
- Identify topic name
- Group related topics in a hierarchical structure

[Economic growth: T9, T2, T3, T10]
Topic modeling

Figure 4.5: Average of estimated topic proportions during 2001-2015
Tones of topics

- Positive/ negative word list used in Hansen and McMahon (2015)

**Table 1: Example of Contraction and Expansion Words**

<table>
<thead>
<tr>
<th>Contraction</th>
<th>Expansion</th>
</tr>
</thead>
<tbody>
<tr>
<td>decreases*</td>
<td>increases*</td>
</tr>
<tr>
<td>decelerate*</td>
<td>accelerate*</td>
</tr>
<tr>
<td>slow*</td>
<td>fast*</td>
</tr>
<tr>
<td>weak*</td>
<td>strong*</td>
</tr>
<tr>
<td>low*</td>
<td>high*</td>
</tr>
<tr>
<td>loss*</td>
<td>gain*</td>
</tr>
<tr>
<td>contract*</td>
<td>expand*</td>
</tr>
</tbody>
</table>
Results of topics and tones

- Proportion of topics & tones ~ macroeconomic variable i.e. economic growth, inflation and policy interest rate
  - FOMC: estimated growth topic proportion is related negatively to GDP growth and positively to unemployment rate
  - ECB: both estimated net tones of inflation and growth in the ECB policy statements are more informative when compared with the FOMC statements.
  - BOT: Monetary Policy Committee has tendency to talk more about growth in the statement when the observed GDP growth declines
- .....
Can we do that on Thai texts?
Thai Text Mining
Is it possibile?

• Basic assumptions of the tools
  • Bag of words
  • Financial stability risks in China remained
  • China stability remains risky in financial
  • Stability risks remain in Chinese finance
  • Financial stability risk to remain in China
  • *Risk remain financial stability China
Thai Text Mining
Is it possible?

- Nature of Thai language
- No word, sentence boundary
- Vague / ambiguity
  - มีคนขับรถมาอยู่หน้าบ้าน
  - มีคนขับรถมารออยู่หน้าบ้าน
  - มีคนขับรถมาจอดอยู่หน้าบ้าน
  - มีคนขับรถมาจอดอยู่หน้าบ้าน
มี คนขับรถ มา รถ อยู่ หน้า บ้าน
มี คนขับรถ มา รถ อยู่ หน้า บ้าน
Thai Text Mining
Is it possible?

- Nature of Thai language
  - No word, sentence boundary
  - Vague / ambiguity
  - Polysemy: one word has many senses
  - Serial verb construction
  - Proper names, transliterated words
  - Discourse oriented: zero pronoun, ellipsis
Thai Text Mining
Is it possible?

• Bag of words: OK to use?

• Polysemy: => one sense per domain
  •ทรงตัว
    • อัตราเงินเฟ้อพื้นฐานทรงตัวstable
    • พอร์ตชีบยืนก็ทรงตัวไม่อยู่balance

• Topic modeling: identify only word co-occurrence
  • price, increase, oil, consume, rise, food, commodity, mainly, energy, decline, …

• identify thematic roles by users
Thai Text Mining
Is it possible?

- Monetary Policy Committee Press release:
  - sublanguage
  - formal style
  - very few pronoun
  - know about the language and texts
  - know how to interpret the results
Thai Text Mining
Is it possible?

- X who understands the subject field
- X who understands tools and statistical analysis
- X who knows how to choose useful Thai linguistic features
- X who can explain why the outcome is like that
- => It is collaborative work