

# 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:

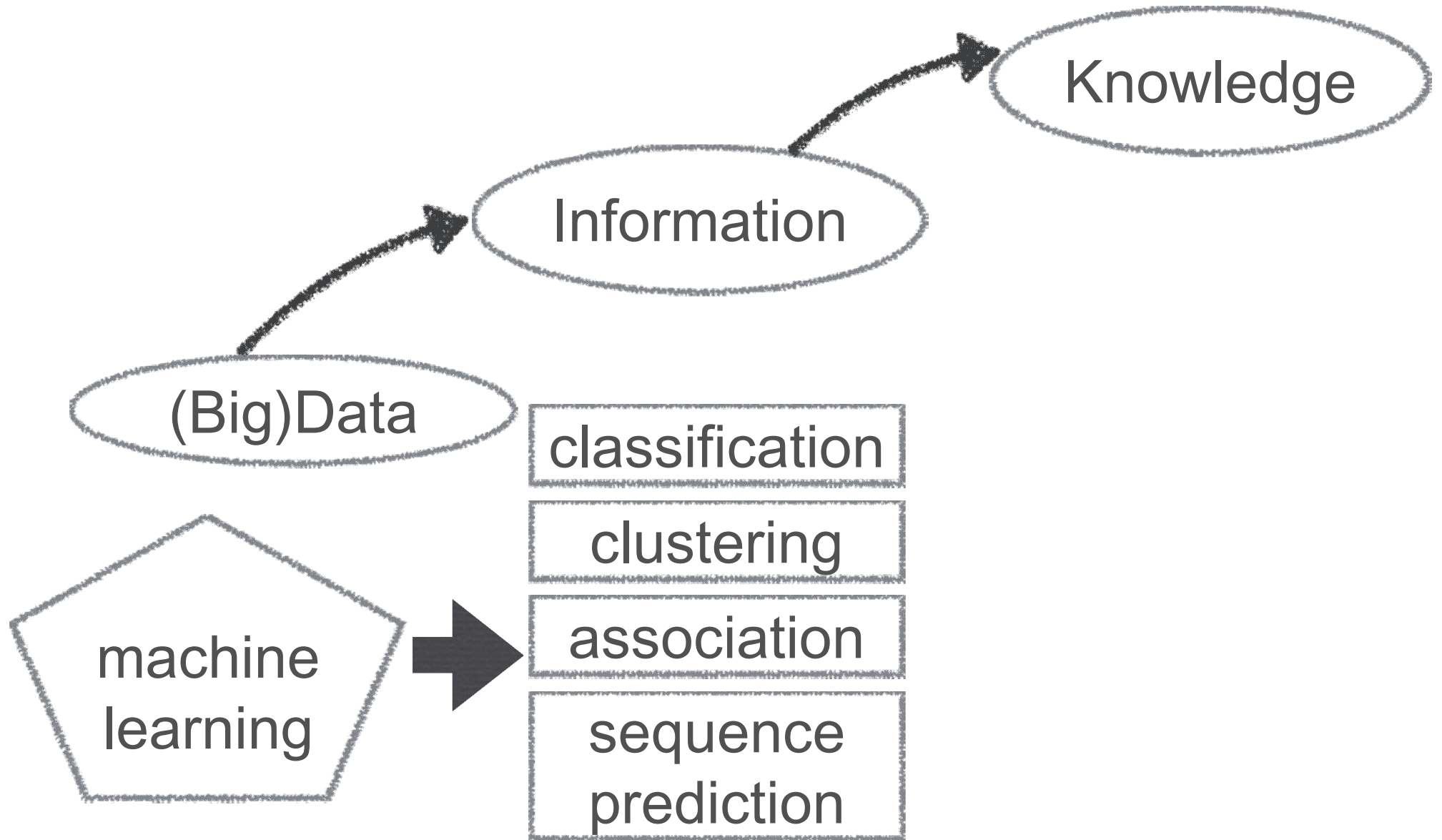
- » An explosion in the volume of data, the speed with which data are produced, the number of producers of data, the dissemination of data, and the range of things on which there is data, coming from new technologies such as mobile phones and the “internet of things”, and from other sources, such as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

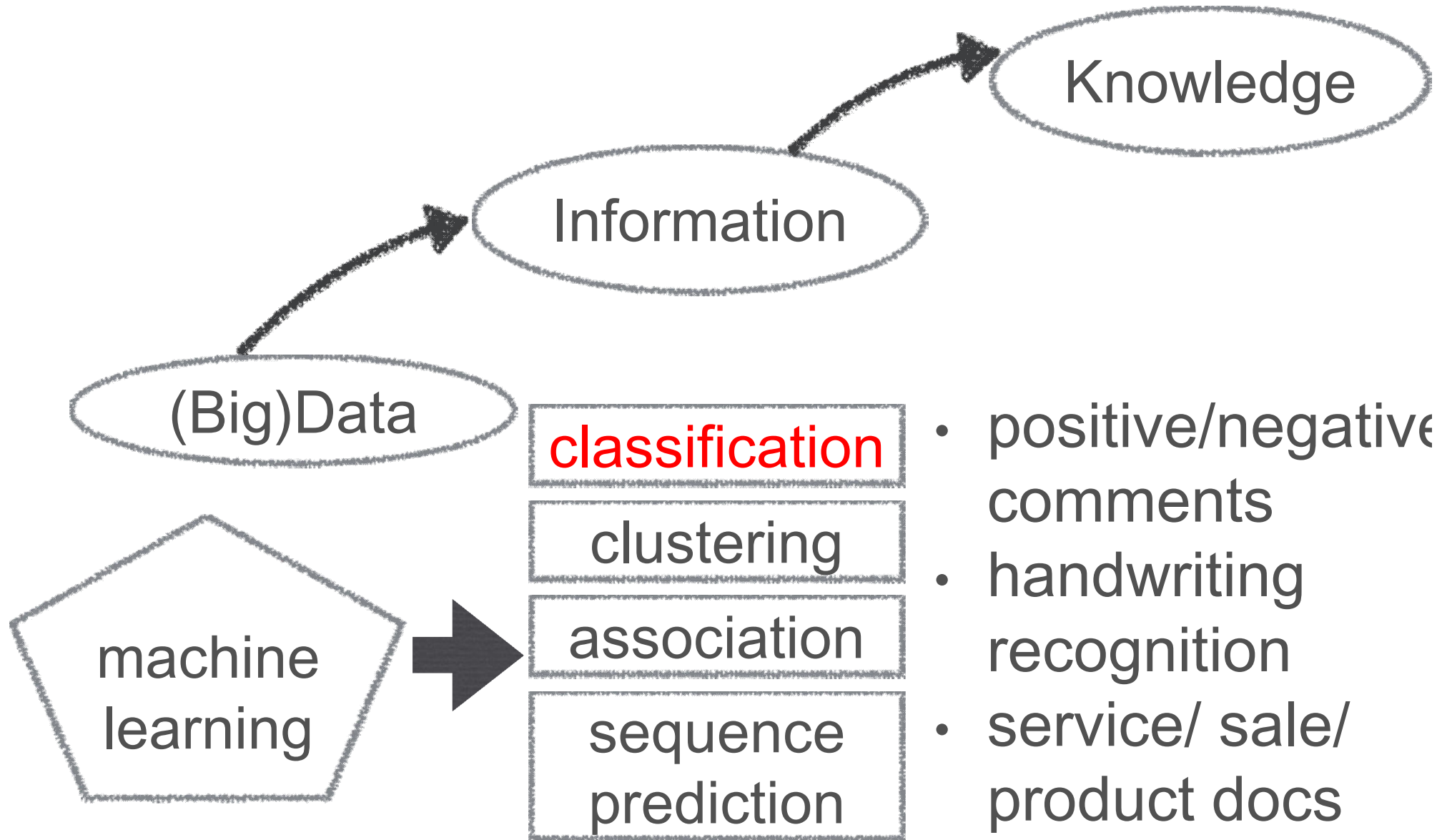
- » The integration of these new data with traditional data to produce high-quality information that is more detailed, timely and relevant for many purposes and users, especially to foster and monitor sustainable development;
- » The increase in the usefulness of data through a much greater degree of 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.

<http://www.undatarevolution.org/report/>

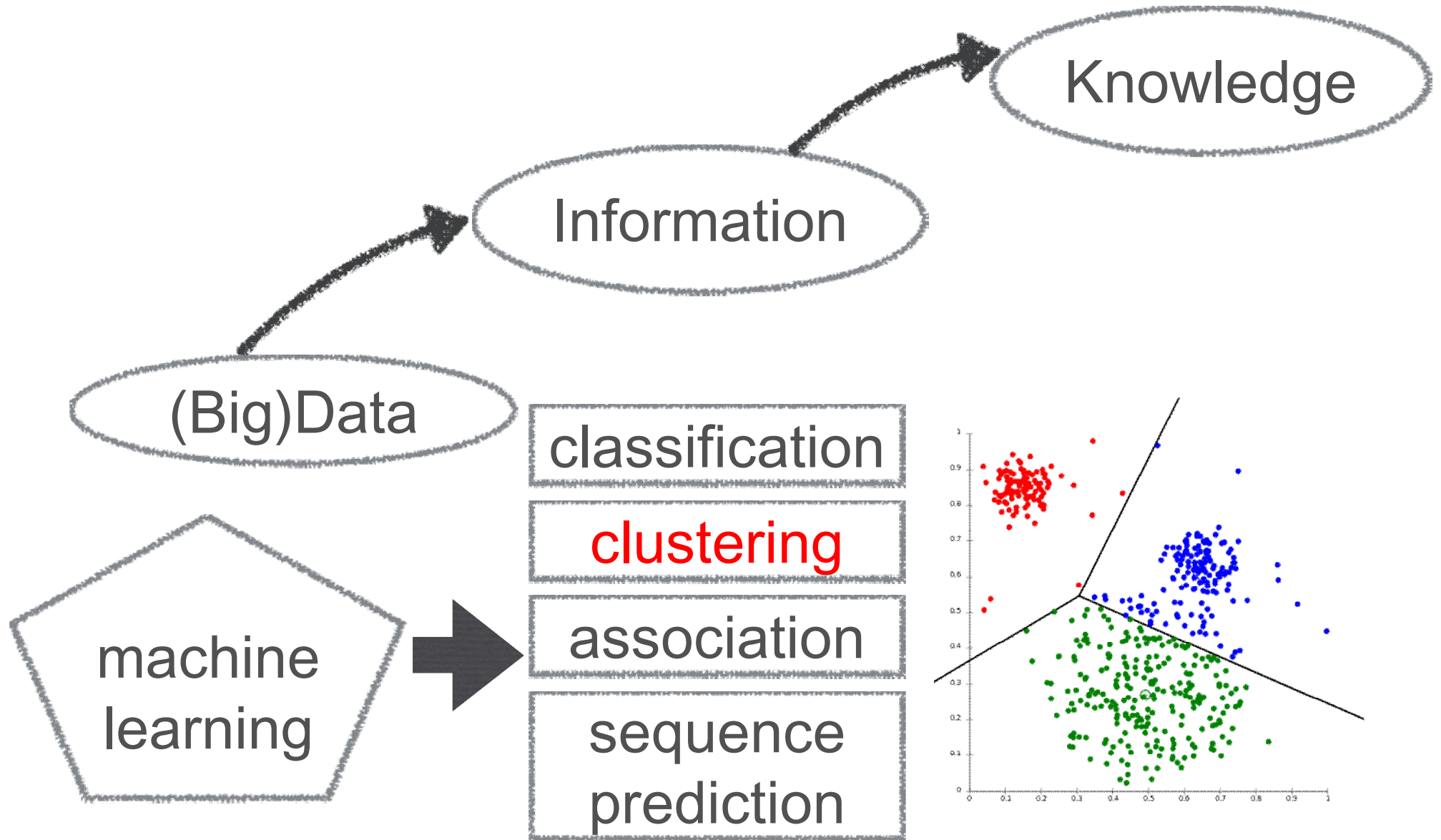
# Data revolution



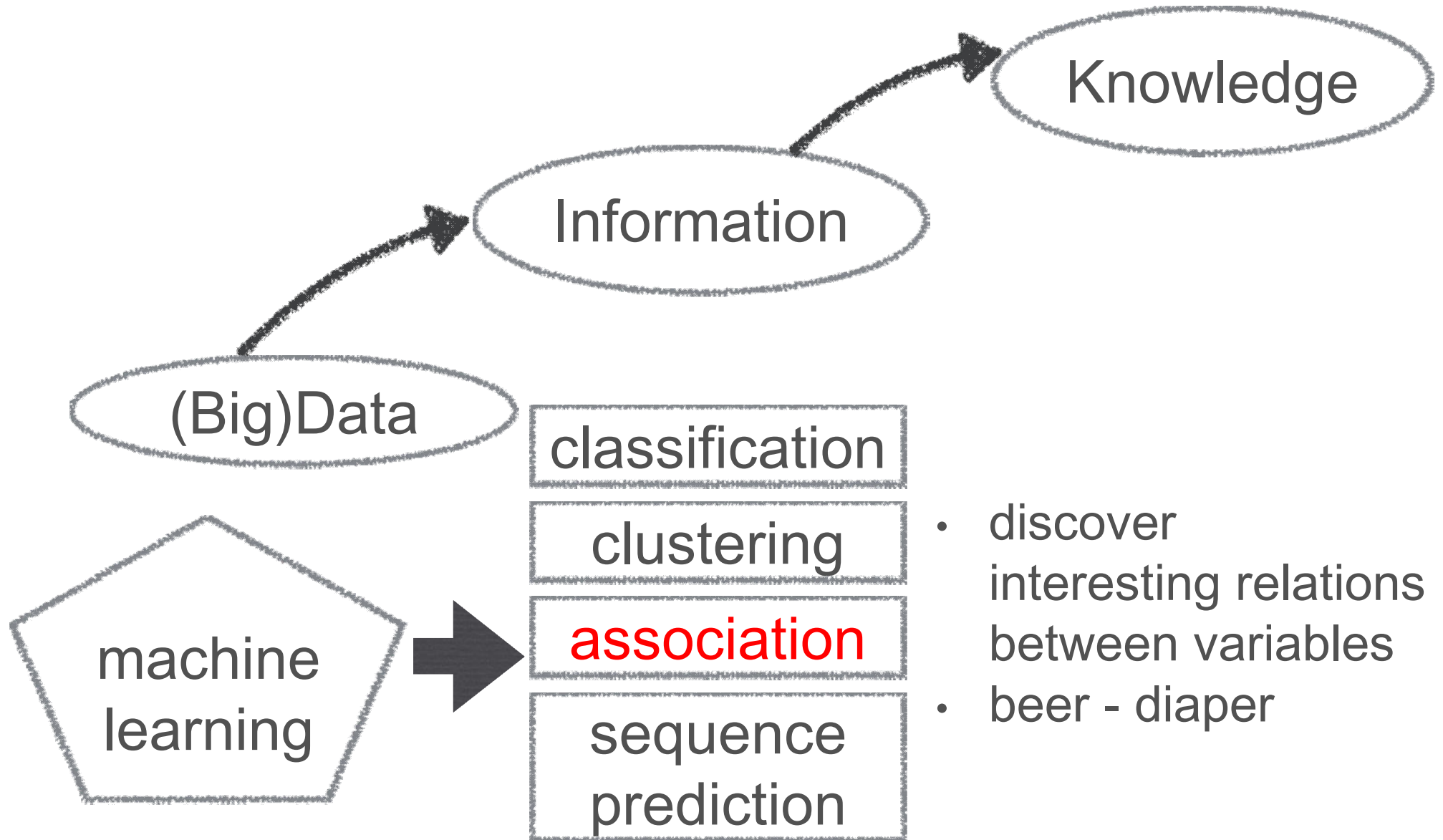
# Data revolution



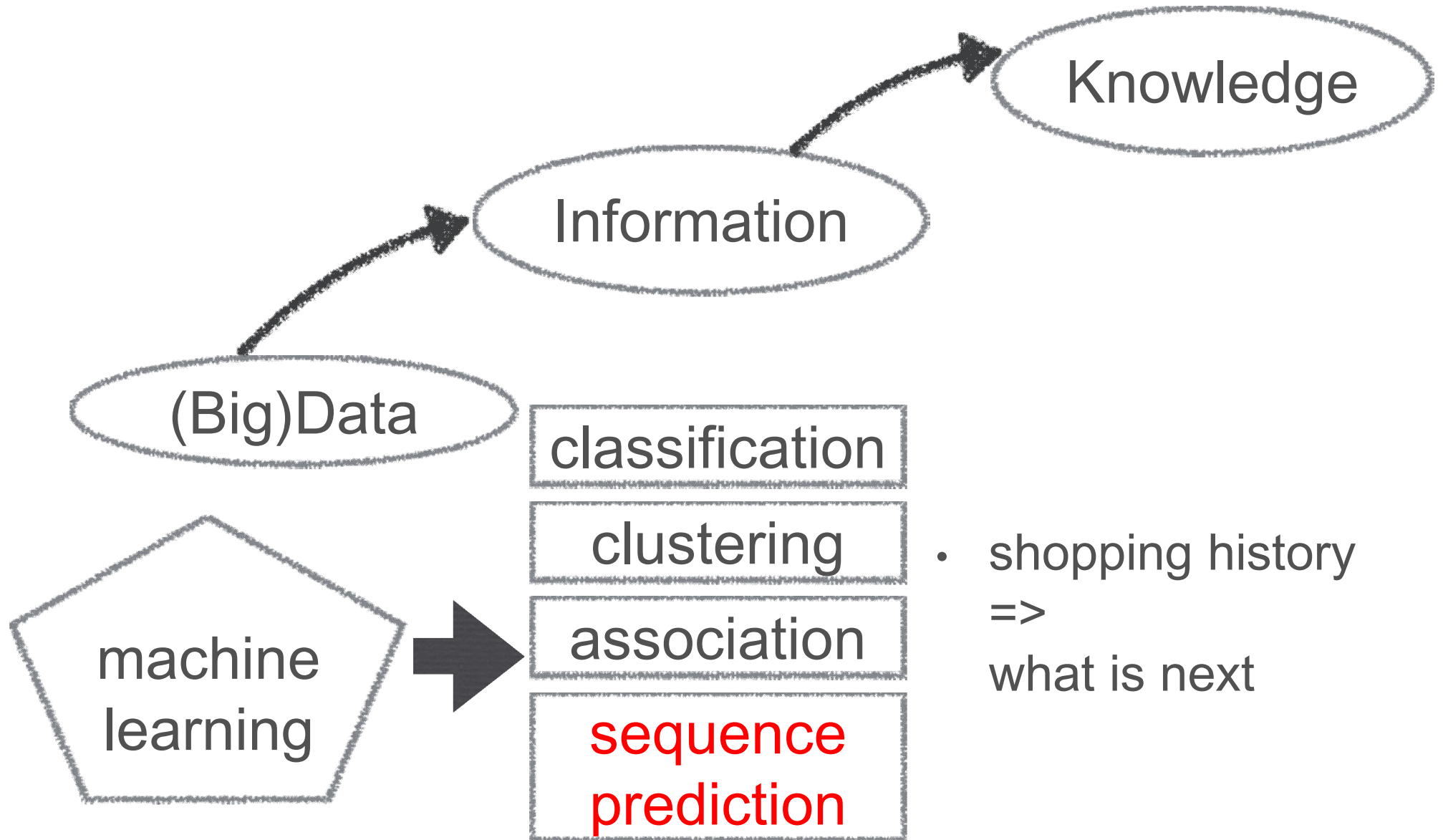
# Data revolution



# Data revolution

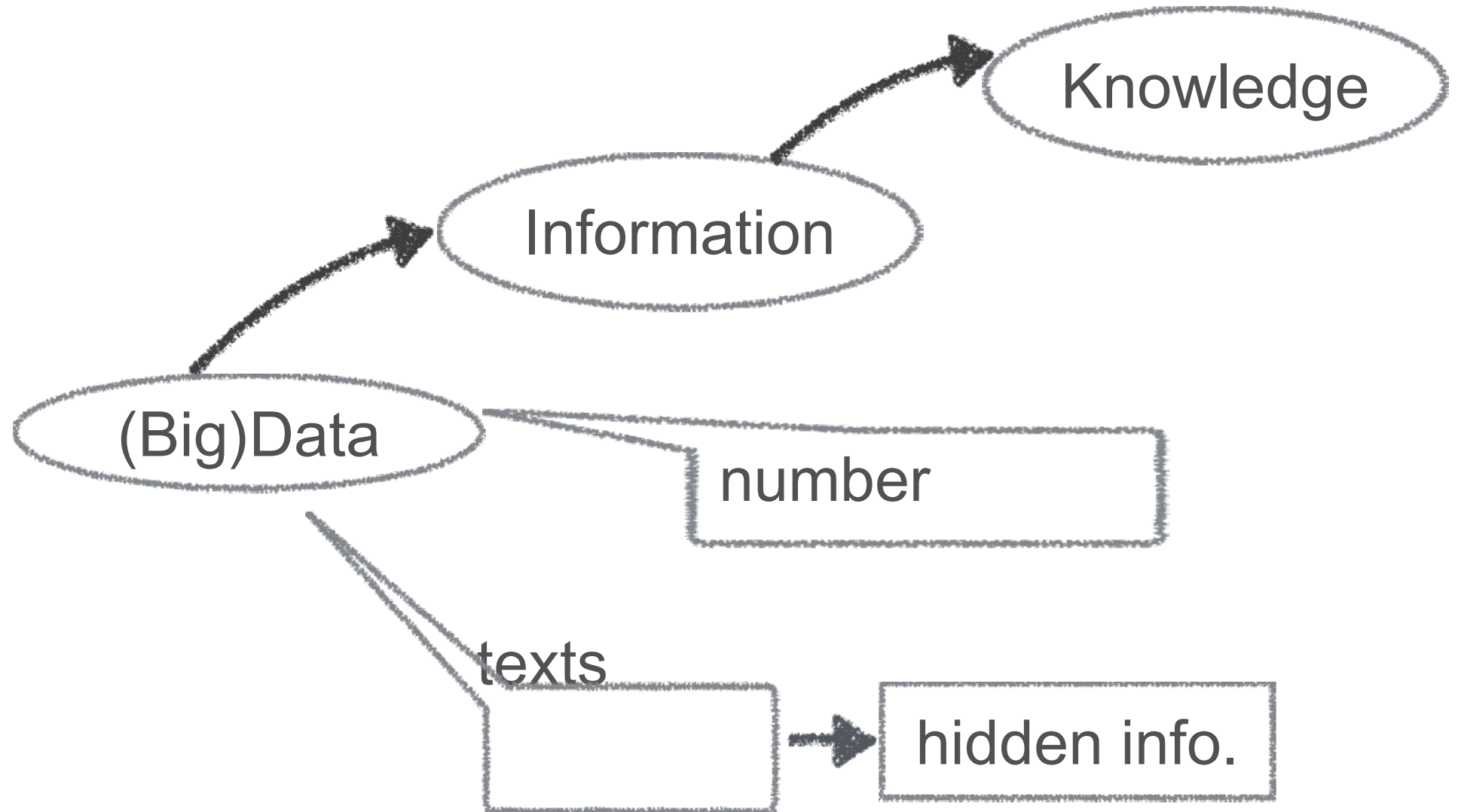


# Data revolution





# Data revolution



# Text mining



- character sequence
  - word sequence
- 

- part of speech
- syntactic structure
- semantics
- meaning in situation
- discourse structure

# Comments on paper

- Readability :
  - Flesch-Kincaid (FK) grade level
  - ETS TextEvaluator
- Topics :
  - Topic modeling using LDA
- Tones : positive / negative outlook
  - Dictionary method

# ETS TextEvaluator

Dimension of Variation/Component Score	Effect on Complexity
<b>Sentence Structure</b>	
<b>Syntactic Complexity</b> <i>(Increases Complexity)</i>	↑
<b>Vocabulary Difficulty</b>	
<b>Academic Vocabulary</b> <i>(Increases Complexity)</i>	↑
<b>Word Unfamiliarity</b> <i>(Increases Complexity)</i>	↑
<b>Concreteness</b> <i>(Decreases Complexity)</i>	↓
<b>Connections Across Ideas</b>	
<b>Lexical Cohesion</b> <i>(Decreases Complexity)</i>	↓
<b>Interactive/Conversational Style</b> <i>(Decreases Complexity)</i>	↓
<b>Level of Argumentation</b> <i>(Increases Complexity)</i>	↑
<b>Organization</b>	
<b>Degree of Narrativity</b> <i>(Decreases Complexity)</i>	↓

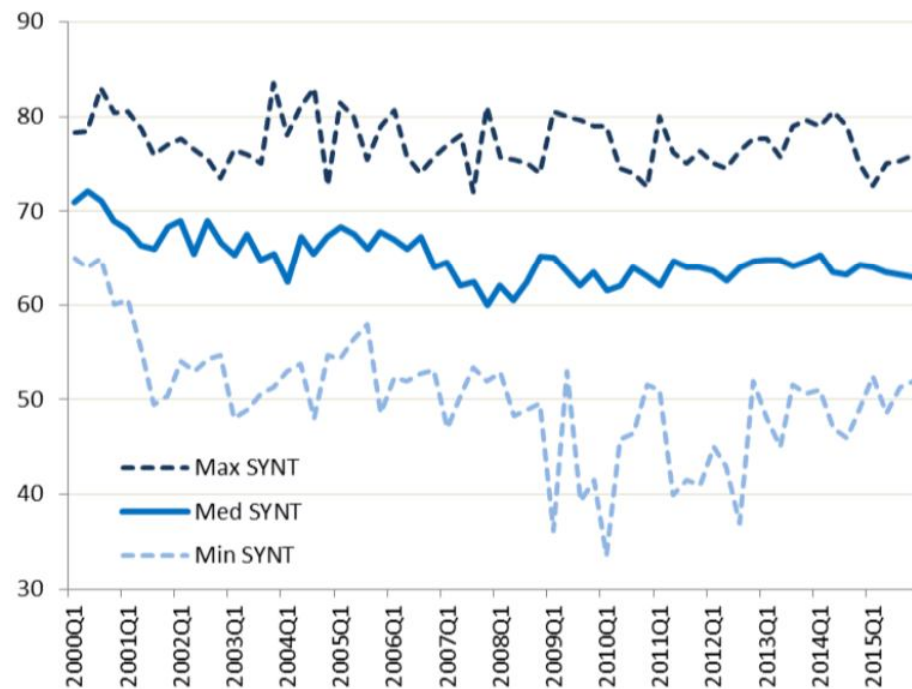
\*\*\*

\*\*\*

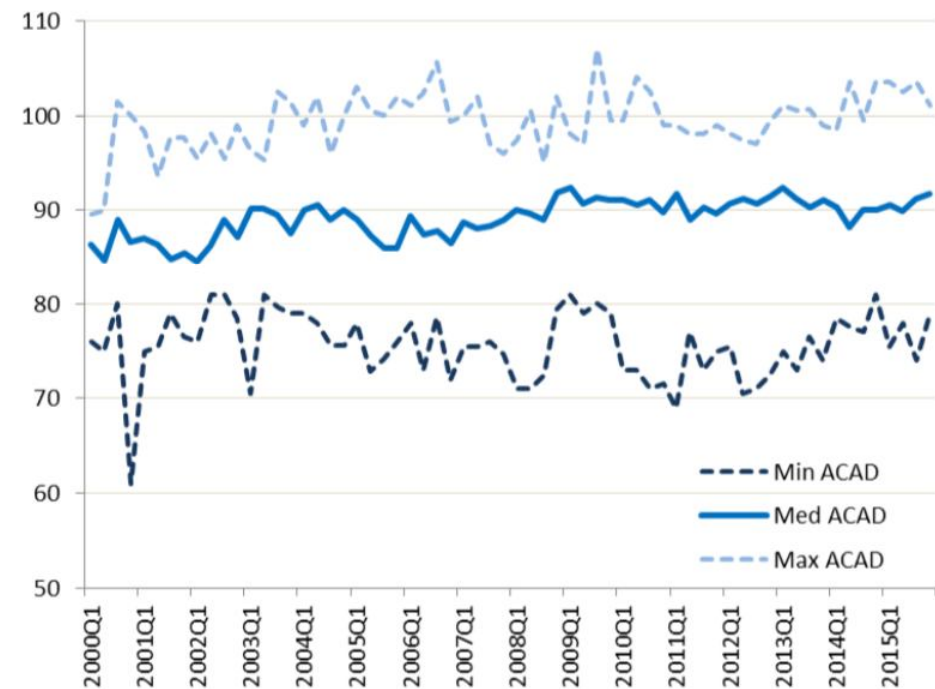
# ETS TextEvaluator

Figure 3.4: Contributors to ETS Complexity Scores

a) *Syntactic complexity score*

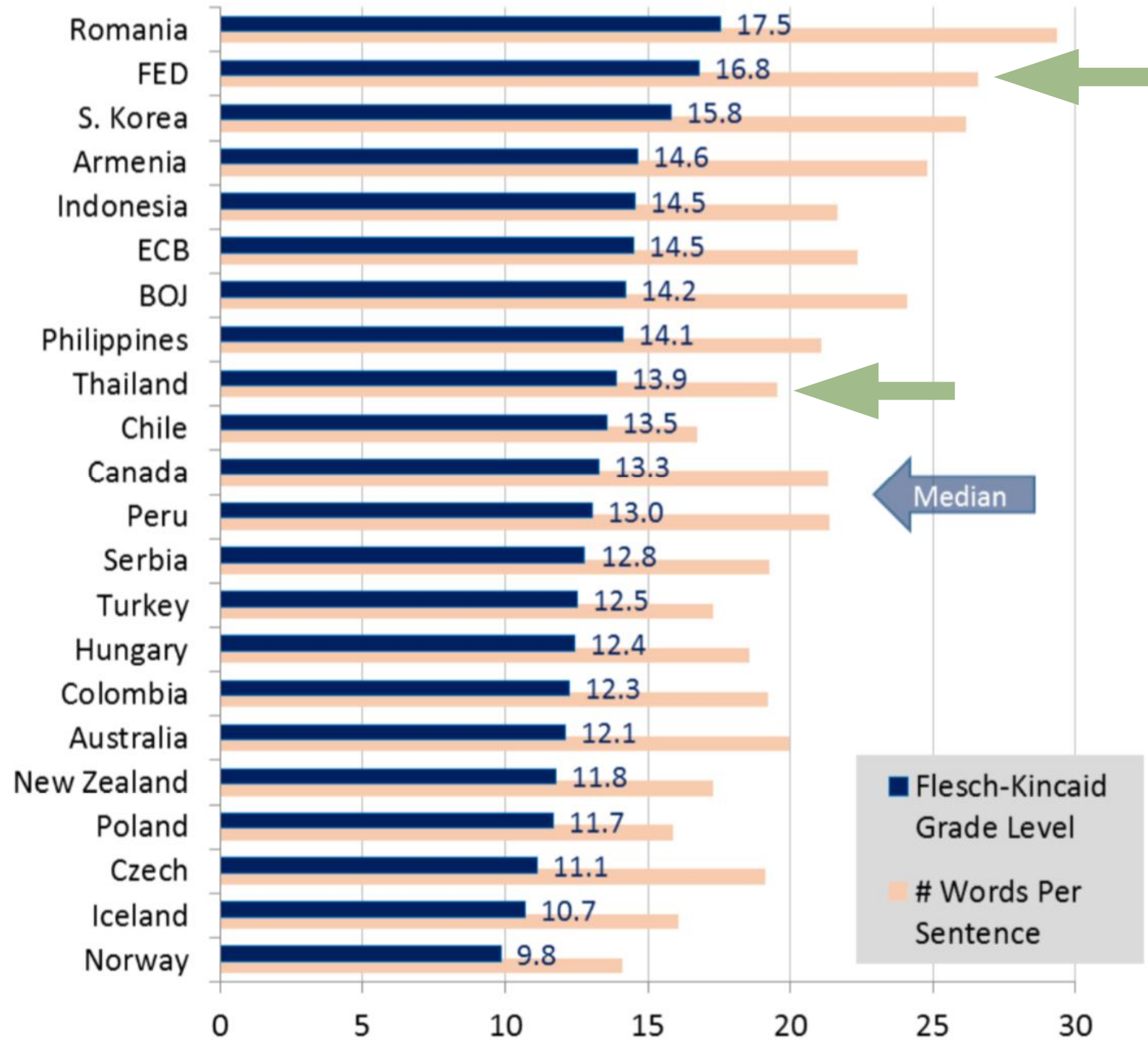


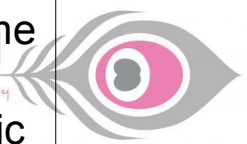
b) *Academic words score*



changes in policy statements (in terms of language)

Figure 3.3: Flesch-Kincaid Grade Level Prediction (Avg. 2008-2015)

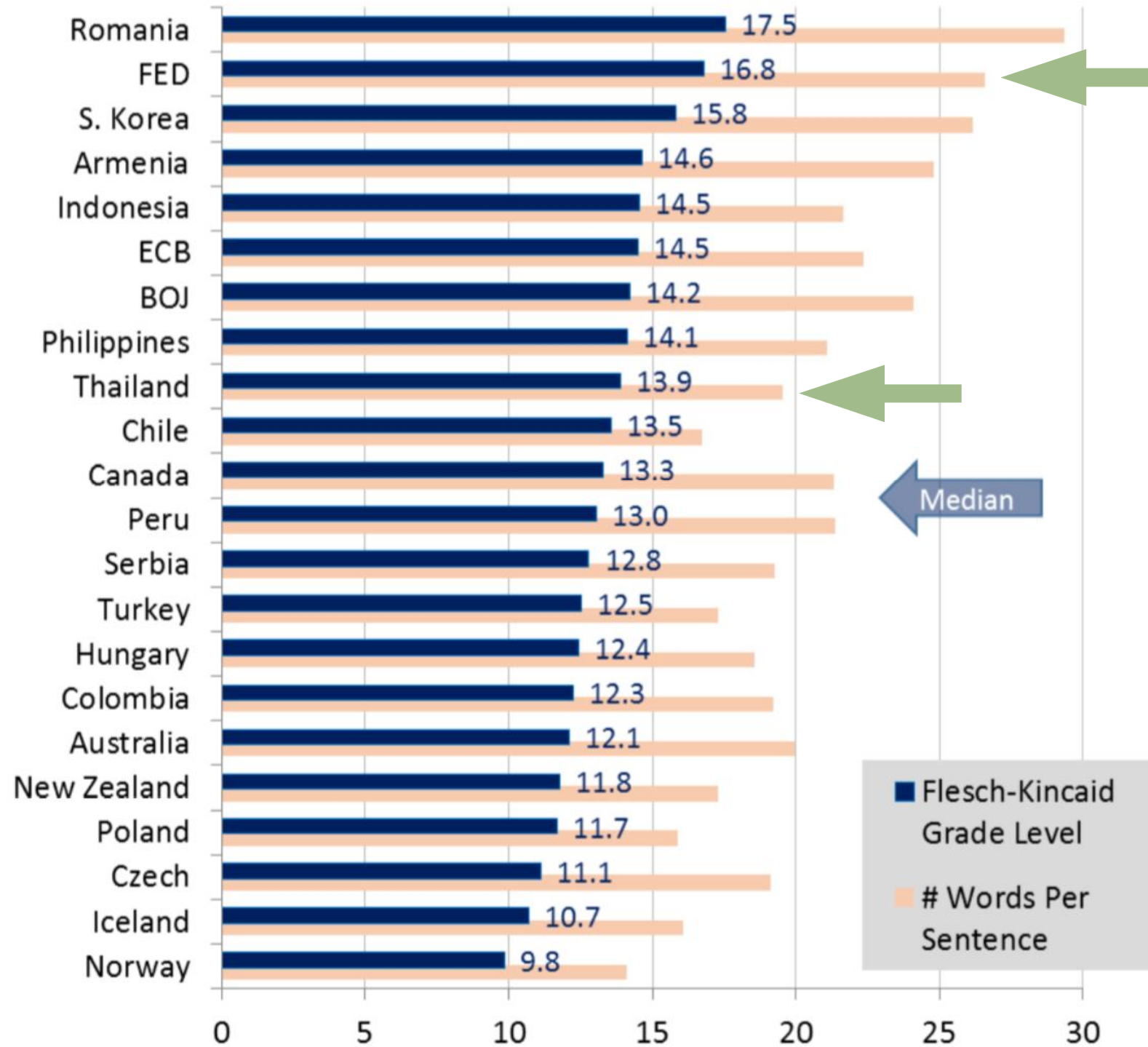




<p>อย่างไรก็ตาม ค่าเงินบาทในช่วงที่ผ่านมา โน้มหนักค่าขึ้นเมื่อเทียบกับเงินสกุลหลักบางสกุล และอาจไม่เป็นผลดีต่อการฟื้นตัวของเศรษฐกิจไทย</p>	<p>Nonetheless, the baht appreciated against some major currencies over the recent period, which might not be beneficial to the ongoing economic recovery.</p>
<p>นอกจากนี้ คณะกรรมการฯ เห็นว่ายังคงต้องติดตามความเสี่ยงต่อเสถียรภาพการเงิน รวมทั้งพฤติกรรมแสวงหาผลตอบแทน ที่สูงขึ้น (search for yield) จาก การที่อัตราดอกเบี้ยอยู่ในระดับต่ำเป็นเวลานาน</p>	<p>In addition, financial stability risks from the prolonged low interest rate environment, including search-for-yield behavior, continued to warrant close monitoring.</p>
<p>คณะกรรมการฯ เห็นว่าการรักษาขีดความสามารถในการดำเนินนโยบาย (policy space) ยังมีความสำคัญ เพราะในระยะข้างหน้า เศรษฐกิจไทยยังจะต้องเผชิญกับความไม่แน่นอนเพิ่มขึ้น โดยเฉพาะ การฟื้นตัวของเศรษฐกิจโลกที่ยังเปราะบาง และความไม่แน่นอนของทิศทาง การดำเนินนโยบายการเงินของประเทศอุตสาหกรรมหลักที่จะส่งผลให้ เงินทุนเคลื่อนย้ายและอัตราแลกเปลี่ยนผันผวนมากขึ้น</p>	<p>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.</p>
<p>ในระยะต่อไป คณะกรรมการฯ เห็นว่านโยบายการเงินควรอยู่ในระดับผ่อนปรนต่อเนื่อง และพร้อมที่จะใช้เครื่องมือเชิงนโยบายที่มีอยู่อย่างเหมาะสม เพื่อให้ภาวะการเงินโดยรวมเอื้อต่อการฟื้นตัวของเศรษฐกิจ ควบคู่กับการรักษาเสถียรภาพการเงินของประเทศ</p>	<p>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.</p>

Syntactic complexity caused from Thai-to-English translation?

Figure 3.3: Flesch-Kincaid Grade Level Prediction (Avg. 2008-2015)







<b>FOMC</b>	<b>BOT</b>
<p>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.</p>	<p>Mr. Jaturong Jantarangs, Secretary of the Monetary Policy Committee (MPC), announced the outcome of the meeting on 22 June 2016 as follows.</p>
<p>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.</p>	<p>The Committee voted unanimously to maintain the policy rate at 1.50 percent. One MPC member was unable to attend this meeting.</p>
<p>The CCyB would then be available to help banking organizations absorb shocks associated with declining credit conditions.</p>	<p>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.</p>
<p>Implementation of the buffer could also help moderate fluctuations in the supply of credit.</p>	<p>Meanwhile, monetary conditions remained accommodative and conducive to the economic recovery.</p>
<p>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.</p>	<p>Hence, the policy rate should be kept on hold at this meeting.</p>

FOMC	BOT
<p>Because economic and financial risks are constantly evolving, the range of indicators and models that the Board may consider is likely to change over time.</p>	<p>The Thai economy continued to recover on the back of public expenditure and broad-based improvement in the tourism sector, while private consumption expanded as anticipated.</p>
<p>In response to comments, the final policy statement clarifies that the Board expects that the CCyB will be activated when systemic vulnerabilities are meaningfully above normal and that the Board generally intends to increase the CCyB gradually.</p>	<p>However, expansion in private investment remained low, and merchandise exports still contracted in tandem with other Asian economies which were weaker than expected.</p>
<p>The final policy statement also emphasizes that the Board expects to remove or reduce the CCyB when the conditions that led to its activation abate or lessen and when the release of CCyB capital would promote financial stability.</p>	<p>Overall, the economic momentum from domestic demand and the tourism sector offset the impact from contraction in merchandise exports.</p>
<p>The Board generally would expect to provide notice to the public and seek comment on the proposed level of the CCyB as part of making any final determination to change the CCyB.</p>	<p>As such, the economy was expected to expand at the same rate as previously assessed.</p>
<p>The CCyB applies to banking organizations that are subject to the advanced approaches capital rules, generally those with more than \$250 billion in assets or \$10 billion in on-balance-sheet foreign exposures, and to any depository institution subsidiary of such banking organizations.</p>	<p>Nevertheless, there are remaining downside risks from slower-than-expected growth of trading partners, and private sector's confidence which is still fragile. However, concerns pertaining to the impact of the drought have subsided, and prices of some agricultural commodities have shown sign of recovery.</p>

# Result of Readability

- Readability tends to fall when central banks lower their policy interest rates
- What is the implication?

# Topic modeling

## Topics

gene	0.04
dna	0.02
genetic	0.01
...	

life	0.02
evolve	0.01
organism	0.01
...	

brain	0.04
neuron	0.02
nerve	0.01
...	

data	0.02
number	0.02
computer	0.01
...	

## Documents

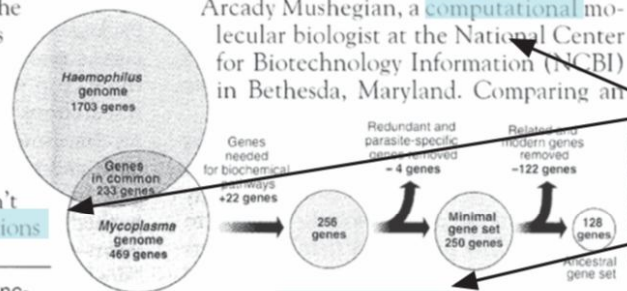
### Seeking Life's Bare (Genetic) Necessities

COLD SPRING HARBOR, NEW YORK— How many **genes** does an **organism** need to **survive**? Last week at the genome meeting here,\* two genome researchers with radically different approaches presented complementary views of the basic genes needed for **life**. One research team, using **computer** analyses to compare known **genomes**, concluded that today's **organisms** can be sustained with just 250 genes, and that the earliest life forms required a mere 128 **genes**. The other researcher mapped genes in a simple parasite and estimated that for this organism, 800 genes are plenty to do the job—but that anything short of 100 wouldn't be enough.

Although the numbers don't match precisely, those **predictions**

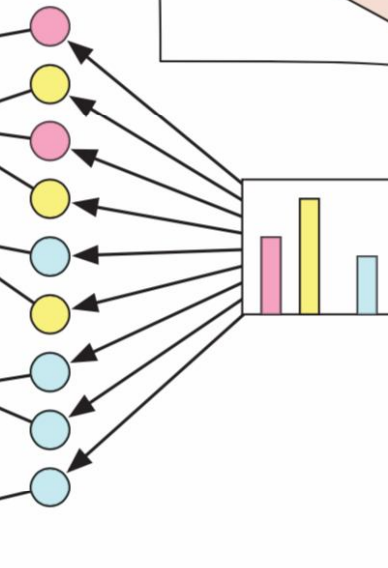
\* Genome Mapping and Sequencing, Cold Spring Harbor, New York, May 8 to 12.

"are not all that far apart," especially in comparison to the 75,000 **genes** in the human genome, notes Siv Andersson of Uppsala University in Sweden, who arrived at the 800 number. But coming up with a consensus answer may be more than just a **genetic numbers** game, particularly as more and more **genomes** are completely mapped and sequenced. "It may be a way of organizing any newly **sequenced genome**," explains Arcady Mushegian, a **computational** molecular biologist at the National Center for Biotechnology Information (NCBI) in Bethesda, Maryland. Comparing an

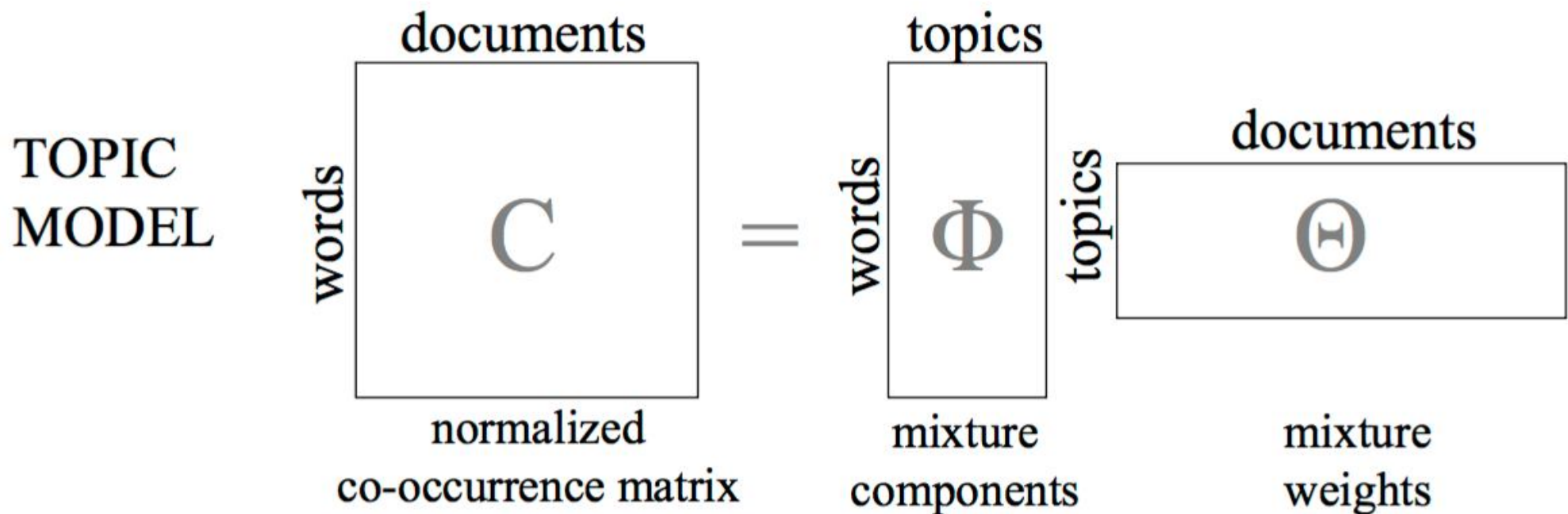


Stripping down. **Computer analysis** yields an estimate of the minimum modern and ancient genomes.

## Topic proportions and assignments



# Topic modeling



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

# Topic modeling

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

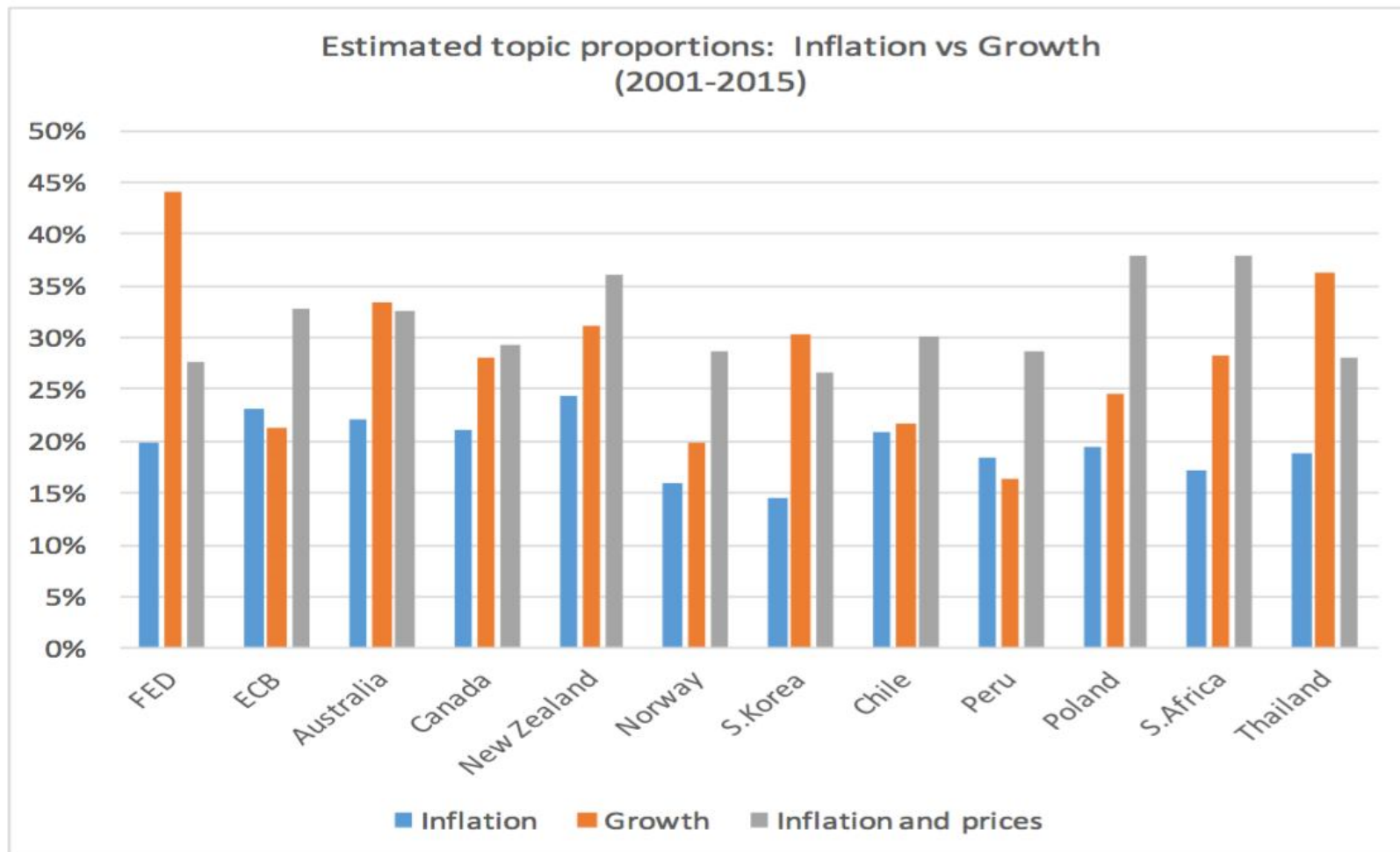
Topic 1 : remain, pressure, low, recent, modern, data, continue, growth, inflationary, confirm

= inflation



# 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

<b>Contraction</b>	<b>Expansion</b>
decreas*	increas*
decelerat*	accelerat*
slow*	fast*
weak*	strong*
low*	high*
loss*	gain*
contract*	expand*

# Results of topics and tones

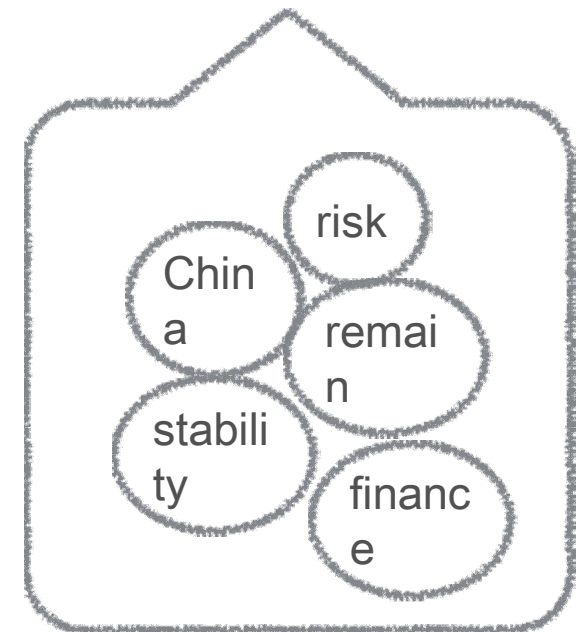
- Proportion of topics & tones ~ macroeconomic variable  
i.e. economic growth, inflation and policy interest rate
- FOMC : estimated growth topic proportion is related  
negatively to GDP growth and positively to  
unemployment rate
- ECB : both estimated net tones of inflation and growth  
in the ECB 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 possible?

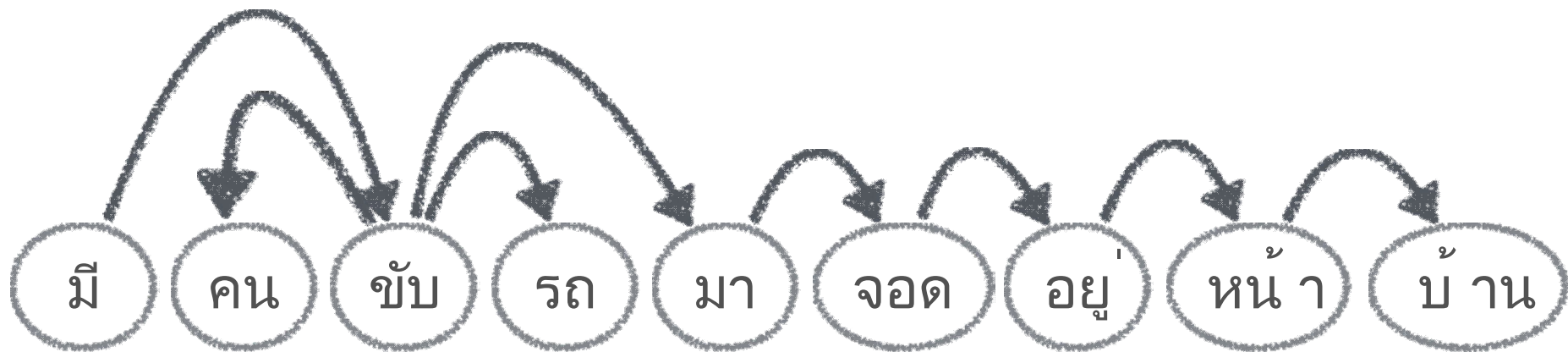
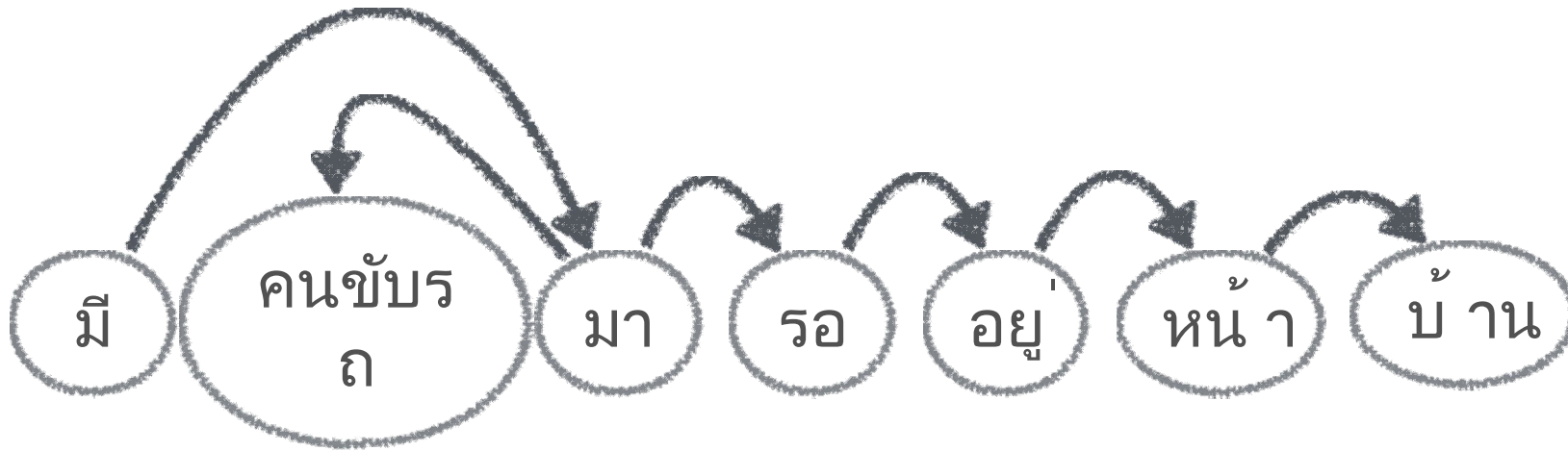
- Basic assumptions of the tools
  - Bag of words
  - Financial stability risks in China remained
  - China stability remains risky in financial
  - Stability risks remain in Chinese finance
  - Financial stability risk to remain in China
  - \*Risk remain financial stability China



# Thai Text Mining

## Is it possible?

- Nature of Thai language
  - No word, sentence boundary
  - Vague / ambiguity
    - มีคนขับรถมา รอ อยู่ หน้า บ้าน
    - มี คนขับรถ มา รอ อยู่ หน้า บ้าน
    - มีคนขับรถมา จอด อยู่ หน้า บ้าน
    - มี คน ขับ รถ มา จอด 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?

- 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

