

# The Use of Natural Language & Large Language Models for Analysis: Economics, Markets & Geopolitics

---

Alvaro Ortiz  
Head of Big Data Analysis  
BBVA Research

Predicción Económica Con Big Data (Funcas)

November 7<sup>th</sup>, 2023

# Index

- A “Short” history on Natural Language Models
- The Bags of Word Models: Just Words
- The Topic Models: The Global Structure of the Text
- Word Embeddings: “Local Context Matters”
- The Origin of Large Language Models (LLM)
- Tentative Conclusions



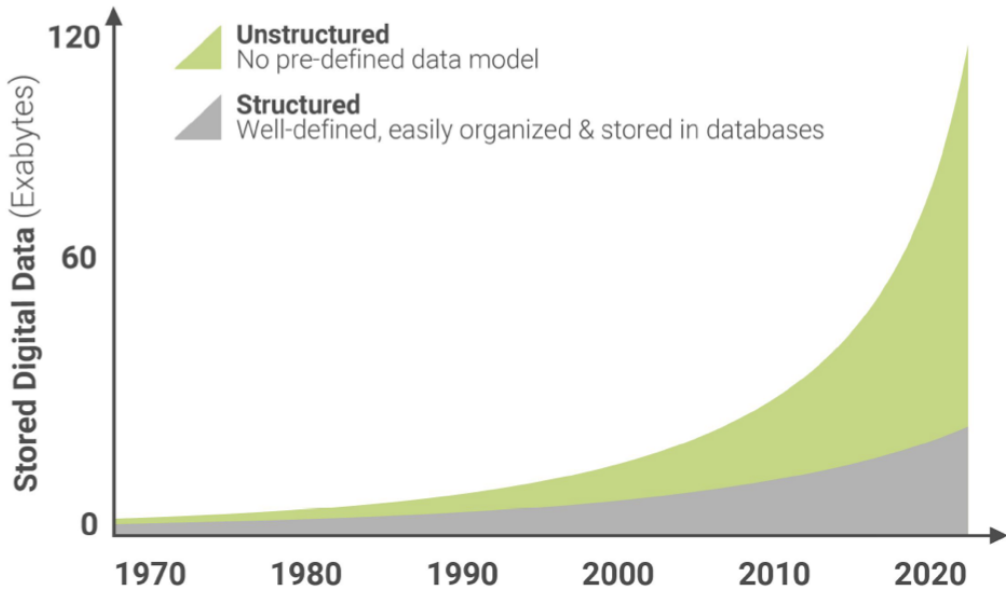
# A Short History of Natural Language Models

---



# ... generating a Vast Volume of information in Non-Structured way

A huge increase in Non-Structured Data since 2000s



Source: Hansen



# “Text as Data”: A short history of the origin of Large Language Models

## Bag of Words Model

It is individual Words  
What Matters...

Boolean Searches  
Dictionary Assisted  
Sentiment Analysis  
Networks...

## Topic Models

...But structural  
Words dependence too  
(Global  
Co-occurrence)...

LDA  
Dynamic Topic Models  
(Factor Models of Text)

## Word Embeddings

... Local Context  
is a more Natural Guide  
(Synonymy Vector)...

Word2Vec

## Large Language Models (LLM)

... Addressing Polysemy.  
requires “Attention”  
(weighing embeddings) for  
Prediction Tasks

BERT  
LLAMA  
Chat-GPT

# What is this for: The value of “Unstructured Data” in Social Sciences

- Measure important phenomena Economic, Social & geopolitics.
- Particularly **useful for some applications(\*)** on:
  - **Concept Detection** (ie Sentiment, Uncertainty, Risk..)
  - **Measuring Document Similarity** (ie Polarization of Political Documents)
  - **Relation of Concepts** (ie Risk vs Uncertainty)
  - **Associating Text with Metadata** (Using Topic to Forecasts conflicts)
- **Complement existing or create new measures for Economics & Social Sciences**
- **Powerful for some Social Sciences** with scarce quantitative (Politics, Geopolitics..)

(\*) [See Ash and Hansen \(2023\). Text Algorithms in Economics](#)



# The Bags of Words Models: Just Words

---



# Bag of Words Models: First “Clean” then convert “Text to Numbers”

- What is Text: **Cleaning or Pre-Processing** to Tokens
  - Latin Letters
  - Punctuations
  - Spaces, tabs, newlines
  - Numbers & Non- Alphanumeric: @,#
- **Construction of a term matrix: Text to Numbers**
  - Matrix can be very Large
  - Sparse
- **Order and Context** in same document is **irrelevant**
- **The most frequent the less relevant**

## A simple Example: 3 docs

Doc 1: Pilar is Econometrician

Doc 2: Eva is also Econometrician

Doc 3: Daniel is Statistician

Index	Pilar	is	Econometrician	Eva	also	Daniel	Statistician
1	1	1	1	0	0	0	0
0	1	1	1	1	1	0	0
0	1	0	0	0	0	1	1

Doc1:  $w1 = (1,2,3)$  Doc2:  $w2 = (4,2,5,3)$  Doc3:  $w3 = (6,2,7)$

Document Term Matrix (X)

$$X = \begin{bmatrix} 1110000 \\ 0111100 \\ 0100011 \end{bmatrix}$$

## Imagine a Complex one !!

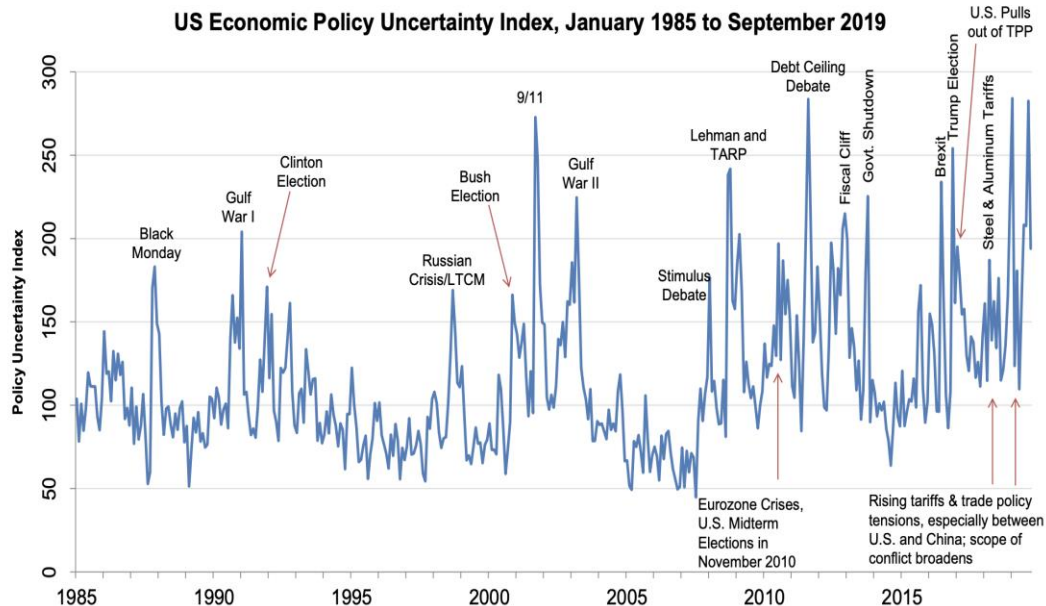
“Transparency and Deliberation FOMC Transcripts”

149 meetings (1987-2006)

46502 document

Associated metadata

# Concept Detection using Boolean Searches in Economics: The Economic Policy Uncertainty Index



Source: "Measuring Economic Policy Uncertainty" by Scott R. Baker, Nicholas Bloom and Steven J. Davis, as updated at [www.policyuncertainty.com](http://www.policyuncertainty.com). Monthly data normalized to 100 prior to 2010.



Latest **Most Read** Most Cited

Measuring Economic Policy Uncertainty

The Economic Policy Uncertainty (EPU) index of [Baker et al., 2016] (<http://www.policyuncertainty.com/>) is based on key word search applied to news paper articles from major US EU Newspapers. Search logic is the result of extensive manual audits of news paper articles. For each paper on each day since 1985, submit the following query:

1. Article contains "uncertain" OR "uncertainty", AND
2. Article contains "economic" OR "economy", AND
3. Article contains "congress" OR "deficit" OR "federal reserve" OR "legislation" OR "regulation" OR "white house"

Take resulting article counts and normalize by total news pa per articles that month.

# Supervised (Dictionary Assisted) Sentiment Analysis and Event codes

## General Sentiment Dictionaries (Harvard IV)

## Financial Dictionaries (Loughran and McDonald Dictionary)

## Financial Stability Dictionaries (Federal Reserve Financial Stability Dictionary)

## Geopolitical Events Data Sets (CAMEO & POLECAT)



## A Financial Stability Dictionary



Correa, Ricardo, Keshav Garud, Juan M. Londono, and Nathan Mislav (2017). Sentiment in Central Banks' Financial Stability Reports. International Finance Discussion Papers 1203.

# Geopolitics: Coding Events to “Extract” real time & HD Geopolitical information from News

## Geopolitical Events (Cameo) & Goldstein Scale

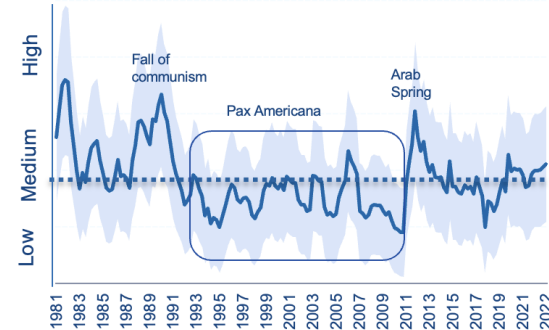
	Cameo Quads	Cameo Events	Code	Goldstein
Cooperation +	Material Cooperation	ENGAGE IN MATERIAL COOPERATION (6)	6	6
		PROVIDE AID (7)	7	7
		YIELD (26)	8	5
		INVESTIGATE (6)	9	-2
Verbal Cooperation	Verbal Cooperation	MAKE PUBLIC STATEMENT (11)	1	0
		APPEAL (28)	2	3
		EXPRESS INTENT TO COOPERATE (29)	3	4
		CONSULT (8)	4	1
Verbal Conflict	Verbal Conflict	ENGAGE IN DIPLOMATIC COOPERATION (9)	5	3.5
		DEMAND (27)	10	-5
		DISAPPROVE (13)	11	-2
		REJECT (27)	12	-4
Conflict +	Material Conflict	THREATEN (23)	13	-6
		PROTEST (27)	14	-6.5
		EXHIBIT FORCE POSTURE (6)	15	-7.2
		REDUCE RELATIONS (14)	16	-4.0
		COERCE (13)	17	-7.0
		ASSAULT (14)	18	-9.0
		FIGHT (8)	19	-10
		USE UNCONVENTIONAL MASS VIOLENCE (8)	20	-10

### Example: Inside Protest

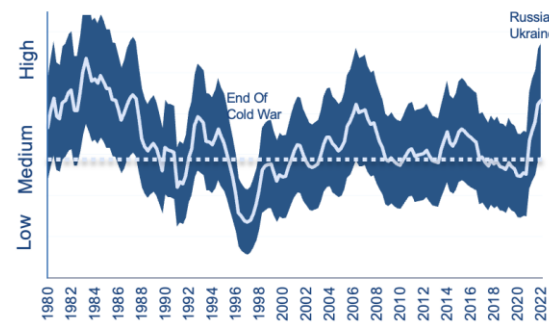
PROTEST	Goldstein
Engage in political dissent, not specified below	-6.5
Demonstrate or rally	-6.5
Demonstrate for leadership change	-6.5
Demonstrate for policy change	-6.5
Demonstrate for rights	-6.5
Demonstrate for change in institutions, regime	-6.5
Conduct hunger strike, not specified below	-6.5
Conduct hunger strike for leadership change	-6.5
Conduct hunger strike for policy change	-6.5
Conduct hunger strike for rights	-6.5
Conduct hunger strike for change in institutions, regime	-6.5
Conduct strike or boycott, not specified below	-6.5
Conduct strike or boycott for leadership change	-6.5
Conduct strike or boycott for policy change	-6.5
Conduct strike or boycott for rights	-6.5
Conduct strike or boycott for change in institutions, regime	-6.5
Obstruct passage, block	-7.5
Obstruct passage to demand leadership change	-7.5
Obstruct passage to demand policy change	-7.5
Obstruct passage to demand rights	-7.5
Obstruct passage to demand change in institutions, regime	-7.5
Protest violently, riot	-7.5
Engage in violent protest for leadership change	-7.5
Engage in violent protest for policy change	-7.5
Engage in violent protest for rights	-7.5
Engage in violent protest for change in institutions, regime	-7.5

Dissent & Rally  
Violent Riot+

### World Protest Index



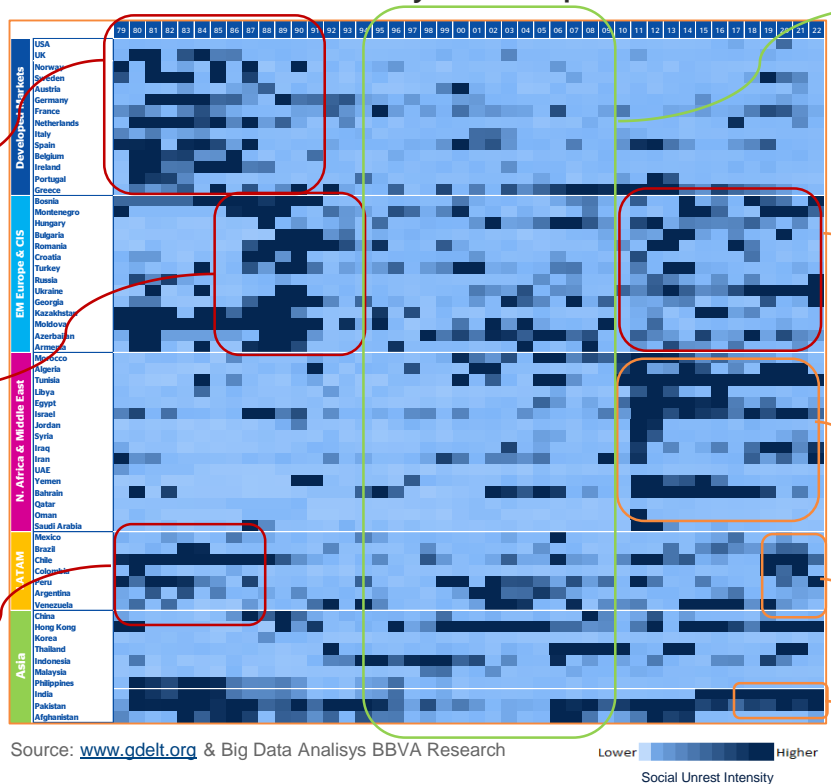
### World Conflict Index



Source: Big Data Analysis Unit (BBVA Research) using CAMEO

# The BBVA Research "Social Unrest" Heat Map from News Worldwide: How the World Feels & The Need to Remember history

World Social Unrest Intensity Heat Map 1979- 2022



Western Europe enjoyed a long period of "Pax", but don't forget that "stability" was not always the case and is Far from guaranteed

The Fall of Communism triggered some "Adaptation Noise " in East Europe but short lived

The 80s & 90s were also tumultuous for Latin America

There is also margin for Optimism...!!!

The Financial Crisis and tectonic geopolitical shifts have triggered some "Clouds" In East Europe than need to be monitored closely

The Long Lasting "Artificial Calm Provided by Autocrats in MENA came to a halt during Arab Spring.. Nothing has been the same since then

Latin America is experiencing some Winds of change

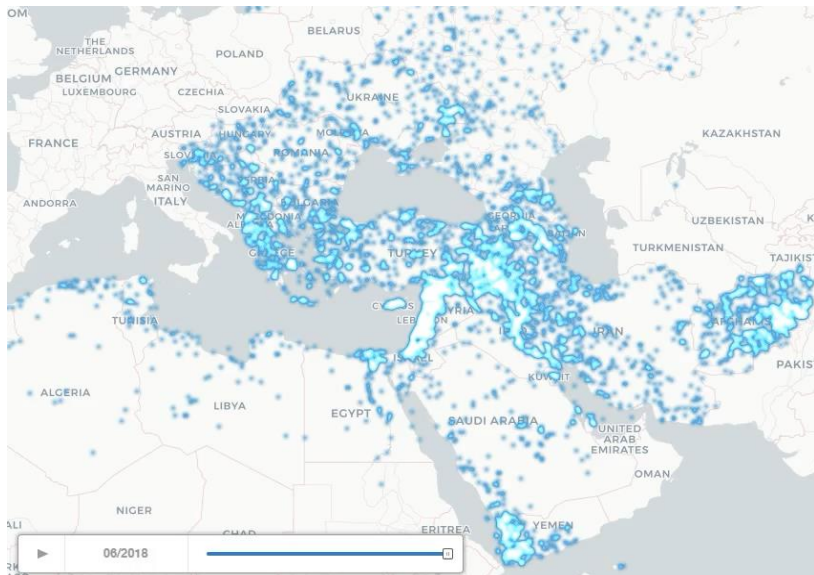
Still some chronic conflicts with the potential to contaminate neighbors

Source: [www.gdelt.org](http://www.gdelt.org) & Big Data Analys BBVA Research

Lower Higher  
Social Unrest Intensity

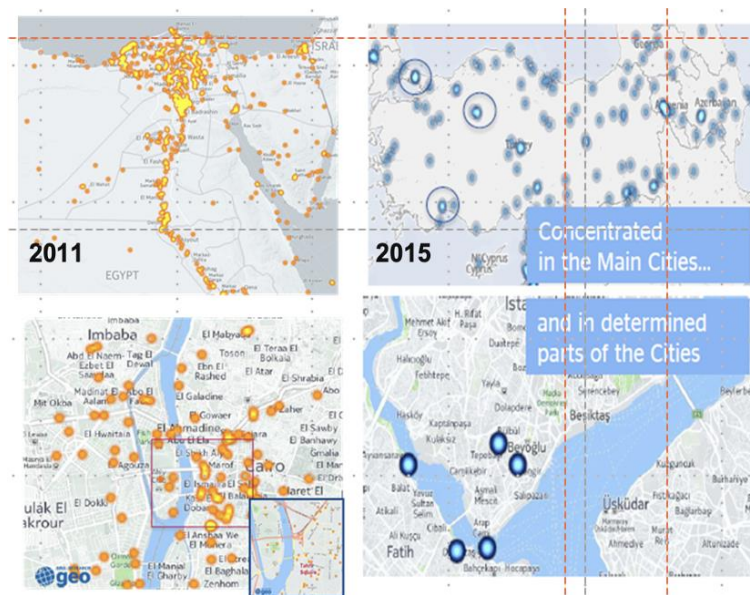
...remember that the Information is Geo-Localized offering High-Definition advantages including “zooming” potential...

**World Conflict Intensity Map : Inside The Middle East 2017-18**  
(number of conflicts/ total events)



● Conflict Intensity

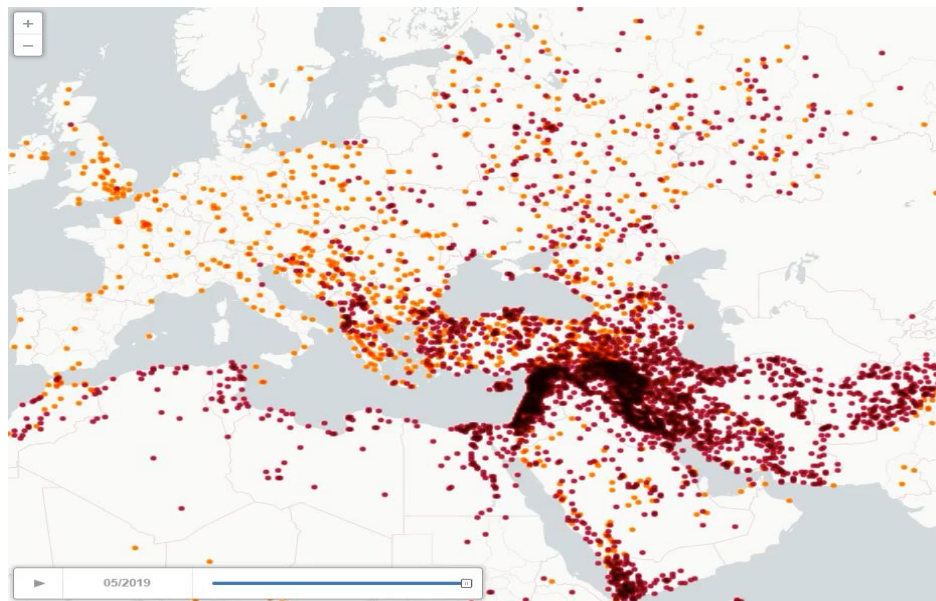
**Zooming Out: Social Unrest Events in Real Time HD**  
(Events: El Cairo 2011, Istanbul 2015, Hong Kong 2014)





## ... or tracking migrant flows from origin to destiny ...


**BBVA Research Refugees Flows Map 2015-19**  
(Number of Media Citations of Migrant inflows and outflows)




● Refugees Outflows    ● Refugees Inflows

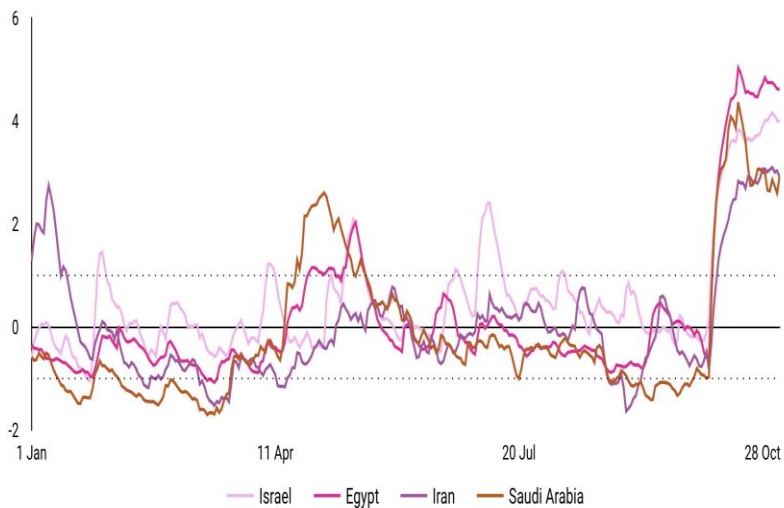
Source: Big Data Analysis Unit (BBVA Research)

# The BBVA Research Geopolitics Monitor ... is being useful to track recent Middle East tensions...

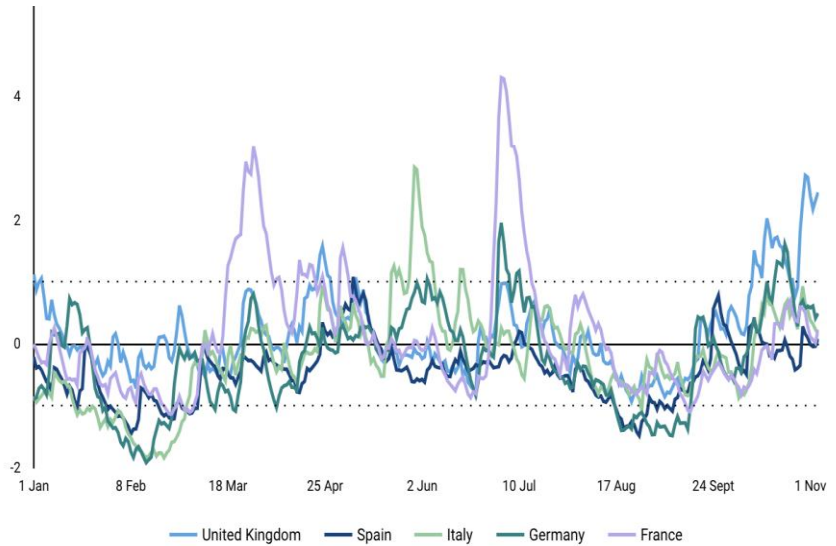

BBVA Research Geopolitics Monitor  
Using text sentiment analysis from the Media

 Geopolitical Risk
Political Tensions
Economic Policy Uncertainty
Protest
Conflict
Bilateral Tensions

## Middle East: Geopolitical Risk



## Europe: Protest Intensity Index Risk



Source: Big Data Analysis Unit (BBVA Research)



# And follow the potential "collateral" Damage for the Rest of the world...

## BBVA Research Geopolitics Monitor main "Signal" changes after the Hamas-Israel Conflict

**Geopolitics Monitor Summary Before Hamas-Israel Conflict**

(News Sentiment Map. Sentiment relative to average sentiment in each country)

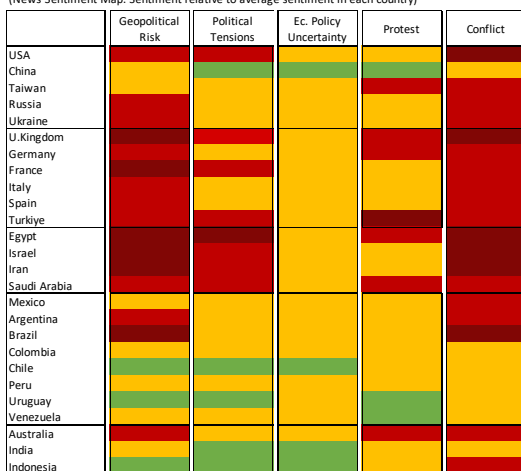


Source: BBVA Research

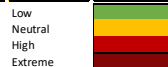


**Summary of Signals Before Hamas-Israel Conflict (October 23)**

(News Sentiment Map. Sentiment relative to average sentiment in each country)

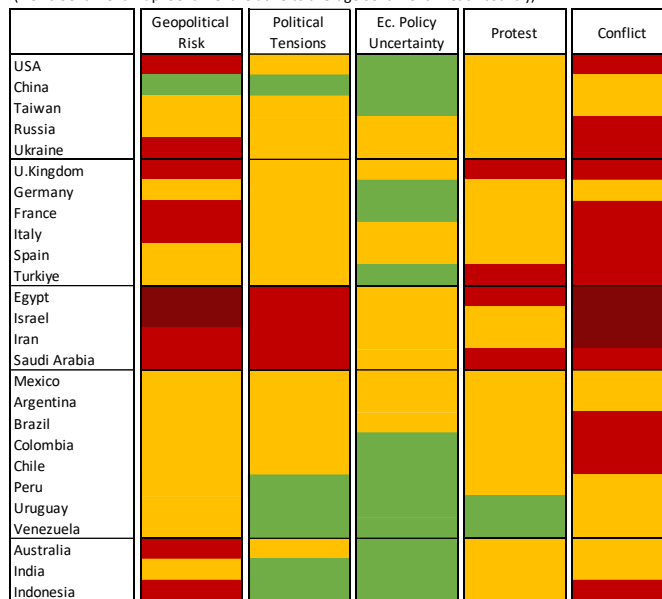


Source: BBVA Research

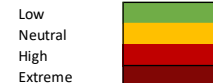


**Summary of Signals After Hamas-Israel Conflict November 5th**

(News Sentiment Map. Sentiment relative to average sentiment in each country)



Source: BBVA Research





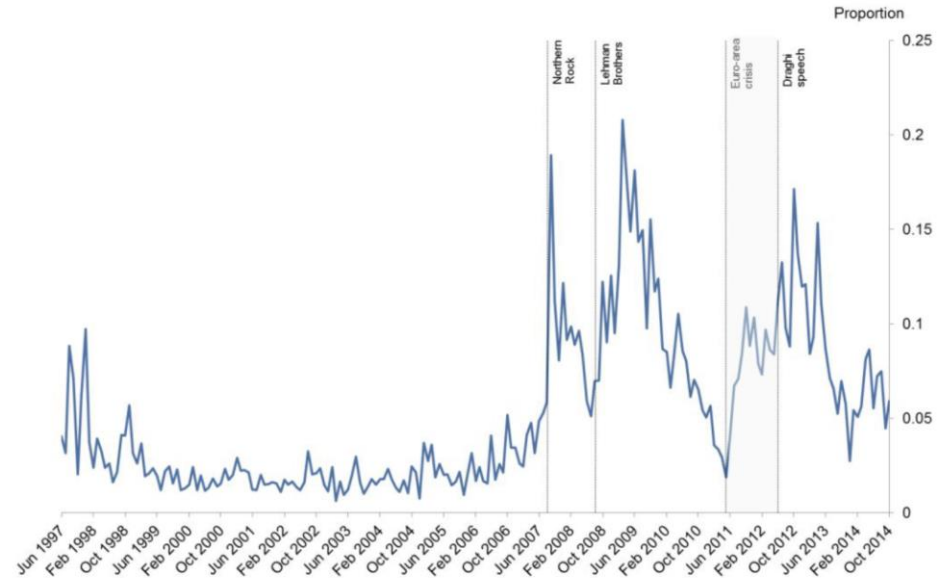
# Topic Models: The Global Structure of Text

---

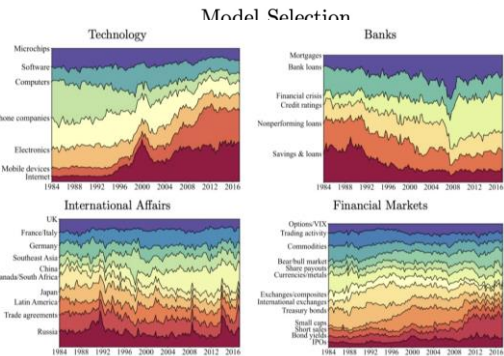
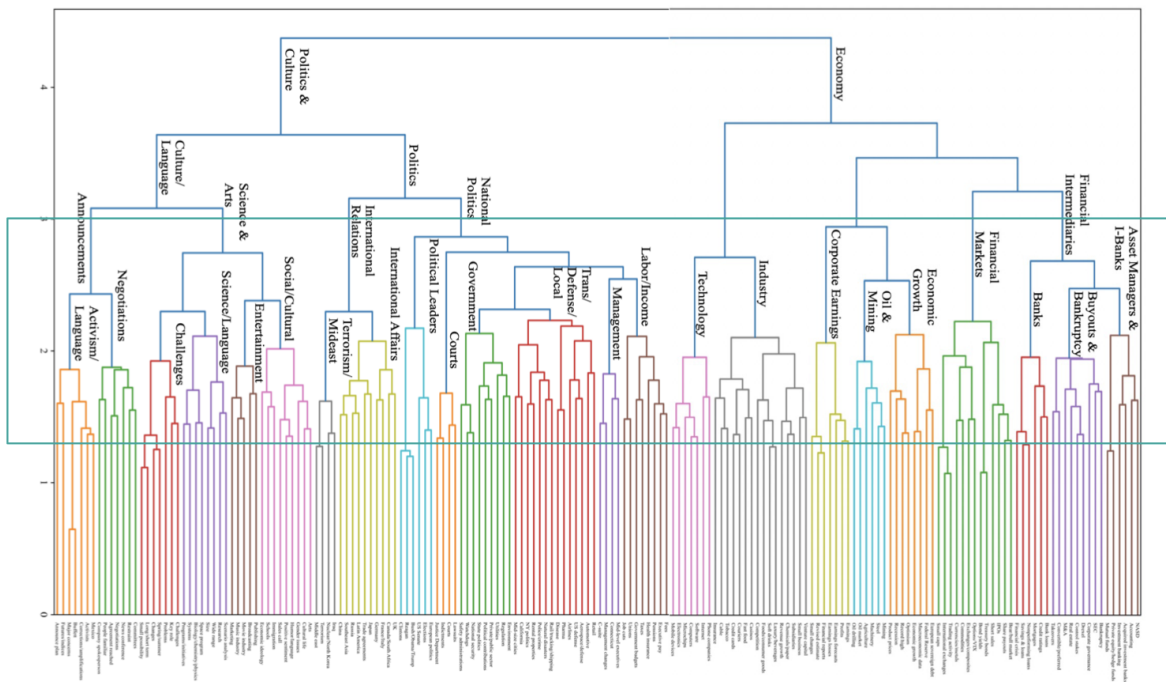
# The Structure of the Test: Topic Models (Factor Models of Discrete Text) and Networks

- The bag-of-words model ignores the **strong dependence structure among words**.
- **Topic Models reduce the dimensionality (“Factor Models”)** keeping heterogeneity
- Analyze the **“Global Documents “Co-Occurrence”**
- They have **many applications in Economics, Geopolitics & Social Sciences**

**Chart 13:** Estimate of the proportion of MPC minutes covering the topic of 'banking'



# The Structure of News: A classification of Topics and Meta-topics of News in 800000 WSJ articles (1984-2017)



Note. Stacked area plot of attention allocation to baseline topics within four of the first-level metatopics.

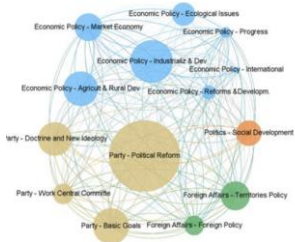
**23** Metatopics  
**250** Topics ( $\eta$ )



# Inside the China Communist Party Congress Files: "Tectonic Shifts" of China's Long Run Strategy...

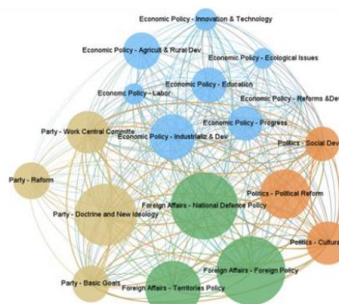
Chinese Communist Party congresses national documents: Network Topic analysis

2002



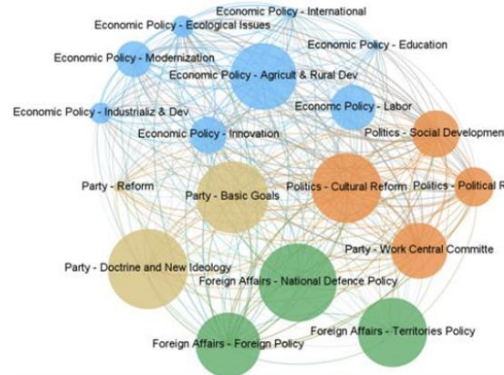
Reforms

2007



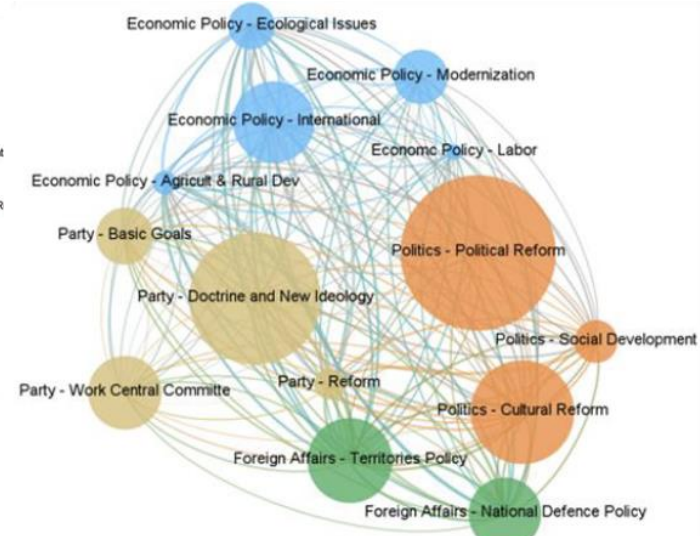
Economic Reforms &  
Foreign Affairs dominance

2012



Doctrine & Politics  
Gaining Momentum

2017



Doctrine & Politics &  
Territorial Aspirations

-  Size of the Node = Relevance
-  Economic Policies
-  Doctrine-Ideology-Party
-  Foreign Affairs
-  Politics





# Word Embeddings: “Local Context Matters”

---



# Word Embeddings: Words with similar meanings lie close together

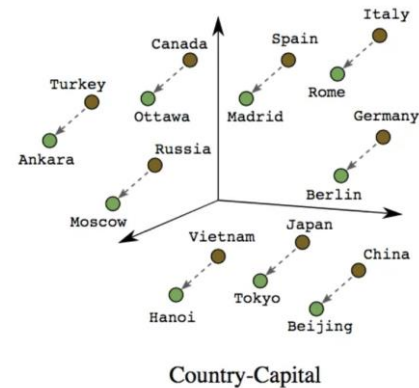
- **Local context can a more natural guide to meaning** than Global co-occurrence.
- Words present many sentences away would generally be less informative (**Distance matters**)
- **Distributional Hypothesis: “You shall know a word by the company it keeps”** (Firth).
- A word embedding is a low-dimensional vector representation of a word Ideally including words with similar meanings (“synonymy”).
- **Word2vec** was developed by Google (2013) and was an important milestone. It was invented for training Word Embedding and is based on a Distributional Hypothesis.

Every workday I make a break in the morning to have a [???]. Contrary to my colleagues I like Black (no Milk), with a lot of Ice and no Sugar

Please identify the word for '???'



The word for ??? based on the context provided is "coffee."



# Word Embeddings have been used to explore “Disruptive Technologies” in Patent, Earning Calls and Job Posts...

## Identifying Novel Technologies finding bigrams technologies in Patent-Calls-Jobs

### US Patent Fillings

- They Use bigrams less ambiguous
- Read US Patent Fillings 1976-2016 (3 million Patents)
- Cleaning
- Decompose Text in bigrams (less ambiguous)
- Isolate novel and influential inventions (>1000 citations)

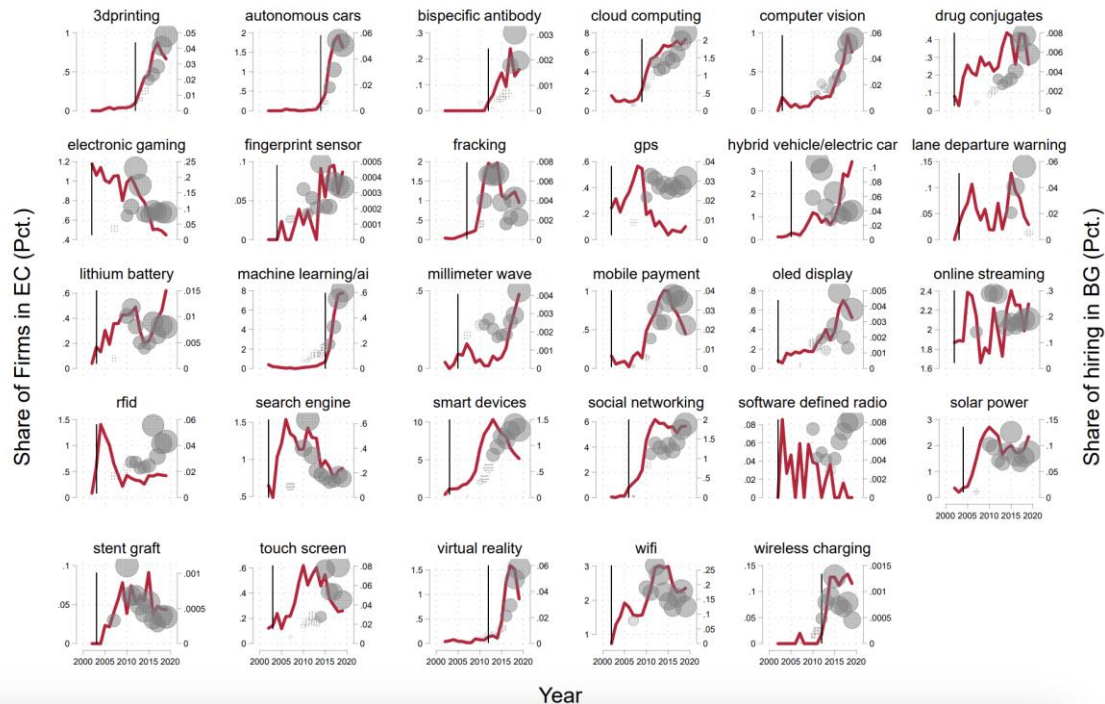
### Quarterly Earning Calls

- Bigrams Appearing > 100 times
- Disrupting: Increasing Incidence in Calls
- Supervised: Refine manually (Cloud Compute= Cloud Services)
- Using Word2Vec to find proximate or similar bigrams (AI=ML=DL)
- 221 bigrams associated to 29 tech

### Tech Job Postings

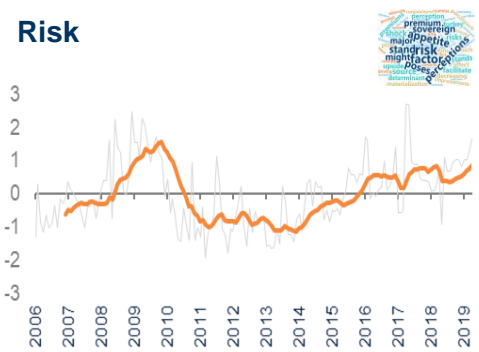
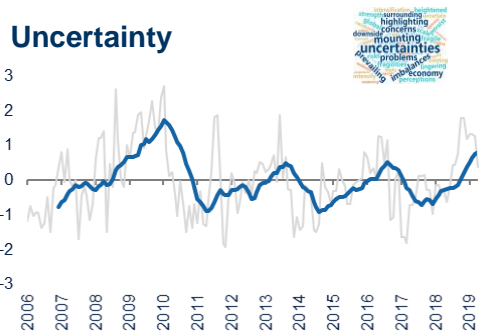
- Find tech bigrams in offers
- Identify Skill, Location, Industry, Firm

## Exposure of Disruptive Bigrams in Earning Calls Job Offers

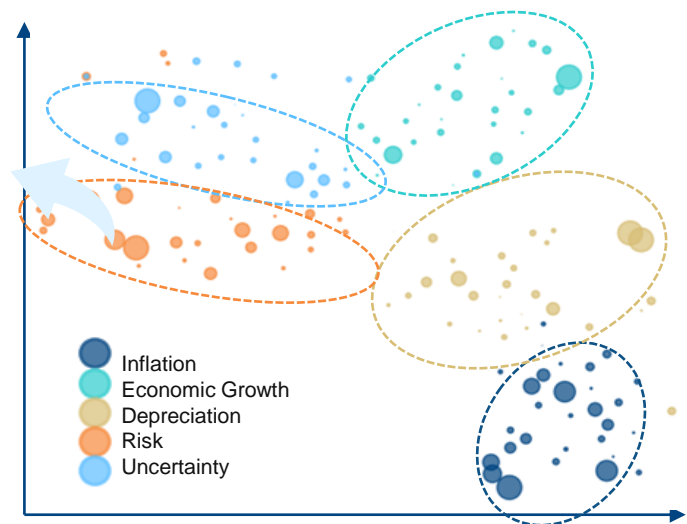


# ... To Distinguish “Uncertainty and Risk” inside Central Bank Reports and Transcripts...

**TURKEY: UNCERTAINTY AND RISK**  
(ESTIMATED THROUGH BIG DATA TECHNIQUES USING WORD TO VEC MODEL)



Words distribution given selected keywords



Terms Close to Uncertainty in FOMC Transcripts

term	sim	term	sim
uncertainties	0.741	challenges	0.415
anxiety	0.48	fragility	0.405
pessimism	0.479	clarity	0.401
skepticism	0.465	concerns	0.4
optimism	0.445	risks	0.397
caution	0.442	disagreement	0.387
gloom	0.437	volatility	0.384
uncertain	0.433	tension	0.383
sensitivity	0.427	certainty	0.382
angst	0.426	skepticism	0.38

Terms Close to Risk

term	sim	term	sim
risks	0.737	misdirected	0.385
threat	0.609	odds	0.379
danger	0.541	uncertainty	0.375
dangers	0.463	concern	0.371
vulnerability	0.457	prospect	0.37
chances	0.451	instability	0.363
breakout	0.433	potentially	0.352
probability	0.426	concerns	0.352
possibility	0.409	challenges	0.346
likelihood	0.406	risking	0.342

Source: Big Data Analysis Unit (BBVA Research)



# The Origin of Large Language Models (LLM)

---

# Sequence of Word Embeddings to Forecast: LLM Generative Models

- **Problem:** Words Embedding address the problem of “synonymy” NOT “polysemy” (many meanings for the same word).
- **Solution:** To deal with Polysemy (and limit Hallucination !!!) word embeddings have to **move to sequence embeddings**.
- The Large Language Models focus on “attention” by weighing embeddings using Neural Networks with Transformer Architecture
- LLM models combine **attention and feedforward layers to perform Prediction tasks**

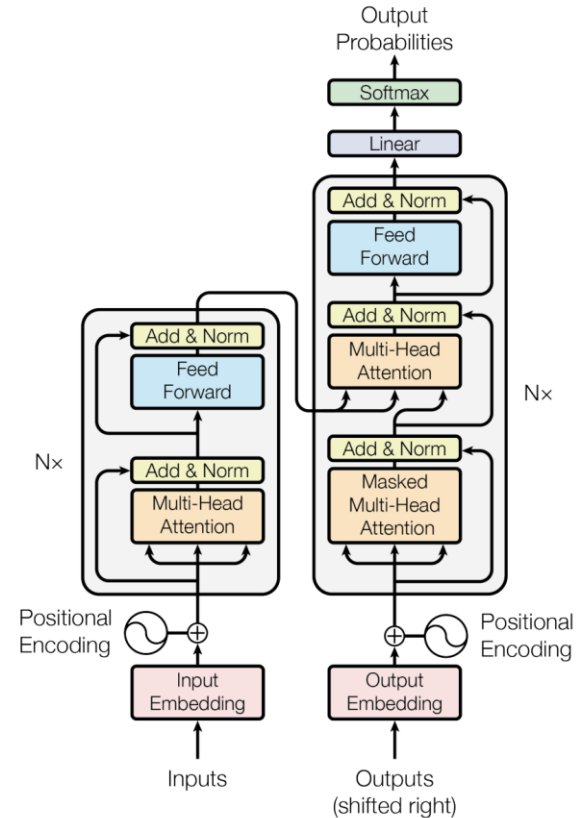


Figure 1: The Transformer - model architecture.

# Training Matters: You do better what you are trained for !!!

- **Different LLM Models:**

- BERT, OPT,
- GPT-2, GPT-3,
- RoBERTa , Chat-GPT

- **Different Tokenization Strategy**

- **Different Transformer Architecture**

- **Different Hyperparameters**

- **Pre-Training on Different Large Datasets:**

Wikipedia, Common Crawl, Web Text

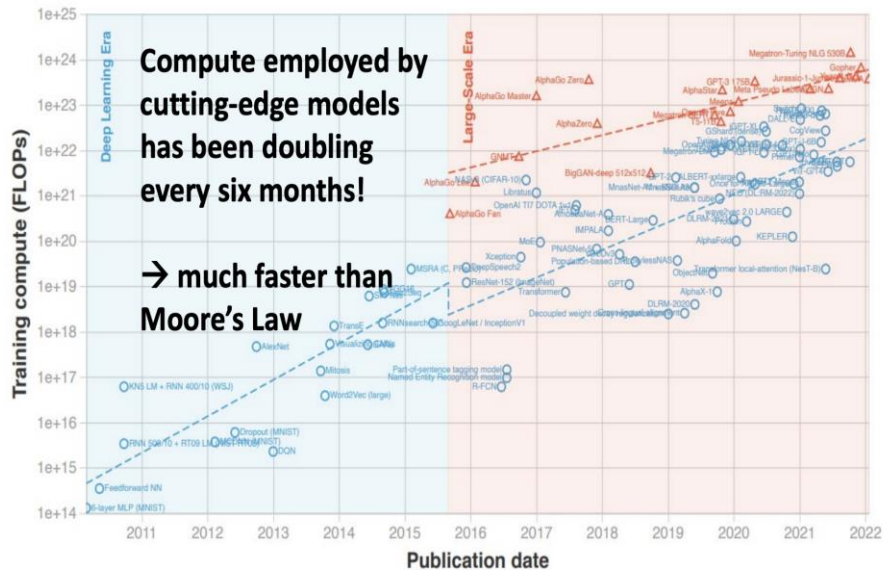
- **LLM can be trained on Specific Ad-Hoc Data bases**

- **Different Tasks:**

- BERT : Masked Language Model (MLM) & Next Sentence Prediction (NSP)
- RoBERTa: Masked Language Model Only
- OPT and GPTs : Next Token Prediction (NTP) or Autoregressive Language Modelling (ALM)

Training compute (FLOPs) of milestone Machine Learning systems over time

n = 102



Copyright © 2022 by Sevilla et al under a CC-BY-4.0 license



# All what you need is Attention: Example on Remote Work

## Remote Work across Jobs, Companies, and Space

Stephen Hansen (UCL), Peter Lambert (LSE), Nick Bloom (Stanford),  
Steven Davis (Hoover), Yabra Muvdi (ETH), Raffaella Sadun (Harvard),  
Bledi Taska (Lightcast)

'As a leading firm in the [MASK] sector, we hire highly skilled software engineers.'

'As a leading firm in the [MASK] sector, we hire highly skilled petroleum engineers.'

### Prediction for Masked Words in the Sentences Example

'software engineers' Sentence		'petroleum engineers' Sentence	
Word	Prob.	Word	Prob.
it	0.08	energy	0.279
automotive	0.079	oil	0.27
technology	0.072	petroleum	0.088
healthcare	0.058	mining	0.035
insurance	0.053	defence	0.021
software	0.041	automotive	0.02
engineering	0.031	construction	0.017
public	0.03	gas	0.017
infrastructure	0.028	engineering	0.016
financial	0.028	water	0.012

This table displays masked word prediction probabilities for the two example sentences above. The training corpus for estimating these probabilities is English-language online job postings provided by Lightcast (formerly Emsi Burning Glass). The Transformer model estimated for the task is DistilBERT (Sanh et al. 2020). See Hansen et al. (2023) for more details.

**Software Developer**  
Pearson ★★★★★ 2,739 reviews  
Australia  
Remote  
Full-time  
You must create an Indeed account before continuing to the company website to apply

[Apply on company site](#)

**Our purpose:** At Pearson we 'add life to a lifetime of learning' so everyone can realise the life they imagine. We do this by creating vibrant and enriching learning experiences designed for real-life impact.

**Our company:** Pearson was founded in 1844 and has been built on our ability to grow with and adapt to a constantly evolving market. Our 20,000+ employees are dedicated to creating the high-quality, digital-first, accessible and sustainable resources for lifelong learning.

**Flexible working:** Pearson is committed to hybrid working practices. When you are not working from home, you'll be based in our Nunawading office that has free parking and is walking distance to 2 train stations. This is a great location for those that are not a fan of the city commute.

**The Role:** As a Software Engineer, you will be joining one of our cross-functional scrum teams and will play a key role in the development of our online assessment platform. Reporting to our Engineering Manager, you will work from home and collaborate via telecommuting platforms.

**What you will do:**

**Expense Processor (Remote U.S.A.)**  
Plus Relocation ★★★★★ 17 reviews  
Minneapolis, MN 55426 • Remote  
Full-time  
You must create an Indeed account before continuing to the company website to apply

[Apply on company site](#)

**Job details**

**Job Type:** Full-time  
**Work From Home:** Allowed  
**Location:** Anywhere

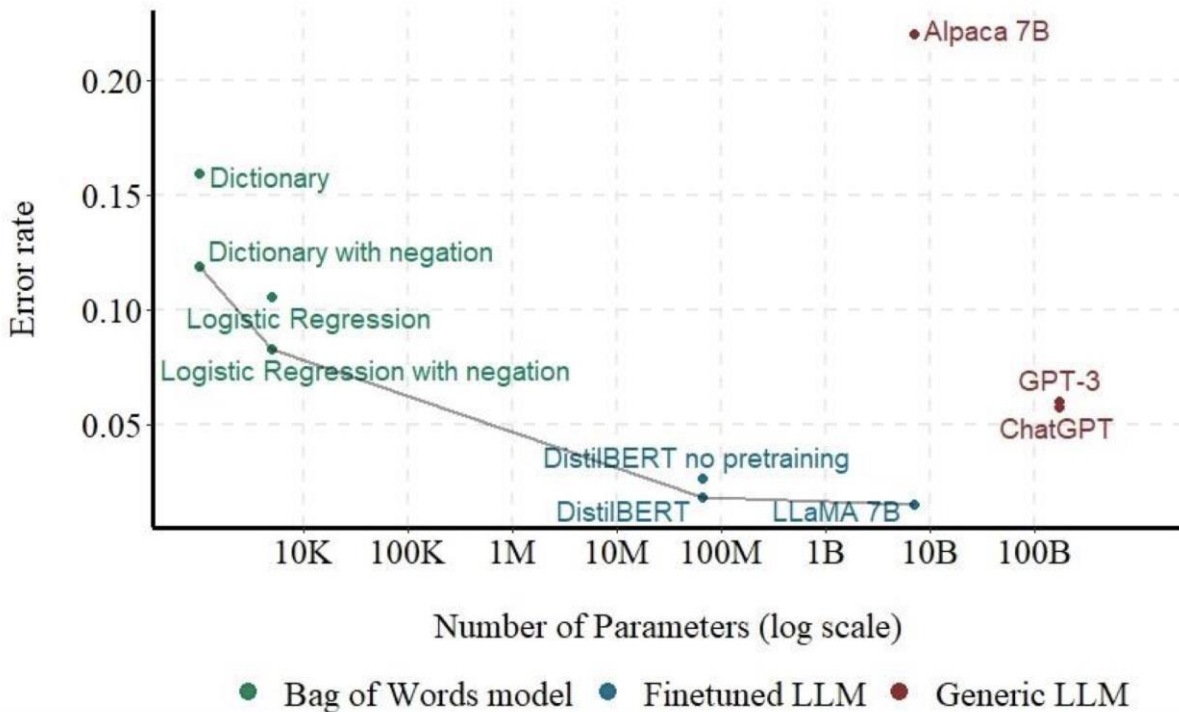
**Full Job Description**

Plus Relocation is looking for a numbers driven, detail



# Wise Prompting, Fine Tuning and Customization helps

## Trade-off in model choice



# True ... They are better and better and still margin to improve with Fine Tuning Example Powell Recent Speech

Use the LM financial sentiment to evaluate the statement (exclude the Q&A paragraph). How is the sentiment of the speech?

The Loughran and McDonald (LM) financial dictionary is a lexicon specifically designed for analyzing the sentiment of financial documents. It categorizes words into positive, negative, uncertainty, litigious, strong modal, and weak modal categories.

What is the balance os positive vs negative tokens of the statement ?

To determine the sentiment of the statement using the LM need to analyze the text for the presence of words from that have a direct integration with the LM dictionary in this env

That said, I can provide a general sentiment analysis base

- Positive Sentiment:** The statement emphasizes the Federal mandate, the robustness of consumer spending, an activity. The Fed also acknowledges the challenges and r goals.
- Negative Sentiment:** The statement acknowledges the h need for a restrictive monetary policy, and the potent below-trend growth. It also mentions challenges in the h investment due to higher interest rates.
- Uncertainty:** The statement mentions the evolving outlo uncertainties in gauging the stance of policy, and the pot on incoming data.

Given this analysis, the sentiment of the statement appea Reserve acknowledges challenges and uncertainties, it al indicators and its commitment to achieving its goals.

For a more detailed sentiment analysis using the LM finan financial analysis software or tools that integrate the LM c

However, I don't have a direct integration with the LM dictionary in this environment. Instead, I can provide a general analysis based on the content:

**Positive Tokens:**

- "focused"
- "maximum employment"
- "stable prices"
- "strongly committed"
- "price stability"
- "strong labor market conditions"
- "expanding"
- "above expectations"
- "robust"
- "picked up"

**Negative Tokens:**

- "hardship"
- "high inflation"
- "tightened"
- "reducing"
- "weighing"
- "cooled"
- "well below"
- "restrictive"
- "headwinds"
- "tighter credit conditions"

**Conclusion:**  
Based on this general analysis, there seems to be an even balance between positive and negative tokens in the statement. However, it's important to note that this is a high-level analysis, and a more detailed sentiment analysis using a specialized tool or software that integrates a specific sentiment dictionary would provide a more accurate assessment.

**Margin to Improve with Financial Dictionary !!!**

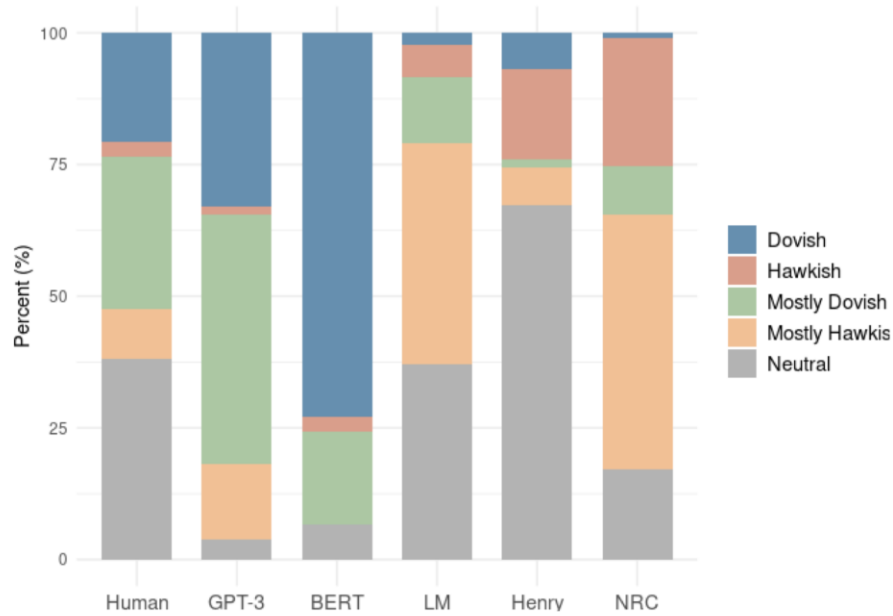


Figure 1: Distribution of categories by method

Hansen & Kazinnink (2023) Can ChatGPT Decipher Fedsspeak?

# ... Summing Up: How can LLM Generative Models can help us in our daily lives ...

Category	Task	Usefulness
Ideation and Feedback	Brainstorming	●
	Feedback	⦿
	Providing counterarguments	⦿
Writing	Synthesizing text	●
	Editing text	●
	Evaluating text	●
	Generating catchy titles & headlines	●
	Generating tweets to promote a paper	●
Background Research	Summarizing Text	●
	Literature Research	○
	Formatting References	●
	Translating Text	●
Coding	Explaining Concepts	⦿
	Writing code	⦿
	Explaining code	⦿
	Translating code	●
	Debugging code	⦿
Data Analysis	Creating figures	⦿
	Extracting data from text	●
	Reformatting data	●
	Classifying and scoring text	⦿
	Extracting sentiment	⦿
Math	Simulating human subjects	⦿
	Setting up models	⦿
	Deriving equations	○
	Explaining models	⦿

The third column reports my subjective rating of LLM capabilities as of September 2023:

○: experimental; results are inconsistent and require significant human oversight

⦿: useful; requires oversight but will likely save you time

●: highly useful; incorporating this into your workflow will save you time

## Ideation and Feedback (generating, selecting, developing of ideas):

- Brainstorming ideas
- Offering feedback
- Providing counterarguments
- Evaluating ideas

## Writing:

- Synthesizing text
- Providing examples
- Evaluating text
- Editing text (for mistakes, style, clarity, simplicity, ...)
- Generating catchy titles & headlines
- Generating tweets / promotional materials

## Math (LaTeX):

- Setting up models
- Deriving equations
- Explaining models

## Background Research:

- Researching the literature
- Summarizing text
- Formatting references
- Translating text
- Explaining concepts

## Coding:

- Writing code
- Explaining code
- Translating code
- Debugging code

## Data Analysis:

- Creating figures
- Extracting data from text
- Reformatting data
- Classifying and scoring categories
- Extracting sentiment
- Simulating human subjects

<b>3 – green</b>	capability is very well-developed & quite robust, highly useful – recommended
<b>2 – yellow</b>	capability works in limited circumstances, occasionally useful – good to experiment
<b>1 – red</b>	capability is emerging, useful with serious limitations – experiment at your own risk



# Conclusions

---

## “Tentative” Conclusions: The Medium-Term Outlook

- Some of the NLP Models are very useful: **Sometimes is better to focus on specific than generalist LLM Models**
- Large Language Models (LLM) will become better and better assistants and Tutors reducing cognitive tasks
- There is a margin to Improve with Fine-Tuning and they will be adapted to different use cases (Customization)
- It can take some time but LLM Models will enhance productivity of analysts and has the potential to change how we conduct Research

# The Use of Natural Language & Large Language Models for Analysis: Economics, Markets & Geopolitics

---

Alvaro Ortiz  
Head of Big Data Analysis  
BBVA Research

Predicción Económica Con Big Data (Funcas)

November 7<sup>th</sup>, 2023