Data Driven Public Affairs
bridge_ci

About Me

- Data Science & Policy Adviser - EU Commission, 2010
- Data Science & Strategy Adviser - EU Parliament, EPP Group, 2011
- Head of Intelligence & Analytics – Intrasoft-International, 2012-2014
- Director of Analytics, Walmart, 2014-2018
- Head of Product, Machine Learning - Wells Fargo, 2018
- Global Head of Data Science Innovation – HSBC, 2018-2020
- Lecturers – Northwestern Kellogg School of Business, Turing Institute, University of Birmingham, Columbia Business School, Tel Aviv University, London Business School.
- Chief Data Scientist/Operating Partner – Datavest Partners, 2020—2021
- Special Advisor – America's Frontier Fund, 2023-present
- Senior Advisor – Alvarez and Marsal, 2023-present
- Founder – bridge_ci, 2020- present

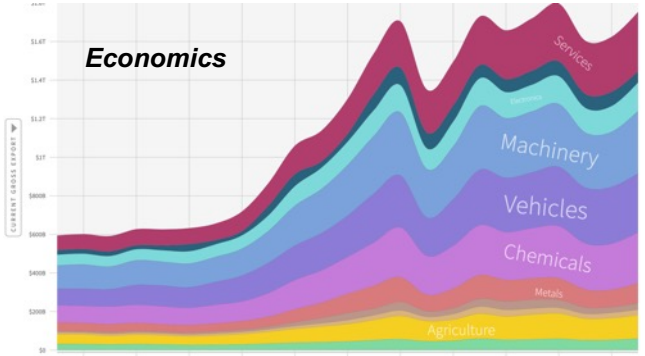
Contact ct@bridgeci.co
bridgeci.co
https://osintdrip.substack.com/

OSINT + Alternative Data (?)

- **OSINT** = News articles, 10k, earnings, social media, economic, markets, and Google search trends. In short, anything freely available.
- **Alternative data** = data that is often harmonized, formatted, and then for further insights or processing. E.g., Credit card purchases and product category, website visits of competition or vendors, product pricing.
- **Internal data** = campaign data, sales, marketing analytics & ad spend, budgets/asset allocations.

10Ks

Economics

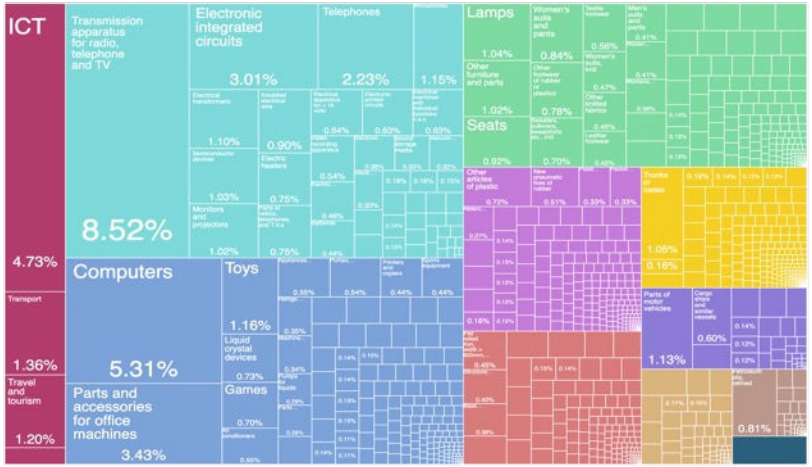


Online News

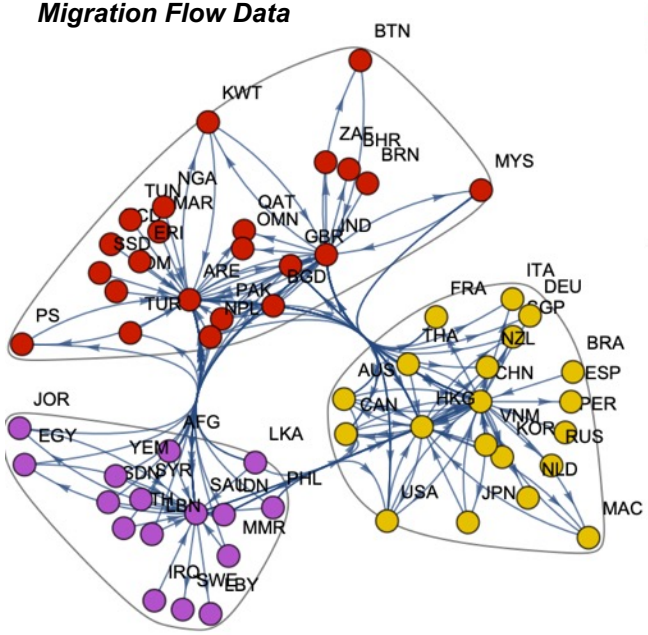
China Exports Rebound

Exports from China rose unexpectedly by 1.1% year-on-year to USD 213.85 billion in May, recovering from a 2.7% decline in April and defying market expectations of a 3.8% fall. The rebound in overseas sales came in amid efforts from companies to rush out shipments to avoid higher US tariffs that US President Donald Trump is threatening to impose in a rapidly escalating trade conflict.

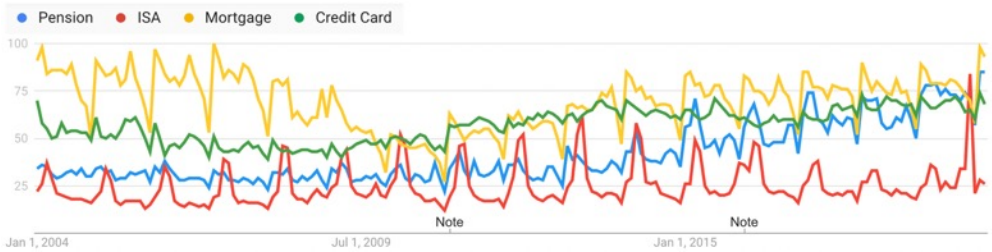
Trade

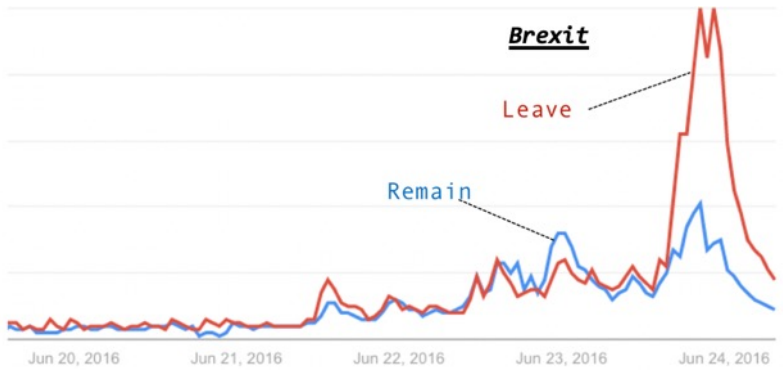


Migration Flow Data



Google Search Trends

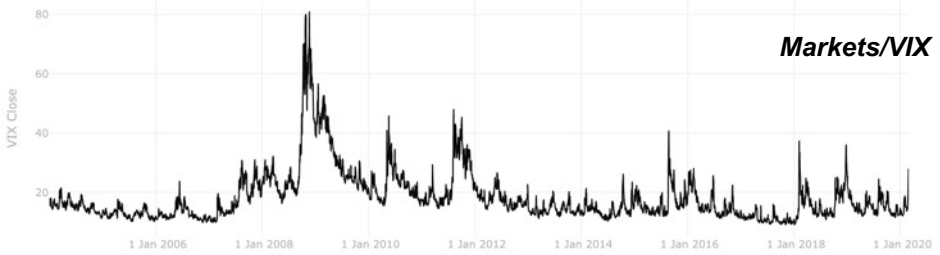




Earnings Reports

Analyst Claire Kane continued: "Our detailed review of HSBC's cost base leads us to conclude that management has yet to address the structural inefficiencies within the group. This presents a key opportunity for the new management team to drive return on tangible equity above 11% by the 2022 full year, leading to a re-rating of the shares."

Markets/VIX

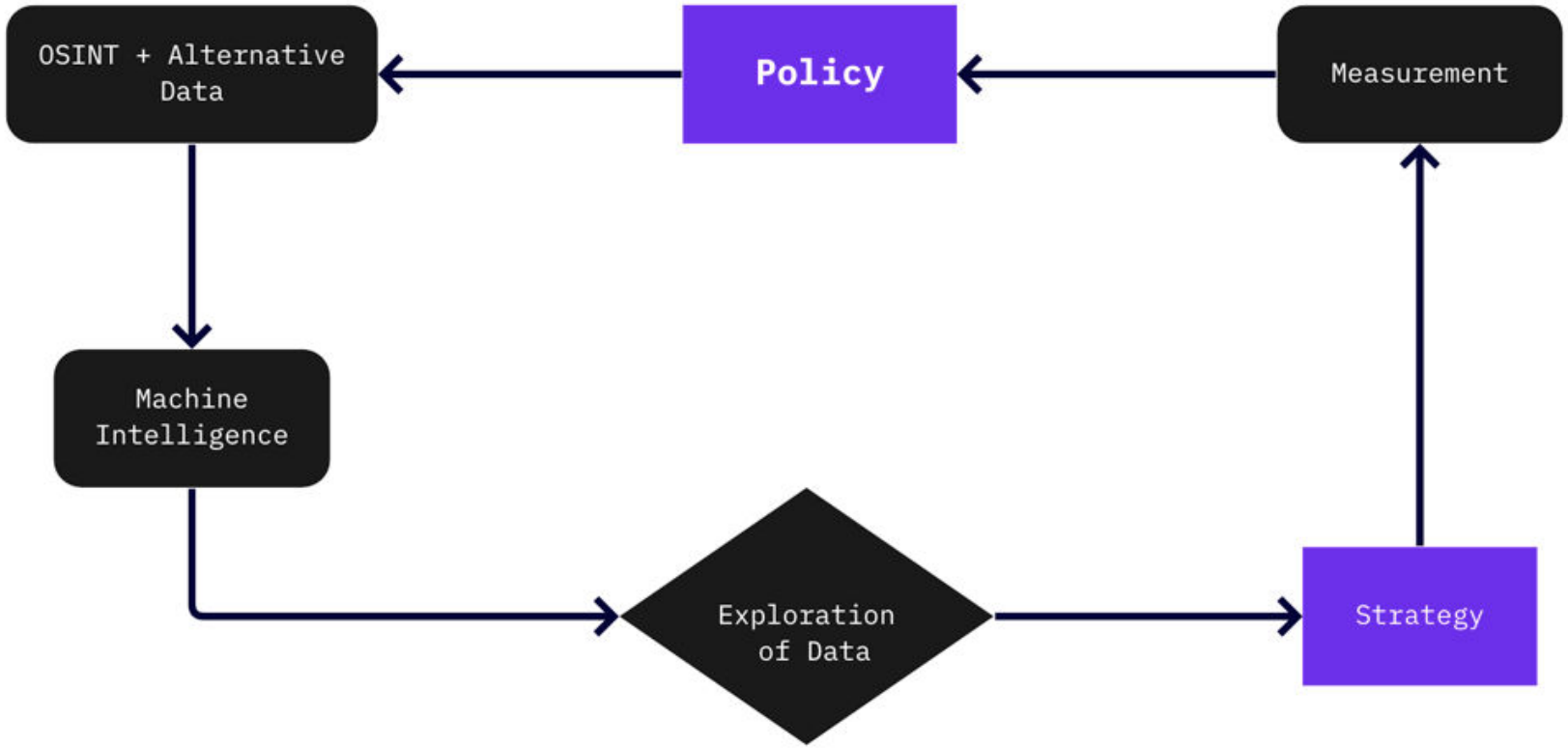


Data Driven Strategy

A domain analysis™ leverages machine intelligence, OSINT and alternative data to extract insights on any policy, institution, policy makers, or organization from millions of data points - before making assumptions about what narratives and trends are influential.

Advantage

- See all the intelligence on topic or issue structured in the context of the information available, not just a few news articles or data points.
- Remove biases - analytical model and measure political, regulatory, market, or social trends.
- “Soft” trends, such as political, regulatory, or social risk, are measured analytically rather than by intuition.

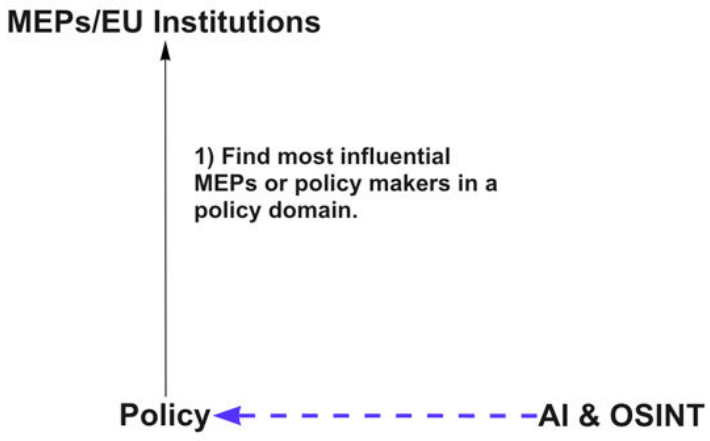


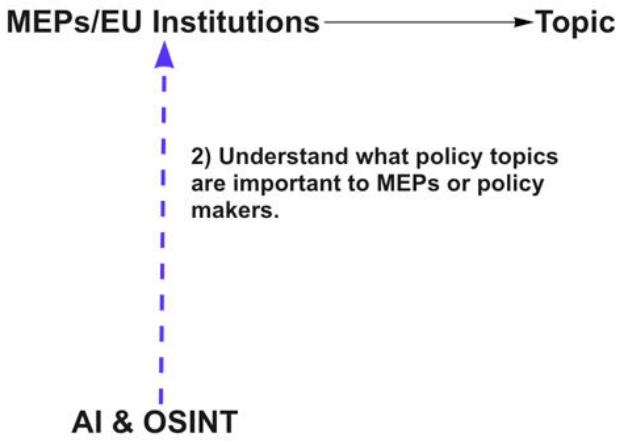
Next Generation Public Affairs Strategies

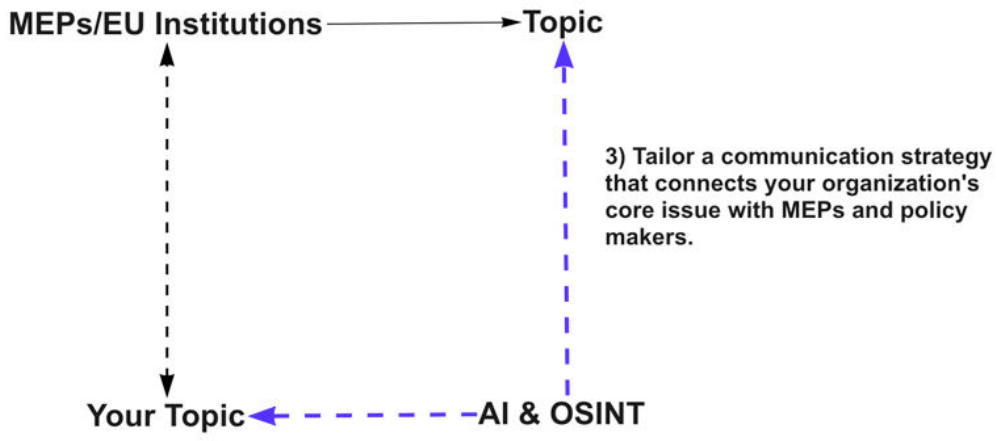
With AI and alternative data systems organizations can look at the world at 1000x faster speeds, then focus on higher value & creative tasks.

Outcomes

•Real-time monitoring of MEPs, institutions, briefings, memos, policy areas, and the current “trending” policy areas.
•More empathetic and efficient communications.
•Get ahead of risks by seeing through the noise others must wade through.
•Quantitatively rank of key topics in any member state, MEP or institution.
•Find latent adjacent policies and market trends that are influential.







MEP or Institutional Twitter accounts, policy papers or blog post can be analyzed with natural language processing with mathematical precision to create engagement strategies which are empathetic to those very MEPs, policy makers or institutions, or used as hedge against regulatory policy which are not aligned with organizational interest.

Machine Learning

Types of Data

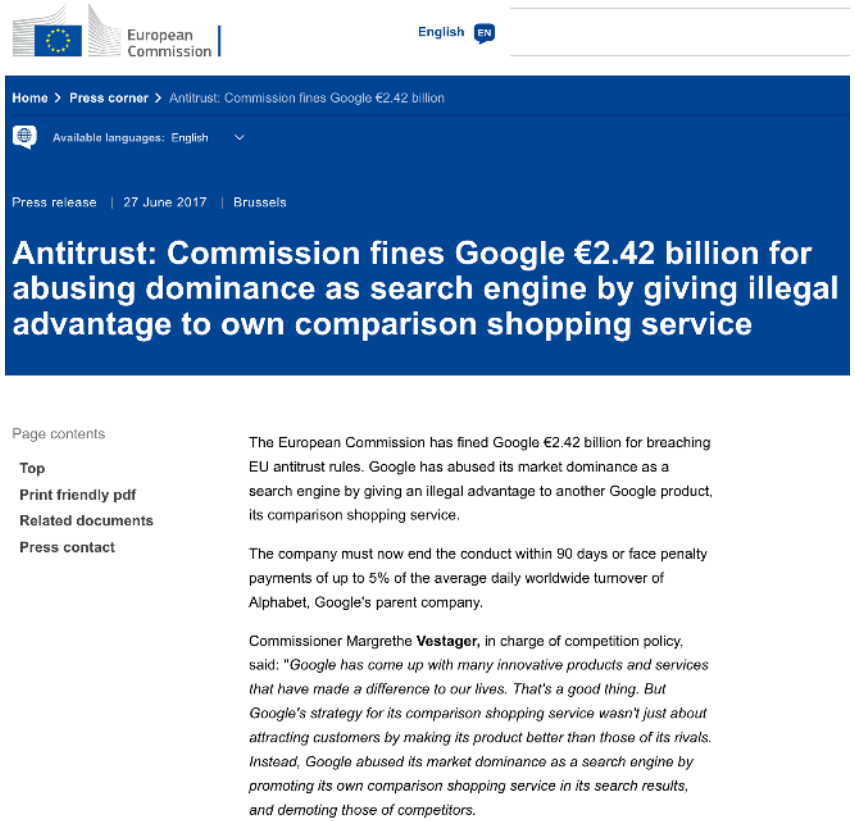
- **Structured data** is organized into a predefined format and typically stored in databases, spreadsheets, and tables. Structured data machine learning uses algorithms such as neural networks and decision trees to classify items into categories or predict future trends. By leveraging predictive analytics, businesses can make informed decisions based on past patterns in the data.

- **Unstructured data** can come in many forms, including text files, images, videos, audio recordings, social media posts, and other types of digital content. This data is often generated in real-time and in large quantities, making it difficult to manage and analyze using traditional database management systems.

Structured data

days POSIXct	Transatlanticism # Numeric	Collective security # Numeric	Multilateralism # Numeric	Nationalism # Numeric	Global Trade # Numeric
2022-02-22 05:00:00 UTC	1	0	1	3	3
2022-02-23 05:00:00 UTC	2	1	0	3	7
2022-02-24 05:00:00 UTC	18	0	2	4	8
2022-02-25 05:00:00 UTC	4	0	7	14	9
2022-02-26 05:00:00 UTC	0	0	1	2	1
2022-02-27 05:00:00 UTC	0	0	0	4	5
2022-02-28 05:00:00 UTC	1	0	2	12	5
2022-03-01 05:00:00 UTC	2	1	2	12	4
2022-03-02 05:00:00 UTC	1	7	1	25	13
2022-03-03 05:00:00 UTC	0	0	13	9	5
2022-03-04 05:00:00 UTC	31	0	2	9	4
2022-03-05 05:00:00 UTC	1	0	1	8	4
2022-03-06 05:00:00 UTC	1	0	0	5	37
2022-03-07 05:00:00 UTC	0	0	4	5	4

Unstructured data



European Commission

English EN

Home > Press corner > Antitrust: Commission fines Google €2.42 billion

Available languages: English

Press release | 27 June 2017 | Brussels

Antitrust: Commission fines Google €2.42 billion for abusing dominance as search engine by giving illegal advantage to own comparison shopping service

Page contents

Top
Print friendly pdf
Related documents
Press contact

The European Commission has fined Google €2.42 billion for breaching EU antitrust rules. Google has abused its market dominance as a search engine by giving an illegal advantage to another Google product, its comparison shopping service.

The company must now end the conduct within 90 days or face penalty payments of up to 5% of the average daily worldwide turnover of Alphabet, Google's parent company.

Commissioner Margrethe **Vestager,** in charge of competition policy, said: "*Google has come up with many innovative products and services that have made a difference to our lives. That's a good thing. But Google's strategy for its comparison shopping service wasn't just about attracting customers by making its product better than those of its rivals. Instead, Google abused its market dominance as a search engine by promoting its own comparison shopping service in its search results, and demoting those of competitors.*

Machine Learning

1.Collect and clean the data: This step involves collecting data relevant to the problem you are trying to solve, and then cleaning and preprocessing the data to get it ready for analysis.

2.Split the data into training and test sets: Once the data is cleaned and prepared, it is typically split into two sets: a training set and a test set. The training set is used to train a machine learning model, while the test set is used to evaluate the performance of the model.

3.Train the model: During the training phase, a machine learning algorithm is used to learn from the training data. The goal is to learn the patterns and relationships present in the data in order to make predictions about new, unseen data.

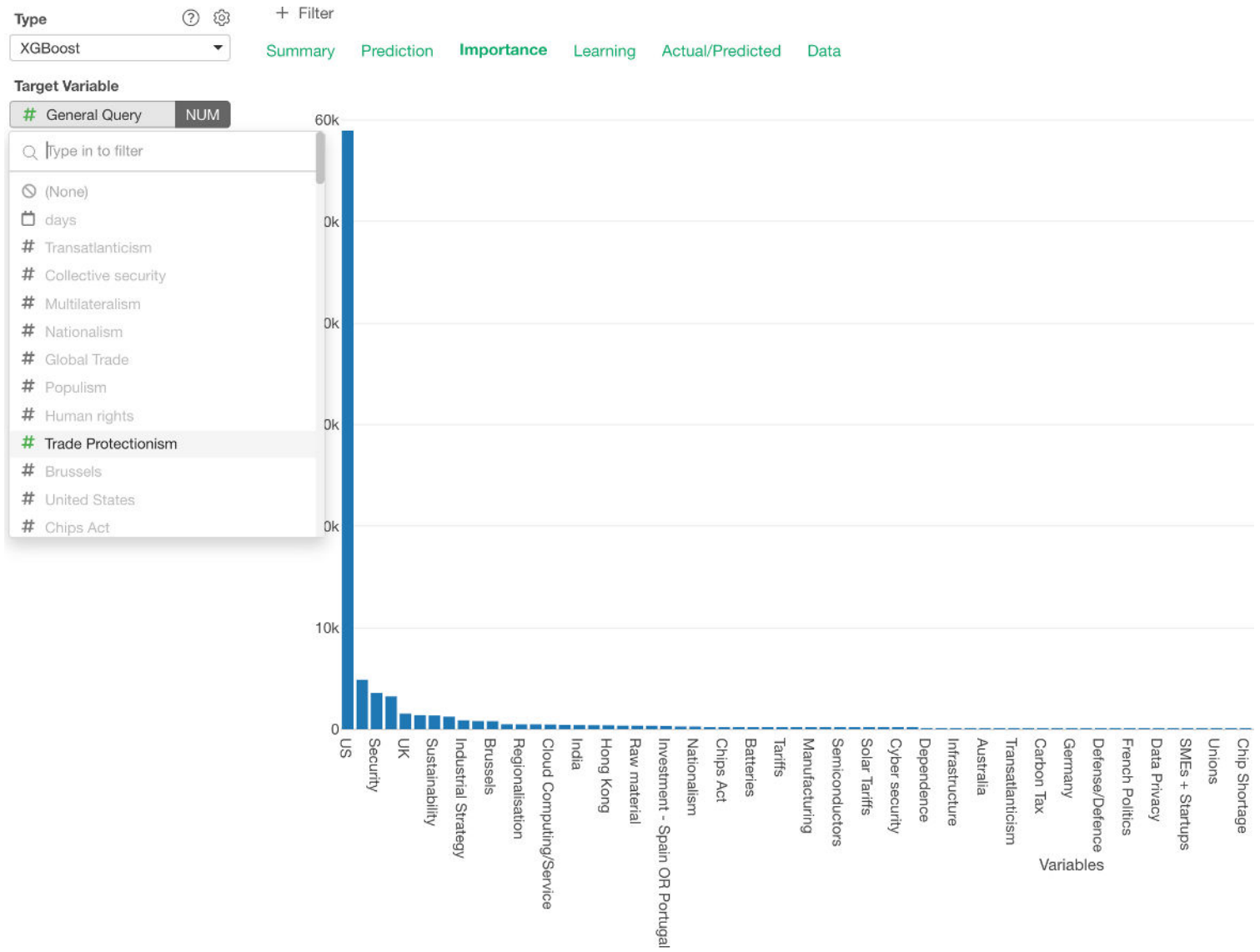
4.Evaluate the model: After the model has been trained, it is evaluated on the test set. This is done to assess the model's performance and determine how well it generalizes to new data.

5.Fine-tune the model: If the model's performance is not satisfactory, the model's hyperparameters (e.g. learning rate, regularization strength) can be adjusted and the model can be retrained. This process is known as fine-tuning the model.

6.Deploy the model: Once the model is performing satisfactorily, it can be deployed in a production environment where it can be used to make predictions on new, unseen data.

Supervised Machine Learning

SL is used when there is a specific outcome that we want to model. The algorithm iteratively makes predictions on the supplied training data. Then the learning stops when the algorithm achieves an acceptable level of performance.



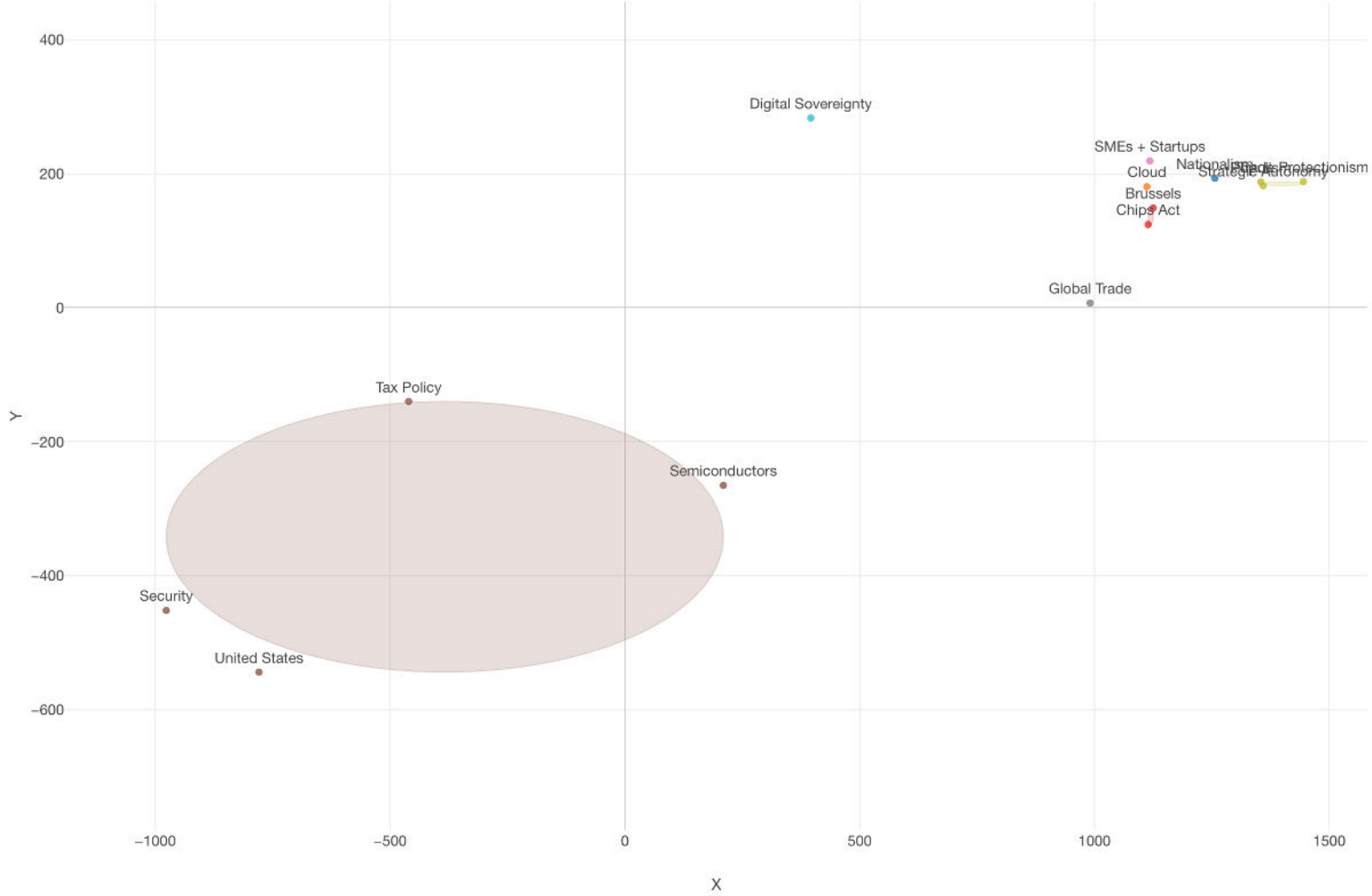
Semi- Supervised Machine Learning

SS sits between both supervised and unsupervised learning. The process starts by having an analyst or domain expert tag articles that consist of themes such as “mergers” or “stock buy backs”, thereafter leaving to train the machine to tag the same themes on their own within the rest of the data. Many real-world machine learning problems fall into this area. It can be expensive or time-consuming to label all data, as it may require access to domain experts, in addition to a lot of training time.

Title	Top Location	Start Date	Significance	Total Docs ↓	Sentiment	Event Types
UK, EU strike new Northern Ireland post-Brexit trade deal	Northern Ireland, United Kingdom	Feb 26, 2023	5.0	178	neut	Diplomacy, Poor Financial Performance, & 3 more
European Commission bans TikTok on staff devices	Brussels, Belgium	Feb 22, 2023	5.0	103	neut	Disputed Information, Outsourcing or Non-U.S. Company, & 3 more
Rishi Sunak says ‘No deal yet’ on Northern Ireland Protocol	Northern Ireland, United Kingdom	Feb 16, 2023	4.7	98	neut	Diplomacy, Disputed Information, & 2 more
Canada Bans Tiktok On All Government Devices	Beijing, China	Feb 26, 2023	4.2	54	neut	Disputed Information, Corruption (TikTok), & 3 more
UK, EU Agree On Post-Brexit Trade Agreement For Northern Ireland	Northern Ireland, United Kingdom	Feb 27, 2023	4.0	45	neut	Diplomacy, Outsourcing or Non-U.S. Company, & 1 more
Rishi Sunak To Meet Ursula Von Der Leyen For Talks On Monday	Northern Ireland, United Kingdom	Feb 26, 2023	4.5	42	neut	Strategic Partnership, Diplomacy

Unsupervised Machine Learning

US offers the most upsides of all the techniques, in addition to requiring the user to place the most trust in the machine. The goal for unsupervised learning is to model the underlying structure or distribution in the data in order to learn more about the data. It's called unsupervised learning because unlike supervised learning there is no correct answers and there is no teacher. Algorithms are left to their own devises to discover and present the interesting structure in the data. It's important to develop these capacities and trust them due to the growth rate of data and computational power. It's exceedingly unrealistic to think humans have the mental capacity to develop high quality hypothesis given the amount of signals that a modern business produces.



Natural Language Processing (NLP)

NLP enables machines that can extract information and insights contained in the documents as well as categorize and organize the documents themselves.

Natural Language Generation (NLG)

NLG is the process of creating phrases and sentences in the form of natural language with a machine. It can automatically generate narratives that describe, summarize or explain input structured data in a human-like manner at the speed of thousands of pages per second.

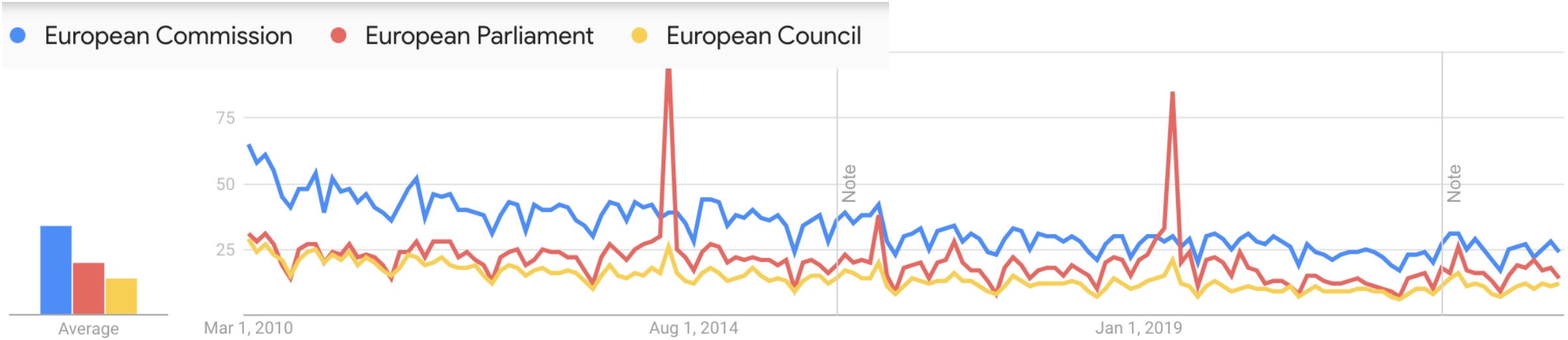
Natural Language Understanding (NLU)

NLU reads human language and turns it into structured data understandable to computers.

Google Search Data

Google Trends is a tool that provides insights into the popularity and frequency of search terms on Google. It allows users to see the relative popularity of search terms over time and across different geographic locations. Google Trends data can be used for lobbying intelligence and public affairs in a number of ways, including:

1.Identifying key issues
2.Monitoring engagement
3.Geographic targeting



Google Search Data Advantage

While traditional polling and surveys can still be valuable tools in certain situations, Google Trends data can provide valuable and complementary insights into public behavior and attitudes that may not be captured by traditional polling and surveys. By leveraging this data, organizations can make more informed decisions and optimize their communication and advocacy strategies.

1.Large sample sizes: Google Trends can analyze the behavior of a larger sample size than traditional polling or surveys, which typically rely on a smaller sample size of respondents. This can provide more accurate and representative insights into the behavior and attitudes of a larger population.

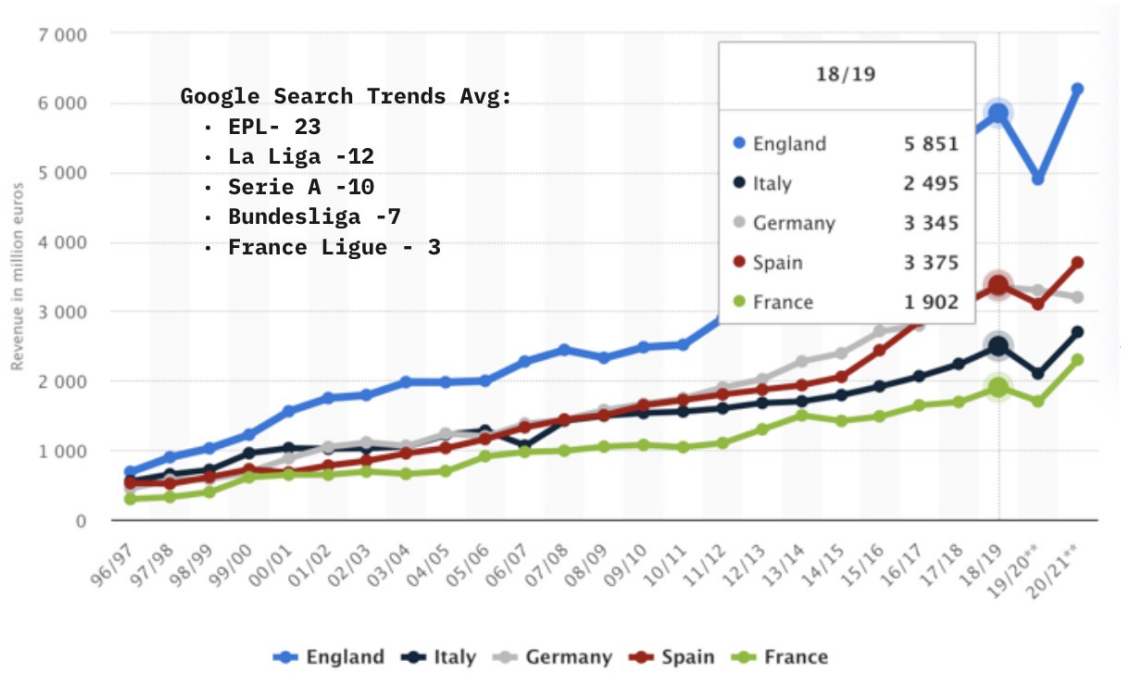
2.Real-time data: Google Trends data is updated in real-time, providing immediate insights into the latest trends and behaviors. Traditional polling and surveys can take days or weeks to collect and analyze data, which may not provide timely insights for decision-making.
3.Cost-effective: Google Trends data can be collected and analyzed at a fraction of the cost of traditional polling and surveys, making it a more cost-effective option for organizations with limited resources.

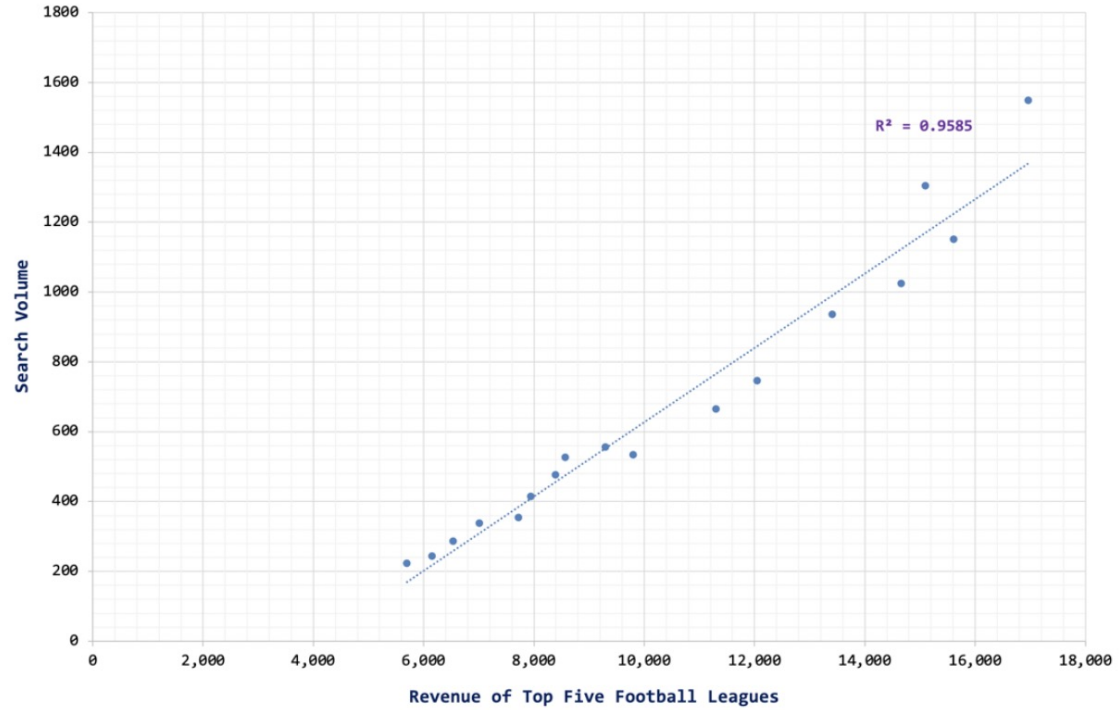
4.Unbiased data: Google Trends data is based on actual behavior, rather than self-reported opinions, which can be subject to bias or social desirability effects. This can provide more accurate and unbiased insights into public behavior and attitudes.

5.No survey fatigue: Traditional polling and surveys can lead to survey fatigue, which can reduce response rates and accuracy. Google Trends data is based on passive data collection, which does not require active participation from respondents, reducing the likelihood of survey fatigue.

Google Trends Accuracy

Using real world data to contrast search interest and results show that search is 95% correlated with revenue of the top football leagues.

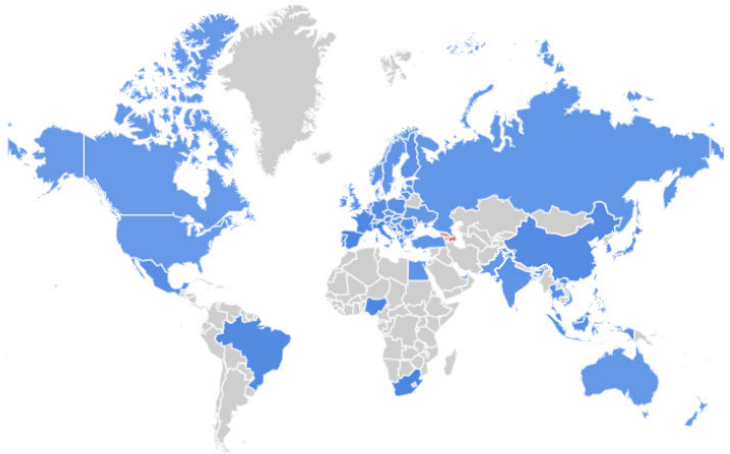




Google Search Data

Below are the key regions searching for EU institutions. The same data can be used to rank policy or trend interest to determine if the issue is populist or niche.

European Commission European Parliament European Council

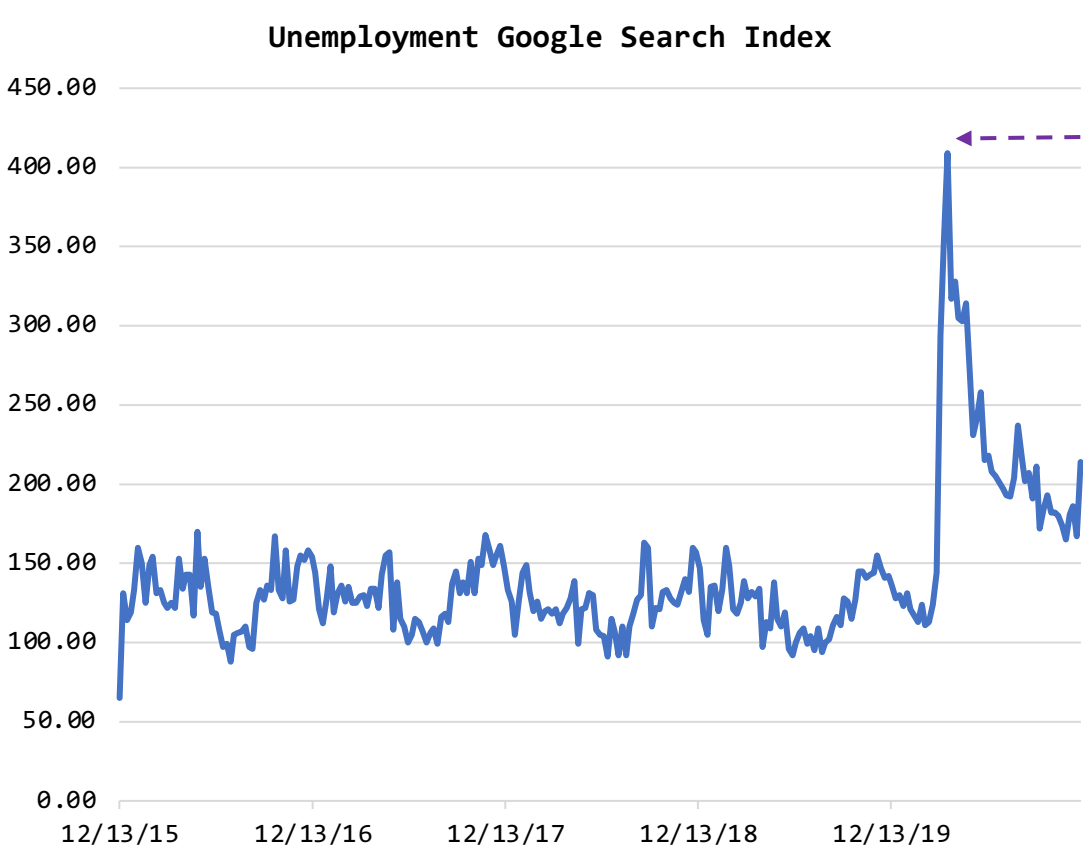


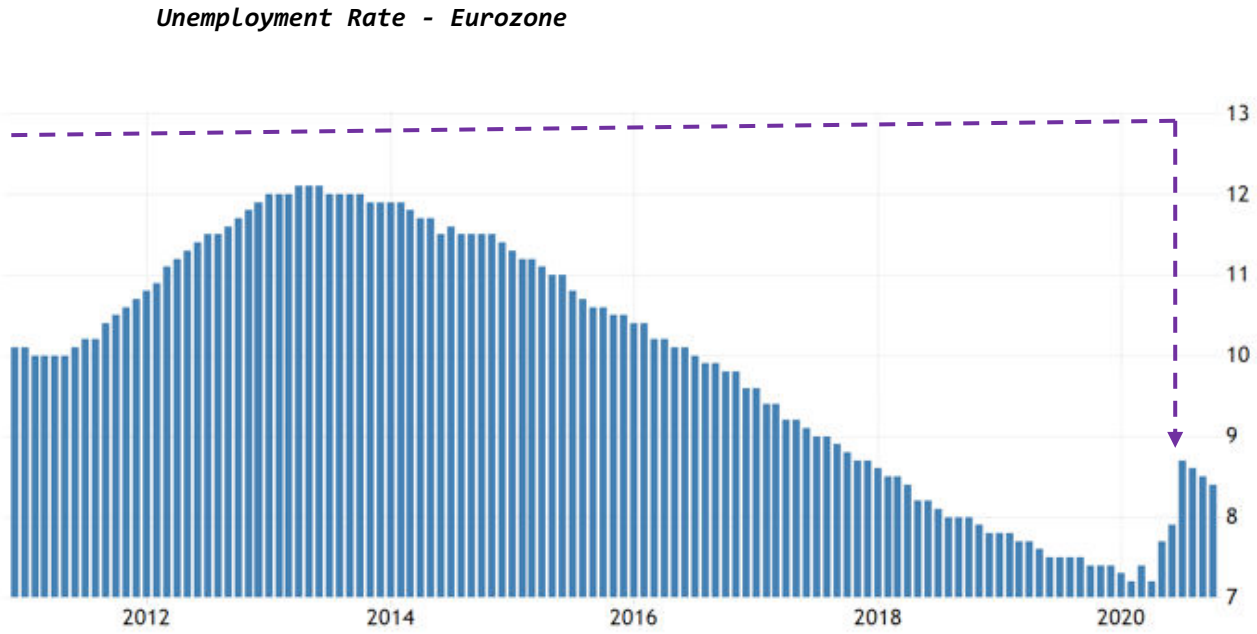
Sort: Interest for European Commission

1	Italy	
2	Israel	
3	Bosnia & Herzegovina	
4	St. Helena	
5	Spain	

G Trends – Leading Indicator

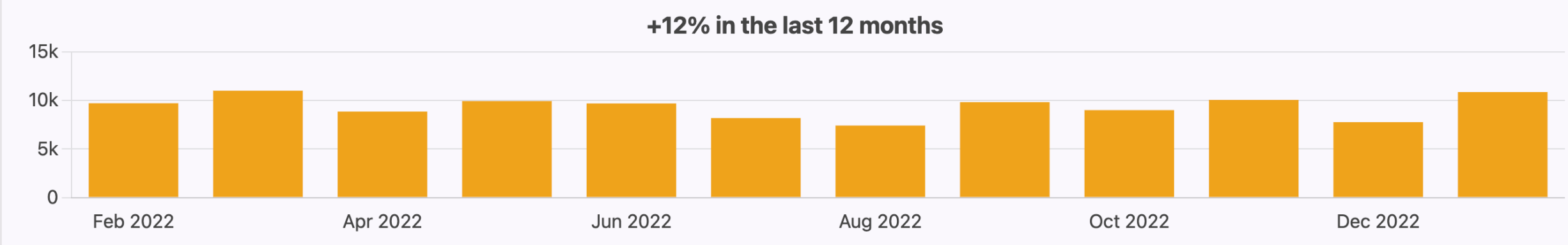
Alternative data from Google Trends can provide firms with leading, not lagging indicators about policy and macro economic trends.





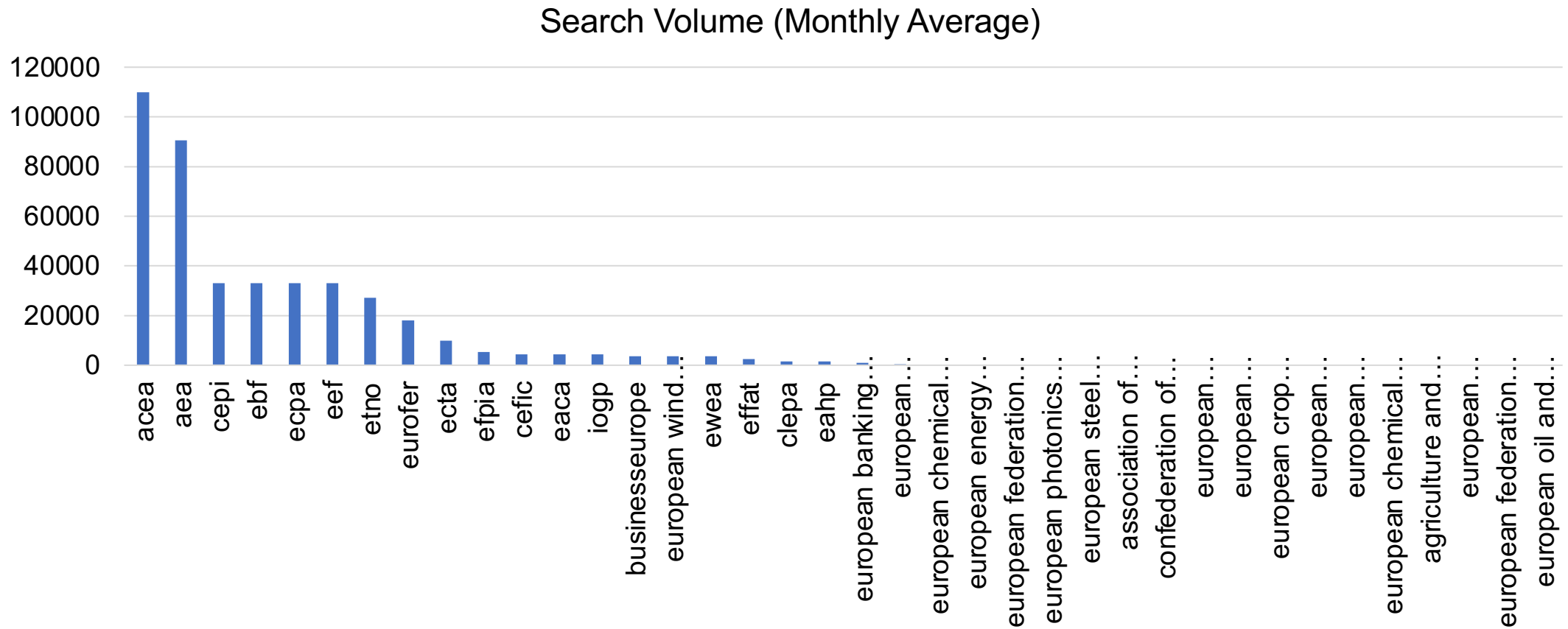
Google Search Data – Lobbying Group Indices

An aggregation of the top lobbying groups in Brussels show that there are approximately 10k search a month



Google Search Data – lobbying Groups

Google Trends is a tool that provides insights into the popularity and frequency of search terms on Google. It allows users to see the relative popularity of search terms over time and across different geographic locations. Google Trends data can be used for lobbying intelligence and public affairs in a number of ways, including:




Media Monitoring

Media Monitoring

Media monitoring tools are software applications that are designed to track and analyze media coverage across various types of media sources, such as news articles, social media posts, and broadcast media. These tools typically use natural language processing and machine learning algorithms to identify and categorize relevant content, and can provide insights into the tone and sentiment of media coverage, as well as key themes and trends.

1.Tracking media coverage: Media monitoring tools can help public affairs professionals track media coverage of their organization, competitors, and industry. By tracking media coverage across a variety of sources, they can stay informed about the latest developments and issues, as well as potential risks and opportunities.

2.Identifying key influencers: Media monitoring tools can also help public affairs professionals identify key influencers and thought leaders in their industry or area of interest. By tracking mentions and coverage of these individuals, they can identify opportunities to engage with them and potentially build relationships.

3.Measuring sentiment: Media monitoring tools can provide insights into the tone and sentiment of media coverage, which can be valuable for understanding public perception and identifying potential reputational risks.

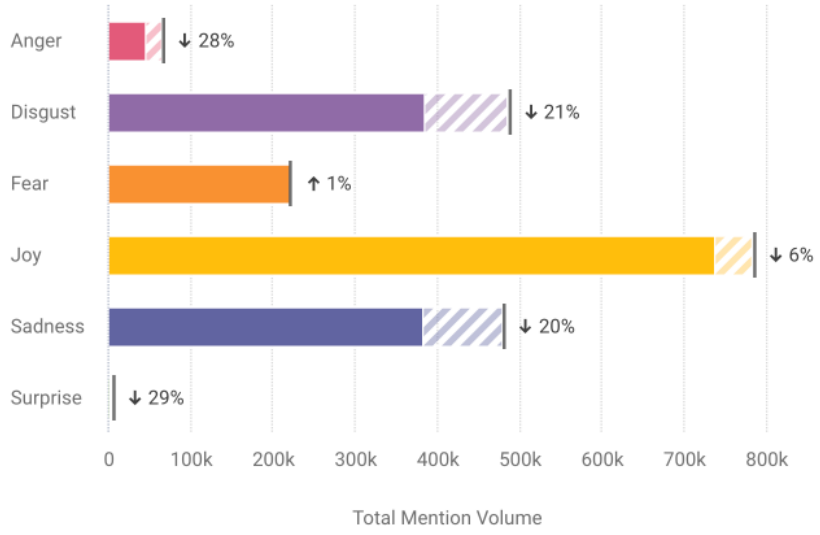
4.Monitoring trends: Media monitoring tools can also help public affairs professionals track trends and themes in media coverage over time. This can provide valuable insights into the issues and topics that are resonating with the public and can inform communication and advocacy strategies.

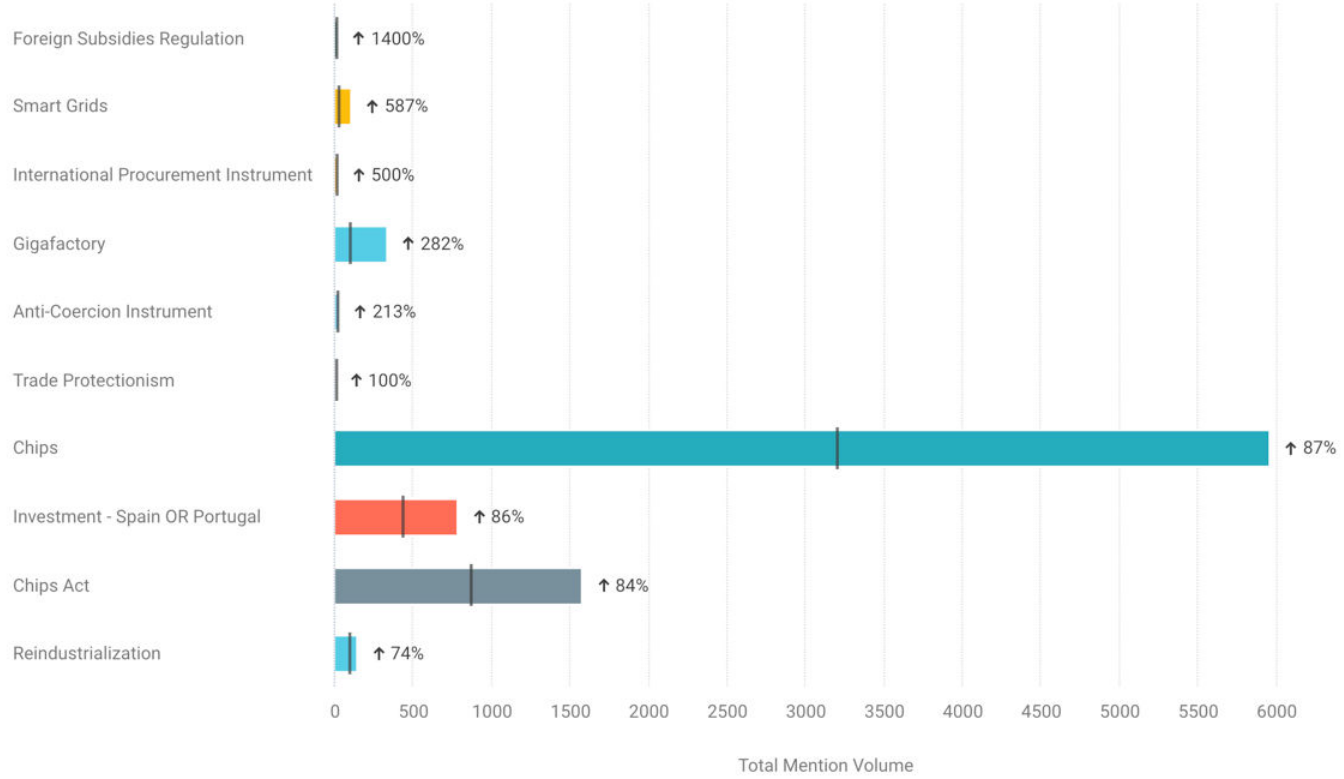
Build Categories & Classifiers

Currently, we are closely monitoring all events occurring within the European Union (EU) and utilizing automatic categorization to swiftly model influence and identify significant trends that are associated with specific policy concerns. To achieve this, we use advanced methods such as AI models trained on text to develop a classifier, as well as the use of specific keywords to build categories.

Our tracking efforts include the following categories:

- The European Commission (EC)
- The European Parliament (EP)
- Members of the European Parliament (MEPs)
- Industry Associations and Non-Governmental Organizations (NGOs)
- Key Policy Themes
- EU Member States

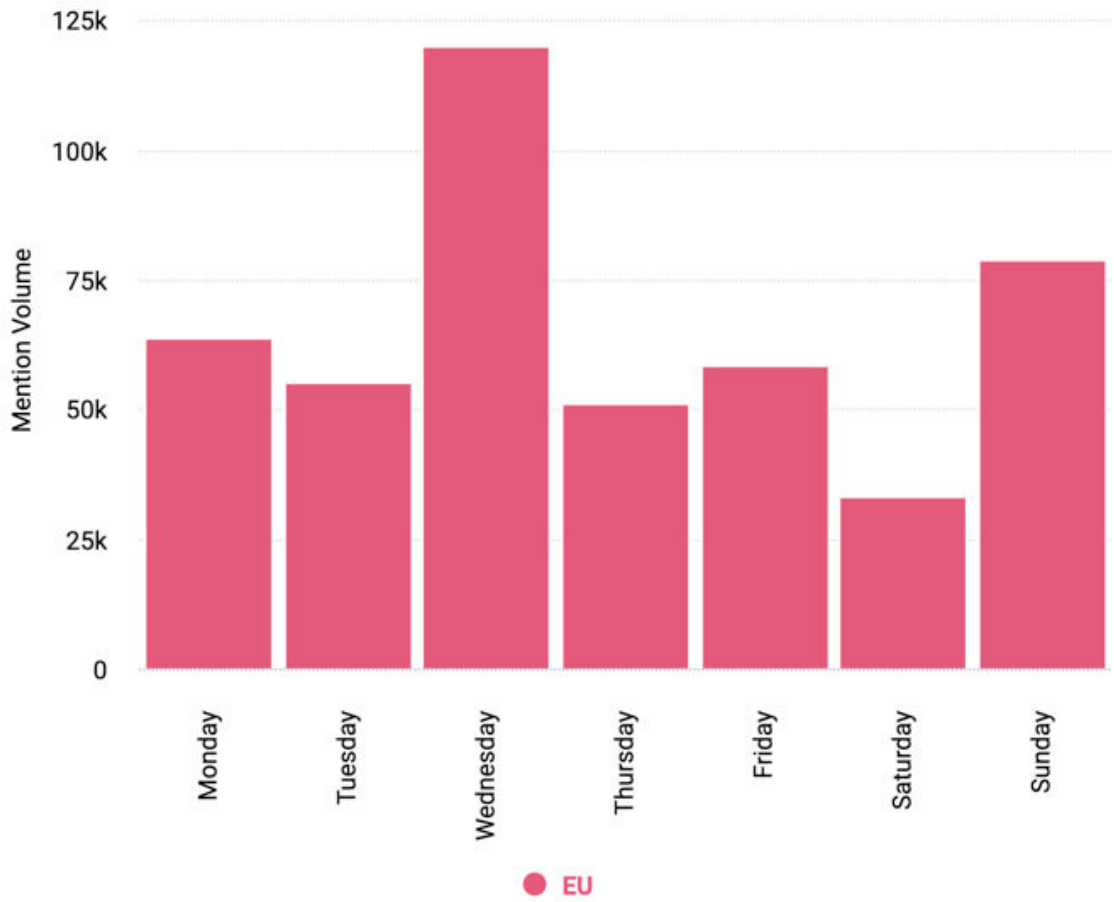




Content Optimization

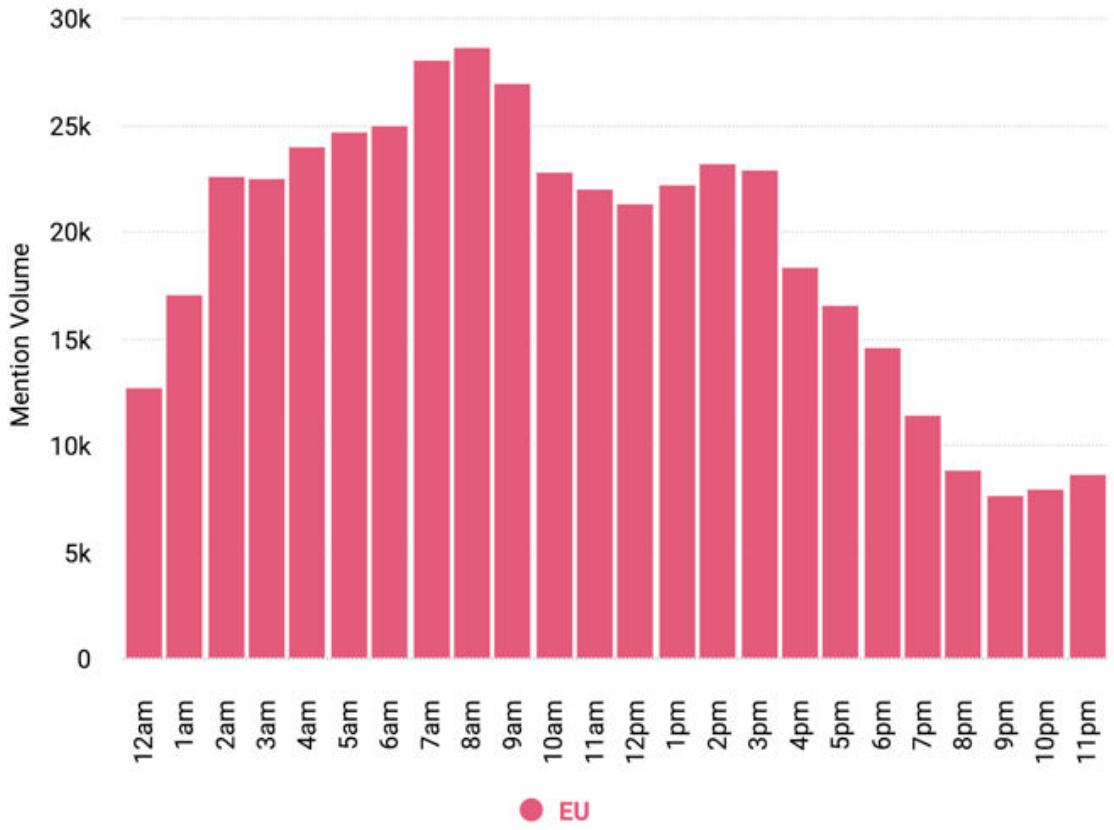
Mentions by Day of Week | EU | Last 7 days

Mention Volume for **Day of Week** broken down by **Queries**

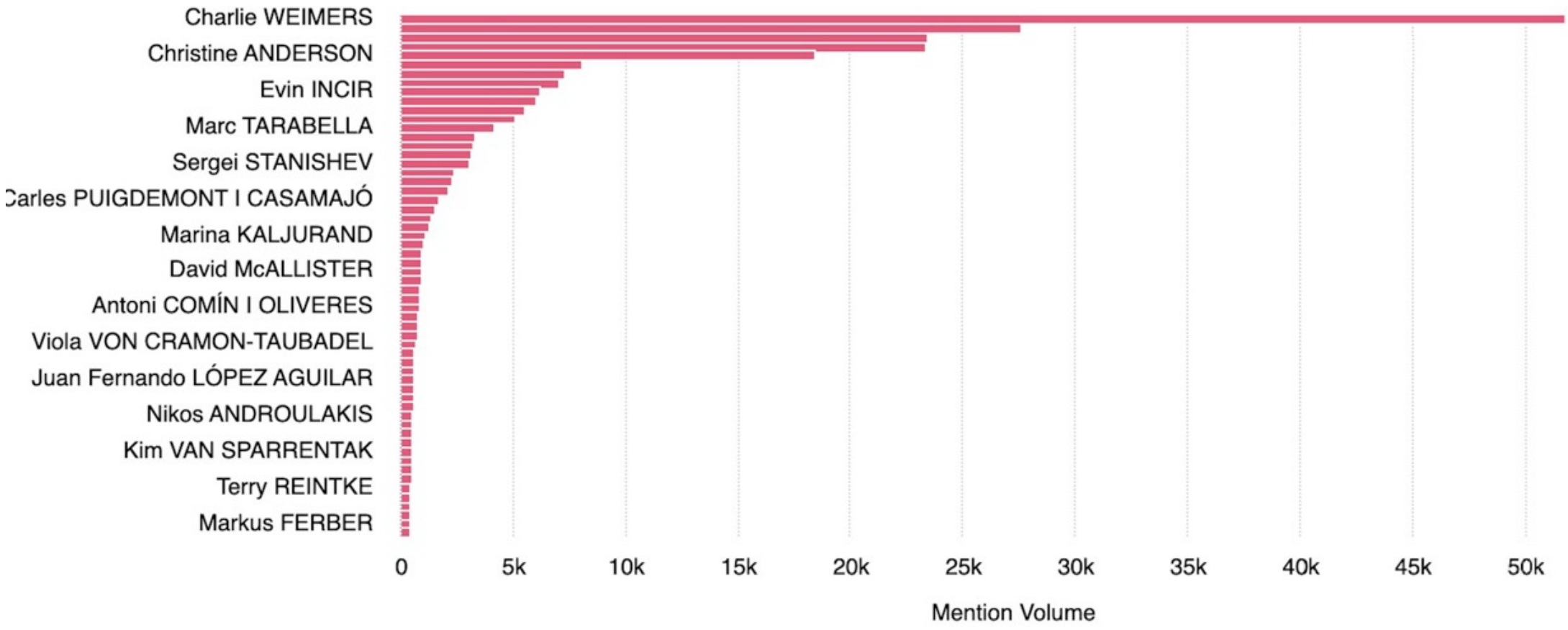


Mentions by Hour of Day | EU | Last 7 days

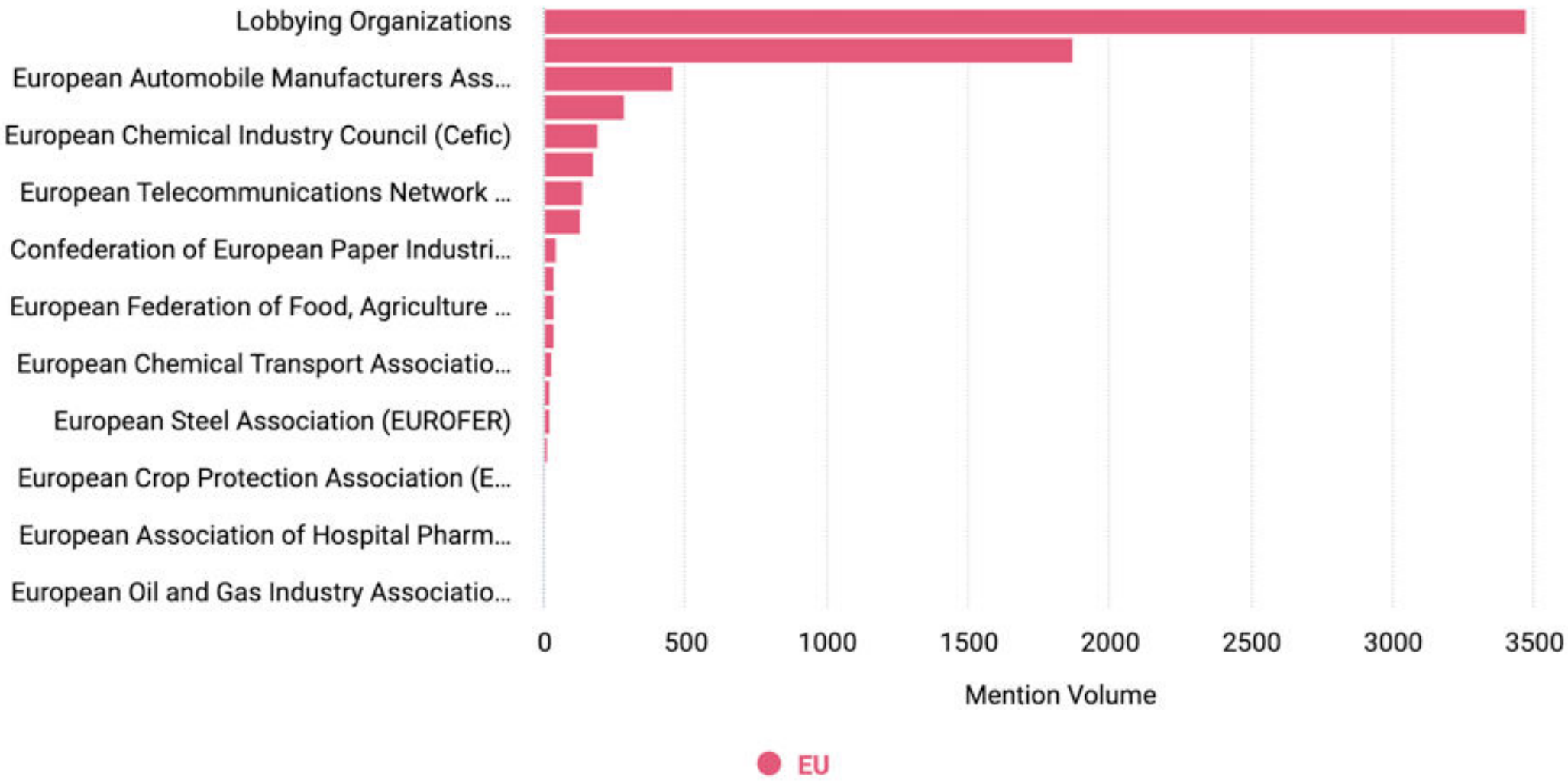
Mention Volume for **Hour of Day** broken down by **Queries**



Mentions – MEPs



Mentions – Lobbying Groups



Monitoring Metrics

1.Volume: This metric measures the total number of mentions of a brand, organization, or topic across different media channels, such as news articles, social media posts, and broadcast media.

2.Reach: This metric measures the potential audience size of media coverage. For example, the number of readers, viewers, or followers who may have seen a particular news article or social media post.

3.Sentiment: This metric measures the tone or sentiment of media coverage, such as positive, negative, or neutral. This can be useful for assessing public perception and identifying potential reputational risks.

4.Share of voice: This metric compares the volume of coverage of a brand or organization to that of its competitors or peers. It can provide insights into how a brand or organization is performing relative to its competitors in terms of media coverage.

5.Key message penetration: This metric measures the extent to which key messages or themes are being communicated in media coverage. This can help to assess the effectiveness of communication and advocacy strategies.

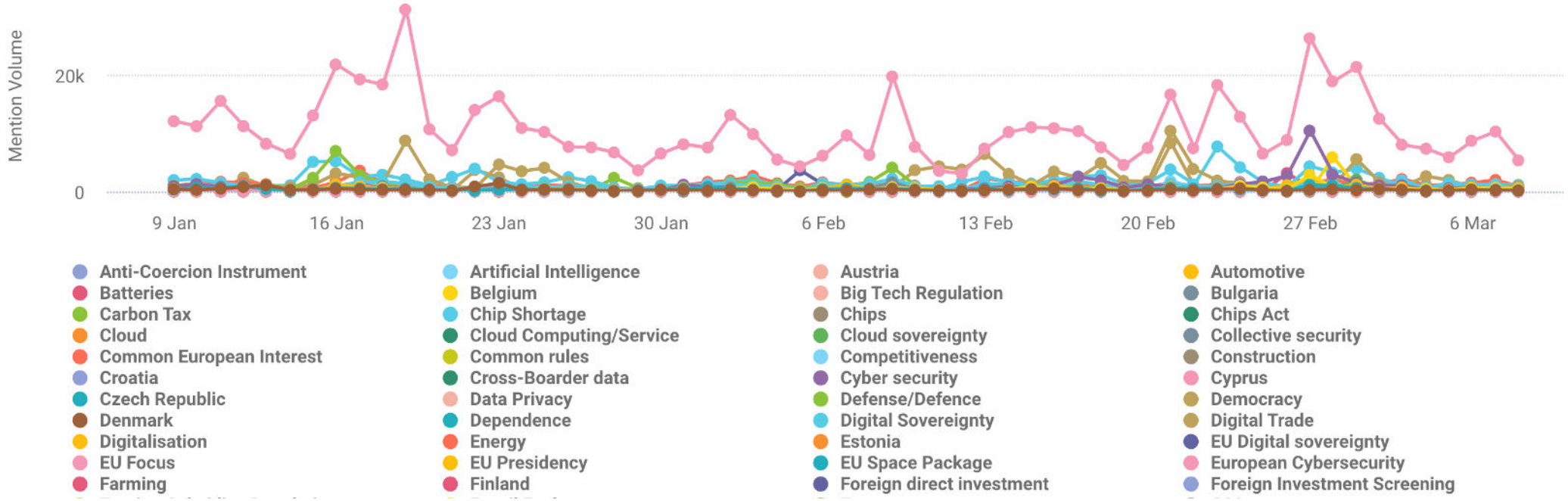
6.Influencer mentions: This metric measures the number of times key influencers or thought leaders in a particular industry or topic are mentioned in media coverage. It can help to identify potential allies and build relationships with influential stakeholders.

Policy Indices

Similar to how stock traders analyze asset and stock prices, pre-built classifiers can be used to create what is known as "policy indices." These indices monitor online mentions related to a particular policy, allowing public affairs teams to rank and model emerging trends. The data used to build these indices can be sourced from a variety of platforms including Twitter, news outlets, blogs, forums like Reddit, and Google Search Trends. Overall, the use of policy indices is an effective tool for staying ahead of public affairs issues and ensuring that policies are effective, relevant, and aligned with public sentiment.

The use of policy indices provides several advantages in tracking issues related to public affairs.

- Enable policymakers and stakeholders to stay up-to-date with emerging trends and public sentiment related to a particular policy. The information can be used to make more informed decisions and adjust policies accordingly.
- identify potential issues or controversies before they escalate, allowing public affairs teams to proactively address them.
- Allow organizations to monitor and benchmark their own performance and engagement against competitors or other organizations in their industry.



Correlate Policy Issues

Taking the same data, we can apply regression to find which domains (topics) are most correlated to one another. Matrix correlation charts (like the one below) are often used in finance and economics to understand the relationships between different variables and to make informed decisions based on the data. They can be especially useful for identifying trends and patterns in large datasets and help researchers and analysts better understand the underlying relationships between dozens of variables (think multiple columns in Excel).

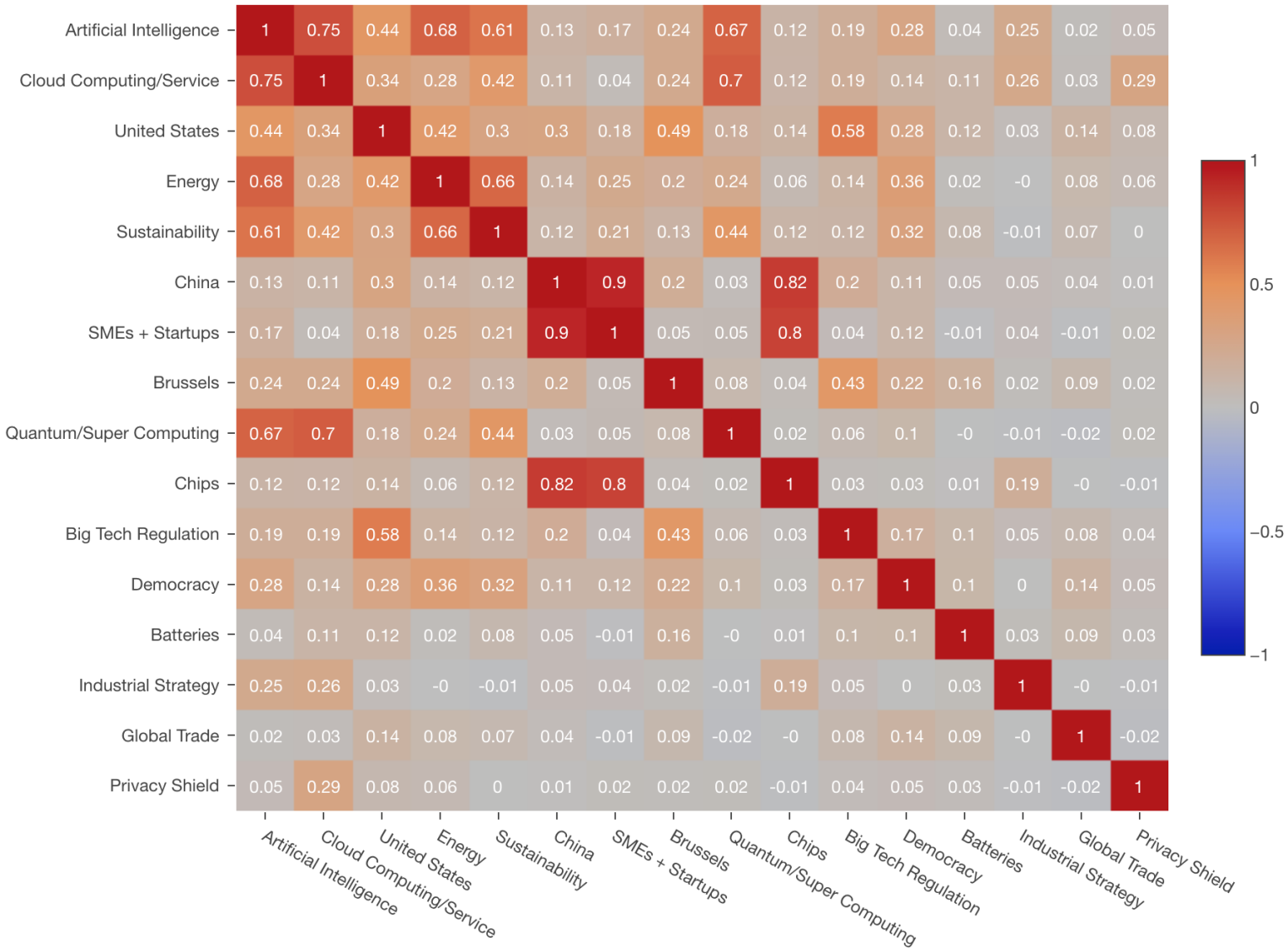
The correlation coefficient measures the strength and direction of the relationship between two variables and can range from -1 to 1.

•A value of 1 indicates a perfect positive correlation, meaning that the two variables are strongly related and change in the same direction.

•A value of -1 indicates a perfect negative correlation, meaning that the two variables are strongly related and change in opposite directions.

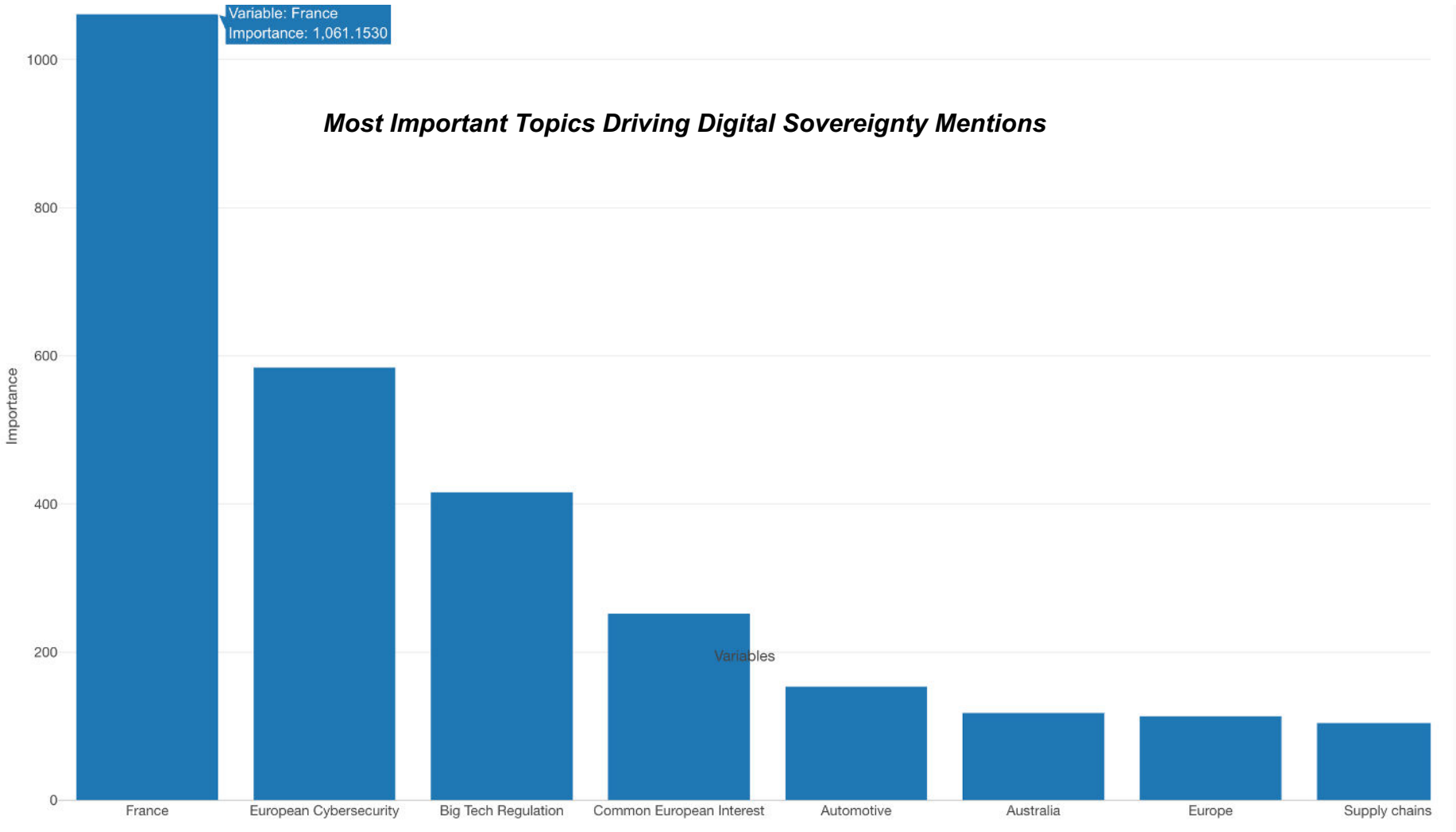
•A value of 0 indicates no correlation.

Policy, Regional & Thematic Correlations



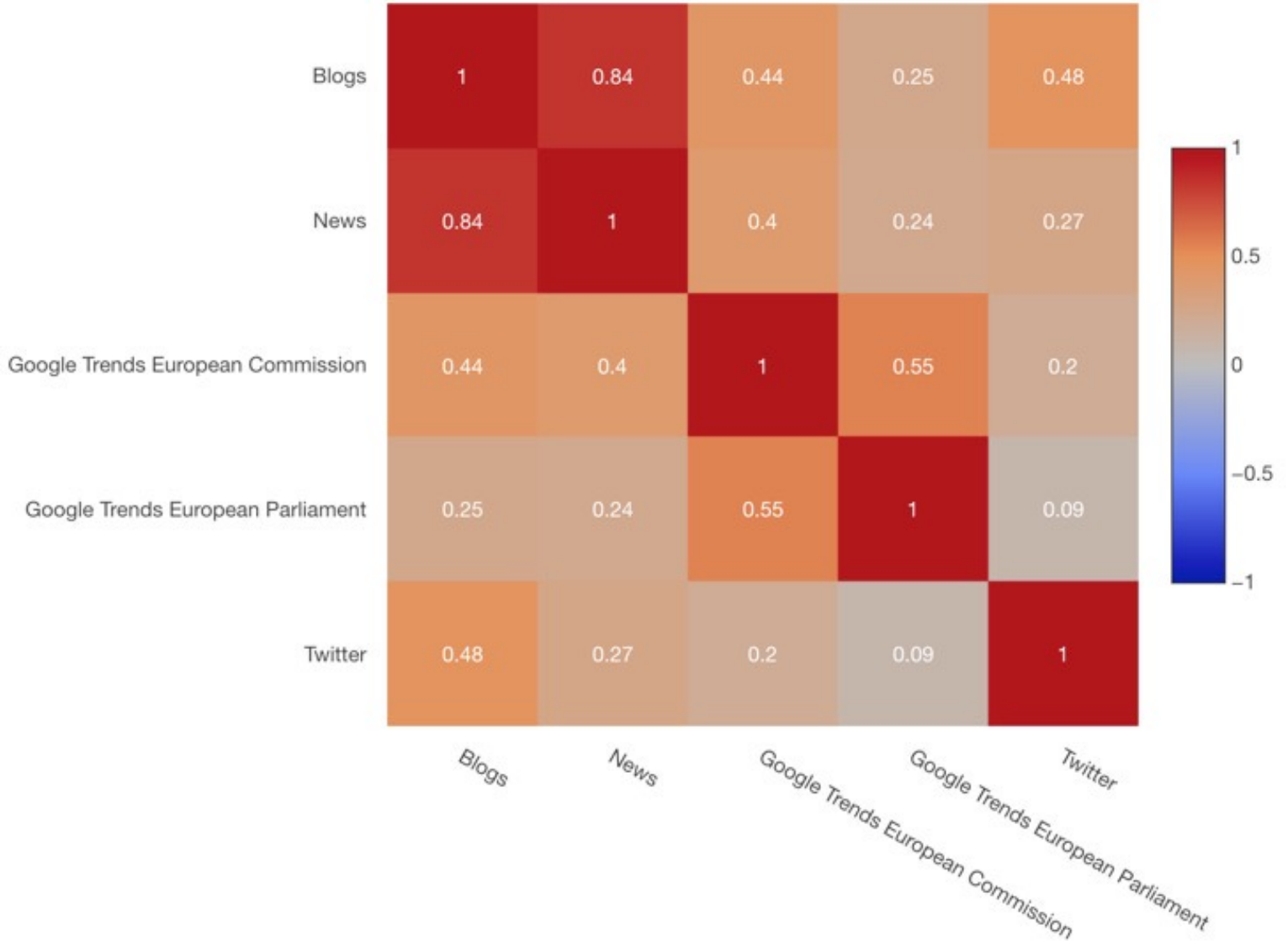
Isolate Key Drivers

Machine learning can be applied to the same data to model which topic is the main driver of the issue. For example, the left graph below shows the US, Chips and China are the driving thematics within Digital Sovereignty .



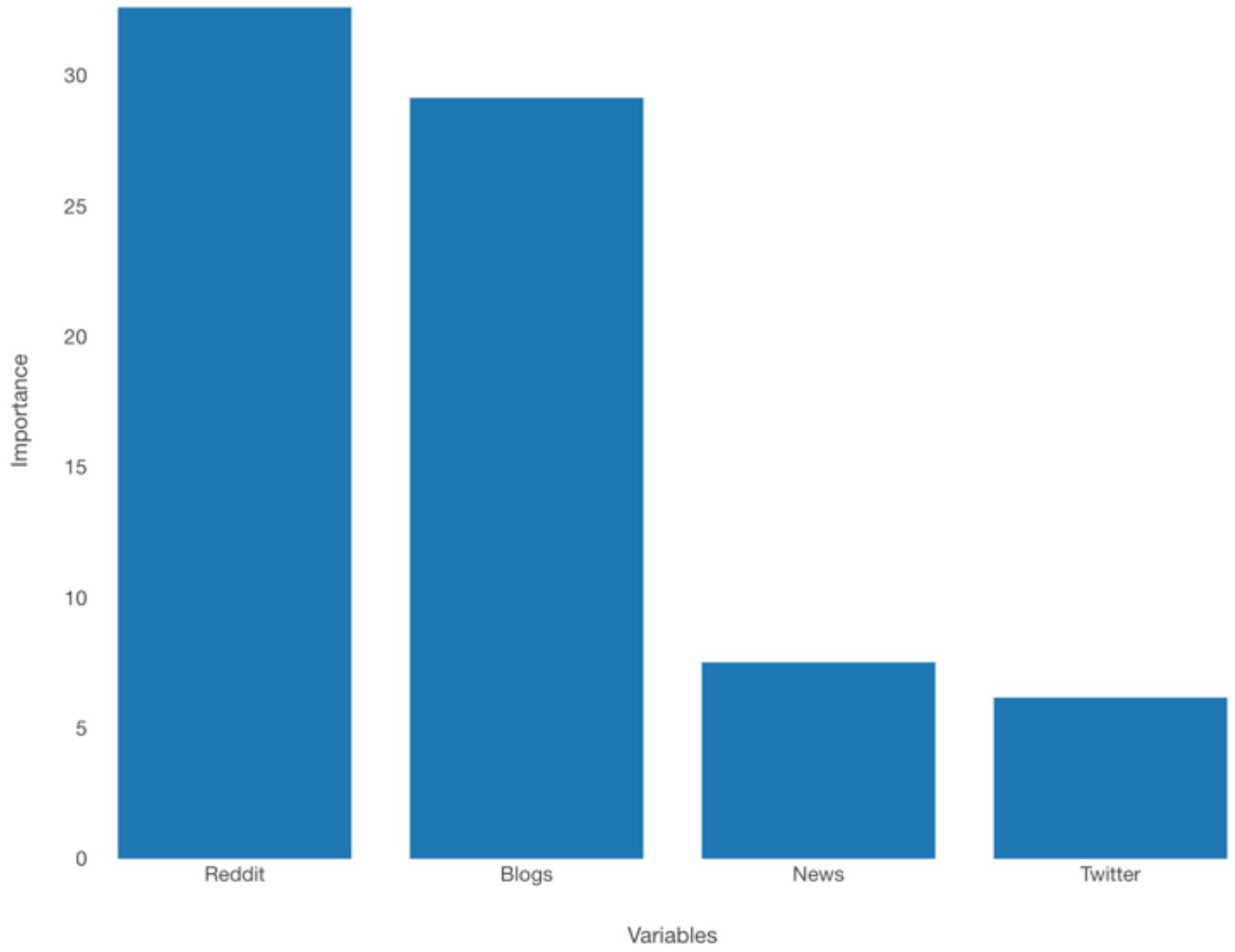
Correlate Channels

The same data can be used to correlated different communications channels to see which drive wanted outcomes or reputational risks.



Calculate Channel Influence

Just like the policy indices, Machine learning can be applied to the same data to model which channel is the main driver of engagement.

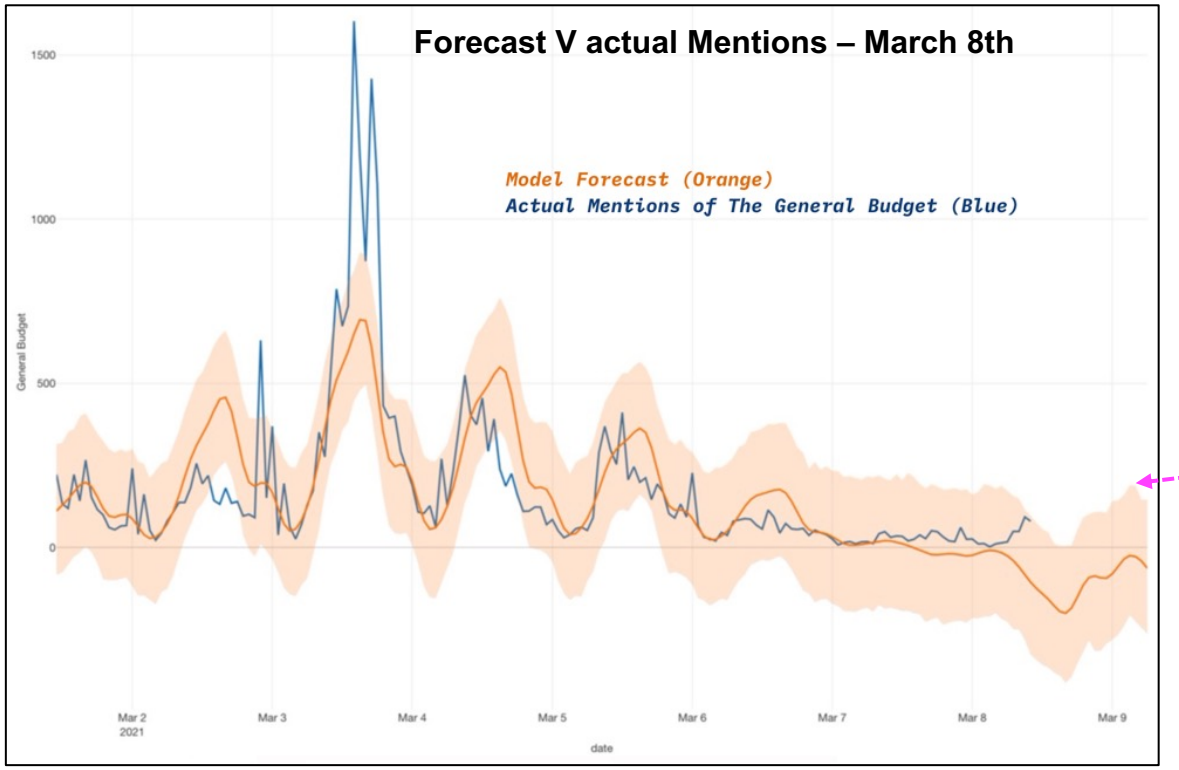


Forecast Future Mentions

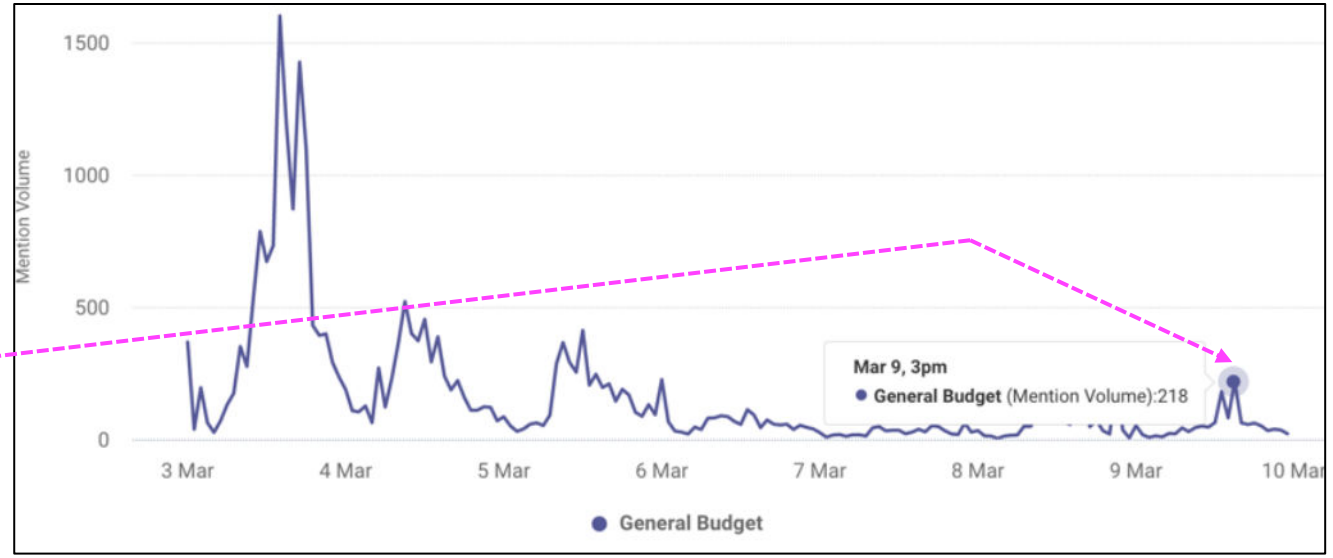
Opposed to guessing what is “smoke” or “fire”, algorithms can model momentum using variables such as volume, sentiment and channel (Twitter, Web, or Google Search Trends) data to accurately predict the direction of an issue in near real-time. Below are the mentions of Rishi Sunaks’s budget speech in 2021.

1. Upper range of the model created on March 8th predicted around 210 mentions on March 9th.
2. On March 9th, which is how far out the model forecasted the outcome was 218 mentions showing the model was directionally accurate.

Forecast V actual Mentions – March 8th



Actual Data



Natural Language Processing (NLP)

NLP enables machines that can extract information and insights contained in the documents as well as categorize and organize the documents themselves.

Natural Language Generation (NLG)

NLG is the process of creating phrases and sentences in the form of natural language with a machine. It can automatically generate narratives that describe, summarize or explain input structured data in a human-like manner at the speed of thousands of pages per second.

Natural Language Understanding (NLU)

NLU reads human language and turns it into structured data understandable to computers.

Natural Language Processing M.O.

Natural Language Processing (NLP) is a subfield of artificial intelligence that deals with the interaction between computers and human language. It aims to enable computers to understand, interpret, and generate human language.

NLP has many useful applications within EU public affairs, including:

- Sentiment analysis: NLP can be used to analyze the sentiment expressed in public opinion and social media posts related to EU policies and political issues. This information can be valuable for policymakers and public affairs professionals to better understand the public's views and to adjust their messaging and policies accordingly.

- Text classification: NLP can be used to classify large volumes of text, such as news articles or social media posts, into categories such as topic, source, or sentiment.

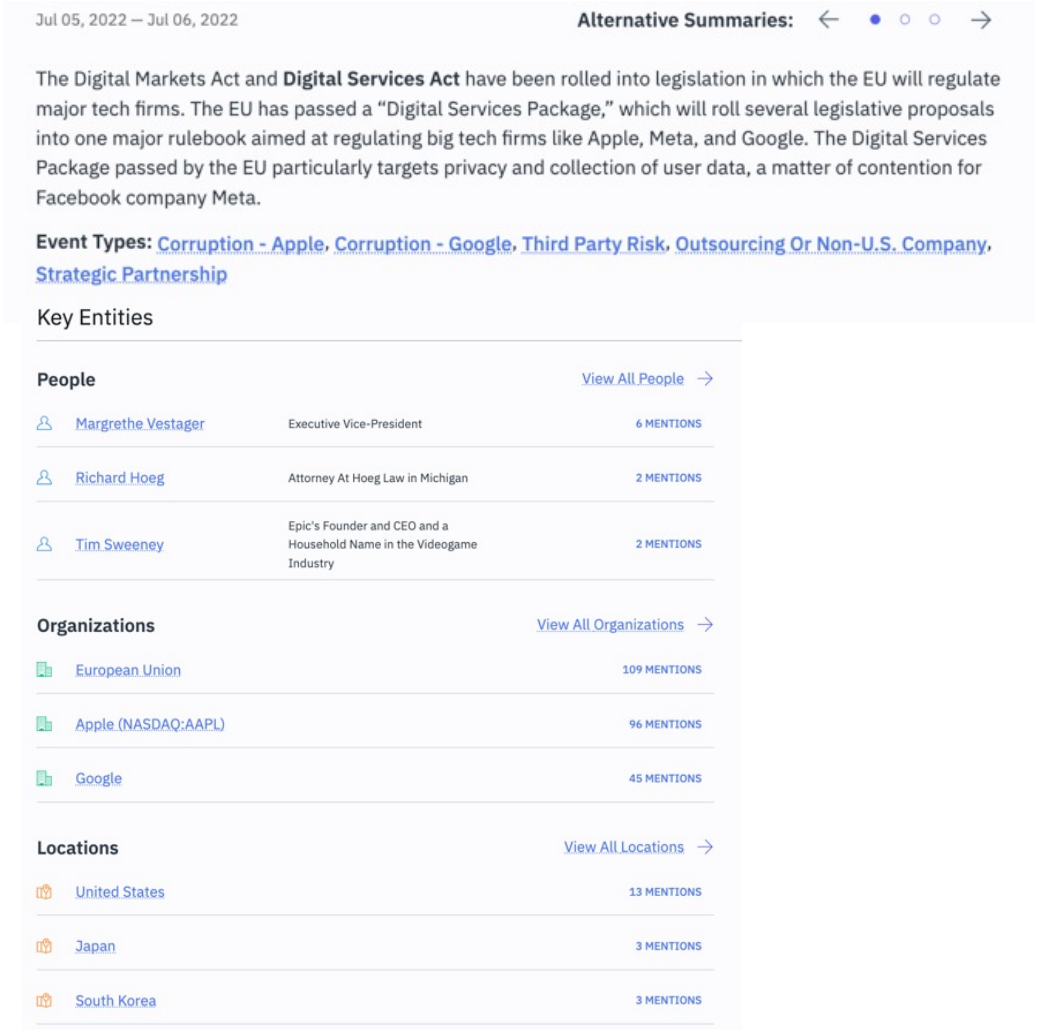
- Named entity recognition: NLP can be used to identify and extract key entities from text, such as people, organizations, and locations.

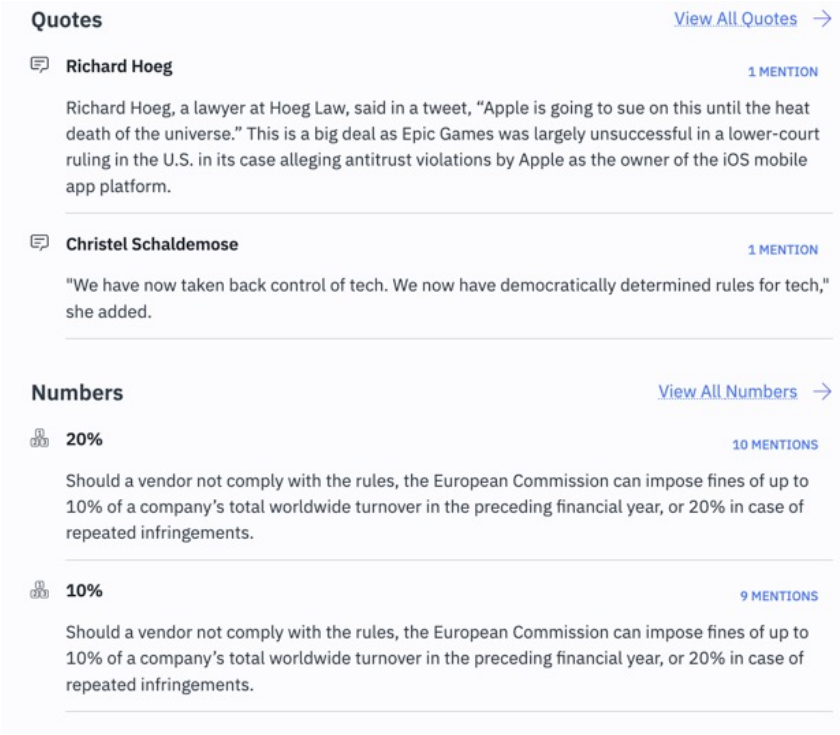
- Content creation: NLP can be used to generate natural language content, such as summaries or reports, based on structured data or other inputs.

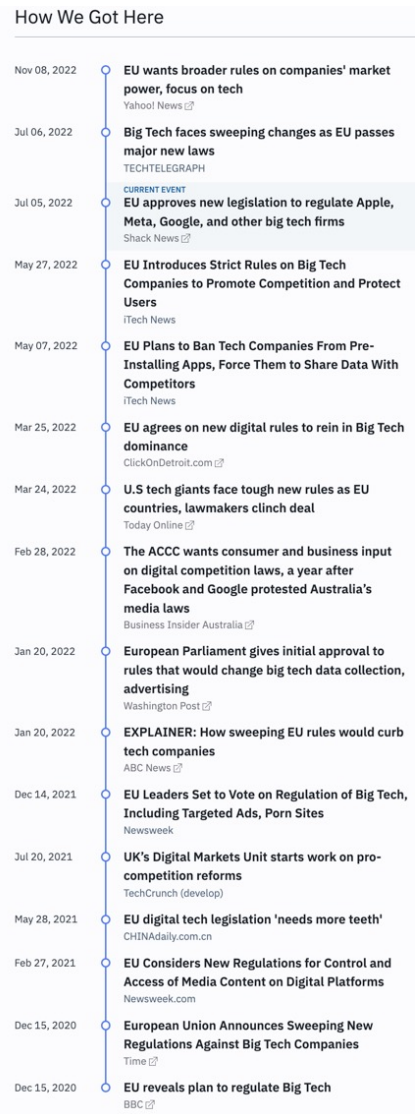
Real Time Monitoring/Summarization

Always on AI alerting and monitoring systems can be set up to provide immediate context & insights to any public issue.

- Summarize hundreds of documents to get a more holistic picture on the issue.
- Identify misinformation, key people, and places.
- Structure the order in how events happed in near real-time.

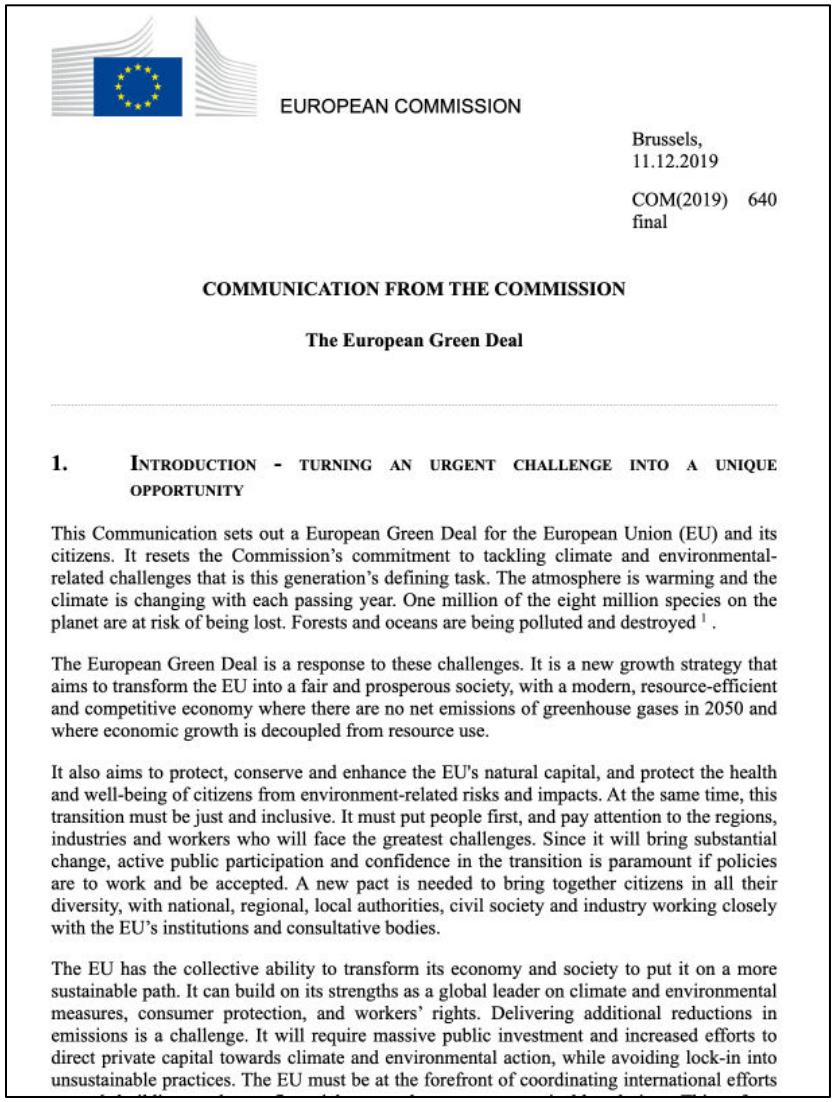






Summarize Policy Documents

Using large language models policy documents can be summarized in seconds, leaving time to be think deeper about a given position, as well as more consistency within an organization's information flow, and therefore reducing blind spots and bias.



EUROPEAN COMMISSION

Brussels,
11.12.2019

COM(2019) 640
final

COMMUNICATION FROM THE COMMISSION

The European Green Deal

1. INTRODUCTION - TURNING AN URGENT CHALLENGE INTO A UNIQUE OPPORTUNITY

This Communication sets out a European Green Deal for the European Union (EU) and its citizens. It resets the Commission's commitment to tackling climate and environmental-related challenges that is this generation's defining task. The atmosphere is warming and the climate is changing with each passing year. One million of the eight million species on the planet are at risk of being lost. Forests and oceans are being polluted and destroyed [1].

The European Green Deal is a response to these challenges. It is a new growth strategy that aims to transform the EU into a fair and prosperous society, with a modern, resource-efficient and competitive economy where there are no net emissions of greenhouse gases in 2050 and where economic growth is decoupled from resource use.

It also aims to protect, conserve and enhance the EU's natural capital, and protect the health and well-being of citizens from environment-related risks and impacts. At the same time, this transition must be just and inclusive. It must put people first, and pay attention to the regions, industries and workers who will face the greatest challenges. Since it will bring substantial change, active public participation and confidence in the transition is paramount if policies are to work and be accepted. A new pact is needed to bring together citizens in all their diversity, with national, regional, local authorities, civil society and industry working closely with the EU's institutions and consultative bodies.

The EU has the collective ability to transform its economy and society to put it on a more sustainable path. It can build on its strengths as a global leader on climate and environmental measures, consumer protection, and workers' rights. Delivering additional reductions in emissions is a challenge. It will require massive public investment and increased efforts to direct private capital towards climate and environmental action, while avoiding lock-in into unsustainable practices. The EU must be at the forefront of coordinating international efforts

Machine Summarization

- *The European Green Deal is a new growth strategy that aims to transform the EU into a fair and prosperous society, with a modern, resource-efficient and competitive economy where there are no net emissions of greenhouse gases in 2050 and where economic growth is decoupled from resource use.*
- *It also aims to protect, conserve and enhance the EU's natural capital, and protect the health and well-being of citizens from environment-related risks and impacts.*
- *The Green Deal is an integral part of this Commission's strategy to implement the United Nation's 2030 Agenda and the sustainable development goals.*

Communicate With Machine Made Content

Firms can use machines to create communication strategies that are 10x faster, more empathetic, objective and cheaper than any traditional approach. These technologies can also be applied to create twitter posts, blogs, web content or emails and predict the outcomes with mathematical precision.

Training Input

Carbon Capture and Storage (CCS) can contribute to meeting Paris Agreement GHG emission ambition.

- The Paris Agreement and the IPCC (Intergovernmental Panel on Climate Change) 1.5°C report a very ambitious global trajectories that can only be achieved **with a combination of a wide range of technologies** and new business models. Possible technology pathways have been investigated for the EU chemical industry[1] and for EU's energy-intensive industries[2,3]
- Different options and technologies are needed to address climate change. Besides avoiding emissions in the most effective ways, remaining emissions need to be reused/recycled (CCU), stored (CCS or deposition of carbon in solid forms) or compensated to reach carbon neutrality.
- Carbon Capture Storage (CCS) involves the capture, transport and permanent storage of CO2.
- CCS has a role to play in meeting the EU's climate commitment and is recognized by the IPCC, IEA and in SET plans. According to the IEA (2017), CCS will need to provide 14% of all global cumulative CO2 reductions by 2060 to meet the IEA's '2ºC' (2DS) scenario and will provide one third of the incremental CO2 reductions in the IEA's 'Below 2ºC' (B2DS) scenario. Moreover, the IPCC (2014) reported that 2ºC scenarios, which use alternatives to CCS, would on average cost more than twice as much (a 138% increase).
- CCS is also a multi-pathway and cross-sector climate solution for ensuring deep emissions reduction in power, decarbonization of natural gas to hydrogen, in industrial processes, and for bioenergy (bio CCS to enable negative emissions). These solutions enable low carbon heat, low carbon transport and low carbon industries.

Machine Output

Title: EU's Stance on Carbon Capture and Storage

Blog Introduction: The Paris Agreement and the IPCC (Intergovernmental Panel on Climate Change) 1.5°C report a very ambitious global trajectories that can only be achieved with a combination of a wide range of technologies and new business models. Possible technology pathways have been investigated for the EU chemical industry1 and for EU's energy-intensive industries2,3. Different options and technologies are needed to address climate change. Besides avoiding emissions in the most effective ways, remaining emissions need to be reused/recycled (CCU), stored (CCS or deposition of carbon in solid forms) or compensated to reach carbon neutrality.

What Is Carbon Capture and Storage? Carbon Capture and Storage, or CCS, involves the capture, transport, and storage of anthropogenic carbon dioxide (CO2). CO2 can be captured from various emission sources including power plants, industrial processes and even from the air. Once captured, the CO2 is transported to a storage site where it is injected into geological formations deep underground, such as depleted oil and gas fields or saline aquifers. CCS is a key technology in helping to mitigate climate change as it reduces atmospheric CO2 concentrations whilst allowing continued use of fossil fuels.

How Can CCS Help Reduce Emissions? There are two main ways in which CCS can help reduce GHG emissions; firstly, by reducing emissions from power generation and secondly by enhancing oil recovery while preventing emissions from flaring gas associated with oil production.

In terms of power generation, CCS can play a role in reducing emissions from both existing coals fired power stations as well as new build projects. It is estimated that if CCS was deployed on all newly built coal fired power stations around the world, then global CO2 emissions could be reduced by up to 8 GtCO2/year (). This would be the equivalent of taking all the cars off every road in America and Europe combined. For existing coal fired power stations, CCS could play an important role in extending their operating life whilst at the same time significantly reducing their emissions footprint. Indeed, if retrofitted to just half of the world's existing coal fired capacity, then annual CO2 reductions of up to 4 GtCO2 could be achieved ().

When it comes to oil production, CCS can help reduce emissions in two ways; firstly, by preventing flaring of associated gas and secondly by enhancing oil recovery rates through EOR (Enhanced Oil Recovery). Globally around 140 billion m3 (5 trillion ft3) of natural gas is flared each year which is equivalent to approximately 300 million tonnes of CO2 being emitted into the atmosphere (). If this gas was instead captured and stored underground, then these substantial flaring emissions could be eliminated. In addition, through EOR operations it is possible to increase ultimate oil recovery rates from around 35% to 55% (). As a result, more economically, viable reserves can be extracted whilst at the same time reducing GHG intensity per barrel of oil produced. This technique is however still in its infancy with only around 60 EOR projects currently operational around the world ().

conclusion:
To meet future GHG emission targets it is clear that multiple technologies will need to be deployed including Carbon Capture and Storage (CCS). When used correctly CCS can play an important role in mitigating climate change by helping to reduce emissions from both power generation as well as oil production. Furthermore, due to its flexibility CCS can also be used in conjunction with other mitigation measures such as solar PV or wind power to provide a more holistic approach in tackling climate change.

Stakeholder Mapping – People and Organization Affiliation

Machine can quickly extract all people associated with the policy area, institution, and the organization they are affiliated with. The data can be exported via CSV files or entered into a database for further analysis or enrichment.

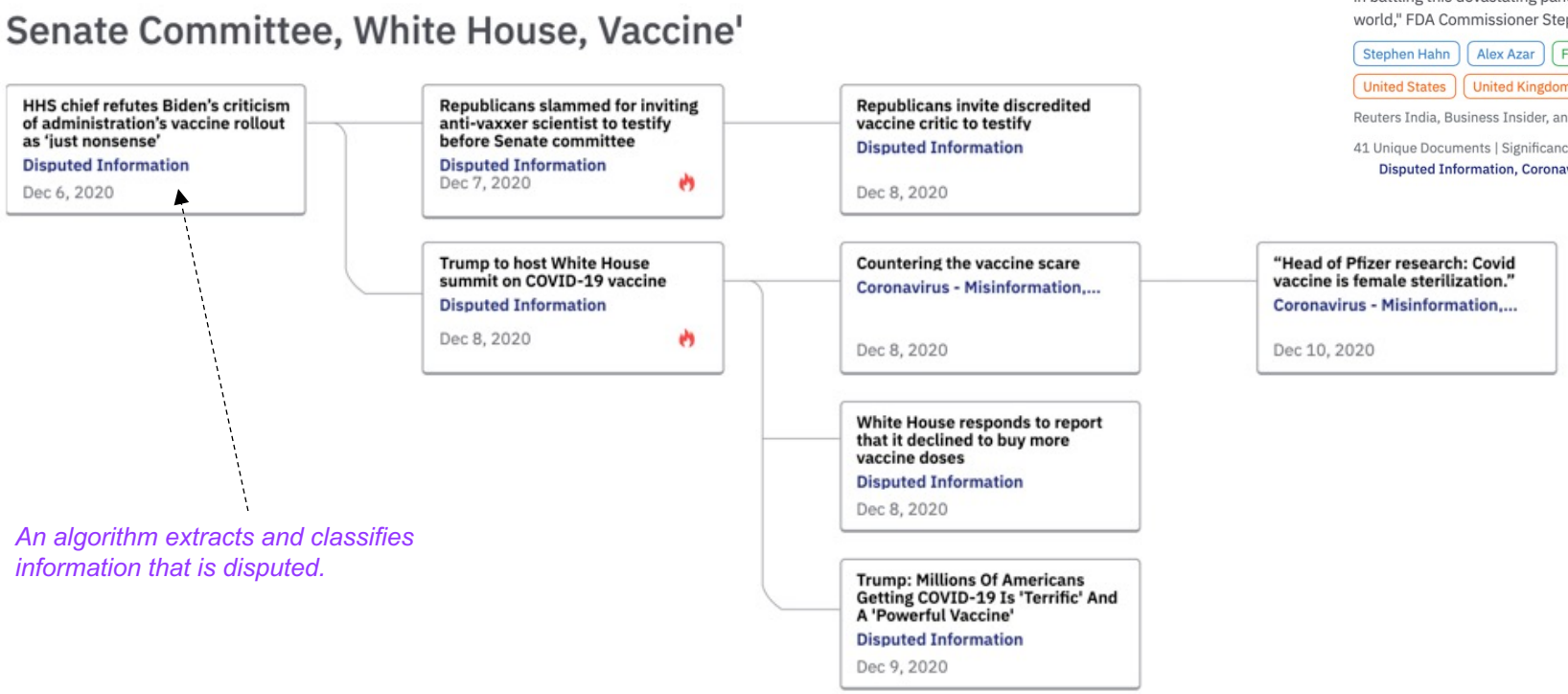
Name	Affiliation	Events ↓	Related Affiliations	Quotes	Related People	Related Organizations	Related Locations
Bruno Le Maire	-	16	0	8	5	10	6
Narendra Modi	Prime Minister	6	1	4	0	5	4
Joe Biden	President	5	2	0	1	3	1
Nirmala Sitharaman	Finance Minister	4	1	5	4	13	3
Donald Trump	Former President	4	2	2	1	1	2
Aynne Kokas	-	4	0	4	1	3	1
Elon Musk	-	3	0	1	1	4	3
Jerome Powell	-	3	0	0	1	6	1
Emmanuel Macron	French President	3	1	1	1	2	2
Rajesh Kandaswamy	Gartner Analyst	3	1	15	0	2	0
Shaktikanta Das	RBI Governor	3	2	0	3	9	1
Mike Gallagher	Republican Legislator	3	1	1	1	3	2
Alexander Kalinin	Founder of NCC	3	1	0	0	1	0
Sarah Kreps	Director of Cornell University's Tech Policy Institute	2	1	0	0	4	2
Antony Blinken	U.S. Secretary Of State	2	1	0	1	1	2

Classify Events and Misinformation

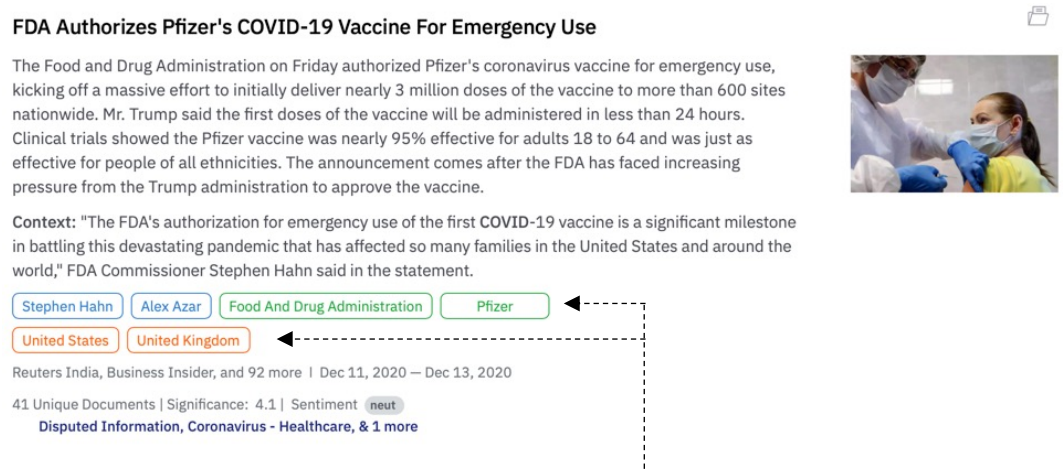
With massive global events affect everything it's vital for firms to have clarity of information and data. Only through machines is analysis is possible due to exponential amounts of data and misinformation being created everyday.

Finding Misinformation

- Indexing the entire corpus of information enable people to see what isn't connect thus higher probability of being false.
- Using data to filter high quality sources
- Concurrent citations of numbers, facts and figures



An algorithm extracts and classifies information that is disputed.



An algorithm classifies information & event categories, as well as key institutions, locations, corporations and people involved using world leaded (named entity resolution - NER).

Value Proposition

- 8 mins - time it takes to read 1000 words
- 1890 - average word count per article on Google's first page result
- 4 - documents per hour, @15 mins per
- 96 - documents in 24 hours
- 160 per 40-hour work week, 672 @ 24 hours of reason per day
- 35k - documents read a year by a human 24/7/365…

One analyst that leverages the machine & OSINT ecosystem can summarize 9.6 years of information in one day (circa 1 million documents)

About $316,000,000 in intellectual value per year at 150k per year per analyst.

The EU Commission and Parliament employs roughly 6,000 analyst types according to LinkedIn at a cost of around $1B per year. 25% efficiency gains are realistic at the enterprise level, thus an amplification of $250m in value creation per year.

Or the amplification is equivalent to buying one analyst at $.29 per day and $65 per year.

Network Analysis

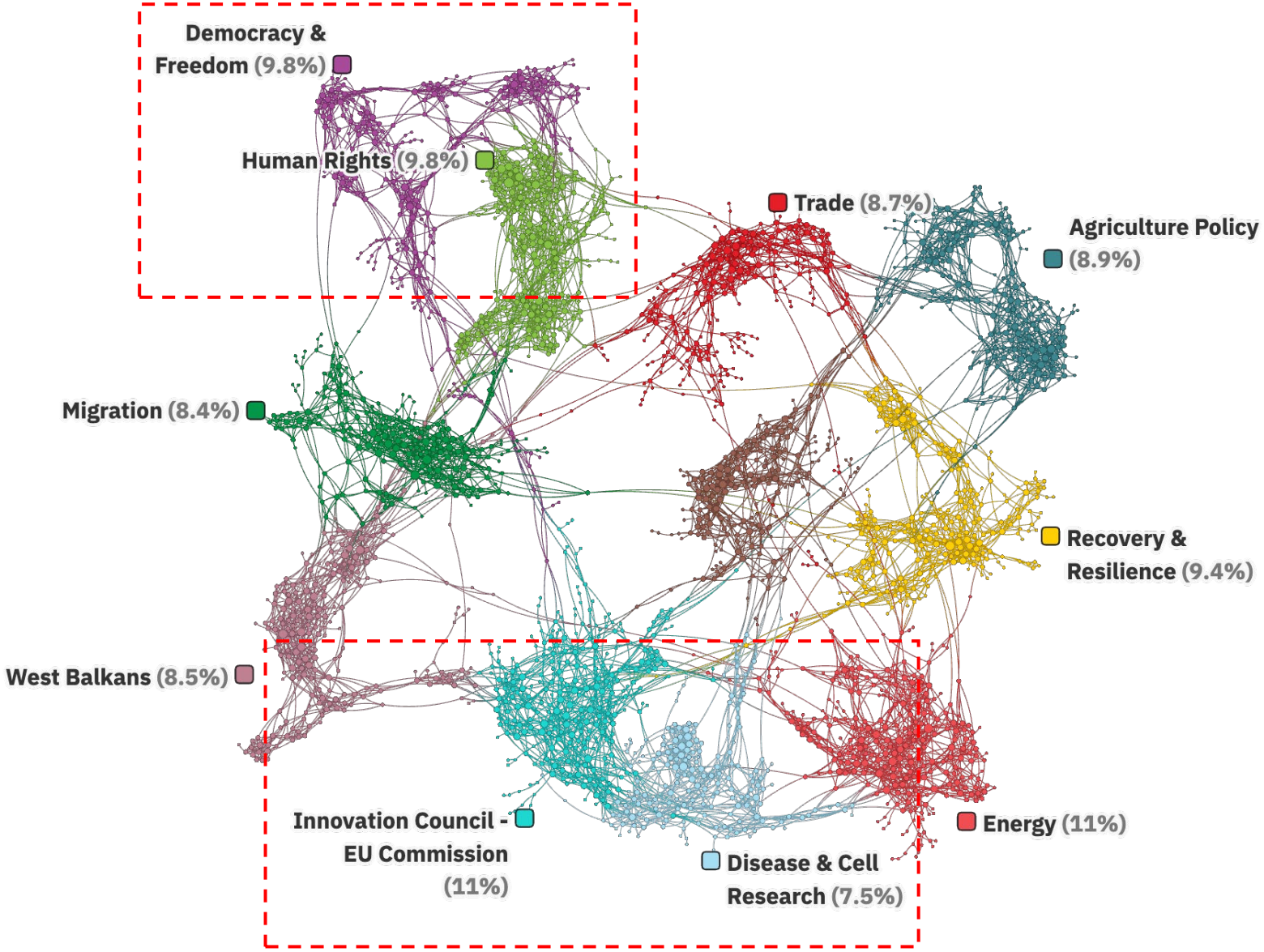
Networks excel at exploring how policies, people, regions, and institutions and companies are connected, allowing organizations to find risks, blind spots, and opportunities not possible with traditional analytics.

The network example is data mining communications from from the EU to quantify how policy priorities, and concerns do or don't overlap.

- Closer the policy areas equals more of a relationship via issues and connection. For example, innovation and disease research are connected policy areas to one another.

- Policy areas that are further away from one another highlight a lack of connection in their current state. For example, Human rights and energy discussions are siloed, while HR is connected to migration and trade.

Network map of EU Communication themes over the last five years



Network Analysis Metrics

In summary, centrality measures such as degree centrality and inter-cluster connectivity are useful tools in network analysis for understanding the importance and influence of nodes in a network, as well as the level of cohesion or fragmentation within a network. These measures can be used to identify key nodes or groups of nodes that are critical to the functioning of a network, and to evaluate the overall structure and health of a network. Data that can be analyze is both unstructured, and structured data.

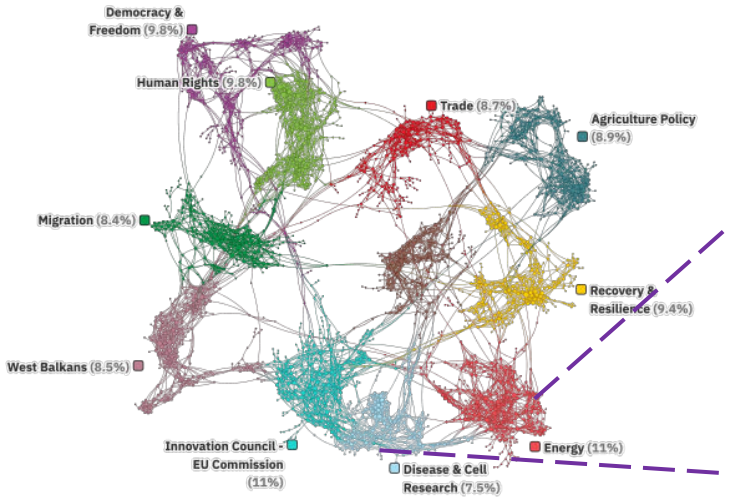
Key Metrics	**Description**
Centrality	Centrality is a concept in network analysis that refers to the importance or influence of a node (or vertices) in a network. There are several types of centrality measures, including degree, betweenness, closeness, and eigenvector centrality, that can be used to determine the level of centrality of a node within a network.
Degree	Degree centrality is the simplest form of centrality and is calculated based on the number of connections (edges) a node has with other nodes in the network. Nodes with higher degrees are considered more central, as they have more connections and are thus more likely to have a greater influence over the network.
Inter-cluster connectivity	Inter-cluster connectivity is another aspect of network analysis that measures the connectivity between clusters or groups of nodes within a network. It refers to the extent to which nodes in different clusters are connected to each other, and is used to evaluate the level of cohesion or fragmentation within a network. High inter-cluster connectivity suggests that nodes from different clusters are well-connected, indicating a more cohesive network structure. Conversely, low inter-cluster connectivity indicates that nodes are more fragmented and less well-connected, indicating a less cohesive network structure.

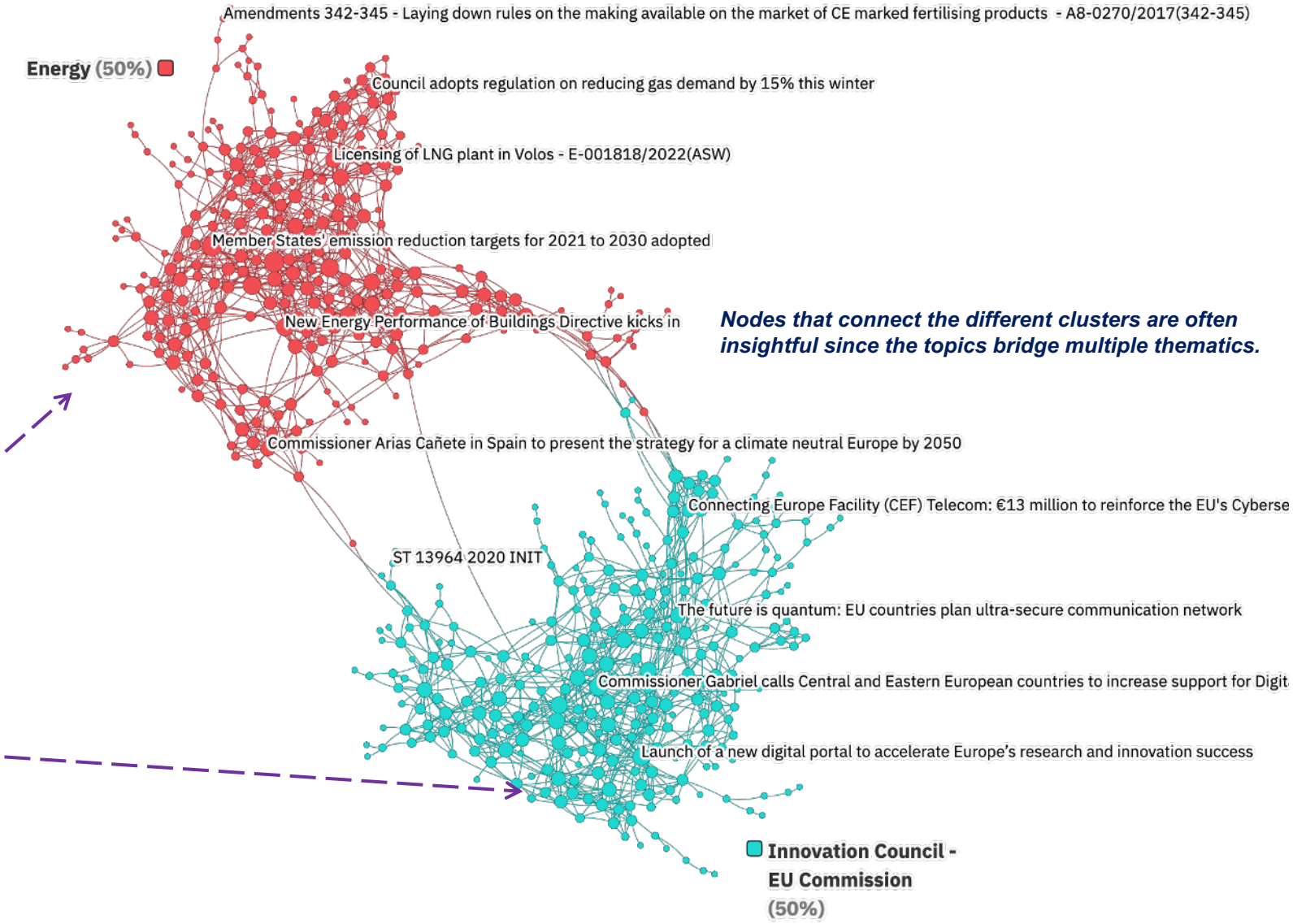
Zoom In For Context

For more detail it's possible to zoom in on any topic within the networks for granular insights.

Zoom in on the Innovation & Energy Clusters (Colors) for more context on how

Same Network from Prior Page



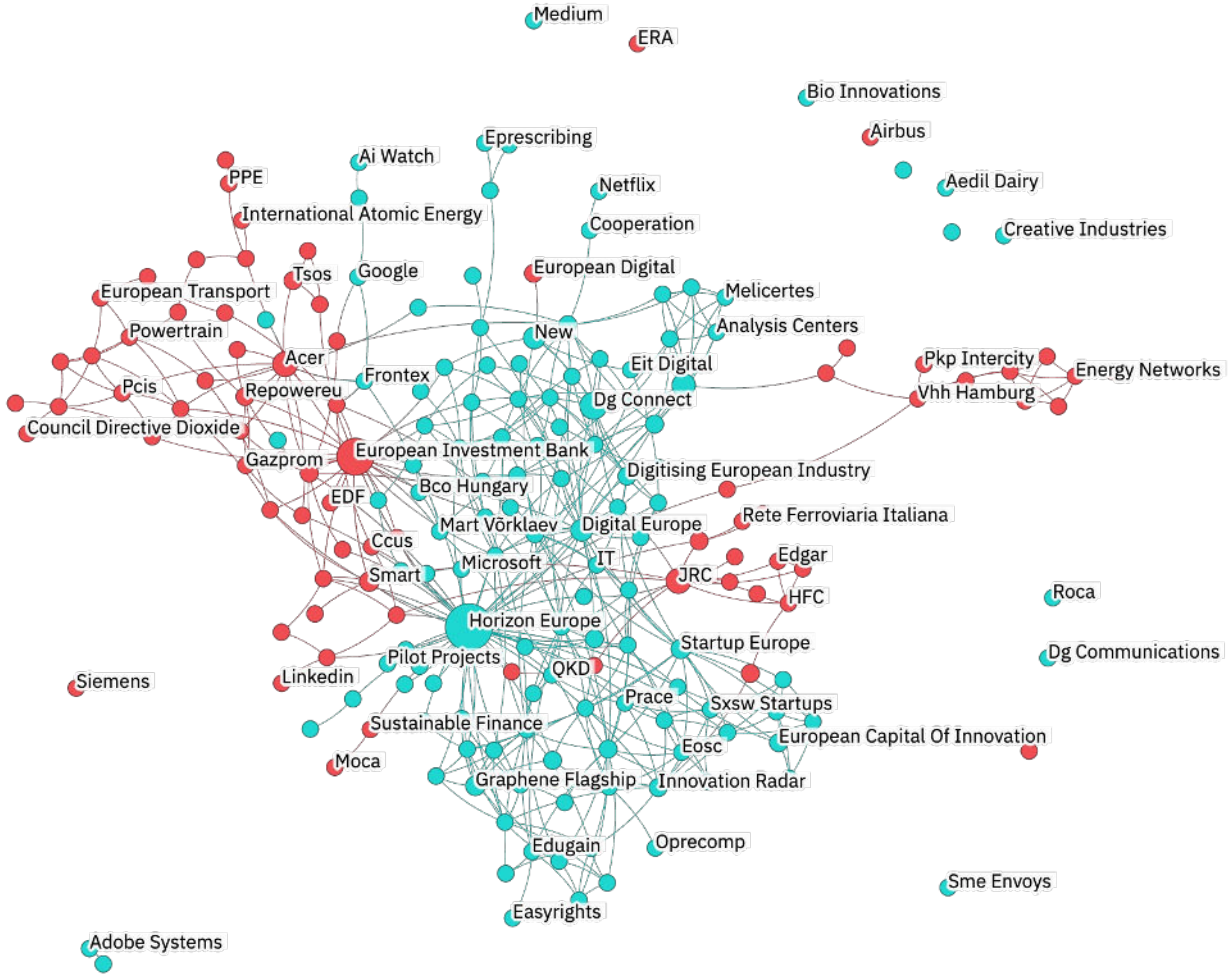


Nodes that connect the different clusters are often insightful since the topics bridge multiple thematics.

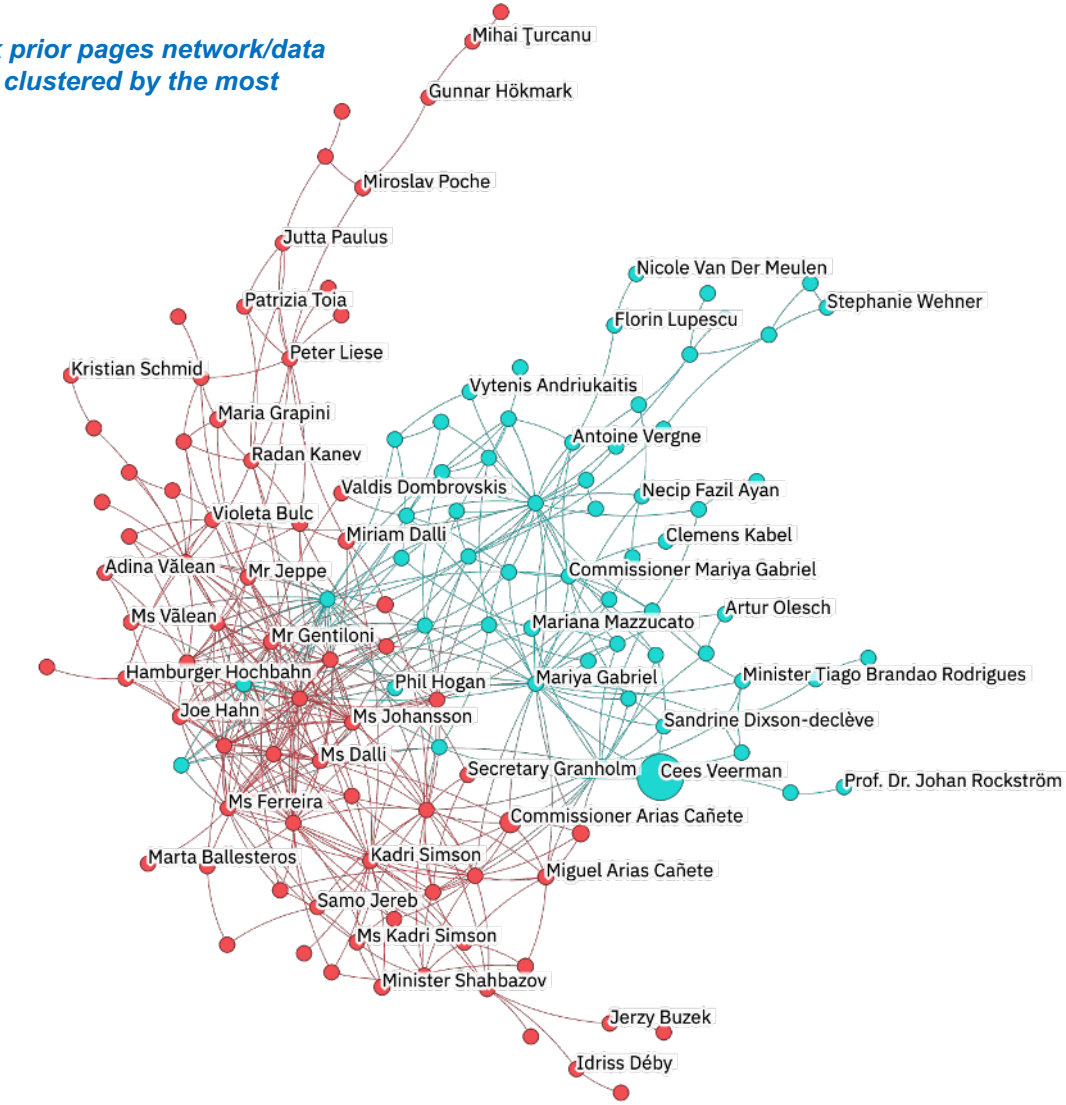
Next Level Stakeholder Mapping

The network graphs below are the same date from the prior page. In addition to emerging trends, algorithms can extract the entities - people, institutions, companies, and locations, most associated with domain or any content (such as policy white papers, news articles or minutes services i.e., like Dods) which in effect makes stakeholder mapping 100-1000x faster than traditional desk research.

Same data from the network prior pages network/data (Innovation and energy) but clustered by key Institutions

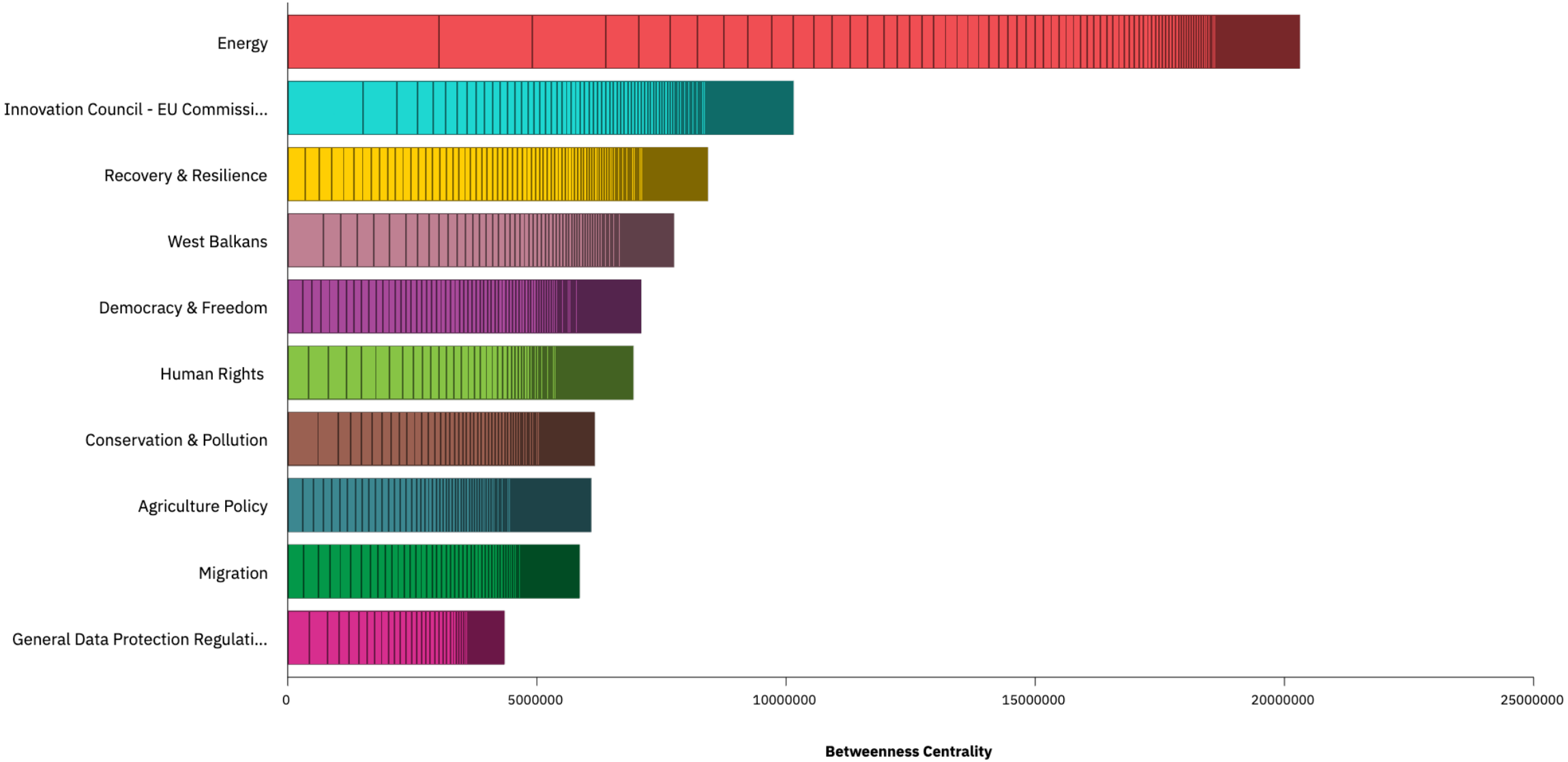


Same data from the network prior pages network/data (Innovation and energy) but clustered by the most central person extracted.



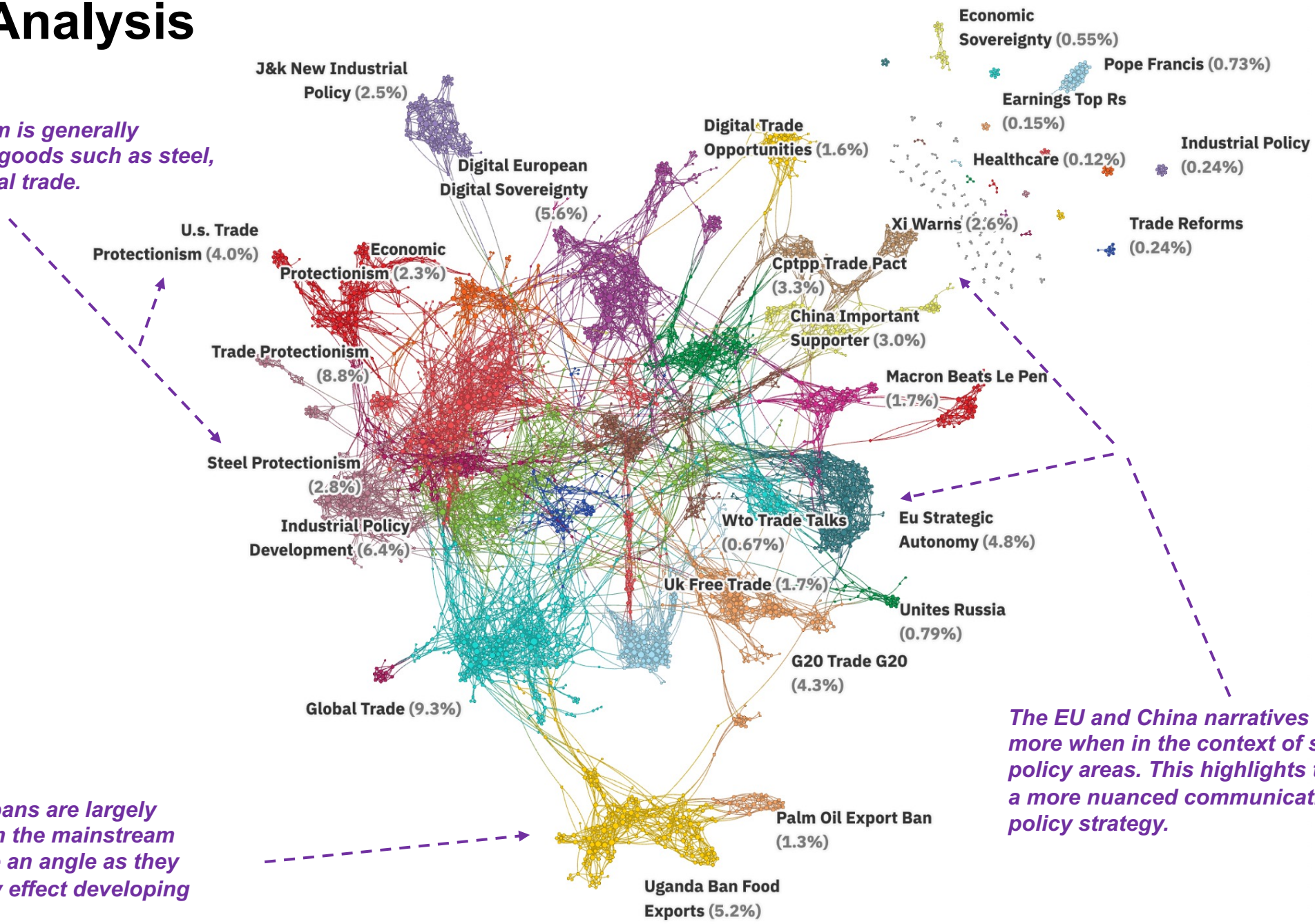
Rank Order Policy Areas With Advanced Metrics

Network analytics can also provide metrics that provide more context than social sharing, mentions volume or sentiment only (which we can also do). Further, going a step forward these advanced metrics can be used to build more accurate forecasts and predictions of emerging trends.



Network Analysis

US trade protectionism is generally focused on industrial goods such as steel, and then links to global trade.

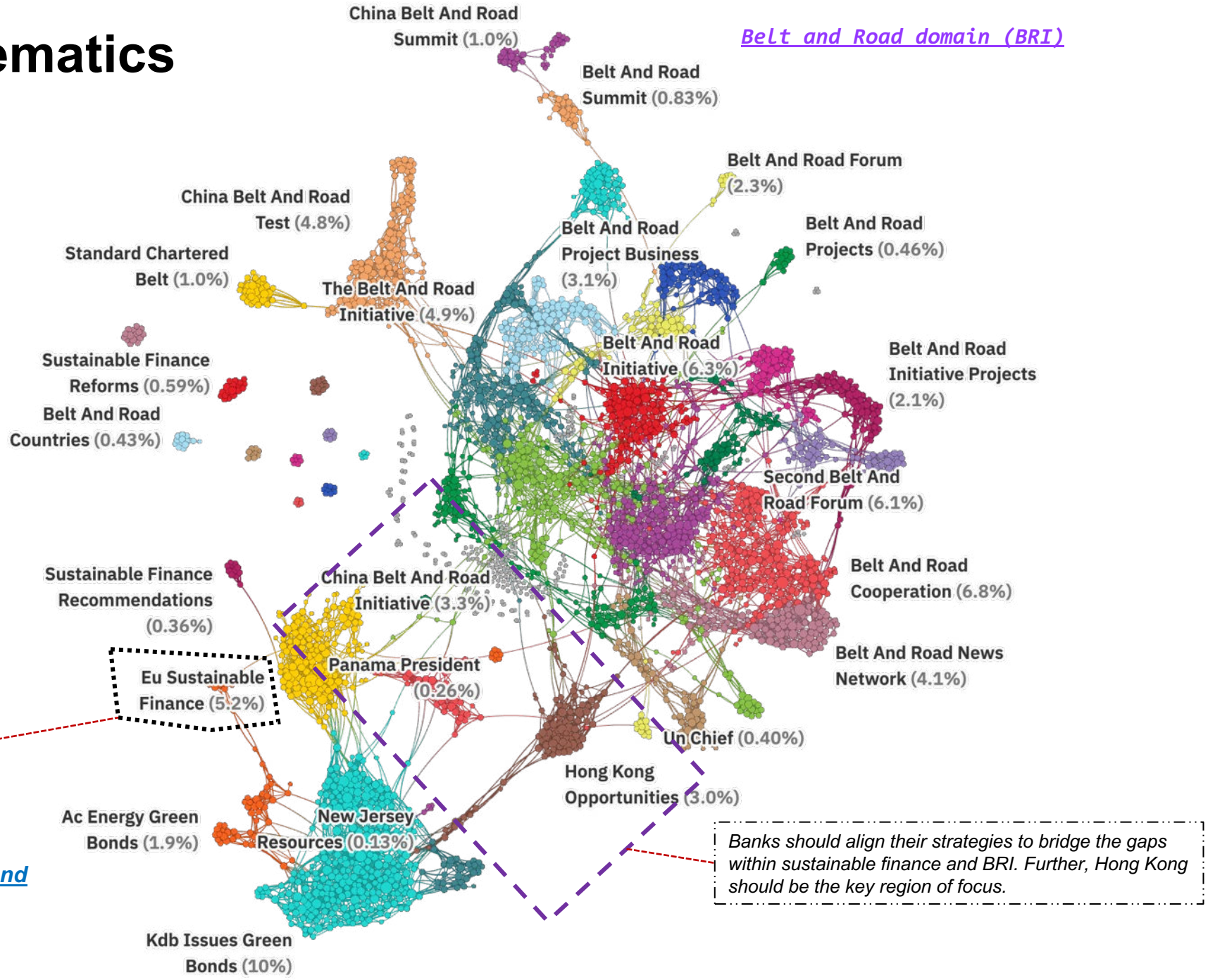


The EU and China narratives intersect more when in the context of specific policy areas. This highlights the need for a more nuanced communication and policy strategy.

Resource export bans are largely disconnected from the mainstream debate and maybe an angle as they disproportionately effect developing societies.

Combine Multiple Thematics

This graph models two domains, Belt and Road and Sustainable finance - two core banking initiatives. Banks should align strategy to bridge the gaps within sustainable finance and BRI. The fact it's lightly connected signals an opportunity. Note that Hong Kong is the link between the two clusters thus; firms in that region should be in a perfect position to facilitate investments between Green Finance and BRI within HK.



Belt and Road domain (BRI)

The connectedness of EU and China within context of sustainable finance and green Bonds should be the investment and product area to focus on.

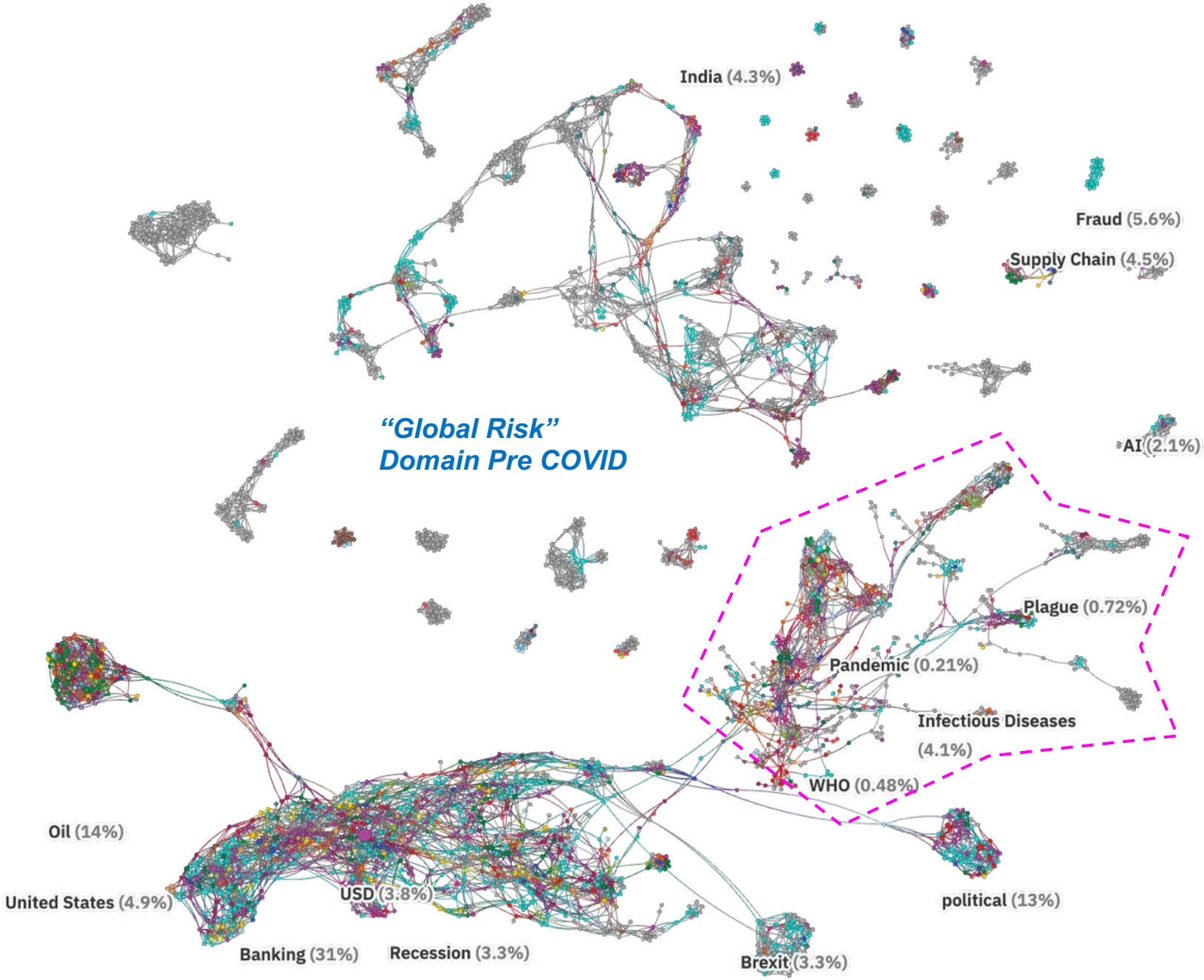
Sustainable Finance and Green Bond Domain

Banks should align their strategies to bridge the gaps within sustainable finance and BRI. Further, Hong Kong should be the key region of focus.

COVID

The graph is a domain analysis of “Global Risk” data from from April 2017 to December 2019 (Pre-COVID). **Pandemics were surfaces by machines all along as a major risk. So, why were organizations blindsided?**

- Firms look at information in siloes and non-strategically.
- No use of open source and alternative data sets.
- No machine intelligence infrastructure to quickly analyze all information streams in context.

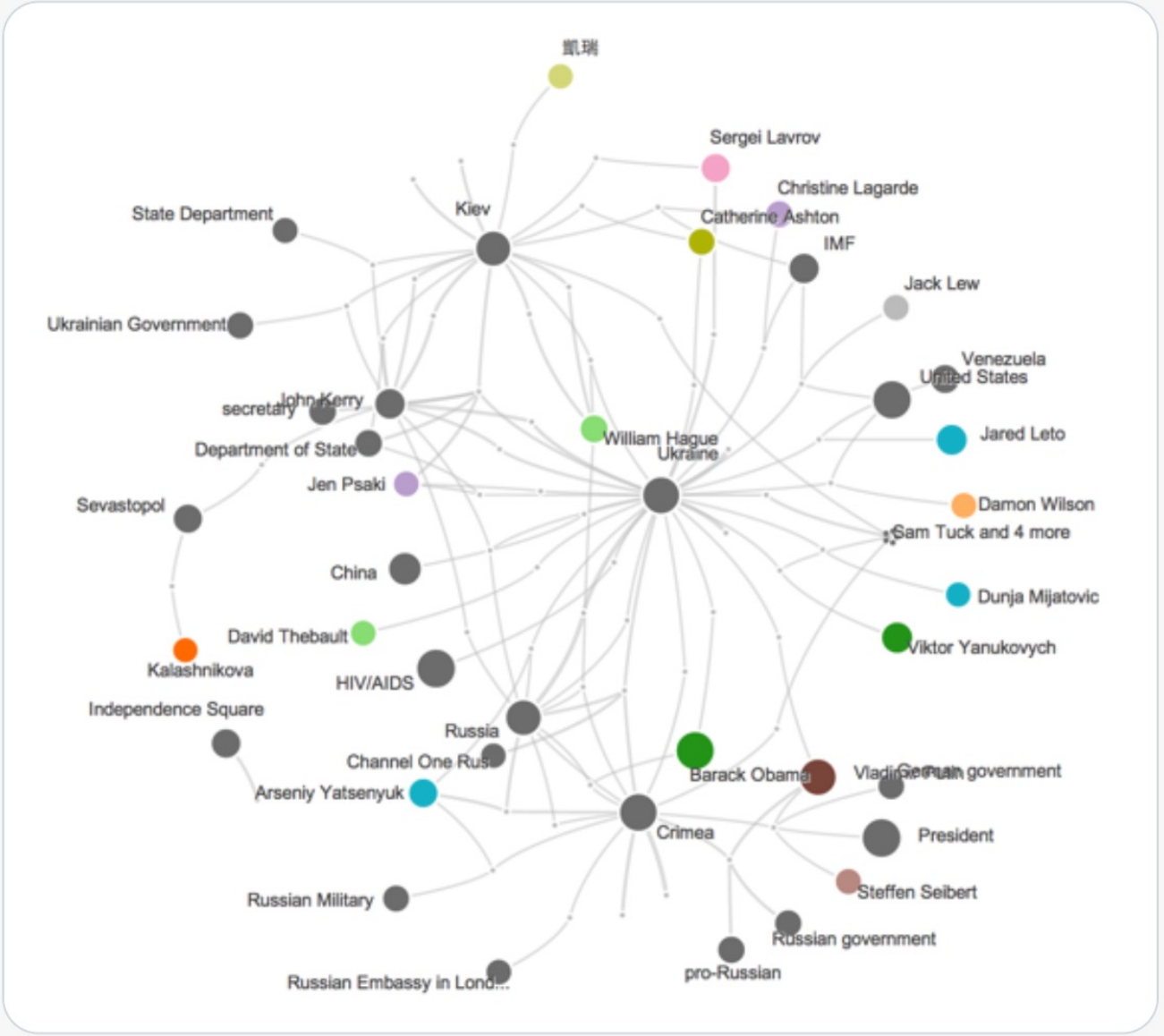


Ukraine

The graph of OSINT mentions of Ukraine in 2014 show the lack of EU institutions signaling a power vacuum.

Chandler T Wilson @chandlertwilson · Mar 3, 2014

Daily Network graph on Ukraine. Surprised by lack of EU ppl in the mix.
#measure #EU #EP2014



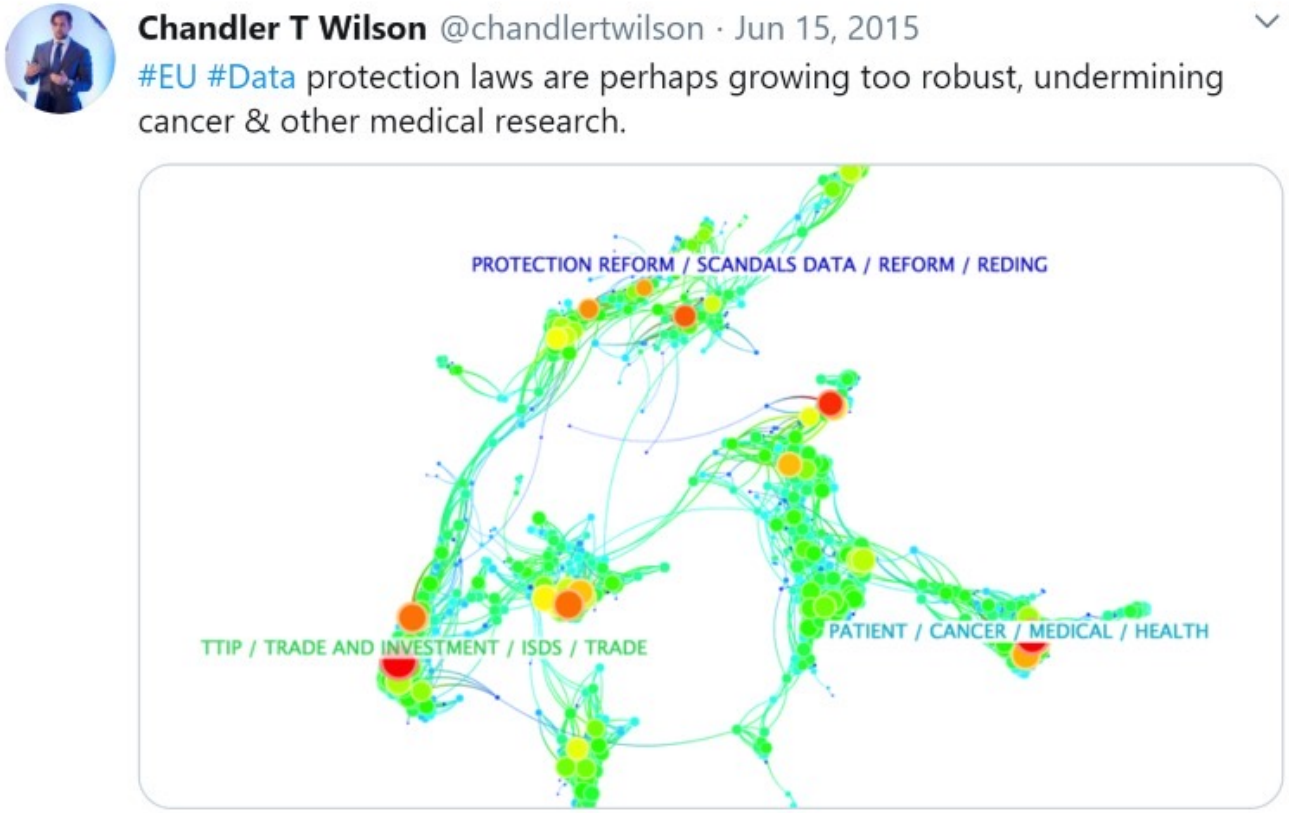
Ukraine - 2022

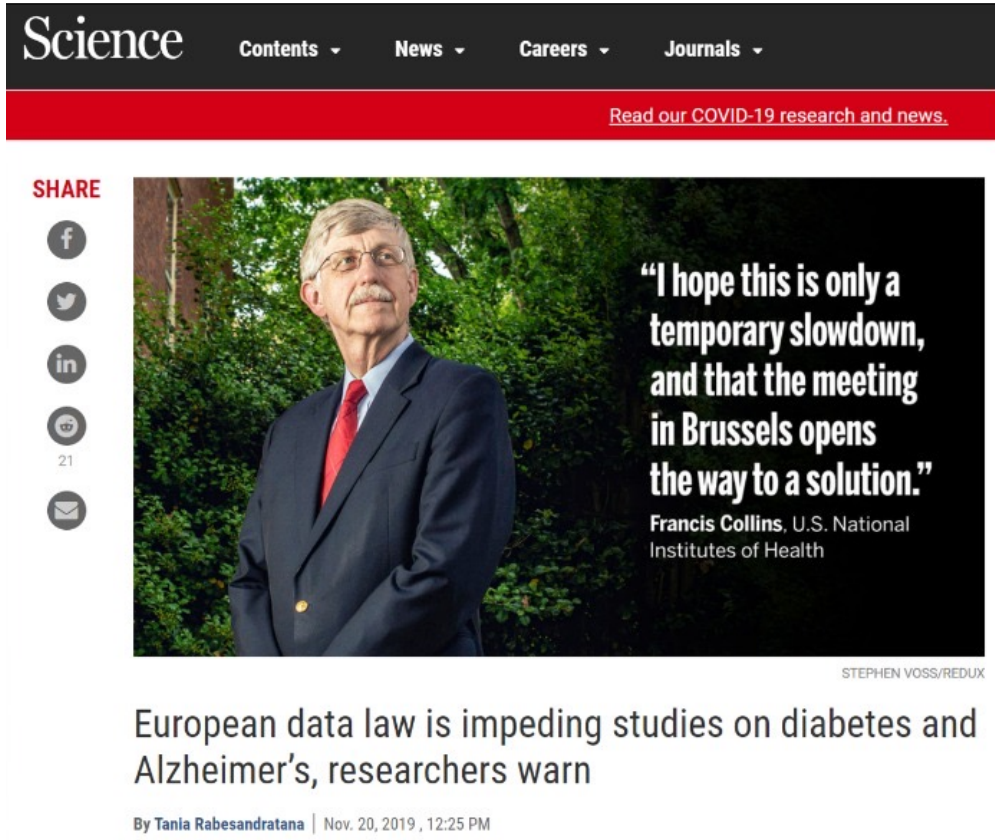
The biggest difference is the presence of the EU actors - Olaf Scholz and Ursula von der Leyen which for the most part were non-existent prior (Bottom center right of the network). From my lens, this quantifies what we've seen - the EU finally coming together in a unified manner.

GDPR

- Connections to GDPR and Health was surfaced in 2015 machine learning few if any people were making the connection between Health and Digital Privacy at the time.

- It wasn't until 2018 when top researchers in the US called out the issues GDPR was causing in modernize health research. Even people working in those domains can't make the unknown connection in OSINT data that a machine can. In 2015 most professions failed to realize that influence is borderless.





Quantitative Communications

Once only the realm of humans, language analysis and strategy can also be analyzed with AI to determine if a message will resonate with a given EU institution or policy domain, or with a particular policy maker. This involves using natural language processing (NLP) and machine learning techniques to analyze the content of text data, such as speeches, position papers, or social media posts, and identify patterns in language and communication that may be more effective at influencing the target audience.

1.Collect and preprocess data: The first step is to collect relevant text data, such as speeches or policy documents, from the EU institution, policy domain, or policy maker of interest. This data should be preprocessed to clean and transform it for analysis.

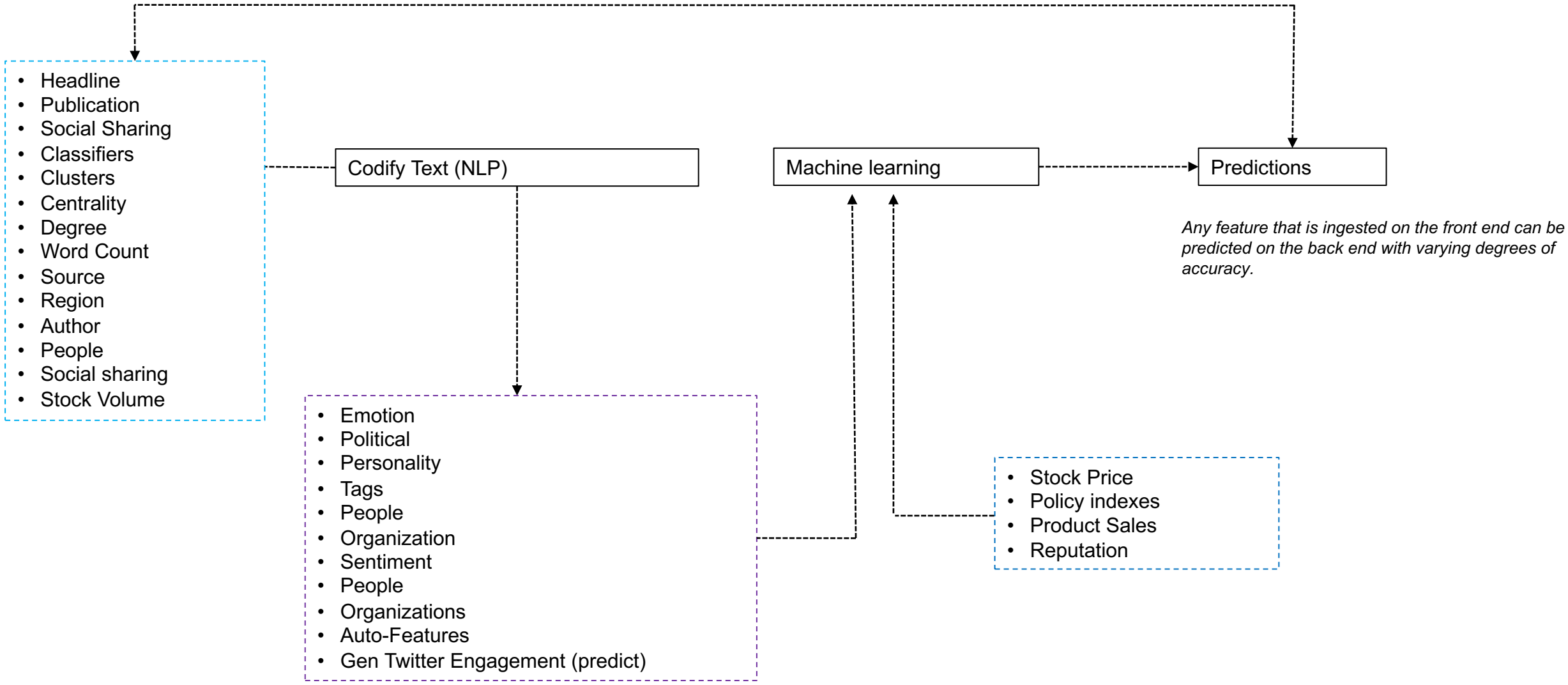
2.Identify key language features: Next, key language features that may be relevant to the target audience should be identified. This might include keywords or phrases that are commonly used in the policy domain or by the target audience, as well as sentiment or tone of language.

3.Train a machine learning model: Once key language features have been identified, a machine learning model can be trained to identify those features in new text data. This typically involves using a supervised learning approach, where the model is trained on labeled data to learn how to recognize the relevant language features.

4.Test and evaluate the model: After the model is trained, it should be tested and evaluated to ensure that it is accurate and effective at identifying the relevant language features. This typically involves testing the model on a separate set of labeled data that was not used for training, and evaluating its performance using metrics such as precision, recall, and F1 score.

5.Apply the model to new data: Once the model is trained and tested, it can be applied to new text data to identify the key language features that may be most effective at influencing the target audience. This can help inform communication and advocacy strategies that are more likely to resonate with the target audience.

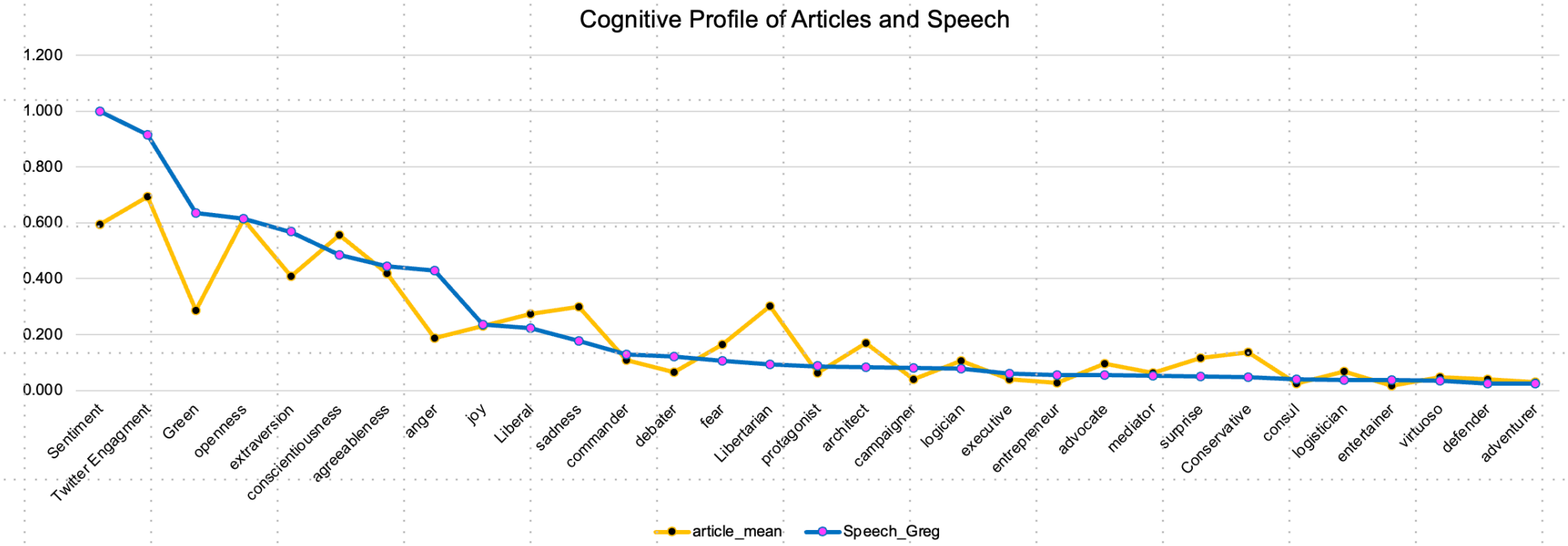
Quantitative Communications



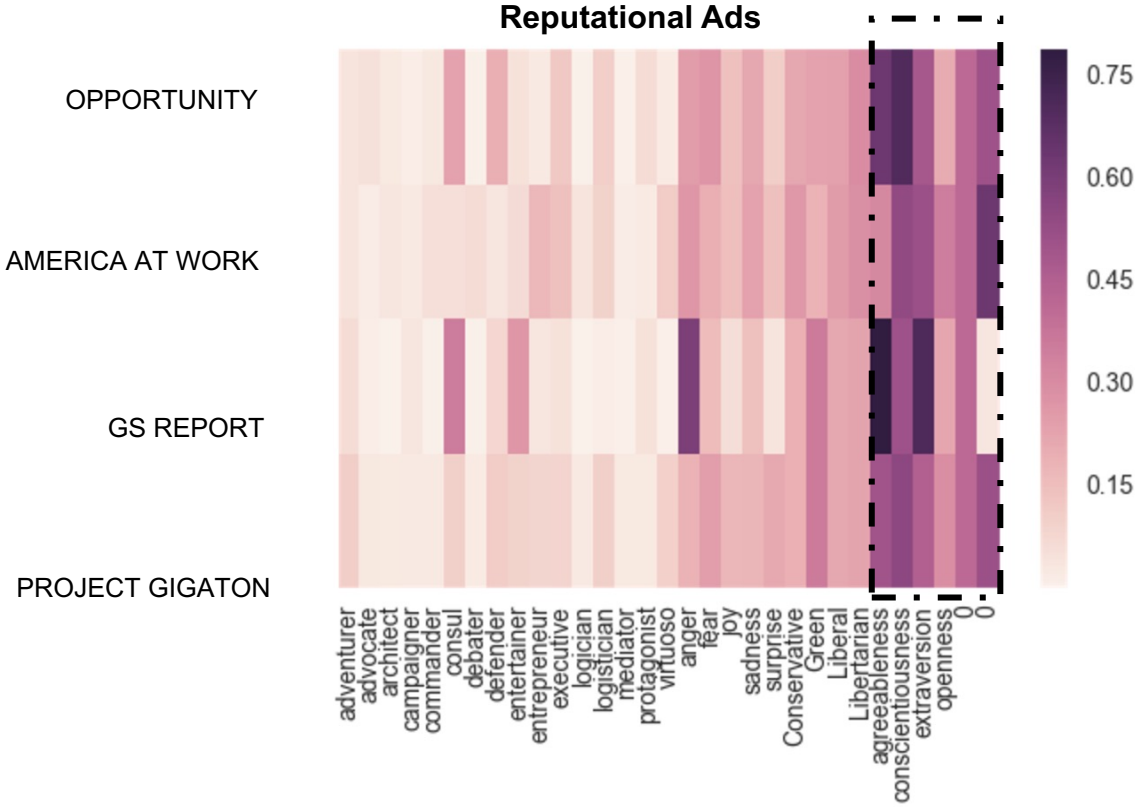
Transference Score

By analyzing articles and speech copy and identifying the differences and similarities in how they relate, a "Transference Score" can be calculated. This score represents the message pull-through at the cognitive and emotional level, which can be quantified separately for each key performance indicator (KPI).

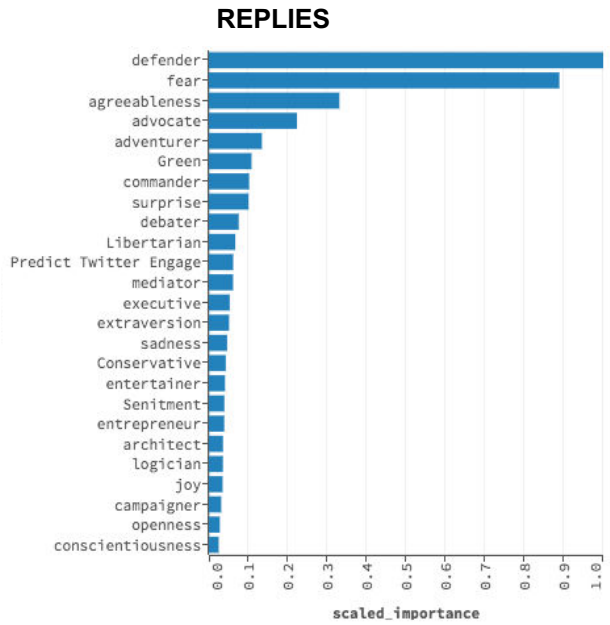
A score of zero represents emotional equilibrium between the communication and the media, while a negative score indicates deficits, where emphasis was not placed on an emotion or cognitive feature that the media that was output and framed strategically was focusing on. Conversely, a positive score indicates that the press or media is emphasizing or keying in on an thematic that the communication is not.
Using this approach, insights can be gained into how messages are being received and interpreted, and areas for improvement in communication and messaging strategies can be identified.

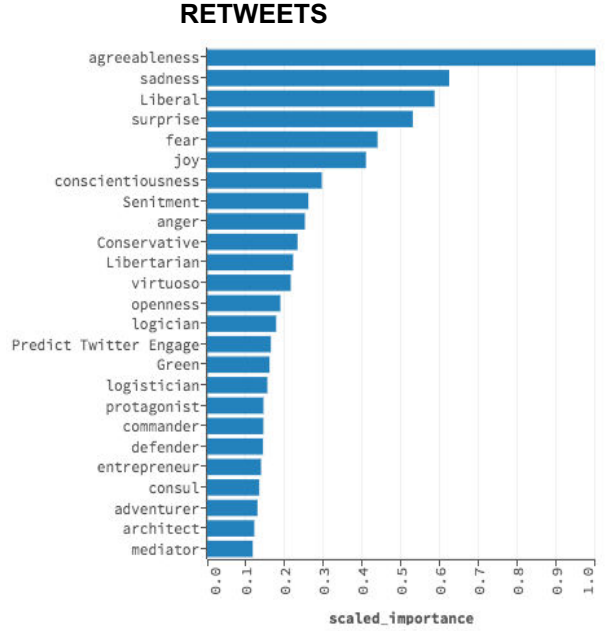


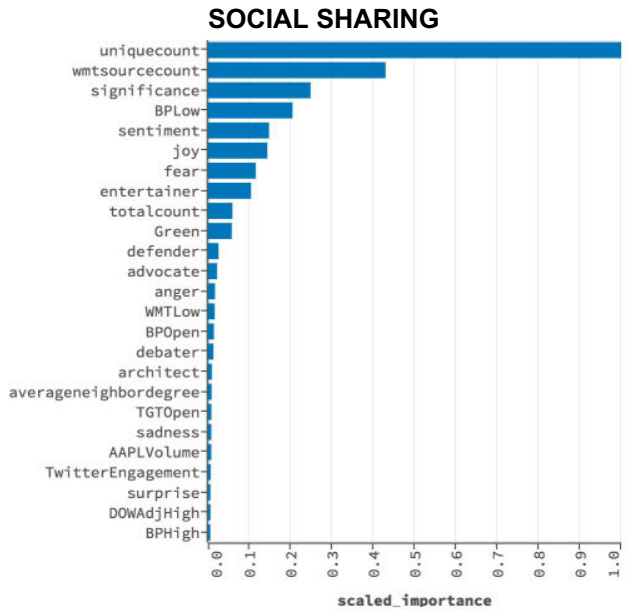
Model Outcomes

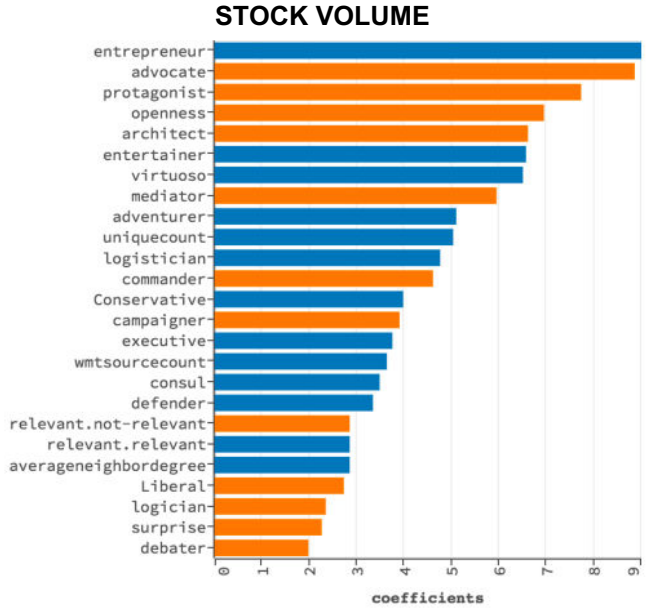


The darker the color the more associated the content is with the cognitive attribute.



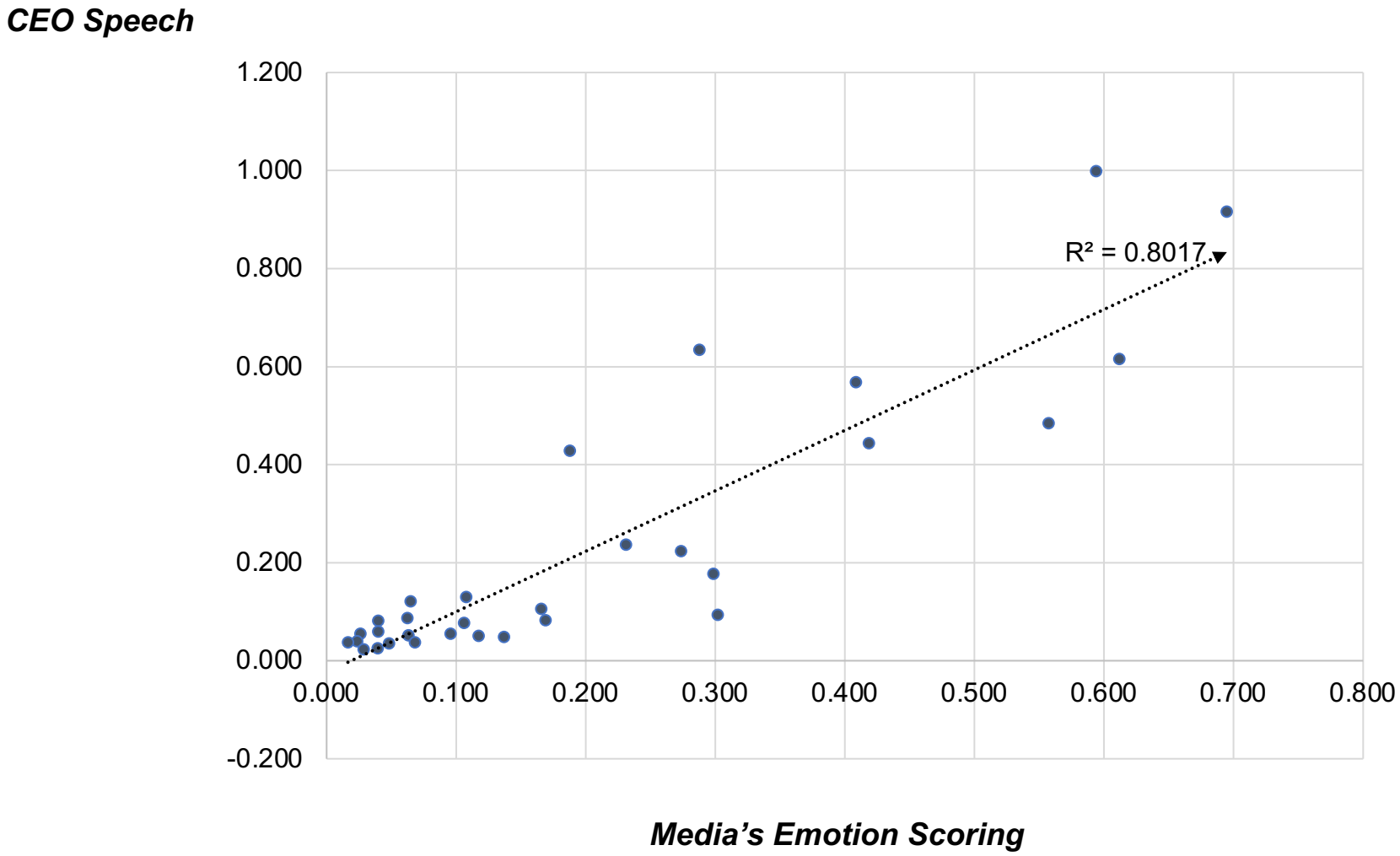




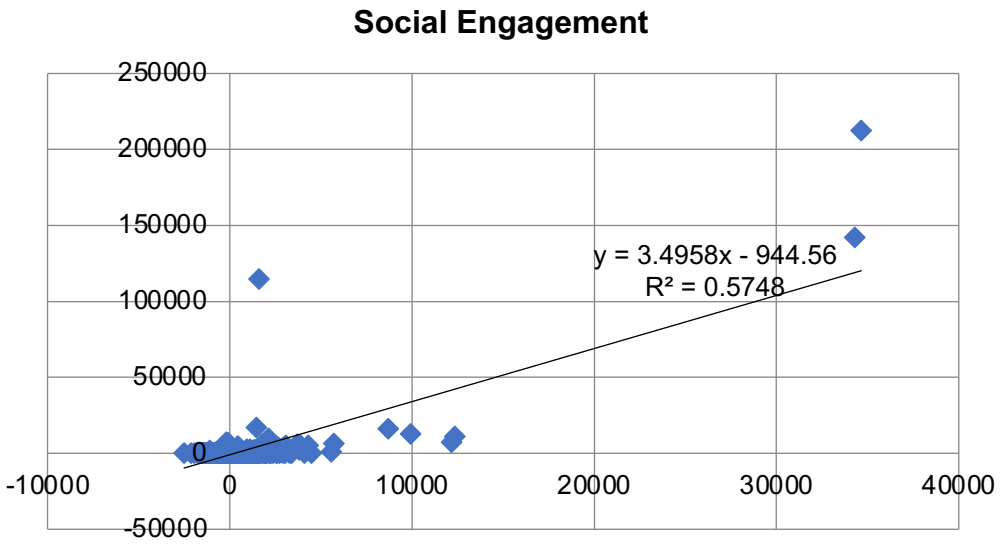


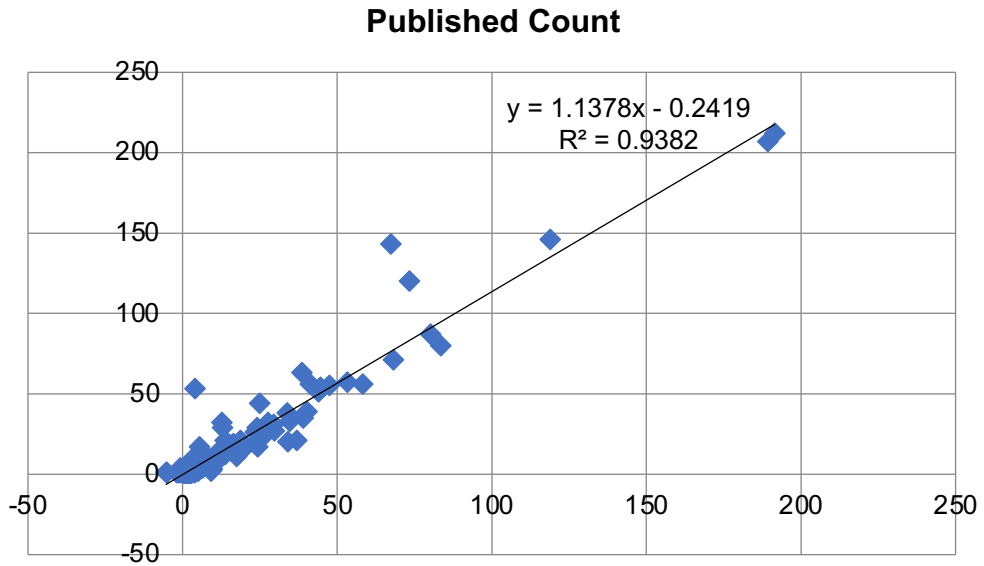
Transference Score

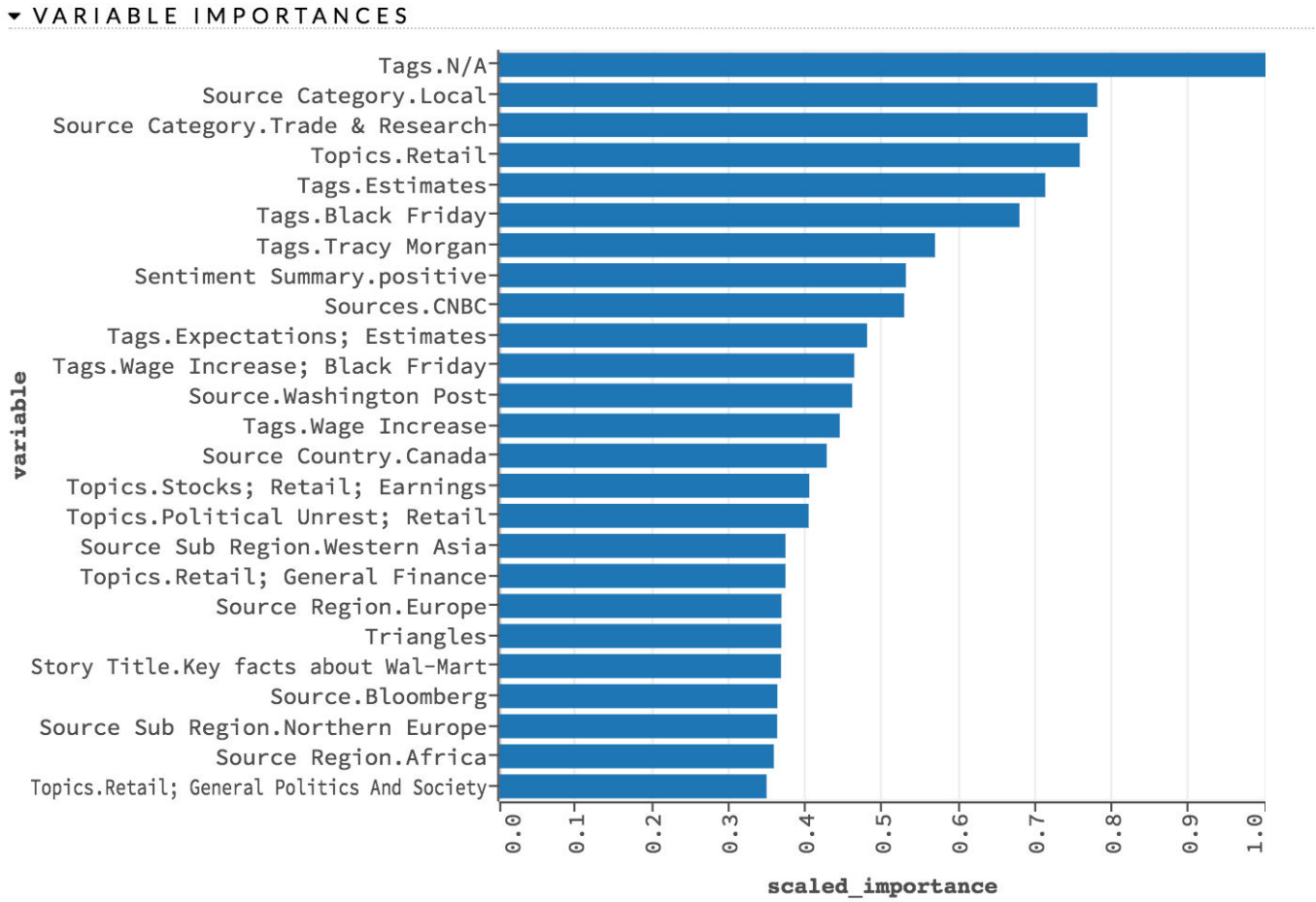
Plotting the articles and speech copy and quantifying the deficits & similarities in how they relate gives a “Transference Score”. For the CEO speech then contrasting it with how it was covered gives us a score of .80. This means there is about 80% message pull through at the cognitive and emotional level.



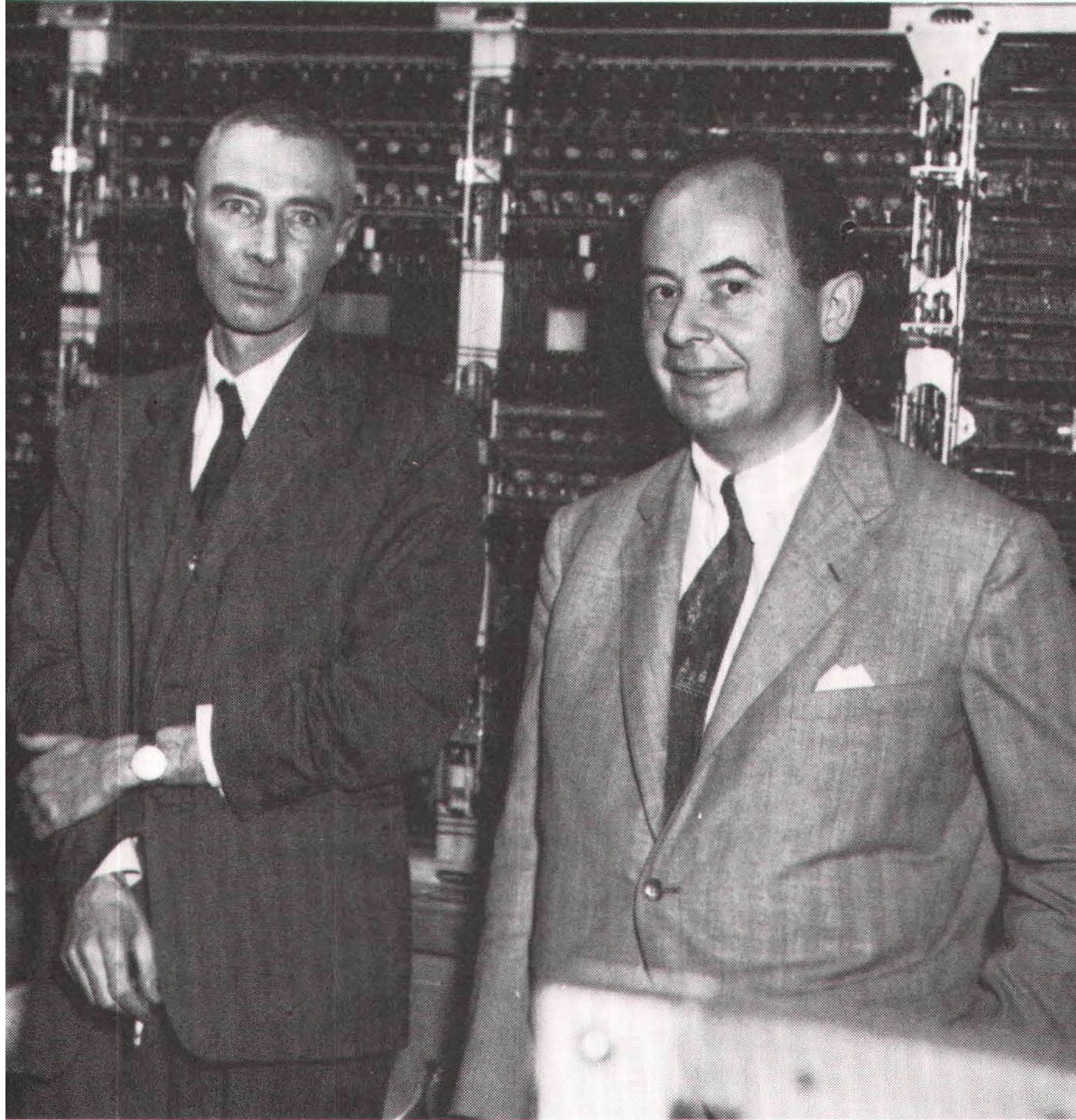
Predicting Engagement







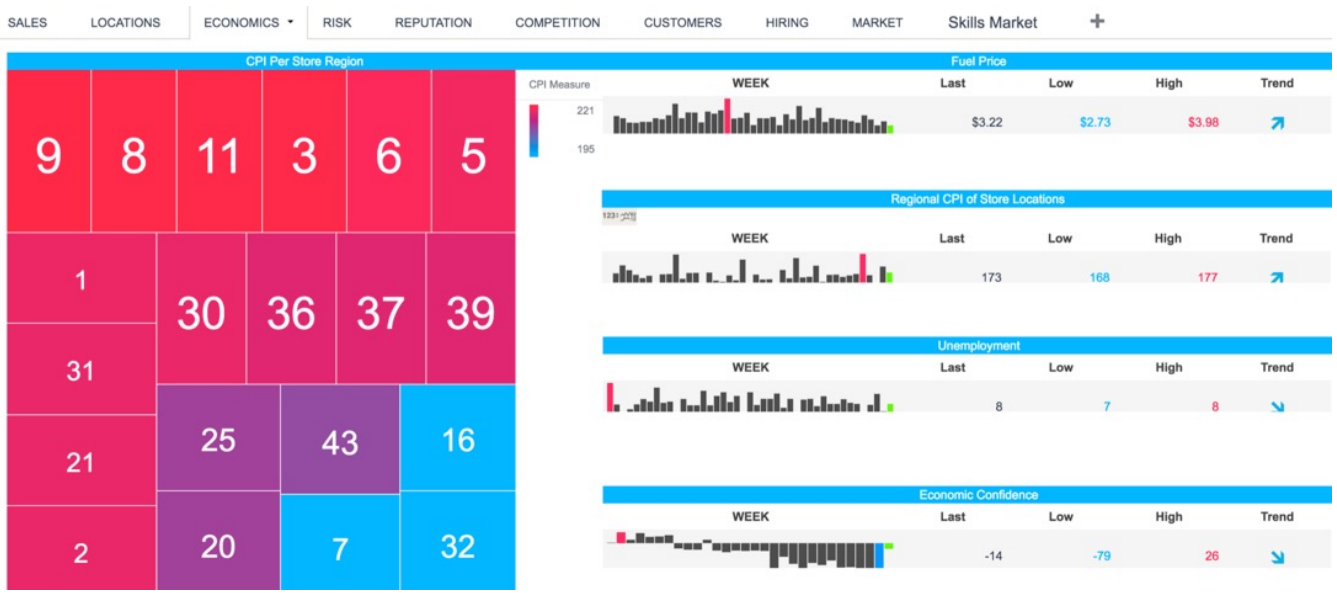
MGMT

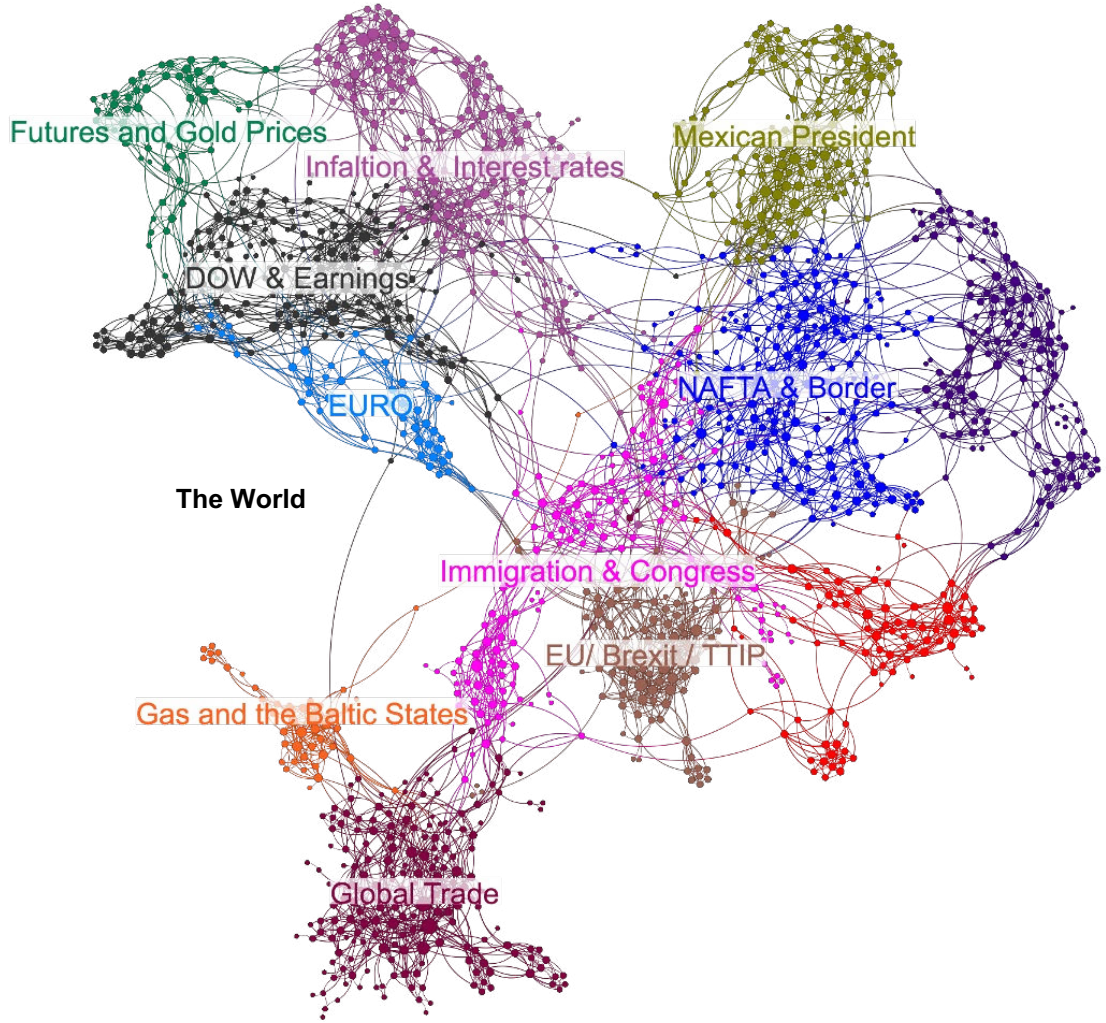


Defer Judgment

The world is interconnected, ambiguous, and complex most of the time. As such, so are the data and outputs, as illustrated by the network below, which uses natural language processing to connect thematics from OSINT. **Decision-makers should learn to defer judgment -** not immediately revert to their heuristics if they do not understand machine-derived outputs. While humans crave black-white classification (like dashboards), accepting ambiguity and probabilistic thinking leads to better decisions.

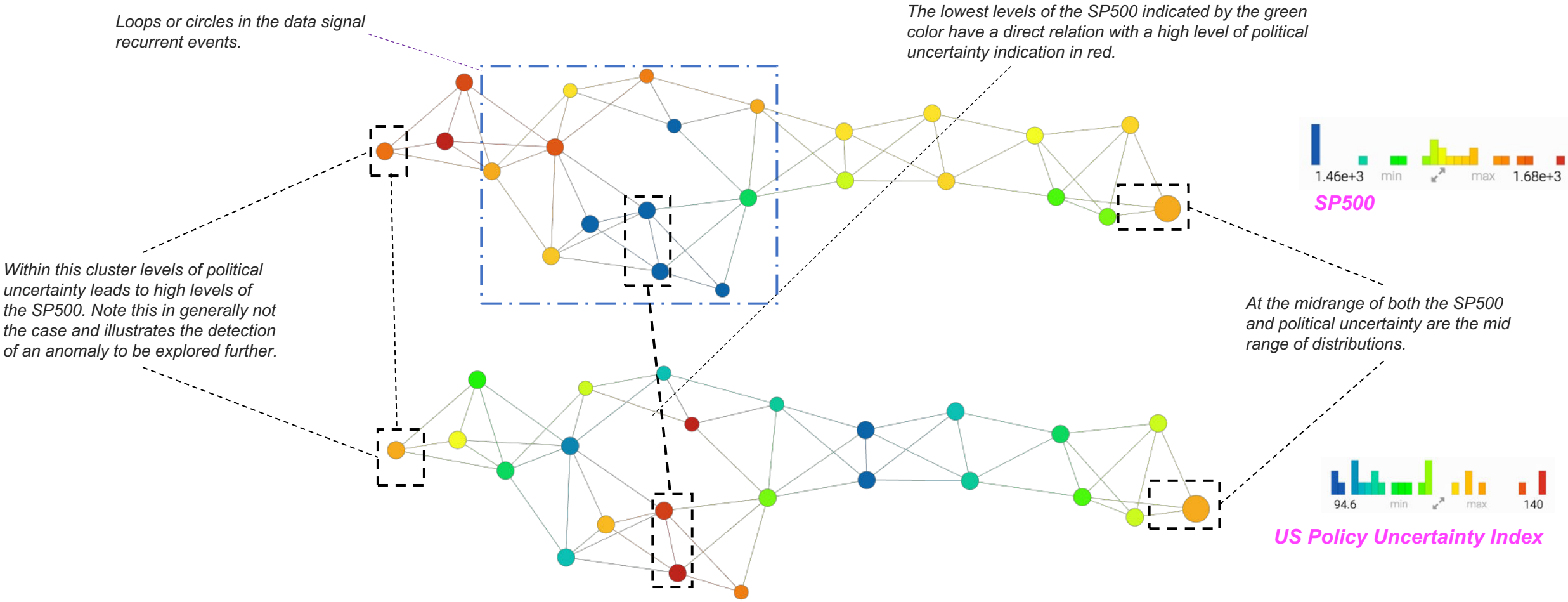
How we want the world





Advanced Analysis – Topological Data Analysis

Topological Data Analysis (TDA) explores the underlying structure within high-dimensional datasets, drawing from the concept that data possesses shape, and that shape carries significance. The technique generates a condensed representation of a broader dataset, unveiling patterns by examining the geometric relationships among all the data points. This approach enables political analysts to extract insights from intricate, multidimensional data more rapidly, reducing dependency on human-generated hypotheses. As a result, analysts can concentrate on strategically applying the acquired insights rather than struggling to interpret the data.



How to Think

- Only look to apply information where it can be affected – there are drastic biased consequences to people looking at too much information (most of which is irrelevant) without context (emails, twitter, insider information).

- It's important to understand that the best operations and information that need this level of speed and quality are often times abstract and uncongenial to standard processes or thinking.

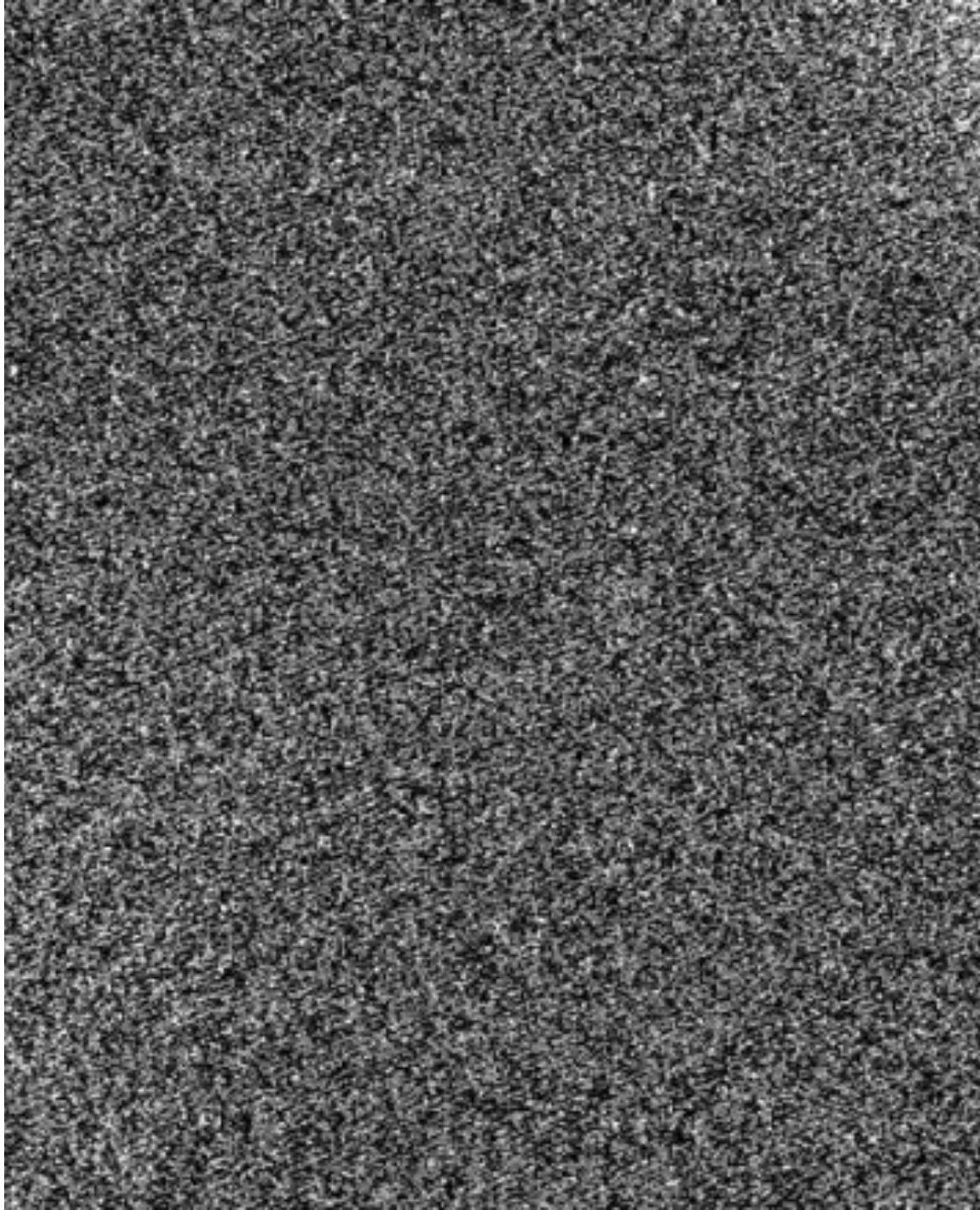
- In most cases the unorthodox and differentiation are the only things drive a competitive advantage - not best practices.

Understand Algorithm Aversion

Research shows that evidence-based algorithms more accurately predict the future than do human forecasters. Yet when forecasters are deciding whether to use a human forecaster or a statistical algorithm, they often choose the human forecaster.

- people are especially averse to algorithmic forecasters after seeing them perform, even when they see them outperform a human forecaster.

- people more quickly lose confidence in algorithmic than human forecasters after seeing them make the same mistake.

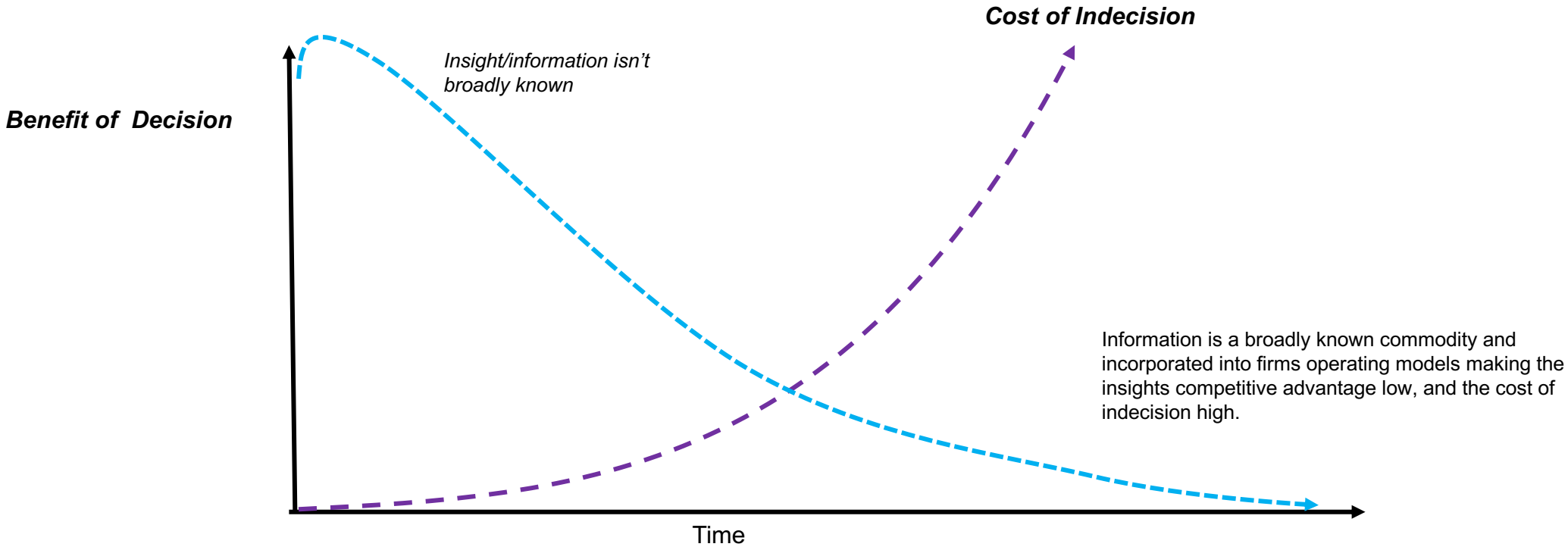


Information to Action

Due to machine learning and data abundance, the flip side is the rapidly diminishing value of extracted insights. As a result, firms need to create operations that can execute on insights/intelligence as quickly as possible.

•Margins of a competitive edge will shrink but will become exponentially more valuable (think search engine adoption of Google V Bing while they are technically on par, Google's market capture is unrivaled).
•Machines also reframe what business and strategy mean. It must be more creative, precise, and mindful of how an organization's assets can be configured to create new value, not optimize traditional processes.

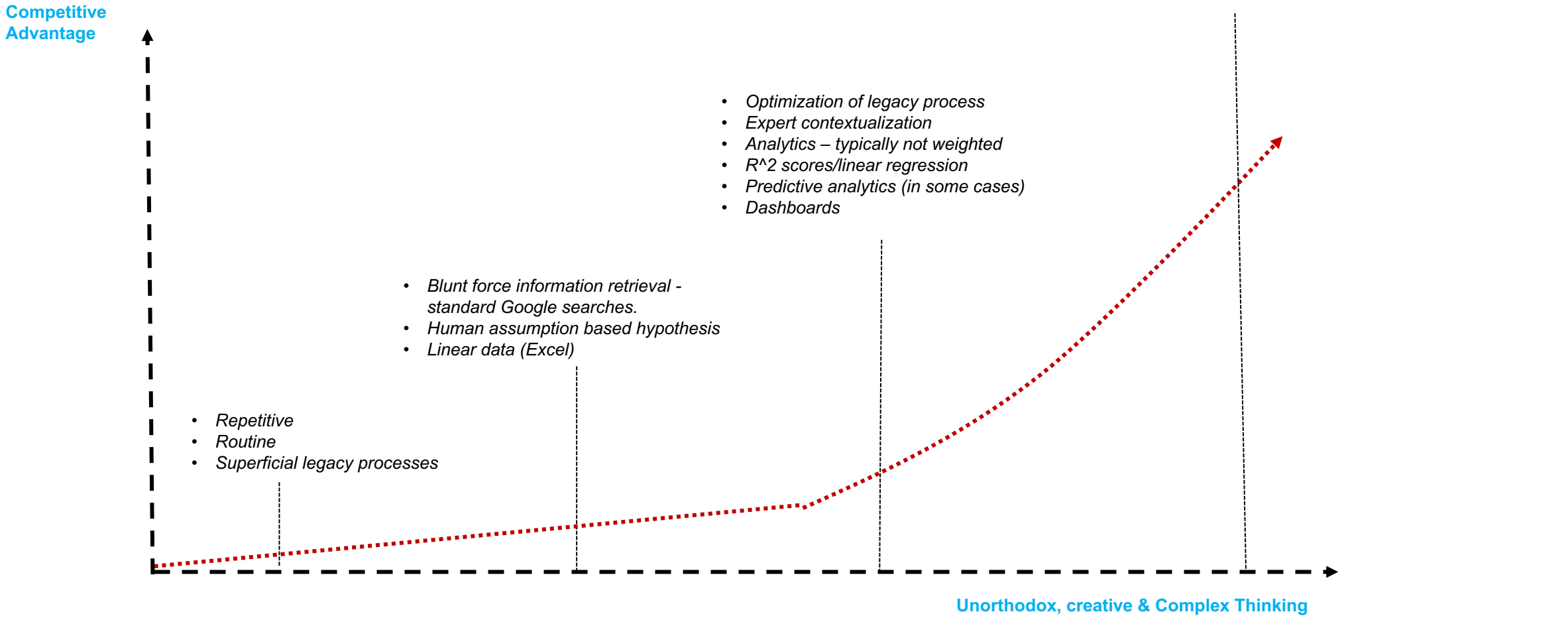
A decision's value exponentially highest in the beginning – and often more more valuable than a perfect decision made later.



Unorthodoxy Is The Advantage

Using "best practices" and optimizations create a marginal competitive advantage - they are known commodities. Upside exists at the convergence of unestablished markets, technologies or actors.

- *Contextual weighting of disparate data streams and knowledge to develop new processes and solutions.*
- *Ok with the ambiguous/unorthodox/divergent*
- *Take advantage of uncertainty and complexity, not run from it.*



Datafying Decisions and Process

Datafying involves using different data sets to replicate your decision-making architecture so it can be used for modeling and statistical analysis.

1. Think about the process and intelligence you need to making current decisions.

2. Identify the relevant data sets that can be used as a proxy your decision-making processes, such as meeting data, key policy topics, customer data, or operational data.

3. Once you have identified the relevant data sets, you need to clean and preprocess the data to prepare it for analysis.

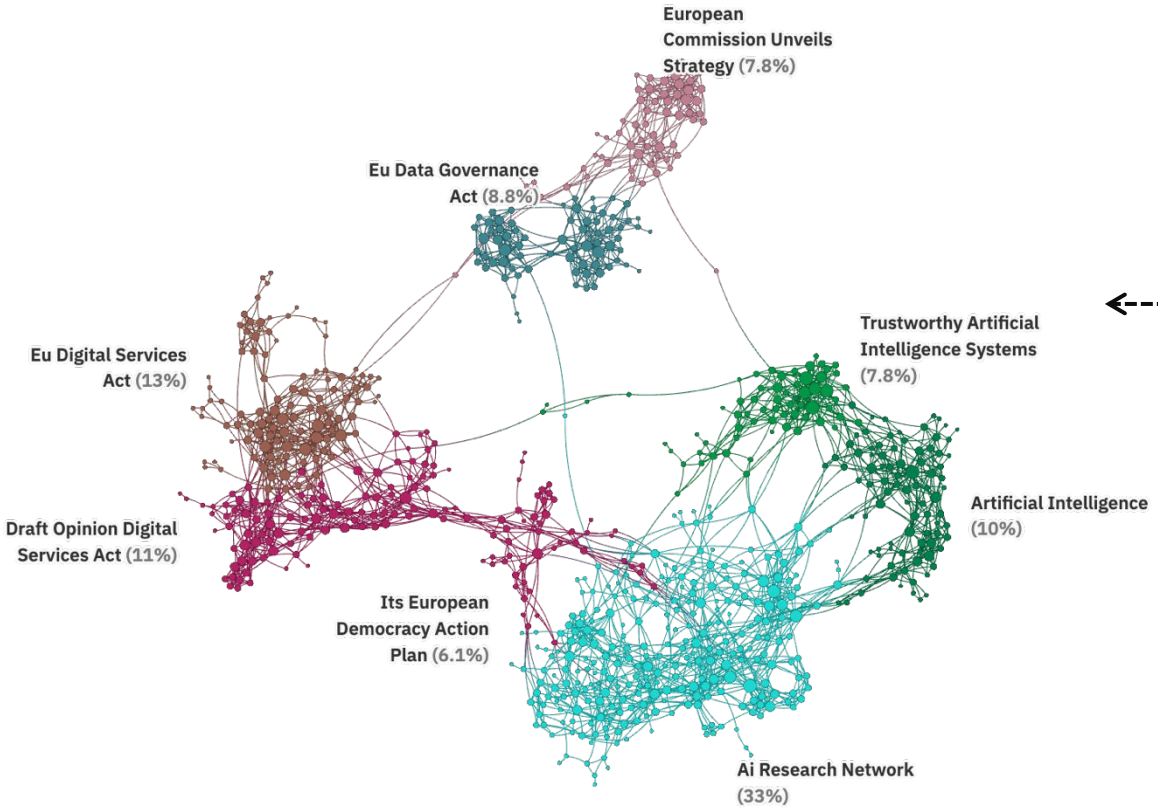
4. Statistical tools and machine learning algorithms can be used to analyze the data and build predictive models that replicate your decision-making processes.

Get Started On The Cheap

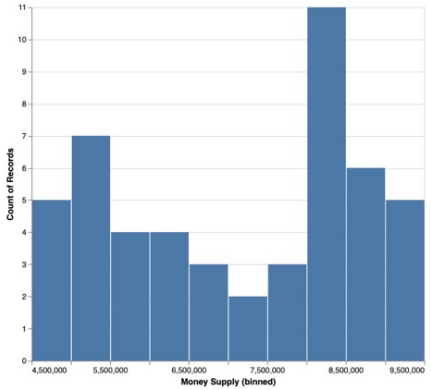
Tool	What it does	Costs
Google Trends	Google that allows users to see how often certain keywords or phrases have been searched for over time. It provides data on the popularity of keywords and phrases, and allows users to compare the popularity of different keywords and phrases. Google Search Trends API shows how often certain keywords or phrases have been searched for on Google over time. It provides data on the popularity of keywords and phrases, and allows users to compare the popularity of different keywords and phrases on a scale of 1-100.	Free
Keyword tool	keyword tool is a type of software or online application that helps users to find and analyze relevant keywords and key phrases for their website or marketing campaign. It is commonly used in search engine optimization (SEO) and pay-per-click (PPC) advertising.	100
Wolfram Alpha	Wolfram Alpha's database contains information on a wide variety of topics, including mathematics, science, geography, history, sports, and more. Users can enter natural language queries, mathematical expressions, or data sets, and Wolfram Alpha will provide a response that is often presented in the form of a graph, chart, or other visual representation.	69 per year
Jasper AI	Jasper is an AI platform that helps teams create content tailored for their brand quickly and efficiently. It offers more than 52 writing templates for various types of content, such as personal bios, Instagram captions, product descriptions, and real estate listings. It also uses an algorithm to generate content related to the input provided.	69 per month
ChatGPT	ChatGPT is an artificial intelligence language model developed by OpenAI, one of the leading AI research organizations in the world. It is designed to understand and generate human-like natural language responses to a wide range of questions and topics. It can be used in a variety of applications, including chatbots, virtual assistants, and customer service automation. It can understand and generate responses to questions on a wide range of topics, from general knowledge questions to specialized technical queries.	20 per month
Brand Watch Media Monitoring	Brandwatch is a social media monitoring and analytics platform that provides businesses and organizations with insights into their brand reputation, customer sentiment, and industry trends. It is used by companies, governments, and non-profit organizations to monitor and analyze social media conversations and engagement across a variety of platforms, including Twitter, Facebook, Instagram, and YouTube.	200-15,000 depending on topics tracked.
Exploratory	Exploratory.io is a data analysis platform that allows users to explore, visualize, and share their data. It is designed to be easy to use, and includes a range of tools for data cleaning, transformation, and visualization. Exploratory.io is often used by data analysts, data scientists, and business analysts who need to quickly understand and communicate their data. It supports a wide range of data sources, including CSV, Excel, and SQL databases, and allows users to combine and merge data from multiple sources. Exploratory.io also includes tools for collaboration and sharing, so users can work together on data analysis projects or share their results with others.	49-120 per month

Iterate With Fly Wheel

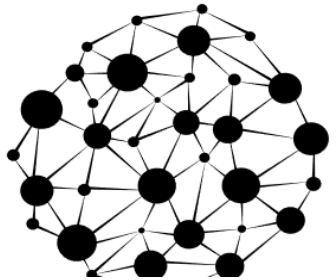
Unstructured Data + NLP = Augmented Intelligence
Augment our lens a domain with the use of machines and Open-Source Intelligence (OSINT) to surface trends and or inform what structured data sets may be of interest to interrogate.

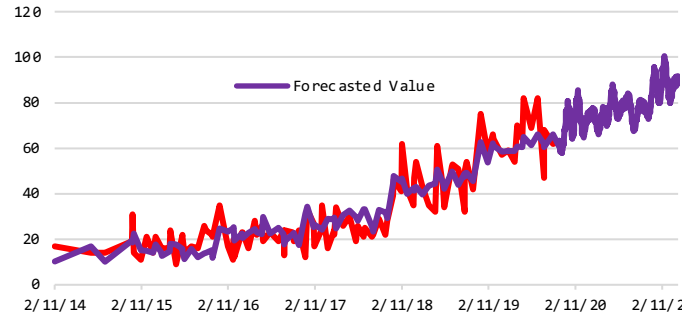


Alternative & Traditional Structured Data
Insights surfaced from OSINT guides which structured data (financial markets, economic, internal) maybe relevant to the domain.



Machine & Data Driven Policy Research & Foresight
Insights surfaced from OSINT connects core themes and issues within a policy and produces more accurate policy thematic and economic forecasts.





This Stuff Is Real

AI systems and frameworks are proven at the biggest companies in the world. Firms that embrace these technologies can see the world with 1000x faster speed and context, so they can focus on higher value & more creative tasks.

Outcome

1. Reduce labor costs by 1000%-10,000% while getting better/faster outcomes over manual analysis.
2. Real-time monitoring of policies, institutions, briefings, memos, policy areas, and “trending” policy areas equals more empathetic communications.
3. Quantitatively rank key policy areas for any member state, politician or institution (great when asking for budgets).
4. Find latent adjacent policies, risks and market trends that are influential.