

RAISE-25

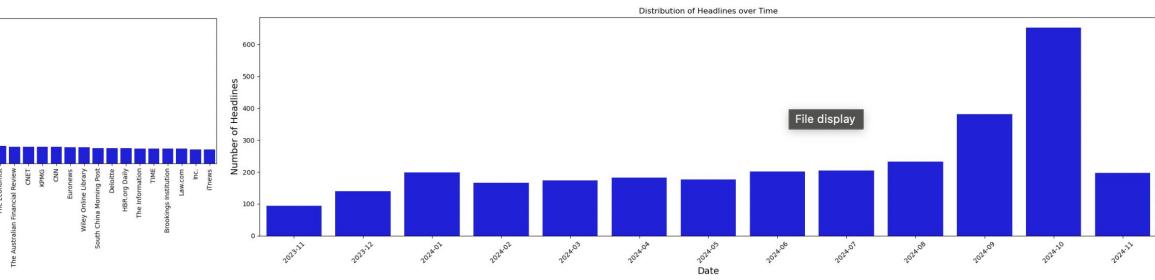
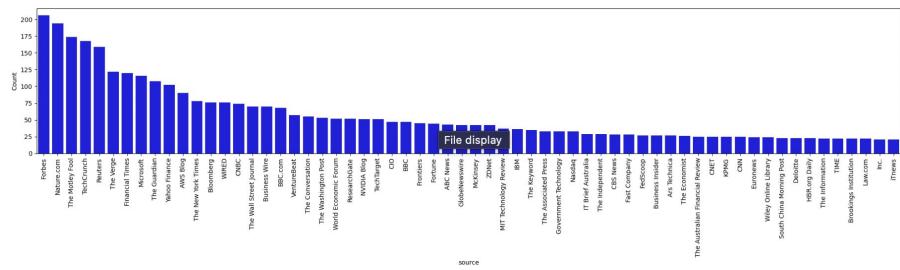
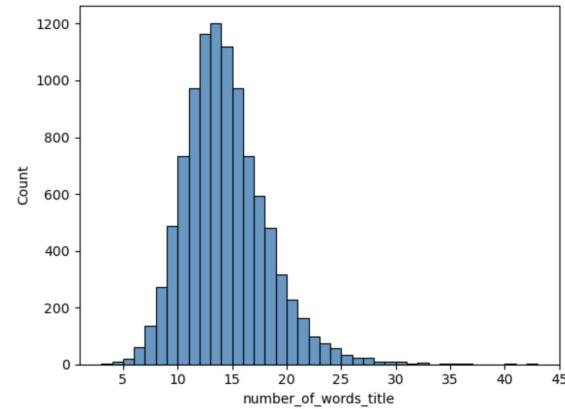
Our Future with AI: Utopia or Dystopia

Discovering the promises and challenges of a future with AI by analyzing news headlines.

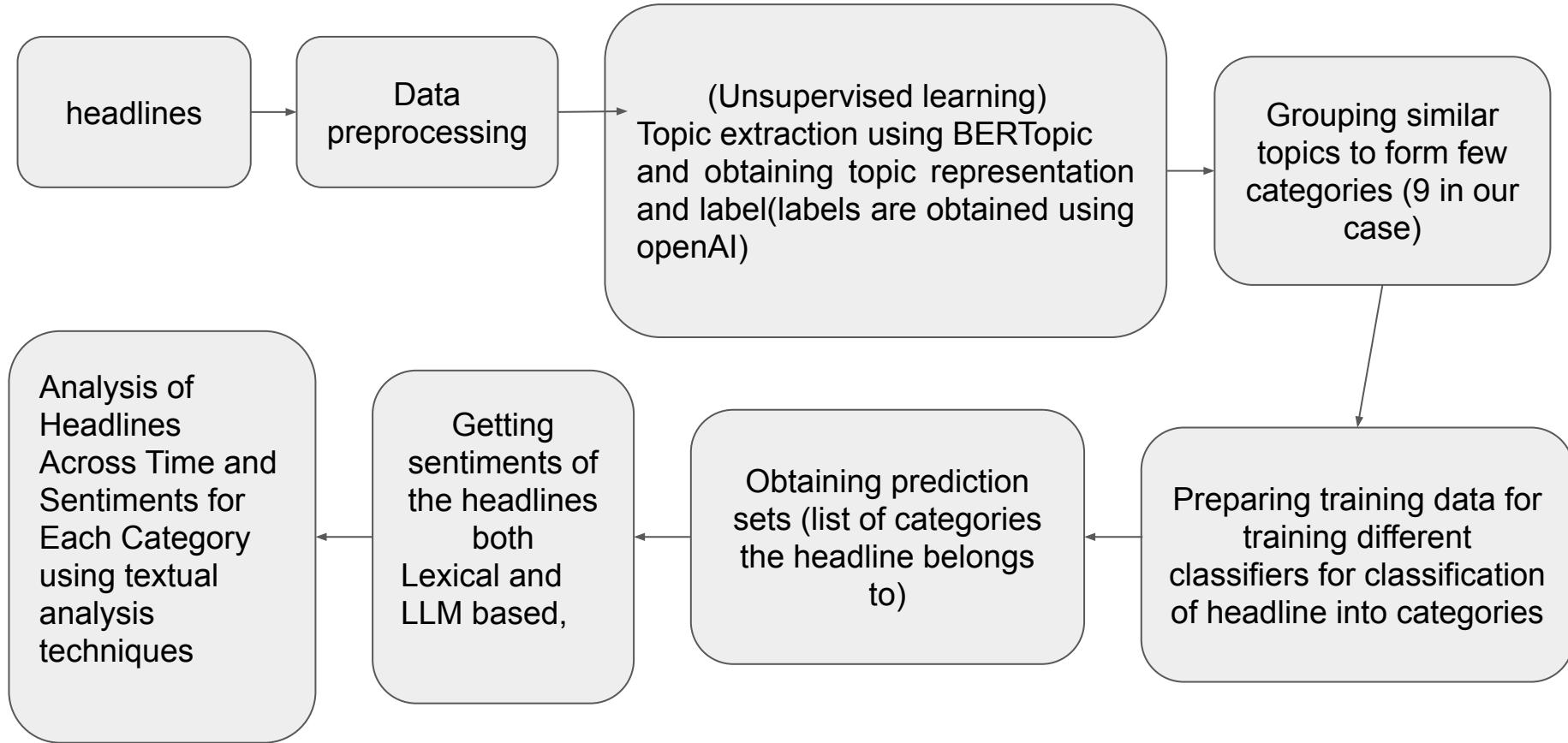
Amrutha Karuturi

Initial Observations From Data

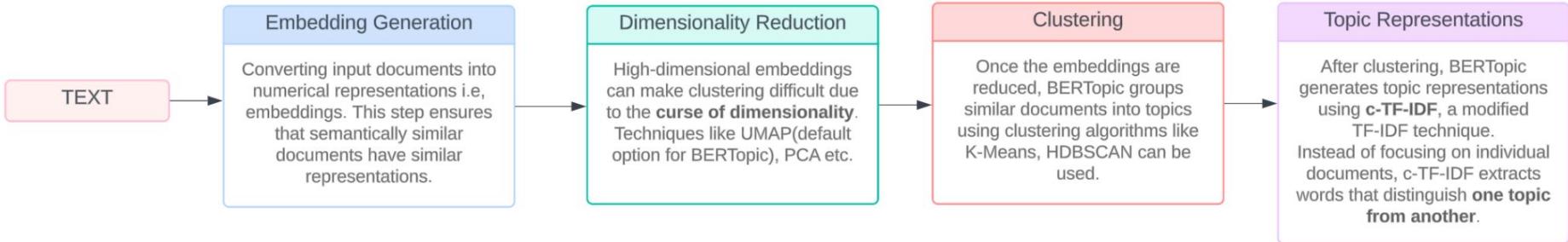
1. The no.of news headlines increasing at a faster pace in sept and oct 2024 (reasons to be investigated). The sudden drop in count for Nov 2024 is because the dataset has headlines only till Nov-11-2024 which is only $\frac{1}{3}$ of the month..
 2. The average no.of words per headline seem to be 13.
 3. Some of the main sources for the headlines are Forbes, Nature.com, The Motley Fool etc.
 4. Given category column do not form good representation for the headlines as one headline may cover multiple categories so some unsupervised methods need to be used to identify and classify the headlines into appropriate categories.



Project WorkFlow



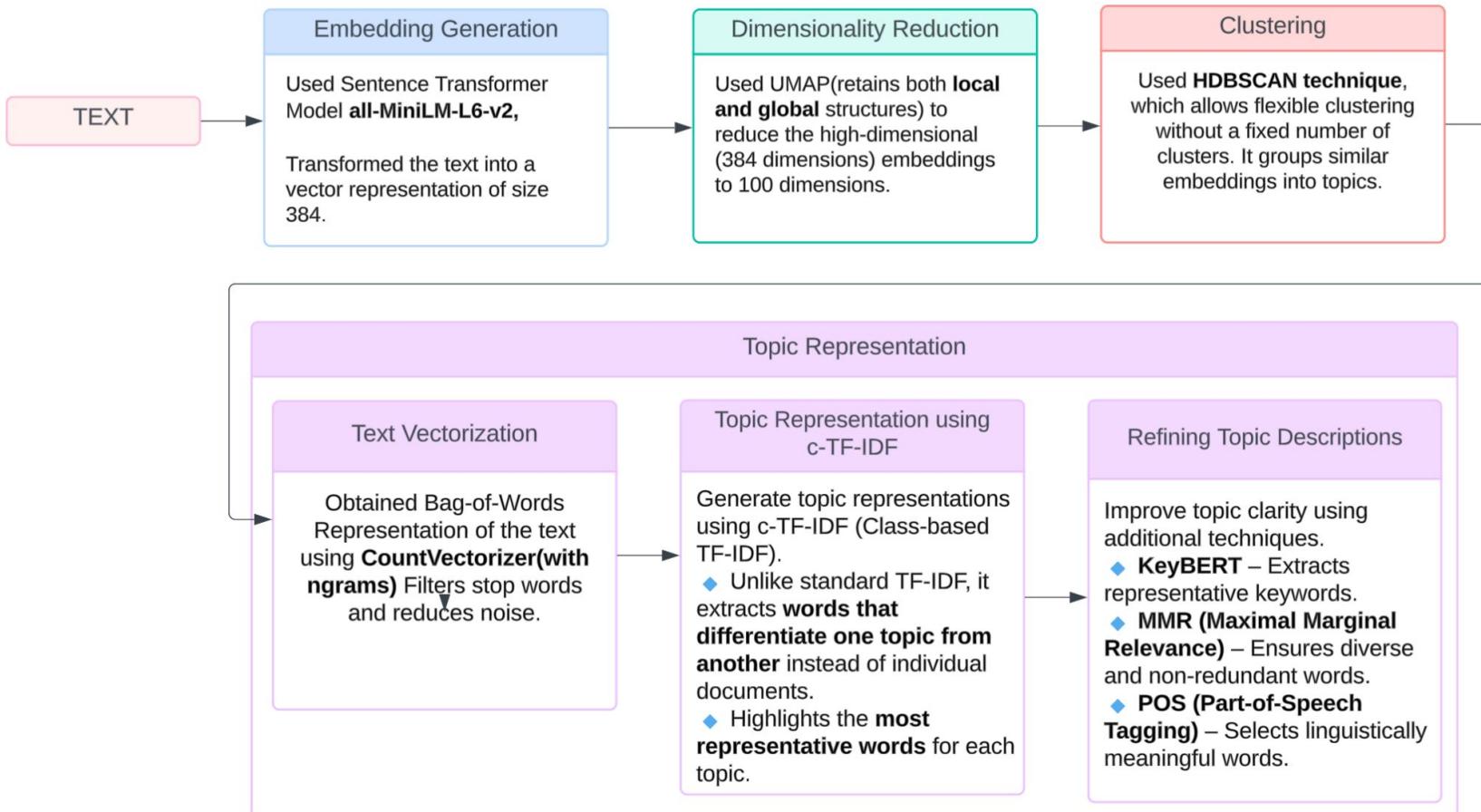
BERTopic



Key features of BERTopic:

- Modularity:** You can customize each step independently.
- Flexibility:** Different embedding models, clustering methods, and topic representations can be used.
- Interpretability:** c-TF-IDF enhances the understanding of topics.

BERTopic Workflow: From Text to Topics



Outputs of BERTopic

Topic	Count		Name	Representation	KeyBERT	MMR	POS	Representative_Docs
0	-1	3310	-1_human_chief_good_ibm	[human, chief, good, ibm, officer, ai officer,...]	[ai officer, ai policy, launches ai, chief ai,...]	[human, ibm, ai officer, chief ai, sap, white ...]	[human, chief, good, officer, house, machine, ...]	[News ESSC UAH scientist earns National ai...]
1	0	808	0_health_care_healthcare_medical	[health, care, healthcare, medical, medicine, ...]	[healthcare ai, ai healthcare, ai medical, hea...]	[healthcare, ai healthcare, ai health, healthc...]	[health, care, healthcare, medical, medicine, ...]	[ai in Health Care, AI in Health Care: Powerin...]
2	1	412	1_education_schools_school_students	[education, schools, school, students, classro...]	[ai education, ai schools, ai academic, ai cla...]	[education, schools, students, ai education, a...]	[education, schools, school, students, classro...]	[Exploring the impact of ai on higher educatio...]
3	2	281	2_generative_ai_generative_ai generative_use g...	[generative ai, generative, ai generative, use...]	[ai generative, generative ai, ai creative, br...]	[generative ai, ai generative, use generative,...]	[generative, usage, enterprise, creative, diff...]	[Executive Conversations: Putting generative A...]
4	3	260	3_jobs_job_hr_hiring	[jobs, job, hr, hiring, workers, ai jobs, empl...]	[ai workplace, ai jobs, workforce ai, ai job, ...]	[jobs, hr, ai jobs, employers, ai workplace, w...]	[jobs, job, hiring, workers, employers, employ...]	[AI will affect 40% of jobs and probably worse...]

- BERTopic returns the **topic number, topic name, topic representation** (top distinguishing words from c-TF-IDF), and the **top 3 representative documents** for each topic.
- Topic **-1** indicates **outliers**, documents that don't fit into any topic.
- BERTopic formed **88 topics**, with each topic containing **headlines** that share similar themes or characteristics, providing a cohesive group of related headlines within that topic.
- The model can be found the folder with name `tmod2`.
- The name for each topic can be further refined by prompting **OpenAI**, to label the topic based on **keywords** and **representative documents** to generate more meaningful and descriptive topic names.

Labeling each Topic with openAI

The **representative documents** returned by BERTopic are typically limited to 3, which may not always be sufficient to generate a comprehensive label for the headlines within a topic. To address this, the **top 20 representative documents** for each topic were used by using `._extract_representative_docs()` on the BERTopic model. Additionally, various representation methods, such as **KeyBERT**, **MMR (Maximal Marginal Relevance)**, and **POS (Part-of-Speech)**, were combined to form a set of **keywords** that more effectively capture the essence of each topic.

The following prompt was then used to generate topic labels based on the combined **keywords** and **representative documents**:

```
Prompt : "I have topic that contains the following headlines:+>docs+ '. The topic is described by the following keywords: +'  
kwords+ '. Based on the above information, can you give a short label of the topic?'
```

Some of the examples for the topic labels:

response for prompt 0: "AI Innovations in Healthcare: Enhancing Patient Care, Diagnosis, and Treatment"

response for prompt 1: "Integrating AI in Education: Opportunities, Challenges, and Innovations"

response for prompt 2: "Exploring the Impact and Adoption of Generative AI Across Industries"

response for prompt 3: "Impact of AI on Job Security and Workforce Dynamics"

response for prompt 4: "AI Chip Market Competition: Intel, AMD, and Alternative Stocks to Nvidia"

response for prompt 5: "Impact of AI on Search Engines: Google's Innovations and Challenges"

response for prompt 6: "Wall Street Insights: Promising AI Stocks and Investment Opportunities"

response for prompt 7: "AI Chatbots: Innovations, Controversies, and Legal Challenges"

All this information can be in 'topic_info2.csv'

Analyzing the Topics

After Analysis the topic labels we have grouped the possible labels into following categories:

1. Education - AI's role in education, including its impact on learning, teaching methods, and academic institutions.
2. Career - AI's influence on jobs, employment trends, small business and professional development.
3. Society:
 - a. AI impact: From industries to enterprises :
 - i. AI in various industries - The adoption and application of AI in various industries.
 - ii. AI in companies & Enterprises- The adoption and application of AI in various enterprises, including AI-driven product innovations, developments, and challenges faced by businesses.
 - b. AI overview, risks & impact - Broad perspectives on AI, societal impact, energy consumption, and associated risks.
 - c. AI Governance & Geopolitics - Government actions and policies on AI, including national AI strategies, regulations, global AI summits, international collaborations, diplomatic discussions, and AI-driven geopolitical competition
 - d. AI Investments & Market Trends - Funding, grants, investments, market evaluations, and stock trends related to AI and AI-driven products.
 - e. AI Ethics, Law & Policy - Ethical concerns, legal challenges, privacy issues, copyright disputes, bias in AI, compliance with regulations, and societal implications of AI laws.
4. Other

For categorizing the topics, we manually reviewed and selected the best topics for each category. To ensure accuracy, we cross-referenced the topics by prompting **OpenAI** to assign them to the predefined categories. After carefully checking the results, we selected a few topics that we were confident belonged to these categories.

Some example topics for each category

Topic number : openAI label

<p>Education</p> <p>1 : "Integrating AI in Education: Opportunities, Challenges, and Innovations"</p>	<p>AI in various industries</p> <p>0 : "AI Innovations in Healthcare: Enhancing Patient Care, Diagnosis, and Treatment"</p> <p>13 : "AI-Driven Innovations in Retail and Customer Experience"</p> <p>19 : "AI in Financial Services: Opportunities, Challenges, and Future Trends"</p> <p>28 : "AI's Transformative Impact on Sports: Innovations, Challenges, and Future Directions"</p> <p>32 : "AI Innovations in Wildlife Conservation and Biodiversity Monitoring"</p> <p>56 : "AI Innovations and Applications in Advanced Manufacturing"</p> <p>58 : "AI's Transformative Impact on Fashion and Beauty Industries"</p>	<p>AI in companies & Enterprises</p> <p>5 : "Impact of AI on Search Engines: Google's Innovations and Challenges",</p> <p>12 : "Generative AI Innovations and Applications on Amazon Bedrock and AWS",</p> <p>16 : "Apple's AI Innovations and iOS 18.1 Features for iPhone 16"</p>
<p>Career and Workforce</p> <p>3 : "Impact of AI on Job Security and Workforce Dynamics"</p>		
<p>AI Investments & Market Trends</p> <p>8 : "AI Startup Funding and Market Growth Trends"</p> <p>35 : "AI Investment Trends and Opportunities",</p> <p>67 : "Funding and Grants for AI Research and Training Initiatives"</p>		
<p>AI Ethics, Law & Policy</p> <p>9 : "AI in the Legal Field: Impacts, Challenges, and Ethical Considerations"</p> <p>21 : "AI Regulation and Legislative Developments",</p> <p>53 : "AI and Copyright: Legal Challenges and Industry Implications"</p>	<p>AI overview, risks & impact</p> <p>31 : "AI's Energy Impact: Balancing Efficiency, Emissions, and Sustainability"</p> <p>52 : "AI-Driven Sustainability and Waste Management Innovations"</p>	<p>AI Governance & Geopolitics</p> <p>30 : "China-US AI Competition and Diplomacy",</p> <p>37 : "Global AI Governance: Frameworks, Challenges, and Collaborative Approaches",</p> <p>44 : "Global Initiatives and Collaborations for AI Safety and Security",</p>
		<p>Other</p> <p>41 : "AI Training and Data Strategies: Opportunities and Challenges"</p> <p>42 : "Comprehensive Overview of AI: Applications, Terminology, and Implications"</p>

Training classifiers

1. Training Data Formation:

1. Forming the labelled data:

- Extracting representative documents: The extract_representative_docs() function from BERTopic was used to retrieve the most representative documents for each topic.
- This function selected documents that uniquely represented each topic based on their closeness to the topic representation.
- The initial dataset obtained after this process consisted of: Society: 3,797 samples, Education: 300 samples, Career & Workforce: 300 samples, Other: 100 samples

2. Train-Test Split (Unbalanced Dataset):

- The dataset was split into an 80-20% train-test split, keeping the test set the same for all experiments.
- The resulting training dataset contained: Society: 3,037 samples, Education: 240 samples, Career & Workforce: 240 samples, Other: 80 samples
- The test set remained fixed across different experiments.

3. Balanced Training Data Formation

- To create a balanced dataset, 240 samples were selected for each category.
- This was done by sampling 240 data points from the 3,037 Society training samples, ensuring equal distribution across its subcategories, Up-sampling the Other category to reach 240 samples.

This process ensured that the training data was tested in both unbalanced and balanced scenarios while keeping the test set consistent across different methods.

Total Labelled data

broad_category	count
{Society}	3797
{Education}	300
{Careers & Workforce}	300
{Other}	100

Unbalanced training data

broad_category	count
{Society}	3037
{Careers & Workforce}	240
{Education}	240
{Other}	80

Balanced training data

broad_category	count
{Other}	240
{Careers & Workforce}	240
{Society}	240
{Education}	240

2. Training classifiers:

We trained 2 classifiers:

1. The first classifier determines whether a headline belongs to one of the broad categories: **Education**, **Career**, **Society**, or **Other**.
2. If a headline is classified under **Society**, a second classifier is used to assign it to one of the **Society subcategories**.

Models and Training Strategies:

1. The models tested were: Logistic Regression, Decision Tree, Multi-label Classifier
2. For each model, different training strategies were evaluated: Imbalanced Training Data, Balanced Training Data, Imbalanced Training Data with Class Weights

Performance evaluation: Accuracy is evaluated both with and without prediction sets. For the multi-label classifier, prediction sets are formed using a default threshold of 0.5, meaning categories with probability values greater than 0.5 are included in the prediction set, as each probability is an independent value. In the case of the logistic model, where the sum of probabilities equals 1, threshold values of 0.3 and 0.4 were used instead.

Table 1: Performance Comparison of Different Models

broad Model	accuracy	soc_model	accuracy
Logistic model with balanced training data	0.853	Logistic model with imbalanced data & class_weights with prediction set	0.8368 (threshold < 0.3)
Logistic model with imbalanced training data	0.894		0.6868 (threshold = 0.3)
Logistic model with imbalanced data & class_weights	0.9488		0.4579 (threshold = 0.4)
Decision tree with balanced training data	0.746	Decision tree with balanced training data	0.7671
Decision tree with Imbalanced data	0.9344	Decision tree with Imbalanced data	0.7315
Decision tree with class weights	0.916	Decision tree with class weights	0.6986
Multi-label classifier balanced data	0.9088	Multi-label classifier Balanced data	0.8882
Multi-label classifier balanced data with prediction set	0.9844	Multi-label classifier Balanced data with prediction set	0.9513
Multi-label classifier imbalanced data	0.9644	Multi-label classifier imbalanced data	0.8579
Multi-label classifier imbalanced data with prediction set	0.9778	Multi-label classifier imbalanced data with prediction set	0.9211
Multi-label for all categories balanced training data	0.7555		
Multi-label for all categories balanced data with prediction set	0.9222		
Multi-label for all categories imbalanced data	0.8078		
Multi-label for all categories imbalanced data with prediction set	0.9267		

3. Performance evaluation:

1. Logistic Regression:

- Performed best when trained on imbalanced data with class weights.
- Society model struggled with finer distinctions; accuracy dropped as the threshold increased for the prediction sets.

2. Decision Tree:

- Achieved 0.9344 accuracy when trained on imbalanced data. Performed similarly when trained on balanced data (0.916), avoiding overfitting.
- Could not generate prediction sets due to binary category assignments.
- Best society model accuracy: 0.7671 when trained on balanced data.

3. Multi-label Classifier:

- Performed best when trained on balanced data with prediction sets.
- Society model achieved highest accuracy (0.9513) when trained on balanced data.
- A single multi-label model for both broad and society categories performed little less(0.9267) when trained on imbalanced data, 0.9222 when trained on balanced data).

Final conclusion: The best-performing approach was to use separate broad and society models with a multi-label classifier trained on balanced data and prediction sets. This approach effectively captured multiple labels while ensuring better classification performance.

Sentiment classification using lexical, LLM, openAI

With the categorization of the headlines complete, the next step is sentiment analysis. We have approached sentiment classification using two main methods:

1. Lexical-Based Sentiment Analysis:

- **TextBlob**: A simple tool that provides polarity (positive/negative sentiment) and subjectivity (objective/subjective) scores.
- **VADER**: A lexicon and rule-based model designed for social media text, which returns sentiment scores for positive, neutral, and negative sentiments.
- **AFINN**: A list-based approach that assigns sentiment scores to words, summing these scores to determine overall sentiment.
- **FLAIR** : classification model en-sentiment has been trained on IMDB dataset.

2. LLM-Based Sentiment Analysis:

- **RoBERTa**: A transformer-based model fine-tuned for sentiment analysis, providing a context-aware prediction of sentiment.
- **OpenAI Prompting**: Using OpenAI's GPT-based models, we prompt the model to analyze sentiment by leveraging the **category prediction sets obtained from classification discussed in previous slide**. This method combines category information with context to generate sentiment labels more accurately.

These approaches are used to provide both **baseline sentiment analysis** (via lexical methods) and **more advanced sentiment analysis** (via LLM-based methods), improving the accuracy and depth of sentiment understanding for the categorized headlines.

```
openAI prompt_template = """Analyze the sentiment of the following news headline and classify it as Positive, Negative, or Neutral, considering the listed categories.
```

```
Headline: "{headline}" , Categories: {categories}
```

```
Instructions:
```

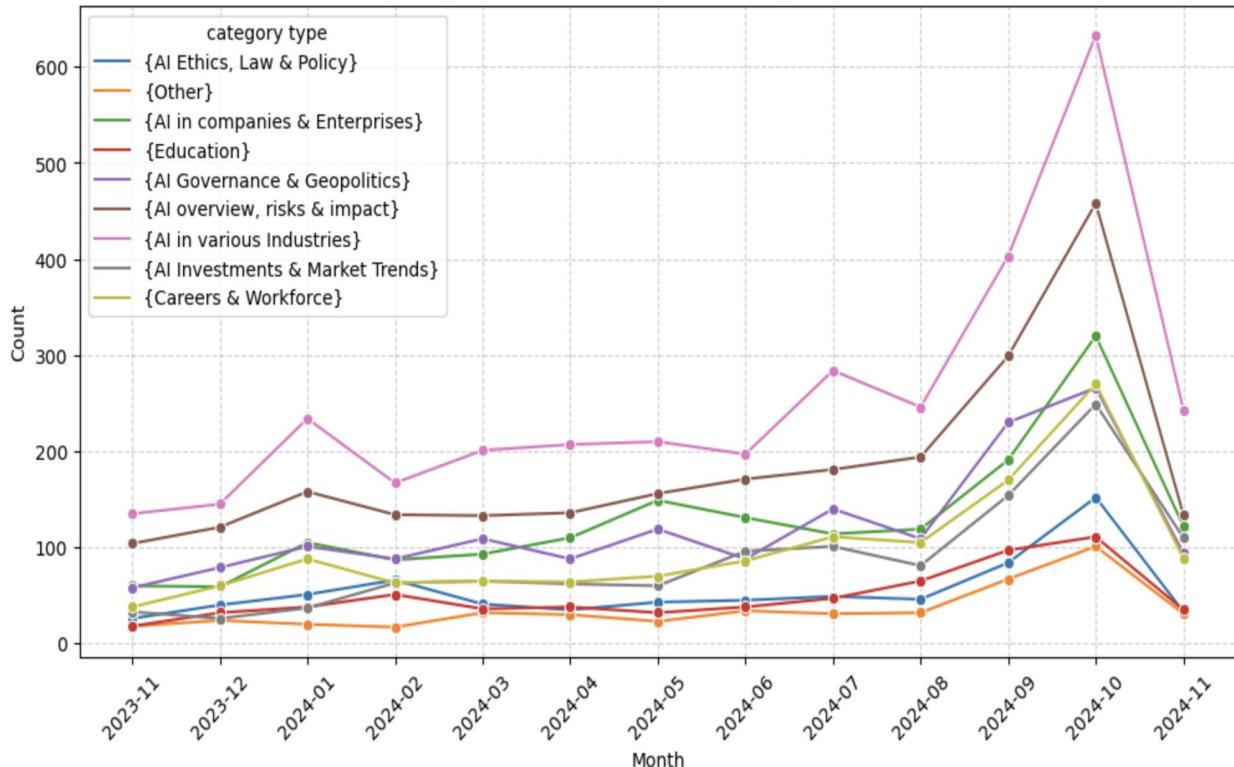
- 1.Focus on the sentiment in the context of the given categories.
- 2.Provide only one sentiment label: Positive, Negative, or Neutral.
- 3.Respond with just the sentiment label-no explanation needed."

```
"""
```

Some examples where openAI sentiment gave better result

Headlines	Other sentiments (Roberta / vader sentiment scores)	openAI sentiment
<p>Using AI to detect heart failure. Using artificial to diagnose infectious keratitis. Cutting edge AI, spots cancer and viruses and cells before symptoms appear. Government funds research into VR and AI to tackle drug deaths.</p> <p>EY claim success in using AI to find audit frauds. How AI is transforming the fight against fraud. Tax authorities adopt AI for tax frauds and efficiencies. Swift uses AI to enhance cross-border payment fraud detection. The audit revolution integrating artificial intelligence and detecting accounting frauds.</p> <p>I thought I would hate AI in home security, is just the opposite. Argentina implements, artificial intelligence to find crime before it happens. AI system to protect athletes from online abuse during Paris, 2024. Ethical applications of AI to public sector problems</p> <p>Coming soon, AI, finding the most painful parts of the white collar jobs. Noise cancelling headphones use AI to let a single voice through.</p>	Negative	Positive
How art artificial intelligence might help healthcare or harm it. AI conflict zones and crisis care takes Centre stage at Geneva Cancer Congress. Testing toxicity using stem cells and AI	Negative	Neutral
Generative AI is coming for healthcare and not everyone's thrill. AI generator songs are getting longer, not necessarily better. The AI industry is obsessed with chart about Arena, but it might not be the best benchmark	Positive/neutral	Negative

Distribution of AI-Related Headlines Across Categories Over Time

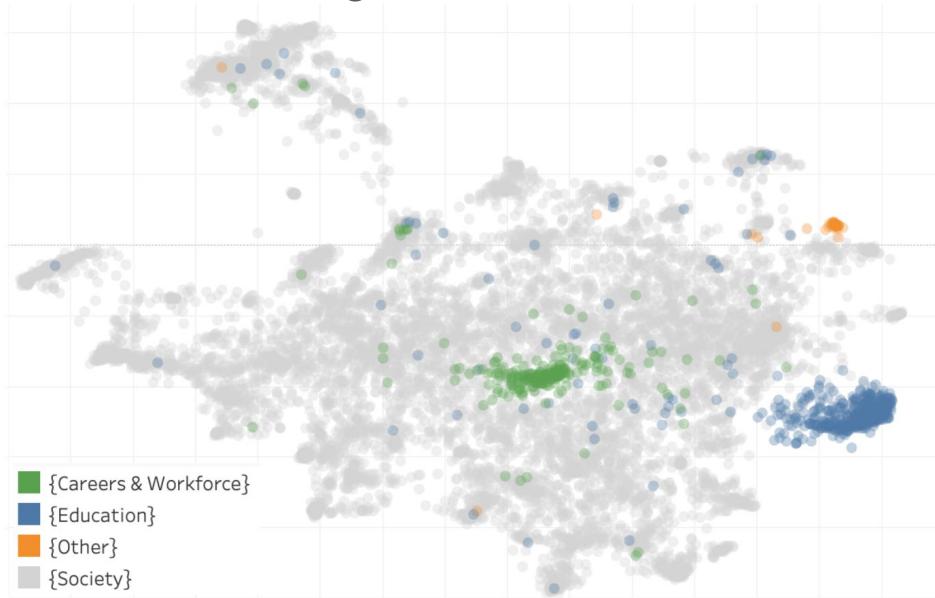


There are more headlines related to **AI in different industries** and **AI overview & impact**, followed by **AI in companies & enterprises**, **AI governance & geopolitics**, **career & workforce**, **AI investment trends**, **AI ethics, law & policy**, and **AI in education**, all showing an increasing trend over time.

The sudden drop in **November 2024** is due to the dataset covering headlines only until **November 11**, representing just one-third of the month.

2D Projection of news headlines

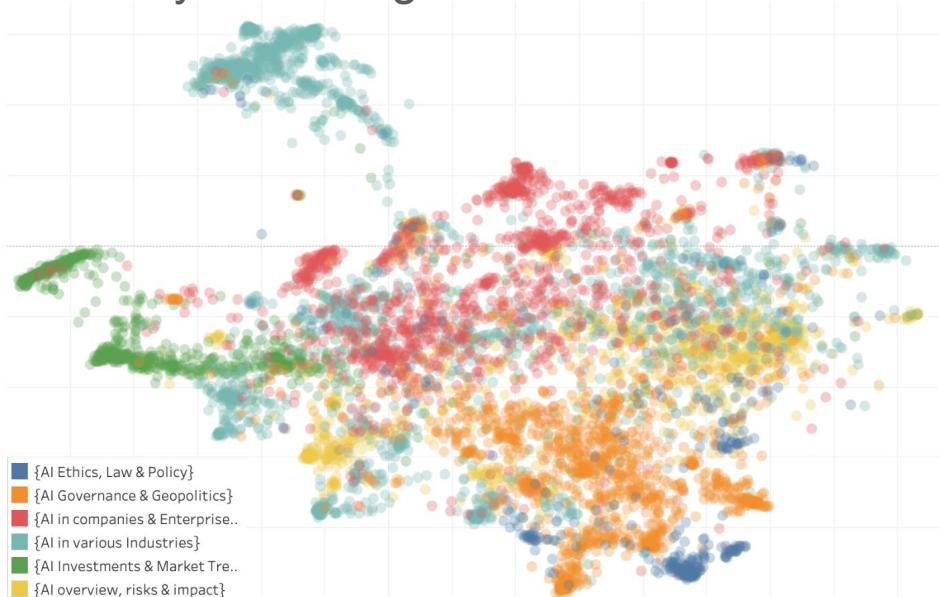
Broad categories



This slide visualizes the 2D projection of news headline embeddings, obtained using the all-MiniLM-L6-v2 sentence transformer—the same embedding model used for BERTopic clustering. By applying dimensionality reduction, the projection provides insight into how well the model captures thematic distinctions. The clear separation of categories suggests that the chosen embedding model effectively differentiates AI-related news themes, reinforcing its suitability for BERTopic-based clustering.

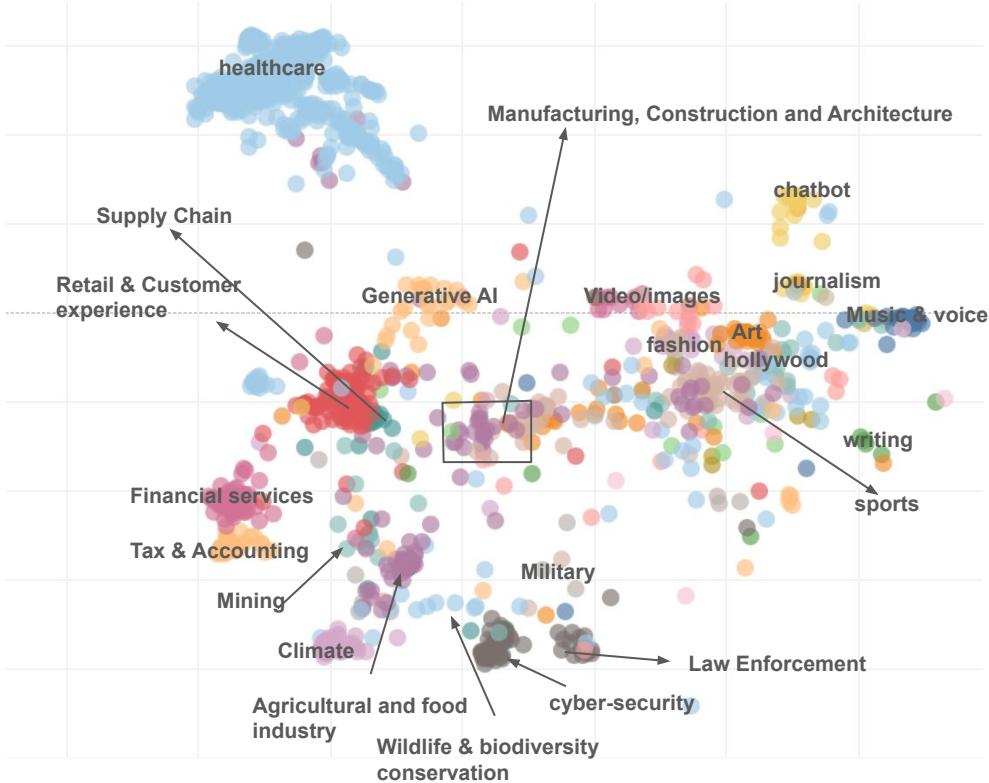
2D Projection of news headlines

Society sub categories



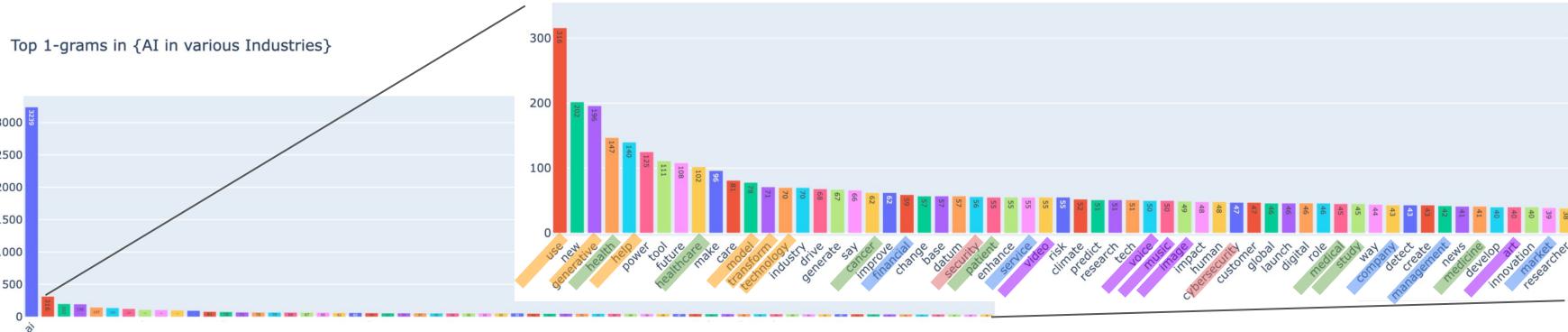
This slide visualizes the 2D projection of AI-related news headlines, now colored based on Society subcategories. While some categories, such as AI Governance & Geopolitics, AI in Companies & Enterprises, AI Ethics, Law & policy, AI investments & Market trends, exhibit distinct clustering, others—like AI Overview, Risks & Impact and AI in Various Industries—are more dispersed, overlapping with multiple subcategories. This distribution suggests that certain AI-related topics span multiple societal aspects, reflecting the interconnected nature of AI discussions in news coverage.

AI in various Industries



This slide provides a detailed exploration of the AI in Various Industries, highlighting key themes identified through an interactive Tableau dashboard. By selecting specific data points, the topics covered in AI-related news headlines were analyzed.

The results reveal natural clustering of related topics—industries with similar applications of AI appear closer together. For example, video, image, art, fashion, and Hollywood cluster near each other, while music and voice are slightly farther but still related. Similarly, financial services, tax, and accounting are clustered closer, as do supply chain, retail, and customer experience. Notably, healthcare stands apart, indicating its unique AI applications, while cybersecurity and law enforcement cluster closer together due to their overlapping themes. This visualization demonstrates how AI's influence spans diverse industries while maintaining thematic connections.



Key Themes & Industry Presence:

- AI frequently appears alongside terms like “**use**,” “**help**”, “**generative**,” “**model**,” “**technology**,” and “**transform**”, indicating discussions on its application and industry impact.
- Healthcare Focus:** Frequent occurrence of words like “**health**,” “**healthcare**,” “**cancer**,” “**patient**,” “**medical**,” suggests strong discussions around AI’s role in medical research, disease detection, and patient care.
- Finance & Business:** Terms like “**financial**,” “**market**,” “**service**,” “**company**,” and “**management**” indicate AI’s relevance in financial operations, investment strategies, and business processes.
- Security & Risk Management:** Occurrences such as “**cybersecurity**”, “**security**” reflect AI’s use in **fraud detection, threat analysis, and security enhancement**.
- Creative & Media Applications:** Words like “**video**,” “**music**,” “**image**,” “**art**,” and “**voice**” appear together in discussions about AI’s role in content creation, digital media, and entertainment.
- Sustainability & Global Impact:** “**climate**,” “**global**,” “**research**”, “**impact**,” and “**innovation**” highlights AI’s role in climate applications, environmental impact and global technology advancements.
- “**New**” indicates advancements in AI tools, models, and industry applications. “**Power**” and “**Tool**” highlight AI-powered solutions in healthcare, cybersecurity, and finance. “**Research**”, “**Technology**” and “**Innovation**” emphasize AI’s influence on global progress.

Bigrams: Phrases such as “use AI,” “generative AI,” “AI power,” “use AI tool,” “AI drive,” “AI generate”, “AI help”, “AI make”, “AI transform,” “new AI”, “AI model, “AI improve”, “role AI”, “ai enhance”, “ai assist”, “ai assistant”, “ai solution”, “harness ai”, “ai shape future”, “ai model predict” and “leverage AI” suggest a strong focus on AI’s role in present innovation, transformation and applications of the industries.

Trigrams: Some Thematic Clusters:

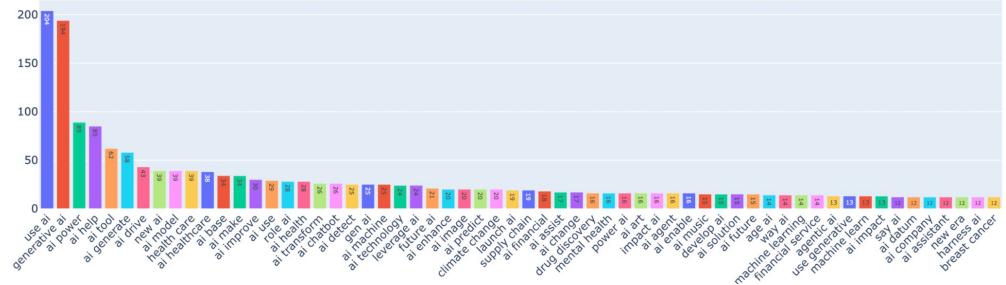
Healthcare, Finance & other Businesses: “AI healthcare”, “AI mental health”, “AI medicine” “breast cancer,” “digital medicine”, “medical imaging AI”, “AI financial service,” “AI financial reporting”, “AI financial Market”, “supply chain management”, “ai supply chain”

Impact, Security & Risk: “threat report AI,” “cyber security”, “ai help prevent”, “use ai detect”, “AI climate change”

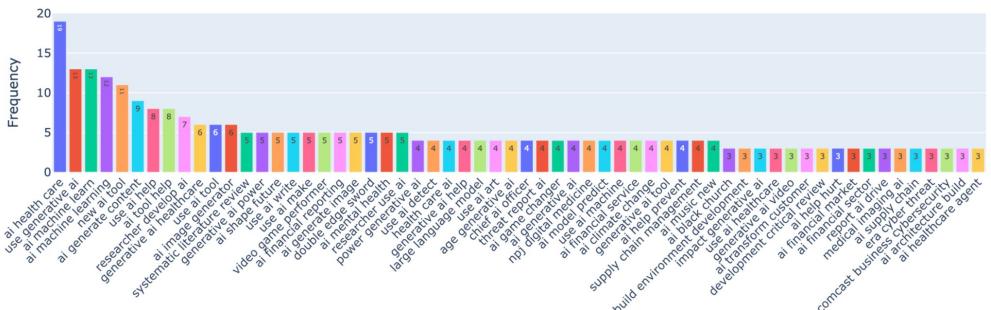
Technology & Creativity: “AI image generator,” “AI music,” “generative AI tool”, “AI generate content”, “use ai write”, “generative ai video”, “use ai art”, “systematic literature review”

The n-gram analysis reveals how AI is widely discussed across industries, with emphasis on transformation, innovation, impact and sector-specific advancements. The recurring phrases confirm AI's expanding role in healthcare, finance, security, and creative fields while also highlighting emerging trends such as generative AI, agentic AI and AI-driven automation. The term "climate" reflects discussions on both AI-driven climate modeling for predicting climate change and the environmental impact of AI models themselves.

Top 2-grams in {AI in various Industries}



Top 3-grams in {AI in various Industries}



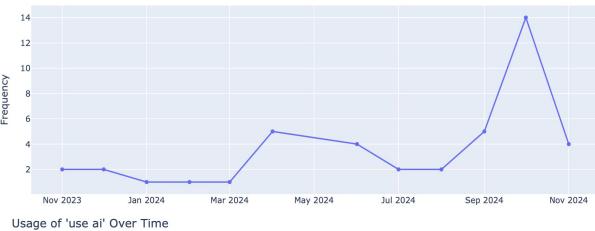
The graphs depict the frequency of few n-grams over time, showing a **notable rise in September and October**. There is also a potential increase in November, though the observed drop may be due to only **one-third of the November headlines being available**.

A closer look at industry trends reveals:

Drug Discovery: The growing occurrence of terms like “**drug discovery**” highlights AI’s expanding role in **medical research and pharmaceutical innovation**.

Overall, these patterns suggest a **rising interest in AI across industries**, reflecting its growing influence and applications.

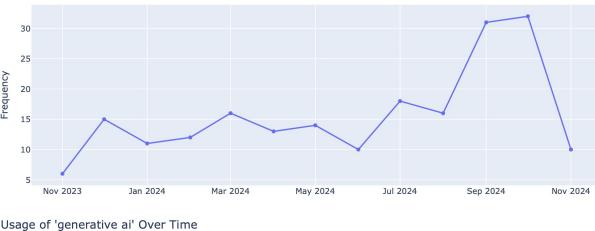
Usage of 'ai drive' Over Time



Usage of 'ai healthcare' Over Time



Usage of 'use ai' Over Time



Usage of 'ai tool' Over Time



Usage of 'generative ai' Over Time



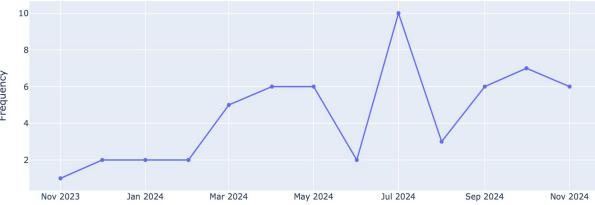
Usage of 'ai help' Over Time



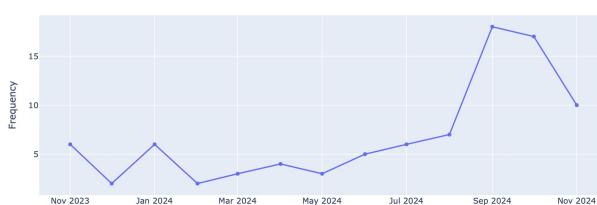
Usage of 'drug discovery' Over Time



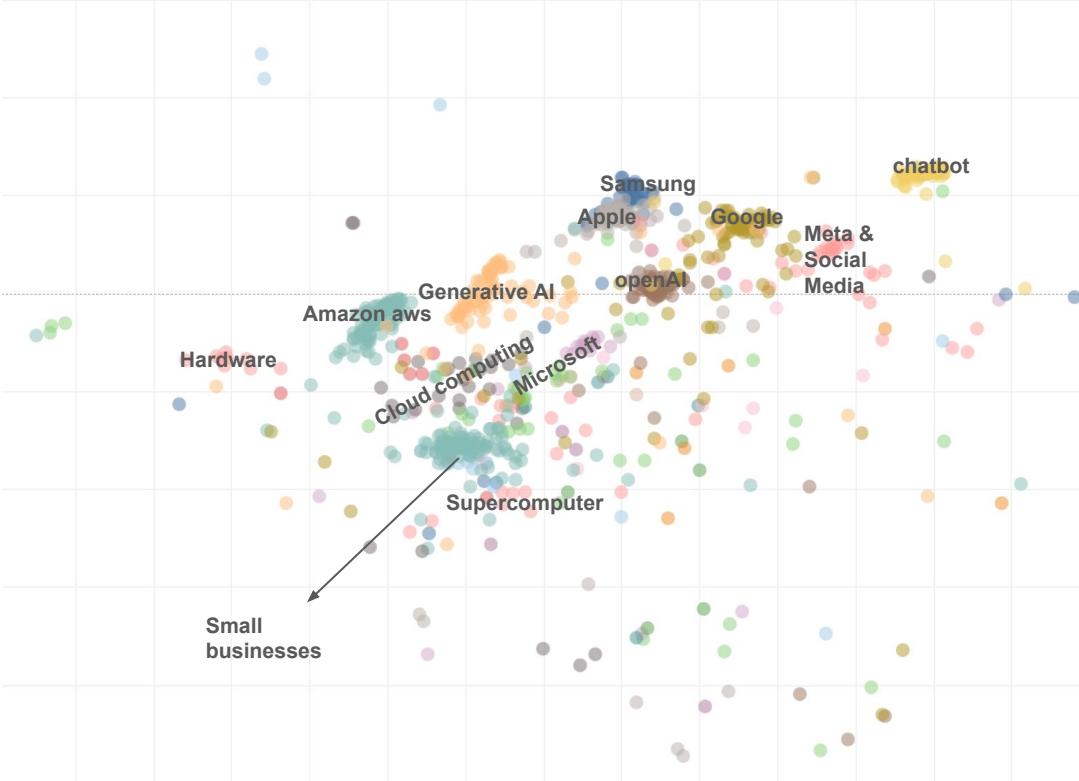
Usage of 'ai generate' Over Time



Usage of 'ai power' Over Time

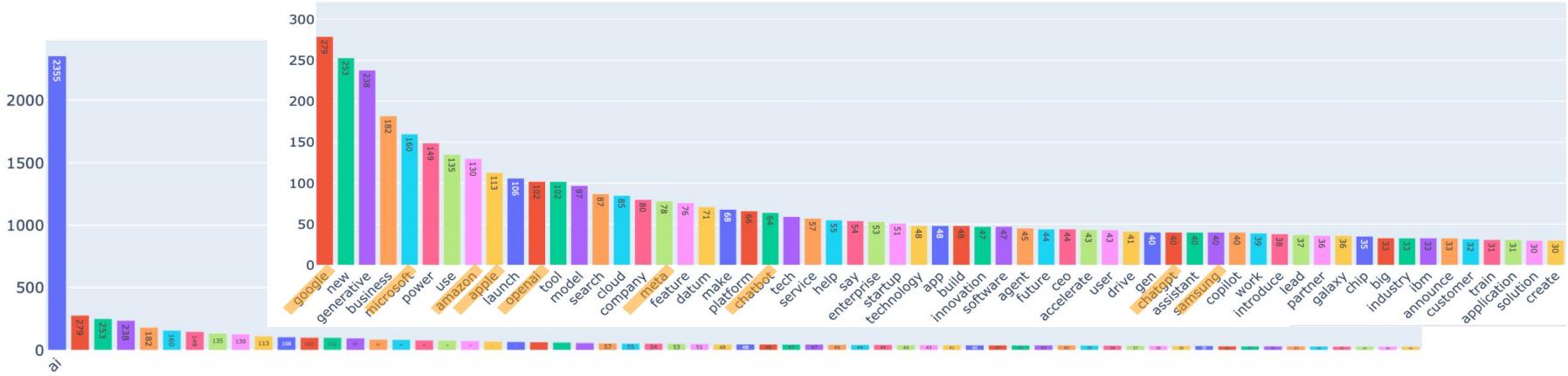


AI in companies and Enterprises



The scatter plot presents 2D projections of AI-related news headlines categorized under **AI and Companies & Enterprises**. By analyzing the data points on the Tableau dashboard, several key topics emerged, including **Samsung, Apple, OpenAI, Google, Chatbots, Meta, Social Media, Amazon, AWS, Microsoft, Small Businesses, Generative AI, and Hardware**, highlighting AI's role in various companies. Additionally, an **n-gram analysis** was conducted to further examine company-specific AI involvement.

Top 1-grams in {AI in companies & Enterprises}



- Major Tech Companies: Frequent mentions of companies like Google, Microsoft, Amazon, Apple, OpenAI, Meta, IBM, and Samsung suggest a strong presence of discussions around their AI advancements, products, or strategies.
- AI Development & Usage: Words like AI, Generative, Model, Chatbot, Copilot, ChatGPT, Assistant, Agent indicate a focus on AI-driven innovations, including generative AI models, virtual assistants, and AI-powered tools.
- Business & Industry Impact: Terms like Business, Company, Enterprise, Startup, Innovation, Industry, Service, Solution, Customer reflect AI's role in transforming businesses and industries.
- Product & Feature Launches: Words such as Launch, Tool, Platform, Build, Software, app, Application, Train, Introduce, Announce suggest ongoing AI-related product releases, updates, and developments.
- Cloud & Computing Power: The presence of Cloud, Power, Drive, Chip implies discussions on AI infrastructure, cloud services, and hardware advancements.
- Future Outlook: Words like Future, Accelerate, Lead, Big, Make indicate that AI's role in shaping future industry trends is a key topic.

Tech Giants Driving AI Innovation: Companies like Google, Microsoft, Amazon, Apple, Meta, IBM, and Samsung are actively developing AI-powered tools, cloud solutions, and business applications, as seen in terms like Google AI, Microsoft AI, Amazon web service, Apple Intelligence, apple AI, Samsung Galaxy AI.

Enterprise AI Adoption: Businesses are integrating AI for automation, customer engagement, and operational efficiency, reflected in terms like AI adoption, Enterprise AI, AI Business, AI Startup, AI Platform, AI Solution, AI Assistant, AI Application.

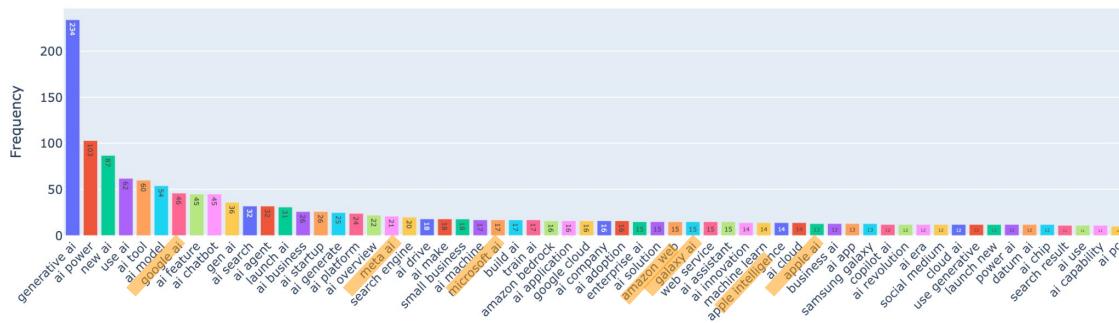
AI in Search and Cloud Computing: AI is transforming search engines and cloud-based services. Terms like AI Search Engine, AI Power Search, Google AI Overview, Generative AI Cloud, AI Cloud, AI Cloud Computing, Amazon Web Service, Google Cloud, AI PC indicate so.

AI-Powered Productivity and Assistants: AI assistants and copilots are becoming integral to business workflows, as highlighted by terms like Chatbot, Copilot AI, AI Assistant, AI Tool, OpenAI, ChatGPT, AI Work, AI Feature, AI Revolution.

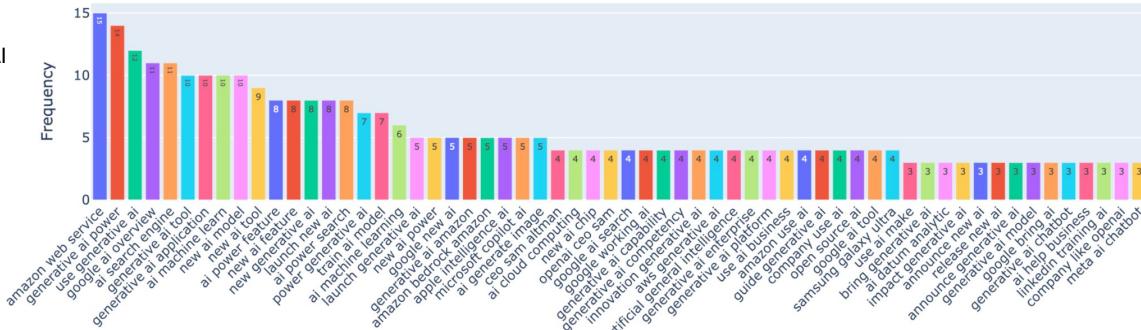
Generative AI's Expanding Influence: AI is shaping multiple industries through generative capabilities, seen in Generative AI Tool, Generative AI Application, Generative AI Platform, AI Generate Image, AI Power Feature, Innovation Generative AI, AI Machine Learning.

Corporate AI Announcements and Investments: Companies are frequently launching new AI features, partnerships, and training initiatives, as indicated by Launch New AI, Announce Generative AI, LinkedIn Training AI, AI Help Business, AI Adoption.

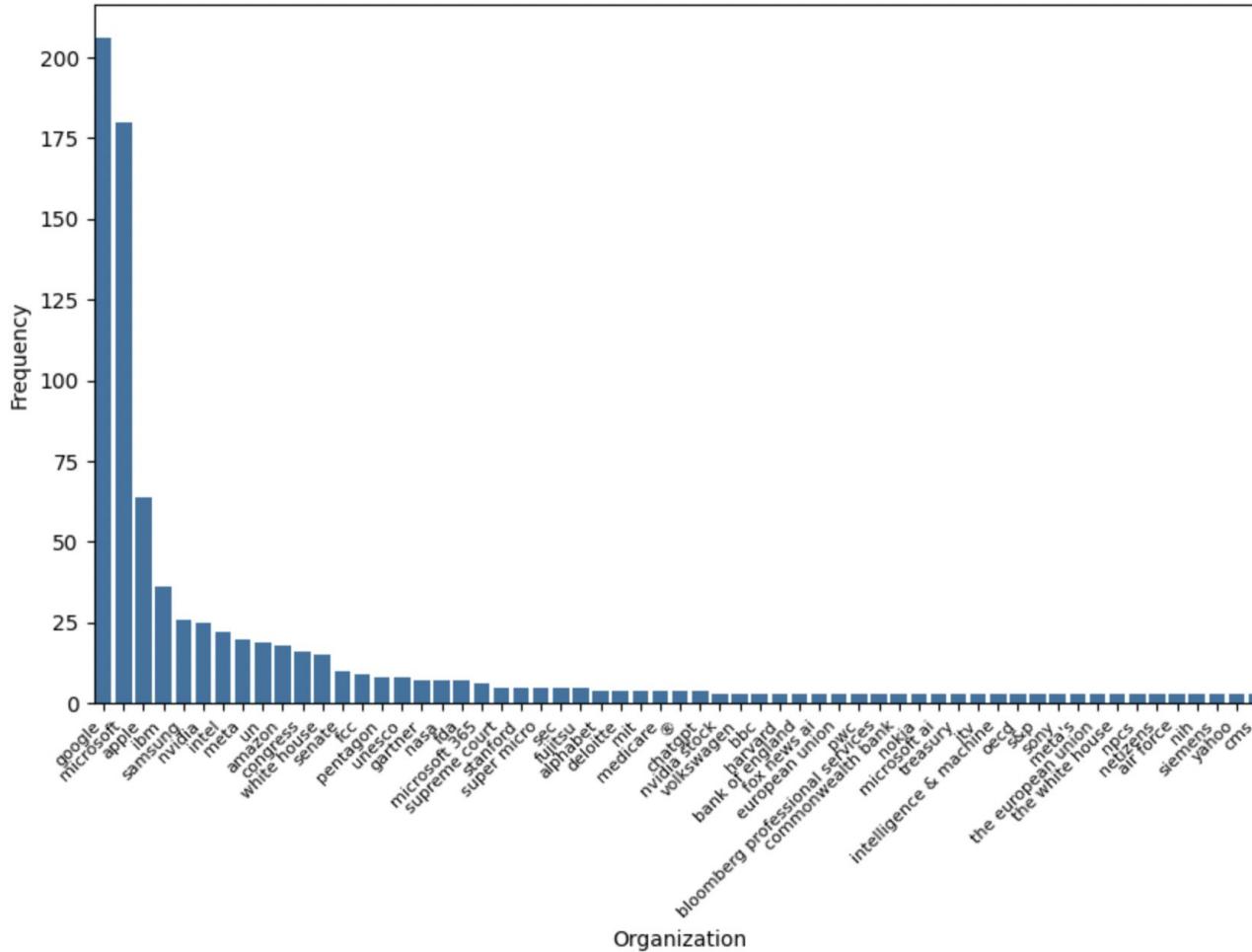
Top 2-grams in {AI in companies & Enterprises}



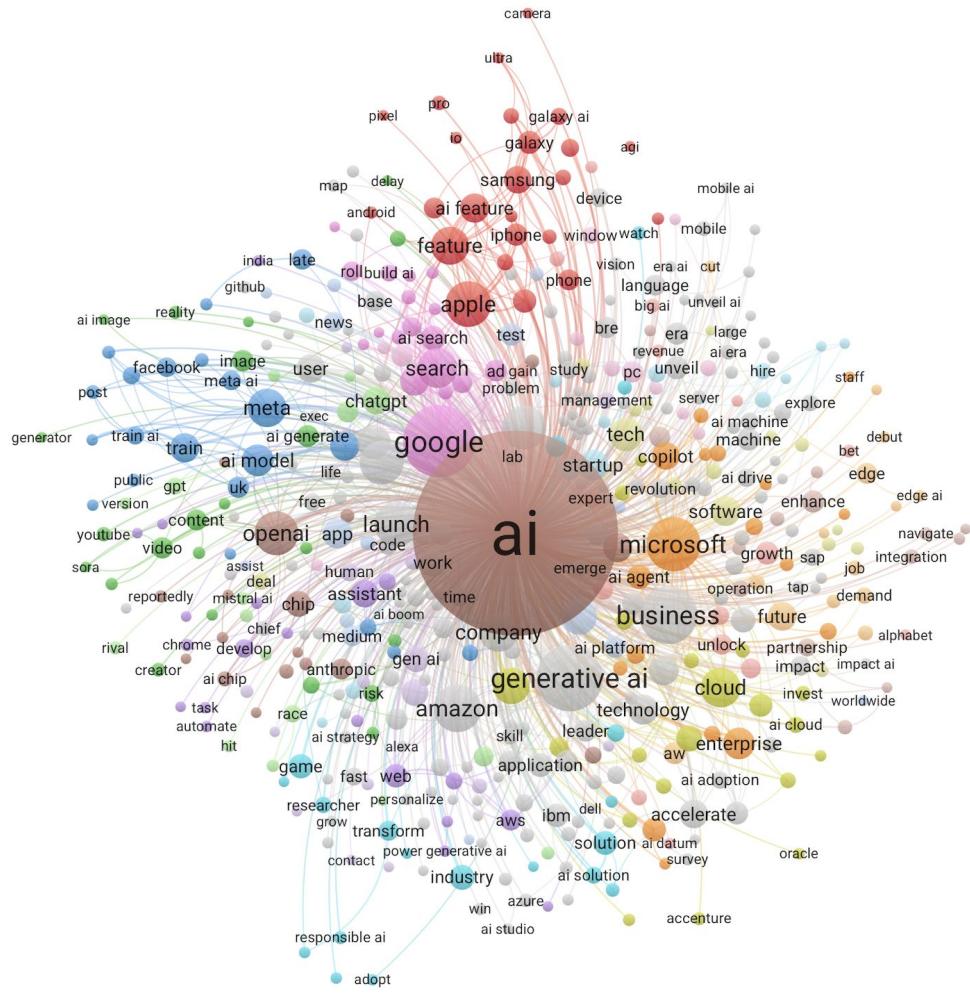
Top 3-grams in {AI in companies & Enterprises}



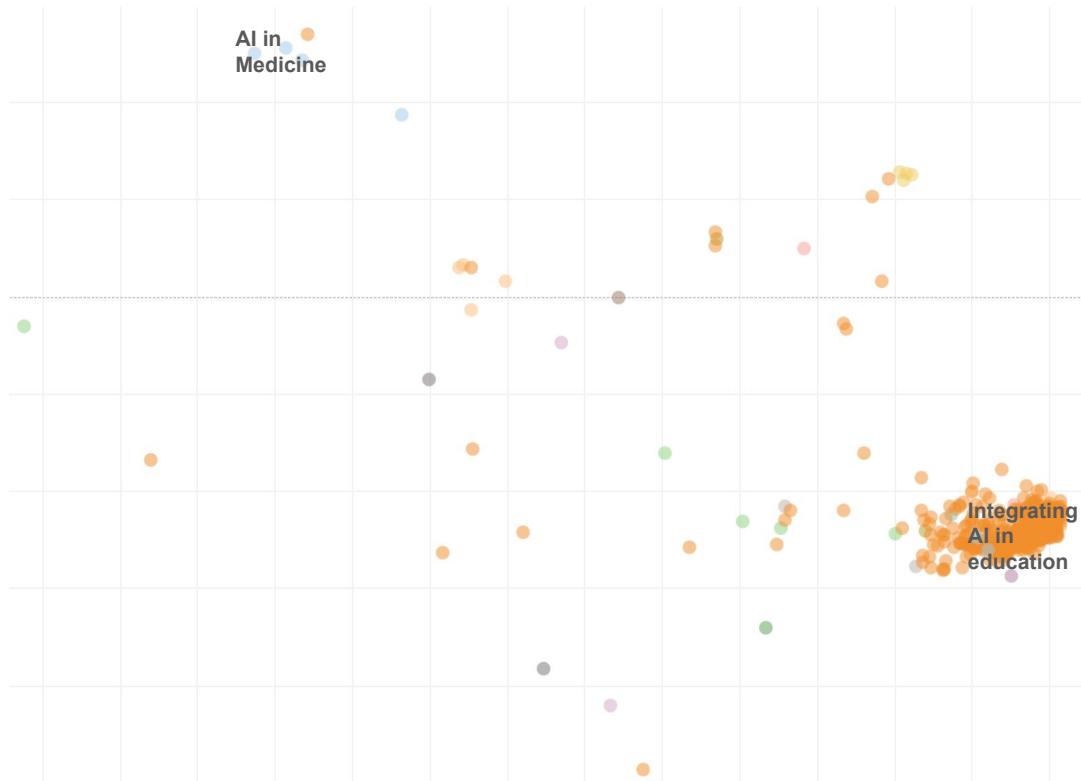
Frequency of Organizations in Text Data



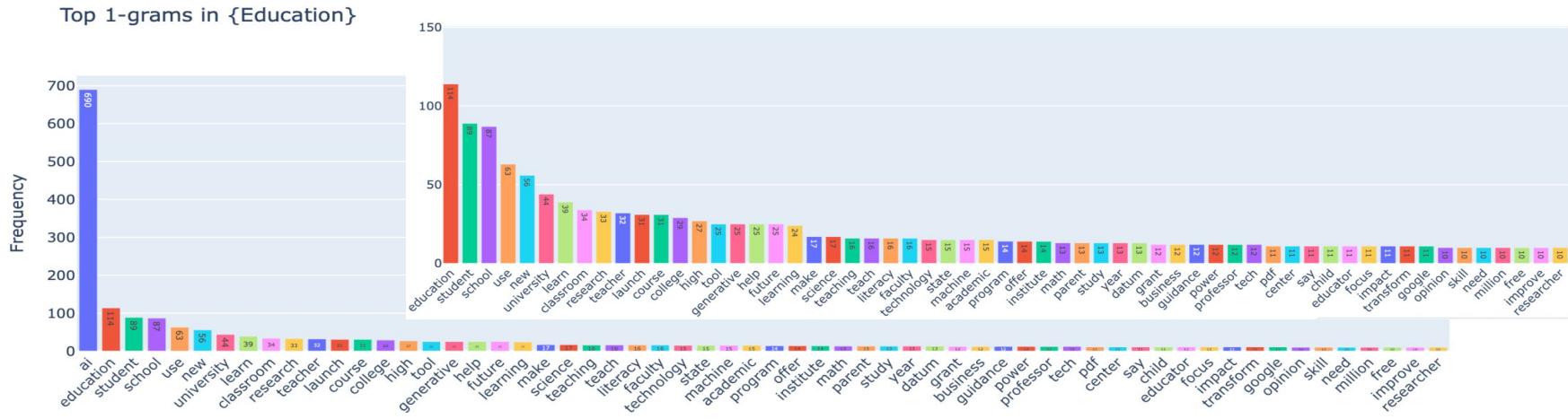
Got this plot by
performing Named Entity
Recognition



Education



The news headlines related to AI in education cover a wide range of key topics, with a major focus on the integration of AI in education, emphasizing both the opportunities and challenges it presents, alongside innovations in teaching and learning. Other prominent themes include the use of AI in medical education, funding and grants for AI research, the role of AI chatbots, and the importance of AI in language preservation.



Some key observations from the unigram analysis

- Words like education, school, university, teacher, and student indicate AI's relevance in both primary-secondary and higher education.
- Terms like learning, teaching, and skill suggest both the need to develop AI-related skills and AI's role in enhancing education and skill improvement.
- Technological Advancements: The presence of AI, generative, tool, technology, machine, science, program, and tech points to AI-driven innovations, particularly in tools and platforms for learning
- Funding and Research Initiatives: Unigrams like research, grant, researcher, study, center, and million highlight discussions on AI funding, academic research, and investments in AI-driven educational initiatives.
- Business and Government Involvement: Words like state, business, power, and policy suggest that government policies, educational institutions, and businesses are involved in shaping AI's role in education.
- Parental and Societal Perspectives: The presence of parent, child, opinion, and impact indicates discussions on how AI is influencing students, parental concerns, and broader societal implications.

The bigram and trigram analysis of AI-related news headlines in education reveals several key themes:

1. AI Integration in Education: Phrases like AI education, AI classroom, AI school, AI course, AI teaching, AI college, AI institute, AI academy, and AI revolutionize education indicate that AI is being actively integrated into various educational settings, from schools to higher education institutions.

2. AI's Role in Learning and Teaching: Bigrams such as AI tool, AI literacy, AI learn, learning AI, college admissions, AI guidance, AI enhance, AI improve, AI help, use AI classroom, AI save teacher and AI potential highlight how AI is being used to enhance learning, improve literacy, and provide academic guidance.

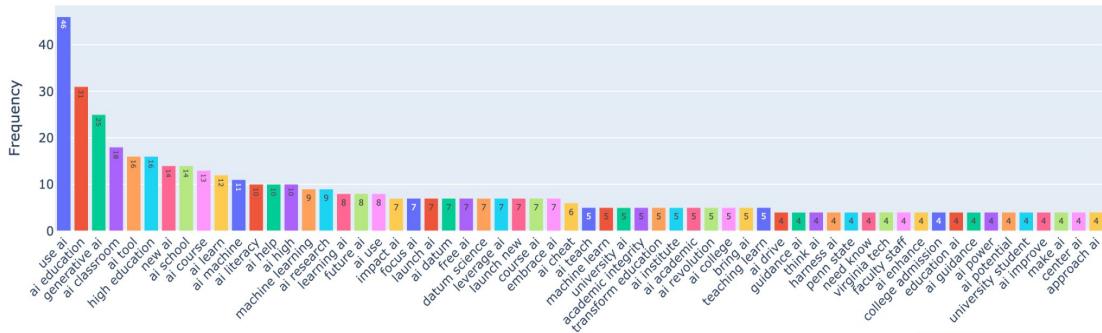
3. Generative AI in Education: The presence of generative AI, generative AI course, generative AI education suggests that AI-generated content and tools are becoming a significant part of modern education.

3. AI and Academic Integrity: The phrases AI cheat, academic integrity, use AI cheat indicate concerns about AI's impact on ethical learning practices, particularly in assessments and coursework.

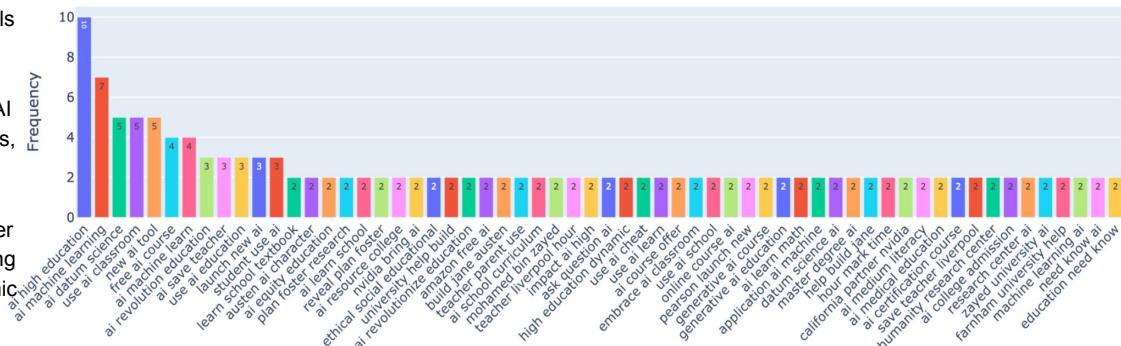
4. AI in Research and Innovation: Terms such as AI research, research center, AI, plan foster research, and university help build suggest that AI is being explored not only as a teaching tool but also as a subject of academic research and innovation.

5. Corporate and Institutional Involvement: Mentions of Amazon free AI, launch new ai, Pearson launch new, free ai course, AI school curriculum and Zaire University AI suggest that both tech companies and academic institutions are playing a role in driving AI adoption in education.

Top 2-grams in {Education}



Top 3-grams in {Education}

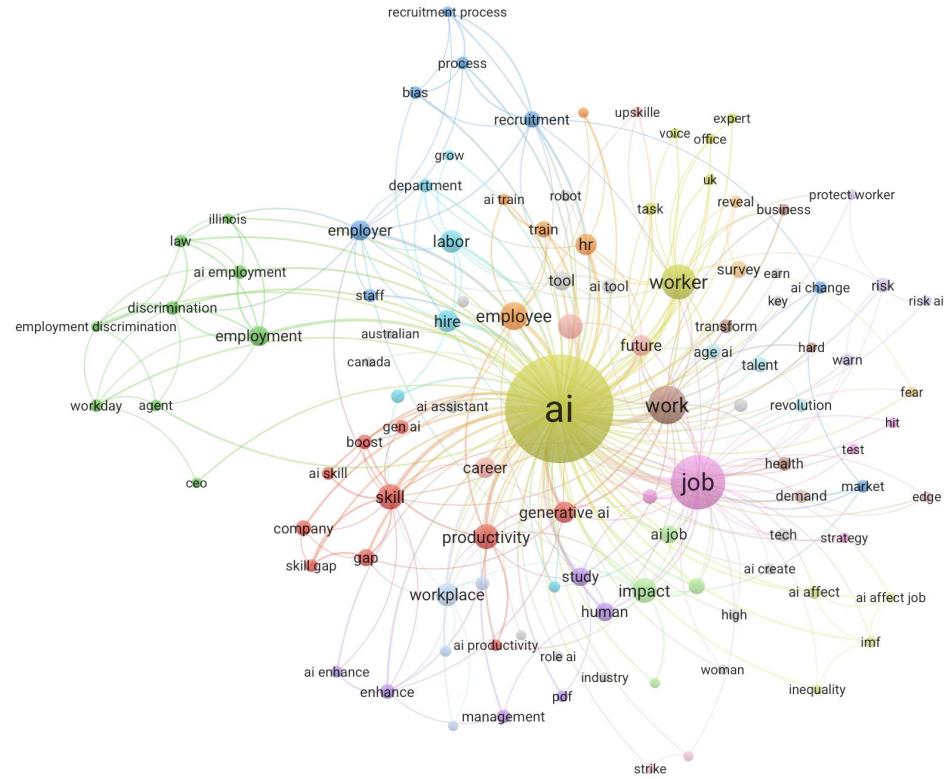


Overall, these n-grams highlight AI's growing influence in education, spanning classroom integration, research, ethical concerns, specialized training, and institutional involvement.

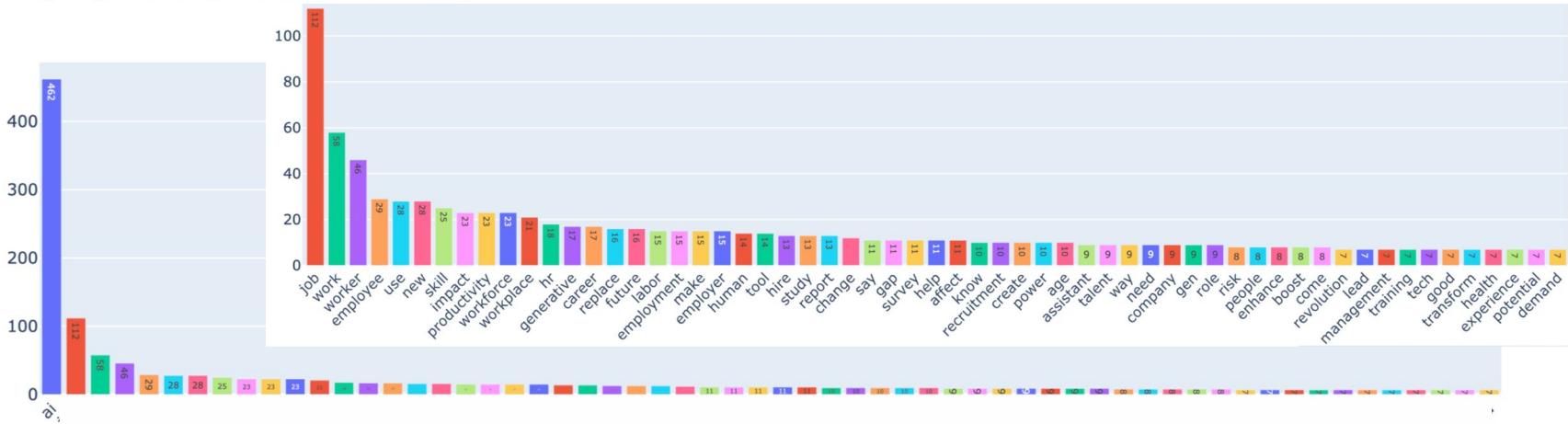
few ngrams frequencies over time



Career & Workforce



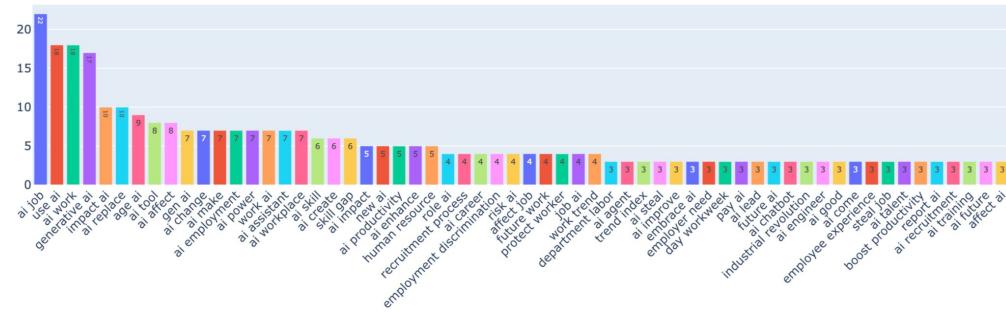
Top 1-grams in {Careers & Workforce}



Key Insights from AI and Career/Workforce-Related Headlines

1. **AI's Impact on Jobs and Employment** – *job, work, worker, employee, employment, employer, workforce, workplace, labor, recruitment* – AI is transforming job markets, affecting employment dynamics, and reshaping workforce structures.
2. **Skills and Career Development** – *skill, career, training, talent, education, gap* – Upskilling and reskilling are crucial for adapting to AI-driven job market changes.
3. **AI in Workplace Productivity and HR** – *productivity, HR, management, enhance, boost, assistant, tool, transform* – AI is improving efficiency, optimizing HR processes, and redefining workforce management.
4. **Concerns About Job Replacement and Risks** – *replace, risk, effect, change, lead, revolution, future* – Discussions focus on AI's potential to replace jobs, reshape industries, and create workforce uncertainties.
5. **Industry Adoption and Trends** – *company, report, study, demand, experience, tech, good, potential* – Industry reports highlight AI's influence on careers, workforce trends, and business strategies.

Top 2-grams in {Careers & Workforce}



Based on the bigrams and trigrams related to AI in Careers & Workforce, here are some key insights:

1. AI's Impact on Jobs and Employment

- AI job, AI replace, AI steal job, AI effect job worker, AI employment decision AI's role in employment is a major concern, particularly regarding job displacement. There are fears that AI could replace human workers, especially in repetitive or automatable roles, leading to employment uncertainty.
- AI talent, AI recruitment, AI career, AI engineer: While AI may eliminate some jobs, it is also creating new career opportunities, particularly in AI-related fields like engineering and data science.

2. AI's Role in Workplaces and Productivity

- AI workplace, AI assistant, AI tool, AI enhance, AI productivity gain, AI boost productivity: AI is being integrated into workplaces to improve efficiency, automate tasks, and enhance productivity. Many companies are adopting AI tools to streamline workflows and reduce manual effort.
- AI skill gap, AI training, AI certification edge, AI enhanced skill: The rise of AI is increasing the demand for new skills, leading to discussions on training and upskilling the workforce. Workers who adapt to AI-driven changes may have a competitive advantage.

3. Ethical & Legal Issues in AI-Driven Workplaces

- Employment discrimination, AI employee discrimination, employment discrimination law, court hold AI: AI's use in hiring and employment decisions raises concerns about bias and discrimination. Legal and ethical debates focus on ensuring fairness in AI-driven recruitment and workplace decision-making.
- Human resources, recruitment process, AI recruitment: AI is transforming human resources and hiring practices, with automated recruitment tools helping employers screen candidates more efficiently. However, there is concern over transparency and fairness in AI-driven hiring decisions.

4. The Future of Work & Changing Workforce Trends

- Future work, work trend index, AI work come, work age AI: The concept of “work” is evolving with AI, influencing trends such as hybrid work, automation, and digital transformation. Companies are reassessing how AI fits into their long-term workforce planning.
 - Protect worker, step protect worker, AI pay AI, AI effect freelance: As AI disrupts job markets, discussions on worker protections, fair pay, and AI’s impact on freelancing and gig economy jobs are becoming more prominent.

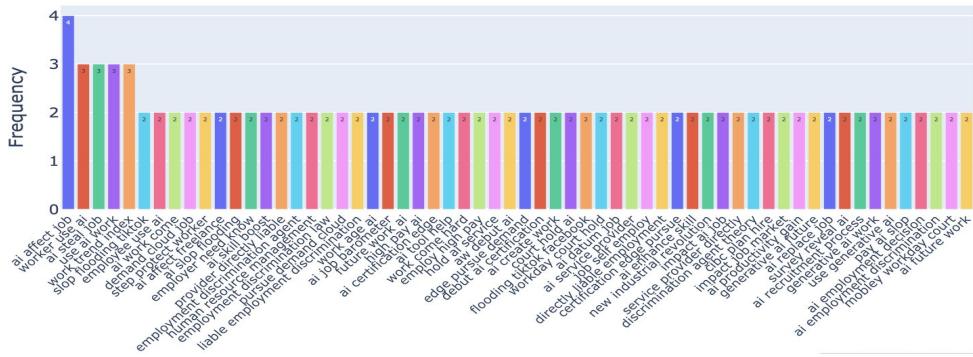
5. AI and Industry-Specific Changes

- Industrial revolution, new industrial revolution, AI impact job: AI is driving a modern industrial revolution, reshaping industries by automating tasks, optimizing processes, and changing job structures.
 - AI service provider, AI datum job, AI cloud job, demand cloud job: AI is increasing demand for jobs in cloud computing, data science, and AI service management, highlighting new career paths in AI-driven industries.

Conclusion

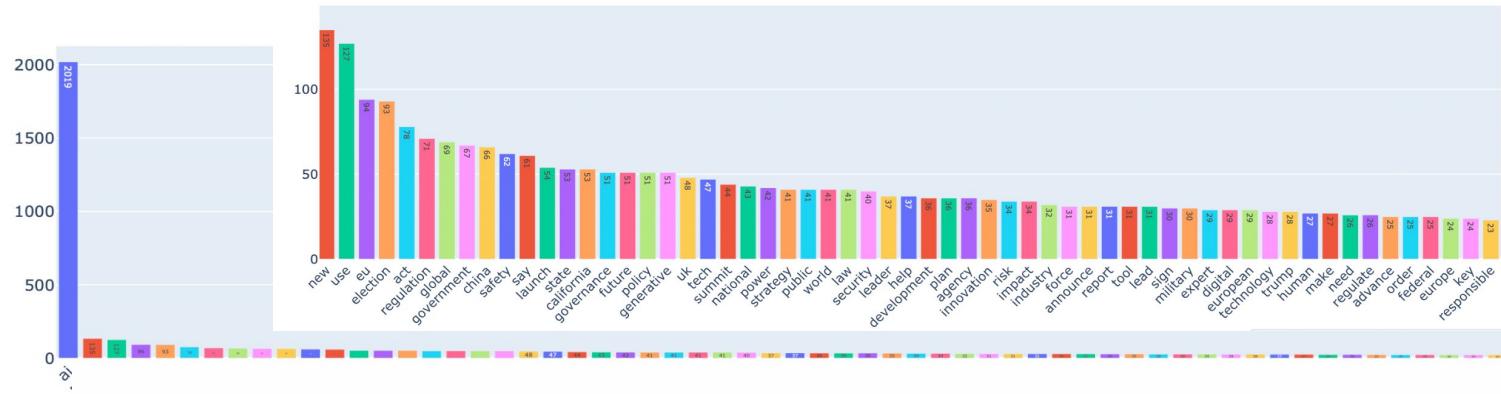
AI is reshaping careers and the workforce by automating jobs, boosting productivity, and creating new opportunities. However, concerns around job displacement, ethical hiring, and skill gaps remain. The future of work will likely require continuous adaptation, with AI playing a central role in defining job roles and workplace dynamics.

Top 3-grams in {Careers & Workforce}



AI Governance & Geopolitics

Top 1-grams in {AI Governance & Geopolitics}



- Regulation and Policy Development – Terms like Regulation, new, Act, Law, Policy, Governance, and Federal indicate active efforts by governments and agencies to establish AI-related policies.
- Global and Regional AI Strategies – Mentions of EU, UK, European, China, California, National, Public, and World suggest that AI governance is a global concern, with different regions adopting their own strategies to regulate and advance AI technologies.
- AI in Political and Governmental Decision-Making – Words like Election, Government, Summit, Strategy, Power imply the impact of AI in political landscapes, global summits
- AI Safety, Security, and Risks – Terms such as Safety, Security, Military, Risk, Digital, Human, and Responsible highlight concerns over AI's potential risks, including cybersecurity threats, military applications, and ethical implications
- Innovation and Industry Impact – Mentions of Technology, Development, Plan, Innovation, Industry, Advance, and Tool suggest that while regulation is a focus, there is also an emphasis on fostering AI innovation within industries under structured governance.

These insights suggest that AI governance is a rapidly evolving and balancing the need for innovation with regulatory measures to ensure ethical, secure, and globally coordinated AI development.

The bigrams and trigrams from AI governance and geopolitics headlines reveal several key themes:

1. Government Regulation & Security:

Many phrases, such as *White House*, *Executive Order*, *National AI, Federal Agency*, *European Union*, and *California Governor Veto*, highlight government involvement in AI regulation. Terms like *AI Legislation*, *AI Act Publish*, and *AI Executive Order* suggest active policy-making, while *National Security*, *AI Task Force*, and *AI National Security* reflect concerns about AI's role in defense.

2. Global AI Governance & Geopolitics:

Phrases like *Global AI Governance*, *European Union AI* and *Global Regulatory Tracker* indicate a global push for AI regulation, with various countries shaping their own policies. The presence of *Israel Use AI*, *eu ai act* and *Union AI Act* underscores the geopolitical complexities of AI governance.

3. Ethics & Responsible AI:

Recurring terms such as *Responsible use AI*, *AI Ethics*, *Safe Responsible AI*, and *AI Safety Institute* emphasize the importance of ethical AI development.

Concerns about bias are evident in phrases like *Perpetuate Historic Bias* and *Historic Bias Emerge*, reflecting ongoing debates on fairness and accountability in AI systems.

4. AI Innovation & Technological Advancements:

Phrases like *AI Innovation*, *AI Model*, *AI Development*, *AI Research*, and *AI Chip* point to continuous advancements in AI technologies, particularly in machine learning and generative models. The mention of *AI Machine Learn* and *Generative AI* highlights key areas of progress.

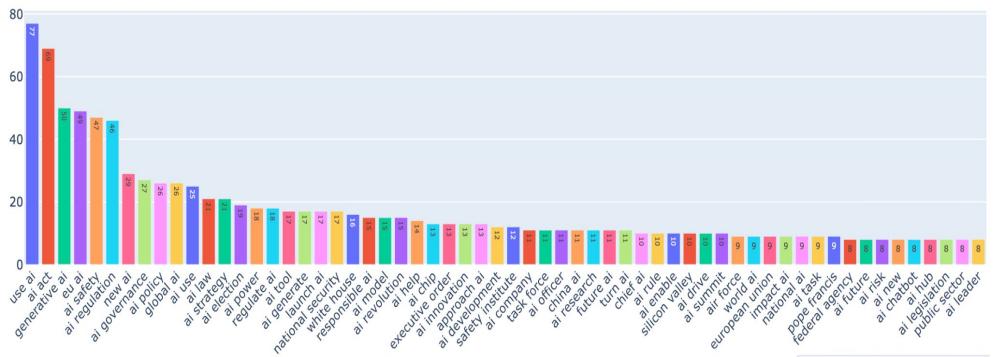
5. AI's Societal Impact & Future Preparedness:

Terms like *AI Future*, *AI Game Changer*, and *AI Preparedness Index* highlight AI's transformative potential and the need for strategic planning, while *AI Revolution* and *AI Root* reflect both optimism and caution.

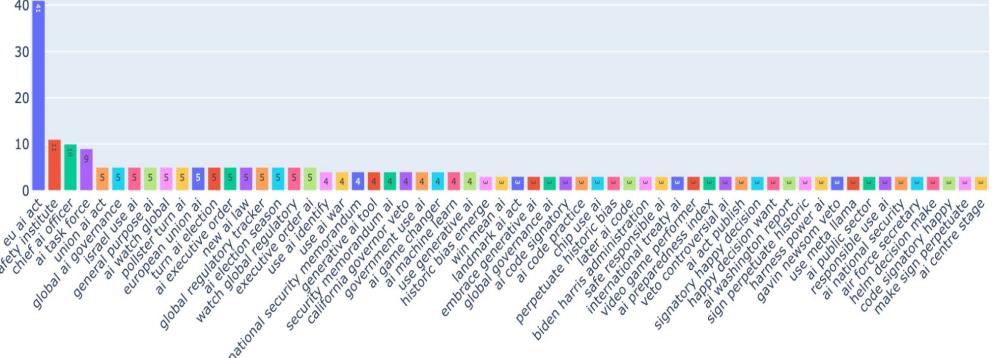
6. AI in Politics & Elections:

Phrases like *AI Election Season* and *Use AI War* indicate AI's role in elections and security. Mentions of *National Security Memorandum* and *Generative AI Tool* suggest its use in sensitive areas.

Top 2-grams in {AI Governance & Geopolitics}

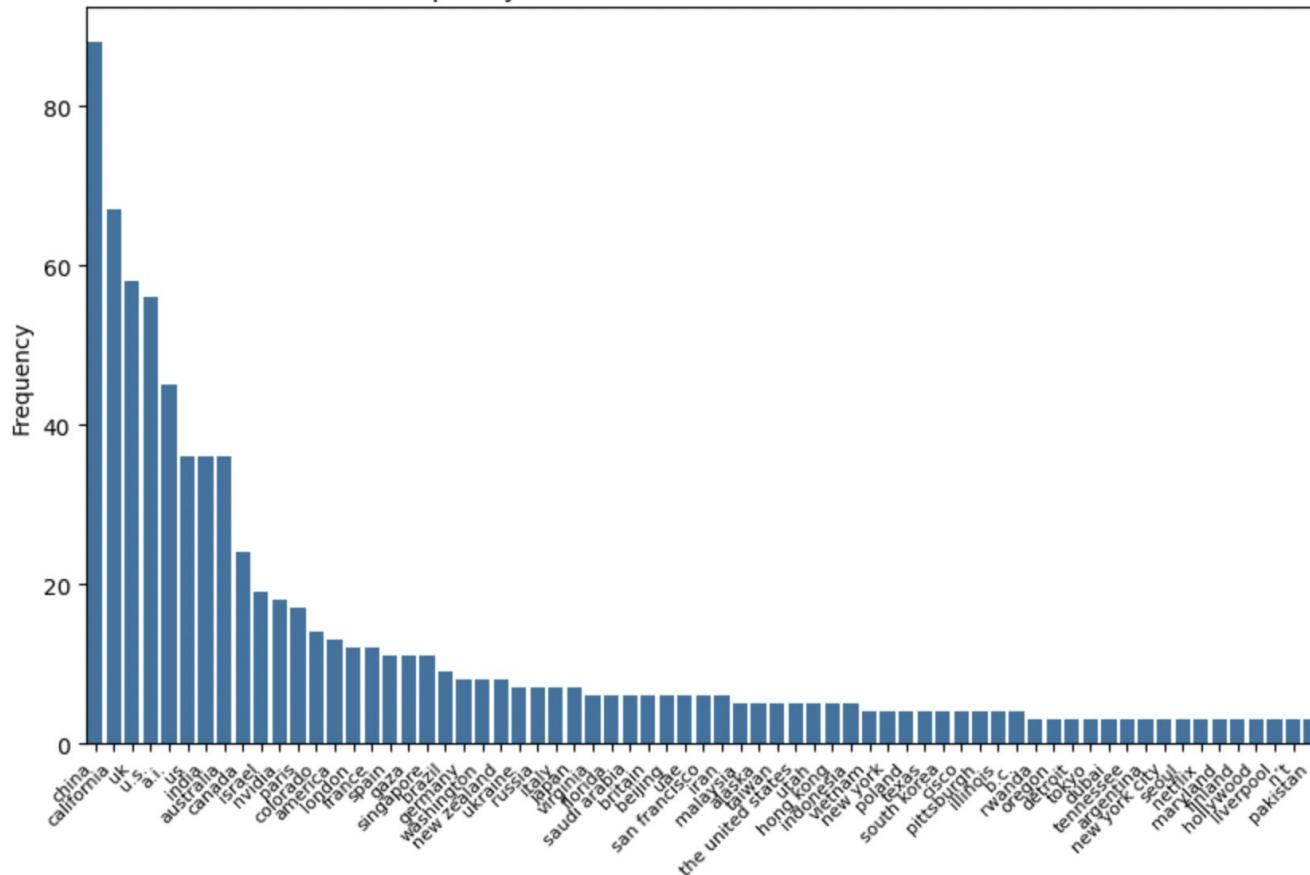


Top 3-grams in {AI Governance & Geopolitics}



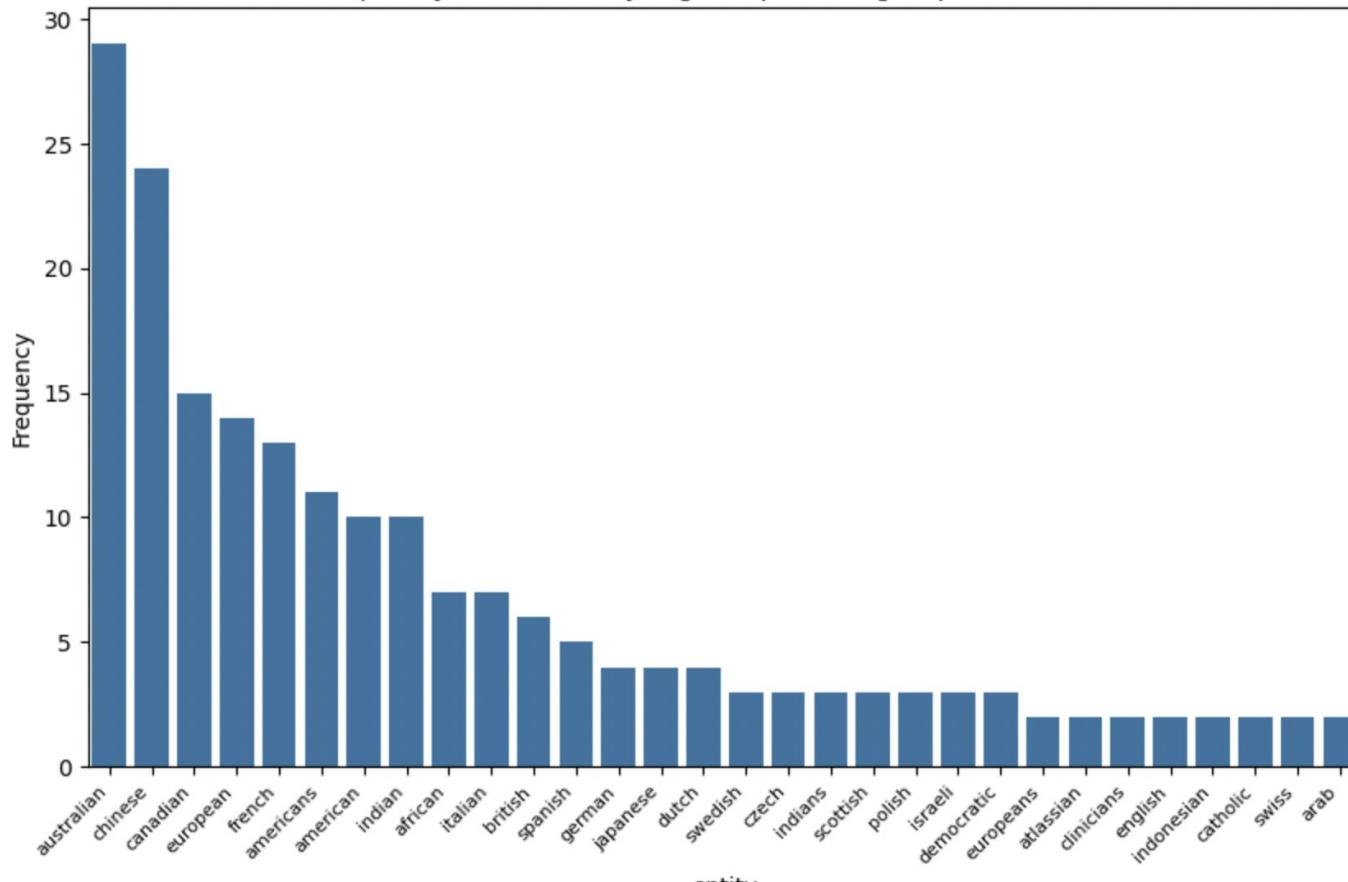
Overall, these n-grams highlight AI's growing influence in governance, ethics, security, innovation, and global cooperation.

Frequency of Countries/cities/states in Text Data



Got this plot by
performing Named Entity
Recognition

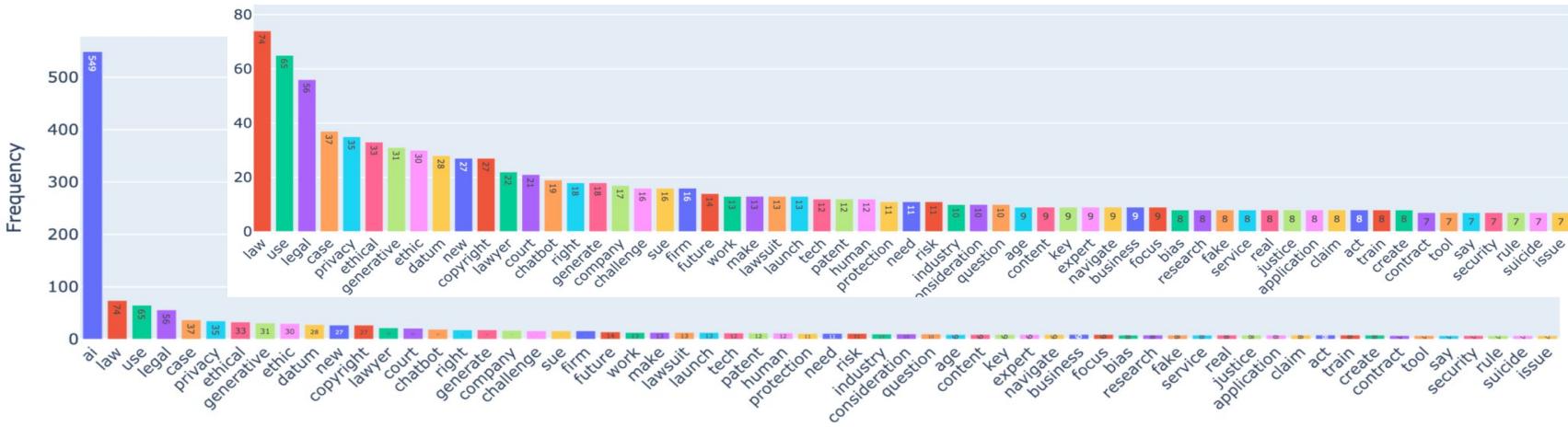
Frequency of Nationality/regions/political groups in Text Data



Got this plot by
performing Named Entity
Recognition

AI ethics, Law & policy

Top 1-grams in {AI Ethics, Law & Policy}



- Legal & Regulatory Frameworks:** Mentions of *AI law*, *lawsuit*, *patent*, *copyright*, and *rule* indicate discussions on evolving AI legislation, intellectual property disputes, and legal challenges surrounding AI technologies.
- Ethical Considerations:** Terms like *bias*, *risk*, *ethical challenge*, and *human* highlight concerns about AI fairness, decision-making transparency, and potential harm to vulnerable populations.
- Privacy & Security:** Words such as *privacy*, *protection*, *security*, and *data* suggest a focus on AI's impact on personal data, digital identity, and cybersecurity.
- Generative AI & Misinformation:** Terms like *generate*, *chatbot*, *fake*, and *content* suggest concerns over AI-generated media, deep fakes, and misinformation risks.
- Justice & Accountability:** Unigrams such as *justice*, *sue*, *lawsuit*, and *challenge* point to legal disputes over AI transparency, and ethical responsibility.
- Social Impact & Bias:** Words like *bias*, *age*, *human*, and *question* indicate discussions on AI's societal impact, including concerns about discrimination and inclusivity.

Legal and Regulatory Landscape

- **N-grams:** AI law, AI act, data protection law, legal AI, EU AI act, AI copyright lawsuit, AI copyright infringement, AI law firm, AI legal, AI intellectual property, AI patent
 - **Indicates:** Development of laws and regulations for AI systems, focusing on **copyright, data protection, intellectual property, and patents**. The **EU AI Act** is a key regulatory development.

Ethical Considerations and Business Applications

- **N-grams:** Ethical AI, ethical consideration, AI ethical business, ethical AI, AI ethics, generative AI, generative AI attempt, AI generate content, AI generate image, AI ethical business, AI tool, AI service provider
 - **Indicates:** The growing importance of ethical development and deployment of AI technologies. This includes ensuring AI systems are fair, transparent, accountable and addressing ethical issues in generative AI applications like content creation and deep fakes.

Privacy, Data Protection, and Security

- **N-grams:** *Datum privacy, data privacy security, AI privacy, AI data, data protection, data protection law, AI privacy, AI use case industry, AI train model*
 - **Indicates:** Focus on **privacy and data protection**, with legal frameworks to ensure **security** and **compliance** in AI systems, especially concerning **personal data** processing and usage.

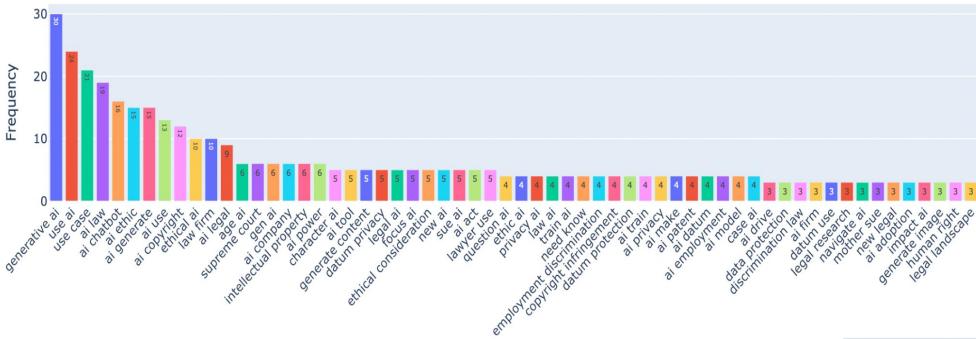
Litigation and Legal Cases Involving AI

- **N-grams:** *AI chatbot lawsuit, AI copyright lawsuit, AI employment discrimination, AI copyright infringement, AI firm, mother sue, AI lawsuit*
 - **Indicates:** The emergence of **legal challenges** and **litigation** concerning **AI-driven discrimination, copyright issues, and chatbot misuse**, reflecting increasing legal scrutiny over AI's impact.

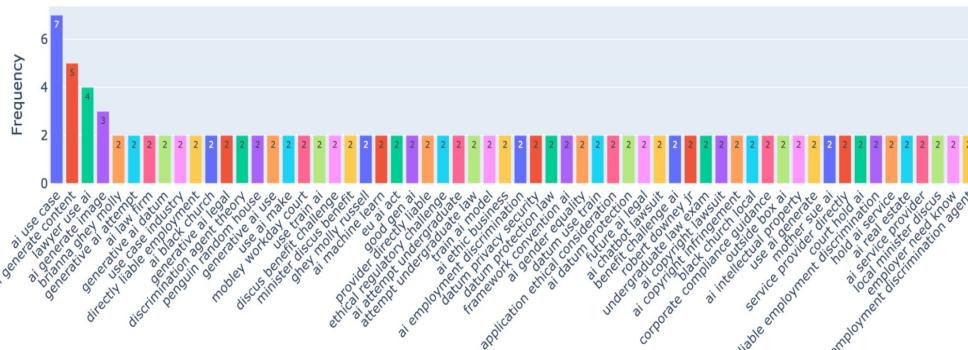
AI in Employment and Industry

- **N-grams:** AI employment, employment discrimination, AI model, AI drive, directly liable employment, AI gender equality, use case industry, AI use case
 - **Indicates:** Concerns over AI's impact on the workforce, focusing on discrimination and bias in employment-related decisions. Discussions on AI applications across industries and workplace accountability are also highlighted.

Top 2-grams in {AI Ethics, Law & Policy}

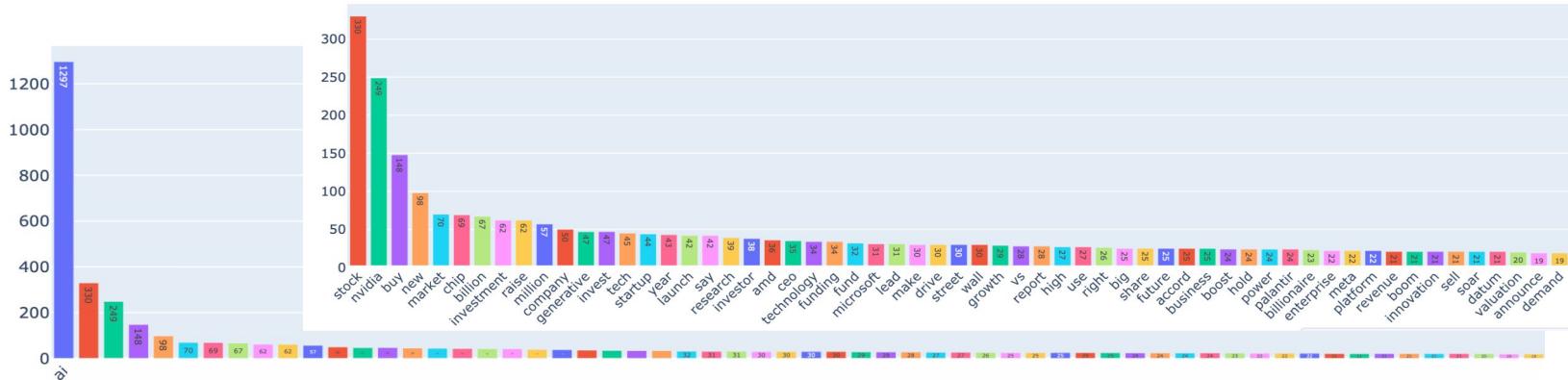


Top 3-grams in {AI Ethics, Law & Policy}



AI Investments & Market Trends

Top 1-grams in {AI Investments & Market Trends}



- **NVIDIA, AMD, Microsoft, Meta:** These are major players in the AI space. Their presence suggests that they are central to the investment conversations around AI, likely due to their leadership in AI hardware (e.g., GPUs), software, and AI-related technologies.
- **Company, Startup, Tech, Platform:** These terms point to the growth of AI startups and established tech companies focusing on AI. There is an emphasis on innovation and the commercialization of AI technology.
- **Investment Activity:** Words like **Investment, Raise, Funding, Billion, and Growth** point to significant capital flowing into AI ventures, with **Valuation and Fund** indicating high investments in AI startups.
- **Market Sentiment and Dynamics**
 - Boom, Soar, High, Demand: These words indicate a strong positive sentiment around AI investments, suggesting that the market for AI-related technologies is booming, with demand for AI products and services at an all-time high.
 - Stock, Share, Report, Revenue: These terms point to the financial performance of AI companies and how their stock prices and market shares are affected by investor sentiment. There is a focus on the financial growth and potential returns of investing in AI.
- **Technologies & Innovations:** **Generative, Technology, Chip, Innovations** reflect the focus on cutting-edge AI technologies like generative AI and hardware (e.g., AI chips) attracting investment.

Investment Activity & Trends

- **AI investment, Invest AI, Funding round, Raise million –** AI companies secure major funding rounds, attracting substantial capital.
 - **Buy AI, Buy whole, AI boom, AI growth –** Rising investment in generative AI indicates strong market optimism.

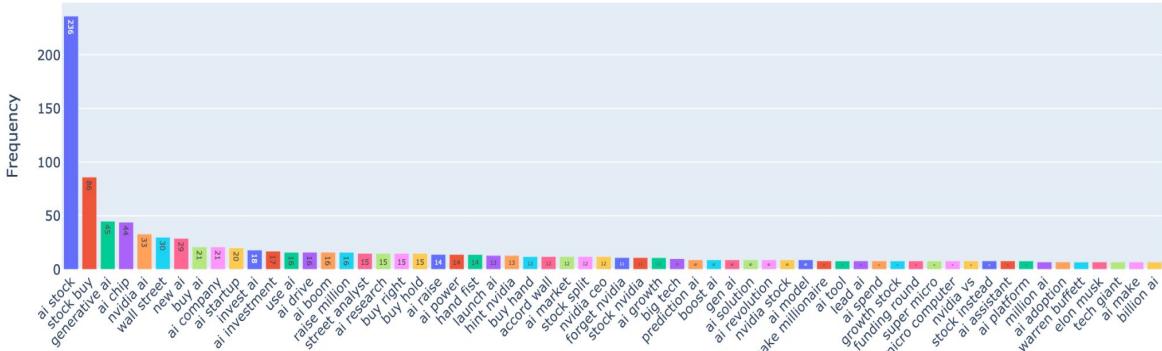
Technological Innovations & Applications

- **Generative AI, AI chip, AI model, AI tool** – Advances in AI models and chips fuel industry growth.
 - **AI platform, AI assistant, AI solution** – AI applications gain traction, attracting strategic investments.
 - **AI semiconductor stock, Super microcomputer, AI chip NVIDIA** – AI chip innovations, particularly NVIDIA's, shape market direction.

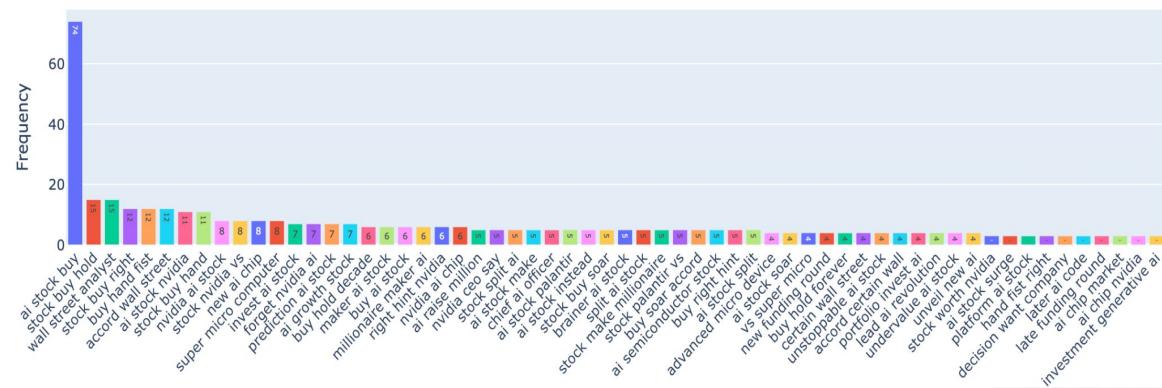
Market Sentiment & Predictions

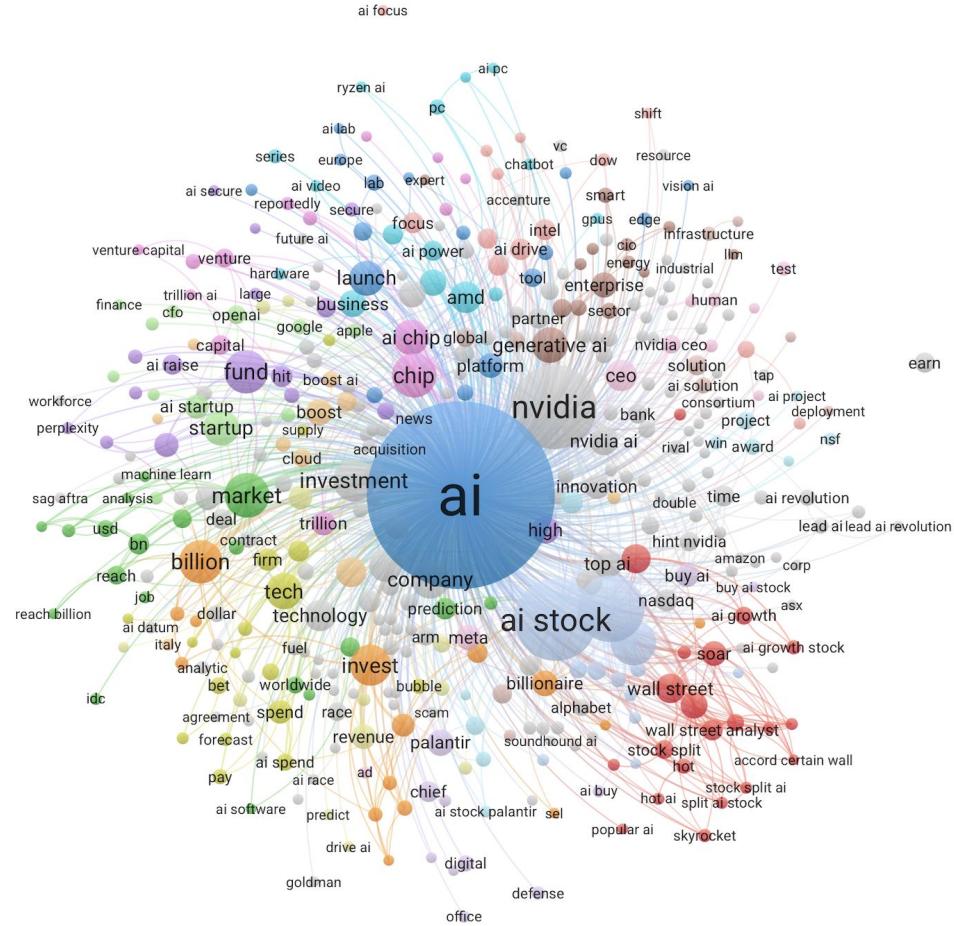
- **AI growth stock, Prediction AI, AI revolution** – AI is seen as a high-growth sector with transformative potential.
 - **AI raise million, AI stock soar, AI adoption** – AI adoption drives stock surges and rising valuations.

Top 2-grams in {AI Investments & Market Trends}



Top 3-grams in {AI Investments & Market Trends}

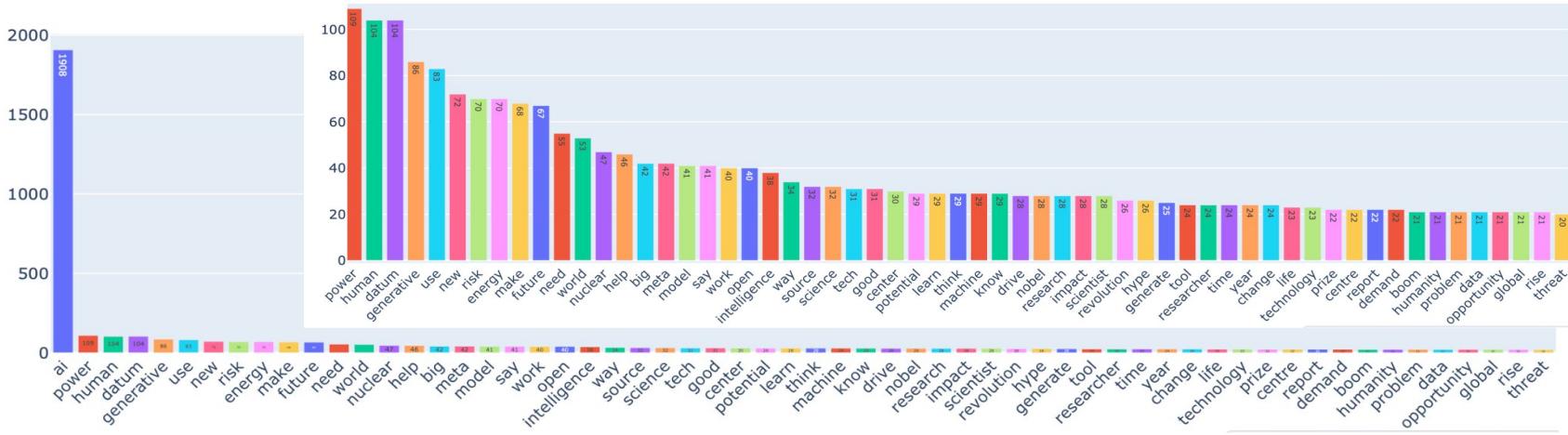




goldman sach

AI overview

Top 1-grams in {AI overview, risks & impact}



- **AI power, AI potential, Generative AI, Artificial General Intelligence, AI model, AI revolution** – AI's transformative potential spans industries, with AGI and generative AI at the forefront.
- **AI help, AI tool, AI change, AI make** – AI enhances human capabilities, driving innovation and problem-solving.
- **AI risk, AI black box, AI replace human, AI threat, AI risk management** – Concerns over AI replacing humans, opacity, and existential risks necessitate responsible management.

AI's Societal Impact

- **AI change world, AI impact, AI revolution, AI make life, AI future** – AI is reshaping society, work, and human interaction.
 - **AI drive change, AI boom, AI industry** – The AI boom is driving industrial transformation and tech innovations

Energy and Resource Demands

- **AI energy, power-hungry AI, nuclear power AI, energy drive** – AI's high computational needs raise sustainability concerns.
 - **AI data center, compute drive energy** – Data centers require vast energy resources, driving discussions on efficiency.

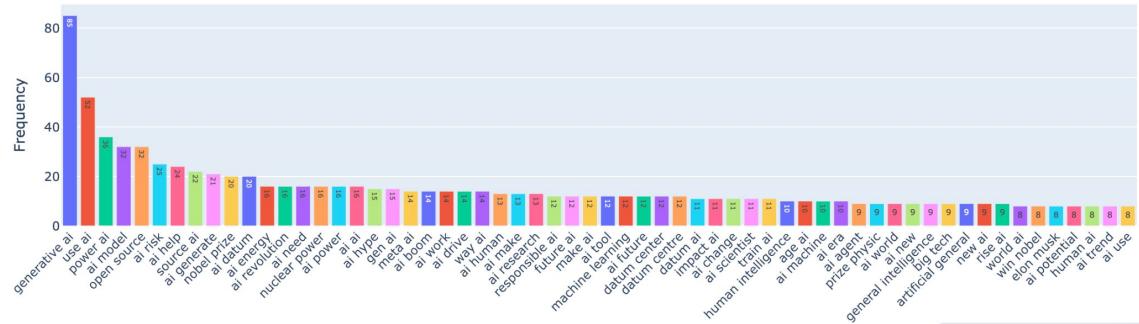
Existential Risks & AI Hype

- **AI super intelligence destroy, destroy humanity really, AI threat, AI hype** – Fears of AGI surpassing human intelligence spark debates on existential threats.
 - **AI scientific research, AI researchers say, AI researchers use** – Researchers analyze AI's risks and societal impacts.

AI Applications & Open Source Movement

- **AI model, AI machine learning, AI tool, AI agent** – AI tools are deployed across healthcare, education, and tech.
 - **AI open source, open source AI, AI datum center** – Open-source AI promotes transparency and collaboration.

Top 2-grams in {AI overview, risks & impact}

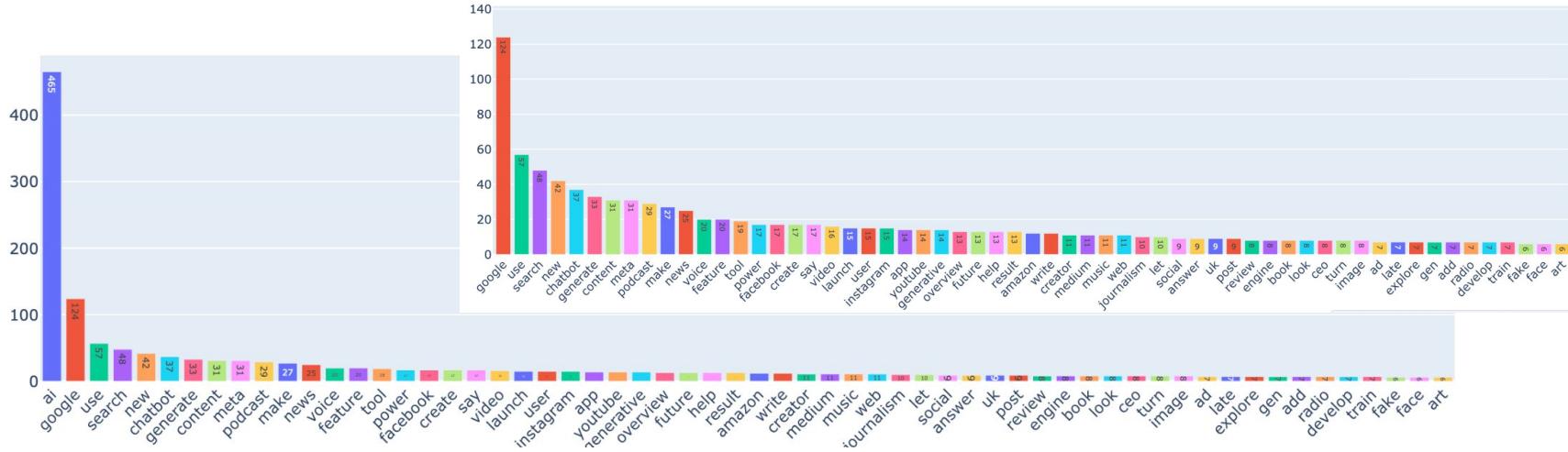


Top 3-grams in {AI overview, risks & impact}

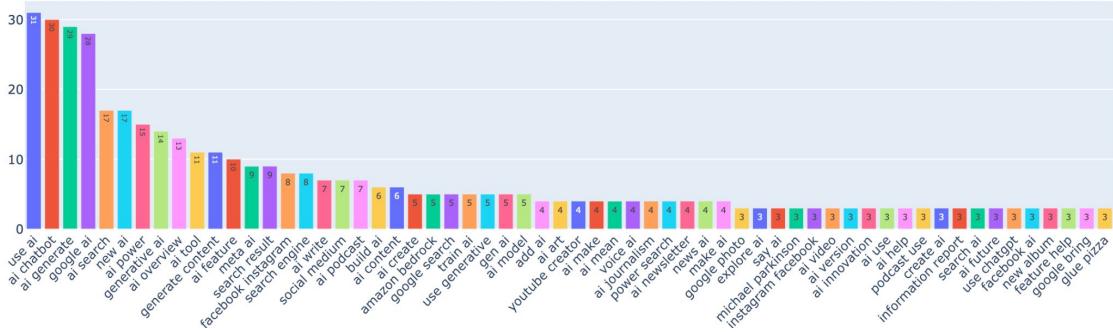


Other

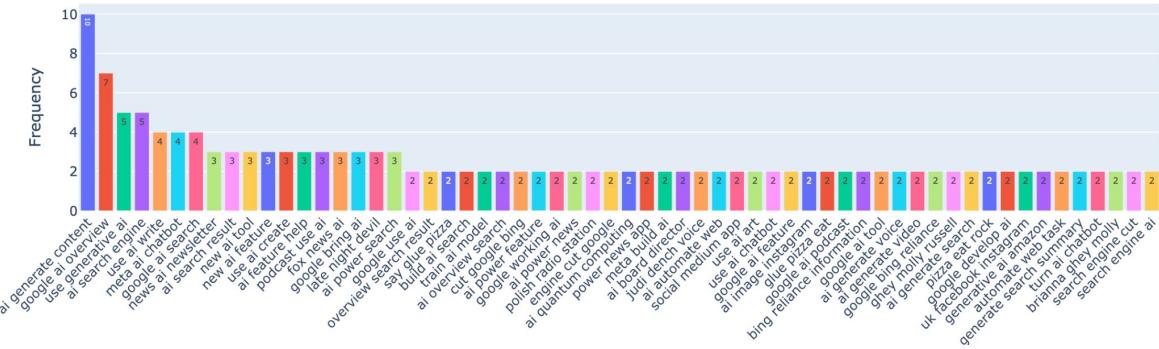
Top 1-grams in {Other}



Top 2-grams in {Other}



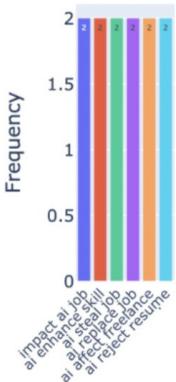
Top 3-grams in {Other}



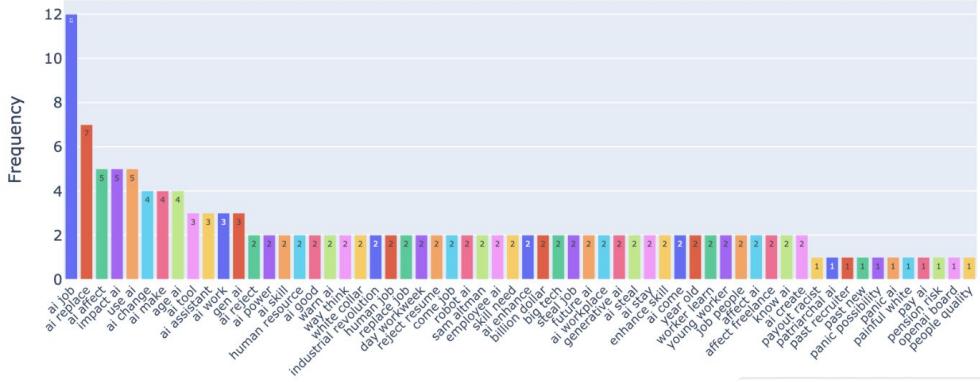
Fear prediction analysis

Used a pretrained model to predict the fear emotion of the headlines and analyzed the ngrams of the fear filtered data.

Note : the model is not fine tuned on this dataset



Top 2-grams in {Careers & Workforce}



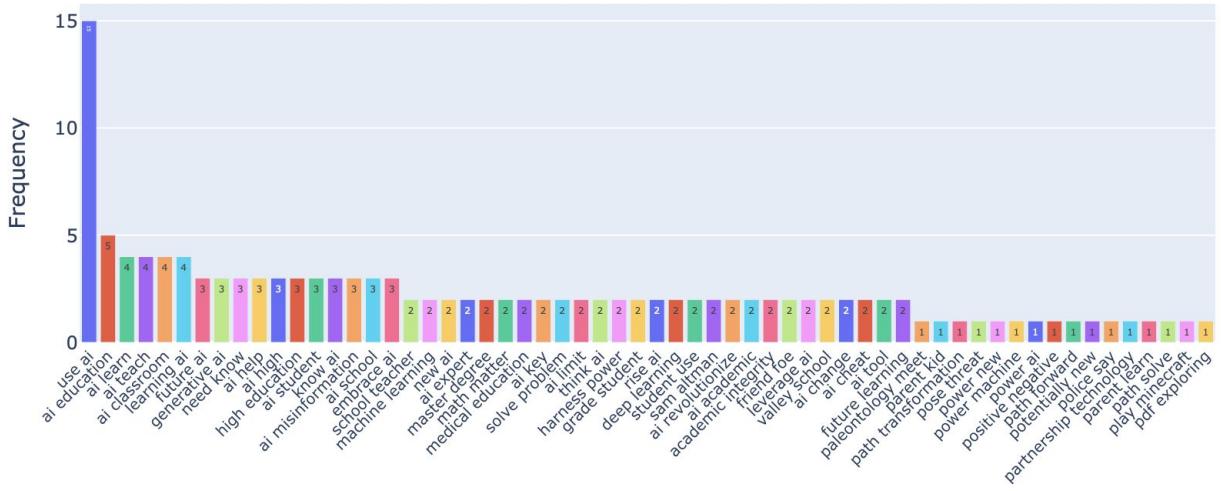
Fear-Based Insights from N-Grams (Career & Workforce)

1. Job Displacement & Automation → N-Grams: AI replace, AI steal job, replace job, AI effect job, AI work, robot AI
 - Fear that AI will take over jobs, especially white-collar and freelance work.
 2. Bias & Discrimination in Hiring → N-Grams: AI reject resume, reject resume, payout racist, human resource, AI hiring bias
 - Concerns about AI unfairly filtering candidates, leading to discrimination.
 3. Workplace Transformation & Skill Shifts → N-Grams: AI enhance skill, AI assistant, AI workplace, AI power, AI tool
 - Anxiety over the need for new skills and AI's growing influence in workplaces.
 4. Industrial & Societal Impact → N-Grams: industrial revolution, AI age, AI change, AI make, human job
 - AI is seen as driving a major industrial shift, creating uncertainty about the future of work.

Top 2-grams in {Education}

Academic Integrity & Misinformation

- N-Grams: AI cheat, AI misinformation, AI academic, academic integrity, student use
- Fears that AI enables cheating and spreads misinformation, threatening educational credibility.



Top 2-grams in {AI Ethics, Law & Policy}

1. Legal & Copyright Issues

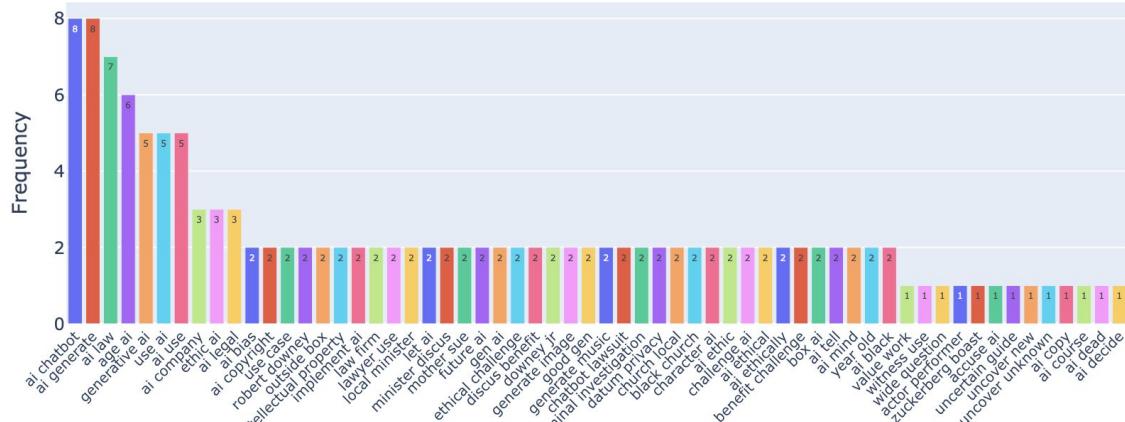
- N-Grams: AI copyright, intellectual property, AI law, lawyer use, plagiarism controversy
 - Concerns about AI-generated content violating copyright laws and legal uncertainties in ownership.

2. Privacy & Security Risks

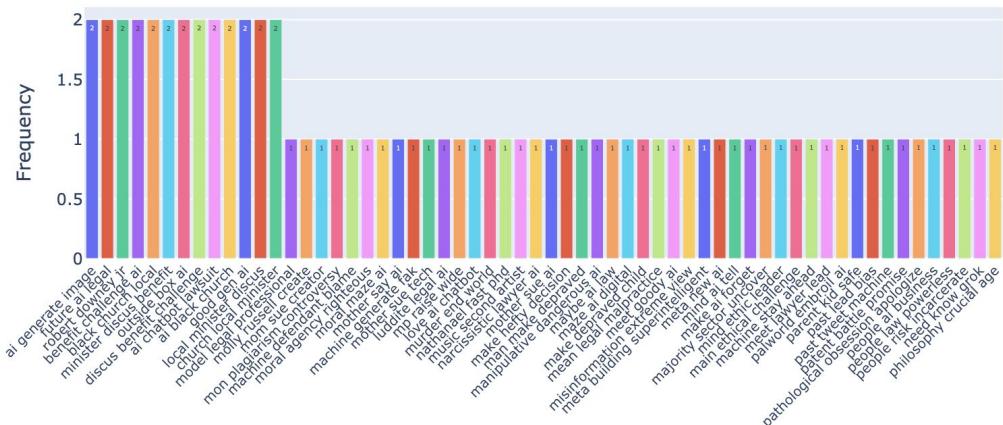
- N-Grams: data privacy, criminal investigation, challenge AI
 - Worries about AI misuse in surveillance, legal investigations, and personal data security.

3. AI-Generated Content & Accountability

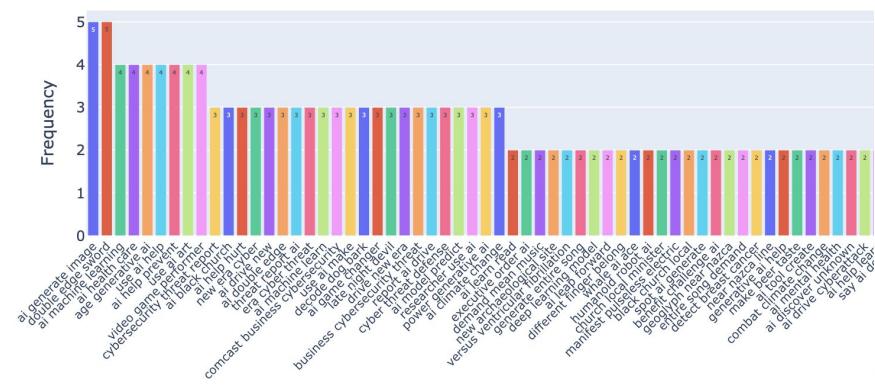
- N-Grams: AI generate image, generate music, chatbot lawsuit, AI generate
 - Concerns over AI-generated media leading to misinformation, deep fakes, and legal disputes.



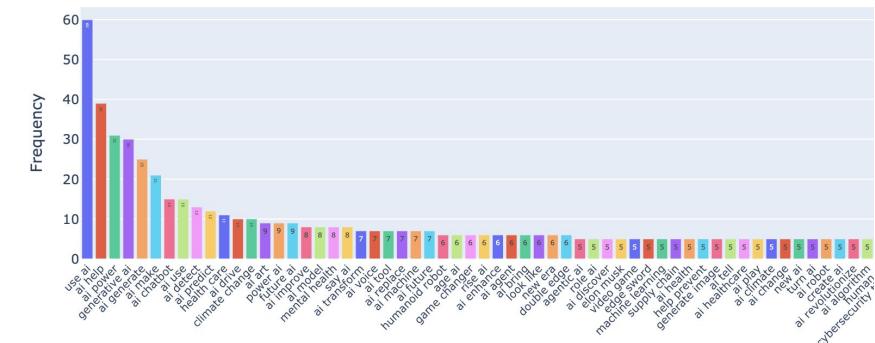
Top 3-grams in {AI Ethics, Law & Policy}



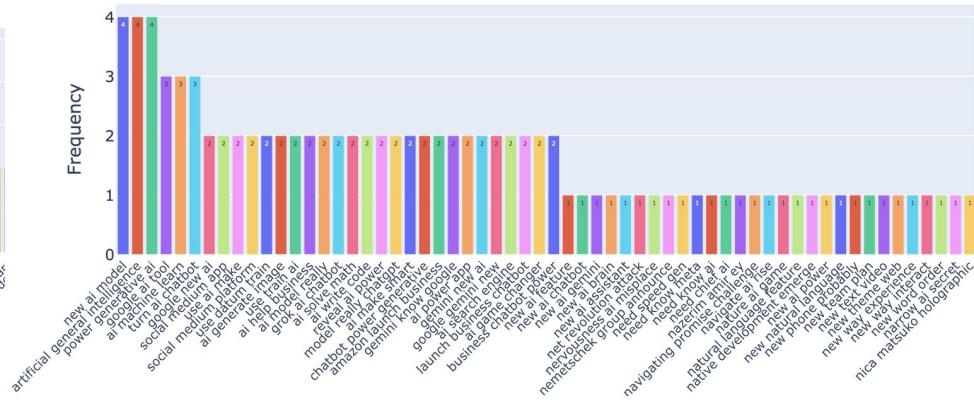
Top 3-grams in {AI in various Industries}



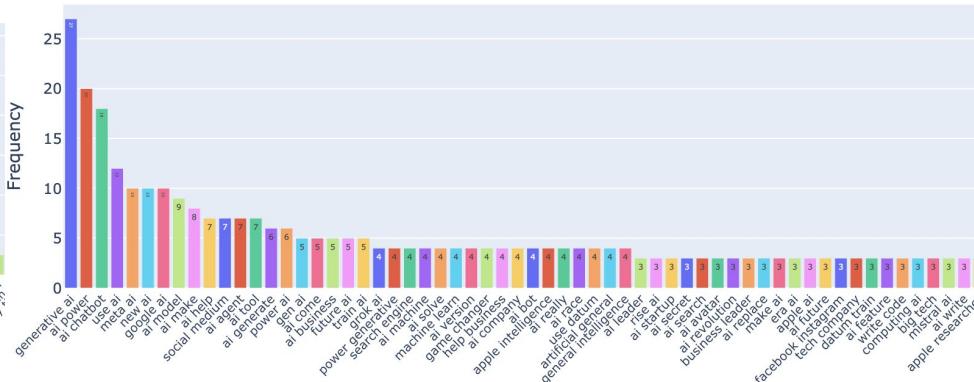
Top 2-grams in {AI in various Industries}



Top 3-grams in {AI in companies & Enterprises}



Top 2-grams in {AI in companies & Enterprises}



Criticism of AI Products from Tech Giants:

- N-Grams: Google AI, search engine, Meta AI, Apple AI, AI Chatbot
- The AI tools are facing scrutiny, possibly for misinformation, biases, or underperformance.

.AI's Role in Social Media & Manipulation:

- N-Grams: Social Media AI, Social Media Platform, Power AI, Use AI Meta
- Anxiety over AI's influence on social media, potentially spreading misinformation or manipulating public opinion.

AI-Generated Content & Ethical Concerns

- N-Grams: AI Generate Image, Use AI Make, ai art
- Skepticism about AI-generated images and content, with worries about accuracy, ethics, and the potential misleading visuals without copyright considerations

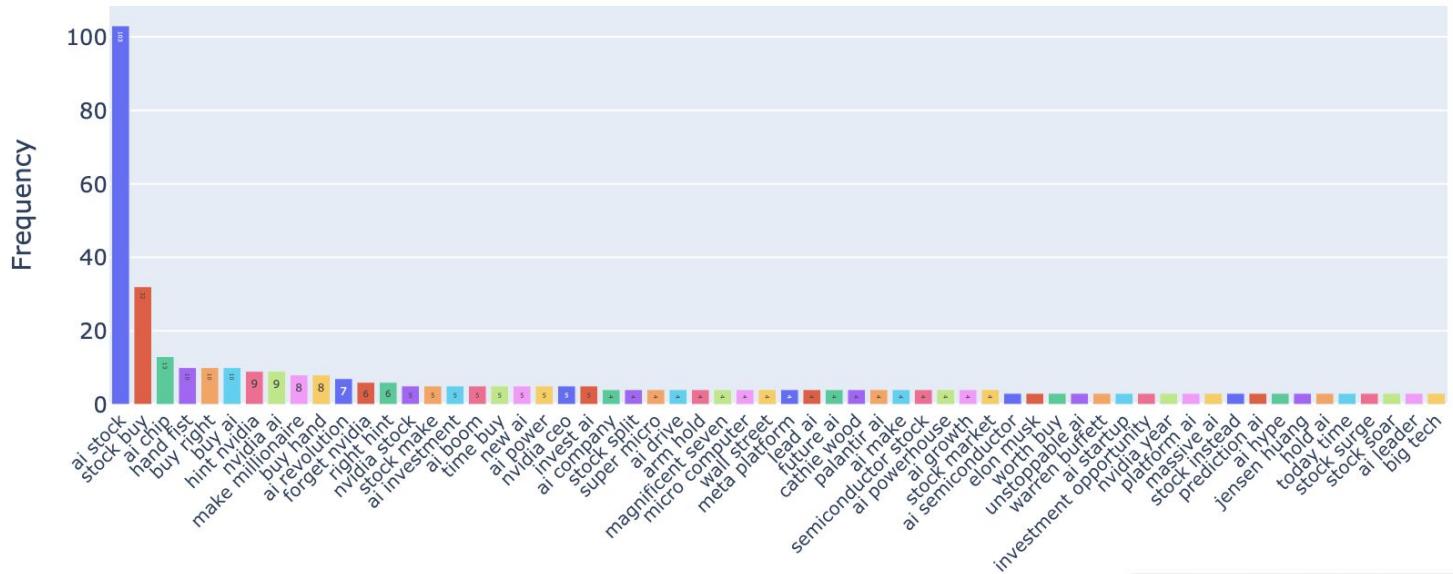
General AI & Automation Concerns

- N-Grams: AI Replace, AI Avatar, AI Computing
- Fears of AI surpassing human intelligence, automating jobs, and reducing human involvement in decision-making.

Cybersecurity Threats:

- N-Grams: Cyber Security Threat, AI Machine Learning, Business Cyber Security Threat, Era Cyber Threat
- Fears revolve around the use of AI in enhancing cyber threats, particularly in the context of business and cybersecurity, with concerns about the growing sophistication of attacks.

Top 2-grams in {AI Investments & Market Trends}



Our Future with AI – Utopia and Dystopia

1. AI's Promising Role in a Utopian Future

- **Transforming Industries & Enhancing Life:** AI improves efficiency, solves complex problems (e.g., climate change, mental health), and creates opportunities in healthcare, education, and business.
- **Boosting Productivity & Reducing Skill Gaps:** AI enhances work processes, personalized experiences
- **Human-AI Collaboration:** AI serves as a partner in research, business, and science, driving innovation (e.g., AI in climate science, medical research).

2. Dystopian Concerns About AI

- **Job Displacement & Economic Inequality:** Automation threatens jobs,
- **Ethical Risks & Misinformation:** AI bias, privacy issues, and misinformation pose risks
- **Surveillance & Privacy Violations:** AI-powered surveillance raises concerns about personal freedoms and data security.

3. Ethical & Regulatory Challenges

- **Need for Responsible AI Development:** Ethical frameworks and regulations are essential to prevent misuse (e.g., AI in legal, health, and security sectors).
- **Uncertainty of AI's Impact:** AI's long-term effects remain unclear, with both promises and risks shaping future societal changes.

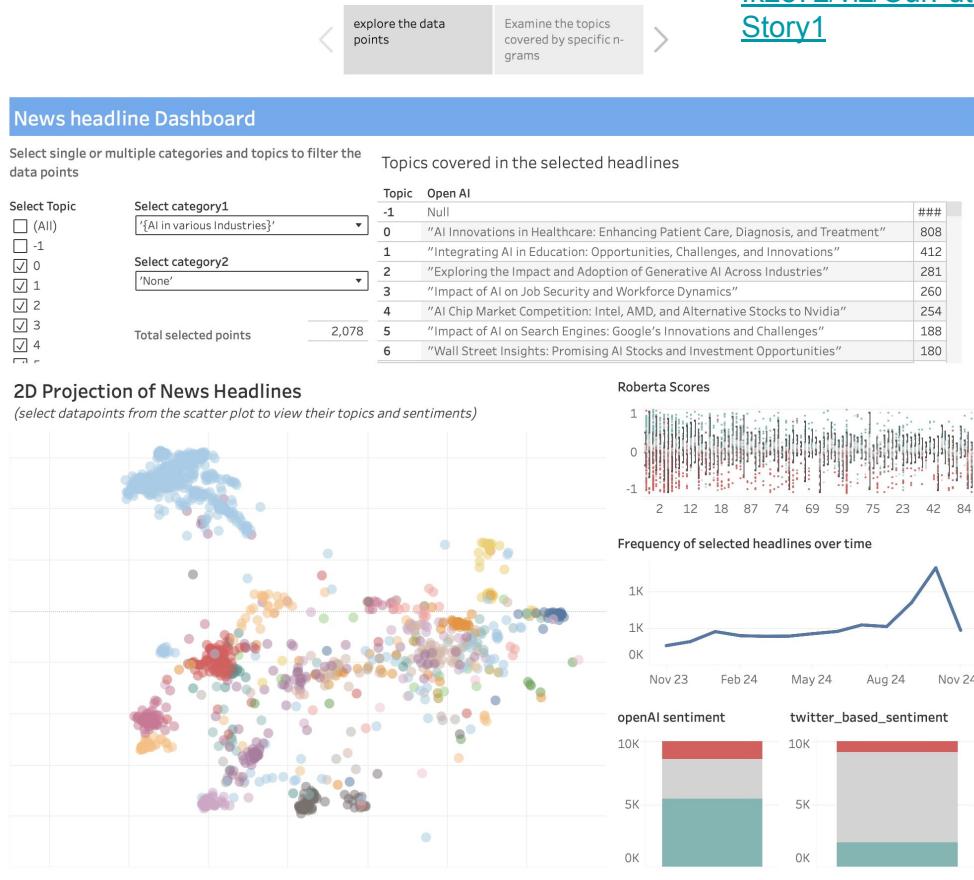
4. AI as a Double-Edged Sword

- **Mixed Future of AI:** AI offers both opportunities (efficiency, innovation) and risks (bias, privacy invasion, job loss).
- **Human Oversight is Crucial:** AI should augment, not replace, human decision-making, ensuring ethical and responsible use.

Conclusion: Toward a Balanced AI Future

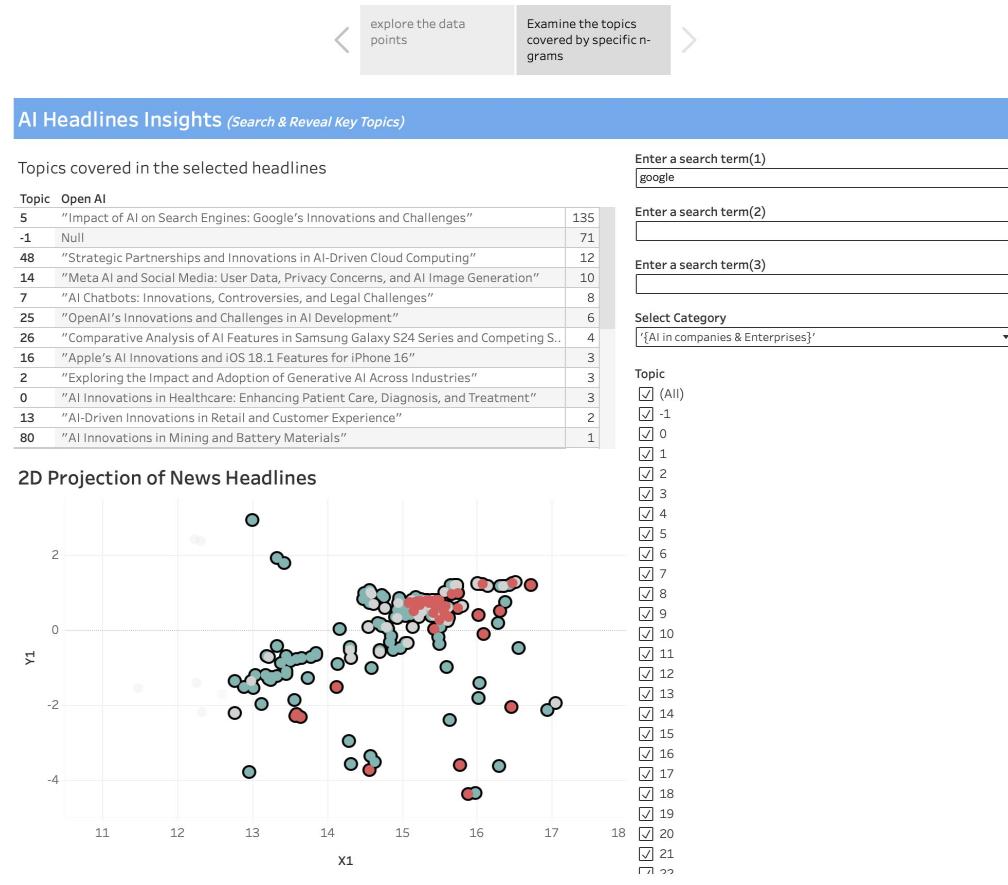
AI's impact depends on balancing innovation with accountability. Strong regulations, ethical frameworks, and human oversight are essential to harness AI's benefits while mitigating risks.

Tableau Dashboard 1



https://public.tableau.com/app/profile/amrutha_k2872/viz/OurFuturewithAIUtopiaorDystopia/Story1

Tableau Dashboard 2



THANK YOU