

REDEFINING NEWS CONSUMPTION

CLEAR, CONCISE, AND NOISE-FREE

JOURNEY 2023 - 2024 :



Founded in October 2023, Wat'sTheStory is an emerging startup transforming the way news is consumed. By simplifying content delivery, it focuses on providing users with concise, focused, and noise-free updates. The platform covers a diverse range of topics, catering to audiences who value clarity and actionable information. Through innovative strategies, Wat'sTheStory empowers users to stay informed without the overwhelm of traditional news formats

Working with Wat'sTheStory,

Introduction and Context

This project analyzes the Wat'sTheStory dataset, a collection of news articles on cybersecurity, technology, marketing, stock markets, and Irish news, spanning from October 2023 to 2024. The goal was to uncover meaningful trends, sentiment patterns, and thematic overlaps, aligning with Wat'sTheStory's mission of delivering clear and impactful news.

Analysis Techniques

The dataset was transformed and analyzed using:

- Data Cleaning: Handled missing values, standardized text, and added features like word count and cleaned summaries.
- Keyword Extraction: Used Scikit-learn's CountVectorizer to identify dominant themes.
- Sentiment Analysis: Applied TextBlob to classify sentiments into positive, negative, and neutral categories.
- Visualizations: Created charts, word clouds, and trends using libraries like Matplotlib and Plotly.

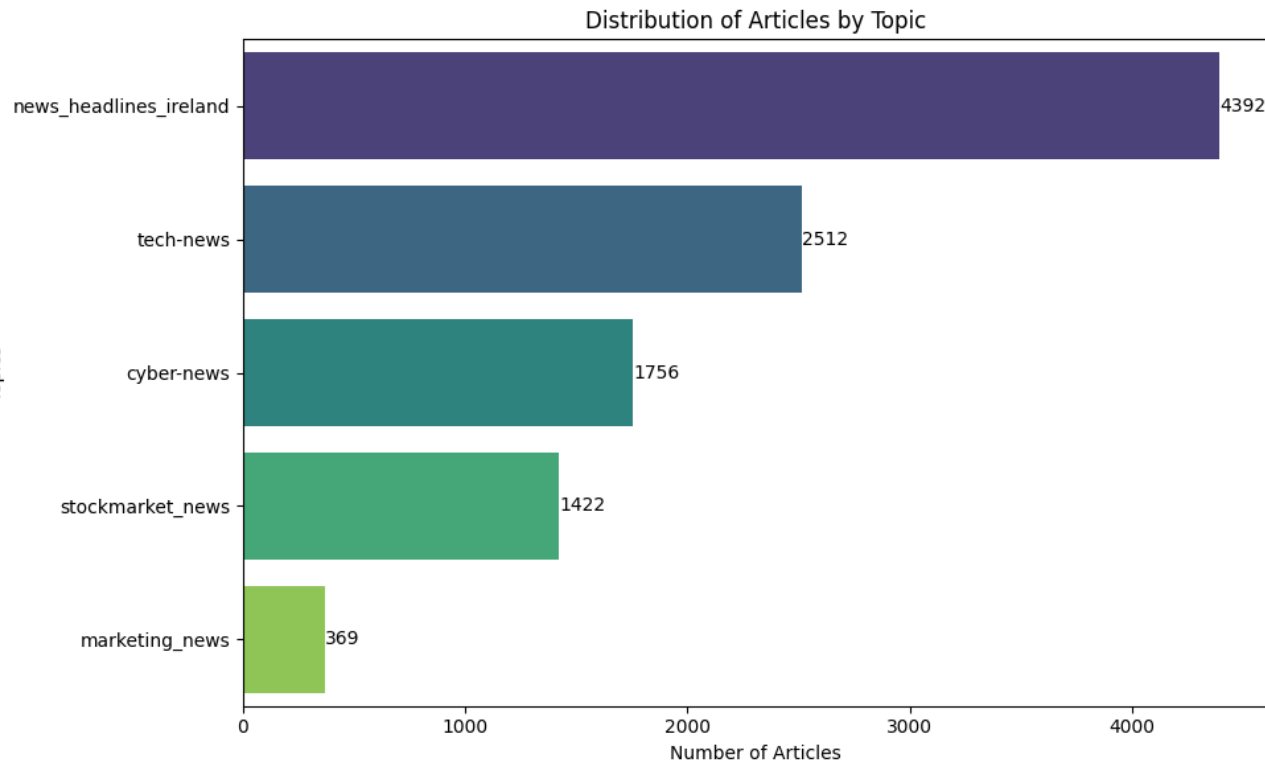
Raw Dataset

id	topic	collection_date	collected_data	comment
1	cyber-news	2023-10-08	Title: A Vulnerability in Apple Products Could Allow for Privilege Escalation Summary: <p>A vulnerability has been discovered in Apple products, which could allow for privilege escalation. Successful exploitation of this vulnerability could allow for privilege escalation in the context of the logged on user. Depend Title: Bounty offered for secret NSA seeds behind NIST elliptic curves algo Summary: A bounty of \$12,288 has been announced for the first person to crack the NIST elliptic curves seeds and discover the original phrases that were hashed to generate them. [...] Title: D.C. Board of Elections confirms voter data stolen in site hack Summary: The District of Columbia Board of Elections (DCBOE) is currently probing a data leak involving an unknown number of voter records following breach claims from a threat actor known as RansomedVC. [...]	
2	tech-news	2023-10-08	Title: The cheapest 4TB PCIe 4.0 SSD on Amazon has been tested ahead of Prime Day — and it reveals a worrying trend Summary: The Fanxiang S660 retails for just \$158, but isn't all it's cracked up to be. Title: A tiny startup has helped Intel trounce AMD and Nvidia in critical AI tests — is it game over already? Summary: This is the latest twist in the AI hardware arms race, with Numenta applying a novel approach to boosting CPU performance Title: And that's how it all started: AI managed to design a robot that walks — from scratch, and in less than 60 seconds Summary: This is the first algorithm that can design robots that actually work in the real world. Title: Don't expect a budget Google Pixel phone anytime soon Summary: A Google exec explains why we won't be getting an ultra-cheap Google Pixel phone in the near future. Title: 'It's systemic': Loki season 2 director speaks up for Marvel's overworked VFX artists Summary: One of Loki season 2's directors has opened up on the working conditions of Marvel's VFX team, who recently voted to unionize. Title: The OnePlus Watch 2 is being tipped to launch next year Summary: It's been a bit quiet on the OnePlus wearable front recently, but it seems the OnePlus Watch 2 is on the way. Title: Beware - GoldDigger malware will drain your bank accounts without you even realizing Summary: A new and dangerous Android malware has emerged, stealing people's money and cryptocurrencies. Title: Quordle today - hints and answers for Sunday, October 8 (game #622) Summary: Looking for Quordle clues? We can help. Plus get the answers to Quordle today and past solutions. Title: Intel wants its GPUs to take a bigger slice of the VDI market by eliminating licensing fees Summary: Intel's Flex GPUs are tailor-made for the VMware Horizon suite of VDI tools and services. Title: Are hyper-specialized processing units the future of computing? Meet the company who wants to be the Nvidia of data queries Summary: Neuroblade's SPU for data analytics workloads can accelerate performance by up to 100 times. Title: Prime Video's new horror movie is Totally Killer with no filler - and so are these 6 great slasher films Summary: Totally Killer isn't the only terrifically terrifying slasher film on the world's best streaming services. Title: This massive new spoofing campaign is targeting job seekers, so watch out. Summary: A sophisticated scam involving cryptocurrency exchanges is rampant across the globe. Title: The Samsung Galaxy Ring may launch with the Galaxy Z Fold 6 Summary: A new report suggests the Samsung smart ring is going to show up later in 2024 or early in 2025. Title: Next-gen Windows is coming in 2024, Intel exec confirms (without mentioning Windows 12) Summary: Intel has been the source of a leak around Windows 12 previously, and the spillage keeps coming. Title: The Samsung Galaxy S24 is tipped to use both Exynos and Snapdragon chipsets Summary: A new report says only the S24 Ultra model will use a Snapdragon CPU in all the regions where it's sold. Title: ICYMI: the 8 biggest tech stories of the week, from smart tags to the Pixel 8 Pro Summary: It's been an incredibly busy week in the world of tech, but we've recapped everything that's happened for you. Title: The inside of the world's largest spherical screen is jaw-dropping — but wait till you see how the Las Vegas Sphere looks on the outside Summary: At 54,000 square meters, the Las Vegas Sphere is the largest display on the planet. Title: A new ChatGPT dating scam is looking to catfish lonely AI fans Summary: Generative AI is causing more havoc as AI chatbots such as ChatGPT pop up in dating apps, catfishing unsuspecting lovers.	
3	cyber-news	2023-10-09	Title: GNOME Linux systems exposed to RCE attacks via file downloads Summary: A memory corruption vulnerability in the open-source libcue library can let attackers execute arbitrary code on GNOME Linux systems. [...] Title: Over 17,000 WordPress sites hacked in Balada Injector attacks last month Summary: Multiple Balada Injector campaigns have compromised and infected over 17,000 WordPress sites using known flaws in premium theme plugins. [...] Title: Hackers modify online stores' 404 pages to steal credit cards Summary: A new Magecart card skimming campaign hijacks the 404 error pages of online retailer's websites, hiding malicious code to steal customers' credit card information. [...] Title: Hackers hijack Citrix NetScaler login pages to steal credentials Summary: Hackers are conducting a large-scale campaign to exploit the recent CVE-2023-3619 flaw in Citrix NetScaler Gateways to steal user credentials. [...]	

Final Dataset

id	topic	collection_date	Title	word_count	year	month	day	YearMonth	Cleaned_Summary
1	cyber-news	2023-10-08	A Vulnerability in A	84	2023	10	8	2023-10	a vulnerability has been c
1	cyber-news	2023-10-08	Bounty offered for	30	2023	10	8	2023-10	a bounty of 12288 has b
1	cyber-news	2023-10-08	D.C. Board of Elec	32	2023	10	8	2023-10	the district of columbia b
2	tech-news	2023-10-08	The cheapest 4TB	15	2023	10	8	2023-10	the fanxiang s660 retails
2	tech-news	2023-10-08	A tiny startup has l	21	2023	10	8	2023-10	this is the latest twist in t
2	tech-news	2023-10-08	And that's how it a	16	2023	10	8	2023-10	this is the first algorithm
2	tech-news	2023-10-08	Don't expect a buc	18	2023	10	8	2023-10	a google exec explains w
2	tech-news	2023-10-08	'It's systemic': Lok	22	2023	10	8	2023-10	one of loki season 2s dir
2	tech-news	2023-10-08	The OnePlus Watc	22	2023	10	8	2023-10	its been a bit quiet on the
2	tech-news	2023-10-08	Beware - GoldDigg	13	2023	10	8	2023-10	a new and dangerous an
2	tech-news	2023-10-08	Quordle today - hir	17	2023	10	8	2023-10	looking for quordle clues
2	tech-news	2023-10-08	Intel wants its GPU	15	2023	10	8	2023-10	intels flex gpu are tailor
2	tech-news	2023-10-08	Are hyper-specializ	14	2023	10	8	2023-10	neuroblades spu for data
2	tech-news	2023-10-08	Prime Video's new	15	2023	10	8	2023-10	totally killer isnt the only
2	tech-news	2023-10-08	This massive new	11	2023	10	8	2023-10	a sophisticated scam inv
2	tech-news	2023-10-08	The Samsung Gale	20	2023	10	8	2023-10	a new report suggests th
2	tech-news	2023-10-08	Next-gen Windows	17	2023	10	8	2023-10	intel has been the source
2	tech-news	2023-10-08	The Samsung Gale	21	2023	10	8	2023-10	a new report says only th
2	tech-news	2023-10-08	ICYMI: the 8 bigge	19	2023	10	8	2023-10	its been an incredibly bu
2	tech-news	2023-10-08	The inside of the w	15	2023	10	8	2023-10	at 54000 square meters t
2	tech-news	2023-10-08	A new ChatGPT da	20	2023	10	8	2023-10	generative ai is causing r
3	cyber-news	2023-10-09	GNOME Linux sys	20	2023	10	9	2023-10	a memory corruption vuln
3	cyber-news	2023-10-09	Over 17,000 Wordf	20	2023	10	9	2023-10	multiple balada injector c
3	cyber-news	2023-10-09	Hackers modify on	25	2023	10	9	2023-10	a new magecart card ski
3	cyber-news	2023-10-09	Hackers hijack Citr	21	2023	10	9	2023-10	hackers are conducting e
3	cyber-news	2023-10-09	HelloKitty ransomv	32	2023	10	9	2023-10	a threat actor has leaked
3	cyber-news	2023-10-09	Google Chrome's c	47	2023	10	9	2023-10	in a bid to upgrade user
3	cyber-news	2023-10-09	Microsoft 365 adm	26	2023	10	9	2023-10	microsoft urged microsof

Article Distribution Across Topics:



Dominant Topic:

News_Headlines_Ireland is the most covered topic with 4,392 articles, which highlights the focus on Ireland's local and global news coverage.

Second Highest:

Tech-News comes second with 2,512 articles, showing significant interest in technology updates and innovations.

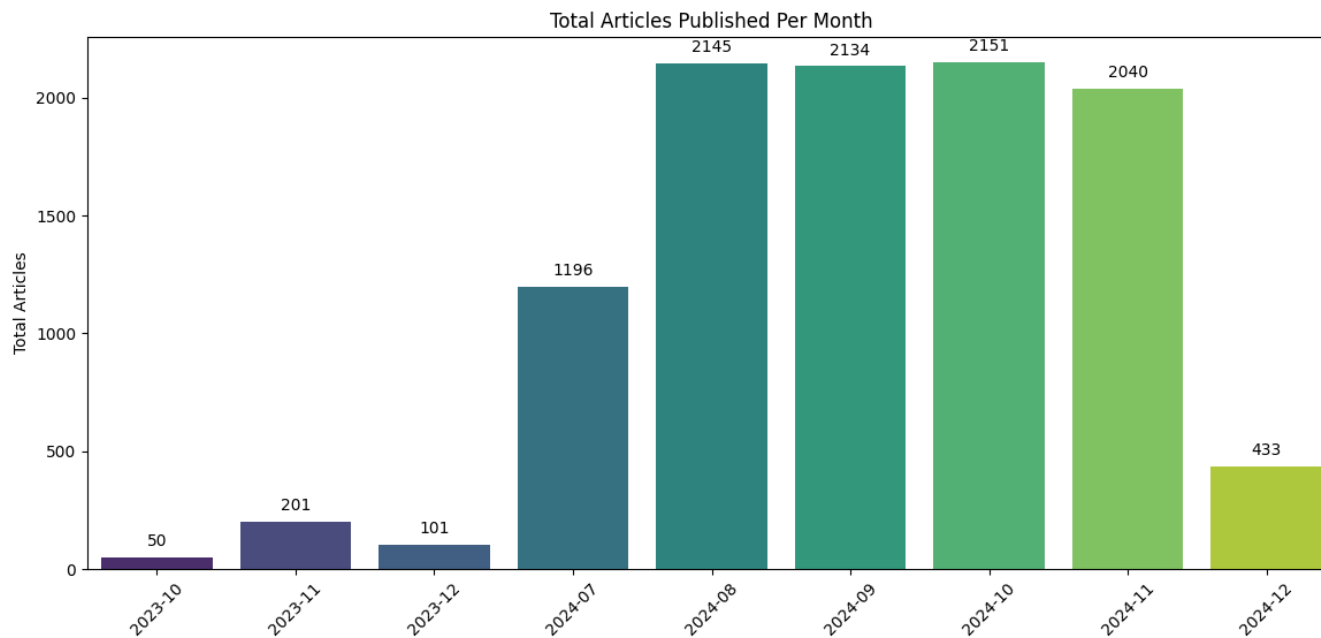
Niche Focus:

Topics like Marketing News have only 369 articles, indicating limited coverage or a niche audience.

This emphasis highlights the platform's focus on delivering news that resonates with both local and global audiences, ensuring relevance and engagement.

Overview of Articles:

Monthly Article Distribution :



Trend: The number of articles has significantly increased from October 2023 (50 articles) to peak months like August 2024 (2145 articles), showing steady growth in content production.

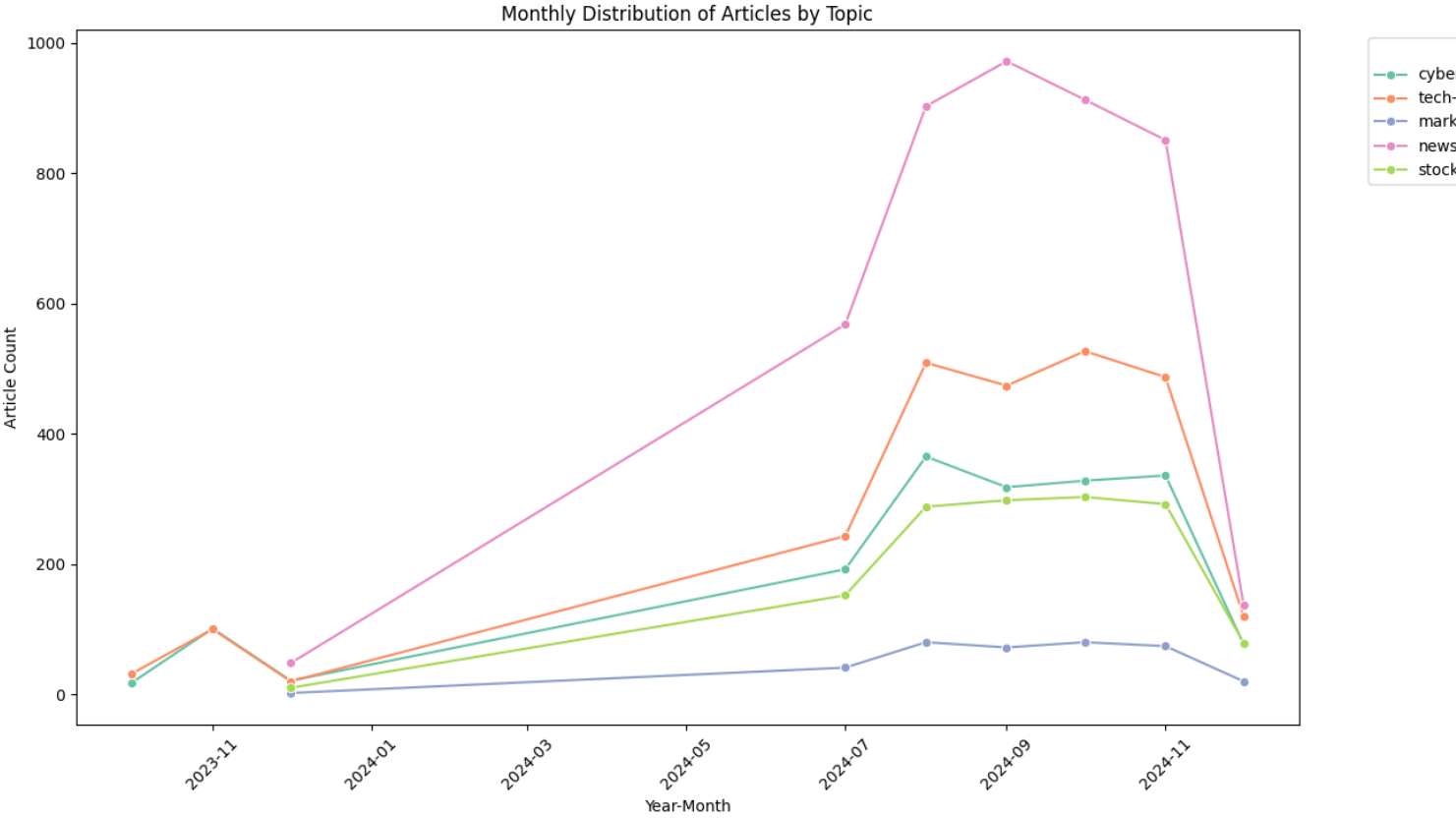
Key Period: The months from August to October 2024 were highly productive, with over 2000 articles each month

This steady growth reflects Wat'sTheStory's capability to scale its content production and cater to an expanding audience base over time.

Monthly Distribution of Articles by Topic:

This visualization illustrates the growth and consistency of article contributions across various topics over time. It highlights the diversity in content creation with:

Significant increase in content production across all topics during mid-2024, showcasing progress and engagement.

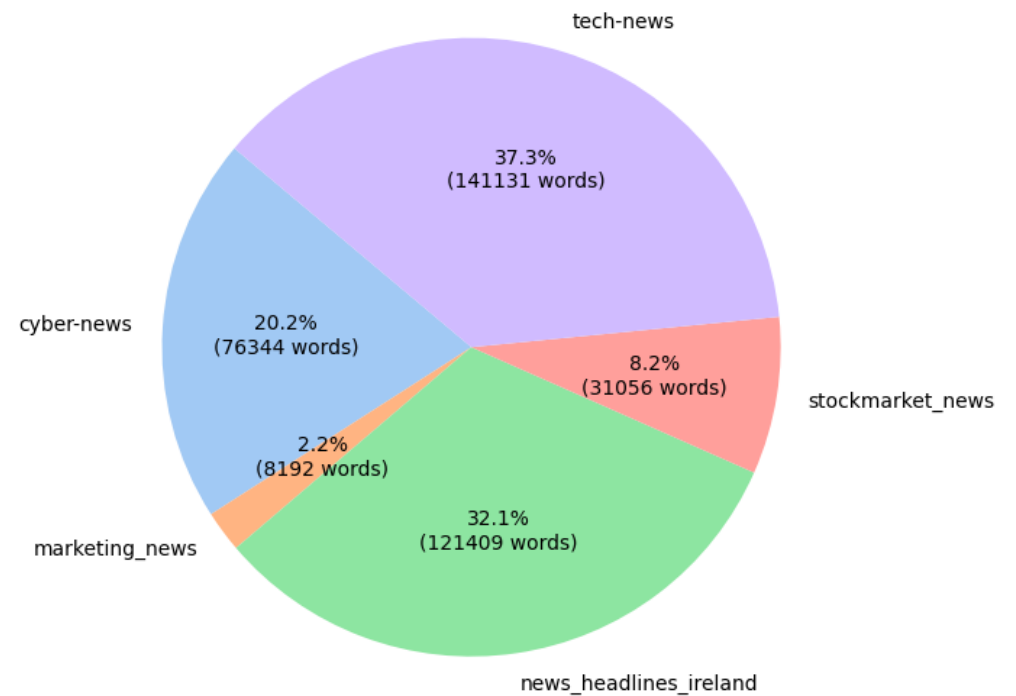


Total Word Count/percentage by Topic

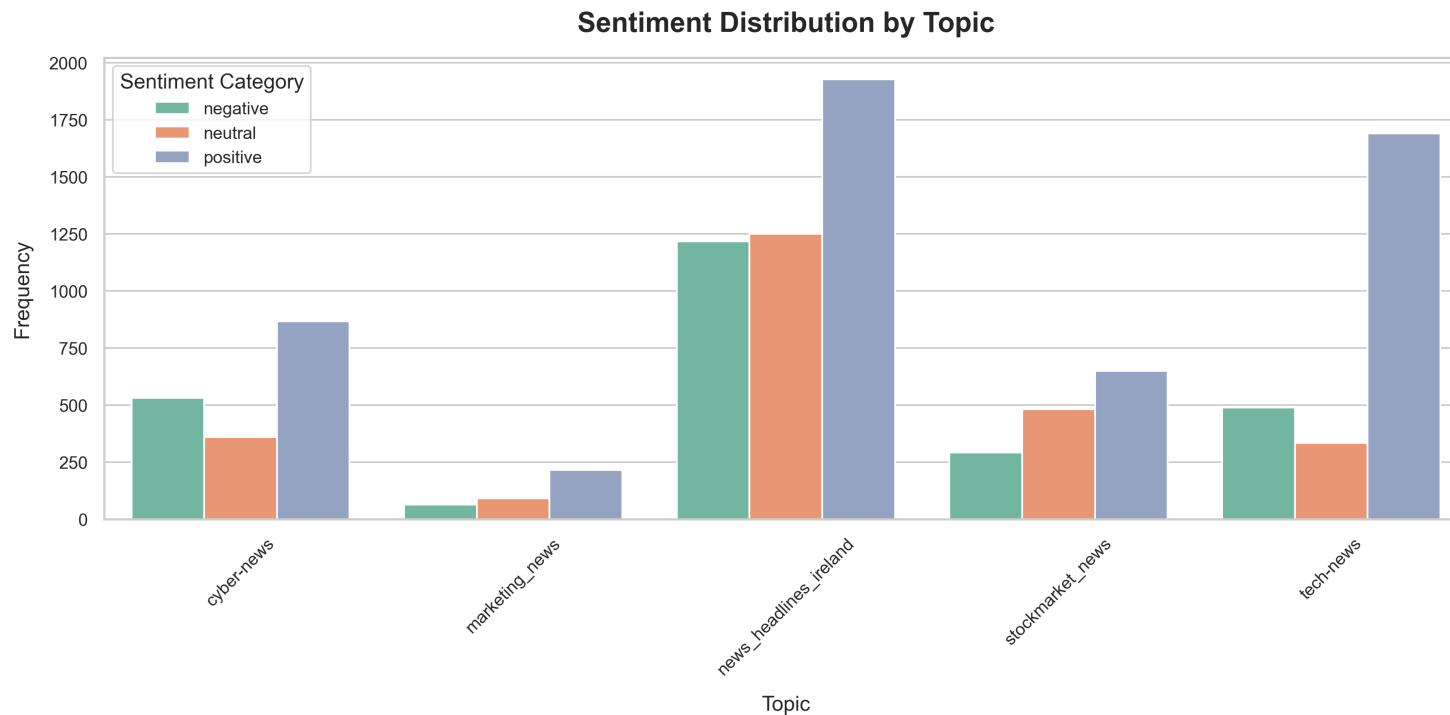
This pie chart provides a comprehensive overview of the word distribution across topics, highlighting the content focus within the project:

The distribution reflects a well-rounded approach, with a strong emphasis on Tech-News and News Headlines Ireland.

Proportion of Total Word Count by Topic (Counts and Percentages)



Insights on Sentiment Distribution by Topic:



1. Tech News: Predominantly positive, reflecting optimism around advancements and innovations.

2. News Headlines Ireland: A mix of negative and positive, likely tied to diverse socio-political events.

3. Cyber News: Mostly negative, emphasizing concerns over threats and vulnerabilities.

4: Stock Market News: Balanced between neutral and negative,

reflecting cautious market reporting.

5. Marketing News: Primarily neutral, focusing on industry updates and strategies.

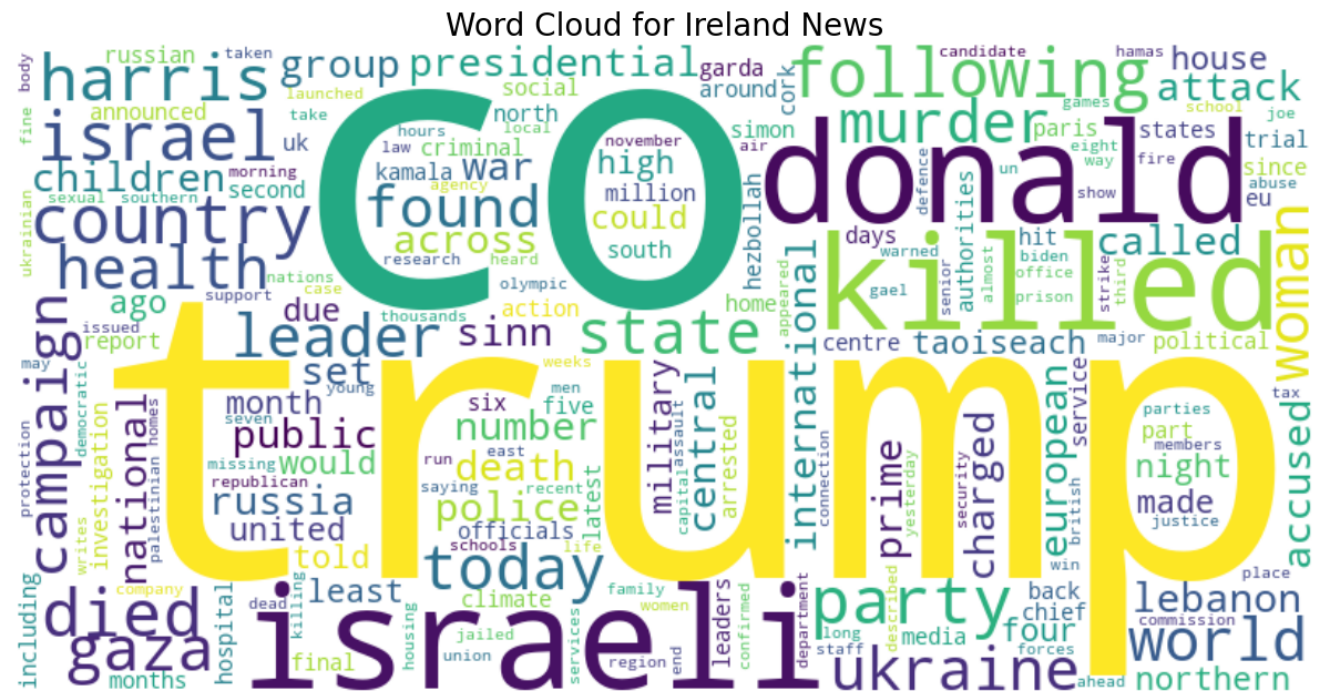
Ireland News Word Cloud:

Key Highlights:

The word cloud highlights major themes like Trump, Israel, Gaza, and Ukraine, reflecting Ireland's engagement in global conflicts and political discourse. Domestic issues such as crime, health, and leadership (e.g., Taoiseach, police) also stand out, showcasing a balance between international and local priorities.

Note:

This snapshot provides clarity on Ireland's news focus, helping users grasp key topics shaping public and policy discussions.



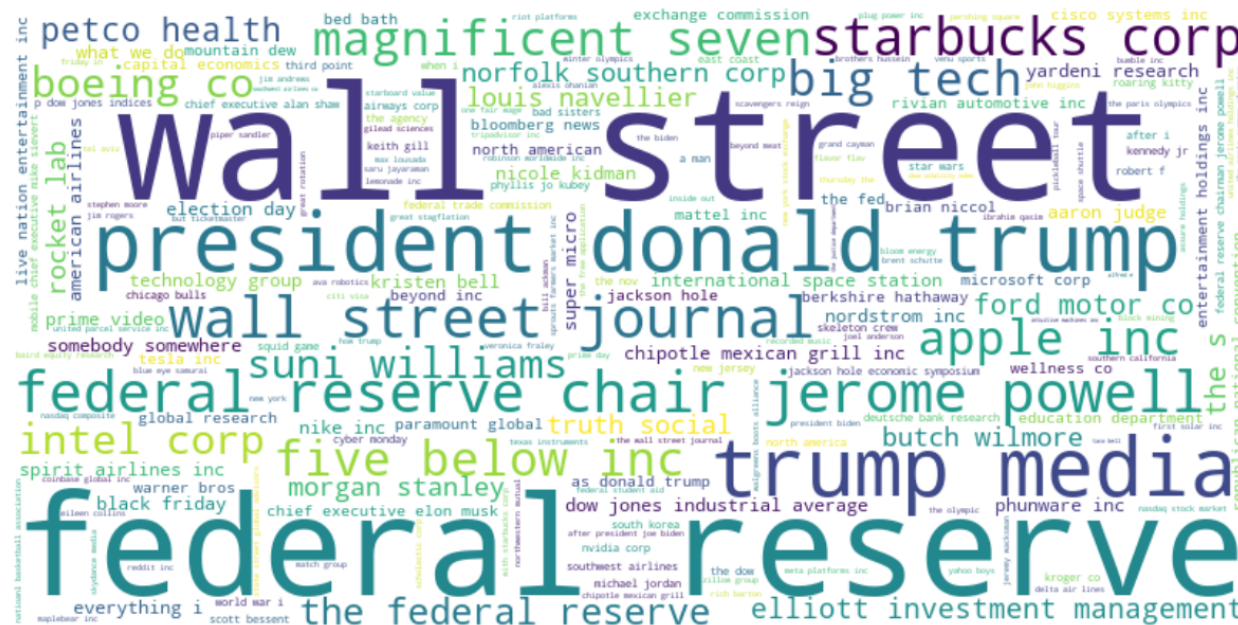
Stock News Word Cloud:

Key Highlights:

- "Wall Street" and "Federal Reserve" dominate the discussion, reflecting their central roles in shaping financial markets and monetary policies.
- Mentions of "Donald Trump" and "President" indicate the influence of political figures and decisions on the stock market.
- Companies like "Apple Inc.", "Intel Corp.", and "Starbucks Corp." highlight their relevance and news coverage in financial contexts.

- “Five Below Inc.” and “Boeing Co.”
showcase specific corporate activities and
impacts being closely monitored.

Word Cloud for Stock News



Note:

This visualization underscores the intersection of politics, corporate actions, and monetary policies in shaping stock market trends. It provides an informative snapshot of key players and influences in the financial domain.

Cyber News Word Cloud:

Key Highlights:

- Prominent mentions of “Microsoft” and “Windows” underline the tech giant’s pivotal role in cybersecurity updates and software vulnerabilities.
- Keywords like “malware”, “ransomware”, and “threats” emphasize the ongoing cyber risks organizations and individuals face.
- Terms such as “critical”, “vulnerability”, and “discovered” indicate active monitoring and identification of system flaws.
- Mentions of “phishing”, “accounts”, and “credentials” suggest a focus on protecting sensitive user information against exploitation.

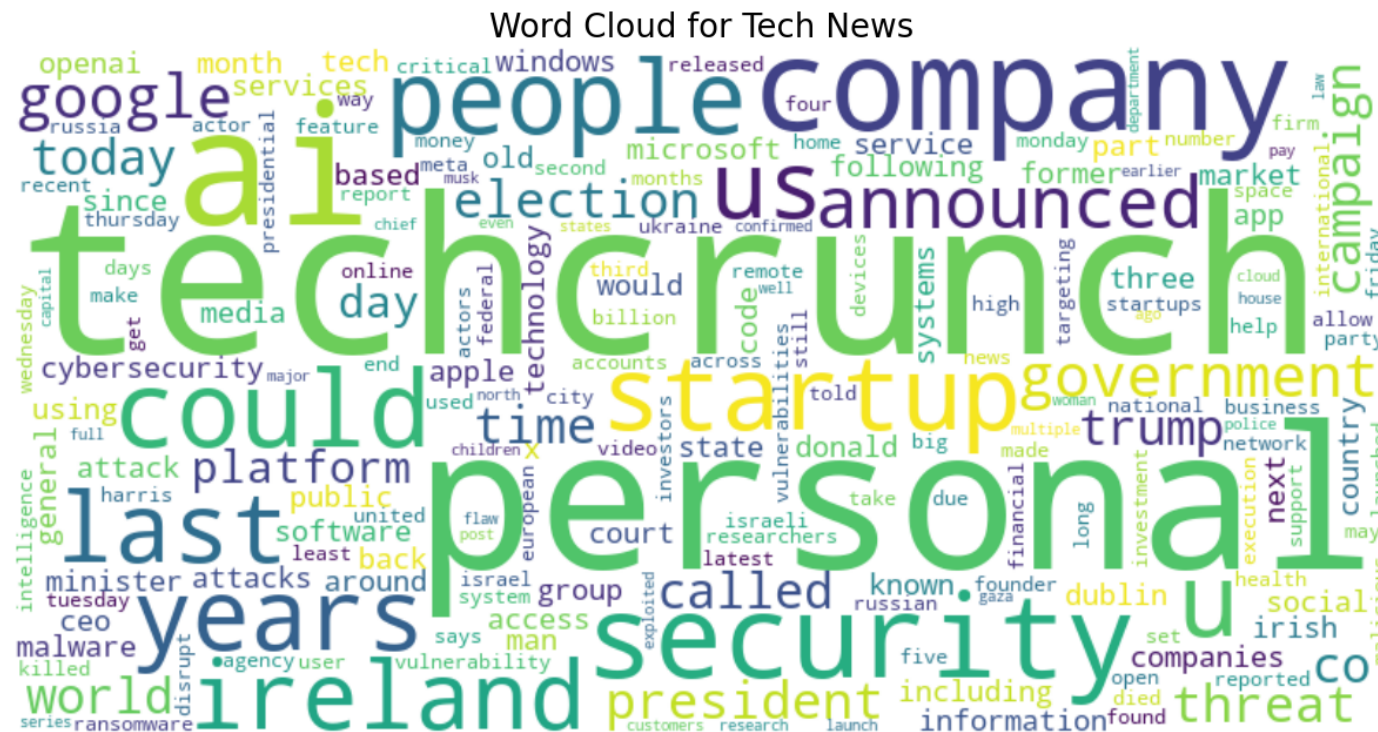
Note:

This visualization underscores the importance of maintaining robust cybersecurity measures, with active updates and vigilance against vulnerabilities being central to digital safety strategies.



Tech News Word Cloud:

Key Highlights:



- Dominant keywords like “TechCrunch,” “startup,” and “security” reflect a focus on technology innovation and cybersecurity.
- Mentions of “election” and “government” indicate intersections of tech and policy discussions.
- Companies like “Google” and “Apple” appear prominently, showcasing their impact on the tech landscape.

Note:

The word cloud provides a glimpse into the broad spectrum of tech-related topics, from startups to cybersecurity and tech giants influencing daily life.

Marketing News Word Cloud:

Key Highlights:

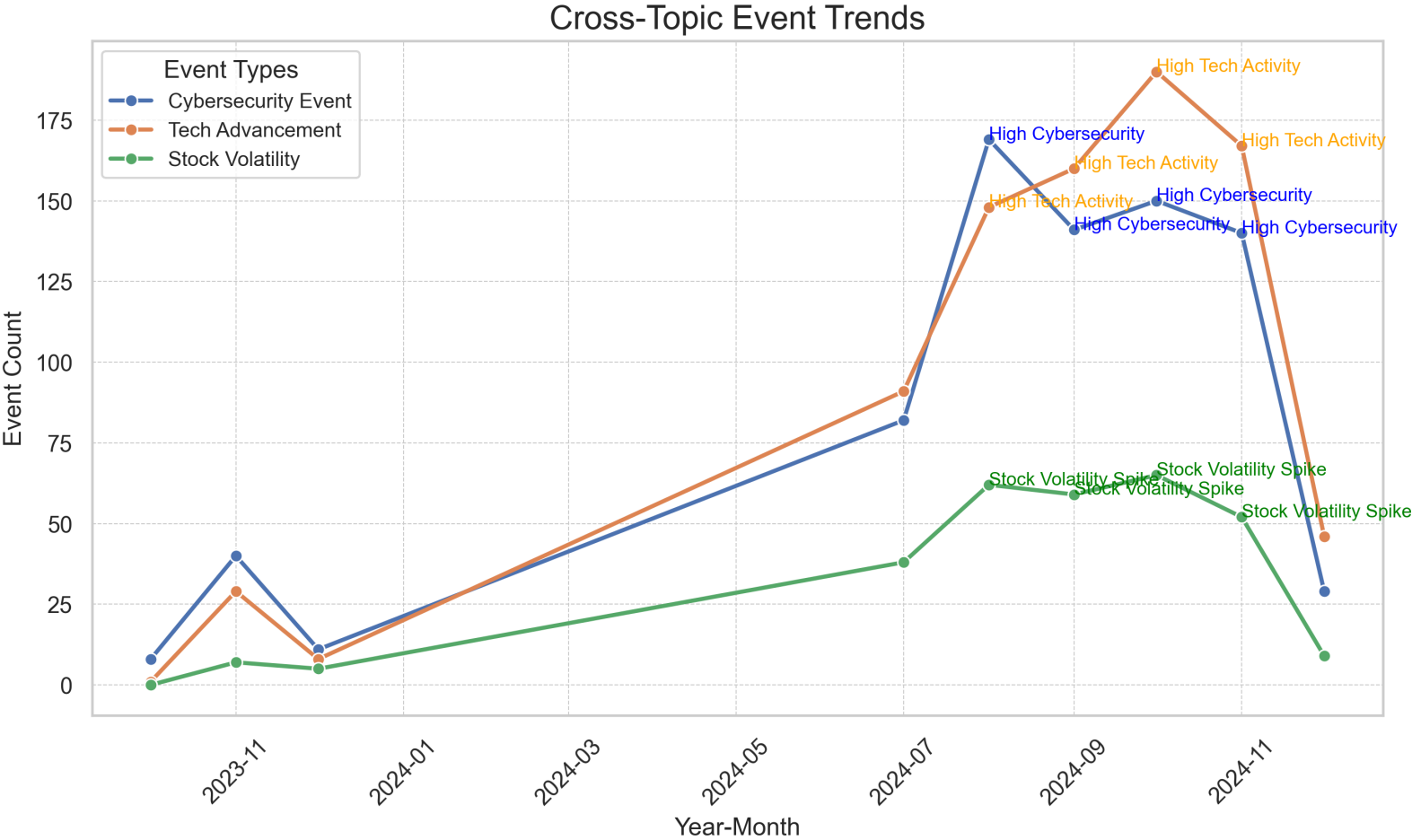
- Keywords like “advertising week,” “pizza chains,” and “urban future” highlight trends in marketing campaigns and consumer engagement.
- Notable mentions of individuals like “Angela Zepeda” and “George Dewey” suggest influential personalities in recent campaigns.
- Topics like “Gen Alpha” and “Dentsu Creative” show focus areas for future market strategies.

Note:

This visualization emphasizes the evolving strategies of brands and agencies to adapt to changing consumer preferences and emerging demographics.

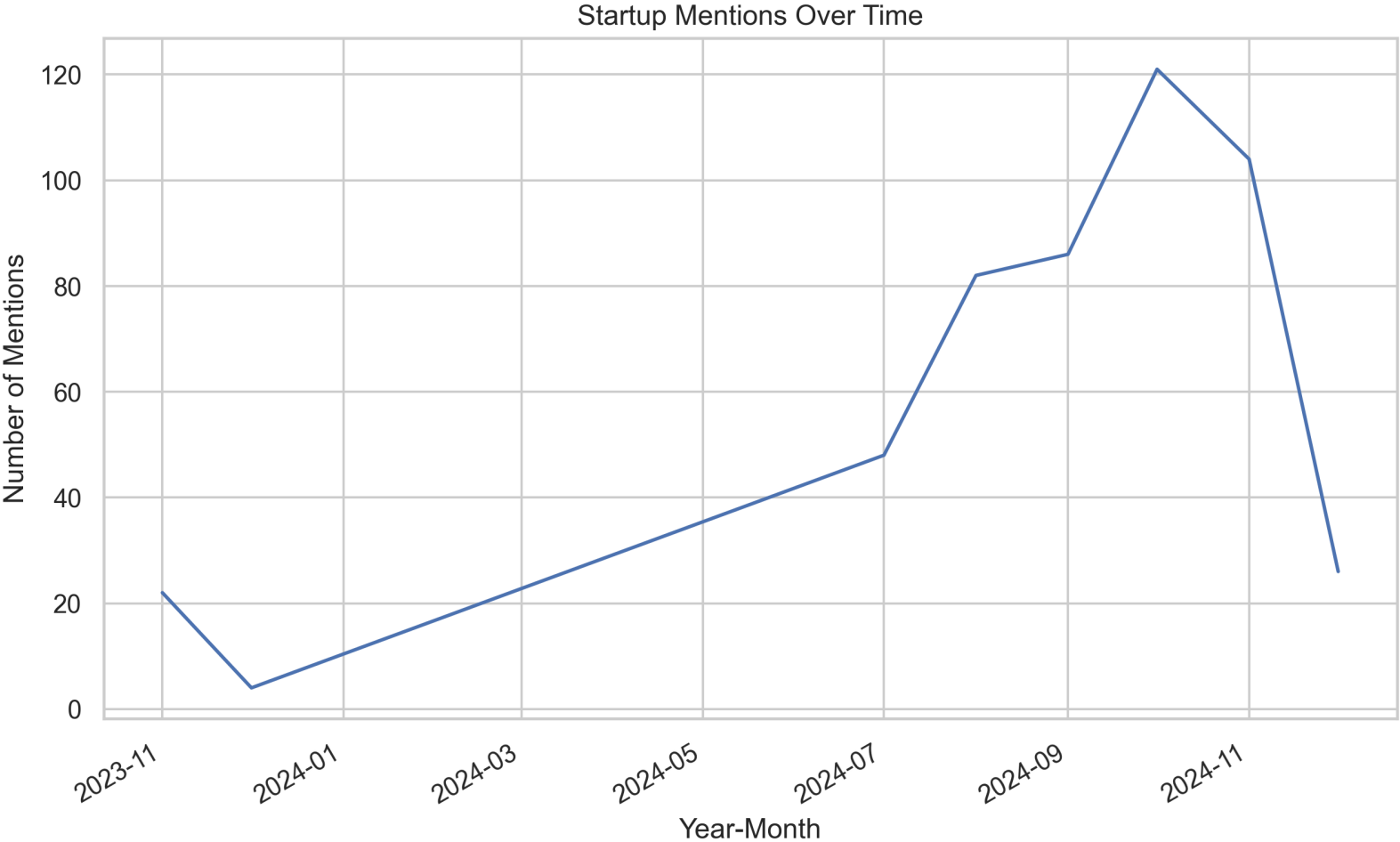


Interlinked Dynamics of Cybersecurity, Technology, and Market Behavior:



This analysis reveals a significant interconnection between cybersecurity events, technological advancements, and stock market volatility. Discussions around these topics often occur in tandem, with technological innovations frequently driving conversations about cybersecurity risks. Interestingly, stock market volatility tends to trail these discussions, suggesting a delayed response from financial markets to key technological or security developments.

Trend Analysis: Fluctuations in Startup Mentions Over Time:



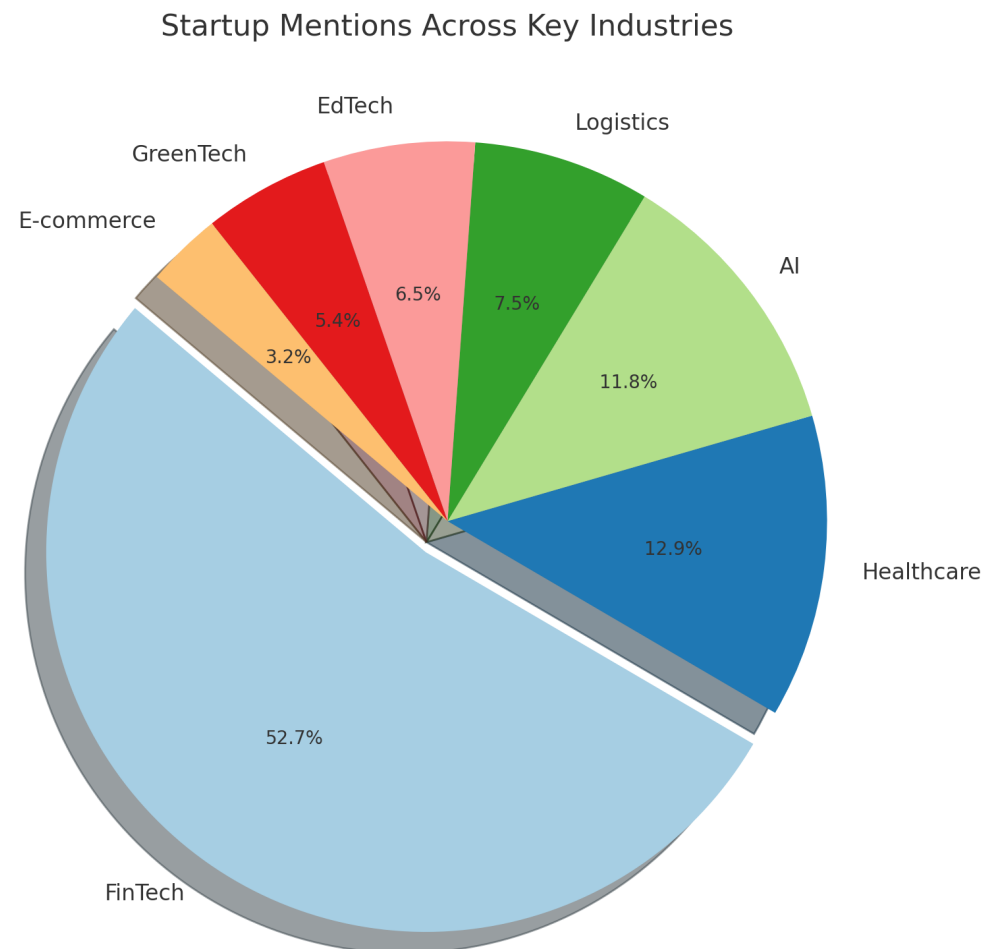
A noticeable decline in startup mentions in early 2024, followed by a steady rise from March 2024.

- The mentions peaked significantly in September 2024, potentially indicating increased attention due to notable events or announcements.

- Post-peak, a sharp drop is observed in November 2024, possibly reflecting a shift in focus or reduced reporting on startups.

Key Industry Trends in Startup Ecosystems

This analysis highlights the most-discussed industries in startups, helping you see where innovation is happening. FinTech leads due to advances in digital payments and blockchain, Healthcare grows with telemedicine and biotech, and AI remains a transformative force across industries. This gives a clear picture of where trends are shaping the future.



DOMINANCE OF FINTECH:

FinTech leads the startup discussions, highlighting its critical role in the innovation ecosystem.

This dominance could be attributed to the rapid adoption of digital wallets, blockchain technology, and cryptocurrency, which continue to reshape traditional financial services.

EMERGING HEALTHCARE SECTOR:

Healthcare is a significant contributor, driven by advancements in biotech, telemedicine, and health tech.

The emphasis on healthcare startups reflects increased demand for innovation in medical technology, particularly post-pandemic.

CONSISTENT INTEREST IN AI:

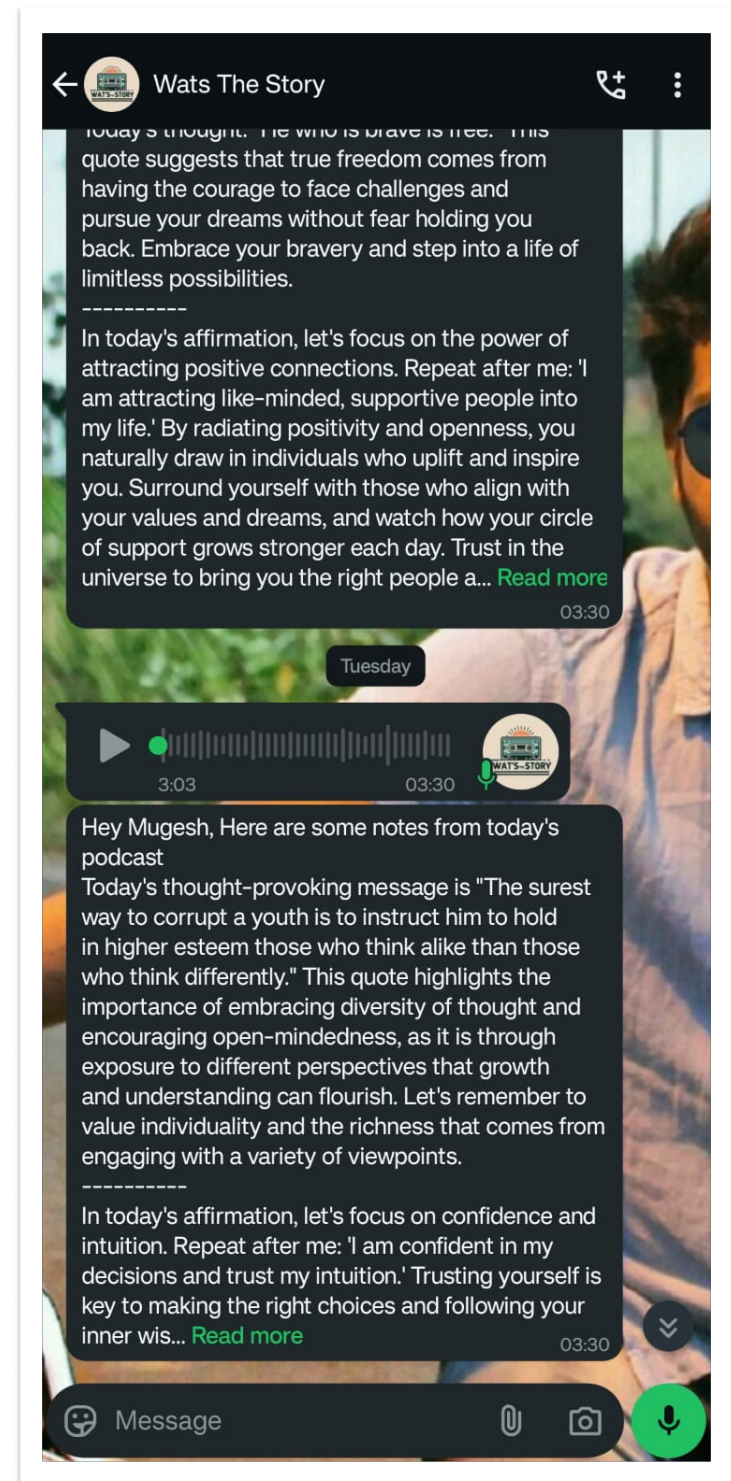
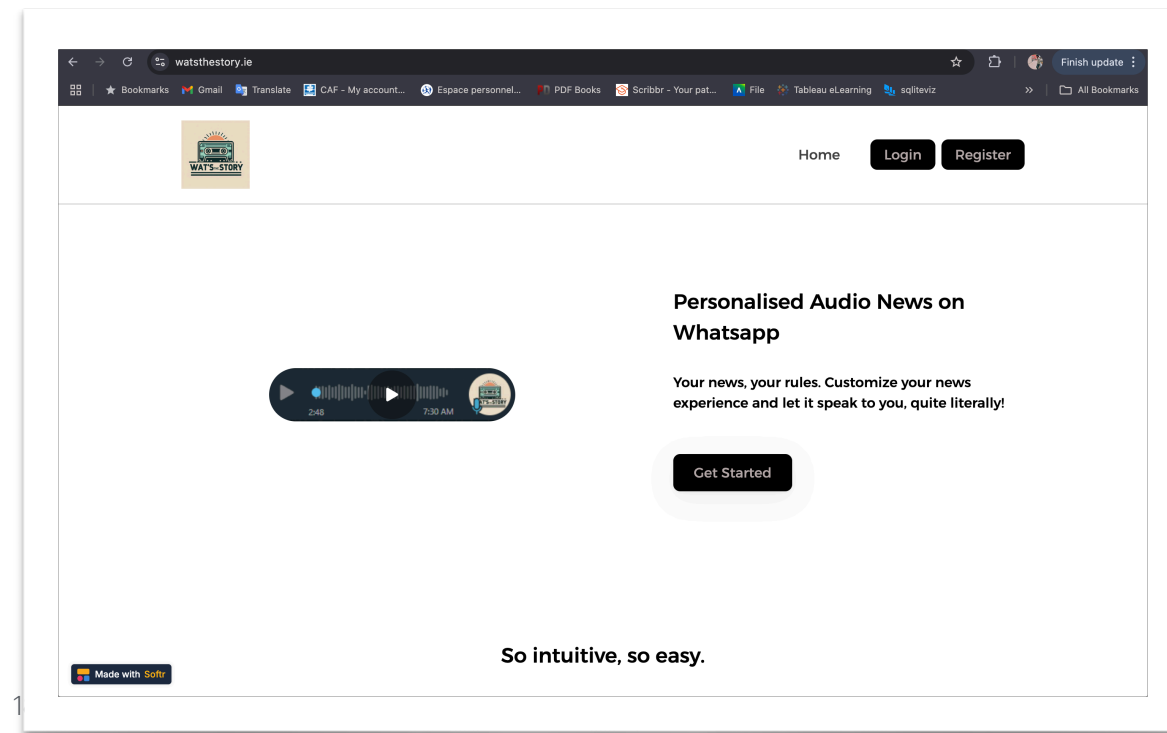
AI maintains a strong presence in startup discussions, with its applications spanning multiple industries, including FinTech, healthcare, and logistics.

The consistent interest underscores its transformative potential in automation, machine learning, and decision-making.

To showcase the user-friendly and intuitive experience of Wat'sTheStory, screenshots of its website and WhatsApp integration have been added. These visuals highlight the application's unique ability to deliver personalized audio news and concise updates directly to your mobile device.

I encourage everyone to try this amazing application to stay updated effortlessly. Experience news in a new, more accessible way—anytime, anywhere. Let Wat'sTheStory revolutionize how you consume information!

Check out this page link - <https://www.watsthestory.ie/>



Final Conclusion:

This project marked an initial exploration of the Wat'sTheStory dataset, analyzing trends across cybersecurity, technology, stock markets, and Irish news. Key findings include:

- Word Cloud Analysis:
- Cybersecurity focused on vulnerabilities and threats.
- Tech news highlighted AI advancements and startups.
- Stock market trends reflected financial growth and policy shifts.
- Irish news balanced local and global socio-political discussions.
- Emerging Startups: A significant spike in FinTech and AI-related startups in September 2024 reflected innovation trends.
- Sentiment Analysis:
- Positive sentiment dominated tech news, driven by optimism about innovation.
- Cybersecurity leaned negative, highlighting ongoing concerns.

This project is a starting point for deeper exploration. Future work will leverage real-time data, advanced NLP models, and thematic exploration to uncover more actionable insights, aligning with Wat'sTheStory's mission to deliver impactful and noise-free news.

THANK YOU FOR THE OPPORTUNITY TO CONTRIBUTE TO THIS INCREDIBLE PROJECT AND UNCOVER VALUABLE INSIGHTS THAT REFLECT



WAT'S THE STORY'S INNOVATIVE VISION.