

# Exploring the Digital Landscape of STEAM Learning Discourse: Insights from GPT-4 based Twitter Data Analysis

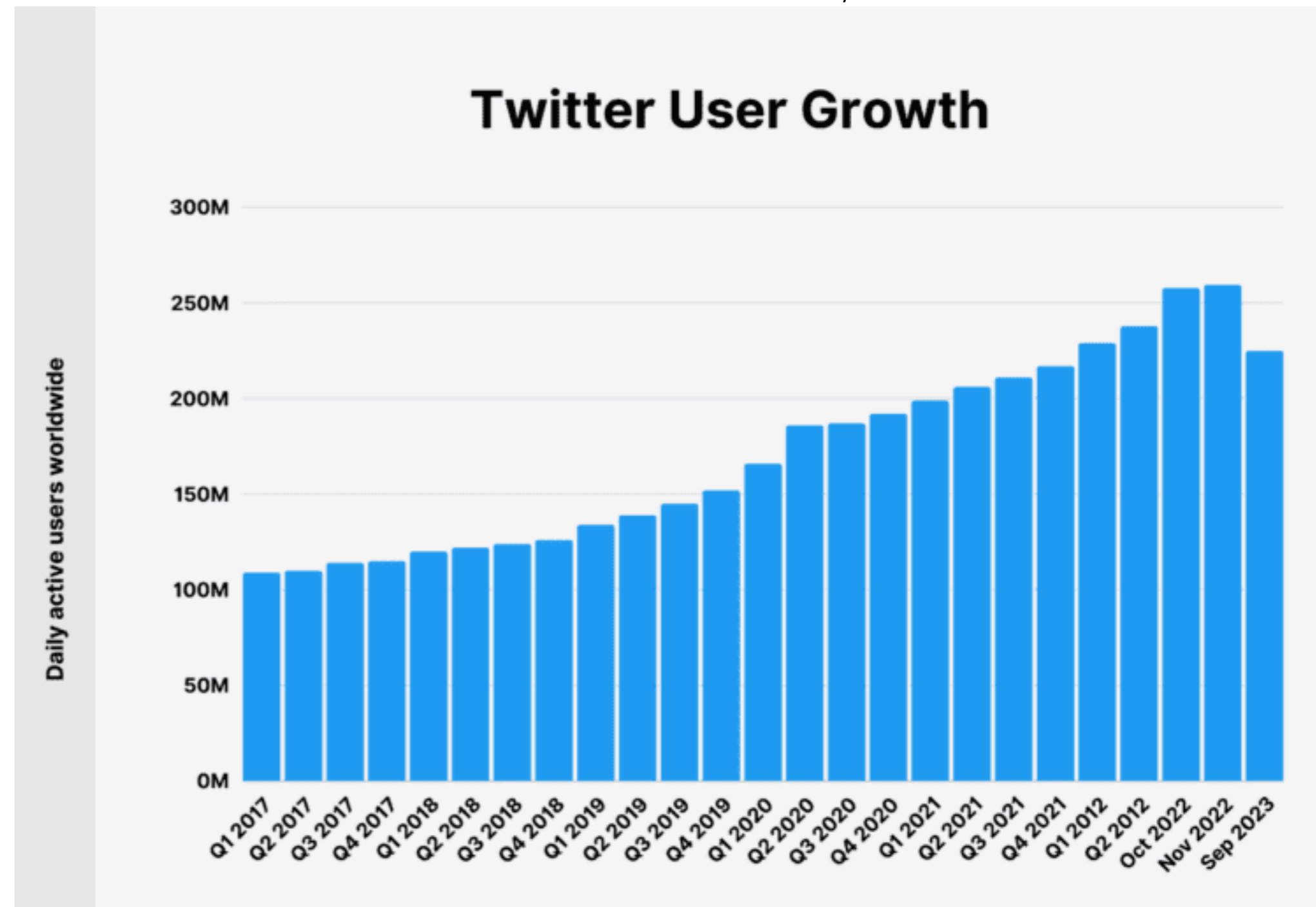
Sherif Abdelhamid, Elijah Bass

---

# Introduction

Each day witnesses the dispatch of millions of tweets, encapsulating vast quantities of information across diverse subjects.

Source: Twitter/ X

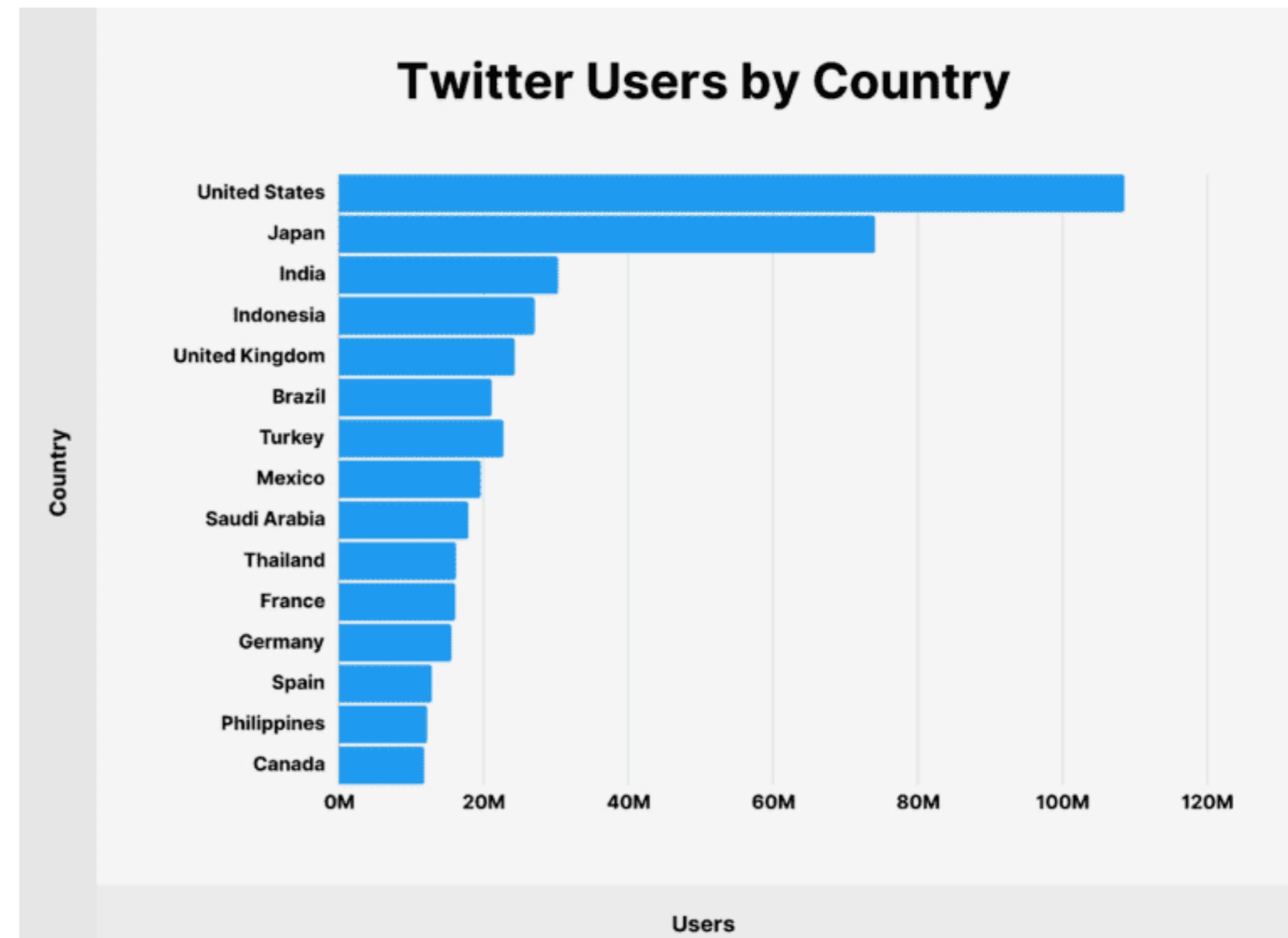


---

# Introduction

The largest number of users happens to be from their country of origin (USA). However, Twitter is quickly penetrating international markets

Source: DataReportal



X/Twitter is predominantly used for getting the news. Other commonly reported activities on the platform include researching brands, looking for entertaining content

# Introduction

Source: DataReportal

Activity	Share of respondents
Keep up to date with news and current events	60.6%
Follow or research brands and products	34.9%
Look for funny or entertaining content	34.8%
Post or share photos or videos	27%
Message friends and family	19.2%



---

# Twitter for STEAM

A significant portion of the tweets pertains to the domain of (Science, Technology, Engineering, Arts, and Mathematics) STEAM education.

Originating from various sources such as **students, academic institutions, governmental bodies, and policymakers.**

The tweets' content includes updates on **daily occurrences, pivotal announcements, educational materials, discussions on pertinent topics, geographical data, and user interaction** through retweets, comments, and likes.



# Problem

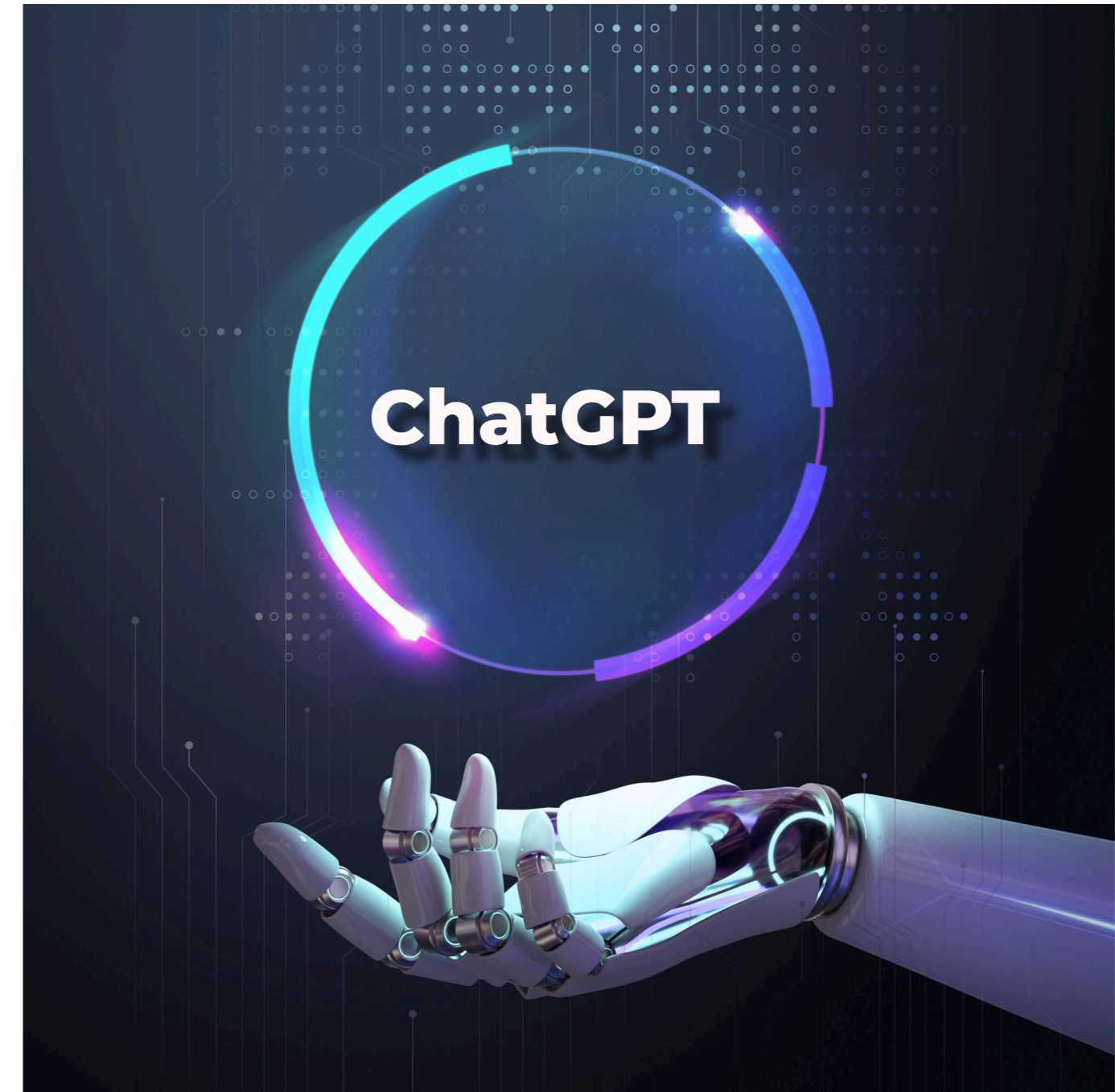
The high complexity, unstructured format, and large volume of this data often pose significant challenges for researchers seeking to extract meaningful insights using qualitative or quantitative approaches.





## Solution

Adapt the Generative Pre-trained Transformer 4 (GPT-4), an advanced multimodal large language model (LLM), to analyze tweet data.



## Why GPT-4

- Advanced natural language processing capabilities.
- Ability to infer meaning from limited text makes it ideal for analyzing such concise tweets.
- Semantic analysis capabilities, identifying themes, topics, and sentiments within tweets.
- Trained on a diverse large text corpus, which includes content from various cultures, enabling it to understand and analyze tweets from a wide range of global users.

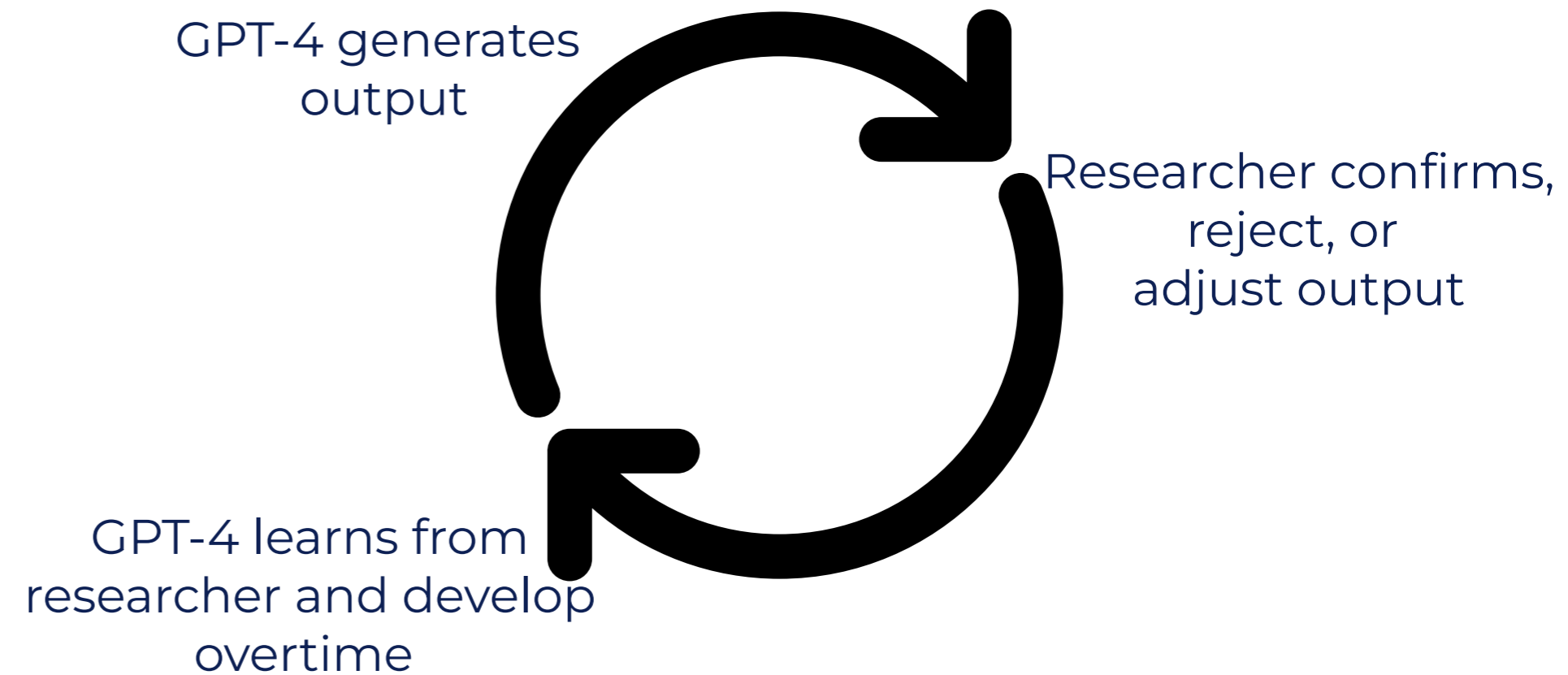




---

## Methodology

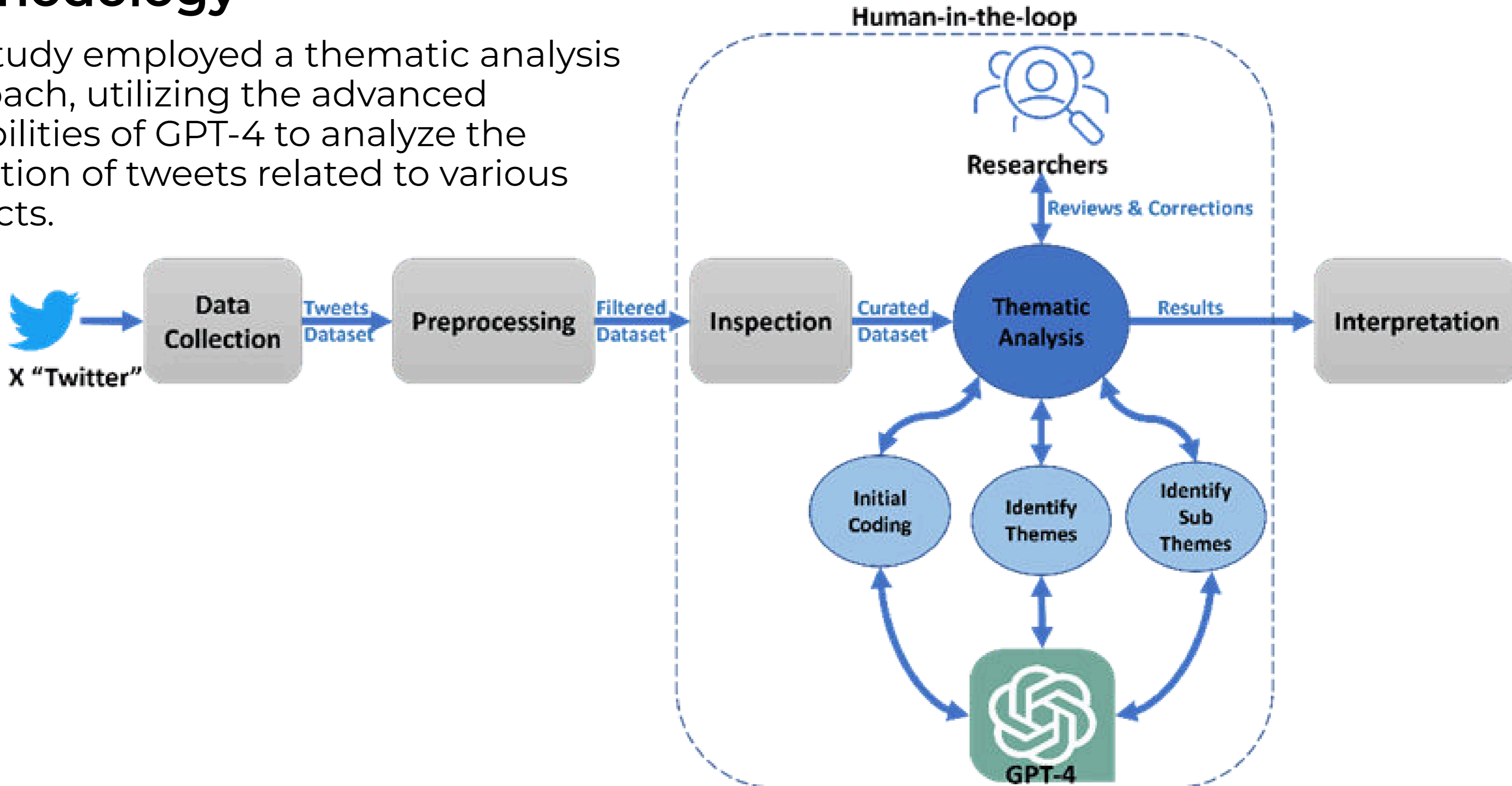
The study employed a thematic analysis approach, utilizing the advanced capabilities of GPT-4 to analyze the collection of tweets related to various subjects.



Our methodology follows the human-machine loop approach, where researchers and ChatGPT mutually augment each other. This type of is referred to as Artificial Augmented Intelligence.

# Methodology

The study employed a thematic analysis approach, utilizing the advanced capabilities of GPT-4 to analyze the collection of tweets related to various subjects.



---

# Results and Findings

## Search by Subject

The data gathering yielded datasets of 22,337 tweets.

The primary objective of this search was to find tweets pertinent to a wide range of STEAM topics.

The process may capture tweets that do not align with the search criteria, known as false positives.

To maintain the integrity of the study, these tweets are reviewed and filtered out to ensure the dataset's relevance and accuracy

Subject	Number of Tweets	Hashtag
Arts Education	1002	#ArtsEd
Computer Science	1225	#CompSci
Engineering	563	#EngChat
Language	4070	#LangChat
Literacy	2414	#Literacy
Physical Education	5765	#PhysEd
Science	1558	#SciChat
Social Science	5740	#SSchat



## #LangChat Key Themes

- **The diverse use of digital platforms** like Anki, Quizlet, and Kahoot for interactive learning activities highlights the significant role of technology in enhancing language education.
- **Professional networking** emerged as a central theme, with tweets emphasizing the value of conferences and professional learning networks in fostering knowledge exchange and continuing education.
- **Advocacy and leadership within language education** were underscored, reflecting a commitment to advancing the field and supporting educational programs.

---

## Results and Findings



# #CompSci Key Themes

- **Technological Advancements**, highlighting discussions on AI, machine learning, and new tech developments.
- The second most prevalent theme was the **importance of STEAM education**, underscoring the **significance of computer science in educational contexts**, teaching methodologies, and learning experiences.
- **Educational Achievements** were the third discussed topic, showcasing individual projects and the application of specific technologies like NumPy and Python.
- While **Resources and Events** had 140 mentions, pointing to sharing educational platforms, conferences, and resources.
- The **Personal Experiences** and Humor theme appeared in some tweets.

---

## Results and Findings





## #Literacy Key Themes

- "**Reading Practices and Engagement**" is the most prominent theme with occurrences in most tweets.
- The "**Educational Contexts of Literacy**" theme, evident in 511 tweets, underscored discussions around integrating technology, digital literacy, and educator strategies within literacy education.
- Less prevalent but still significant, the "**Literacy Promotion and Advocacy**" theme appeared in 239 tweets, reflecting efforts in advocacy, support for literacy initiatives, and the importance of charitable contributions to literacy.

---

## Results and Findings





## #ArtsEd Key Themes

- **Art Appreciation and Education** emerged as a prominent theme, with tweets highlighting individual artists, artworks, and their historical and cultural significance.
- **Educational Resources and Activities** were also highly represented, focusing on practical resources, teaching strategies, and creative projects to enhance arts education.
- **Research, Policy, and Advocacy** themes highlighted the importance of supporting arts education through research findings, policy discussions, and funding opportunities.

---

## Results and Findings



# #PhysEd Key Themes

- The most dominant theme was "**Educational Activities**," with 1655 tweets emphasizing the diversity of physical activities, sports, and educational games.
- Following this, "**Engagement and Innovation**" and "**Health and Well-being**" themes were prominent, indicating a strong emphasis on engaging students with innovative teaching methods.
- "**Professional Development**" emerged in 441 tweets, highlighting the value placed on educators' growth and knowledge sharing.
- "**Resources and Tools**" were discussed in 220 tweets, pointing to using various educational resources and tools.
- The "**Inclusivity and Accessibility**" theme, though less prevalent, underscored the importance of making physical education accessible and inclusive for all students

---

## Results and Findings





## #SciChat Key Themes

- The thematic analysis of the science tweet dataset revealed encompasses themes such as:
- **"Technology and Innovation," "Education and Outreach," "Historical Notes," "Astronomy and Space," "Environmental Insights," "Scientific Research," and "Wildlife."**
- The analysis underscored a strong focus on **environmental concerns, conservation efforts,** and the implications of climate change.
- The dataset also reflected a keen interest in wildlife conservation, genetic research, and ecological studies, showcasing science-related conversations' wide-ranging and multi-faceted nature.

---

## Results and Findings





# Results and Findings

## Search by Education Level

In this part of the study, we used GPT-4 to conduct a thematic analysis of tweets related to higher education, high school, and middle school.

The analysis revealed distinct priorities, concerns, and discussions pertinent to each education level.

Education Level	Main Themes	Number of Tweets	Hashtag
Higher Education	Career and Professional Development, Educational Trends and Insights, Technology and Privacy in Education, Academic Research and Policy, Security and Financial Aid	710	#HigherEd
High School	Educational Opportunities and Achievements, Athletic and Extracurricular Activities, Career and College Readiness, Personal Milestones and Celebrations, Support and Community, Technology and Online Engagement, Social and Political Issues	602	#HighSchool
Middle School	General Education and School Life, Libraries and Librarians, Teacher Identity and Merchandise, Specific School Programs, <a href="#">College</a> and Sports	4106	#MiddleSchool

---

## Conclusion

In this study, we utilized GPT-4 to analyze Twitter discussions related to STEAM education in various subjects and at different education levels.

We identified key themes and sub-themes within each subject.

The human-in-the-loop approach integrated the analytical strengths of GPT-4 with human cognitive abilities for more reliable and relevant results.

We will explore other social networking platforms like LinkedIn or Reddit for future work. Additionally, we will use the growing capabilities of GPT-4 with the multimodal nature of Twitter to incorporate an analysis of visual content such as infographics, videos, and images.





---

# Thanks!

*Do you have any questions?*

abdelhamidse@vmi.edu

---

