

How Does a University's Computer Science Strength and Location Impact Its Total ChatGPT News?

Eileen Zhan¹, Sumith Gunasekera², Yong Xu^{2*}

¹Forest Hills Central High School, Grand Rapids, USA

²Marketing Department, Ferris State University, Big Rapids, USA

Email: kideileenzhan@gmail.com, sumigunasekera@ferris.edu, *xuy1@ferris.edu

How to cite this paper: Zhan, E., Gunasekera, S. and Xu, Y. (2024) How Does a University's Computer Science Strength and Location Impact Its Total ChatGPT News? *Open Journal of Applied Sciences*, 14, 2973-2984.

<https://doi.org/10.4236/ojapps.2024.1410194>

Received: August 10, 2024

Accepted: October 27, 2024

Published: October 30, 2024

Copyright © 2024 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

As AI, starting with ChatGPT has become increasingly prevalent in academic discussions, school especially, colleges have become hotspots of AI activities and debates. Colleges have the responsibility of addressing not only the academic, integrity-based concerns of students using AI for their homework, but also as the forebearers of new learning and technology, how AI will change their students' futures and careers. In this study, we will explore the different factors, such as Computer Science Score and location, that might affect how much a college discusses AI, ChatGPT specifically. To demonstrate the validity of our research, we used self-collected data with our methods detailed below.

Keywords

AI, ChatGPT, Computer Science, Statistical Analysis, Location Impact

1. Introduction

ChatGPT and other AI-assisted chatbots (computer programs that simulate human conversation with an end user) like it represent a major recent technological leap. Widely regarded as a historical breakthrough in AI, ChatGPT has seized the attention of both the public and academic communities. Like other fields, studies, discussions, research, articles, and even policies about this technology have exploded at colleges and universities across the country since the chatbot's launch on Nov 30, 2022.

This project seeks to evaluate the influence of ChatGPT on universities by

analyzing each university's total number of articles mentioning ChatGPT on the Google News platform published during the last complete calendar year (2023). News reports of recent events [1] while Google News is the world's largest news aggregator, therefore the count of ChatGPT news connected to each university on the Google News platform reflect the impact of ChatGPT on each campus, including ChatGPT reports about, by, or for the university or personnel (faculty, students, and alumni) with the university's name in the title. Furthermore, compared to general information, forums, videos, Facebook, and other media data, Google news data is approachable, collectable, and diggable.

Since ChatGPT is an IT breakthrough started in Silicon Valley, the authors were curious whether universities with more robust CS programs and those closer to Silicon Valley are more likely to be involved in ChatGPT-related activities. Next, the authors will explore how the strength of CS programs and the location of colleges—whether in the Midwest, South, Northeast, West, and Southwest—affects the number of news articles related to ChatGPT. The authors believe this will help us understand the influence of college region and CS program strength on ChatGPT-related activity.

2. Literature Review

In their study [2] “Chatting about ChatGPT: How may AI and GPT impact academia and libraries”, Lund & Wang (2023) explore the influence of AI technologies on academic institutions and libraries. They emphasize the potential of these technologies to revolutionize information retrieval, academic research, and educational methods. Libraries are poised to benefit from more efficient data management and increased user engagement by using AI-driven tools. The study also discusses the ethical and practical challenges associated with the integration of AI, highlighting concerns about data privacy, algorithmic bias, and the need for digital literacy among users.

Dempere *et al.* (2023) [3] examine the implications of ChatGPT on higher education, presenting a nuanced view of its potential and limitations. The study highlights the practical applications of ChatGPT in administrative support and academic assistance but also underscores the challenges, such as accuracy and ethical considerations.

Firat (2023) [4] examines the implications of ChatGPT on higher education, presenting a nuanced view of its potential and limitations. The study highlights the practical applications of ChatGPT in administrative support and academic assistance but also underscores the challenges, such as accuracy and ethical considerations.

Sahadevan (2023) [5] examines the impact and utilization of ChatGPT among college students, emphasizing the tool's effects on learning motivation, productivity, and mental health support. The varied findings across studies illustrate both the benefits and challenges of integrating AI-driven technologies like ChatGPT in academic settings.

Fabella (2023) [6] explores first-year college students' attitudes towards ChatGPT, highlighting mixed sentiments. Students recognize the tool's potential to enhance productivity and learning but express concerns about its misuse, ethical implications, and accuracy. The study suggests integrating AI ethics discussions into educational curriculums to guide responsible AI use.

Scimeca *et al.* (2023) [7] investigate AI's impact on education, noting that while AI can support learning with instant feedback and personalized assistance, it raises concerns about plagiarism and academic integrity. They advocate for robust AI literacy programs to help students use AI ethically.

Qi *et al.* (2023) [8] focuses on improving academic search engines through SLaTe, a large-scale benchmark for scholarly literature search, emphasizing the importance of high-quality, domain-specific data in enhancing AI's effectiveness in academic research.

Research by Akgun *et al.* (2023) [9] addresses the ethical implications of AI in educational contexts. This study provides a framework for understanding the potential biases in AI models and emphasizes the need in their deployment. The ethical and societal drawbacks of these systems are rarely fully considered in K-12 educational contexts. They discuss the ethical challenges and dilemmas of using AI in education.

Roe and Perkins (2023) [10] analyze UK news media headlines, revealing a paradoxical portrayal of AI that oscillates between promising societal solutions and cautioning against systemic risks. This study underscores the media's role in shaping public perceptions and calls for a deeper understanding of the social, cultural, and political contexts influencing AI representation.

Lastly, the Bio-Conferences article (2024) [11] examines AI's role in medical and healthcare education. It discusses AI tools like ChatGPT in medical training, emphasizing their potential to provide real-time feedback and support decision-making while addressing challenges related to data privacy, accuracy, and ethics. Collectively, these studies highlight AI's transformative potential across various fields and the critical need for balanced, informed approaches to its integration and use. They emphasize the dual nature of AI's promise and peril, the importance of responsible use, and the ongoing research and education required to navigate the ethical and practical challenges posed by AI technologies.

3. Methodology

To create a comprehensive dataset for this analysis, the authors gathered information on the top 116 universities listed on the U.S. News website [12], which are rated annually based on their computer science (CS) programs. Each university's CS program is assigned a score on a scale from 1 (marginal) to 5 (outstanding), based on a survey of academics at peer institutions [13]. After excluding three graduate-only programs (Toyota Technological Institute at Chicago, Naval Postgraduate School, and The City University of New York Graduate School and University Center) from further discussion, the authors collected the CS scores and

geographic locations (states and regions) for the remaining 113 universities and recorded them in an Excel file. They then searched and recorded the longitude and latitude of each university. The authors also included whether schools were public or private and if they were rural, urban or suburban. For example, Harvard University is a private school in an urban area.

Next, the authors collected the ChatGPT news data for the year 2023 from each university using the following steps:

Step 1: Search “ChatGPT” + “(university names)” + “2023” on Google News.

Step 2: Use the “Tools” option to “Sort by date”.

Step 3: Count the number of relevant news articles and remove any that are outside the criteria.

For example, when searching for Harvard University, the authors initially found six pages of results, including 10 records on the first five pages and five on the sixth. However, the authors also found one news article from January 2024 and another from December 2022, which should not be counted towards the year 2023. After removing these two articles, the total count for Harvard in 2023 was $5 \times 10 + 5 - 2 = 53$.

To account for alternate university names and avoid duplicate results, the authors conducted searches using all known names. For instance, the Georgia Institute of Technology, also known as Georgia Tech, yielded 35 and 40 news items respectively, with 11 duplicates subsequently identified and removed. Similarly, the University of South Carolina, alternatively known as South Carolina University, and the California Institute of Technology, also known as Caltech. A search for Dartmouth College returned 35 news items, while Dartmouth University showed an additional five articles. For universities with multiple branches, the authors used various search terms, such as “University of Texas San Antonio”, “University of Texas at San Antonio”, and “UT San Antonio”, adding them together to capture all relevant news. To ensure quality, the authors manually screened the results for relevance. To do this, they clicked links to articles to verify the content and determine if it should be included or not.

4. Results

First, the authors use a Box-and-Whisker Plot to compare total news by region. **Figure 1** below, shows that the regional differences in central tendency are minor compared with others. The Northeast and West regions show a higher variability in Total News compared to the Midwest and South, while the Midwest has a notably smaller variability. The low variability in the Midwest could be explained by the more spread-out locations of those universities, leaving local news outlets less choice for which universities to approach for a story. Illinois Institute of Technology (IIT) is the outlier in the Midwest. A possible reason for this may be IIT’s smaller student population (3125 students in 2024) and proximity to the University of Chicago and Northwestern University.

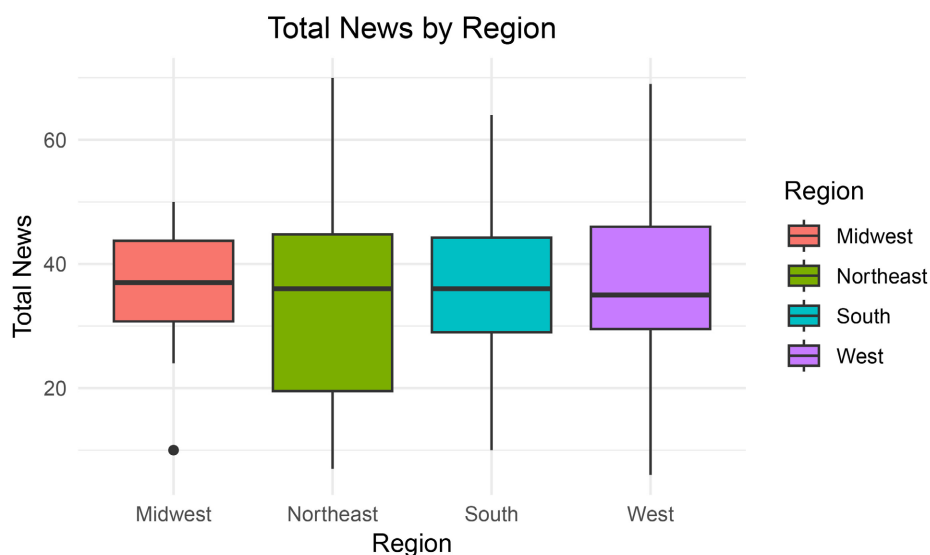


Figure 1. Box-Whisker plot of total news by region.

To further determine if the stated variables are a significant factor in influencing Total News, the authors performed an analysis of variance (ANOVA) test for each of the variables independently and found the following data presented in **Table 1** below.

Table 1. ANOVA summary.

Test	Sum Sq	H ₀ .	F	P	Conclusion
Region	317.54	No difference in Total News across Rural/Urban categories	19.36	3.50e-10	Reject H ₀ , significant difference
S Score	10955.72	CS Score does not affect Total News	111.87	1.243 × 10 ¹⁸	Reject H ₀ , significant difference
Social Mobility	68.87	No relationship between SMS and Total News	0.35	0.5527	Fail to reject H ₀ , no significant relationship
Overall Ranking (Categories)	5915.27	No difference in Total News across Ranking categories	47.52	1.05e-15	Reject H ₀ , significant difference
Public/Private	527.39	No significant difference for Public vs Private colleges	2.77	9.86e-02	Reject H ₀ , significant difference
Rural/Urban	7584.94	No, significant difference for Rural vs Urban colleges	19.36	3.5e-10	Reject H ₀ , significant difference

Now, knowing all these things, the authors wanted to see if the different factors in an analysis would give us different results. The model the authors used for analyzing the influence of Total News on CS Score and Region can be written in Equation (1).

$$\text{Total_News} = \beta_0 + \beta_1 \times \text{CS_Score} + \beta_2 \times \text{Region} + \beta_3 \times \text{Rural/Urban} + \beta_4 \times \text{Public/Private} + \beta_5 \times \text{Ranking_Cat} + \beta_6 \times \text{SMS} \quad (1)$$

The authors used R software to run the regression analysis. An R-squared value of 0.495 indicates that above regression model explains approximately 49.5% of the variability in Total News. The regression results are summarized in **Table 2** below.

Table 2. Regression analysis results for total news vs CS score and region.

Variable	Coefficient	Standard Error	t-Value	p-Value	Null Hypothesis (H ₀)	Conclusion
Intercept	-6.7515	4.0574	-1.6640	0.0989	$\beta_0 = 0$	
CS_Score	12.9604	1.2253	10.5770	1.24e-18	$B_1 = 0$	Reject H ₀
Region	0.114	0.91	0.125	0.901	$B_2 = 0$	Fail to reject H ₀
Suburban	-2.6379	7.7458	-0.339	0.7351	$B_3 = 0$	Fail to reject H ₀
Public/Private	-3.2848	2.4324	-1.35	0.1798	$B_4 = 0$	Fail to reject H ₀
SMS	-10.8908	20.5718	-0.529	0.5977	$B_5 = 0$	Fail to reject H ₀

The test results for CS Score from **Table 2** indicate a significant positive relationship between CS Score and Total News. For each unit increase in CS Score, Total News increases by approximately 12.96 articles.

Regarding regional effects, the p-values for region (the baseline Region AKA the Midwest), Northeast, South, and West are ≥ 0.05 , which indicates the coefficient of Midwest and the coefficient of region differences between the Midwest and other regions (Northeast, South, and West) are all not statistically significant. Therefore, the region is not a significant predictor of Total News even when considering CS Score. We can see the pattern from **Figure 2** below, of total News versus CS score with region information.

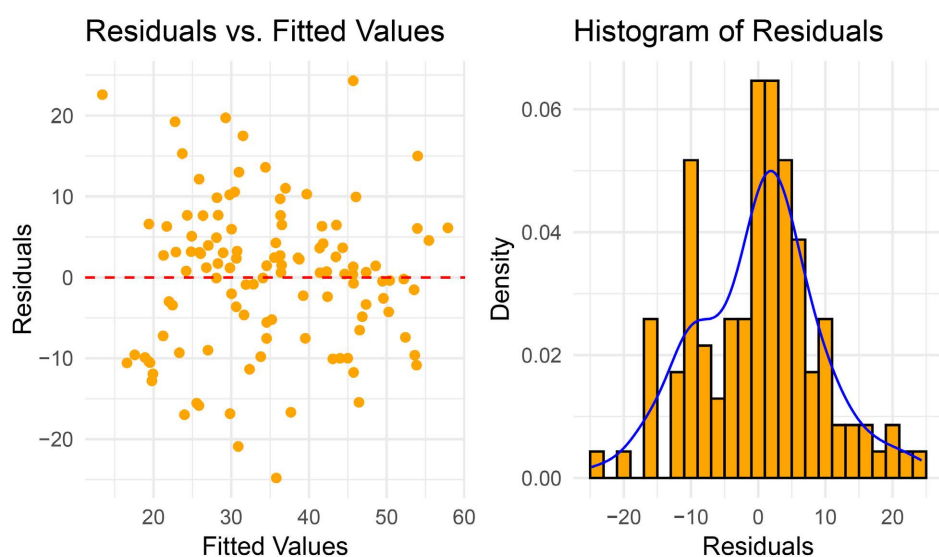


Figure 2. Plots to check homoscedasticity, linearity, and normality.

The plot of Residuals vs. Fitted Values (Figure 2) above, shows no obvious pattern, indicating that the residuals have constant variance and that the relationship between the independent variables and Total News is reasonably linear. In addition, the Histogram of Residuals shows that the residuals are approximately normally distributed; therefore, the assumptions of homoscedasticity, linearity, and normality required for linear regression seem to be reasonably met.

Figure 3 shows a clear positive correlation between CS Score and Total News, and Figure 4 shows a correlation between higher SMS category and total news.

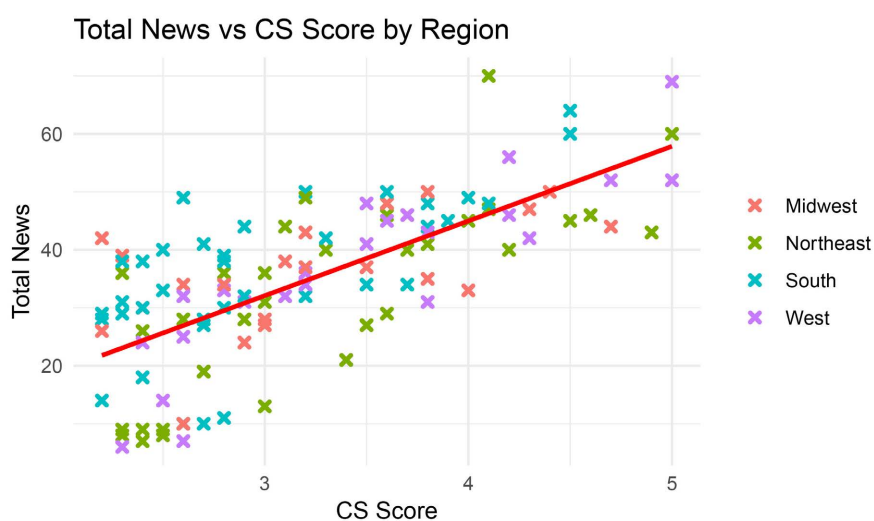


Figure 3. Graph of total news vs CS score with region.

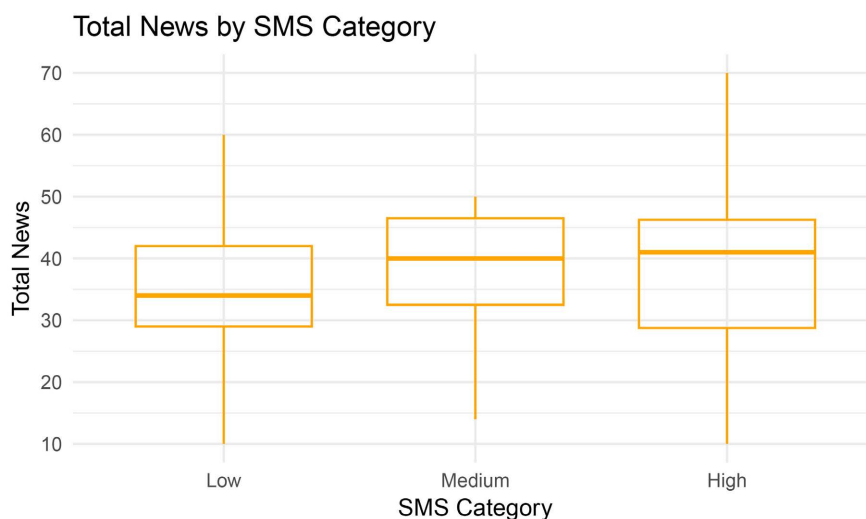


Figure 4. Total news vs SMS.

Figure 5 shows a strong positive correlation between ranking and Total News. Figure 6 shows a difference between rural and urban collages, with suburban in the middle.

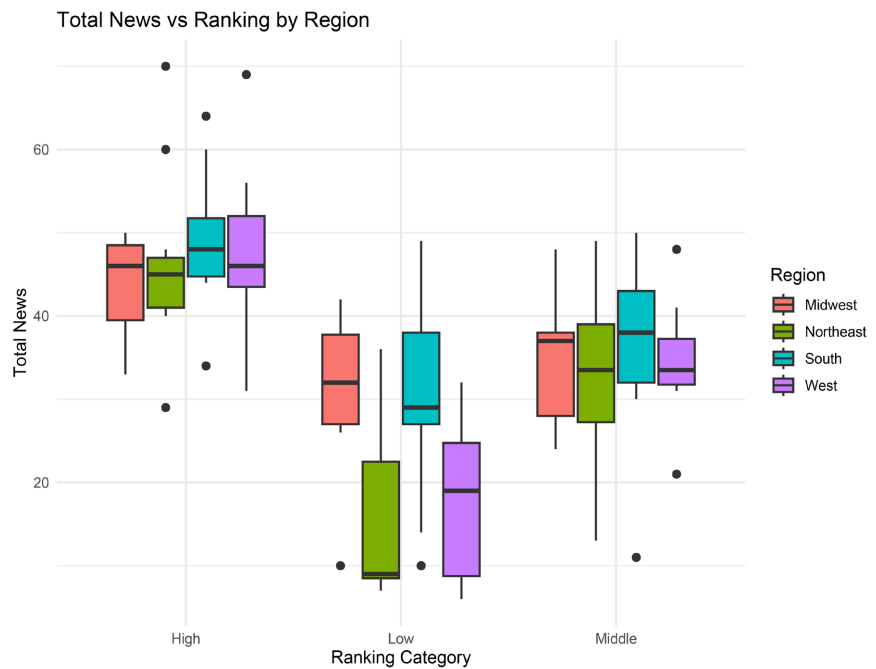


Figure 5. Total news vs ranking by region.

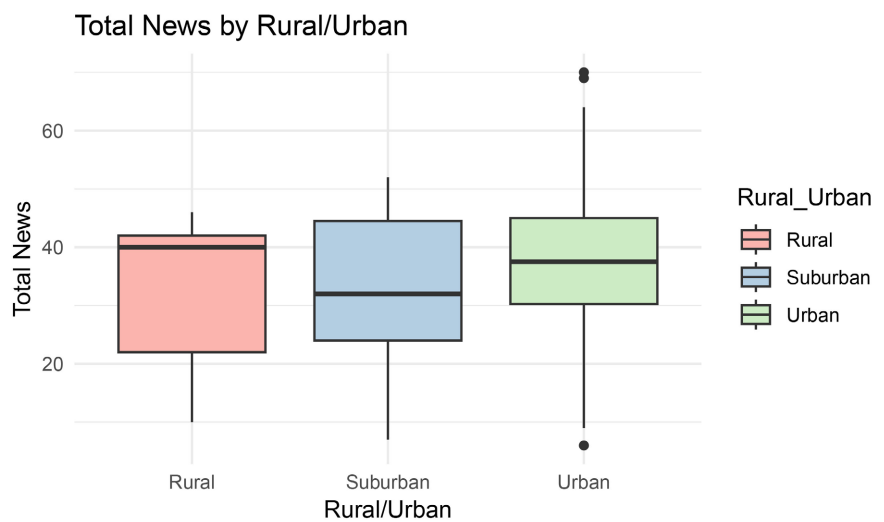


Figure 6. Total news by Rural/Urban.

According to **Table 3** below, all Slope p-values are less than or equal to 0.05, which indicates that the corresponding slope coefficients are significant. From the plots and regressions by region, the authors notice that there is a positive correlation between CS Score and Total News across all regions. Universities with higher CS Scores tend to have higher numbers of ChatGPT-related news.

However, the strength of this relationship varies by region. The West, Northeast, and South, in that order, show more positive trends and have stronger significance compared to the Midwest. The West shows the strongest relationship between CS Score and Total News, with the highest R-squared value of 0.73 and

most statistically significant slope and intercept.

Table 3. Regression analysis results for total news vs CS score for different regions.

Region	Equation	Slope p-value	R-squared
Northeast	Total News = $15.79 \times \text{CS Score} + -19.38$	$1.60e-07$	0.58
West	Total News = $16.09 \times \text{CS Score} + -19.50$	$1.19e-07$	0.73
Midwest	Total News = $7.70 \times \text{CS Score} + 11.50$	$5.41e-03$	0.33
South	Total News = $13.13 \times \text{CS Score} + -3.57$	$4.99e-06$	0.46

On the other hand, the Midwest has the weakest relationship, showing the most variability with a lower R-squared value of 0.33. This indicates that factors other than CS Score might more strongly influence Total News in this region. Next, the authors used a heat map to visually detail the distribution of universities throughout the country and their corresponding number of ChatGPT news articles. When some universities are too close to be distinguished on the map, the average total news from these universities is calculated.

From **Figure 7** heat map, below, the authors noticed that major metropolitan areas and regions known for their educational institutions generally have higher news coverage. Higher concentrations of top-scoring universities are seen in the metropolitan areas of the Northeast corridor, West Coast, and parts of the Midwest and South. However, there is noticeable variability within these larger regions.



Figure 7. Heat map of total news in top 119 universities.

For example, while the Northeast overall has high coverage, some schools with lower coverage are scattered throughout the region. This is consistent with the larger variability in total ChatGPT news in the Northeast as seen in **Figure 5**. Colleges with strong computer science programs in the central region are more geographically dispersed, making it easier for them to make the local news. In

contrast, in the Northeastern region, where colleges with strong computer science programs are more concentrated, universities with slightly weaker programs and fewer students may have relatively less access to resources and less impact on news coverage.

5. Conclusions

5.1. Summary

This research investigates how the strength of Computer Science (CS) programs and the geographic location of universities in the U.S. affect the number of news articles that mention ChatGPT alongside the institution. Analyzing Google News data from 2023 for 113 universities, it was found that universities with stronger CS programs tend to appear in more ChatGPT-related news. Although geographic region was also studied, its impact was less significant. Statistical analysis confirmed that the strength of the CS program is a key predictor, while location has a smaller effect.

5.2. Discussions

CS Score can be used to predict ChatGPT news presence effectively. The consistent, strong positive pattern across regions demonstrates the importance of academic strength in computer science for ChatGPT-related media visibility. Universities in the Northeast are more influenced by the CS Program rankings due to their geographic proximity and experience more news exposure and competition for resources. Conversely, the Midwest shows the weakest relationship, suggesting that other factors might play a more significant role in influencing news coverage there.

Even though there is no statistically significant difference among the Northeast, Midwest, South, and West, notable regional differences exist in news coverage variability, with the Northeast and West showing higher variability compared to the Midwest and South. The Midwest's low variability is likely due to geographically dispersed universities accessing local media and resources more effectively. Notably, major metropolitan areas generally have higher news coverage, but variability persists within regions, particularly in the Northeast.

5.3. Limitations

1) Data Source Limitations: The reliance on Google News as the sole data source could lead to biased results, as it may not fully capture all relevant media coverage or reflect broader public discourse. Including additional data sources, such as social media mentions, academic publications, or regional news outlets, could provide a more comprehensive view. While Google News is a comprehensive platform, it may not capture all relevant publications, especially those behind paywalls or in less accessible databases. It will be more ideal to capture the more board spectrum of media coverage or accurately represent the broader public discourse. Additionally, the study focused on a specific timeframe (2023) and only the top

116 CS universities are included in the study. Since this paper focuses only on U.S. universities, it limits the generalizability of the findings to other countries or educational systems.

2) Regional Analysis: While the study considers geographic location, the analysis of regional differences is somewhat superficial. A more detailed examination of local factors, such as differences in media ecosystems, cultural attitudes toward AI, or the presence of tech hubs, could enrich the findings. The limitation comes from the relatively modest impact of geographic location on the results, which may be due to an oversimplification of regional differences or an insufficient consideration of other factors such as local media environments or cultural attitudes toward technology.

3) Statistical Methods: The researchers employ regression analysis to determine the variables, but it does not fully explore potential confounding factors or interactions between variables. Including a more sophisticated statistical approach, such as multilevel modeling, could provide deeper insights into the complexities of the data.

5.4. Future Research

Researchers could incorporate multiple news aggregators and academic databases to obtain a more complete picture. Sample size and time length can further be extended to achieve more comprehensive results. The authors can also look for more statistically significant factors in the regression model. Furthermore, the authors can further investigate the news content, categorize it, and study the content focus and changes over time, which would improve our understanding.

5.5. Implications and Applications

The findings of this study have important implications for universities, policymakers, and AI researchers. Understanding the regional factors that influence media visibility can help universities tailor their strategies to enhance their research output and public engagement. Policymakers can use these insights to allocate resources more effectively and support regional centers of excellence in AI research.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Merriam-Webster (n.d.) News. Merriam-Webster Dictionary. <https://www.merriam-webster.com/dictionary/news>
- [2] Lund, B.D. and Wang, T. (2023) Chatting about ChatGPT: How May AI and GPT Impact Academia and Libraries. *Minds and Machines*, **30**, 681-694. <https://doi.org/10.1108/LHTN-01-2023-0009>
- [3] Dempere, J., Modugu, K., Hesham, A. and Ramasamy, L.K. (2023) The Impact of Chatgpt on Higher Education. *Frontiers in Education*, **8**. <https://doi.org/10.3389/educ.2023.1206936>

- [4] First, M. (2023) What ChatGPT Means for Universities: Perceptions of Scholars and Students. *Journal of Applied Learning & Teaching*, **6**, 1-22.
<https://doi.org/10.37074/jalt.2023.6.1.22>
- [5] Sahadevan, S. (2024) An Analysis on Impact and Utilization of ChatGPT among College Students. *Journal of Emerging Technologies and Innovative Research*, **11**, b1-b7.
https://www.researchgate.net/publication/379824157_AN_ANALYSIS_ON_IMPACT_AND_UTILISATION_OF_CHATGPT_AMONG_COLLEGE_STUDENTS
- [6] Fabella, F.E. (2023) Attitudes toward the Positive and Negative Features of ChatGPT by Selected First-Year College Students. *International Research Journal of Modernization in Engineering Technology and Science*, **5**, 1632-1639.
<https://doi.org/10.56726/IRJMETs45426>
- [7] Scimeca, S., et al. (2023) The Impact of AI Tools on Education: Learning Outcomes and Academic Integrity. Proceedings of the European Conference on E-Learning.
https://www.researchgate.net/publication/376260972_Impact_of_AI-Powered_Learning_Tools_on_Student_Understanding_and_Academic_Performance
- [8] Qi, H., Liu, J., McAuley, J. and Riedel, S. (2023) SLaTe: A Large-Scale Benchmark for Text-Based Academic Literature Search. arXiv Preprint.
- [9] Akgun, S. and Greenhow, C. (2021) Artificial Intelligence in Education: Addressing Ethical Challenges in K-12 Settings. *AI and Ethics*, **2**, 431-440.
<https://doi.org/10.1007/s43681-021-00096-7>
- [10] Roe, G. and Perkins, R. (2023) "What They're Not Telling You about ChatGPT": Exploring the Discourse of AI in UK News Media Headlines. *Humanities and Social Sciences Communications*, **10**, Article No. 753.
<https://doi.org/10.1057/s41599-023-02282-w>
- [11] Bio-Conferences (2024) The Role of AI in Medical and Healthcare Education. BIO Web of Conferences.
- [12] U.S. News & World Report (2024) Best National University Rankings.
<https://www.usnews.com/best-colleges/rankings/national-universities>
- [13] U.S. News & World Report (2024) Best Computer Science Schools.
<https://www.usnews.com/best-graduate-schools/top-science-schools/computer-science-rankings>