The data collected from both case researchers and science influencers were analyzed quantitatively as follows.

The content of Twitter (rebranded as "X" since 2023) was subjected to text mining analysis using natural language processing (NLP) techniques. NLP facilitates the examination of words or lemmas that convey meaning in a sentence. Common stop words, such as “and” and “the,” were excluded from the analysis to focus on relevant content [1].

The current study employed R programming (R Project for Statistical Computing) to perform text mining and data analysis. In particular, dendrograms were employed as tree-like structures that indicated the relationships between the most frequently used words. These were subsequently clustered using the k-means algorithm, which assisted in identifying the main topics present in the messages on Twitter/X.

Current research has focused on text mining and analyzing messages’ emotional and sentimental aspects. This analysis employs computational techniques to provide a quantifiable measure of discourse structure and identifies potential controversies related to specific topics [2].

The primary objective of this study was to obtain essential information regarding the content and themes communicated in the messages shared on Twitter/X by case researchers and scientific influencers. Researchers have used analytical techniques to evaluate engagement and communication practices quantitatively. Through the analysis of data obtained from the platform messages, the researchers could discern the types of content shared by the users, which could provide insight into the role of science influencers and case researchers in disseminating scientific knowledge.

Graphical analysis using network theory can offer valuable insights into the relationships and interactions among the users of Twitter/X. By examining retweet relationships, this study aims to identify patterns of information diffusion and clusters or communities of users who frequently engage with one another's content.

To visualize the network, the researchers employed Gephi ver. 0.9.4 (Gephi Consortium) and used the Louvain algorithm for community detection. This algorithm recognizes clusters of nodes with a higher density of connections within each group, depicted in grayscale in the graph. This allows researchers to distinguish between various communities or groups of users based on their retweet patterns.

The present investigation sought to disclose the configuration of Twitter/X’s platform network, pinpoint influential users or nodes, and explore the emergence of communities or subgroups in the academic field using network analysis methods. This examination aims to offer a more in-depth understanding of the interactions and knowledge dissemination within Twitter/X’s platform environment, thereby identifying crucial figures and groups that significantly impact academic discourse and knowledge sharing in the educational domain.

This study utilized the Kardashian index (K-index) [3] to examine and assess the influence of scientific researchers on social media platforms, with a particular emphasis on Twitter/X’s platform. The K-index is a measure calculated by determining the ratio of retweets to followers for a researcher, which serves to quantify the level of engagement and attention that their messages receive relative to their follower count. This calculation is reminiscent of the phenomenon observed in celebrity culture, where individuals often attain considerable attention and followers despite having limited substantive contributions or expertise in their respective fields. The term “Kardashian” was adopted to represent this phenomenon.

Researchers may employ the K-index to evaluate the impact of their social media presence and assess whether their engagement and reach are proportional to the number of followers. A lower K-index value may indicate inadequate interaction with followers, whereas a higher value (K > 5) suggests greater engagement and attention relative to the follower count.

The K-index is a beneficial metric for evaluating researchers’ social media performance and ability to disseminate scientific knowledge. It aids in identifying researchers who effectively engage their audiences and utilize social media for substantial science communication.

This study employed statistical analysis techniques, specifically Pearson correlation and regression tree modeling, to explore the relationships between variables and predict the behavior of the dependent variable. The Pearson correlation technique assessed the significance of any substantial correlations between the K-index and related independent variables, such as message content, user behavior, and other factors, such as retweets, likes, mentions, followers, and users. This analysis provides critical insights into the relationships between these variables and their potential impacts on the dependent variable.

The application of Pearson correlation analysis and regression tree modeling presents a comprehensive approach for investigating the connections and predicting the behavior of the K-index by considering various message and user behavior variables. This approach enables a thorough examination of the relationships between the variables and the capacity to make predictions based on these connections.

https://www.escienceediting.org
References

