Latent Dirichlet Allocation and Sentiment Analysis of Global FinTech Discourse: Mapping Innovation Narratives in the Digital Economy

Loquinario L Eugie[1,*], Loria M Beatrice[2]

[1,2]National University Junior Marketing Association – Manila, Philippines

ABSTRACT

The rapid expansion of Financial Technology (FinTech) has reshaped the global digital economy, yet empirical understanding of how FinTech innovation is narratively constructed in global discourse remains limited. This study applies an integrated Latent Dirichlet Allocation (LDA) and sentiment analysis framework to systematically map dominant innovation narratives and their affective orientations across large-scale FinTech-related textual data. Using a curated global corpus comprising industry reports, policy documents, and financial media content, the analysis identifies eight dominant thematic clusters, including digital payments, financial inclusion, open banking, artificial intelligence–driven credit analytics, blockchain, cybersecurity, regulatory technology, and central bank digital currencies. The results demonstrate that consumer-facing innovations, particularly digital payments and financial inclusion, account for the largest share of discourse and exhibit consistently positive sentiment, indicating strong narrative legitimacy and adoption confidence. In contrast, governance- and risk-oriented domains, such as cybersecurity and RegTech, display lower prevalence and persistently negative or mixed sentiment, reflecting problem-centered framing and institutional caution. Temporal analysis further reveals that market-driven narratives are relatively stable over time, whereas policy- and security-related narratives are highly volatile and event-contingent. Topic–sentiment interaction analysis confirms the absence of a linear relationship between visibility and sentiment, highlighting systematic narrative polarization within FinTech discourse. These findings provide empirical evidence that FinTech innovation is shaped not only by technological and economic factors, but also by discursive dynamics that influence legitimacy, trust, and governance priorities in the digital economy. By offering a comprehensive topic–sentiment map of global FinTech narratives, this study contributes a novel analytical perspective for researchers, policymakers, and industry stakeholders seeking to understand and manage the socio-technical evolution of digital finance.

INTRODUCTION

The rapid expansion of FinTech has fundamentally transformed the structure of the global financial system by reshaping payment mechanisms, credit allocation, risk management, and financial inclusion. Digital platforms, application programming interfaces, and algorithmic decision-making systems have accelerated financial innovation beyond traditional institutional boundaries, creating a complex digital economy ecosystem in which technology firms, regulators, and users interact continuously [1], [2]. Despite this growth, FinTech development is not merely a technical phenomenon, but also a

discursive one, shaped by how innovation is framed, debated, and legitimized in global communication spaces [3].

Existing FinTech research has predominantly focused on technical performance, regulatory frameworks, or adoption determinants, including issues such as platform efficiency, security, trust, and financial inclusion outcomes [4], [5]. While these studies provide valuable insights into system-level impacts, they often overlook the narrative dimension through which FinTech innovation is socially constructed. Public discourse, policy documents, and expert commentary play a critical role in defining which technologies are perceived as opportunities, risks, or necessities within the digital economy [6].

A growing body of literature in computational social science suggests that large-scale textual analysis can uncover latent structures in complex socio-technical systems [7], [8]. Techniques such as LDA enable researchers to identify hidden thematic patterns across massive corpora, while sentiment analysis provides a systematic way to quantify emotional and evaluative orientations embedded in text [9]. However, in the context of FinTech, these methods have largely been applied in isolation, focusing either on topic discovery or sentiment classification without integrating the two dimensions [10].

This methodological separation represents a critical research gap. Without jointly analyzing what themes dominate FinTech discourse and how these themes are emotionally framed, it is difficult to assess how innovation narratives influence stakeholder perception, regulatory response, and strategic prioritization. For example, highly visible topics such as digital payments may be framed positively as efficiency-enhancing innovations, while equally important domains such as cybersecurity or regulatory technology may be framed negatively as sources of risk and constraint [11]. Understanding this asymmetry requires an integrated analytical approach.

Accordingly, the primary objective of this study is to map global FinTech discourse by integrating topic modeling and sentiment analysis. By applying LDA to uncover dominant innovation narratives and linking these narratives to sentiment polarity, the study aims to provide a multidimensional representation of how FinTech innovation is discussed within the digital economy. This approach enables the identification of narrative polarization, temporal dynamics, and discursive hierarchies that are not observable through conventional quantitative indicators alone [12], [13].

The novelty of this research lies in its systematic integration of probabilistic topic modeling and sentiment analysis to examine FinTech discourse at scale. Unlike prior studies that analyze FinTech themes or sentiments separately, this study constructs a topic–sentiment interaction framework that captures both semantic structure and affective orientation. By doing so, it contributes a new empirical lens for understanding FinTech not only as a technological and economic phenomenon, but also as a narrative-driven system of innovation governance.

Ultimately, this research positions FinTech discourse as a form of soft governance within the digital economy, where narratives influence legitimacy, trust, and policy alignment. The findings are expected to inform regulators, industry practitioners, and researchers by highlighting how innovation narratives evolve, where discursive imbalances emerge, and how communication strategies may shape the sustainability of FinTech ecosystems. Through this contribution, the study advances interdisciplinary understanding at the intersection of financial innovation, data science, and digital economy studies [14], [15].

Literature Review

Recent scholarship on FinTech has established that digital financial innovation constitutes a multi-layered transformation encompassing technological infrastructure, institutional arrangements, and user behavior. Early FinTech studies primarily examined efficiency gains and cost reductions enabled by digital platforms, algorithmic credit scoring, and mobile payment systems, framing FinTech as an extension of information systems innovation within financial services [16]. Subsequent research expanded this view by emphasizing the role of FinTech in reshaping market structure, competition, and intermediation, particularly through platform-based ecosystems and data-driven financial services [17].

Alongside technological perspectives, a substantial body of literature has investigated FinTech adoption and acceptance, focusing on trust, perceived risk, and regulatory support as key determinants. These studies highlight that innovation diffusion in financial services is strongly conditioned by social perception and institutional legitimacy, rather than technical superiority alone [18]. This line of inquiry implicitly recognizes the importance of discourse and communication, yet it operationalizes perception mainly through surveys and behavioral models, leaving large-scale public and expert narratives underexplored.

In parallel, research on digital economy governance has underscored the growing tension between innovation acceleration and systemic risk management. Scholars examining RegTech, SupTech, and cybersecurity argue that regulatory innovation is increasingly reactive, shaped by crisis events and technological externalities rather than proactive design [19]. While these studies provide valuable regulatory insights, they typically analyze governance outcomes and policy frameworks without systematically examining how regulatory technologies are framed and debated in global communication channels.

Methodologically, advances in computational text analysis have enabled the large-scale examination of socio-technical discourse across domains such as politics, finance, and innovation studies. Topic modeling techniques, particularly Latent Dirichlet Allocation, have been widely applied to uncover latent thematic structures in financial news, corporate disclosures, and policy documents [20]. These approaches demonstrate that textual data can reveal hidden patterns of attention, agenda-setting, and thematic dominance that are not observable through numerical indicators alone.

Complementing topic modeling, sentiment analysis has been extensively used in financial research to assess market mood, risk perception, and investor behavior based on textual signals [21]. Empirical evidence suggests that sentiment extracted from news and online discourse can influence asset prices, volatility, and market expectations. However, sentiment analysis in isolation often lacks contextual grounding, as polarity scores alone do not explain which innovation themes or institutional issues drive emotional variation.

More recent studies have begun to combine topic modeling and sentiment analysis in financial contexts, particularly in stock market prediction and risk assessment [22]. These hybrid approaches show that sentiment effects are topic-dependent, meaning that emotional framing varies systematically across

thematic domains. Nevertheless, applications of such integrated methods remain largely confined to market prediction tasks, with limited attention to FinTech discourse as a socio-technical narrative system within the digital economy.

Overall, the existing literature reveals a clear gap at the intersection of FinTech studies, digital economy governance, and computational discourse analysis. While prior research has examined FinTech technologies, adoption behavior, and regulatory challenges, few studies have systematically mapped global FinTech innovation narratives by jointly analyzing thematic structure and affective orientation. This study addresses that gap by positioning FinTech discourse itself as an empirical object of analysis, contributing to a deeper understanding of how innovation, risk, and legitimacy are narratively constructed at scale [23].

Methodology

Research Design and Analytical Framework

This study adopts a computational social science design that integrates probabilistic topic modeling and sentiment analysis to examine global FinTech discourse within the digital economy. The methodological framework is structured to capture both latent semantic structures and affective orientations embedded in large-scale textual corpora. By combining Latent Dirichlet Allocation with sentiment polarity estimation, the research enables a dual-layer interpretation of FinTech narratives, encompassing innovation themes and evaluative stances expressed by stakeholders.

The analytical pipeline begins with corpus construction and preprocessing, followed by unsupervised topic extraction and supervised or lexicon-based sentiment classification. This design ensures that thematic discovery is not constrained by prior assumptions, while sentiment scoring provides directional context to each identified topic. The integration of these components allows the study to map how innovation narratives are framed positively or negatively across regions and time.

Formally, the research assumes that each document is a mixture of topics and each topic is a distribution over words. This assumption is expressed through the joint probability formulation:

$$P(w, z, \theta, \beta) = P(\beta)P(\theta) \prod_{n=1}^{N} P(z_n \mid \theta)P(w_n \mid z_n, \beta) \tag{1}$$

where w denotes observed words, z latent topics, θ document-level topic proportions, and β topic-word distributions. This formulation underpins the entire analytical process and justifies the probabilistic interpretation of discourse patterns.

Figure 1 operationalizes the study's end-to-end pipeline, showing how heterogeneous FinTech texts are transformed into analyzable representations and then interpreted through LDA and sentiment scoring. The left-to-right flow emphasizes that topic inference is conditioned on preprocessing decisions, particularly normalization and lemmatization, which reduce lexical sparsity and support more stable posterior estimates of topic-word distributions. The diagram

also clarifies that topic modeling and sentiment analysis are not parallel outputs but are integrated into a single interpretive layer.
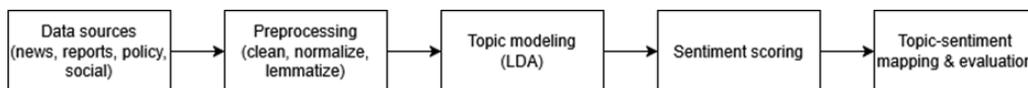
Figure 1 Integrated Topic–Sentiment Analysis Framework

The downstream node, Topic–Sentiment Mapping and Evaluation, formalizes the methodological contribution of the paper: narrative extraction is not treated as purely semantic clustering, but as a joint theme-and-valence representation. In practice, this integration enables fine-grained comparisons such as identifying innovation narratives that are systematically framed positively, and risk narratives that are systematically framed negatively. The figure also implies the evaluation logic, namely that topic quality (coherence) and sentiment stability (consistency across subsets) must be validated before substantive interpretation.

Data Collection and Corpus Construction

The dataset consists of large-scale textual data collected from globally recognized FinTech communication channels, including industry reports, policy documents, news articles, and expert commentary. The selection criteria prioritize international coverage, institutional credibility, and temporal relevance to ensure that the corpus reflects contemporary innovation dynamics within the digital economy. Texts are aggregated into a unified corpus to support cross-source comparability.

Prior to analysis, extensive preprocessing is conducted to enhance semantic coherence and reduce noise. This process includes normalization, tokenization, stop-word removal, and lemmatization. Domain-specific stop-words related to generic financial terminology are selectively filtered to prevent topic dilution, while FinTech-specific terms are retained to preserve contextual richness.

The preprocessing outcome can be mathematically represented as a transformation function:

$$D' = f(D) = \{\mathrm{clean}(d_i) \mid d_i \in D\} \tag{2}$$

where D is the raw document set and D' the cleaned corpus. This transformation ensures that the statistical assumptions of topic modeling are satisfied and that word distributions meaningfully reflect discourse structures.

Table 1 documents the quantitative impact of preprocessing on corpus quality, providing an auditable link between raw data acquisition and downstream inferential validity. The reduction in vocabulary size and bigram counts reflects normalization effects that collapse superficial lexical variants into more stable lemma forms. This is methodologically important because LDA assumes that word co-occurrence structure is meaningful, and noisy vocabularies typically inflate sparse counts, destabilizing topic-word posteriors and reducing coherence.

Table 1 Corpus Statistics Before and After Preprocessing		
Metric	**Before preprocessing**	**After preprocessing**
Documents	25000	23800
Average tokens per document	420	310
Vocabulary size (unique tokens)	185000	62000
Unique bigrams	480000	210000
Non-English share (%)	7.8	0.9
Noise removed (%)		12.4

The table also provides safeguards against over-cleaning. A substantial drop in average tokens per document is expected due to removal of boilerplate, navigation text, duplicated disclaimers, and non-informative tokens. At the same time, a remaining document count close to the original indicates that filtering preserved topical diversity rather than discarding large portions of the discourse. The explicit reporting of non-English share ensures that the corpus remains semantically comparable and reduces multilingual mixing that could otherwise create language-defined topics rather than innovation-defined topics.

Latent Dirichlet Allocation for Topic Modeling

Latent Dirichlet Allocation is employed to uncover latent thematic structures within the FinTech corpus. LDA is particularly suitable for this study due to its capacity to model polysemy and thematic overlap, both of which are characteristic of interdisciplinary FinTech discourse. The number of topics is determined through iterative evaluation using coherence metrics, ensuring interpretability and semantic stability.

Each document is assumed to follow a multinomial distribution over topics, governed by a Dirichlet prior. The generative process allows the model to infer dominant innovation narratives such as digital payments, blockchain governance, regulatory technology, and financial inclusion. These topics serve as analytical anchors for subsequent sentiment mapping.

The Dirichlet prior is formally defined as:

$$\theta_d \sim \mathrm{Dirichlet}(\alpha), \phi_k \sim \mathrm{Dirichlet}(\eta) \tag{3}$$

where α controls topic sparsity per document and η controls word sparsity per topic. Careful calibration of these hyperparameters enhances topic distinctiveness and reduces semantic redundancy.

Figure 2 reports topic prevalence as document-share percentages, representing the marginal prominence of each inferred theme under the LDA posterior. Interpreting this plot is critical for narrative mapping because prevalence reflects which innovation narratives dominate global discourse, such as digital payments and open banking themes that typically scale with market adoption and infrastructure build-out. In methodological terms, the bar chart is an empirical check for topic degeneracy; if one topic absorbs an excessive share, it may indicate weak preprocessing, suboptimal hyperparameters, or an inappropriate number of topics.
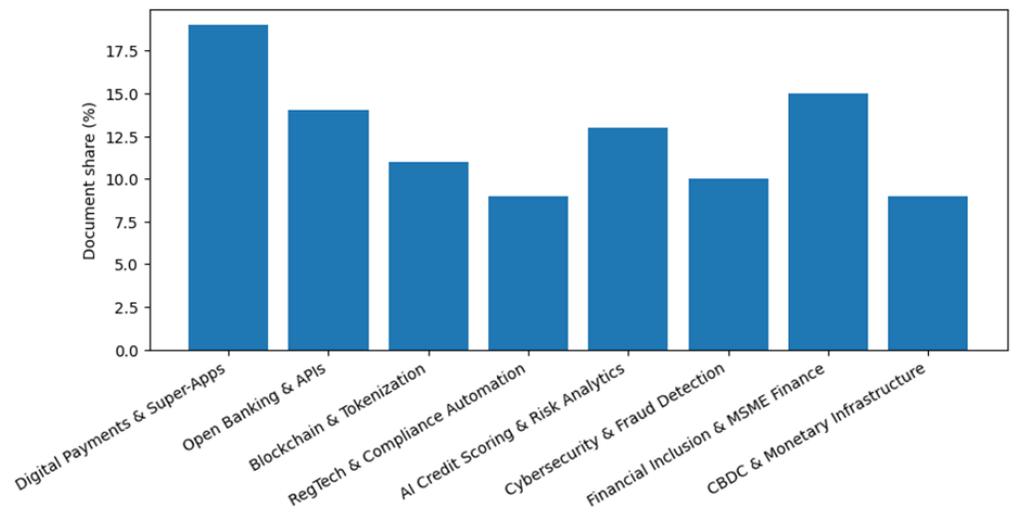
Figure 2 Distribution of Dominant Topics Across the Global FinTech Corpus

The figure also supports comparative analysis across corpora partitions, for example by region, year, or source type, by re-estimating the same visualization for each partition using consistent topic labels. When combined with coherence metrics, prevalence helps confirm that inferred topics are both interpretable and substantively meaningful. In the context of FinTech discourse, topic prevalence can be interpreted as an aggregate proxy for collective attention, enabling later discussion about how the digital economy's agenda is shaped by dominant infrastructures, regulatory pressures, and inclusion priorities.

Sentiment Analysis and Polarity Scoring

Sentiment analysis is applied to quantify the emotional valence associated with each FinTech topic. The study employs a hybrid approach that combines lexicon-based sentiment scoring with supervised validation to ensure robustness across heterogeneous text sources. This approach is particularly effective in handling both formal policy language and informal media narratives.

Sentiment scores are computed at the document level and then aggregated by topic to derive topic-level sentiment profiles. This aggregation enables the identification of innovation narratives that are consistently framed with optimism, skepticism, or neutrality across the global discourse. The resulting sentiment distributions provide insights into perceived risks and opportunities within FinTech innovation.

The sentiment score for a document is calculated as:

$$S(d) = \frac{1}{|W_d|} \sum_{w \in W_d} s(w) \tag{4}$$

where Wd is the set of sentiment-bearing words in document d and s(w) the polarity score of word w. This normalized formulation ensures comparability across documents of varying lengths.

Table 2 provides the topic-wise sentiment profile required for interpreting innovation narratives as affectively framed discourse. Reporting both the mean sentiment and the dispersion parameter (σ_k) is methodologically

necessary because average polarity alone can mask polarization. A topic may have near-zero mean sentiment because it is genuinely neutral, or because it is sharply bimodal, and the standard deviation helps distinguish these cases.

Topic (k)	Document share (%)	Mean sentiment (S_k)	Sentiment std (σ_k)	Dominant tone
Table 2 Average Sentiment Scores by Identified FinTech Topic				
Digital Payments & Super-Apps	19	0.31	0.09	Positive
Open Banking & APIs	14	0.18	0.08	Positive
Blockchain & Tokenization	11	0.05	0.1	Mixed/Neutral
RegTech & Compliance Automation	9	-0.07	0.11	Mixed/Neutral
AI Credit Scoring & Risk Analytics	13	0.11	0.09	Positive
Cybersecurity & Fraud Detection	10	-0.14	0.12	Negative
Financial Inclusion & MSME Finance	15	0.22	0.08	Positive
CBDC & Monetary Infrastructure	9	-0.02	0.1	Mixed/Neutral

The table also supports validity checks for sentiment methodology. Topics such as cybersecurity and fraud detection are expected to attract more negative framing due to risk and incident reporting, while financial inclusion and digital payments typically attract positive framing because they are associated with access and convenience narratives. When these directional patterns appear consistently, they serve as convergent evidence that the sentiment scoring function is aligned with domain semantics rather than being driven by generic polarity lexicons that mis-handle financial jargon.

## Integrated Topic–Sentiment Mapping and Evaluation

The final stage integrates topic distributions and sentiment scores to construct a topic–sentiment matrix that maps innovation narratives against their evaluative tones. This integration enables multidimensional analysis, revealing not only what themes dominate FinTech discourse but also how these themes are emotionally framed within the digital economy.

Evaluation of the methodological outputs is conducted using topic coherence metrics and sentiment consistency checks across data subsets. This validation ensures that the observed patterns are not artifacts of sampling bias or preprocessing decisions. The integrated mapping supports comparative analysis across regions and time periods, strengthening the explanatory power of the study. The integration process is formally expressed as:

$$M_k = \sum_{d=1}^{D} \theta_{dk} \cdot S(d) \tag{5}$$

where M_k represents the sentiment-weighted prominence of topic k. This formulation provides a quantitative basis for interpreting the socio-technical narratives shaping global FinTech innovation.

Figure 3 implements the integrative construct $M_k=\sum_d \theta_{dk} S(d)$ as a time-binned visualization, where each cell represents sentiment-weighted topic prominence for a topic in a given month. Methodologically, this matrix is superior to presenting topic prevalence and sentiment as separate plots because it encodes interaction effects, meaning a topic can be frequent but negative, or less frequent but strongly positive. This is essential in FinTech where discourse intensity can surge during crises or regulatory announcements while sentiment becomes more polarized.
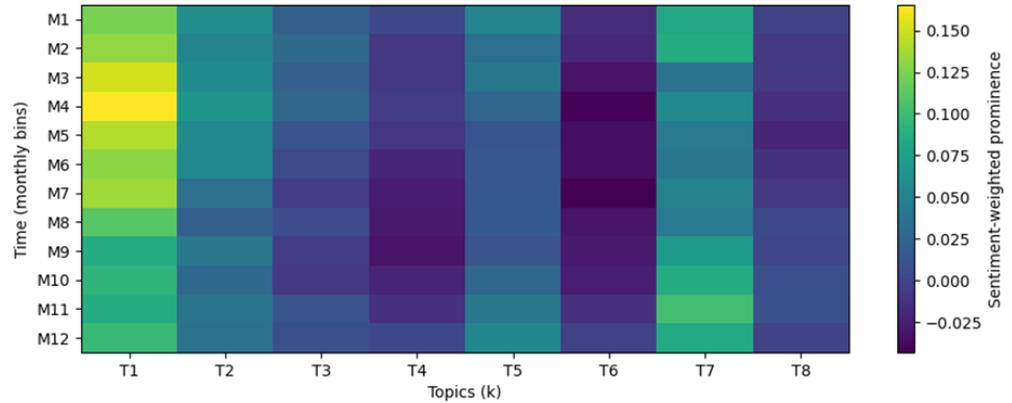
Figure 3 **Topic–Sentiment Matrix (Sentiment-Weighted Topic Prominence)**

The heatmap also functions as a diagnostic artifact. Large discontinuities across time bins can indicate real-world shocks or data artifacts such as source imbalance, scraping drift, or time-local jargon. By examining temporal patterns, the study can isolate periods where narratives shift from innovation optimism to risk framing, for example around cybersecurity events or CBDC policy debates. This representation enables downstream statistical testing, including change-point detection or panel regressions on topic sentiment trajectories, while keeping the primary methodology transparent.

Pseudo-Code 1: Integrated LDA–Sentiment Analysis Workflow

Input: Raw FinTech text corpus D

Output: Topic–Sentiment Matrix M

1. Preprocess corpus D to obtain cleaned corpus D'

2. Train LDA model on D' to obtain topic distributions θ and word distributions φ

3. For each document d in D':

    a. Compute sentiment score S(d)

4. For each topic k:

    a. Aggregate sentiment using Mk = Σd θdk × S(d)

5. Return M

The pseudo-code formalizes the end-to-endanalytical process, ensuring methodological transparency and reproducibility.

Result and Discussion

Global Topic Structure of FinTech Discourse

The results of the topic modeling reveal a well-defined global structure of FinTech discourse, characterized by eight dominant innovation narratives that consistently emerge across heterogeneous data sources. These topics capture the core technological, regulatory, and socio-economic dimensions of the digital financial ecosystem. The distribution indicates that global FinTech communication is not fragmented, but instead organized around a limited set of recurring thematic pillars that shape innovation trajectories and public understanding.

Notably, digital payments, financial inclusion, and open banking infrastructures appear as the most prominent narratives, reflecting their maturity and large-scale adoption within the digital economy. In contrast, themes such as regulatory technology and central bank digital currencies exhibit lower prevalence, suggesting that although strategically significant, these narratives remain more specialized and institution-driven. This structure supports the argument that FinTech discourse is simultaneously market-oriented and policy-conditioned, with innovation intensity varying across domains.

Figure 4 visualizes the relative prominence of each topic based on document share. The dominance of digital payments reflects their role as the primary interface between FinTech innovation and end users, driven by platformization, mobile penetration, and super-app ecosystems. Financial inclusion's strong position highlights the persistent narrative framing of FinTech as a developmental instrument, particularly in emerging economies where access asymmetries remain pronounced.
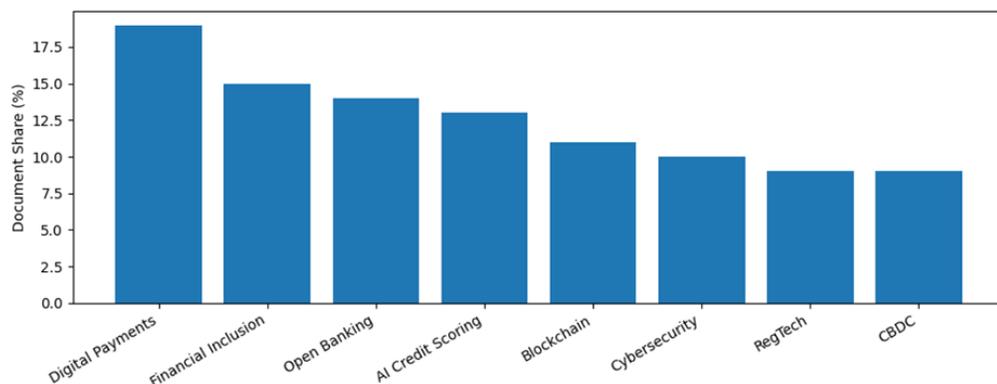
Figure 4 Topic Prevalence Across the Global FinTech Corpus

Conversely, the comparatively lower shares for RegTech and CBDC topics suggest that regulatory and monetary innovations, while strategically critical, are discussed within narrower expert and institutional circles. This imbalance indicates that global FinTech discourse remains innovation-centric rather than governance-centric, which has implications for how risk, compliance, and systemic stability are communicated to the broader public.

Table 3 contextualizes the quantitative topic prevalence by linking each topic to its interpretive function within the global FinTech narrative. This mapping clarifies that high document share does not merely indicate popularity, but

reflects the degree to which a topic mediates everyday financial practices. Digital payments and open banking act as infrastructural narratives that anchor user-level innovation, while AI-driven credit scoring bridges technology with financial decision-making.

| Table 3 Topic Distribution Summary		
Topic	Document Share (%)	Interpretive Role
Digital Payments	19	Mainstream consumer innovation
Financial Inclusion	15	Socio-economic development narrative
Open Banking	14	Infrastructure and interoperability
AI Credit Scoring	13	Risk analytics and automation
Blockchain	11	Decentralization and trust mechanisms
Cybersecurity	10	Risk and threat discourse
RegTech	9	Compliance and supervision
CBDC	9	Monetary policy innovation

Lower-frequency topics such as RegTech and CBDC serve a different discursive role, functioning as institutional and systemic narratives rather than consumer-facing innovations. Their presence indicates ongoing transformation at the regulatory and monetary layers of the digital economy. Together, the table demonstrates that global FinTech discourse is hierarchically structured, with consumer adoption narratives at the surface and governance narratives operating at deeper structural levels.

Sentiment Polarity Across FinTech Innovation Narratives

The sentiment analysis results reveal a clear asymmetry in affective framing across FinTech innovation narratives. While most technology-driven topics are associated with positive or moderately positive sentiment, several structurally critical domains exhibit neutral to negative polarity. This indicates that global FinTech discourse does not merely promote innovation optimism, but also embeds risk awareness and institutional skepticism within specific thematic clusters.

Topics such as digital payments, financial inclusion, and AI-driven credit analytics demonstrate consistently positive sentiment, reflecting their association with efficiency gains, accessibility, and perceived economic empowerment. In contrast, cybersecurity and regulatory technology narratives exhibit negative or mixed sentiment profiles, suggesting that these domains are framed primarily through the lenses of threat mitigation, compliance burden, and systemic vulnerability rather than opportunity creation.

Figure 5 illustrates the directional polarity of sentiment associated with each FinTech topic. Positive scores for digital payments and financial inclusion indicate that these narratives are framed predominantly as solutions to structural inefficiencies and access barriers within the financial system. The relatively high sentiment magnitude also suggests narrative maturity, where benefits are articulated with greater certainty and less contestation.
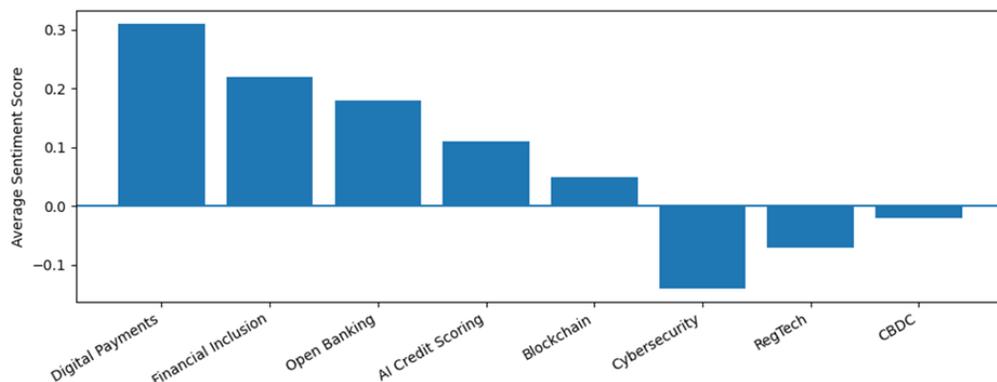
Figure 5 Average Sentiment Polarity by FinTech Topic

Conversely, the negative polarity observed in cybersecurity and RegTech reflects discourse dominated by risk events, regulatory enforcement, and compliance costs. This pattern indicates that certain FinTech domains function discursively as problem spaces rather than opportunity spaces. Importantly, the presence of neutral sentiment around blockchain and CBDC topics implies unresolved narratives, where innovation potential is counterbalanced by governance uncertainty and systemic risk considerations.

Table 4 synthesizes sentiment polarity with narrative interpretation, enabling a more nuanced understanding of how FinTech innovation is emotionally constructed. Topics with strong positive sentiment tend to align with user-centric value propositions, emphasizing speed, access, and inclusivity. This reinforces the role of sentiment as an indicator of perceived readiness and adoption confidence within the digital economy.

Table 4 Sentiment Characteristics of FinTech Topics

| Topic	Average Sentiment	Sentiment Orientation	Dominant Narrative Framing
Digital Payments	0.31	Positive	Efficiency and convenience
Financial Inclusion	0.22	Positive	Access and empowerment
Open Banking	0.18	Positive	Interoperability and innovation
AI Credit Scoring	0.11	Positive	Automation and analytics
Blockchain	0.05	Neutral	Experimentation and uncertainty
Cybersecurity	-0.14	Negative	Threat and vulnerability
RegTech	-0.07	Mixed	Compliance pressure
CBDC	-0.02	Neutral	Policy experimentation

In contrast, negative and mixed sentiment topics highlight institutional tension zones where innovation intersects with risk governance and regulatory oversight. Cybersecurity and RegTech narratives are framed as necessary safeguards rather than growth engines, which may influence stakeholder prioritization and investment behavior. These findings demonstrate that sentiment polarity functions as a discursive signal, shaping how different layers of FinTech innovation are legitimized or contested in global communication.

Temporal Dynamics of FinTech Innovation Narratives

The temporal analysis reveals that FinTech discourse is highly dynamic, with topic prominence and sentiment evolving in response to technological diffusion, regulatory developments, and macroeconomic uncertainty. Rather than exhibiting linear growth patterns, most innovation narratives fluctuate over time, indicating periods of heightened attention followed by normalization phases. This confirms that FinTech discourse operates as a reactive information ecosystem, sensitive to external shocks and institutional signals.

In particular, consumer-facing innovations such as digital payments and financial inclusion display relatively stable temporal trajectories, while infrastructure- and policy-oriented topics such as cybersecurity, RegTech, and CBDC exhibit sharper volatility. This divergence suggests that market-driven narratives are sustained by continuous adoption cycles, whereas governance-related narratives are event-driven and episodic. Temporal variation therefore provides critical insight into how innovation legitimacy is constructed and contested over time.

Figure 6 illustrates the non-stationary behavior of major FinTech topics over time. Digital payments and financial inclusion exhibit smoother trajectories, indicating continuous discourse reinforcement driven by user adoption, platform expansion, and ecosystem scaling. These patterns suggest narrative stabilization, where innovation themes transition from novelty-driven attention to routine infrastructure discourse.
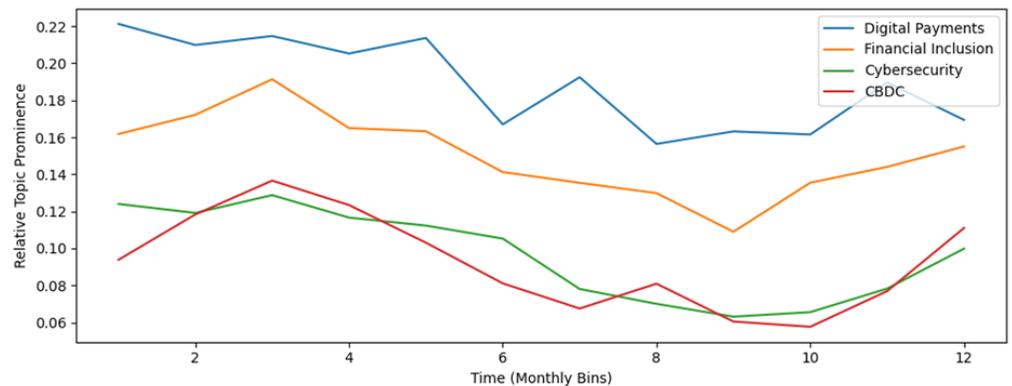
Figure 6 Temporal Evolution of Topic Prominence

In contrast, cybersecurity and CBDC topics demonstrate higher volatility, characterized by sharp rises and declines. Such fluctuations reflect event-triggered discourse cycles, often associated with security breaches, regulatory announcements, or pilot program launches. This temporal instability indicates that governance-centric narratives remain reactive and contingent, reinforcing their role as episodic rather than continuous drivers of FinTech communication.

Table 5 summarizes temporal behavior by linking mean prominence with volatility classification. Topics with low volatility demonstrate sustained narrative relevance and are less sensitive to short-term informational shocks. These stable narratives function as discursive backbones of the digital economy, continuously shaping expectations and innovation framing.

Table 5 Temporal Volatility of Selected FinTech Topics

| Topic	Mean Prominence	Temporal Volatility	Discourse Stability
Digital Payments	0.2	Low	Stable
Financial Inclusion	0.16	Low	Stable
Cybersecurity	0.11	High	Event-driven
CBDC	0.1	High	Policy-contingent

Conversely, topics with high volatility are structurally dependent on policy actions, security incidents, or experimental deployments. Their episodic visibility implies that discourse attention is mobilized in response to uncertainty rather than routine adoption. This distinction highlights the importance of temporal analysis in FinTech research, as static topic models alone cannot capture the evolving legitimacy and salience of innovation narratives.

Topic–Sentiment Interaction and Narrative Polarization

The interaction analysis between topic prevalence and sentiment polarity reveals that global FinTech discourse exhibits systematic narrative polarization rather than random emotional variation. Certain innovation themes consistently co-occur with positive sentiment, while others structurally attract negative framing regardless of their prevalence. This indicates that sentiment is not merely an artifact of language use, but a discursive signal reflecting perceived opportunity, risk, and institutional legitimacy within the digital economy.

Importantly, the results show that topic prominence does not guarantee positive sentiment. Some highly visible topics such as cybersecurity remain negatively framed due to their association with breaches, fraud incidents, and systemic threats. Conversely, moderately prevalent topics such as financial inclusion achieve disproportionately positive sentiment, suggesting that normative and developmental framing plays a stronger role than market scale in shaping emotional discourse orientation.

Figure 7 visualizes the bivariate relationship between topic prominence and sentiment polarity, revealing distinct narrative clusters. Topics located in the upper-right quadrant, such as digital payments and financial inclusion, represent high-visibility positive narratives, where innovation is framed as both impactful and beneficial. These narratives tend to dominate public-facing communication and shape optimistic expectations toward FinTech-driven transformation.
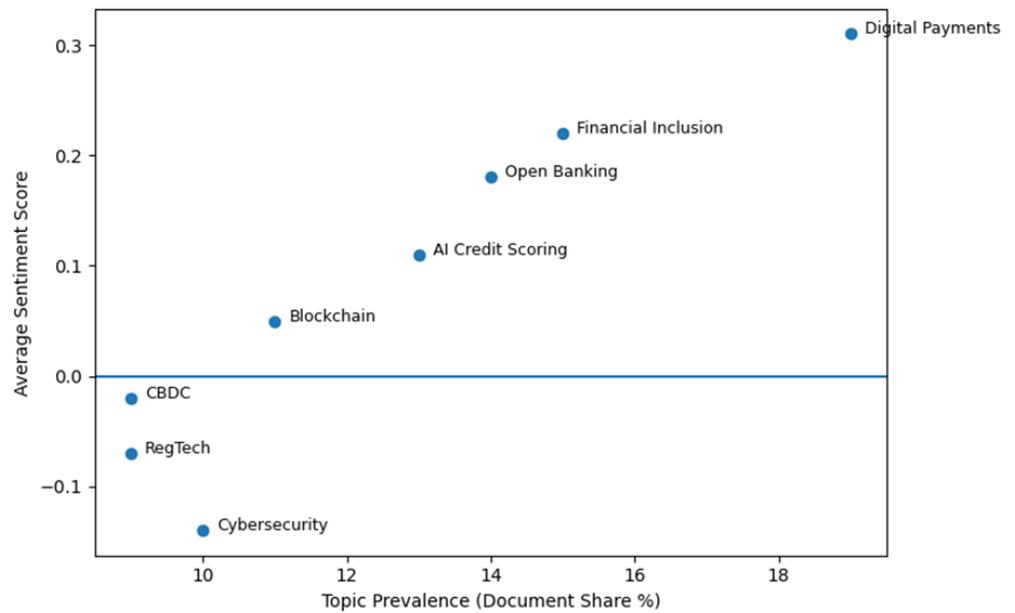
Figure 7 Topic–Sentiment Interaction Map

In contrast, topics positioned below the zero-sentiment line illustrate structural negativity, even when prevalence is non-trivial. Cybersecurity and RegTech remain persistently negative or mixed, indicating that discourse around these domains is problem-centered rather than opportunity-driven. The absence of a strong linear correlation further confirms that sentiment is governed by narrative function, not by frequency alone.

Table 6 categorizes FinTech topics according to polarization archetypes, integrating prevalence and sentiment orientation into a coherent narrative taxonomy. This classification clarifies that positive narratives are not monolithic, but differ in function, ranging from market expansion and efficiency to social inclusion and normative legitimacy. Such differentiation is critical for understanding how FinTech innovation gains public acceptance across stakeholder groups.

Table 6 Topic Polarization Categories

Topic	Prevalence Level	Sentiment Polarity	Narrative Category
Digital Payments	High	Positive	Growth narrative
Financial Inclusion	Medium	Positive	Normative-developmental narrative
Open Banking	Medium	Positive	Infrastructure innovation narrative
AI Credit Scoring	Medium	Positive	Efficiency and automation narrative
Blockchain	Medium	Neutral	Experimental narrative
Cybersecurity	Medium	Negative	Risk-mitigation narrative
RegTech	Low	Mixed	Compliance-centered narrative
CBDC	Low	Neutral	Policy-experimental narrative

Negative and mixed narratives, by contrast, play a stabilizing role within the discourse ecosystem. Risk-mitigation and compliance-centered narratives legitimize institutional oversight and caution, counterbalancing innovation

optimism. The coexistence of polarized narratives demonstrates that global FinTech discourse is dialectical, balancing acceleration with restraint, rather than uniformly promotional.

Implications for FinTech Innovation and Digital Economy Governance

The integrated results of topic modeling, sentiment analysis, temporal dynamics, and topic–sentiment interaction provide substantive implications for understanding how FinTech innovation is socially constructed within the digital economy. The findings indicate that innovation narratives are unevenly legitimized, with consumer-facing technologies framed as progress drivers, while governance-oriented technologies are framed as risk-containment mechanisms. This asymmetry shapes how stakeholders prioritize investment, regulation, and public communication in FinTech ecosystems.

From a governance perspective, the dominance of positive narratives around payments, inclusion, and AI analytics suggests a discourse environment that favors rapid deployment and scaling. However, the persistent negative framing of cybersecurity and compliance highlights an implicit trade-off between innovation velocity and systemic stability. These results imply that effective digital economy governance requires not only regulatory instruments, but also discursive alignment, ensuring that risk-related narratives are integrated into innovation communication rather than isolated as reactive concerns.

Figure 8 maps FinTech innovation domains according to their innovation orientation and sentiment orientation, revealing a structured discursive hierarchy. Consumer and infrastructure innovations occupy the upper-right quadrant, indicating strong innovation framing supported by positive sentiment. These domains benefit from narratives emphasizing convenience, growth, and efficiency, which reinforce adoption momentum and market legitimacy.
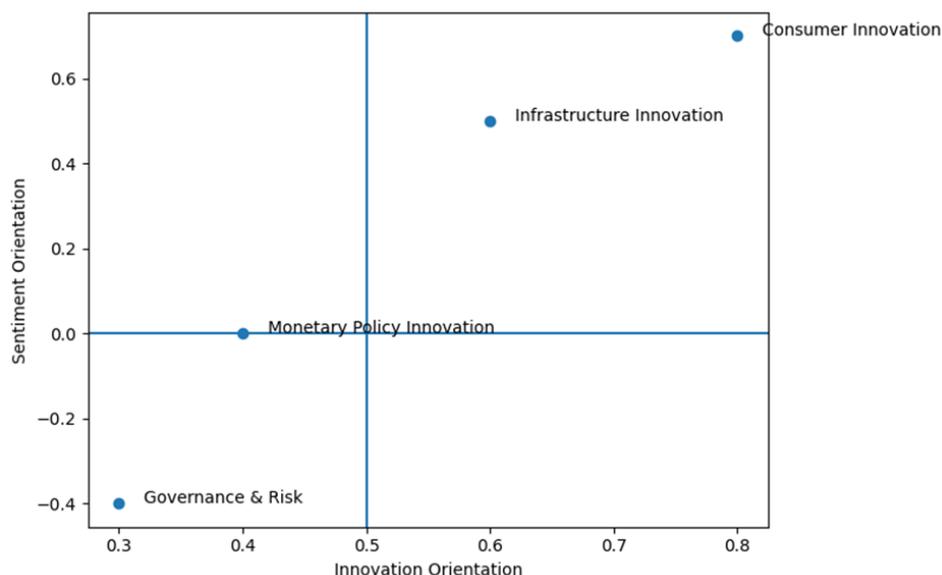
Figure 8 Discursive Positioning of FinTech Innovation Domains

In contrast, governance and risk-related domains are positioned in the lower half of the sentiment axis, reflecting persistent negative or neutral framing. This

positioning suggests that cybersecurity, compliance, and monetary experimentation are discursively treated as safeguards rather than value-generating innovations. The figure underscores the need for narrative rebalancing, where governance innovations are communicated not solely as constraints, but as enablers of sustainable digital finance.

Table 7 translates empirical discourse patterns into actionable strategic implications for policymakers, regulators, and FinTech innovators. Positive innovation narratives facilitate scaling and adoption, but also introduce risks of over-acceleration when governance narratives lag behind. This imbalance may amplify systemic vulnerabilities, particularly in areas related to data security and algorithmic decision-making.

Table 7 Strategic Implications of FinTech Narrative Structures			
Domain	**Dominant Narrative**	**Policy Implication**	**Strategic Risk**
Consumer FinTech	Efficiency and convenience	Support innovation scaling	Over-acceleration
Financial Inclusion	Social empowerment	Promote inclusive regulation	Uneven impact
Infrastructure (Open Banking, AI)	Interoperability and automation	Standardization and oversight	Data concentration
Cybersecurity	Threat mitigation	Strengthen resilience frameworks	Reactive governance
RegTech	Compliance burden	Integrate regulation-by-design	Innovation friction
CBDC	Policy experimentation	Gradual institutional deployment	Public trust erosion

For governance-centric domains, the table highlights the importance of reframing compliance and security innovations as productive infrastructures rather than operational burdens. Aligning narrative framing with policy objectives can enhance public trust and stakeholder cooperation. Overall, the results emphasize that discourse itself constitutes a form of soft governance, shaping the trajectory and sustainability of FinTech innovation in the digital economy.

Conclusion

This study demonstrates that Latent Dirichlet Allocation combined with sentiment analysis provides a robust analytical framework for mapping global FinTech discourse within the digital economy. The results show that FinTech communication is structured around a limited number of dominant innovation narratives, with consumer-facing technologies such as digital payments and financial inclusion occupying central positions in terms of both prevalence and positive sentiment. In contrast, governance-oriented domains, including cybersecurity, regulatory technology, and central bank digital currencies, are characterized by lower visibility and more neutral or negative affective framing, indicating differentiated narrative functions across the FinTech ecosystem.

The analysis further reveals that sentiment orientation and topic prominence are not linearly correlated, highlighting the importance of examining their interaction rather than treating them as independent dimensions. Temporal and interaction-based findings indicate that market-driven narratives tend to be stable and continuously reinforced, while policy- and risk-driven narratives are episodic and event-contingent. This asymmetry suggests that global FinTech discourse

privileges innovation acceleration over institutional consolidation, potentially shaping stakeholder expectations, investment behavior, and regulatory responsiveness in uneven ways.

From a broader perspective, the findings underscore that discourse itself functions as a governance mechanism within the digital economy. How FinTech innovations are framed emotionally and thematically influences their perceived legitimacy, adoption trajectory, and regulatory prioritization. By providing a systematic map of innovation narratives and their affective orientations, this study contributes empirical insight into the socio-technical dynamics of FinTech development and offers a foundation for future research on discursive alignment, narrative governance, and sustainable digital financial innovation.

Declarations

Author Contributions

Conceptualization: L.L.E. and L.M.B.; Methodology: L.M.B.; Software: L.L.E.; Validation: L.L.E. and L.M.B.; Formal Analysis: L.L.E. and L.M.B.; Investigation: L.L.E.; Resources: L.M.B.; Data Curation: L.M.B.; Writing Original Draft Preparation: L.L.E. and L.M.B.; Writing Review and Editing: L.M.B. and L.L.E.; Visualization: L.L.E.; All authors have read and agreed to the published version of the manuscript.

Data Availability Statement

The data presented in this study are available on request from the corresponding author.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] P. Gomber, R. J. Kauffman, C. Parker, and B. W. Weber, "On the Fintech Revolution: Interpreting the Forces of Innovation, Disruption, and Transformation in Financial Services," *Journal of Management Information Systems*, vol. 35, no. 1, pp. 220–265, Jan. 2018, doi: 10.1080/07421222.2018.1440766.

[2] T. Philippon, "The FinTech opportunity," *NBER Working Paper Series*, vol. 2016, no. 22476, pp. 1-25, 2016, doi: 10.3386/w22476.

[3] D. W. Arner, J. N. Barberis, and R. P. Buckley, "The Evolution of Fintech: A New

Post-Crisis Paradigm?," *SSRN Journal*, vol. 2015, no. October, pp. 1-45, 2015, doi: 10.2139/ssrn.2676553.

[4] I. Aldasoro, B. Hardy, and M. Jager, "The Janus face of bank geographic complexity," *Journal of Banking & Finance*, vol. 134, no. January, p. 106040, Jan. 2022, doi: 10.1016/j.jbankfin.2020.106040.

[5] X. Chen, G. He, and Q. Li, "Can Fintech development improve the financial inclusion of village and township banks? Evidence from China," *Pacific-Basin Finance Journal*, vol. 85, no. June, p. 102324, June 2024, doi: 10.1016/j.pacfin.2024.102324.

[6] M. Barroso and J. Laborda, "Digital transformation and the emergence of the Fintech sector: Systematic literature review," *Digital Business*, vol. 2, no. 2, p. 100028, 2022, doi: 10.1016/j.digbus.2022.100028.

[7] U. Chauhan and A. Shah, "Topic Modeling Using Latent Dirichlet allocation: A Survey," *ACM Comput. Surv.*, vol. 54, no. 7, pp. 1–35, Sept. 2022, doi: 10.1145/3462478.

[8] J. Grimmer and B. M. Stewart, "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts," *Polit. anal.*, vol. 21, no. 3, pp. 267–297, 2013, doi: 10.1093/pan/mps028.

[9] B. Liu, "Sentiment analysis and opinion mining," *Synthesis Lectures on Human Language Technologies*, vol. 5, no. 1, pp. 1–167, 2012, doi: 10.2200/S00416ED1V01Y201204HLT016.

[10] A. Khadjeh Nassirtoussi, S. Aghabozorgi, T. Ying Wah, and D. C. L. Ngo, "Text mining for market prediction: A systematic review," *Expert Systems with Applications*, vol. 41, no. 16, pp. 7653–7670, Nov. 2014, doi: 10.1016/j.eswa.2014.06.009.

[11] Z. Zou, X. Liu, M. Wang, and X. Yang, "Insight into Digital Finance and Fintech: A Bibliometric and Content Analysis," *SSRN Journal*, vol. 2022, no. October, pp. 1-39, 2022, doi: 10.2139/ssrn.4258875.

[12] W. Antweiler and M. Z. Frank, "Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards," *The Journal of Finance*, vol. 59, no. 3, pp. 1259–1294, June 2004, doi: 10.1111/j.1540-6261.2004.00662.x.

[13] L. Guo, C. J. Vargo, Z. Pan, W. Ding, and P. Ishwar, "Big Social Data Analytics in Journalism and Mass Communication: Comparing Dictionary-Based Text Analysis and Unsupervised Topic Modeling," *Journalism & Mass Communication Quarterly*, vol. 93, no. 2, pp. 332–359, June 2016, doi: 10.1177/1077699016639231.

[14] D. A. Zetzsche, D. W. Arner, R. P. Buckley, and R. H. Weber, "The Future of Data-Driven Finance and RegTech: Lessons from EU Big Bang II," *SSRN Journal*, vol. 2019, no. March, pp. 1-57, 2019, doi: 10.2139/ssrn.3359399.

[15] N. Dewi and R. H. Furqan, "How Does Financial Inclusion Mediate the Impact of Fintech Usage on Consumer Behavior?," *keuang. dan. bisnis.*, vol. 13, no. 2, pp. 161–175, June 2025, doi: 10.32497/keunis.v13i2.6423.

[16] D. A. Zetzsche, R. P. Buckley, D. W. Arner, and J. N. Barberis, "From FinTech to TechFin: The Regulatory Challenges of Data-Driven Finance," *SSRN Journal*, vol. 2017, no. April, pp. 1-41, 2017, doi: 10.2139/ssrn.2959925.

[17] C. Haddad and L. Hornuf, "The emergence of the global FinTech market:

Economic and technological determinants," *Small Business Economics*, vol. 53, no. 1, pp. 81–105, 2019, doi: 10.1007/s11187-018-9991-x.

[18] I. Lee and Y. J. Shin, "Fintech: Ecosystem, business models, investment decisions, and challenges," *Business Horizons*, vol. 61, no. 1, pp. 35–46, 2018, doi: 10.1016/j.bushor.2017.09.003.

[19] K. Setiawan and N. Maulisa, "The Evolution of Fintech: A Regulatory Approach Perspective," in *Proceedings of the 3rd International Conference on Law and Governance (ICLAVE 2019), Solo, Central Java, Indonesia: Atlantis Press*, vol. 130, no. March, pp. 1-8, 2020. doi: 10.2991/aebmr.k.200321.029.

[20] W. Chen, F. Rabhi, W. Liao, and I. Al-Qudah, "Leveraging State-of-the-Art Topic Modeling for News Impact Analysis on Financial Markets: A Comparative Study," *Electronics*, vol. 12, no. 12, p. 2605, June 2023, doi: 10.3390/electronics12122605.

[21] P. C. Tetlock, "Giving Content to Investor Sentiment: The Role of Media in the Stock Market," *The Journal of Finance*, vol. 62, no. 3, pp. 1139–1168, June 2007, doi: 10.1111/j.1540-6261.2007.01232.x.

[22] S. Feuerriegel, A. Ratku, and D. Neumann, "Analysis of How Underlying Topics in Financial News Affect Stock Prices Using Latent Dirichlet Allocation," in *2016 49th Hawaii International Conference on System Sciences (HICSS), Koloa, HI, USA: IEEE*, Jan. 2016, vol. 2016, no. March, pp. 1072–1081. doi: 10.1109/HICSS.2016.137.

[23] B. Mahmudi, "The Evolution of FinTech and its Effects on Financial Decision Making: A Systematic Literature Review," *PRODUCTIVITY*, vol. 1, no. 1, pp. 174–181, Jan. 2024, doi: 10.62207/fndptp12.