A Cyber Threat Intelligence Model for Identification of Crime-Facilitating Elements on Darknet Marketplace Using Guided LDA-Based Topic Modeling Weights

Gowri Priya[1], *Manohar Naik S [2]

[1]Research Scholar, Department of Computer Science, Central University of Kerala, Kasaragod,India
[2]Assistant Professor, Department of Computer Science, Central University of Kerala, Kasaragod,India

[1]gowri.2500707005@cukerala.ac.in
[2*]manoharamen@cukerala.ac.in

Abstract:

The dark web is a hidden part of the internet that provides anonymity and privacy to users by masking their identities through multiple layers of encryption. This structural nature enables it to function as a covert platform for various illicit activities. Recent data indicate that stolen and compromised personal information, as well as financial fraud, remain dominant forms of cybercrime. However, existing research mainly concentrates on identifying crime categories rather than uncovering the hidden background mechanisms that drive these high-impact cybercrimes. To address this gap, we propose a novel framework that employs Guided Latent Dirichlet Allocation (Guided LDA) technique to identify hidden activities behind stolen-financial data related crimes on darknet marketplaces. This study utilizes a publicly available dataset from Kilos, a darknet search engine that indexes information from six darknet markets. The findings reveal that the Guided LDA model produces coherent and meaningful topics across multiple coherence metrics (UMass, UCI, CNPMI, and CV), and effectively identifies the recurring activity patterns and key facilitating elements present in darknet communications. By using topic modeling, this study offers deeper insight into the ecosystem of darknet crimes, assisting researchers and law enforcement in better understanding the underlying dynamics and developing more effective cyber threat detection strategies.

Keywords: Darknet; Cyber Threat Intelligence; LDA Topic Modeling; NLP; Coherence Metrics; Machine Learning.

D. R. Reddy et al. (eds.), *Proceedings of the First International Conference on Advances in Forensics and Cyber Technologies (ICFACT 2025)*, Advances in Computer Science Research 127,
https://doi.org/10.2991/978-94-6239-610-4_7

1 Introduction

The dark web serves as a platform for both legitimate and illegal activities. While it enables journalists, activists, and privacy-seeking individuals to maintain anonymity and protect their privacy, the same features are exploited by criminals to conceal their identities and carry out illicit operations, including trading drugs, weapons, forged documents, stolen data, and malware. Unlike the surface web, which can be accessed through standard browsers such as Edge and Chrome, the dark web requires specialized browsers like Tor, I2P, and Freenet (Jin et al., 2024) . Among these, Tor Browser (The Onion Router) is the most widely used browser for accessing dark web's .onion sites. Tor uses 'onion routing,' where data is encrypted in multiple layers and passed through several relays, with each relay removing one layer to reveal only the next destination. This design ensures that no single node knows both the source and the final endpoint, preserving user anonymity. However, Tor has limitations, including slower speeds due to multi-hop routing and potential vulnerabilities at exit nodes, where unencrypted data can be intercepted (Gond, 2025).

In cybersecurity research, these underground marketplaces play a crucial role because they provide direct visibility into the tools and services used by cybercriminals. Investigators and analysts closely monitor them to understand how threat actors operate, how their pricing and business models change over time, and how quickly they adopt new technologies. Studying this marketplace activity helps security professionals spot upcoming attack patterns early and prepare defensive measures in advance. This kind of insight is essential for weakening the financial networks and supply chains that keep global cybercrime running (Temara, 2024).

Dark Web marketplaces create major obstacles for investigators because they constantly shift and transform. New markets surface while others shut down or relocate without warning, preventing stable long-term observation. Sellers frequently change their identities, product descriptions, and selling habits to stay hidden, which breaks the continuity of intelligence gathered over time(Devarajan et al., 2024). These platforms post listings in multiple languages and diverse formats, often filled with slang, codewords, and encrypted communication, making large-scale information extraction extremely demanding (Zhang et al., 2025). Transactions between buyers and sellers are deliberately structured to protect anonymity through cryptocurrencies, escrow mechanisms, and privacy-focused tools like Tor, making it even harder to trace activity or collect solid evidence. Altogether, this continuous instability, linguistic diversity, and built-in anonymity severely hinder efforts to map supply chains, track criminal collaborations, and spot emerging cybercrime trends(Owen et al., 2025).

Cyber Threat Intelligence (CTI) has become essential because cybercriminals continuously innovate new attack methods, tools, and services that slip past conventional security measures until it is too late. Darknet markets often serve as the earliest distribution hubs for fresh malware strains, exploit packages, stolen data, and "cybercrime-for-hire" services. Systematic tracking and analysis of these platforms enable the early detection of such malicious resources before they are deployed in widespread attacks. This early awareness strengthens response capabilities, improves strategic security decisions, and helps disrupt illicit cyber operations more effectively. CTI is not only about detecting criminal activities but also

about understanding the deeper mechanisms that allow them to grow and adapt. Yet most existing research on darknet marketplaces has focused only on basic crime categorization, grouping listings into broad illegal segments without exploring the operational nuances embedded within them (Santos et al., 2025). This type of surface-level analysis fails to capture the subtle behavioural patterns, strategic interactions, and economic signals that influence how cybercrime develops and spreads across these platforms. To address this gap, the present study goes beyond simple classification and aims to uncover the latent structures and thematic patterns within marketplace content. This approach provides a more practical and insightful form of CTI that explains not just what is being traded but also how the cybercriminal ecosystem evolves over time.

In the proposed framework, Guided Latent Dirichlet Allocation (LDA) is used to identify the underlying thematic structures within darknet marketplace listings and extract the most influential terms that define each topic. LDA is typically applied to document collections either to extract predefined topics based on an existing theoretical framework (a deductive approach) or to discover new, latent themes that can refine that framework (an inductive approach) (Watanabe & Baturo, 2024). However, evaluating LDA is difficult because it is an unsupervised method: although metrics like perplexity and coherence are widely used, perplexity often fails to reflect human judgement, coherence captures similarity but not interpretability, and tools such as LDAvis help visualization but offer limited numerical validation (Zhou et al., 2023). As a result, the alignment between machine-generated topics and theoretical concepts frequently occurs only by chance, since LDA identifies topics merely as clusters of co-occurring words in the corpus a limitation that complicates theory-driven research (Watanabe & Baturo, 2024; Zhou et al., 2023). Guided LDA overcomes this issue by allowing researchers to define topics in advance using seed words, thereby guiding the model toward conceptually meaningful outputs (Watanabe & Baturo, 2024).

For every topic produced by the model, the associated term weight probabilities indicate how strongly each word contributes, allowing a clearer interpretation of the hidden cybercriminal themes. These probabilities are then transformed into numerical indicators that quantify how closely a given listing aligns with a particular topic based on the distribution and contextual relevance of its vocabulary. As a result, listings are categorized not just by visible keywords but by the deeper semantic patterns present in their descriptions. Additionally, tracking topic-weight distributions across the dataset makes it possible to identify dominant trends and newly emerging themes over time, providing a scalable method for proactive threat detection rather than merely reactive classification (Basheer & Alkhatib, 2024; Shin et al., 2024).

The major contributions presented in this study are as follows:

- We developed a CTI model that identifies crime-facilitating elements rather than merely classifying crime categories. This shifts the focus from surface-level labeling to uncovering the hidden mechanisms behind darknet-based cybercrimes.
- We applied Guided Latent Dirichlet Allocation (LDA) topic modeling to extract latent activity patterns within darknet marketplace communications. This enabled the

discovery of key tools and techniques repeatedly associated with high-impact cybercrimes.

- We demonstrated that topic-derived features significantly improve topic quality when evaluated using multiple coherence metrics. Compared to standard LDA, the proposed guided LDA approach produced more semantically coherent and interpretable topics, validating the effectiveness of integrating domain-guided topic-level insights into cyber threat intelligence (CTI).

The structure of this paper is as follows. Section **2** outlines the background of darknet marketplaces and explains their significance within the cybercrime landscape. Section **3** details the dataset and research methodology, covering both the architectural components and the execution process of the proposed framework. Section **4** presents and discusses the experimental results, highlighting the adaptability of the proposed framework for analyzing the Dark Web ecosystem, along with the major findings, their implications, and the limitations of the study. Lastly, Section **5** offers the conclusion and suggests potential directions for future work.

2 Literature Review

This section reviews prior research related to Dark Web analysis and cyber threat intelligence. This section is organized into four main areas: Darknet Marketplaces: Structure and Criminal Ecosystem, Darknet Intelligence Landscape, Challenges in Data Acquisition and Ethical Constraints, and LDA-Based Cyber Threat Intelligence Models. Together, these studies provide insights into the structure, analysis techniques, limitations, and topic-modeling approaches used to understand illicit activities within Dark Web environments.

2.1 Darknet Marketplaces: Structure and Criminal Ecosystem

Darknet marketplaces have evolved into complex criminal ecosystems, enabled primarily by anonymizing tools such as Tor, VPNs, proxy chains, and multi-layer encryption. These technologies conceal user identities and server locations, making surveillance and enforcement extremely challenging (Adebowale, 2025; Jin et al., 2024). The introduction of Silk Road (The first major market) marked a turning point combining Tor, Bitcoin payments, and escrow systems which inspired a surge of similar platforms. Over time, these markets expanded from drug trading to offering stolen data, fraud services, exploit kits, and other cybercrime-related goods, supported by highly active crime-focused forums that allow offenders to exchange knowledge, tactics, and illicit tools (Chiang et al., 2025; Covrig et al., 2022). Because these environments remain hidden from conventional search engines and operate with high anonymity, they provide safe havens for cybercriminal networks to organize illegal operations with minimal risk of identification (Chiang et al., 2025).

Recent research shows growing academic attention toward understanding the Dark Web's structure, user behaviour, and associated risks. Bibliometric studies highlight four major research clusters: network security and attacks, cybercrime and cryptography, machine learning and social-media analysis, and drug trafficking via cryptomarkets (Raman et al., 2023). Scholars also emphasize the importance of advanced data-collection and analytical

methods to map the constantly shifting Dark Web landscape and uncover hidden patterns (de-Marcos et al., 2025). At the same time, studies warn that these illicit ecosystems undermine global security efforts, including UN SDG 16 on peace and justice. As a result, emerging literature calls for interdisciplinary strategies combining digital forensics, blockchain, IoT security, NLP, and cybercrime prevention to develop effective monitoring and mitigation approaches that address the evolving challenges posed by Darknet marketplaces (Raman et al., 2023).

2.2 Darknet Intelligence Landscape

Recent research shows rapid advancements in how intelligence is collected and analyzed from the Dark Web, driven by AI, machine learning, and multi-source data integration. Multi-agent LLM frameworks such as MAD-CTI demonstrate that distributing tasks across specialized agents responsible for crawling, filtering, classifying, and identifying cyber-threat categories significantly improves accuracy and automation in Dark Web monitoring (Shah & Madisetti, 2025). Other studies reinforce that illicit ecosystems like marketplaces, cryptomarkets, and associated discussion forums form a rich source of intelligence for law-enforcement and cybersecurity analysts, offering early indicators of cybercrime activities, trends, and community behavior (Basheer & Alkhatib, 2024). Parallel work has examined how darknet infrastructures themselves operate: for example, combined analyses of Tor and i2p reveal that these networks are structurally interlinked and function as a unified ecosystem, with shared discovery paths and information hubs that help surface hidden services across both environments (Cilleruelo et al., 2021). Together, these findings emphasize that understanding darknet networks, communities, and communication channels is essential for generating timely and actionable cyber threat insights.

Another research direction focuses on building intelligent systems capable of detecting illicit activities and differentiating between types of harmful content. Semi-supervised pipelines, such as the two-phase system using self-training ensembles and enhanced XGBoost models, show strong performance in identifying and categorizing marketplace sales including drugs, weapons, and stolen credentials even when labeled data is scarce. This approach is strengthened by EISA, a meta-heuristic method for feature selection and hyperparameter tuning, enabling high F1-scores and adaptability to multiple platforms such as the deep web, Telegram, Reddit, and Pastebin (Yazdanjue et al., 2025). Complementary investigations highlight emerging sectors such as dark web data markets, where large datasets of listings reveal the scale, structure, legality, and harm associated with illicit data trading through a combined legal and empirical analysis (Covrig et al., 2022). At the network level, machine-learning-based traffic classification models such as stacking ensembles tested on the CICDarknet 2020 dataset achieve up to 97–99% accuracy in identifying malicious darknet traffic, demonstrating their potential for proactive threat detection and attack mitigation (Almomani, 2025). Together, these studies illustrate a rapidly evolving intelligence landscape where AI-driven tools, legal frameworks, and cross-network analyses converge to improve the detection, understanding, and mitigation of darknet-based cyber threats.

2.3 Challenges in Data Acquisition and Ethical Constraints

Collecting reliable data from the Dark Web remains one of the most difficult aspects of cybercrime research due to its anonymity, instability, and technical barriers. Many platforms require specialized access tools like Tor and use layered security measures, making even simple navigation challenging (de-Marcos et al., 2025). Anti-crawling defenses such as CAPTCHAs further limit automated scraping and slow down large-scale data collection (Kühn et al., 2024). Dark Web services frequently change URLs, go offline, or migrate to new hosts, forcing researchers to continually update crawlers and adapt to an unpredictable environment (de-Marcos et al., 2025). The anonymity of Tor and i2p networks also complicates investigations users often operate under multiple pseudonyms, and cryptocurrency transactions are increasingly obfuscated through mixers and rotating wallet addresses, making attribution nearly impossible (Jin et al., 2024). Additionally, much of the Dark Web remains hidden or encrypted, meaning publicly available datasets often represent only a small, skewed sample of the true ecosystem (Adel & Norouzifard, 2024).

Beyond technical barriers, Dark Web research carries significant ethical, legal, and psychological risks. Automated tools used for language detection, cleaning, and network analysis can introduce errors or biased interpretations, which become more pronounced when scaling to larger datasets, multiple platforms, or real-time monitoring (de-Marcos et al., 2025). Legal restrictions also limit how deeply researchers can explore hidden services, as interacting with illegal material whether drugs, weapons, stolen data, or CSAM raises concerns around consent, privacy, and criminal exposure(Adel & Norouzifard, 2024). Investigators must also contend with emotional strain; prolonged exposure to violent or abusive content can cause severe psychological stress, making mental health safeguards essential for analysts working in this domain (Jin et al., 2024). Together, these challenges highlight that Dark Web data acquisition is not only technically demanding but also ethically complex, requiring careful methodological design and strong protective measures for researchers.

2.4 LDA-Based Cyber Threat Intelligence Models

Recent developments in CTI are increasingly expanding beyond the surface web, with researchers applying topic modeling techniques to analyze hidden trends within underground communities (Shah & Madisetti, 2025). LDA and its variants play a central role in revealing patterns in dark web forums, where users discuss criminal activities, hacking tools, and extremist content. Studies show that topic modeling helps uncover latent themes that may not be visible through manual inspection, enabling analysts to interpret emerging threats more systematically. This shift reflects a broader movement toward automating CTI pipelines through NLP-based methods to handle vast and dynamic online ecosystems.

One stream of research focuses on using LDA to understand the structure of online communities and their security implications. An important contribution proposes combining LDA-based topic modeling with a modified network algorithm (SLTA) to detect overlapping sub-communities acknowledging that users often participate in multiple interest groups simultaneously. Tested on a Dark Web forum (Islamic Awakening), this method produces clearer and more meaningful clusters associated with homeland security risks, and encourages

further work on temporal topic evolution and improved topic evaluation (Ríos & Muñoz, 2012). Other studies expand the methodological rigor of topic modeling on Dark Web conversations by comparing multiple models LDA, CTM, PAM, PTM and evaluating them using coherence metrics. This structured pipeline uncovers dominant themes in criminal discussions and offers a replicable framework for analyzing malicious behavior in underground networks (Basheer & Alkhatib, 2024).

Another research direction integrates topic modeling with advanced machine learning and authorship analysis. For example, combining LLDA topic weights with TextCNN drastically reduces feature vectors (by nearly 300×) and boosts classification accuracy, showing the benefit of merging topic models with deep learning for Dark Web text classification. Future work aims to move toward real-time threat detection and unified topic-deep learning architectures (Shin et al., 2024). The same author also introduced a hybrid BERTopic–authorship attribution approach to identify users active across both the surface and dark webs, demonstrating potential for linking suspicious identities despite some overlapping stylistic features. Similarly, topic modeling frameworks such as BERTopic and Top2Vec have been incorporated into the OWASP Maryam OSINT tool to extract CTI from hacker forums in multiple languages. This integration improves intelligence gathering and suggests future enhancements by filtering irrelevant non-security terms (Suryotrisongko et al., 2022). Collectively, these studies show how LDA-based and hybrid topic modeling techniques are becoming essential components of modern CTI research.

Guided Latent Dirichlet Allocation (Guided LDA) strengthens unsupervised topic modeling by embedding domain knowledge through seed words, allowing theory-driven yet data-grounded topic discovery. It has been effectively used to extract meaningful Employee Value Proposition (EVP) dimensions from Glassdoor reviews by steering topics toward factors such as coworker interactions, workplace relationships, and non-monetary benefits, thereby improving interpretability for the hospitality sector (Guo et al., 2025). In disaster informatics, Guided LDA has been applied in a fully automated manner where event-specific seed words were generated by contrasting Twitter vocabularies from normal and disaster periods, enabling real-time detection of disaster-related tweets without manual topic interpretation (Ferner et al., 2020). Similarly, during Hurricane Laura (2020), Guided LDA integrated prior knowledge to uncover situational awareness topics from large-scale Twitter data, revealed the multi-topic nature of tweets, and supported daily visualization and expert validation aligned with official reports for timely disaster response (Zhou et al., 2023).

3 Proposed Method

This study follows a multi-stage analytical framework to model crime facilitation in stolen-financial data darknet markets. The proposed methodology integrates crime-specific filtering, facilitation-oriented feature extraction, Guided Latent Dirichlet Allocation (Guided LDA), and statistical validation using multiple coherence metrics. As shown in Figure 1, the complete workflow consists of six major phases: dataset selection, Stolen- financial Data Filtering, text preprocessing, Extraction of Facilitation Indicators, guided topic modeling, and coherence evaluation and visualization.

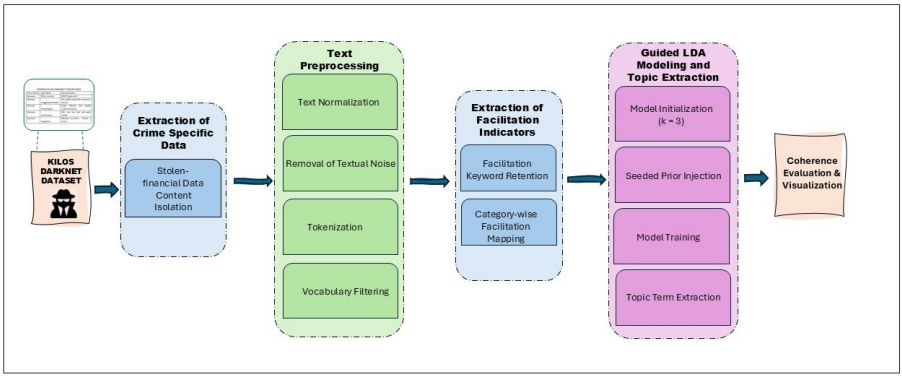

Figure 1: Workflow for Guided Topic Extraction from Dark Web Data

3.1 Dark web Data Selection

The study utilizes an already existing large-scale darknet forum dataset known as the Kilos dataset, which contains user-generated marketplace discussions related to illicit trade activities. The dataset includes forum-level information, user identifiers, timestamps, and textual comments exchanged between buyers and vendors (Branwen, 2013). Since the raw dataset contains heterogeneous discussions, only the most relevant fields required for the present study were selected. Specifically, Forum Name, User Name, and User Comment fields were extracted and retained for further analysis. This ensured that the analytical focus remained on the textual content reflecting illicit operational behavior while preserving contextual and attributional information for potential future extensions.

3.2 Stolen- financial Data Content Isolation

The raw darknet forum dataset contains a wide range of discussions related to multiple illicit activities, including narcotics trade, counterfeit goods, and general marketplace interactions. Since the objective of this study is to analyze stolen- financial data cybercrime, an initial content isolation step was performed to extract only discussions explicitly related to stolen financial and identity information.

A keyword-based filtering approach was adopted due to the absence of labeled data and the highly informal nature of darknet text. A curated list of crime-specific keywords commonly associated with stolen- financial data trading such as *bank, logs, fullz, cvv, card, account, paypal,* and *credentials* was constructed based on prior literature and domain expertise. Each forum comment was converted to lowercase and scanned for the presence of at least one keyword from this list. Comments that contained one or more stolen- financial data indicators were retained, while all others were discarded. As a result of this filtering process, the dataset was substantially reduced in size, isolating a focused subset of crime-relevant discussions. As mentioned in figure 2, from an initial **235,944** darknet forum comments, **28,120** comments were identified as stolen- financial data-related and retained for further analysis, while the

remaining comments were discarded. This output dataset represents a crime-specific textual corpus that forms the foundation for subsequent facilitation analysis and guided topic modeling.

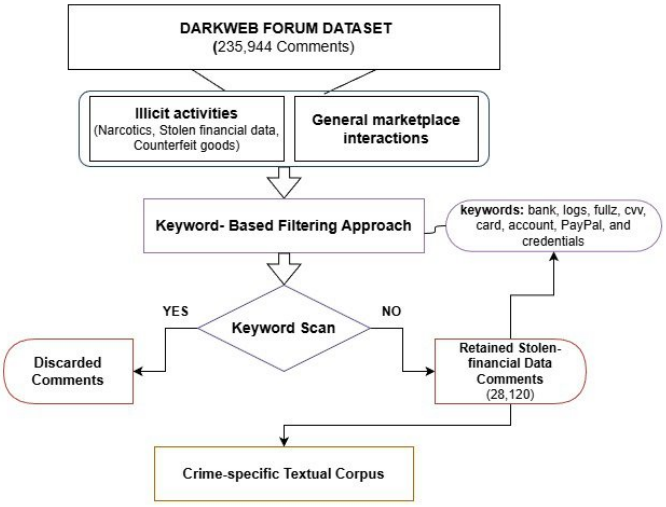

Figure 2: Stolen Data Content Isolation Workflow

3.3 Text Preprocessing

The stolen- financial data-related corpus obtained after key-word based filtering approach was subjected to a comprehensive text preprocessing pipeline to enhance the quality and structure of unstructured darknet forum text and transform it into a standardized, analyzable format (Shin et al., 2024). Technically, each comment was first converted to lowercase to ensure case-insensitive matching and avoid redundant token representations. Regular expression–based filtering was then applied to remove numerals, punctuation, special characters, URLs, and other non-alphabetic symbols commonly present in darknet communications, thereby eliminating textual noise. The cleaned text was subsequently tokenized by splitting each comment into individual word tokens. Following tokenization, vocabulary filtering was carried out through rule-based matching against a predefined keyword set, ensuring retention of only analytically relevant tokens. This preprocessing stage preserved the semantic integrity of crime-related discussions while significantly reducing noise and sparsity. As an outcome, all 28,120 stolen- financial data related comments were converted into clean, normalized, and tokenized textual representations, forming a structured corpus suitable for facilitation indicator extraction and guided topic modeling.

3.4 Extraction of Facilitation Indicators

After preprocessing, the stolen- financial data corpus was examined to identify linguistic indicators that explicitly describe the operational mechanisms enabling stolen- financial data cybercrime (as illustrated in Figure 3). This step aimed to isolate only those comments that

reveal *how* illicit transactions are conducted, rather than *what* is being sold. Facilitation indicators were extracted using a keyword-driven approach grounded in forensic domain knowledge, focusing on payment methods, communication channels, and access infrastructure. This process transforms the cleaned textual corpus into a structured operational dataset suitable for guided topic modeling.

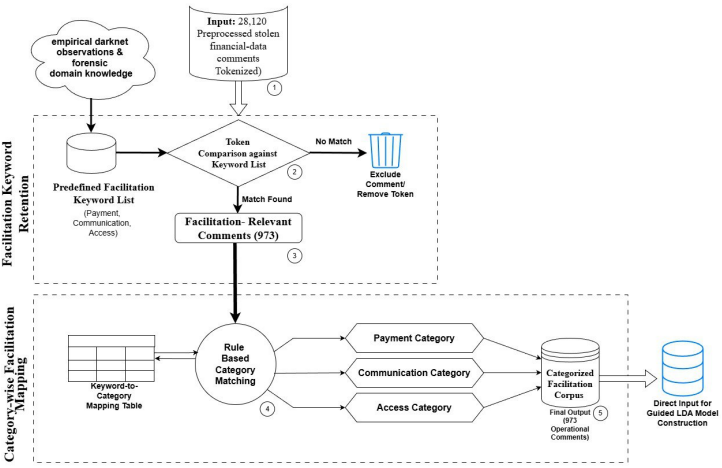

Figure 3: Workflow for the Extraction and Categorization of Facilitation Indicators

3.4.1. Facilitation Keyword Retention

Following text preprocessing, facilitation keyword retention was performed to isolate comments that explicitly describe the operational execution of stolen- financial data cybercrime. Technically, each tokenized comment was compared against a predefined facilitation keyword list encompassing terms related to payment processing (e.g., bitcoin, wallet, escrow), communication methods (e.g., telegram, pgp, signal), and access infrastructure (e.g., vpn, onion, proxy). This keyword set was derived from empirical darknet observations and forensic domain knowledge. During this process, all tokens not matching the facilitation vocabulary were removed. Comments that did not contain at least one facilitation-related keyword were excluded from further analysis, as they lacked operational relevance. As a result, the dataset was substantially reduced from **28,120** pre-processed stolen- financial data comments to **973** facilitation-relevant comments, forming a highly focused corpus that captures explicit criminal facilitation mechanisms and serves as direct input for category-wise mapping and guided topic modeling.

3.4.2 Category-wise Facilitation Mapping

After retaining facilitation-relevant comments, the next objective was to structurally organize the extracted facilitation indicators into meaningful operational categories. While Facilitation Keyword Retention identified *whether* a comment contained facilitation elements, this step focused on determining *which type* of facilitation mechanism was being described. Category-wise facilitation mapping was therefore performed to transform the unstructured facilitation corpus into a structured representation aligned with the core components of stolen-financial data cybercrime operations.

Three facilitation categories were predefined: Payment, Communication, and Access. Each category corresponds to a distinct functional layer of crime execution. Payment-related terms represent financial transaction mechanisms, communication-related terms capture interaction channels between buyers and vendors, and access-related terms describe anonymization tools and infrastructure used to reach darknet platforms. A keyword-to-category mapping table was constructed, where each facilitation keyword was uniquely assigned to one of these categories based on its operational role.

Technically, for each facilitation-retained comment obtained from Facilitation Keyword Retention, the remaining tokens were iteratively examined and mapped to their corresponding facilitation category using rule-based matching. A single comment could contain keywords from one or multiple categories, allowing the representation of multi-faceted crime execution processes within a single transaction. The categorized tokens were stored in structured lists, enabling downstream modeling to capture both individual and co-occurring facilitation mechanisms.

The outcome of this step was a categorized facilitation corpus consisting of **973 operational comments**, where each comment was annotated with one or more facilitation categories. This structured representation explicitly encodes the operational dimensions of stolen-financial data cybercrime and serves as the direct input for the Guided LDA model construction. By introducing category-level organization prior to topic modeling, this step ensures that the discovered topics are both semantically coherent and forensically interpretable.

3.5 Guided LDA Modeling and Topic Extraction

After structuring the facilitation-specific corpus, a Guided Latent Dirichlet Allocation (Guided LDA) model was employed to uncover latent operational patterns underlying stolen-financial data cybercrime. It enables the incorporation of prior knowledge through seeded word–topic priors, allowing the model to converge toward semantically meaningful and operationally interpretable topics rather than depending only on word frequency. In contrast, standard LDA relies solely on probabilistic word distributions, which can lead to the neglect of infrequent but important patterns and the merging of unrelated topics, particularly in imbalanced tweet datasets (Zhou et al., 2023). In this study, the number of topics was fixed at $k = 3$, corresponding to the three predefined facilitation categories: Payment, Communication, and Access.

Technically, the categorized facilitation corpus obtained from Category-wise Facilitation Mapping was provided as input to the Guided LDA model implemented using the Tomotopy framework. Each comment was represented as a bag-of-words document consisting only of facilitation-related tokens. For each topic, representative facilitation keywords were injected as word–topic priors, assigning higher probabilities to category-specific terms and lower probabilities to non-associated topics. This guided the topic learning process toward discovering distinct operational structures while preserving the probabilistic nature of LDA. The model was trained using Gibbs sampling with a predefined burn-in phase to allow stabilization, followed by multiple training iterations until convergence. Upon completion of training, the top-ranked keywords for each topic were extracted based on their posterior probabilities. These dominant terms were analyzed to interpret and label the topics according to their operational significance.

As a result, the Guided LDA model successfully identified three coherent facilitation topics, each dominated by keywords corresponding to one operational dimension. One topic primarily captured financial transaction mechanisms (Payment), another represented communication channels between marketplace participants (Communication), and the third reflected anonymized access infrastructure (Access). These extracted topics form a structured, interpretable representation of the operational workflow of stolen- financial data cybercrime within darknet marketplaces.

3.6 Coherence Evaluation and Visualization

To assess the quality, reliability, and interpretability of the topics generated by the Guided LDA model, a comprehensive coherence evaluation and visualization phase was conducted. Topic coherence measures the degree of semantic similarity among the high-probability words within a topic and is widely used to validate topic modeling results. This step ensured that the extracted facilitation topics were not only statistically sound but also meaningful from an analytical and forensic perspective. Technically, the facilitation-specific corpus used for Guided LDA modeling was converted into a format suitable for coherence computation. A dictionary of unique facilitation terms and a corresponding bag-of-words corpus were constructed using the Gensim framework. The top-ranked keywords extracted from each of the three topics in the Guided LDA Modeling and Topic Extraction step were then supplied as input to the coherence evaluation module.

Four standard coherence metrics were computed to provide a balanced and robust evaluation. The C_V coherence metric was used as the primary measure due to its strong correlation with human interpretability, capturing semantic similarity using a sliding window and normalized pointwise mutual information. The UMass metric evaluated topic coherence based on document co-occurrence probabilities, serving as an internal consistency check. The C_UCI metric employed a pointwise mutual information–based approach to assess statistical association between topic words, while C_NPMI provided a normalized coherence score enabling comparability across topics and datasets.

In addition to numerical evaluation, visual analysis was performed to support qualitative interpretation of the results. Keyword distributions and topic-wise term frequencies were

visualized to examine dominance patterns and overlap across facilitation categories. These visual representations aided in validating the distinctiveness of the Payment, Communication, and Access topics and facilitated intuitive understanding of their relative importance. The coherence evaluation demonstrated that the Guided LDA model produced semantically consistent and operationally meaningful topics. The combined use of multiple coherence measures and visualization techniques provided strong validation of the extracted facilitation structures.

4 Results and Discussion

This section presents the experimental results obtained from the proposed Guided LDA-based crime facilitation modeling framework and discusses their forensic and analytical implications. The objective of the study was to uncover the hidden operational infrastructure supporting stolen- financial data cybercrime activities within darknet marketplaces by modeling payment, communication, and access mechanisms.

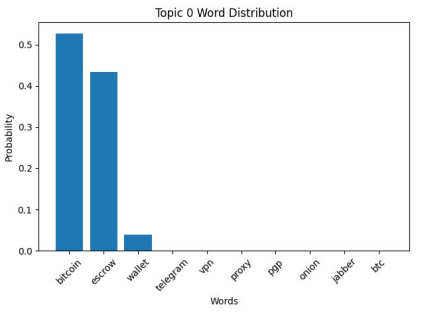

Figure 4(a): Payment Facilitation Topic

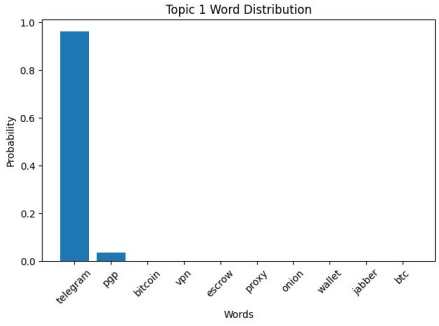

Figure 4(b): Communication Facilitation Topic

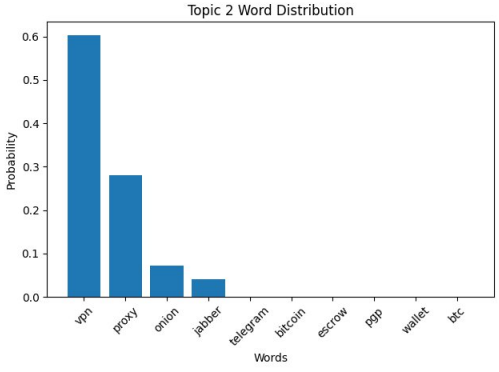

Figure 4(c): Access Facilitation Topic

Figure 4: Top-Word Probability Distribution for Guided LDA Topics

The Guided LDA model successfully extracted three distinct and semantically well-separated topics corresponding to the predefined facilitation classes. The topic–word probability distributions (Figures 4) show a strong correspondence between seeded priors and the learned latent topics. The Payment topic (Figure 4(a)) is dominated by *bitcoin* and *escrow*, reflecting the central role of cryptocurrency-based financial transactions. The Communication topic (Figure 4(b)) is primarily represented by *telegram,* and *pgp*, confirming the reliance on encrypted and privacy-preserving communication channels. The Access topic (Figure 4(c)) is characterized by *vpn, proxy,* and *onion*, highlighting the extensive use of anonymization infrastructures for darknet entry. Topic quality was quantitatively validated using multiple coherence metrics. The C_V score of 0.635 indicates strong semantic coherence, while UMass, UCI, and NPMI values fall within acceptable ranges for sparse guided topic models, confirming the statistical reliability and interpretability of the discovered topics.

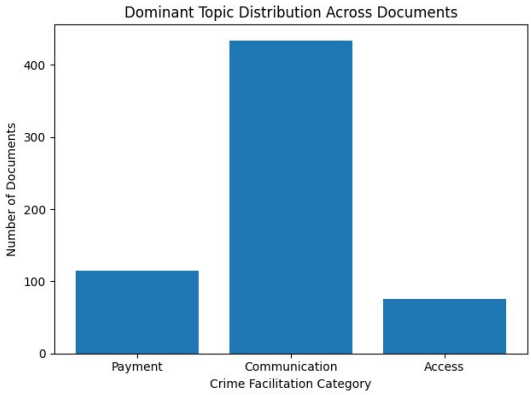

Figure 5: Global Distribution of Crime Facilitation Topics in Stolen- financial Data Posts

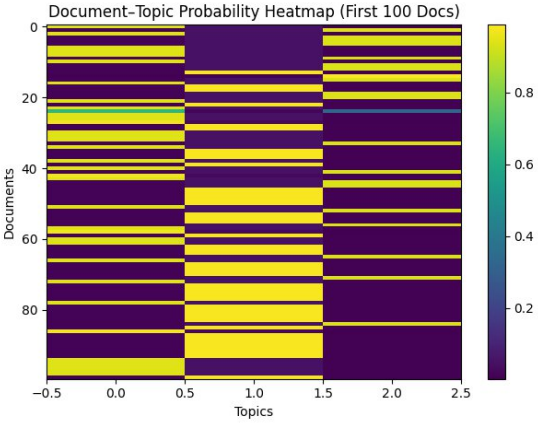

Figure 6: Document–Topic Probability Heatmap for Guided LDA Model

The global topic distribution across all documents (Figure 5) demonstrates that Communication is the most dominant facilitation activity, followed by Access, with Payment appearing least frequent. This imbalance indicates that a substantial portion of stolen- financial data marketplace interactions is devoted to negotiation, coordination, and trust establishment rather than direct monetary exchange. Such behavior reflects the operational reality of darknet economies, where financial details are often shifted to private encrypted channels after initial contact. The document-level topic probability heatmap (Figure 6) further reveals that most documents are strongly dominated by a single topic, indicating high topical purity at the post level. This confirms that darknet communications are typically task-specific, focusing explicitly on one operational stage at a time, which enhances the effectiveness of automated topic-driven surveillance and alerting systems.

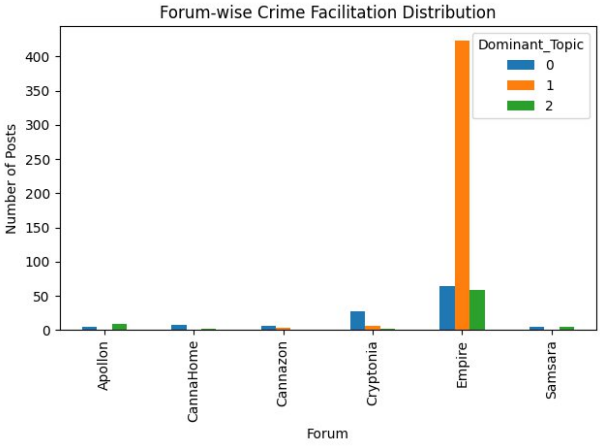

Figure 7: Forum-wise Distribution of Crime Facilitation Categories

The forum-wise topic distribution (Figure 7) provides important structural insights into platform-level specialization within the darknet ecosystem. Certain forums are heavily dominated by Communication-related activity, indicating their function as coordination hubs, while others exhibit higher proportions of Access or Payment topics, suggesting their roles as gateway or transaction-centric platforms. This non-uniform distribution demonstrates that stolen- financial data cybercrime operates through a functionally distributed ecosystem rather than isolated marketplaces. Overall, the combined evidence from guided topic extraction, coherence validation, and multi-level visual analysis confirms that the proposed framework is both statistically robust and operationally meaningful. The results establish the framework's utility for cyber forensic investigations by enabling structured profiling of criminal facilitation mechanisms, platform-level intelligence extraction, and scalable monitoring of underground cybercrime operations.

5 Conclusion and Future Work

This study presented a Guided Latent Dirichlet Allocation (Guided LDA) based forensic intelligence framework for modeling crime facilitation mechanisms in stolen- financial data darknet marketplaces. By integrating domain-driven keyword filtering, guided topic seeding, probabilistic topic modeling, and multi-metric coherence validation, the proposed approach successfully identified three critical operational dimensions of cybercrime: Payment, Communication, and Access. Experimental results demonstrated strong semantic consistency and statistical reliability, with coherent topic formation and clear separation across all levels of analysis. The multi-layered visual analytics further confirmed the practical interpretability of the extracted crime infrastructure, enabling global, document-level, and forum-specific intelligence extraction. These findings validate the effectiveness of guided topic modeling as a scalable and explainable tool for darknet cybersecurity forensics.

Despite its strong performance, the proposed framework exhibits certain limitations. The analysis was restricted to lexically filtered stolen-financial data content, which may exclude implicit or evolving criminal jargon not covered by the predefined keyword list. Additionally, the use of a static topic model limits temporal adaptability in detecting emerging facilitation patterns. The framework also focuses exclusively on textual data and does not incorporate network, transactional, or multimedia evidence, which are increasingly relevant in contemporary cybercrime investigations.

Future work will focus on enhancing the proposed system along multiple dimensions. First, dynamic topic modeling and concept drift detection will be incorporated to track the evolution of cybercrime processes over time. Second, the integration of deep learning–based contextual embeddings (e.g., BERT-based topic modeling) will be explored to improve semantic generalization beyond keyword-level filtering. Third, the framework will be extended to support multilingual darknet content using automatic translation and cross-lingual topic alignment. Finally, the proposed model will be deployed as a real-time cybercrime monitoring and alerting system integrated with threat intelligence feeds and law-enforcement investigation pipelines, enabling proactive detection of emerging illicit trading behaviors.

References

Adebowale, J. (2025). Overview of Internet Anonymizers — Vpns, Tor, Proxy Chains, Encryption Layers, and Darknet Access.

Adel, A., & Norouzifard, M. (2024). Weaponization of the Growing Cybercrimes inside the Dark Net: The Question of Detection and Application. *Big Data and Cognitive Computing*, *8*(8), 91. https://doi.org/10.3390/bdcc8080091

Almomani, A. (2025). Darknet traffic analysis, and classification system based on modified stacking ensemble learning algorithms. *Information Systems and E-Business Management*, *23*(1), 209–240. https://doi.org/10.1007/s10257-023-00626-2

Basheer, R., & Alkhatib, B. (2024). Conceptualizing Discussions on the Dark Web: An Empirical Topic Modeling Approach. *Complexity*, *2024*, 1–24. https://doi.org/10.1155/2024/2775236

Branwen G, Christin N, Décary-Hétu D *et al Dark Net Market archives, 2011–2015.* Available online: *https://www.gwern.net/DNM-archives*

Chiang, E., Kredens, K., & Thornton, J. (2025). Fighting fraud: Corpus-assisted approaches to understanding and disrupting fraud activity on the dark web. *Applied Corpus Linguistics*, *5*(3), 100159. https://doi.org/10.1016/j.acorp.2025.100159

Cilleruelo, C., de-Marcos, L., Junquera-Sánchez, J., & Martínez-Herráiz, J.-J. (2021). Interconnection between darknets. *IEEE Internet Computing*, *25*(3), 61–70. https://doi.org/10.1109/MIC.2020.3037723

Covrig, B., Mikelarena, E. B., Rosca, C., Goanta, C., Spanakis, G., & Zarras, A. (2022). Upside Down: Exploring the Ecosystem of Dark Web Data Markets. In W. Meng, S. Fischer-Hübner, & C. D. Jensen (Eds.), *ICT Systems Security and Privacy Protection* (Vol. 648, pp. 489–506). Springer International Publishing. https://doi.org/10.1007/978-3-031-06975-8_28

de-Marcos, L., Domínguez-Díaz, A., Junquera-Sánchez, J., Cilleruelo, C., & Martínez-Herráiz, J.-J. (2025). Unveiling Dark Web Identity Patterns: A Network-Based Analysis of Identification Types and Communication Channels in Illicit Activities. *Information*, *16*(11), 924. https://doi.org/10.3390/info16110924

Devarajan, S., Panneerselvam, P., Mudigonda, A., & Hemalatha, P. K. (2024). Enhancing Dark Web Classification: A Dynamic Crawler and Robust Classification Framework. *International Journal of Intelligent Systems and Applications in Engineering*, *12*(6s), 01–09.

Ferner, C., Havas, C., Birnbacher, E., Wegenkittl, S., & Resch, B. (2020). Automated Seeded Latent Dirichlet Allocation for Social Media Based Event Detection and Mapping. *Information*, *11*(8), 376. https://doi.org/10.3390/info11080376

Gond, A. K. (2025). Hidden Side of the Internet: Between Crime and Freedom on the Dark-Web. *Journal of Communication and Management*, *4*(03), 55–61. https://doi.org/10.58966/JCM2025436

Guo, Y., Ayoun, B., & Zhao, S. (2025). Is someone listening to me? A topic modeling approach using guided LDA for exploring hospitality value proposition through online employee reviews. *International Journal of Hospitality Management*, *127*, 104114. https://doi.org/10.1016/j.ijhm.2025.104114

Jin, P., Kim, N., Lee, S., & Jeong, D. (2024). Forensic investigation of the dark web on the Tor network: Pathway toward the surface web. *International Journal of Information Security*, *23*(1), 331–346. https://doi.org/10.1007/s10207-023-00745-4

Kühn, P., Wittorf, K., & Reuter, C. (2024). Navigating the Shadows: Manual and Semi-Automated Evaluation of the Dark Web for Cyber Threat Intelligence. *IEEE Access*, *12*, 118903–118922. https://doi.org/10.1109/ACCESS.2024.3448247

Owen, A., Dawson, G., & Reed, O. (2025). *Dark Web Marketplaces and the Role of Cryptocurrency in Criminal Financial Transactions*.

Raman, R., Kumar Nair, V., Nedungadi, P., Ray, I., & Achuthan, K. (2023). Darkweb research: Past, present, and future trends and mapping to sustainable development goals. *Heliyon*, *9*(11), e22269. https://doi.org/10.1016/j.heliyon.2023.e22269

Ríos, S. A., & Muñoz, R. (2012). Dark Web portal overlapping community detection based on topic models. *Proceedings of the ACM SIGKDD Workshop on Intelligence and Security Informatics*, 1–7. https://doi.org/10.1145/2331791.2331793

Santos, P., Abreu, R., Reis, M. J. C. S., Serôdio, C., & Branco, F. (2025). A Systematic Review of Cyber Threat Intelligence: The Effectiveness of Technologies, Strategies, and Collaborations in Combating Modern Threats. *Sensors*, *25*(14), 4272. https://doi.org/10.3390/s25144272

Shah, S., & Madisetti, V. K. (2025). MAD-CTI: Cyber Threat Intelligence Analysis of the Dark Web Using a Multi-Agent Framework. *IEEE Access*, *13*, 40158–40168. https://doi.org/10.1109/ACCESS.2025.3547172

Shin, G.-Y., Jang, Y., Kim, D.-W., Park, S., Park, A.-R., Kim, Y., & Han, M.-M. (2024). Dark Side of the Web: Dark Web Classification Based on TextCNN and Topic Modeling Weight. *IEEE Access*, *12*, 36361–36371. https://doi.org/10.1109/ACCESS.2023.3347737

Suryotrisongko, H., Ginardi, H., Ciptaningtyas, H. T., Dehqan, S., & Musashi, Y. (2022). Topic Modeling for Cyber Threat Intelligence (CTI). *2022 Seventh International Conference on Informatics and Computing (ICIC)*, 1–7. https://doi.org/10.1109/ICIC56845.2022.10006988

Temara, S. (2024). *The Dark Web And Cybercrime: Identifying Threats And Anticipating Emerging Trends*. Computer Science and Mathematics. https://doi.org/10.20944/preprints202410.0147.v1

Watanabe, K., & Baturo, A. (2024). Seeded Sequential LDA: A Semi-Supervised Algorithm for Topic-Specific Analysis of Sentences. *Social Science Computer Review*, *42*(1), 224–248. https://doi.org/10.1177/08944393231178605

Yazdanjue, N., Rakhshaninejad, M., Yazdanjouei, H., Niemelä, M. S., Chen, F., & Gandomi, A. H. (2025). Cyber threat management using semi-supervised ensemble learning and enhanced interior search algorithm: Applications for illicit marketplace classification in deep/dark web and social platforms. *Annals of Operations Research*. https://doi.org/10.1007/s10479-025-06651-3

Zhang, X., Li, P., Tong, X., Chow, K. P., Yuan, D., Chen, S., & Jin, B. (2025). Decoding the shadows: Multi-modal identity profiling in darknet markets using latent behavior feature fusion. *Discover Applied Sciences*, *7*(7), 656. https://doi.org/10.1007/s42452-025-07217-5

Zhou, S., Kan, P., Huang, Q., & Silbernagel, J. (2023). A guided latent Dirichlet allocation approach to investigate real-time latent topics of Twitter data during Hurricane Laura. *Journal of Information Science*, *49*(2), 465–479. https://doi.org/10.1177/01655515211007724