


## Article

# Design of a Trusted Content Authorization Security Framework for Social Media

Jiawei Han <sup>1,2,\*</sup>, Qingsa Li <sup>1</sup> , Ying Xu <sup>3</sup>, Yan Zhu <sup>1</sup> and Bingxin Wu <sup>1</sup>

<sup>1</sup> College of Cyber Security, Changchun University, Changchun 130022, China; lqs20000226@163.com (Q.L.); zhuyannm@126.com (Y.Z.); wbx2465851520@163.com (B.W.)

<sup>2</sup> Digital Identity and Blockchain Joint Laboratory, Peking University, Beijing 100871, China

<sup>3</sup> School of Administration, Changchun University, Changchun 130022, China; yingzi.xu@126.com

\* Correspondence: jason.hjw@gmail.com

**Abstract:** Artificial intelligence-generated content (AIGC) technology has had disruptive results in AI, representing a new trend in research and application and promoting a new era of AI. The potential benefits of this technology are both profound and diverse. However, the benefits of generative tools are accompanied by a series of significant challenges, the most critical of which is that it may cause AI information pollution on social media and mislead the public. Traditional network security models have shown their limitations in dealing with today's complex network threats, so ensuring that generated content published on social media accurately reflects the true intentions of content creators has become particularly important. This paper proposes a security framework called "secToken". The framework adopts multi-level security and privacy protection measures. It combines deep learning and network security technology to ensure users' data integrity and confidentiality while ensuring credibility of the published content. In addition, the framework introduces the concept of zero trust security, integrates OAuth2.0 ideas, and provides advanced identity authentication, fine-grained access control, continuous identity verification, and other functions, to comprehensively guarantee the published content's reliability on social media. This paper considers the main issues of generative content management in social media and offers some feasible solutions. Applying the security framework proposed in this paper, the credibility of generated content published on social media can be effectively ensured and can help detect and audit published content on social media. At the operational level, when extracting key information summaries from user-generated multimodal artificial intelligence-generated content and binding them to user identity information as a new token to identify user uniqueness, it can effectively associate user identity information with the current network status and the generated content to be published on the platform. This method significantly enhances system security and effectively prevents information pollution caused by generative artificial intelligence on social media platforms. This innovative method provides a powerful solution for addressing social and ethical challenges and network security issues.

**Keywords:** AIGC; identity security; network security model; token transfer



**Citation:** Han, J.; Li, Q.; Xu, Y.; Zhu, Y.; Wu, B. Design of a Trusted Content Authorization Security Framework for Social Media. *Appl. Sci.* **2024**, *14*, 1643. <https://doi.org/10.3390/app14041643>

Academic Editors: Mohamed Amine Ferrag, Leandros Maglaras and Helge Janicke

Received: 23 January 2024

Revised: 14 February 2024

Accepted: 17 February 2024

Published: 18 February 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Since the official launch of the generative artificial intelligence model represented by ChatGPT (Chat Generative Pre-trained Transformer) in November 2022 [1], artificial intelligence has greatly contributed to our daily lives and promoted the rapid development of the digital economy. Traditional artificial intelligence (AI) mainly focuses on specific fields, such as data extraction and prediction. In contrast, generative AI models can produce new artificial intelligence-generated content (AIGC). However, this also raises security challenges in multiple respects, such as technology abuse represented by content security and user identity privacy. The Internet information space has long faced challenges of fake information and information content security [2]. Domestic and foreign social platforms [3],

such as Facebook, Twitter, WeChat, Weibo, etc. [4], have continuously improved their governance ability for fake content and information security [5]. With the continuous growth in AIGC volume, the challenges of counterfeit information and information content security are also increasing. In recent years, the open source of domestic large language models has enabled the general public to train personalized large language models with less computing power, which has significantly reduced the threshold for using generative artificial intelligence. This technology is widely used in network platforms to generate short videos, articles, and AI graphics automatically. In today's society, short video software has increasingly begun to adopt generative artificial intelligence technology. These applications use productive artificial intelligence technology to create special effects, allowing users to easily add and edit high-quality effects, further enhancing the appeal of their short videos. However, this has brought much AIGC (artificially generated content) into social media and has spawned many problems [6], including forgery, false information dissemination, moral and ethical issues, excessive attention to inappropriate content, and infringement. Therefore, we need to face the challenges AIGC brings [7]. These problems go beyond the ability of traditional content risk control to address. Illegal actors can use open-source AIGC models or tools to generate a wide range of false information with lower thresholds and higher efficiency, infringing copyright [8] and quickly stealing user identities [9], representing new fraudulent and other illegal activities. In addition, since the training data of AIGC models mainly come from the Internet, the authenticity and accuracy of AI-generated information cannot be guaranteed. Privacy leakage issues have occurred in GPT-2, indicating that personal privacy data have been included in the model training dataset. In this era of digital information explosion, information authenticity, accuracy, and compliance have become urgent problems [10]. Effective control measures will make measuring the potential impact of AIGC on society, individuals, and social media easier. Therefore, this paper proposes an authorization security framework called "spoken", which includes a complete and adequate set of content information control strategies to ensure the security of user identity and the authenticity of the published content to reflect the user's original intention. This method provides a solid and reliable security solution for social media [11]. The contributions of this paper are as follows:

- (1) This study considers a series of issues brought about by the rapid development of artificial intelligence-generated content (AIGC), focusing on managing published content in social media. The research seeks practical solutions to address these issues, protect personal rights, and maintain data privacy.
- (2) To ensure the uniqueness of user identity, this framework extracts the multimodal information users publish on social media. It integrates it with user identity information, timestamp, address label, and network IP address information into a new token. By verifying the token's digital signature, the integrity of the content spread in social media can be effectively ensured. Only by comparing the hash value in the token with the recalculated hash value can the file be confirmed not to have been tampered with.
- (3) In the proposed framework, the digital signature, OAuth2.0, and other mechanisms are adopted, and the concept of zero trust security is introduced to construct a reliable authentication and authorization system to ensure the security of user identity and content.

## 2. Related Work

To solve the multi-faceted problems of AIGC content, including forgery, false information dissemination, ethical issues, excessive attention to inappropriate content, and infringement, Tencent Security provides a full-chain AIGC content security solution. The solution covers four core technologies: review services, security expert consultation, automatic machine review, and copyright content protection. It covers the whole process from model training to content generation to subsequent operational processes, supporting the realization of integrated content security. However, this solution for AI internal operation makes it difficult for content publishers to intervene in the model. In contrast, the Baidu

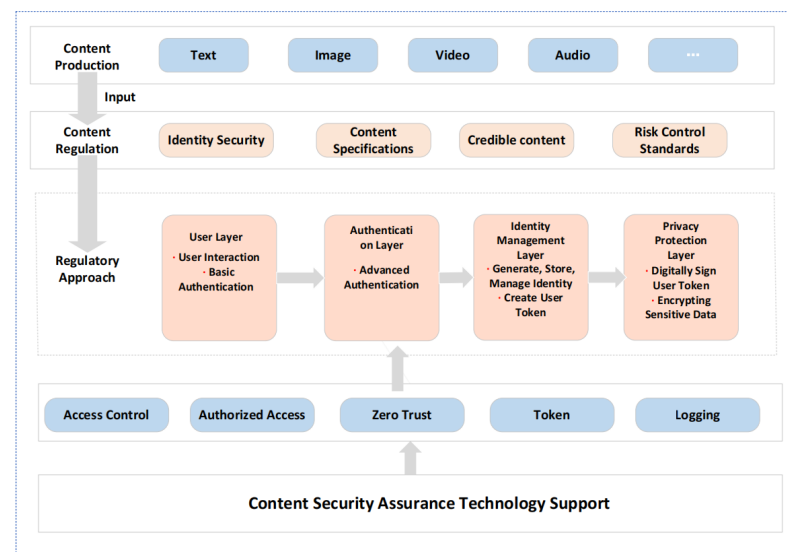
AI open platform has launched a comprehensive content review product, supporting the review of various content formats, such as pictures, text, audio, and video, which helps improve review efficiency and reduces the cost of manual review. However, there is still a potential problem—user identity information (such as timestamp, address label, network IP address, and upcoming information) must be bound and compared in real time, which may affect the accurate assessment of content credibility [12]. In the field of generative artificial intelligence, due to its creative nature, AIGC often presents the characteristics of “fiction”, and even experts find it difficult to clearly distinguish between AI-generated and original works [13,14]. If content security cannot be ensured, this may lead to major problems in the future. Therefore, some researchers, such as Roy et al., have adopted various methods to identify AIGC on social media to prevent the information pollution caused by generative artificial intelligence on social media platforms [15,16]. They focus on dealing with the spread of false news on social media. An automated model is proposed to identify and prevent the spread of false news at an early stage. The model adopts machine learning methods, including random forest, logistic regression, naive Bayes, and K-means as baseline models, and introduces the long short-term memory (LSTM) network based on deep learning. The paper emphasizes the harm of false news spread and the importance of automated models in dealing with this problem. Although the research proposes an automated model that can recognize false news, more is needed to solve the content credibility problem caused by identity security problems in social media. On social media platforms, false information may be published by anonymous or camouflaged users, making it impossible to determine the sender’s identity even if a piece of content is identified as false, and it cannot completely solve the problem of information spread. Therefore, to address the spread of false information, further research and discussion are still needed on how to solve the problem of identity authentication on social media effectively. To solve this problem more comprehensively, this paper proposes a more detailed and comprehensive security framework. The framework adopts multi-level security and privacy protection measures, combined with deep learning and network security-related technologies, to ensure user data security and the credibility of published content. In addition, the framework introduces the concept of zero trust security, provides advanced identity authentication, integrates OAuth2.0, realizes fine-grained access control and continuous identity verification, and comprehensively guarantees the credibility of published content on social media [17]. The proposed framework will help strengthen the effective management of user identity and content and provide a solid foundation for coping with the ever-evolving content security challenges.

### 3. SecToken Framework

Considering the multimodal characteristics of content published on social media, this study covers text, audio, and video information. It aims to extract a summary of users’ posts on social media. Firstly, a network security model based on network security technology and deep learning is constructed; its overall framework is shown in Figure 1.

In the initial stage of the model, access tokens are obtained by adopting the OAuth2.0 [18] authorization code mechanism, and the OpenID Connect (OIDC) authentication process is executed in the authorization server. After completing biometric double authentication, the system obtains the user’s basic identity information. Subsequently, the multimodal content to be published is extracted at the identity management layer, and the user’s current timestamp, address label, and network IP address information are collected simultaneously. This real-time and identity information is packaged together to form a user token and digital signatures are performed on it. By verifying the digital certificate of the user token, it can be determined whether the user themselves publishes the content and whether it has been tampered with during the transmission process on the platform. Since the user token contains the user’s identity information, current geographical location, network IP address, and the summary of published information, it is more secure than a

single identity information binding method. Therefore, this security mechanism can ensure that the content posted on the platform truly reflects the intention of the content publisher.



**Figure 1.** Layered structure of SecToken security framework.

To achieve fine-grained authorization and access process separation, as well as the function of authorizing multiple accesses at once, this paper develops a multi-factor digital token representing data user access rights based on the data structure of JWT (JSON Web Tokens) for access control. The token is called the user token, and its structure is shown in Table 1.

**Table 1.** User token structure.

Field	Name	Description
Header	type	Parameter type
	alg	Signature algorithm
	userIss	Issuer
PayLoad	sub	Abstract
	Exp	Expiration date
	situation	Time stamp, address label, network IP
Sign	secret	Use base64 encoding and add encryption key storage to the server.

The user token designed in this paper contains three fields: Header, Payload, and Sign. In the Header, the token type is JWT, and the signing algorithm uses HS256. The Payload field includes four parts:

- UserIss (publisher).
- Sub (content topic and abstract).
- Exp (Token life cycle).
- Situation (publisher timestamp, address label, network IP information).

In the Sign field, the Base64 encoded Header and Payload are signed with the secret key stored on the server side using an encryption algorithm. The private key is securely stored on the server side. This design provides a certain degree of security, verifies the integrity and authenticity of the token through the signing mechanism, and can be used for information transmission, authentication, and authorization scenarios. At the same time, it is necessary to ensure the secure storage of the secret key and adopt reliable encryption algorithms to ensure the security of the token.

### 3.1. Integration Process and Challenges

Firstly, the location and interface of the SecToken framework in the social media platform need to be determined. Then, the API interface for interacting with the social media platform needs to be designed and developed to realize the functions of the SecToken framework. Secondly, authentication and authorization are required to implement user authentication and authorization mechanisms to ensure that only authorized users can access the functions of the SecToken framework. Finally, data transmission and encryption are required to implement secure data transmission and encryption mechanisms to protect sensitive information transmitted between the social media platform and the SecToken framework. There are multiple challenges in the implementation of the SecToken framework. Firstly, the security challenge is to ensure the framework's security to prevent malicious attacks and data leakage. Secondly, the compatibility challenge is to ensure that the framework can be compatible with different versions and types of social media platforms, and the integration with other security frameworks and components needs to be compatible. The performance challenge is also an important consideration, with the need to ensure that the framework will not negatively impact the social media platform's performance and user experience, including the response time and processing capacity. Finally, the accuracy challenge also needs to be paid attention to, that is, to ensure the accuracy of content judgment and to continuously improve the technical level to ensure its accuracy.

Although the described security mechanisms have achieved particular success in ensuring the authenticity of content posted on social media and the identity of users, there are still some potential areas for improvement. Firstly, collecting personal information, such as user timestamps, address tags, and network IP addresses, may raise privacy issues, and this information needs to be handled more carefully to protect user privacy. Secondly, although technologies such as double authentication and digital signatures enhance security, a balance must be found between security and user experience to ensure that the system is secure and easy to use. In addition, further research is needed to address the potential risks of cyber attacks and digital signature cracking. Data protection and compliance are also key issues to ensure that relevant regulations are followed when processing user information and content summaries. Finally, with the evolution of technology, security mechanisms need to be constantly improved and updated to adapt to new security challenges and threats. Therefore, it is necessary to seek more comprehensive and effective solutions by considering user privacy, the balance between security and convenience, cyber attack response, data protection compliance, and continuous improvement.

### 3.2. Identity Management

The primary task of the identity management layer is to extract key information summaries from the generated multimodal content (including text, audio, and video). For audio content, appropriate feature extraction methods, such as mel frequency cepstral coefficients (MFCCs), are needed to capture the spectrum information of the sound and convert it into a set of real number vectors with fixed length. Bidirectional LSTM is adopted for text content to introduce an attention mechanism to understand the text and focus on crucial information more comprehensively. After optimizing the training data to improve the model and increase accuracy, the model can be applied to generate summaries. For video content, the main task is to extract critical frames that represent the main content of the video and use ResNet to convert each keyframe into a feature vector with a fixed length [19]. At the same time, user information processing is carried out, including embedding vectorization processing of elements, such as timestamps, address labels, and network IP addresses, and merging each keyframe and audio feature vector into a user token comprehensive feature vector, which contains multimodal content and user information summary information. Finally, standardization is carried out to ensure consistency.



### 3.3. Privacy Layer

In the data privacy protection layer, when a user publishes content, the content publisher will use its private key to sign the user token digitally. Only the legal holder of the private key can generate a valid digital signature. This digital signature is attached to the text file or stored in the metadata of the audio and video files together with the content. When the receiver obtains the content, it conveys the digital certificate and the certificate authority's public key (CA). To ensure the authenticity of the digital certificate, the receiver will first use the CA's public key to verify the digital certificate. This step confirms that a trusted CA issues the digital certificate and establishes a trust relationship with the digital certificate. Next, the receiver will use the content publisher's public key to verify the digital signature attached to the content to determine whether the content is complete and from a trusted source. Through such a verification process, it can be ensured that the receiver can confirm that the received content has not been tampered with and is generated by a legitimate publisher during transmission. This complex and rigorous process guarantees data integrity and authenticity and only allows authorized users to access, verify, and trust the published content. Combining digital certificates and signatures can enhance data privacy protection level security, preventing unauthorized personnel from tampering with or accessing data.

### 3.4. Summarizing

In summary, the security framework proposed in this paper differs significantly from traditional network security frameworks regarding performance, scalability, and risk reduction of AI-generated content, as shown in Table 2. The significant improvements in various aspects of the proposed framework in this study are presented by comparing the performance, scalability, and effectiveness of reducing the risk of AI-generated content between the SecToken framework and traditional security frameworks.

**Table 2.** SecToken vs. traditional frameworks.

Angle	SecToken Framework	Traditional Framework
Performance	Combined with deep learning, it has high performance.	Based on rules and signatures, the performance of processing multimodal data is low.
Scalability	Easy to expand and integrate new authentication mechanisms.	Relatively fixed, and integrating new technologies is relatively cumbersome.
Reduce the effectiveness of AI-generated content	Using digital signatures and other methods can effectively reduce the risk of fraud.	Lack of specialized mechanisms to combat the risk of AI-generated content.

## 4. Safety Analysis and Experimentation

### 4.1. Security Analysis

The security framework covers multiple vital layers. Firstly, a user-friendly authentication interface is provided at the user interface layer, using the traditional user name and password method. However, this method has potential risks, such as password cracking, social engineering, and phishing attacks. The identity authentication layer introduces advanced authentication mechanisms to address these risks, combining passwords and biometric recognition technology (especially audio recognition) to improve strictness. However, biometric recognition algorithms must be updated and enhanced regularly because biometric characteristics may be threatened by forgery or simulation attacks. At the identity management layer, the user token is used to achieve the uniqueness of digital identities,

and digital signature technology is used to verify the integrity and authenticity of digital identities. The security of user token generation algorithms and the digital signature process should be paid attention to prevent private key leakage and signature vulnerabilities. Digital signature technology is used to sign content at the privacy protection layer. At this level, remote key management and protection are essential, and effective measures must be taken to prevent private key leakage. The user token is used as a token for user authentication and authorization, and a zero trust security framework is implemented at the access control and authorization layer. The secure transmission and storage of tokens must be ensured and appropriate token management mechanisms adopted to prevent criminals from abusing tokens. Finally, regarding security monitoring and threat detection, the system records user operation logs, detects abnormal behaviors and malicious attacks, and takes timely response measures. To ensure that the monitoring system is not attacked, threat detection rules are regularly updated and adequate measures are taken to ensure the security of the monitoring system.

#### 4.2. Experiments

For the experimental part, the process of conducting token generation and the detailed method of extracting multimodal published content summaries are described. In addition, we also describe the process of generating multi-factor tokens. In further experiments, we conduct multi-stage testing and evaluation of the user token generation and verification process. Firstly, we simulate user authentication requests in different scenarios, including during regular operation, abnormal operation, and malicious attack. Under regular operation, we evaluate the speed and accuracy of user token generation and ensure that the system can generate tokens timely and effectively. For abnormal operations and malicious attacks, we test the fault tolerance and security of the system and verify whether it can correctly identify and reject malicious requests to protect the security of user identity information. Secondly, we thoroughly test the biometric recognition module and evaluate the reliability and stability of different biometric features regarding recognition accuracy, speed, and anti-attack ability in practical applications. Finally, we conduct large-scale testing using simulated and real data to evaluate the performance of the system under different scales. These experimental details can comprehensively demonstrate the effectiveness and potential risks of the system in user authentication and identity management and provide essential references for further optimization and improvement.

##### Token Generation

##### (1) Extracting critical information summary from audio

The following processing flow can be operated for the generated audio content: Firstly, the SeamlessM4T [20] model transcribes the input WAV audio file into text form. Next, a series of preprocessing steps are performed on the generated text, including word segmentation and stop word removal, to better prepare the data. Then, a sequence-to-sequence model is established and a bidirectional LSTM layer is added. By designing a bidirectional LSTM layer the input sequence is processed in two directions to capture contextual information comprehensively [21]. We introduce an attention mechanism to improve the model's focus, which helps focus on the crucial details [22]. The model diagram is shown in Figure 2. In the training phase, we use data with abstract labels and train the model by optimizing the loss function for sequence generation tasks. Once the training is completed, the model can generate abstracts and obtain corresponding abstract sequences through forward propagation. Finally, we perform post-processing steps, such as removing unnecessary punctuation, to improve the overall quality of the abstract and ultimately obtain the text abstract of the input audio content.

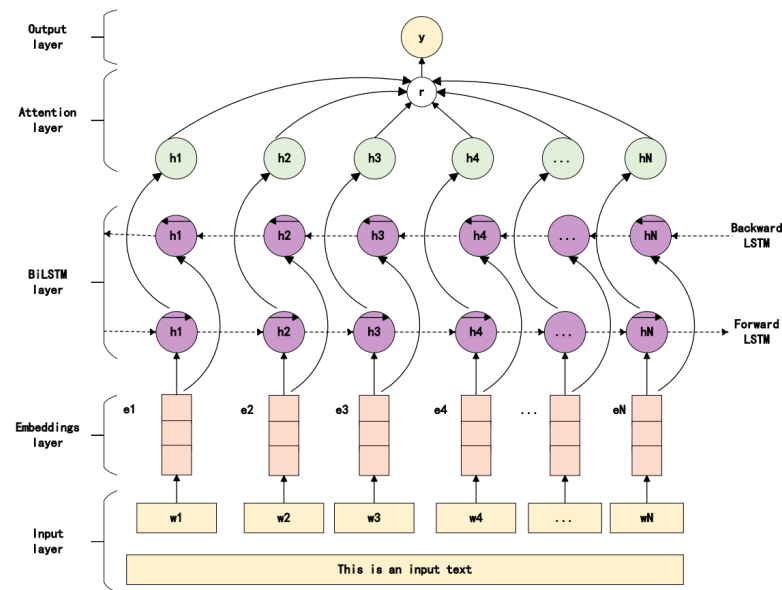


Figure 2. The Att-BiLSTM model architecture.

## (2) Extracting critical information summary from video

For the video information in the generated content, we can use the following steps to process the video content: Firstly, a pre-trained deep learning model, ResNet50, is introduced, whose structure includes the residual blocks shown in Figure 3. OpenCV technology is used to open and frame video files when processing video files. For each frame of the video, a series of necessary preprocessing steps are performed, including adjusting the image size, color channel conversion, and ensuring that the preprocessing meets the input conditions of the ResNet50 model. Next, the preprocessed video frames are input into the ResNet50 model for feature extraction. This step obtains a series of feature vectors and is stored in an array. The output results of the feature extraction array are clearly shown in the figure, which covers the feature vectors of each keyframe in the video. This array can be regarded as the extracted summary vector of the critical information of the video. The detailed steps of the whole processing process are shown in Figure 4, which plays a crucial role in organizing and extracting the critical information of the video in our framework. The organic combination of these steps ensures efficient video content processing and provides rich and accurate information summaries for subsequent tasks.

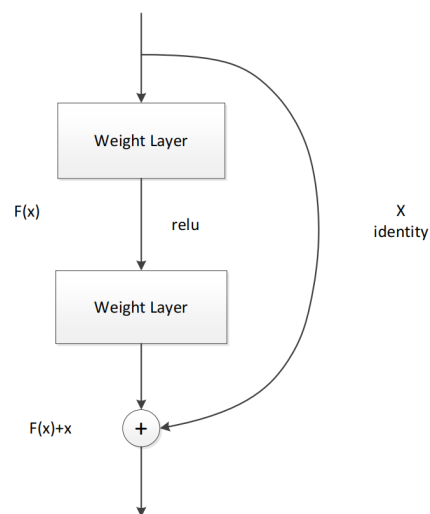


Figure 3. Residual block.

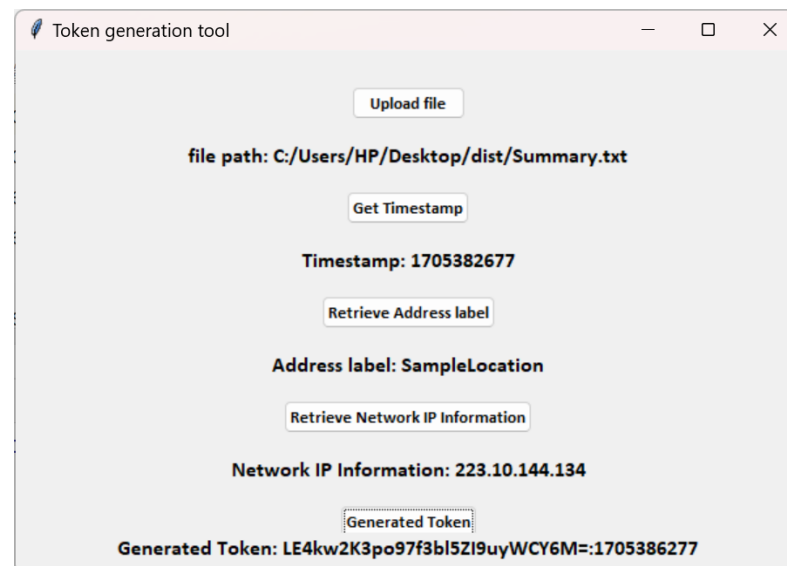




**Figure 4.** Process of extracting feature vectors.

### (3) Multi-factor Token Generation

Firstly, the uploaded text file (critical information summary of the content to be published) is converted into a string and processed using a specific method. Then, the user's current timestamp, address label, and network IP address are obtained. These data will be integrated and spliced in a specific format, including text content, current time, address label, and IP address. Then, the HMAC-SHA1 algorithm calculates the summary of the spliced string as a message, and the corresponding key is used to ensure data security and integrity. Finally, the generated summary is Base64-encoded to form the final token (user token), effectively providing an authentication mechanism to ensure the security and reliability of information transmission. The process and results are shown in Figure 5.



**Figure 5.** User token generation process.

## 5. Discussion

### 5.1. User Experience Impact

The information published on current social platforms comes from various sources, including a large amount of content generated by generative tools, which may be modified and potentially threaten user privacy. After introducing the SecToken framework, social media platforms can control generated content more strictly. By managing the user login method and binding the user identity information, status information, and the summary of the currently published content together, it is helpful to quickly trace the source of the content and ensure that the published content reflects the user's true intention. The SecToken framework provides powerful security mechanisms, including digital signature and user authentication, to ensure the authenticity and integrity of the published content and effectively prevent tampering and impersonation, thereby improving the platform's security and user confidence. However, enhancing security may also impact user experience; for example, double identity authentication and digital signature may increase the number of steps and time for publishing, reducing the convenience of operation. Therefore, the relationship between security and usability needs to be weighed when designing the

SecToken framework, and overemphasis on security causing a decline in user experience, or overemphasis on usability increasing the risk of platform attacks, needs to be avoided. Taking into account the needs of security and usability, appropriate measures should be taken to protect the security of users and content and minimize the adverse impact on user experience; for example, optimizing the authentication process, providing a friendly and easy-to-understand interface, and adopting intelligent algorithms to identify malicious behaviors are all part of the long-term improvement goals of the platform in the future.

### 5.2. Ethical Considerations

Table 3 summarizes the key steps and specific measures of an AI information pollution prevention and control framework that guarantees freedom of speech. The framework includes developing clear and specific filtering standards, ensuring transparent filtering mechanisms, differentiating different types of content, establishing independent supervision and review mechanisms, continuously evaluating and improving the filtering framework, and raising user awareness through education. These measures aim to respect freedom of speech and protect users' right to express themselves while preventing information pollution.

**Table 3.** Key steps and specific measures to ensure freedom of speech.

Steps	Specific Measures
Establishing filtering criteria	Based on objective and impartial principles, such as community norms, laws and regulations or industry standards.
	Clearly define standards
Ensure transparent filtering mechanisms	Users understand the principles and operation of filtering decisions.
	Provide channels for user participation and feedback.
Differentiating between different types of content	Strict censorship of illegal content
	A light touch on political views or controversial topics
Establish an independent monitoring and review mechanism	Third-party organizations or committees are responsible for reviewing filtering decisions periodically
Continuous evaluation and improvement of the filtering framework	Regular audit standards
	Gathering user feedback
	Monitoring the effect
Raise user awareness through education	Users understand how to identify false information or inappropriate content
	Users understand their rights and responsibilities on social media platforms

### 5.3. Privacy Protection

The first principle is data minimization. The framework only collects necessary user information to perform its functions, such as authentication and content abstraction. It does not collect irrelevant personal information. When processing user data, appropriate anonymization and encryption technologies are adopted to protect users' identity information, for example, using hash functions to encrypt sensitive information. The framework gives users complete control over their data, including the right to access, correct, and delete data. Users can choose whether to participate in authentication and data collection

processes. In addition, the framework publishes its privacy policy, clearly explaining the methods and purposes of data collection, use, and sharing to users, ensuring that users fully understand the processing of their data. The framework conducts regular compliance reviews to ensure compliance with global privacy regulations, such as the European General Data Protection Regulation (GDPR) and other relevant regulations. At the same time, appropriate data security measures, including access control, encryption, and secure transmission, should be taken to prevent unauthorized access, disclosure, or tampering with user data. Finally, to further protect user privacy, the framework can provide anonymous posting options, enabling users to post content without exposing their identity.

## 6. Summary and Outlook

This paper introduces a solution to the content trust problem in social media, namely the SecToken framework, which aims to ensure that the generated content published online can accurately reflect the true intention of the content publisher. The framework enhances content security through identity access management to ensure the integrity and independence of the content. It integrates OAuth 2.0 and OpenID Connect technologies to ensure user authentication security. In terms of identity security, this paper proposes an innovative token definition method, which packages the abstract, timestamp, location label, network IP address information, and identity information of the multi-modal content that users will publish as user tokens and uses digital signatures to ensure their integrity. Zero trust policies and real-time monitoring are adopted to ensure the reliability of the overall content. Compared with traditional network security frameworks, the method proposed in this paper shows the best performance on multiple indicators. Therefore, the artificial intelligence-generated content authorization management framework proposed in this paper is of great significance for ensuring the credibility of AIGC, preventing a series of risks that AIGC may bring, and governing the social media information ecology. However, some things could be improved in this study. Firstly, the current data may only partially cover some social media platforms. To solve this problem, the subsequent plan is to obtain data from more foreign online community platforms and generate data using more advanced generative artificial intelligence models to verify the framework's effectiveness further. Secondly, the multi-feature fusion method is considered to improve the accuracy of AIGC recognition. In this process, adopting an end-to-end structure for different feature extraction and fusion types is best. Although the design of this structure is more complex, its effect is more significant. Finally, although the current generative artificial intelligence models, such as ChatGPT, can read pictures and videos, their primary output is text content. Therefore, the subsequent research will explore AIGC recognition methods for multi-modal content to prevent the negative impact of generative artificial intelligence on the social media ecology.

Due to certain constraints, we have not yet had the opportunity to test the framework's effectiveness comprehensively in a real-world environment. We are very interested in conducting case studies in a real-world scenario to better understand the framework's actual performance when facing different challenges. This will enable us to verify the feasibility of the framework and to gain a more comprehensive understanding of its applicability in practical applications.

**Author Contributions:** Conceptualization, J.H. and Q.L.; methodology, J.H.; software, Q.L.; validation, Q.L. and Y.Z.; formal analysis, Y.X.; investigation, Y.X.; resources, J.H.; data curation, B.W.; writing—original draft preparation, Q.L. and Y.Z.; writing—review and editing, J.H. and Q.L.; visualization, Q.L.; supervision, J.H.; project management, J.H.; funding acquisition, J.H. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported in part by the Jilin Provincial Development and Reform Commission Planning Project under Grant 2020C020-2, in part by the Science and Technology Development Center of the Ministry of Education under Grant 2018A04002, in part by the Growth and Climbing Plan Fund Project under Grant ZKP202125, and in part by the Growth Research Fund Project under Grant ZKQD201914.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data are contained within the article.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Carvalho, I.; Ivanov, S. ChatGPT for tourism: Applications, benefits and risks. *Tour. Rev.* **2023**, *79*, 290–303. [\[CrossRef\]](#)
2. Vorobyeva, T.; Mouratidis, K.; Diamantopoulos, F.N.; Giannopoulos, P.; Tavlaridou, K.; Timamopoulos, C.; Peristeras, V.; Magnisalis, I.; Shah, S.I.H. A Fake News Classification Frame Work: Application on Immigration Cases. *Commun. Today* **2020**, *11*, 118–130.
3. Allcott, H.; Gentzkow, M. Social media and fake news in the 2016 election. *J. Econ. Perspect.* **2017**, *31*, 211–236. [\[CrossRef\]](#)
4. Ma, J.; Gao, W.; Wei, Z.; Lu, Y.; Wong, K.F. Detect Rumors Using Time Series of Social Context Information on Microblogging Websites. In Proceedings of the 24th ACM International on Conference on Information and Knowledge Management, Melbourne, VIC, Australia, 18–23 October 2015; pp. 1751–1754. [\[CrossRef\]](#)
5. Centola, D. The Spread of Behavior in an Online Social Network Experiment. *Science* **2010**, *329*, 1194–1197. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Guo, D.; Chen, H.; Wu, R.; Wang, Y. AIGC challenges and opportunities related to public safety: A case study of ChatGPT. *J. Saf. Sci. Resil.* **2023**, *4*, 329–339. [\[CrossRef\]](#)
7. Zhou, X.; Zafarani, R. A survey of fake news: Fundamental theories, detection methods, and opportunities. *ACM Comput. Surv.* (CSUR) **2020**, *53*, 109. [\[CrossRef\]](#)
8. Alsubari, S.N.; Deshmukh, S.N.; Alqarni, A.A.; Alsharif, N.; Aldhyani, T.H.; Alsaade, F.W.; Khalaf, O.I. Data analytics for the identification of fake reviews using supervised learning. *Comput. Mater. Contin.* **2022**, *70*, 3189–3204. [\[CrossRef\]](#)
9. Kim, J. The institutionalization of YouTube: From user-generated content to professionally generated content. *Media Cult. Soc.* **2012**, *34*, 53–67. [\[CrossRef\]](#)
10. Feuerriegel, S.; DiResta, R.; Goldstein, J.A.; Kumar, S.; Lorenz-Spreen, P.; Tomz, M.; Pröllochs, N. Research can help to tackle AI-generated disinformation. *Nat. Hum. Behav.* **2023**, *7*, 1818–1821. [\[CrossRef\]](#) [\[PubMed\]](#)
11. Gupta, A.; Kumaraguru, P.; Castillo, C.; Meier, P. Tweetcred: A real-time web-based system for assessing credibility of content on twitter. *arXiv* **2014**, arXiv:1405.5490.
12. Dwivedi, Y.K.; Kshetri, N.; Hughes, L.; Slade, E.L.; Jeyaraj, A.; Kar, A.K.; Baabdullah, A.M.; Koohang, A.; Raghavan, V.; Ahuja, M. “So what if ChatGPT wrote it?” Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *Int. J. Inf. Manag.* **2023**, *71*, 102642. [\[CrossRef\]](#)
13. Graf, A.; Bernardi, R.E. ChatGPT in research: Balancing ethics, transparency and advancement. *Neuroscience* **2023**, *515*, 71–73. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Woods, L. User generated content: Freedom of expression and the role of the media in a digital age. In *Freedom of Expression and the Media*; Brill Nijhoff: Leiden, The Netherlands, 2012; pp. 141–159.
15. Roy, P.K.; Tripathy, A.K.; Weng, T.H.; Li, K.C. Securing social platform from misinformation using deep learning. *Comput. Stand. Interfaces* **2023**, *84*, 103674. [\[CrossRef\]](#)
16. Hosseinimotlagh, S.; Papalexakis, E.E. Unsupervised content-based identification of fake news articles with tensor decomposition ensembles. In Proceedings of the Workshop on Misinformation and Misbehavior Mining on the Web (MIS2), Los Angeles, CA, USA, 9 February 2018.
17. Ali, G.; ElAffendi, M.; Ahmad, N. BlockAuth: A blockchain-based framework for secure vehicle authentication and authorization. *PLoS ONE* **2023**, *18*, e0291596. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Musliyah, Z.; Satira, A.G.; Dwipayana, M.; Helinda, A. Integrated Email Management System Based Google Application Programming Interface Using OAuth 2.0 Authorization Protocol. *Elkawanie J. Islam. Sci. Technol.* **2020**, *6*, 109–120. [\[CrossRef\]](#)
19. Kang, J.; Zhang, J.; Li, W.; Zhuo, L. Crowd activity recognition in live video streaming via 3D-ResNet and region graph convolution network. *IET Image Process.* **2021**, *15*, 3476–3486. [\[CrossRef\]](#)
20. Seamless Communication; Barrault, L.; Chung, Y.A.; Meglioli, M.C.; Dale, D.; Dong, N.; Duquenne, P.A.; Elsahar, H.; Gong, H.; Heffernan, K.; et al. SeamlessM4T: Massively Multilingual & Multimodal Machine Translation. *arXiv* **2023**, arXiv:2308.11596.
21. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Wang, D.; Liang, Y.; Ma, H.; Xu, F. Refined Answer Selection Method with Attentive Bidirectional Long Short-Term Memory Network and Self-Attention Mechanism for Intelligent Medical Service Robot. *Appl. Sci.* **2023**, *13*, 3016. [\[CrossRef\]](#)

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.