Comparing news and non-news sites in Web3 domain

Jung Min Lee

Cyber Emotions Research Center, YeungNam University, Republic of Korea

Han Woo Park

Department of Digital Convergence Business, YeungNam University,

Department of Media & Communication, YeungNam University,

Department of East Asian Cultural Studies, YeungNam University,

Cyber Emotions Research Center, YeungNam University, Republic of Korea (hanpark@ynu.ac.kr)

Abstract

This study provides a pioneering analysis of the features and topics of news websites in the Web 3.0 era through a comparison with non-news sites. We first classify over 4,600 Web3 sites into news and non-news types to investigate the feature characteristics of decentralized websites as well as semantic and subject categories. The most novel finding is that Web3 news sites have more features than non-news sites in terms of design systems and functions, interactivity, information quality, and hyperlinks. Furthermore, news web3 domains have more terms in semantic networks associated with information provision (e.g., "source"), whereas non-news domains have more terms associated with finance (e.g., "token"). The integration of the Decentralized Autonomous Organization, a fundamental component of the Web3 ecosystem, began in the news industry and is at an early stage in terms of functionality and structure. The study's results are discussed in the context of future Web3 domain development.

1. Introduction

Journalism has undergone significant changes owing to the rapid development of the Internet and digital information technology. One of the most noticeable changes in the Internet environment is the expansion and diversification of traditional news media, often referred to as "legacy media." The increase and diversification of news producers and suppliers have produced a world in which anyone can become a journalist if they have something to report, hear, or say. Consequently, news websites have evolved. Print media, broadcasting, and other news organizations have moved beyond merely disseminating news content online. The emergence of news production, remediation, and distribution through social media networks and the rise of one-person media, as represented by blogs and YouTube, have transformed the structure of the news ecosystem. Consequently, what was once focused on textual content is now widely distributed across social and interactive digital communication networks influenced by the active involvement and interaction of news consumers.

The Web3 ecosystem, which includes technologies such as blockchain, has garnered attention in the media field because it can fundamentally revolutionize centralized media platform-oriented ecosystems. In

markets where journalism has been weakened and faces financial difficulties, the values of transparency and reward make Web3 an attractive breakthrough (Agrawal et al., 2020; Bhatia et al., 2018). Technically, through a decentralized structure, Web3 can distribute the excess profits of media platforms to creators and users, thus forming a media ecosystem that is centered on them.

Interest in Web3 varies among media outlets, from The *New York Times* in the United States to small and medium-sized Korean newspapers in Asia (Roose, 2021). News organizations are approaching Web3 by releasing Non-fungible Tokens (NFTs) for news content. These cases demonstrate that new legitimate ownership and compensation systems have been established for news produced by the press.

This study aims to examine the characteristics and themes of news websites in the current Web3 environment through the comparison of news websites and non-news websites. Is there anything new about Web 3.0 news websites that is different from Web 2.0 news sites? How are the characteristics of Web3 reflected in news websites? Compared to non-news sites, what is the difference in features and topics? While there are several studies on Web 1.0 and Web 2.0 news websites, but only a few of studies are related to news in the Web3 environment (Unzeellah and Memon, 2022; Upadhyay and Baranwal, 2021). Based on these results, we will be able to examine the characteristics of the emerging news media and discuss the structure of the Web3 news media.

2. Background

2.1 Changes in news structure with the development of the Web

After the 1980 debut of the *Columbus Dispatch*, the first online version of a daily U.S. newspaper, many newspapers made efforts to adapt to the online space in the 1990s (Weber, 2017). The offline-to-online transition has various effects on news environments. However, the traditional newspaper business model does not function properly in the online format. Weber and Monge (2011) use Kleinberg's (1999) authoritative model as a theoretical foundation to emphasize the importance of influential intermediaries (H) within the changed online news communication structure. In the SAH (Source – Authority – Hub) structure, hubs serve as the final step of the process in which authority is passed from the source, guiding collected news links to user-preferred topics or related high-relevance information. Google News, Huffington Post, and Yahoo News have taken on this role. The concern is that the dominance of platform media over the traditional press could have negative impacts, potentially undermining the fundamental role of the press (Libert and Binns, 2019; Nielsen, 2015). Through empirical analysis, Libert and Binns (2019) discovered that news websites compromise more personal user information than non-news websites do and lack transparency in their privacy policies. Moreover, they found that news sites rely more on third parties, such as platform companies.

However, there is a two-way relationship between journalism and social media (Kapidzic et al., 2022), given the structure of Web 2.0, which enables the press to use content from social media platforms as news sources. Social media can provide diverse news sources and opportunities to engage the public. Nevertheless, the proliferation of social media platforms in the modern news era has resulted in a decline in journalism diversity, primarily because of a reliance on unverified sources.

Meanwhile, the term "Web3" does not refer solely to specific technologies such as blockchain, decentralized finance (DeFi), or NFTs. More broadly, it refers to a transformation from Web 2.0 to a decentralized Web (Cao, 2022). This involves user-owned, -managed, and -personalized services for vast amounts of data generated on the Internet, such as news and content. The blockchain technology in Web3 can eliminate unfair elements in Web 2.0 (Cao, 2022). Blockchain technology can enhance transparency and process integrity. This feature can effectively address the problem wherein major news organizations and social media giants appropriate news content from small community newspapers and ordinary users without proper citation or compensation. Blockchain's smart contract protocol provides a high level of trust because it guarantees immutability, and all transactions are open and traceable (Zheng et al., 2017). Recent research on the application of blockchain technology to news highlights fake news verification as a notable use case (Unzeelah and Memon, 2022). Research findings have highlighted the necessity of news content verification because intentionally fabricated false information disseminated through traditional or social media can

lead to misunderstandings and even intentional misinterpretations of government policies (Liu et al., 2021). A decentralized distributed trust framework using blockchain adds transparency to news verification and prevents the dissemination of fake news (Qayyum et al., 2019). This research focuses on news content and organizations that are being transformed by the emergence of Web3. We want to explore how the characteristics that will emerge in the next phase of the web will be applied in news media.

2.2 Blockchain domain names

Blockchain domain names are used in a decentralized Web based on blockchain technology. They serve functions similar to those of the Domain Name System (DNS) in Web 1.0 and Web 2.0 but with a significant difference in their operational structure. Unlike traditional systems, there are no central organizations such as the Internet Corporation for Assigned Names and Numbers (ICANN), and it operates in a decentralized manner.

Prominent blockchain domain providers include the Ethereum Name System (ENS) and unstoppable domains (Durand, 2022). The ENS operates using the Ethereum blockchain and uses .eth as the Top-Level Domain (TLD). The ENS system functions in a decentralized manner using smart contracts. To use a designated name, individuals must remit an annual fee for a lease. Furthermore, the execution of smart contracts during the registration and usage phases incurs high gas fees. However, unstoppable domains offer various TLDs, such as .crypto, .nft, .wallet, .x, .bitcoin, .zil, .DAO, .888, and .blockchain. These domains can be based on both polygon (MATIC) and Zilliqa blockchains and are even Ethereum-based. The .eth domains are available for purchase. Flexibility in registering different types of domain is a notable feature. Additionally, several other blockchains—like Bitcoin, Solana, Tezos, and Polkadot—also offer blockchain domain services.

The registration of blockchain domain names has steadily increased worldwide. According to the DNS analysis platform DAP.live, the number of registrations increased significantly, from 1,208,691 in 2021 to 4,365,619 in 2022. As of October 30, 2023, there are over 7.22 million blockchain domain names. Ethereum (ETH) is the most widely used blockchain, with approximately 3.57 million domains, closely followed by MATIC (Polygon), with approximately 3.52 million. The utilization of blockchain domain names varies but can be broadly categorized into three main use cases (Osborn and Alan, 2023). Using them as cryptocurrency wallet addresses, creating censorship-resistant websites through IPFS, and engaging in domain name NFT trading. According to Osborn and Alan (2023), the most common use case is the first. Nevertheless, there is significant potential for using blockchain domains for IPFS-based websites (Ash, 2022; The IPFS Team, 2017). Therefore, this study focuses on ENS, the most widely used blockchain domain, and analyzes empirical cases of its usage for websites, particularly within the context of Web3 news media.

2.3 Analyzing website content and features

Herring (2010) divided web content analysis into two categories. The first comprises traditional content analysis techniques developed by Krippendorff (1980) based on the work of MacMillan (2000), while the second includes non-traditional methods, such as Social Network Analysis (SNA). An evaluation of hyperlinks within the online domain can reveal the reasons for the presence and interconnection of specific website content (Park and Thewall, 2003). Researchers can explore a wide range of analytical opportunities on Web3 by deviating from traditional content analysis methods. It is also possible to use standardized units of analysis such as web features and random sampling. Investigating communicative structures and features goes beyond a rudimentary analysis of textual content and visual representations (Schneider and Foot, 2004). Furthermore, web feature analysis enables the measurement of the immaterial properties of the web space, which reflect the public media space discussed by Bolter and Grusin (2000). It also allows for a broad conceptualization of features, which makes it applicable in various fields across media types.

Benoit and Benoit (2000) examined the attributes of Web 1.0 websites utilized in electoral campaigns, contributing to the body of research on website analytics throughout the various stages of the Web's evolution. They organized their website evaluation questions into eight categories: identification, navigation, readability, irritability, information accessibility, interest level, interactivity, and adaptation to the audience. Basic content analysis focuses on textual and content aspects. Web interactivity is another prominent feature in studies of Web 2.0 sites. Himelboim and McCreery (2012) analyzed user-computer interaction on news sites by dividing Web characteristics into textual and visual types. They found that newspaper

websites were more likely to use text-based features, whereas broadcast websites were more likely to use visual features. Similarly, Rebillard and Touboul (2010) focused on user engagement on Web 2.0 news sites and analyzed their links and content, empirically exploring the potential for digital journalism generated by Web 2.0.

The various website evaluation criteria utilized by previous researchers suggest the usefulness of website content and feature analysis (Cebi, 2013) and reflect the concepts and phenomena derived from the changing Web environment. This study uses Web content analysis as its research methodology. However, a combination of Web feature analysis, which identifies technical web options, and semantic network analysis, which is derived from SNA, is used in conjunction with other techniques, such as topic clustering and ENS cracking. This allows the study to examine the extent to which news sites reflect the changing nature of the decentralized Internet and investigate whether Web3 characteristics are currently operationalized. The results offer new insights useful for preparing for the future of news.

3. Research questions

News does not need to be defined in ordinary conversation because everyone knows what it is, and it is integrated into our lives such that we do not question its existence (Shoemaker, 2006). Nevertheless, the academic definition of "news" has continued to evolve over time as advances in communication technology have changed the primary value of the media. The concept of news has changed significantly amid the development of digital media. Galtung and Ruge (1965) claimed that "News is the product of a news organization, conveying information to an audience about recent events, developments, and issues." However, with the development of social media, where anyone can become a news editor, the agenda-setting function of the mainstream media has weakened. Therefore, news can no longer be defined as a discrete activity constrained by a specific time and place (Hermida, 2016). For instance, social media-based news can be defined as all social stories exchanged during communication, signifying an expansion of the concept of news (Baekdal, 2009).

This study examines the concept of news in this context. Nonetheless, to explore and create opportunities in the new media market, the more expansive concept of news implied in the Web3 context should be scrutinized closely. Studies on news websites have focused on the sites of established media organizations (Rebillard and Touboul, 2010; Hermida et al., 2011; Himelboim and MacCreery, 2012). These studies conducted research on online news users by analyzing news websites in the Web 1.0 and Web 2.0 contexts. Finding a novel mode of news dissemination during the nascent phase of the Web3 era poses a considerable challenge, primarily because of the dearth of recent research on Web3, which hampers the validation of journalism's operational framework. This study began by compiling a list of Web3-based platforms, and it sampled numerous websites, both news and non-news, to allow for a comparative analysis. A comparison of news and non-news sites allows us to understand the role news organizations, including individual owners, play in the Web3 media environment. The following research questions were formulated:

RQ 1: What are the feature-based characteristics of news websites and non-news websites in Web3, and what are the differences between them?

RQ 2: What are the topical themes of content covered by news websites and non-news websites in Web3, and what are the differences between them?

4. Method: Data collection and analytical techniques

4.1 Data collection

The websites under analysis were collected from Almonit.eth and Esteroids.eth, the most popular decentralized web portals. Almonit is a pioneering search engine in the Web3 domain, whereas Steroids is the leading web portal with the most extensive array of Web3 sites. Similar to Google, it assists users in locating Web3 sites by inputting search keywords. The decentralized websites included in Almonit and Esteroids were collected as part of the data-collection procedure. We gathered 4,617 Web3 sites for this study. Next,

we randomly selected news and non-news sites from over 4,000 online sites. During this process, news sites were operationally defined using six keywords: "journal," "broadcast," "magazine," "news," "article," and "post." These keywords were chosen through pilot testing to identify the Web3 news sites. We visited the initial sites and excluded those that were not codable. In total, 208 news and non-news sites were selected, with 104 sites in each category. Random sampling was performed using the Python random library. The coding, which occurred from June to July 2023, was conducted by a graduate student majoring in media communication under the guidance of the principal investigator. Approximately 10% of the 208 decentralized websites were initially coded. The inter-coder reliability between the two coders was 0.71 (Cohen's kappa).

4.2 Analysis method

4.2.1 Analysis procedure for RQ 1

To analyze RQ 1, this study developed evaluation criteria (coding items) for the collected news and non-news sites based on a process used by a global project that introduced the concepts of the "Web sphere" and "Web features" during the early stages of Web 2.0 adoption (Kluver et al., 2007) and designed items suitable for Web3. The items analyzed were categorized into several types: design systems and functions, digital assets, interactivity, information quality, and hyperlinks. Design systems and functions examine the convenience and diversity of the functions provided to users of websites. The feature items used in research on online news audience participation on Asian news websites were adapted for the context of Web3 (Xiaoge, 2008). These items assess the design and functional completeness of Web3 sites. Digital assets assess whether the website reflects the characteristics of Web3, particularly those related to incentivization. Interactivity and information quality items examine whether the site functions as a media source and are constructed based on the evaluation dimensions used in past Web 2.0 news research (Rebillard and Touboul, 2010; Himelboim and McCreery, 2012). The interactivity item assesses the potential for dissemination and the degree of interactivity in Web3 news, whereas the information quality item assesses the credibility of Web3 news. Finally, hyperlinks are related to the expansiveness of the Web3 space and allow for an examination of the patterns of site connections in a decentralized structure. Table 1 presents the evaluation criteria for each item.

Major Categories	Subcategories	Description	Relevance to Web3
Design systems and functions	(1) Search	Does the Website have a search function?	Assessing the design and functional completeness of Web3 sites
	(2) News arrangement	Are news, articles, or new information displayed through divisions (e.g.,	
	(3) Multilingual support	Is there a language change setting on the Website?	
	(4) Subcategories of news	Are news or articles classified using subcategories?	
	(5) Single-Page news site	Does the Web page consist of a single page containing news or information?	
	(6) Sign-up	Is there a sign-up (Sign up) function?	
	(7) Donations	Does the Website accept donations?	
Digital assets	(8) NFT issuance	Does the Website issue its own NFTs?	Assessing Web3 interoperability
	(9) Operator wallet address	Does the Website operator reveal their wallet address?	
	(10) Wallet connection	Is there a wallet connection function?	
	(11) Token issuance	Does the Website issue its own tokens?	
Interactivity	(12) SNS channels	Does the Website have official SNS accounts?	Assessing the potential for dissemination and interactivity in Web3 news
	(13) Sharing	Is there a function to share posts or information?	
	(14) Newsletter service	Does the Website support newsletter services?	
	(15) Contact information	Is there an email address provided for users to contact the operator?	
	(16) Communication space	Does the Website have public communication spaces such as forums, chat rooms, comment sections, etc.?	
Information quality	(17) Personal data protection	Is there information regarding personal data protection?	Assessing the credibility of Web3 news
	(18) Organization/Institution site	Is the Website operated by a company, institution, or organization?	
	(19) Website introduction	Is there information introducing the Website or an About category?	
	(20) Images	Are images present on the Website?	
	(21) Terms of use	Are there terms of use?	
	(22) Date of Information Publication	Is the publication date provided for posts?	
	(23) Latest news provision	Are the most recent posts within the last year?	
	(24) Category depth	Are there two or more levels of categories?	

Major Categories	Subcategories	Description	Relevance to Web3
Hyperlinks	(25) Links to other dWeb	Are there links to other decentralized Web (dWeb) sources?	Assessing the expansiveness of the Web3 space
	(26) Link stability	Are there no link breakages on the Website?	
	(27) Links to Web 2.0 sites	Are there links to Web 2.0 sites other than SNS channels?	

Table 1: Coding evaluation criteria (Source: Developed by the authors for the study based on the literature)

News arrangement items examine the design elements of websites. If a specific decentralized website falls under this category, its news articles are well-organized and displayed, which can assist users in their news consumption. The study classified and categorized types of news subcategories. Decentralized websites falling under this category can be considered similar to conventional Web 2.0 news websites.

A single-page news site refers to a website that comprises a single page and functions as the news, such as websites created by Nimi. Category depth of the information quality item measures how systematically a website provides information. As it is difficult to determine whether a single category contains a substantial amount of information, we set two or more categories as the evaluation criterion.

Finally, we structurally examined the characteristics of 27 news and non-news websites' subcategories using multidimensional scaling (MDS). MDS is a method used to explore the structure of data by measuring distances between variables, represented as points. Each point signifies the similarity and dissimilarity between objects, with closely situated points indicating high similarity. Points positioned at the center of the network hold greater significance (Zhang et al., 2015). In this study, we employed MDS Proxscal based on a frequency matrix of subcategory features. The frequency matrix was transformed into a proximity matrix using the chi-square measure, and Euclidean distance was utilized for measurement. The analysis was conducted using IBM SPSS version 26.0.

4.2.2 Analysis procedure for RQ2

Topic analysis was conducted based on the introductory information provided on the decentralized websites, as it appears in Almonit and Esteroids (see Figure 1).
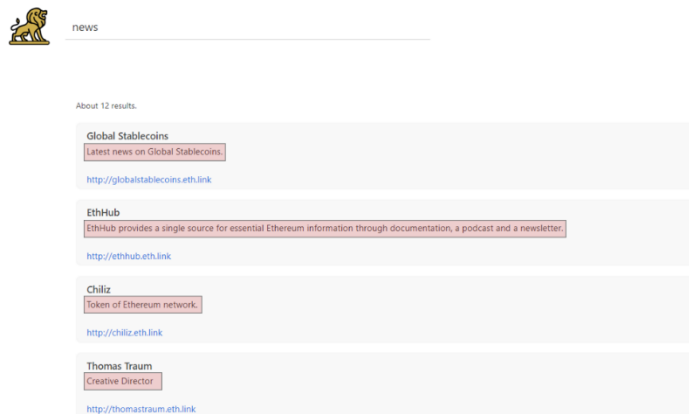

Figure 1: Introductory text on decentralized websites

After the introductory text was preprocessed, a semantic network was created based on the word frequency list. Semantic network analysis applies social network analysis to uncover hidden meanings in a text by setting words as nodes and examining the relationships between them (Jiang et al., 2016). A Haddop-based big data platform was used for data preprocessing, and UCInet was used to calculate semantic network metrics and generate visualizations (Borgatti et al., 2002).

Finally, cluster analysis was performed to compare the detailed topics of the content covered by news and non-news sites. In a network analysis, a cluster occurs when specific subgroups of nodes are more interconnected than subgroups in other areas (Carrington et al., 2005). Cluster analysis was used to identify small-scale groups within the networks. This study examined specific topics within the semantic networks of news and non-news sites. NodeXL Pro's Clauset-Newman-Moore algorithm was employed as the analysis tool (Hansen et al., 2010).

5. Results

5.1 Results for RQ1

5.1.1 Coding results for news and non-news sites

Figure 2 shows the differences in characteristics between news and non-news websites when they are classified into major categories. The figure shows the average frequency of appearance for each item. The most novel finding is that Web3 news sites had more features in most major categories than non-news sites. The proportion of non-news sites was higher than that of news sites only for digital assets. In both cases, the highest number of websites was coded in the hyperlink category. However, news sites had the lowest frequency in the digital assets category, whereas non-news sites had the third-highest. Moreover, the results showed that, overall, news sites appeared more often than non-news sites, except in the digital assets category. This indicates that non-news sites tend to exhibit more characteristics that are unique to Web3, particularly those related to data ownership and incentive activities, as various kinds of Web3 sites engage in such activities through their digital assets.

Table 1 shows that design systems and functions are the major categories that display the most significant differences between news and non-news sites. News sites had an average of 18 appearances, whereas non-news sites had an average of 2.57 appearances. This finding suggests that news websites provide interfaces and features that increase user convenience, which suggests that the user perspective is considered seriously.
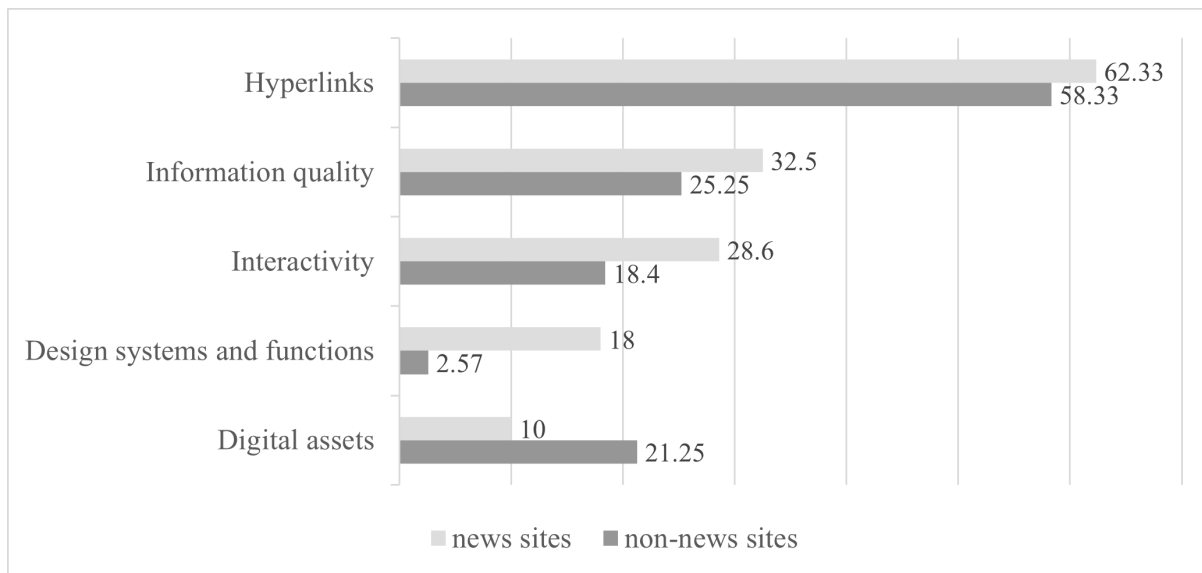

Figure 2: Comparing the characteristics of news and non-news sites across major categories

The results of the detailed coding for news and non-news sites are shown in Table 2. First, for non-news sites, no websites fell into the news or single-page news subcategories. News and non-news sites had the highest proportion of link stability, with 90 and 94 appearances, respectively. Although news sites had a higher average frequency for all items, non-news sites had a slightly more stable web environment. Non-news sites had more appearances than news sites in other subcategories, including organization/institution site, NFT issuance, operator wallet addresses, wallet connection, and token issuance. A significant difference was observed for the organization/institution site category, with 24.05% for news sites and 51.92% for non-news sites, possibly because many DeFi-related sites frequently appear on non-news sites operated by institutions. Moreover, Decentralized Autonomous Organization (DAO) websites were also prevalent among non-news sites. Nevertheless, news sites appear to be run more often by individuals or small groups, indicating that few organized entities are functioning systematically as news outlets in Web3. As mentioned, the categories of NFT issuance, operator walking addresses, walking connections, and token issuance fall under the broad category of digital assets. These results are attributable to the many DeFi-related sites supporting walking connection functions, token issuance, and NFTs.

Conversely, some subcategory items appeared more often on news sites than on non-news sites, such as the date of information publication, latest news provision, search, news arrangement, sharing, and newsletter

services. The date of information publication and the latest news provisions are essential for information dissemination. Search and news arrangements are related to information acquisition. These results suggest the need to deliver accurate information and enhance user convenience during information retrieval.

Contact information was present more frequently on news sites (36.54%) than on non-news sites (18.27%), likely because non-news site operators have opted to share SNS account information as an alternative communication channel instead of providing email addresses.

Categories		News sites		Non-news sites	
Major Categories	Subcategories	Frequency	Percentage	Frequency	Percentage
Design system and functions	Search	12	11.54	3	2.88
	News arrangement	46	44.23	6	5.77
	Multilingual support	5	4.81	2	1.92
	Subcategories of news	7	6.73	0	0
	Single-Page news site	48	46.15	0	0
	Sign-up	2	1.92	2	1.92
	Donations	6	5.77	5	4.81
Digital assets	NFT issuance	10	9.62	18	17.31
	Operator wallet address	19	18.27	25	24.04
	Wallet connection	8	7.69	30	28.85
	Token issuance	3	2.88	12	11.54
Interactivity	SNS channels	78	75	66	63.46
	Sharing	11	10.58	3	2.88
	Newsletter service	13	12.5	3	2.88
	Contact information	38	36.54	19	18.27
	Communication space	3	2.88	1	0.96
Information quality	Personal data protection	8	7.69	4	3.85
	Organization/Institution site	25	24.04	54	51.92
	Website introduction	73	70.19	59	56.73
	Images	56	53.85	47	45.19
	Terms of use	9	8.65	7	6.73
	Date of Information Publication	38	36.54	9	8.65
	Latest news provision	14	13.46	5	4.81
	Category depth	37	35.58	17	16.35
Hyperlinks	Links to other dWeb	13	12.5	4	3.85
	Link stability	90	86.54	94	90.38
	Links to Web 2.0 sites	84	80.77	77	74.04

Table 2: Contrasting the characteristics of news and non-news sites across specific categories

5.1.2 Analysis of coding items for news sites

The "single-page news site" item had the highest frequency in terms of the design system and functions of news sites (48%). Figure 3 shows an example of a single-page news site with the Alexreyes.eth web address. Single-page news sites are often used as personal profiles that provide information about the site operator's background. The "website introduction" item had the highest frequency in terms of information quality, which is likely aimed at increasing the trustworthiness of news sites and the information they deliver. "Operator wallet addresses" had the highest frequency in the digital assets category, possibly because user wallet addresses are included in the default requirements in Nimi (see Figure 4). "SNS channels" had the highest frequency (75%) for interactivity, while "link stability" had the highest frequency in hyperlinks (90%). Both items ranked among the top three in terms of frequency and were commonly found on news sites.

Conversely, "sign up" had the lowest frequency for news sites, with only two instances. This result is due to the high proportion of one-page news sites and suggests that news site establishment and operation in the Web3 ecosystem are not as active as they are in Web2. Further, "personal data protection," "communication space," and "token issuance" were found in only three websites, highlighting the near absence of communication spaces among news consumers. Token issuance, which is associated with news content ownership and is one of the main features of Web3, was rarely observed.

Interesting results were seen regarding the "subcategories of news" and "single-page news site" items, which are linked to the functional completeness of news sites. "Single-page news site" accounted for 46.15% of the total, and "subcategories of news" accounted for 6.73%. This signifies that Web3 news sites differ significantly from the typical appearance of Web2 news sites. In Web3, news content is not as abundant and is typically presented in a one-page format, with fewer categories. However, news sites tend to offer more categories (category depth: 37%) and more information than non-news sites.

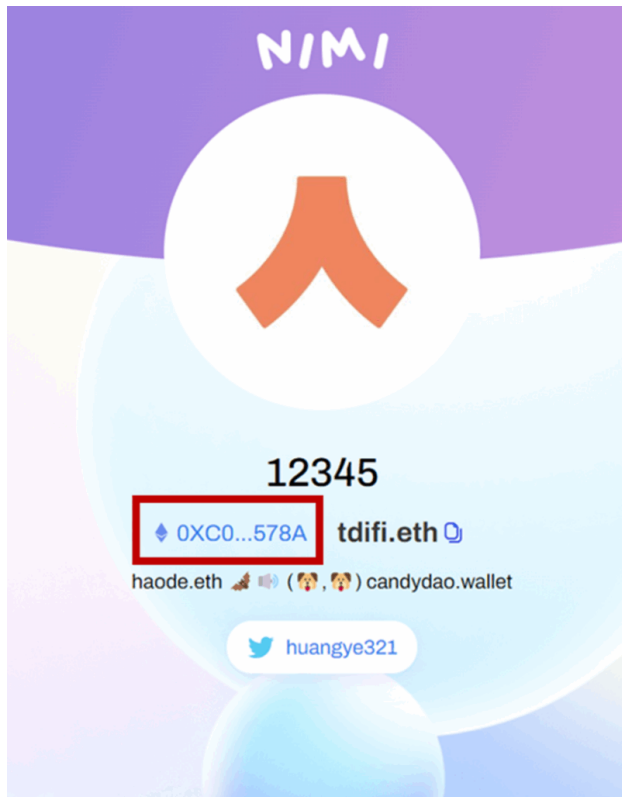**Figure 3:** Single-page news site

Figure 4: Wallet address of Nimi

Nevertheless, instances of "links to Web 2.0 sites" were primarily found in the "website introduction" item, where links to the official IPFS website (https://ipfs.tech/) and ENS official website (https://app.ens.domains/) were included. These links were primarily for explanatory purposes rather than for practical connections.

5.1.3 Analysis of coding items for non-news sites

Non-news sites showed lower frequencies in all major categories except "digital assets." All subcategory items under "design system and functions" had frequencies in the single digits, indicating that non-news sites are relatively simple in their structure and functionality. Among the "digital assets" items, "wallet connection" had the highest proportion (30, 28.85%). This suggests that non-news sites such as DeFi-related services or gaming sites often require or provide walled-connection services. In terms of "interactivity," "SNS channels" had the highest occurrence (66, 63.46%), making it the third most-frequent item overall. This feature is common to both non-news and news sites.

In the "information quality" category, "website introduction" had the highest frequency; this accounted for 59 of 104 sites (56.73%). "Website introduction" remained in the 50% range despite its default setting because many Web3 sites supporting token issuance or exchange services provided little to no description of them. An example of this is shown in Figure 5. The website's address in the picture is "mergeswap.eth," indicating its support for a swapping service as revealed by ENS. However, the website provides no description of the services offered other than the information shown in the picture. Many non-news sites follow this format and offer services without providing introductory information.
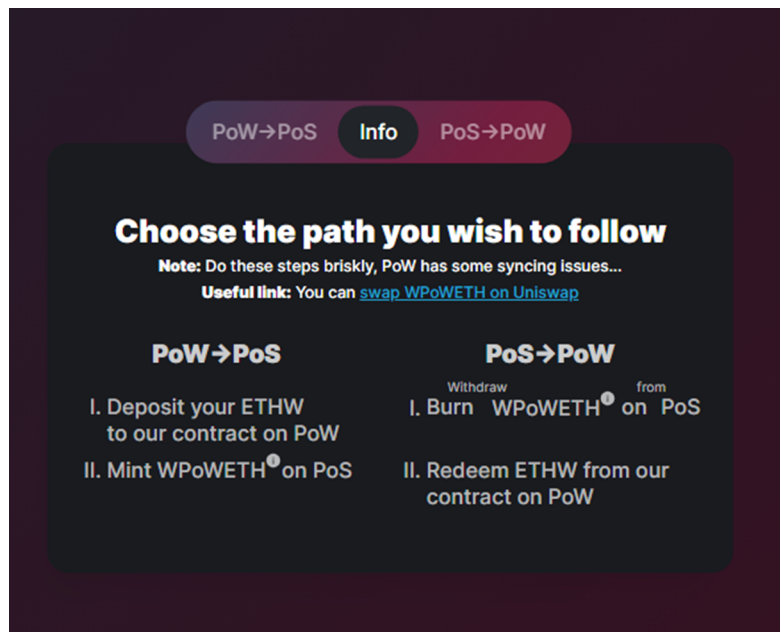
Figure 5: First screen of non-news site

The item that appeared least-frequently among the subcategories was "communication space," which was found only once. The fact that this item occupies the lowest rank (as with news sites) shows that communication spaces in the Web3 environment are still lacking. The "sign up" item also had the lowest frequency, with only two occurrences, placing it at the bottom of the list among all items. The "multilingual support" item was found in two cases, with one website supporting English and Spanish and another supporting nine languages, including Korean. Overall, non-news sites are considered to support specific services, in contrast to news sites, which focus on information delivery. Nonetheless, detailed explanations of service support and user-friendly interfaces are lacking on these sites.

5.1.4 Compare structural characteristics for news and non-news site subcategories

The MDS analysis results revealed that the normalized raw stress values for news sites and non-news sites were 0.09343 and 0.15631, respectively, while Tucker's coefficient of congruence was 0.95214 for news sites and 0.91853 for non-news sites. The values show the appropriateness of using MDS for comparison purposes. Figure 6 shows the results of the MDS analysis for the news and non-news site subcategories. Each entity's number is shown by the numbers in the subcategories (see Table 1). The results show that the plot's major objects appear differently on news and non-news websites. Entities that are centrally located on news sites, like search (N1), sharing (N13), and terms of use (N21) show similarities between these items and highlight features that are given higher importance on news sites. On the other hand, these entities are dispersed around non-news websites. In non-news sites, multilingual support (NN3) is centrally located, excluding subcategories of news (NN4) and single-page news (NN5) with zero frequency. Similarities can also be demonstrated between multilingual support (NN3), token issuance (NN11), and personal data protection (NN17). On the other hand, these entities are dispersed throughout news websites. Furthermore, a multidimensional classification of 27 website features in the analyzed space reveals the following tendencies: In terms of Dimension 1, more complex features for user convenience are typically found below, while functions for basic website usage are typically found above. In Dimension 2, Web3's functions related to finance are on the right, while features related to news content are on the left.
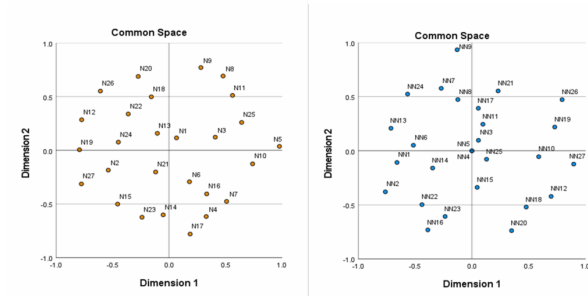
Figure 6: Results of multidimensional scaling analysis of news (left) and non-news sites (right)

5.1.5 Comparison of ENS features between news and non-news sites

The study classified the lengths of the ENS domain names for the collected news and non-news sites. For example, the domain name length of https://ensnews.eth is seven letters, excluding "https://" and "eth." Figure 6 shows that the three-letter ENS domains with the highest annual costs accounted for the smallest percentages of both types of sites (news sites: 0, non-news sites: 2). The four-letter domains appeared on three news sites (2.88%) and two non-news sites (1.92%). The cheapest five-letter ENS domains accounted for 97.12% of the news sites and 96.15% of the non-news sites. The three-letter ENS domains found among non-news sites were "law.eth" and "ada.eth." Of these, "law.eth" was found to be used to sell directory listings, using the rare ENS domain for Web3 sites managed by DAO, while "ada.eth" appeared to be the website of an NFT artist.

Nevertheless, both news and non-news websites had the highest percentages of ENS domains with five or more letters. Beyond these five letters, the domains can include words with specific meanings, such as verbs or proper nouns. In this respect, the usage patterns differed between news and non-news websites. Among the news sites, four had "news" in the domain, one had "post," and one had "journal." However, none of the non-news sites contained those words; instead, seven websites had "NFT," one had "token," and four had "dao." Additionally, none of the domains of the news sites solely comprised numbers or were in languages other than English. For the non-news websites, two domains consisting only of numbers and one domain composed solely of Japanese characters were found. These domains are related to Nimi- or DeFi-related sites.
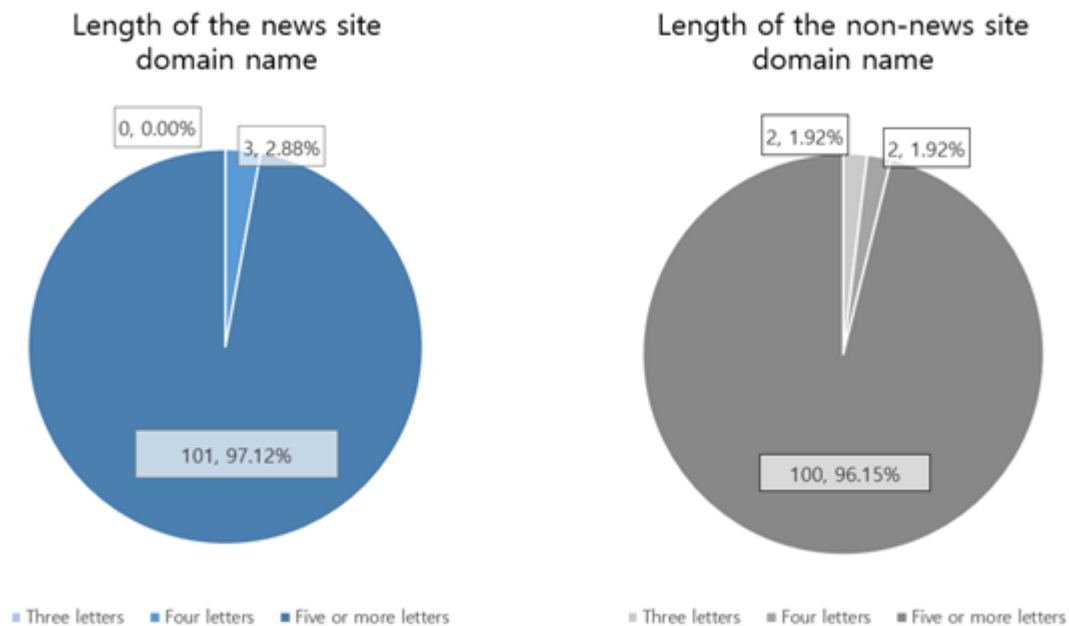
Figure 7: Length of domain name

5.2 Results for RQ2

This study conducted a topic analysis of the introductory information on the collected news and non-news sites to answer RQ2.

5.2.1 Semantic network and centrality analysis

To answer RQ2, we performed a topic analysis of the introductory information on the collected news and non-news sites. First, we extracted 40 words from the websites based on how often they appeared in the introductory information. Centrality analysis was then conducted based on the 40 words. The centrality metrics used were degree and betweenness centrality. Degree centrality is a metric that indicates how well a particular word is connected to other words in the network, while betweenness centrality measures the extent to which a particular word mediates its relationship with other words.

Table 3 shows the top 10 words for each degree of centrality for the news and non-news sites. The table shows differences in the patterns of word occurrences in the introductory information between news and non-news sites. For news sites, "ens" was ranked highest, followed by "news," "website," and "dao." As these are news sites, words related to news appeared frequently, such as "news," "newsletter," and "source." For the non-news sites, "nft" was ranked highest, followed by "ens," "collection," "token," and "subdomain." The appearance of "ens" and "dao" in the top rank is normal and predictable. Nevertheless, the appearance of terms like "nft," "collection," "token," and "defi" suggests that there are a number of sites related to digital assets on non-news sites.

Rank	News	Degree centrality	Non-news	Degree centrality
1	Ens	36	nft	41
2	News	36	ens	30
3	website	25	collection	23
4	Dao	24	token	21
5	general	24	subdomain	21
6	newsletter	23	dao	18
7	blockchain	22	defi	15
8	follow	19	ethereum	15
9	source	18	decentralized	12
10	ethereum	18	blockchain	12

Table 3: Top 10 words based on degree centrality

Table 4 lists the top 10 words based on betweenness centrality. The results for the word rankings differ from those for degree centrality. For news sites, "dao," ranked fourth in degree centrality, has dropped out of the rank, and "post" has appeared as a new entry. This suggests that words with broader and more general applications have increased in rank. As these results are for news sites, it appears that these words were used in the context of news, posts, articles in specific fields, and blogs related to specific individuals. For the non-news sites, many of the words that appeared in the rankings for betweenness centrality also appear in those for degree centrality. Similarly, interest in digital assets is evident on non-news sites.

Rank	News	Betweenness Centrality	Non-news	Betweenness Centrality
1	news	224.955	nft	277.388
2	website	105.453	token	155.785
3	post	90.58	ens	65.247
4	blockchain	69.859	art	41.84
5	ethereum	68.425	decentralized	40.131
6	blog	60.363	defi	39.396
7	based	49.609	website	38

Rank	News	Betweenness Centrality	Non-news	Betweenness Centrality
8	web	45.995	dao	35.589
9	general	38.838	ethereum	34.894
10	defi	38.586	collection	32.084

Table 4: Top 10 words based on betweenness centrality

Figure 8 shows a semantic network visualization of the news sites. There were 39 nodes excluding solitary words and 242 connections. Thicker lines represent a higher co-occurrence frequency between words, and larger nodes indicate higher betweenness centrality. It is evident that the word "news" dominates in size, while words like "token" and "scientific" are significantly smaller. This implies that these words did not play a central role in structuring the introductory information on news sites.
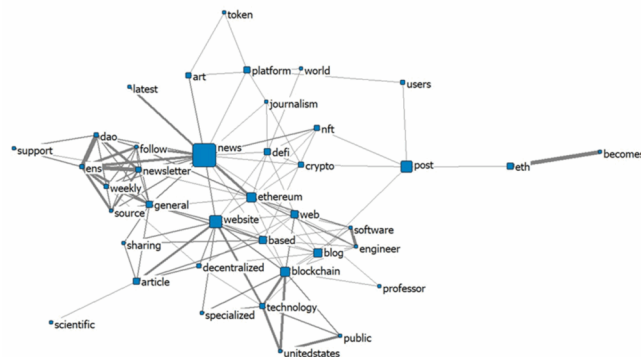

Figure 8: News sites semantic network visualization

Figure 9 shows a semantic network visualization of the non-news sites. There are 40 nodes and 256 connections. You can observe the presence of the word "website" in the non-news sites, but it is located further away from the center, with a smaller node size. For the news sites, "website" has various connections with words like "news," "article," and "ethereum." Conversely, "website" appears to be connected only to "nimi" and "nft." on the non-news sites
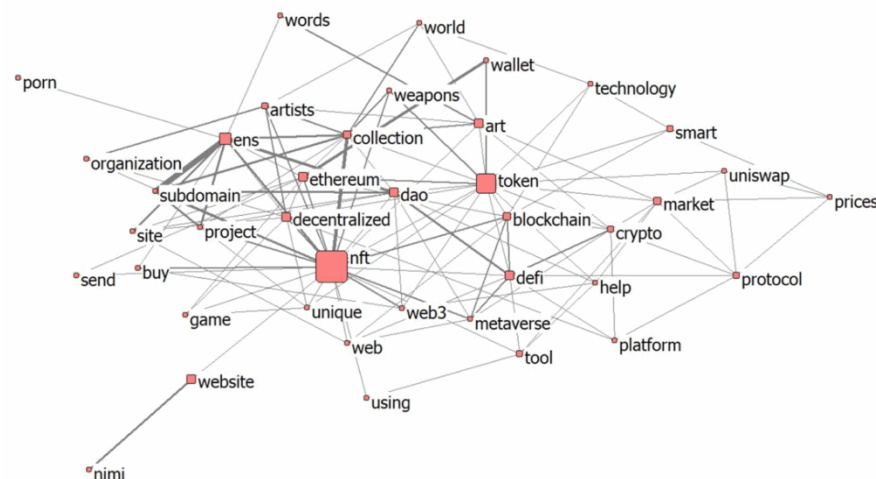

Figure 9: Non-news site semantic network visualization

5.2.2 Cluster analysis

Cluster analysis was conducted using the same words used in the semantic network analysis. Words within the same cluster are represented by the same color. Three clusters emerged for news sites: two large

clusters and one medium-sized cluster. There are 39 nodes and 121 connections. As shown in Figure 10, the words constituting each cluster vary. The largest cluster is related to blogs, indicating websites that post specialized content in specific fields or are operated by experts, as reflected by words like "specialized," "professor," and "scientific." The news-related cluster includes words like "journalism," "latest," and "post," which suggest that these websites are used for delivering the latest news. The newsletter-related cluster includes words like "weekly," "source," "ens," and "dao," indicating websites that periodically deliver news related to decentralized autonomous organizations through newsletters.
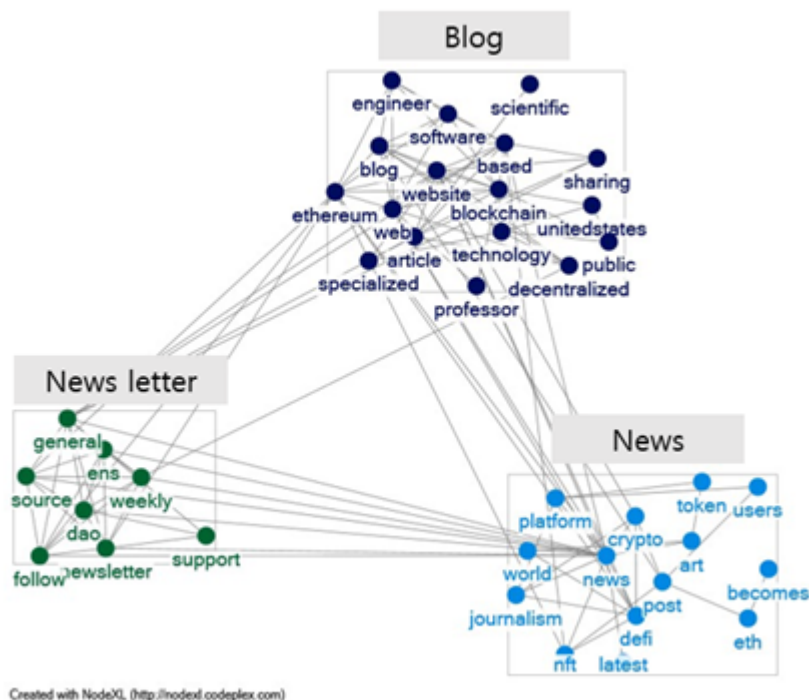

Figure 10: News sites cluster visualization

The non-news sites had six clusters: one large, three medium, and two small. There are 40 nodes and 123 connections. As shown in Figure 11, the largest cluster is the NFT cluster, featuring words like "weapons," "collection," and "buy." This indicates sites related to NFTs, including those showcasing NFTs such as weapons and allowing the purchase of NFT collections. The DeFi cluster includes words like "crypto," "market," "prices," and "uniswap," which relate to digital assets and trading. The token cluster includes words like "wallet," "send," "token," and "blockchain," indicating sites that issue tokens. The game cluster includes words like "game," "unique," and "decentralized," while the other content cluster includes words like "art," "technology," "world," and "words." Both the game and other content clusters appear to be related to various content services supported by decentralized websites in the Web3 environment. Finally, a Nimi cluster was formed, consisting of "website" and "nimi."
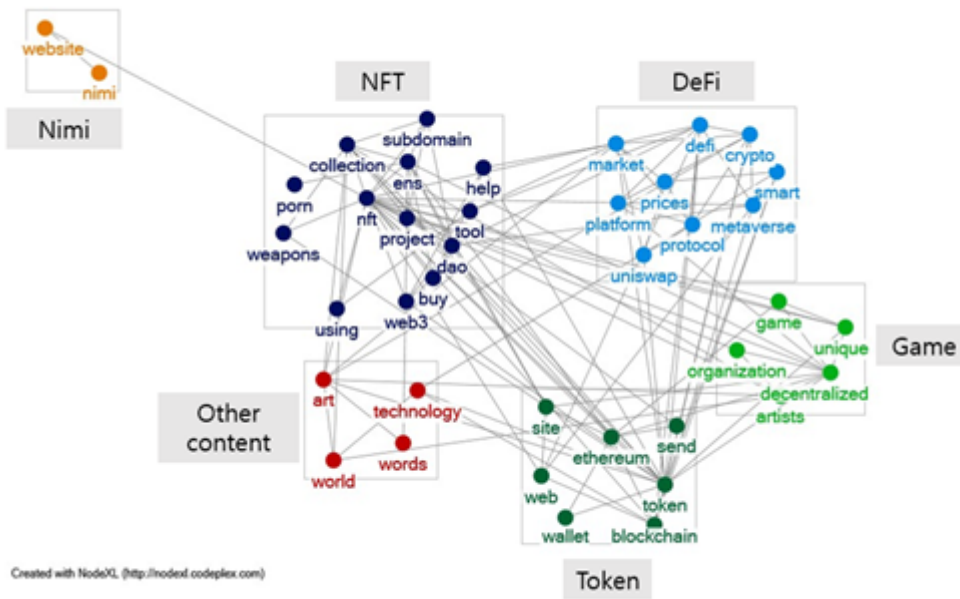
Figure 11: Non-news sites cluster visualization

6. Discussion and Conclusion

This study compared Web3 news websites and non-news websites through feature and topic analyses in anticipation of the upcoming Web3 era. The results show that news sites have more features than non-news sites in terms of design systems and functions, interactivity, information quality, and hyperlinks. This means that news sites are better than non-news sites in terms of website completeness. However, in case of digital assets, non-news sites, including cryptocurrency platforms, have more features than news sites. When analyzing the topics of the websites, we found that news sites had more terms related to information dissemination, such as "source," while non-news sites had more terms related to finance, such as "token." This is also reflected in the structural characteristics of the subcategories of news sites and non-news sites. For news sites, we found that the subcategories were more similar to each other in terms of the movement of information, while for non-news sites, we found that they clustered around the use of financial services, such as token issuance.

Although the results show differences between the two types of websites, focusing on the functional aspects of news media suggests that in the Web3 environment, news websites are just information delivery. Moreover, their content has continued to rely on news websites (in the Web 2.0 environment). Furthermore, specific Web3 features such as digital asset ownership appear to be limited relative to non-news websites. This situation is reminiscent of past transitions from legacy to new media. The content of early Web 2.0 news sites depended on traditional offline media content and merely transferred it to an online format (Boczkowski, 2004). The characteristics of the news sites discussed in this study reflect that phenomenon in a broader context, as they imitate the website design and functions of Web 2.0 and lack their own news content. This is also consistent with the view that the Web3 ecosystem is still dependent on the Web 2.0 structure (Zheng and Lee, 2023).

However, the distinctive feature that sets Web3 news websites apart from their Web 2.0 counterparts is the presence of the organizational operating system known as "DAO." Although more news websites are operated by individuals or small groups, some news sites are operated by organizations or institutions; these organizations and institutions are referred to as "DAOs." We found that "DAO" ranked high in word centrality and was strongly linked with terms like "newsletter" and "ENS" in the semantic network. On news websites, DAOs appear to function as institutional entities, particularly in the governance and

issuance of newsletters. DAOs, which are increasingly being adopted as operational systems in various organizations and companies within the Web3 ecosystem, operate transparently and securely based on blockchain technology in a democratic manner (Ding et al., 2022). This study empirically confirms that DAO systems stemming from the structural characteristics of Web 3.0 also apply to the news domain. These findings suggest that news organizations may change as the Web evolves.

As news media-related research on blockchain domain names is limited and that much of it is focused on technology development, the results of this study can be used as a basis for further Web3 journalism environment research. However, one limitation of this study is that its analysis focused on decentralized websites based on ENS domains. Websites that end with ".eth" are the most representative Web3 cases, but there are other types as well. To address this limitation, the researchers used two types of search portals that index websites using directories. Future research should expand the concept and scope of Web3 news websites beyond ENS to include various other website forms.

References

- Agrawal, V., Agarwal, A., Shah, S., Parmar, D., Rao, U. P., 2020. Decentralised Ecosystem for Journalism based on Blockchain. In IC-BCT 2019: Proceedings of the International Conference on Blockchain Technology. Springer, Singapore, pp. 7-19.
- Baekdal, T., 2009. What the Heck is Social News? [WWW Document]. Baekdal. URL http://www.baekdal.com/insights/social-news-explained (accessed 10.30.23).
- Bolter, J. D., Grusin, R., 2000. Remediation: Understanding new media. mit Press.
- Bhatia, A., Giometti, R., Nicolas, A., 2018. Decentralized social news platform. Tech. Rep, pp. 1-52.
- Boczkowski, P. J., 2004. The mutual shaping of technology and society in videotex newspapers: Beyond the diffusion and social shaping perspectives. The information society. 20(4), 255-267.
- Borgatti, S.P., Everett, M.G. and Freeman, L.C., 2002. Ucinet for Windows: Software for Social Network Analysis. Harvard, MA: Analytic Technologies.
- Cao, L., 2022. Decentralized ai: Edge intelligence and smart blockchain, metaverse, web3, and desci. IEEE Intelligent Systems. 37(3), 6-19.
- Carrington, P. J., Scott, J., Wasserman, S., 2005. Models and methods in social network analysis. Cambridge university press.
- Cebi, S., 2013. Determining importance degrees of website design parameters based on interactions and types of websites. Decision Support Systems. 54(2), 1030-1043.
- Ding, W., Hou, J., Li, J., Guo, C., Qin, J., Kozma, R., Wang, F. Y., 2022. DeSci based on Web3 and DAO: A comprehensive overview and reference model. IEEE Transactions on Computational Social Systems. 9(5), 1563-1573.
- Durand, A., 2022. Challenges with Alternative Name Systems. ICANN Office of the Chief Technology Officer., pp. 1–15.
- Galtung, J., Ruge, M. H., 1965. The Structure of Foreign News. Journal of Peace Research. 2(1), 64-91.
- Hansen, D., Shneiderman, B., Smith, M. A., 2010. Analyzing social media networks with NodeXL: Insights from a connected world. Morgan Kaufmann.
- Hermida, A., 2016. Social media and the news. The SAGE handbook of digital journalism., pp. 81-94
- Hermida, A., Domingo, D., Heinonen, A., Paulussen, S., Quandt, T., Reich, Z., Singer, J. B., Vujnovic, M., 2011. The active recipient: Participatory journalism through the lens of the Dewey-Lippmann debate. In International Symposium on Online Journalism.
- Himelboim, I., McCreery, S., 2012. New technology, old practices: Examining news websites from a professional perspective. Convergence. 18(4), 427-444.
- Jiang, K., Barnett, G.A., Taylor, L.D., 2016. Dynamics of culture frames in international news coverage: A semantic network analysis. International Journal of Communication. 10, 3710-3736.
- Kleinberg, J., 1999. Authoritative sources in a hyperlinked environment. Journal of the ACM. 46, 604–632.
- Kluver, A.R., Jankowski, N.W., Foot, K.A., Schneider, S.M., 2007. The Internet and national elections : A comparative study of web campaigning. Oxfordshire, Routledge.
- Krippendorff, K., 1980. Content analysis: An introduction to its methodology. Newbury Park: Sage.

- Libert, T., Binns, R., 2019. Good news for people who love bad news: Centralization, privacy, and transparency on us news sites. In Proceedings of the 10th ACM Conference on Web Science. pp. 155-164.
- Liu, L., Zhang, W., Han, C., 2021. A survey for the application of blockchain technology in the media. Peer-to-Peer Networking and Applications. 14(5), 3143-3165.
- Nielsen, R. K., 2015. Local journalism: The decline of newspapers and the rise of digital media. Bloomsbury Publishing.
- Osborn, G., Alan, N., 2023. Web3 Disruption and the Domain Name System: Understanding the Trends of Blockchain Domain Names and the Policy Implications. Available at SSRN 4498160.
- Park, H. W., Thelwall, M., 2003. Hyperlink analyses of the World Wide Web: A review. Journal of computer-mediated communication. 8(4), JCMC843.
- Python., 2023. Random — Generate pseudo-random numbers. Python Docs. https://docs.Python.org/3/library/random.html
- Qayyum, A., Qadir, J., Janjua, M. U., Sher, F., 2019. Using blockchain to rein in the new post-truth world and check the spread of fake news. IT Professional. 21(4), 16-24.
- Rebillard, F., Touboul, A., 2010. Promises unfulfilled¿Journalism 2.0', user participation and editorial policy on newspaper websites. Media, Culture & Society. 32(2), 323-334.
- Roose, K., 2021. Buy This Column on the Blockchain!. The New York Times. https://www.nytimes.com/2021/03/24/technology/nft-column-blockchain.html?smid=url-share
- Schneider, S. M., Foot, K. A., 2004. The web as an object of study. New Media & Society. 6(1), 114–122.
- Shoemaker, P. J., 2006. News and newsworthiness: A commentary.
- The IPFS Team., 2017. Uncensorable Wikipedia on IPFS. IPFS Blog & News. https://blog.ipfs.tech/24-uncensorable-wikipedia/
- Unzeelah, M., Memon, Z. A., 2022. Fighting against fake news by connecting machine learning approaches with web3. In 2022 International Conference on Emerging Trends in Smart Technologies (ICETST). IEEE. pp. 1-6.
- Upadhyay, A., & Baranwal, G. (2021, December). Fake news detection using ethereum blockchain. In International Conference on Advanced Network Technologies and Intelligent Computing (pp. 142-152). Cham: Springer International Publishing.
- Weber, M. S., Monge, P., 2011. The flow of digital news in a network of sources, authorities, and hubs. Journal of Communication. 61(6), 1062-1081.
- Weber, M. S., 2017. The tumultuous history of news on the web. UCL Press., London. pp. 83-100.
- Xiaoge, X., 2008. Engaging and empowering news audiences online: A feature analysis of nine Asian news websites.
- Zhang, W., Zhang, Q., Yu, B., Zhao, L., 2015. Knowledge map of creativity research based on keywords network and co-word analysis, 1992–2011. Quality & Quantity. 49, 1023-1038.
- Zheng, J., Lee, D. K. C., 2023. Understanding the Evolution of the Internet: Web 1.0 to Web3. 0, Web3 and Web3. Handbook of Digital Currency: Bitcoin, Innovation, Financial Instruments, and Big Data.
- Zheng, Z., Xie, S., Dai, H., Chen, X., Wang, H., 2017. An overview of blockchain technology: Architecture, consensus, and future trends. In 2017 IEEE international congress on big data (BigData congress). IEEE. pp. 557-564.