ROSA Journal

Contents

Editorial Board

Rafał Smoczyński, Instytut Filozofii i Socjologii Polskiej Akademii Nauk, Poland

Soyoung Park, Chosun University, South Korea

Seyed Mohammad J. Jalali, School of Science, Edith Cowan Unversity, Australia

Umer Shazad, Anhui University of Finance and Economics, China and University of Galway, Ireland

Xanat Meza, University of Tsukuba, Japan

Youjeong Kim, Texas State University, United States

YoungJu Shin, The George Washington University, United States

Publishing Information

Conceptual structure of innovation systems, a systematic approach through qualitative data analysis

Daniel Rocha Jimenez, Universitat Politècnica de València, España (daroc3@doctor.upv.es),
UNIMINUTO, Colombia
Norat Roig-Tierno, Universitat Politècnica de València, España (norat.roig@upv.es)
Alicia Mas Tur, Universidad de Valencia, España (alicia.mas@uv.es)
Pau Sendra-Pons, Universidad de Valencia, España (pau.sendra-pons@uv.es)

Abstract

This article analyzes the conceptual structure of the field of study of innovation systems through the establishment of subfields of knowledge using text mining. The data was sourced from the Web of Science in a search through Boolean operators under the concept "innovation systems" obtaining 6,706 articles in august 2020. For the analysis, a dictionary of categories has been constructed that agglomerate conceptually close families of words comprising keywords, theoretical and geographical approaches. Additionally, the QDA Miner & Wordstat software has been used to exploit the data. As results, it is possible to identify the central role of the concepts of triple helix and open innovation in the field of study of innovation systems; in the same way, it is possible to identify that the concepts of quadruple and quintuple helix related to the greater relevance of social and environmental issues on the agenda emerge strongly. And although the most discussed issues correspond to knowledge transfer, the science, and the economic development; the hottest issues result from the dynamics of innovation systems, its performance, and collaboration actions. Socio-technological changes in the innovation systems, mainly in North America, become less relevant as research interest, while new discussions arise in Europe and Africa on social and environmental sustainability issues. Public policies for innovation systems, management associated with new business models, and sustainability are topics of greatest interest for research. In this way the conceptual structure of the field and subfields of study of innovation systems is explained from the different visualizations through word clusters, the graphic representation of matrices and dendrograms.

KEY WORDS: Innovation Systems, National Innovation Systems, Regional Innovation, Conceptual Structure, Data Mining.

JEL Classification: 030, 039

1. Introduction

Freeman (1997) defines a national innovation system as the "the network of institutions in the public and private sectors whose activities and interactions initiate, import, modify and diffuse new technologies" (Atkinson, 2014), while to Carlsson (2002) a "Innovation systems can be defined in a variety of ways: they can be national, regional, sectoral, or technological. They all involve the creation, diffusion, and use of knowledge. Systems consist of components, relationships among these, and their characteristics or attributes". Anyway, national innovation systems and regional innovation systems are concepts associated with the competitive capacity of a region or territory (Lundvall et al., 2002), which is linked to the search for economic development (Fagerberg & Srholec, 2008) through regional development challenges (Edquist, 2010).

The analysis of the scientific literature for the case of the study of innovation systems has taken several directions: some works use bibliometric studies on innovation systems such as Chao et al. (2018), Uriona-Maldonado et al. (2012), Lopez-Rubio et al. (2020), or Lopez-Rubio et al. (2021). Other studies address the analysis of innovation systems through a literature review. Some examples are the ones conducted by Geels (2004) or Negro et al. (2012). Finally, other research in interdisciplinary studies outside innovation systems have deepened into the conceptual structure to improve the understanding of a given field of study such as Furrer et al. (2008), González-Loureiro et al. (2015) or López-Duarte et al. (2016).

This paper explores the use of data mining to study the conceptual structure of innovation systems using QDA miner and Wordstat, to find and explain the subfields of study and their relationships, based on existing data in the Web of Science.

Thus, the objective of this papers is to identify the role of the triple helix and open innovation in the conceptual structure of innovation systems, as well as identify how new concepts such as the quadruple and quintuple helix are positioned. Also, identify what are the new topics, or emerging topics in the discussions on innovation systems, as well as identify what are the hottest topics in the debates on innovation systems. Finally, identify how discussions about innovation systems change in response to geographic location in continents and subcontinents. This allows understanding the basic theories of the field of study, new trends or emerging topics, and mature or widely studied topics. With the aim of analyzing the conceptual structure associated with innovation systems, a set of concepts organized around three axes have been used: the axis of theoretical approaches; the axis of key elements, understood as the infrastructure for innovation, the macro aspects, processes, activities, and actors of innovation; and the axis of geographic associations classified by continents. The detailed development of the three axes mentioned above is described below.

This document is organized as follows: first, the theorical framework is built; then, the methodology used is reviewed, explaining the source of the data, the methods and the characteristics of the software used. Subsequently, a theoretical analysis of the correlations between the various concepts involved in the discussion on innovation systems is developed. As a next step, data mining software has been used to produce a set of tables and graphs, which facilitate the interpretation of the conceptual structure underlying the search for innovation systems. Finally, the conclusions of the article are outlined.

2. Theoretical framework

As theoretical approaches, the concepts of open innovation, social innovation, and triple helix, together with the term socio technology, which seeks to group different types of connections between technological and social issues. The concept of Open Innovation (Chesbrough, 2006; Christensen et al. 2005), fruit of innovation as a field of study, explores the open use between two or more organizations of resources, knowledge and capital in order to achieve common benefits, which is regulated by some kind of distribution agreement (Dahlander & Gann, 2010; Cooke, 2005) where this way of generating innovation is far from the individual innovator who generates with his own resources the creative destruction enunciated by Schumpeter (1976). Open Innovation thus overcomes the concept of individualism and leads to collective and organizational efforts.

The Triple helix, and later the Quadruple helix, explores the interaction of institutions such as the university, government, state and society (Geels, 2004) in the generation of initiatives that seek innovations linked to new interactions between actors (Etzkowitz & Leydesdorff, 2000; Etzkowitz et al., 2005) and which materialize in structures such as science and technology parks, incubators, and offices for the transfer of research results –although they could also exist through networks of companies in the so-called clusters. The three main agents included by the Triple Helix are the government, industry, and university. The Quadruple Helix includes to the last three agents of the Triple Helix knowledge and civil society (knowledge society). Finally, the Quintuple Helix adds society's natural environments (sustainability), while the Sextuple Helix integrates entrepreneurship.

Thus, the triple helix theories are at the center of the discussion on innovation systems. The interlinkages between actors are explained in the literature through the interaction between industry and science (Kaufmann & Tödtling, 2001), while other discussions focus on the interactions between public policies and

investments made by the state, with research and development, and knowledge transfer by universities (Meyer-Krahmer & Schmoch, 1998; Schartinger, et al. 2006).

On the other hand, a set of new concepts emerges associated with the relationship between technological change, environmental sustainability and the interplay between innovation and society (Carayannis and Morawska-Jancelewicz, 2022; Carayannis et al., 2019; López-Rubio et al., 2021). Thus, the first concept linked here is socio-technical (Geels, 2004), which explores the interaction between society and technology, where the use of technologies are conditioned by the society receiving this technology but at the same time technology can change the structure of society (Geels, 2005). Some approaches to this topic are the connections with social theories (Geels, 2010), the conceptual relationship of socio-technological systems with innovation systems (Coenen & López, 2010), the interaction between technology and institutions (Fuenfschilling & Truffer, 2014), and the dynamics of socio-technical transition (Geels, 2012).

An important role on innovation systems is the transition towards sustainable practices (Markard et al. 2012) in order to generate technological changes (Hekkert et al., 2007) that promote adjustments in activities and processes in companies, in line with the protection of the environment. In this sense, the studies of policies that induce the transition are relevant (Rogge & Reichardt, 2016). Köhler, et al. (2019) makes a review of transitions towards sustainability based on nine axes of analysis that include roles of society, business, and public policies, among others. In turn, social innovation (Phillips et al., 2015; Cajaiba-Santana, 2014) has taken on an important role in analyzing the purpose of actors in the development of society. These roles include new forms of interaction between actors, particularly civil society. In social innovation, the search for solutions and the social impact of solutions goes beyond the mere generation of wealth. This aim of generating not only economic but also social value leads to a new field of study for innovation issues (Van der Have & Rubalcaba, 2016). For the business world, the generation of new business models, linked to social innovation and its connection with sustainable development is a growing topic (Boons & Lüdeke-Freund, 2013).

The elements described above have been explained from the perspective of innovation in the literature through the concept of the quadruple helix, developed in a systematic literature review by Miller et al, (2018), as well as in discussions around the processes of production and transfer of knowledge, studied by Carayannis and Campbell (2012), Miller et al. (2016), or Del Giudice et al. (2017). In the quadruple helix, the role of society has been relevant as a new axis added to the triple helix. Regarding the concept of the quintuple helix, we turn to the works of Carayannis et al., (2018) and Grundel and Dahlström (2016) which imply the introduction of the socio-ecological transition, as a broader and more comprehensive way of adding the natural environment of society. This implies that sustainable development in a knowledge economy must co-evolve with the knowledge society (Carayannis et al. 2012).

Key elements in national and regional innovation systems

For Tödtling and Trippl (2005), innovation systems imply a set of differentiated innovation policies on a given territory. These innovation policies can include various types of instruments (Borrás & Edquist, 2013) where innovation is driven or promoted through the application of investment, the simultaneous presence of actors of the triple helix and the application of knowledge (Asheim et al., 2011). An example of the application of these policies is developed through different structures such as districts, science and technology parks, or others (Moulaert & Sekia, 2003). Specifically, it is common the specialization in certain fields or activities such as industries or industrial sectors particularly intensive in knowledge and where it is sought that the actors in that territory act under the logic of systems.

In this sense, the regional aspect related to the location of knowledge-based companies, the organization in sectoral clusters, the industrial specialization, and the establishment of focused policies are of vital importance to generate the development of innovation systems as part of national innovation capacity (Furman et al., 2002). In turn, regional systems are also a derivation of national systems (Cooke et al., 1998) that implies different varieties of regional innovation systems (Asheim,2007) that go through institutional learning, culture building and financial capabilities (Cooke et al., 1997), and the design and implementation of policy (Uyarra,2010; Uyarra & Flanagan,2010). Regional systems transform sectoral systems that have common characteristics such as technological base, technologies, inputs and demands (Malerba, 2002); comprehensively Asheim et al, (2016) analyses the past present and future of regional innovation systems.

Innovation systems seek localized development in a territory or region (Freeman, 1995) through intensive dynamics in science, the promotion of entrepreneurship, the search for economic, environmental, and social sustainability, with dynamics associated with technological change (Smith et al., 2010; Coenen et al.,2012; Hekkert et al., 2007).

Regarding public policies, Bergek et al. (2008) have criticized the lack of practicality in the design and implementation of these policies, particularly applied to regional innovation systems (Acs et al., 2002), although some other authors point out that it is possible to help policy makers and decision makers in the design, analysis and evaluation of policies applied to innovation systems (Woolthuis et al., 2005).

Thus, national, regional, sectoral, or technological innovation systems (Carlsson et al., 2002) involve a set of science, technology, and innovation activities (Ribeiro et al., 2010) that require structures in order to generate innovation dynamics and results in terms of the creation of new technologies, as well as attracting investment and the generation of wealth (Rodriguez-Pose & Crescenzi, 2008). In this sense, numerous organizations have emerged to provide a global network for scientific and innovation institutions and businesses, contributing to regional and local development. Some well-known international associations are the International Association of Science Parks and the World Technopolis Association, which also contribute to the publication of grey literature made available to society.

Based on the characteristics of innovation systems, it is important to include in the analysis the distribution of activities, organizational boundaries, coordination mechanisms, processes, and the final effectiveness of the systems for the development of technological innovations (Liu & White, 2001), i.e., the consolidation of structures and dynamics of the systems in order to achieve the expected performances, through network activities (Pittaway et al., 2004) where knowledge is produced and transferred (Asheim &; Coenen, 2005) as part of the cluster dynamics.

Under the concepts described above, the industry and industrial sectors developed through clusters become relevant, some of them explained in scientific articles such as tourism clusters (Hjalager, 2010), renewable energies (Jacobsson & Johnson, 2000), knowledge-intensive services (Muller & Zenker, 2001), or the analysis of clusters in countries (Wolfe & Gertler, 2004) or through business development understood from the perspective of the firm as new business models and strategies (Teece, 2010).

Entrepreneurship also has an important role in the discussion on innovation systems so we inquire about the effectiveness of incubators in entrepreneurship (Colombo & Delmastro, 2002), best practices (Bergek & Norman, 2008), the role of ecosystems in terms of their structure and influence (Spigel, 2017), as well as the studies on entrepreneurship that are at the center of this discussion (Acs et al., 2014; Autio et al., 2014).

For the purposes of this article, the countries have been classified by continent (Europe, Asia, Latin America, North America, and Africa). This analysis is important because of the different degrees of interest and maturity of regions in innovation systems and the importance of territory and region in innovation systems studies (Freeman, 2002). The discrimination of regions in developed and developing countries has also been analyzed from the perspective of bibliometric studies on innovation systems (Schmutzler et al., 2017; Toivanen & Ponomariov, 2011).

There are also studies associated with the interest of individual countries analyzed individually in studies of innovation systems, as in the case of Mexico (Solleiro & Castañón, 2005), or continental studies, as in the case of Europe (Cooke, 1992). However, Morgan (2004) dismisses the importance of geographical proximity in the discussion of innovation systems, while introducing other considerations of distance and proximity beyond geographical ones, such as socio-economic distances (Berry et al., 2010).

Over the last decades, the discussion of territorial or geographic aspects associated with regional or national innovation systems has been studied in great depth. More recently, discussions have focused on the role of knowledge in the development of clusters and the promotion of entrepreneurship.

For this article, the research and interest questions are the following:

• What is the role of the Triple Helix theory and Open Innovation in the conceptual structure of innovation systems?

• How are the new concepts of the quadruple and quintuple helix understood in the conceptual structure of innovation systems?

• How does the conceptual structure of innovation systems integrate with territorial logics in terms of continents and subcontinents of the globe? • What are the new discussions in the field of innovation systems and what are the most intense debates on innovation systems issues?

3. Methods and data

For the present work, a search was carried out in the Web of Science Core Collection as of July, 2020, using the respective Boolean operators and the various uses of the term "Innovation Systems" in the period 1900- 2020. The search yields a database consisting of information on scientific articles yielding a total of 6,706 documents structured by title, author, keywords, and abstract. This database grows at a rate of approximately 500 articles per year in the Web of Science database by observing the average annual production of the last decade.

QDA Miner is a qualitative methodological software that allows to study visual, graphical, and textual data, providing categorical and numerical information. This information is examined through exploratory, comparative, and descriptive analyses that determine coding patterns, regularities, relationships, and other properties (Watt, 2015). In this sense, WordStat can add value to the paper by analyzing text content and quantitative content through graphical and statistical tools (Udoh & Rhoades, 2006). The data processing is done by generating a dictionary of categories obtained which correspond to a set of keywords that describe that category following Tranfield et al. (2003). That allows to analyze the conceptual structure of the innovation systems in attention to the emerging topics and gaps, relevant gaps for future, hot topics with relevant gaps, and research of low interest for quality studies according to the figure 4 explained below.

The categories in turn are structured into key elements, theoretical approaches, and geographic location, thus building a dictionary with twenty-nine categories of analysis. As key elements we have used keywords frequently used in articles on innovation systems such as performance, knowledge, cluster, and others. A first filter to select these keywords has been to visualize the connections through the VoS viewer software, and to prioritize the keywords, and subsequently group them into categories or groups of conceptually similar families for analysis. For theoretical approaches, Open Innovation, Socio-technical, Social Innovation and Triple helix have been used. Open Innovation and triple helix have been selected for their relevance in the discussions on innovation systems (Huizingh, 2011; Cooke, 2005), while the transitions approaches (Markard, et al. 2012) here analyzed from the socio-technical transition and social innovation (Howaldt & Schwarz, 2011) have been chosen for their emerging character in the literature of the last two decades, hoping to understand the relationship of these theories with the concepts of quadruple and quintuple helix.

For geographical locations, groups of countries related to continents or subcontinents have been used, thus working with Europe, Asia, Latin America, Australia and Oceania, North America, and Africa. North America has been separated because of its long tradition and high production of scientific articles on the subject.

In this document, for the visualization of the results, the correlation matrices, the tables of principal factor extractions, the frequency distribution and the proximity graphs represented in the dendrograms have been used.

• Frequency distributions show the absolute and relative weight of the various categories in the volume of all words in the 6706 articles referring to the content of the title, abstract, keywords and abstract.

• The tables of extractions executed by means of the factorization analysis of non-negative matrices allow us to extract the main clusters that explain the behavior of the input data. In this case, a set of factors or word clusters, with their respective percentage weights that can explain the analyzed sample.

• The dendrograms show graphically how the words representing the various categories constructed for the dictionary are grouped according to the Jaccard index. The Jaccard index measures the similarity of sets, in this case articles that contain or do not contain a given word or category (word family).

• Correlation matrices show the location on a Cartesian plane of the various words or categories following a matrix factorization analysis, which can be graphically organized through word clusters.

4. Results

In the following paragraphs, the Dictionary of Categories is presented, as well as the various exploitations of the data through dendrograms, frequency distributions, the extraction of factors and, finally, the presentation of matrices, where these tables and graphs allow the conceptual structure of the innovation systems to be explained.

The dictionary of categories was constructed from the use of the key elements, theoretical approaches, and geographic locations described above, resulting in Table 1 in consolidated form as well as in Table A of the annex in detailed form, listing the words included in each category. The letters T and G are placed before the name of the category to indicate respectively Theory and Geography and are thus used in the following graphs and tables.

Number	Category
No.	**Key Elements**
1	Cluster
2	Collaboration
3	Dynamic
4	Development
5	Enterprise
6	Entrepreneurship
1	Cap_Inno (Innovation Capacity)
8	Industry
9	Knowledge
10	Management
11	National innovation
12	Performance
13	Policy
14	Regional innovation
15	Science
16	Sustainability
17	System
18	Transfer
No.	**Theories**
1	T Social innovation
2	T__Open innovation
3	T__Socio technological
4	T__triple helix
No.	**Geography**
1	G__Africa
2	G__Asia
3	G__Australia and Oceania
4	G__Europa
5	G__Latam

Table 1: Dictionary of Categories

Data mining

Below is a set of visualizations resulting from the exploitation of Web of Science data through data mining, with their respective analyses.
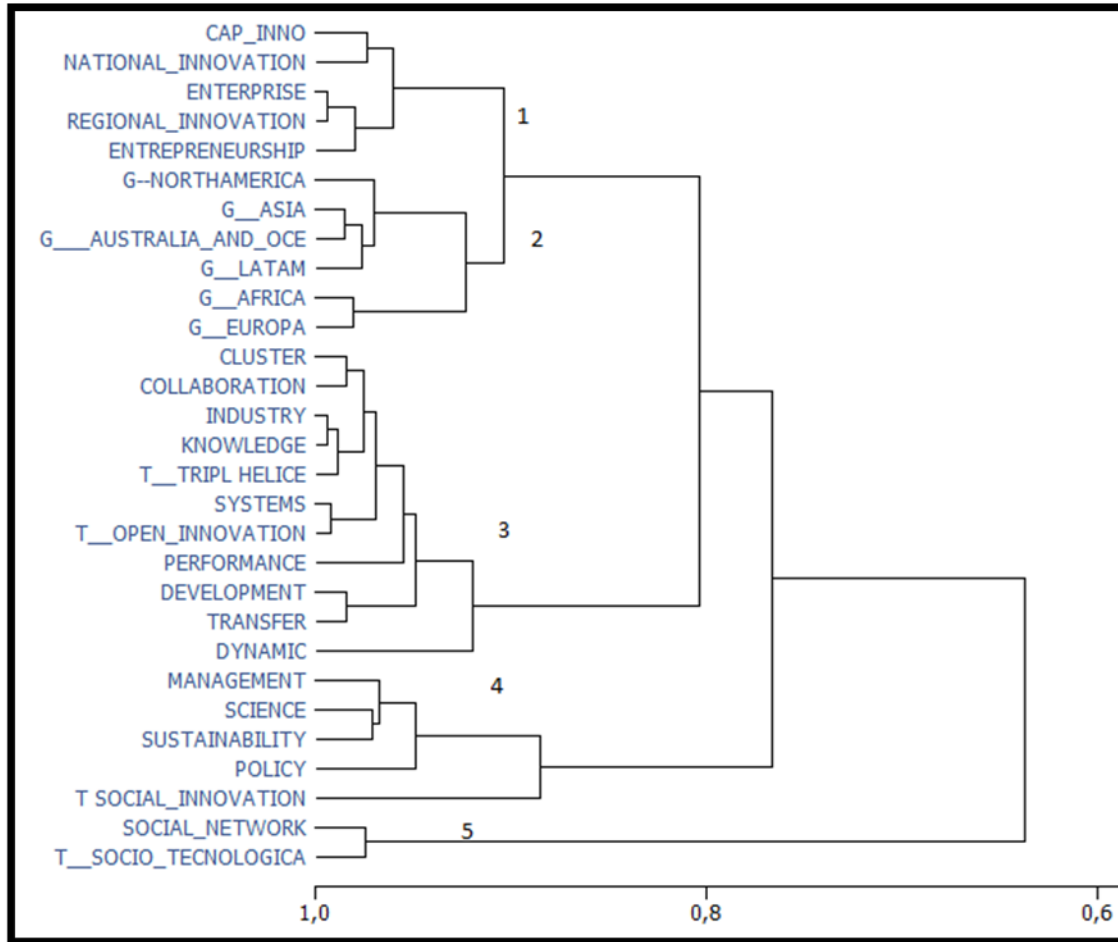
Figure 1: Dendrogram based on Similarity Index. Source: Own elaboration. (Wordstat/ Heatmap / Keyword dendogram).

Figure 1 presents a dendrogram extracted from QDA & Wordstat showing the groupings of the different categories of the dictionary of word families based on the Jacaard similarity index. The dendrogram shows large clusters numbered from 1 to 5 where it is possible to understand the word clustering logic. Thus, it is worth highlighting a grouping associated with the geographical themes indicated in the dendrogram with the number 2. This cluster agglutinates Africa, Europe, Asia, Australia and Oceania and Latin America. It is interesting to note that North America, which includes only the United States and Canada, is found in this cluster 2 separately, which can be explained by the high volume of scientific articles located in these two countries and the maturity of the studies since the initial articles by Freeman (1995) or Furman et al. (2002). In North America (Canada and the United State), discussions have evolved to other topics more related to the dynamics and performance of innovation systems that involve the cluster 1. This Cluster on the other hand, concentrates on topics associated with entrepreneurship and regional and national innovation systems. It is understandable then that, for most of the countries of the different continents and subcontinents, innovation systems, are strongly linked to entrepreneurship and entrepreneurs from the perspective of national and regional systems.

Cluster 4, which includes new trends, comprises the topics of sustainability and social innovation as new discussions around innovation systems and has articles where the main exponents are Markard et al. (2012) and Coenen et al. (2012). The presence of the management category is also coherent, being understood as the search for new business models associated with this trend. The Policy category, on the other hand, explains the interest of public policies to generate more sustainable and inclusive technological changes.

Cluster 5 shows the social issues related to social networks, as well as socio-technical changes, which have

the lowest Jacaard index of the various clusters with an index of approximately 0.65. These two topics are more dissociated than the other clusters in terms of the field of study of innovation systems. Although the value of 0.65 is relevant, the social issues related to socio-technical transitions are not fully articulated with the central discussion of innovation systems. The reason for this situation is that it is an emerging issue. It is interesting to note that sustainability issues, on the other hand, do have a correlation level above 0.9, which indicates that the term appears with a very high frequency with the other key terms that explain innovation systems. Thus, transitions towards sustainability are closer to the discussion of innovation systems than discussions of social issues.

Clusters 1, 2 and 3 represent the categories that explain the processes and expected results of innovation systems and show the heart of the discussion of innovation systems, with Jacaard indices above 0.9, indicating a very high level of correlation. In this sense, the described themes explain with a high level of correlation or connection to the field of study of innovation systems, which in turn indicates how unified the conceptual structure of innovation systems is.

Clusters 1, 2, 3 and 4 together concentrate a large part of the scientific literature on innovation systems with a Jaccard index close to 0.9, leaving out clusters 5, which implies social clusters that are still far from the heart of the discussion on innovation systems.

Frequency distributions

Table 2 shows the frequency distribution for the different words that make up the dictionary of categories previously shown in Table 1.

By drawing an imaginary line in this frequency distribution of Table 2 corresponding to the median value of 6%, it is observed that the most relevant categories are: Cap_Inno (12.56%), industry (10.47%), transfer (7.74 %), development (7.01 %), management (6.30%), systems (6.46%), performance (6.79%). The triple helix as a theoretical approach also appears in this block of categories with the highest frequency with a value of 7.17 %. On the other hand, below this 6%-line, theoretical approaches and geographical locations described by continents and subcontinents are generally less relevant than the other categories of families in the dictionary.

Figure 2 is constructed, which is referred to in the following subsection, by combining the dendrogram in Figure 1 with the frequency Table 2 of the dictionary categories.

Number	Category	Frecuency	% Relative
1	CAP_INNO (Innovation Capacity)	30018	12,56%
2	INDUSTRY	25016	10,47%
3	TRANSFER	18503	7,74%
4	T__TRIPL HELICE	17134	7,17%
5	DEVELOPMENT	16751	7,01%
6	PERFORMANCE	16234	6,79%
7	SYSTEMS	15432	6,46%
8	MANAGEMENT	15046	6,30%
9	SCIENCE	12145	5,08%
10	REGIONAL_INNOVATION	11424	4,78%
11	KNOWLEDGE	9262	3,88%
12	SUSTAINABILITY	6647	2,78%
13	G__ASIA	5868	2,46%
14	G__EUROPA	5749	2,41%
15	ENTREPRENEURSHIP	5432	2,27%
16	NATIONAL_INNOVATION	4594	1,92%
17	POLICY	4403	1,84%
18	COLLABORATION	3854	1,61%
19	DYNAMIC	2779	1,16%
20	T__OPEN_INNOVATION	2751	1,15%

Number	Category	Frecuency	% Relative
21	T SOCIAL_INNOVATION	2379	1,00%
22	ENTERPRISE	2285	0,96%
23	CLUSTER	2199	0,92%
24	G__AFRICA	1315	0,55%
25	G__LATAM	1052	0,44%
26	T__SOCIO_TECNOLOGICA	263	0,11%
27	G___AUSTRALIA_AND_OCE	165	0,07%
28	G—NORTHAMERICA	148	0,06%
29	SOCIAL_NETWORK	138	0,06%

Table 2: Frequency distribution for the categories

For the construction of the figure, the words of each category corresponding to the different branches of the dendrogram have been added, according to the clusters established there. In this way, cluster 1, obtains a percentage of 44.64%, which is obtained by adding the data corresponding to Cap_Inn (Innovation Capacity), National Innovation, Enterprise, Regional Innovation, Entrepreneurship, which as shown in Table 2, have a percentage of participation of 12.56%, 1.92%, 0.48%, 4.78%, and 2.27% respectively.

The figure also shows the corresponding sums to indicate the cumulative percentages of each of the branches and clusters of the dendrogram, as explained in the previous paragraph.
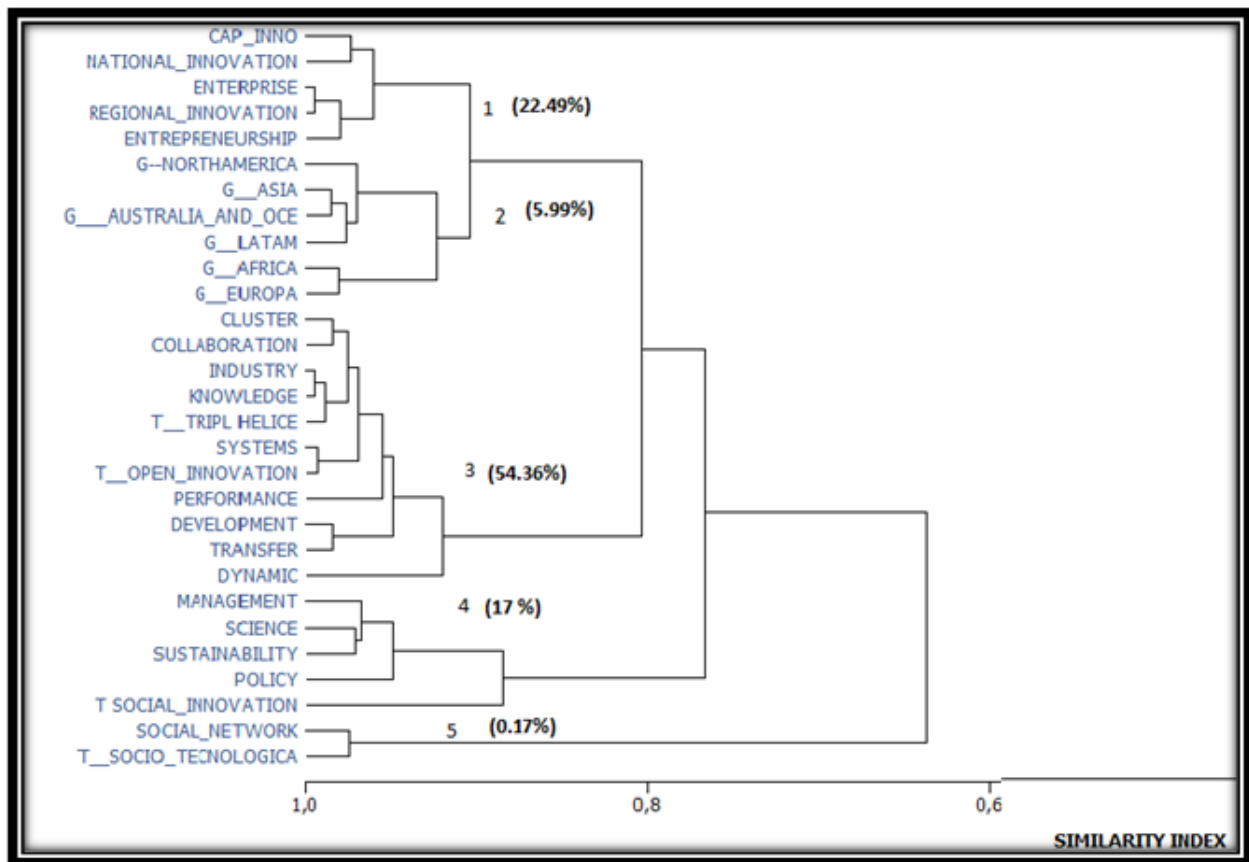

Figure 2: Combined dendrogram and frequency distribution analysis. Source: Own elaboration. (Wordstat/ Heatmap / Keyword dendogram).

Cluster 1 with 22.49%, cluster 3 with 54.36% and finally cluster 4 with 17% should be highlighted. These four clusters represent 93.85 % of the frequencies, which could amply explain the field of study. From

the Non-negative matrix factorization (NMF or NNMF) analysis, executed for the dictionary of categories shown in Table 3 using QDA & Wordstat software, an extraction of the main topics shown in Table 2 has been performed. The resulting extraction comprises 5 main topics which have been nominated for the present article as "New trends", "Collaboration", "Knowledge & regional economy", "Science and Technology & triple helix", and finally Innovation Systems, and which fully correspond to clusters 1, 2,3, 4 and 5 explained above. Thus, the Table 3 indicates the clusters (extracted with QDA miner & Wordstat) that includes the name assigned by the authors to each cluster, the dictionary categories (see Table A) include in each cluster, the frequency and number of cases included, the coherence and the % of the corresponding cases. The presence of these five clusters is explained below, where the "Innovation System" cluster is the result of the search for "innovation systems" in WoS described in the methodology section of this article and therefore comprises 98% of the cases. • The "new trends" cluster explains the current interest in sustainability issues and social change, which comprises about 25 % of the cases in this database. • The word Cooperation (36%) has been used to name this cluster to differentiate it from the word Collaboration, which is already in the dictionary of categories. • "Knowledge & regional economy" is a cluster that comprises 83% of the cases and includes all the families of words related to Regional Innovation, Performance, Entrepreneurship and Development overlapped with the cluster named "Cooperation". • The cluster "Science and Technology, & triple helix", comprises in the center of the cluster the word triple helix, and an associated group of words related to science, technology, and innovation. This corresponds to a value of 93 % of the cases.

With the presence of these four-word clusters, it is possible to explain the structure of the innovation systems, where it is relevant to show the percentage or relative weight of each cluster. The presence of the cluster of words "Science and technology & triple helix" and "Knowledge & regional economy" is consistent with the findings of Cooke (2005) who identifies the triple helix and regionalization as elements of approximation in the discussion of regional innovation systems. Similarly, "new trends" is a word cluster consistent with Markard et al.'s (2012) transitions to sustainability. The cluster of words called "cooperation" is consistent with the logic of the logic of open innovation, but clearly includes other types of cooperation such as social innovation and social networks, among others.

No	Topic	Keywords	Coherence	Freq	Cases	% Case
1	Innovation systems	SYSTEMS; CAP_INNO; PERFORMANCE; INDUSTRY; DEVELOPMENT; INNOVATION SYSTEMS; REGIONAL INNOVATION SYSTEMS; NATIONAL INNOVATION SYSTEMS; SYSTEMS OF INNOVATION	0,399	55978	6538	97,49%

No	Topic	Keywords	Coherence	Freq	Cases	% Case
2	Science and technology & triple helix	SCIENCE; TRANSFER; NA-TIONAL_INNOVATION; G_ASIA; PERFOR-MANCE;"T_TRIPL HELICE"; DEVELOPMENT; INDUSTRY; SCIENCE AND TECHNOLOGY; ECONOMIC DEVELOPMENT; TECHNOLOGY TRANSFER; INNOVATION SYSTEM; RESEARCH AND DEVELOPMENT; KNOWLEDGE TRANSFER; INNOVATION PERFORMANCE; SUSTAINABLE DEVELOPMENT	0406	38301	6246	93.14%

No	Topic	Keywords	Coherence	Freq	Cases	% Case
3	Knowledge & regional economy	REGIONAL_INNOVATION; G_EUROPA; ENTREPRENEUR-SHIP; DEVELOPMENT; PERFORMANCE; ECONOMIC DEVELOPMENT; RESEARCH AND DEVELOPMENT; SUSTAINABLE DEVELOPMENT; INNOVATION PERFORMANCE; REGIONAL DEVELOPMENT; INNOVATIVE DEVELOPMENT; INNOVATION AND ENTREPRENEUR-SHIP; TECHNOLOGY DEVELOPMENT; INNOVATION DEVELOPMENT; TECHNOLOGI-CAL	0,354	22310	5522	82.34%

No	Topic	Keywords	Coherence	Freq	Cases	% Case
4	Cooperation	COLLABORATION; SO-CIAL_NETWORK; T_OPEN_INNOVATION; KNOWLEDGE; CLUSTER; KNOWLEDGE TRANSFER; KNOWLEDGE ECONOMY; KNOWLEDGE CREATION; KNOWLEDGE BASES; KNOWLEDGE PRODUCTION; KNOWLEDGE FLOWS; KNOWLEDGE MANAGEMENT	0,328	6705	2400	35.79 %
5	New trends	T_SOCIO_TECNOLOGICA; SUSTAINABIL-ITY; DYNAMIC; SUSTAINABILITY TRANSITIONS	0,277	3804	1646	24.55 %

Table 3: Five clusters according to the main topics in the innovation systems literature (drawn from the QDA miner & Wordstat software)

Word Cluster Matrix

Figure 3 shows the 4 clusters mentioned above: "New trends", "Cooperation", "knowledge & regional economy", and finally "Science and technology & triple helix".
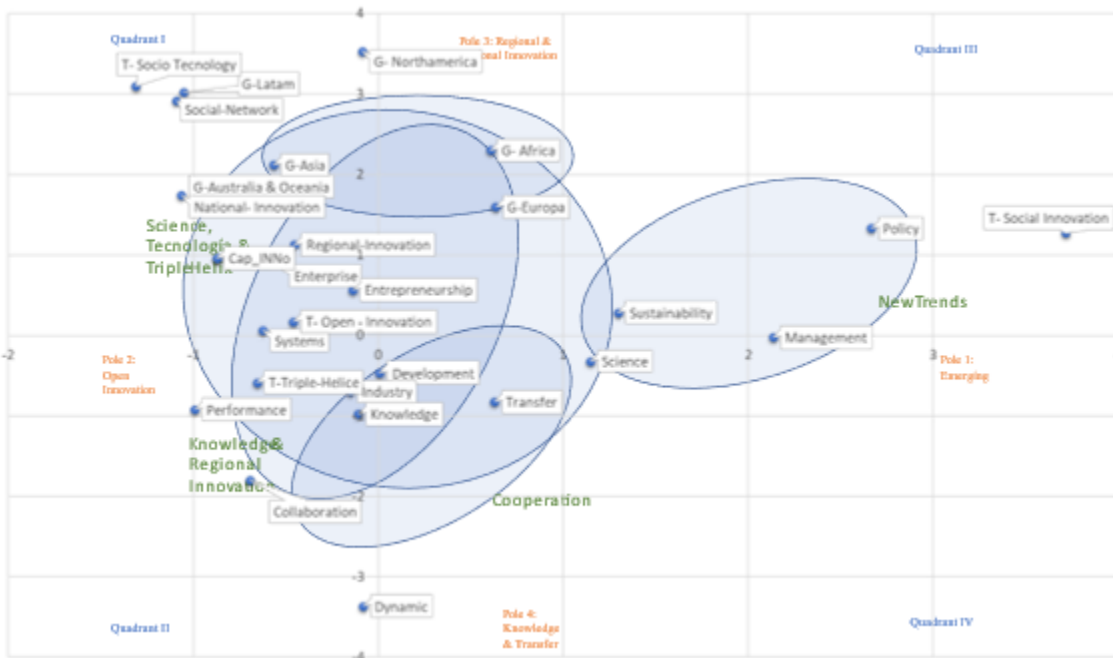

Figure 3: Word Cluster. Source: Own elaboration.

These clusters are complemented in Figure 1 by the presence of the theoretical approaches corresponding to Triple Helix, Open Innovation, Social Innovation and Socio Technical. Thus, it is observed that the concepts Triple Helix and Open Innovation are close in the Cartesian diagram due to their high concurrence in the scientific literature, which in turn is since the development of Open Innovation involves in practice the interaction with various business, state, university, or social actors, in collaboration schemes that correspond to the triple helix approach. Similarly, these two theoretical approaches (Triple Helix and Open Innovation) are located at the center of three of the large clusters in Figure 3.

The topics of transition to sustainability and social innovation are more closely associated with the "new trends" cluster due to their emerging nature, which explains their separation from the other three clusters in Figure 3. This is consistent with the fact that the literature shows that scientific production on these topics corresponds to the last two decades.

On the other hand, the approaches associated with "Socio-technical" transitions appear closer to the concepts of national innovation and regional innovation, since innovation systems are characterized by their focus on technological change.

In Figure 3 we have also added the geographical focus represented by the continents or subcontinents, in this case, Latin America, Europe, Asia, Australia and Oceania, and finally North America. Additionally, the four poles or extremes of the Cartesian diagram have been added, called Pole 1 "Emerging", Pole 2 "Open Innovation", Pole 3 "Regional & National Innovation", and finally Pole 4 "Knowledge &Transfer". Finally, in Figure 2, quadrant designations such as I, II, III and IV have also been added. The above to develop the explanation that follows:

In relation to developments based on geographical aspects, it is possible to identify that Latin America, Australia and Oceania, as well as Asia, are located in quadrant I y IV, which can be understood by the search of these regions to implement and use innovation systems to mobilize innovation indicators. They are also very close to the "Science and Technology, & triple helix" cluster, which is characterized by the development of regional and national innovation systems, the triple helix and innovation structures.

North America is in quadrant IV, closer to the words National Innovation and Regional Innovation, explained by the profusion of research in this field of study in this particular region and by interests closer to the performance of innovation systems.

Europe and Africa appear to be very close. One explanation for this proximity could be the fact of being halfway between the issues of sustainability and social innovation corresponding to the emerging issues of social and environmental concerns, and on the other hand the search for performance and knowledge-intensive industrial clusters. Although in the case of Europe, important developments in innovation systems have been developed in the last decades in contrast to Africa and its underdeveloped countries.

In relation to the extremes or poles of the Cartesian diagram, these have been denominated on the x-axis, "Emerging" and "open innovation", which reflects a tension between the new themes and interests represented here by Social Innovation and Sustainability, and the tradition of innovation systems represented by the theoretical approach of Open Innovation.

On the other hand, the y-axis represents on one side the discussion of regional and national innovation systems dominated by discussions of technological change, and on the other side the expected results or products of the innovation systems measured in knowledge, transfers, and dynamics of the innovation systems that appear loose from the word clusters and very close to the end of the negative axis of the ordinates.

Gaps in qualitative analysis

Figure 4 shows the matrix for detecting gaps in qualitative research that according to González- Loureiro et al, (2015) allows identifying topics of interest for future studies (Quadrant IV), topics of current debate (Quadrant III), and topics in decline (Quadrant I), and emerging topics (Quadrant II). In this way, the various categories can be organized according to this matrix as shown in Figure 4.
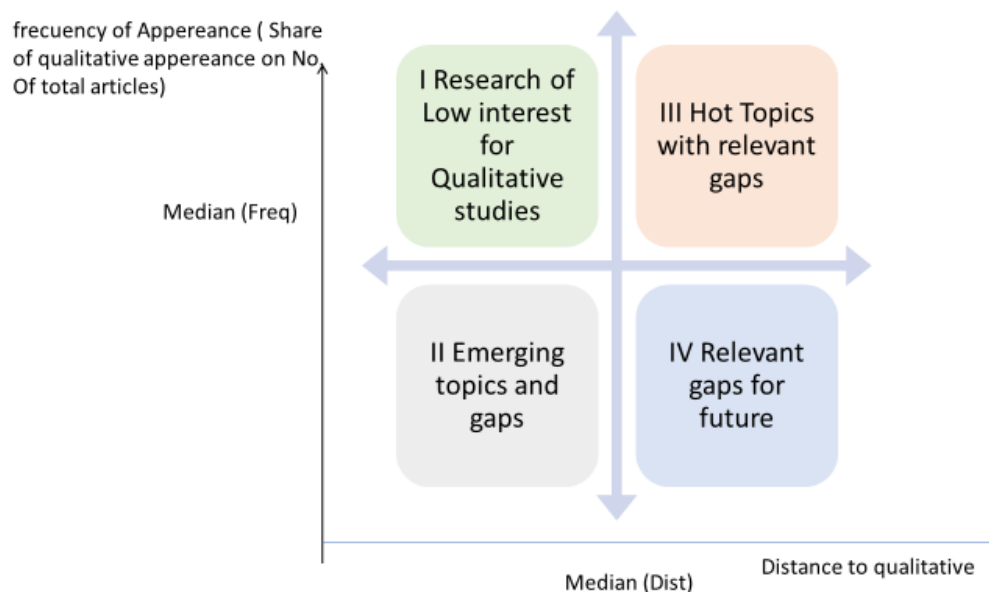

Figure 4: The matrix to detect. Source: Adapted from González-Loureiro et al, (2015).

Through Figures 3 and 4, some conclusions are reached:

• The topics that could be of greatest interest in the future, located in quadrant III, are Policy, Social innovation, Management issues, as well as Sustainability. Additionally, studies associated with Europe and Africa should be considered.

• The topics on which there is currently most debate, corresponding to quadrant IV, are Science, Development, and Transfer.

• The topics of little interest are those associated with North America and related to Socio Technology. These appear in quadrant I.

• Emerging topics include, Performance, Industry, Dynamic and Collaboration. Quadrant II.

It is interesting to observe the central role of the concepts of open innovation and triple helix in the discussions of innovation systems, as result of the efforts of university, enterprise and state interaction, and multiple collaboration schemes. Likewise, it is observed that the discussions on innovation systems move towards aspects of environmental and social sustainability, related to the implementation of public policies for sustainability, and the development of new business models in the logic of the quadruple and quintuple helix. On the other hand, it is observed that socio-technological transitions lose strength, while the central discussions spin around economic development related to the performance and dynamics of innovation systems, in accordance with the essence of innovation systems derived from the transfer of knowledge and the intensive use of science. While discussions in North America on socio-technological issues lose strength, in Europe and Africa sustainability issues are of greater interest in research.

5. Conclusions

According to previous work, the use of data mining software allows the exploitation of data from the Web of Science or other databases through the generation of various types of reports such as dendrograms, factor extraction using the non-negative matrix factorization technique as well as the visualization of word clusters in the Cartesian plane.

In the case of the field of study of innovation systems, the conceptual structure is understood in essence through three large word clusters, which in this article have been called "Cooperation", "Science & technology, & triple helix", and finally "Knowledge & regional innovation". These clusters individually represent 36%, 83% and 93% of the words used respectively with overlaps. The combined cluster of these three clusters explains 97% of the discussions on innovation systems studies. A fourth cluster, called "New Trends" in this article, which represents 24% of the discussions on innovation systems, corresponds to the topics of sustainability and social innovation, which could be understood in association with the "Management" category, linked to new business models in the markets for sustainable and socially responsible products and services, where a meeting point of these terms is found around the circular economy.

Innovation systems pivot around the discussion of the triple helix as a tradition and emerging issues linked to sustainability and socio-technical transition. In this sense, open innovation, and the triple helix form part of the theoretical heart of this field of study, while emerging issues focus on environmental sustainability and social issues. Similarly, innovation systems are torn between regional and national tensions on the one hand, and issues of knowledge transfer and generation on the other.

The link between the socio-technical transition and the socio-ecological transition gives meaning to the presence of the quadruple and quintuple helix in the conceptual structure of innovation systems. Notwithstanding the above, transitions towards sustainability are fully integrated into the discussion of innovation systems, while social issues are emerging.

Entrepreneurship and economic development are anchored in the discussion of innovation systems as the expected end of innovation systems, and therefore at the center of the field of study, consistent with studies since Schumpeter's time.

In relation to geographical issues, discussions in Latin America, Asia, Australia, and Oceania focus on the importance of regional and national innovation systems, possibly linked to the effort to develop innovation systems to boost innovation indicators and regional development. In the case of Latin America, the weak infrastructure in terms of innovation systems makes it necessary to generate enabling conditions and consolidate capacities, so the concern is focused on these terms.North America, a region that has been

extensively researched in terms of innovation systems, debates revolve around Enterprise, Entrepreneurship, Regional and National Innovation. Unlike Latin America, Asia, Australia and Oceania, the focus is on the performance of innovation systems, rather than on the enabling conditions of innovation systems.

In the case of Europe, it is moving towards new discussions on environmental and social issues, probably linked to the Paris Agreement on climate change, the Circular Economy and social innovation agendas. Finally, and related to the emerging issues of socio-technical and socio-ecological transition.

Africa, tied to underdevelopment, is likely to seek to initiate innovation systems based on sustainability agendas and social issues, given the importance of sustainability in a continent with simultaneous abundance and scarcity of natural resources, as well as deep poverty.

It is interesting to note that the role of the theoretical approaches associated with technological transitions or changes remains distant from the three major clusters that define the field of study of innovation systems, which can be interpreted as the gradual abandonment of these approaches, and how the shift towards new social and environmental transitions, explicit in the quadruple and quintuple helix, has been explained. In this way, the importance of technological change, by itself, has been replaced by the emergence of changes or transitions focused on the environmental and social.

In terms of the research question posed in the article, the triple helix and open innovation are at the intersection of the three major word clusters, whereby these theories individually explain each and every one of the major clusters. This places these theories at the center of discussions of innovation systems. Nevertheless, the emergence of the quadruple and quintuple helix concepts linking society and sustainability issues respectively to the old triple helix concept is observed in the conceptual structure as a new layer that is being integrated into the body of knowledge on innovation systems. In this way, the transitions towards sustainability and the social end up being an expected projection of the discussions on innovation systems.North America as Socio Technological change are not of major interest among academics because of the depth of research.

Future topics of interest could be related to Africa, Europe, social innovation, and public policies all related to socio-technical and environmental transition –Africa topics are understood in terms of the low real implementation of innovation systems and related to the low production of scientific articles. Europe is very interesting in terms of the transition of innovation systems towards new sustainability and social issues. Finally, the issues of social innovation and its incorporation into innovation systems in the territories, and the regional could be very interesting.

Another interesting analysis could be how public policies are integrated with innovation systems, especially in relation to the emerging issues of sustainability and social issues, or how public policies feed the impulse of innovation systems based on the quadruple and quintuple helix.

Sustainability is a topic of great discussion and interest, and numerous studies can be foreseen in this field, as well as the social innovation.

However, a new field of discussion is emerging around public policy and management as drivers to achieve sustainability transitions. In the field of management, it could be configured around new business models and strategies, while for public policies, longer-term approaches to include environmental sustainability and inclusive approaches to respond to social demands on productive apparatuses represented by companies and industrial clusters nested in innovation systems.

Footnotes

1. Vos Viewer. https://www.vosviewer.com/

References

- Acs, Z. J., Anselin, L., & Varga, A. (2002). Patents and

- innovation counts as measures of regional production of new knowledge. Research policy, 31(7), 1069-1085.
- Acs, Z. J., Autio, E., & Szerb, L. (2014). National systems of entrepreneurship: Measurement issues and policy implications. Research Policy, 43(3), 476-494.
- Asheim, B. T., & Coenen, L. (2005). Knowledge bases and regional innovation systems: Comparing Nordic clusters. Research policy, 34(8), 1173-1190.
- Asheim, B. (2007). Differentiated knowledge bases and varieties of regional innovation systems. Innovation, 20(3), 223-241.
- Asheim, B. T., Grillitsch, M., & Trippl, M. (2016). Regional innovation systems: Past–present–future. Handbook on the Geographies of Innovation.
- Asheim, B. T., Boschma, R., & Cooke, P. (2011). Constructing regional advantage: Platform policies based on related variety and differentiated knowledge bases. Regional studies, 45(7), 893-904.
- Atkinson, R. D. (2014). Understanding the US national innovation system. ITIF, June.
- Autio, E., Kenney, M., Mustar, P., Siegel, D., & Wright, M. (2014). Entrepreneurial innovation: The importance of context. Research policy, 43(7), 1097-1108.
- Bergek, A., & Norrman, C. (2008). Incubator best practice: A framework. Technovation, 28(1-2), 20-28.
- Bergek, A., Jacobsson, S., Carlsson, B., Lindmark, S., & Rickne, A. (2008). Analyzing the functional dynamics of technological innovation systems: A scheme of analysis. Research policy, 37(3), 407-429.
- Berry, H., Guillén, M. F., & Zhou, N. (2010). An institutional approach to cross-national distance. Journal of international business studies, 41(9), 1460-1480.
- Borrás, S., & Edquist, C. (2013). The choice of innovation policy instruments. Technological forecasting and social change, 80(8), 1513-1522.
- Boons, F., & Lüdeke-Freund, F. (2013). Business models for sustainable innovation: state-of-the-art and steps towards a research agenda. Journal of Cleaner production, 45, 9-19.
- Cajaiba-Santana, G. (2014). Social innovation: Moving the field forward. A conceptual framework.Technological Forecasting and Social Change, 82, 42-51.
- Carayannis, E. G., & Campbell, D. F.. (2012). Mode 3 knowledge production in quadruple helix innovation systems. In Mode 3 knowledge production in quadruple helix innovation systems (pp. 1-63). Springer, New York, NY.
- Carayannis, E. G., Barth, T. D., & Campbell, D. F.. (2012). The Quintuple Helix innovation model: global warming as a challenge and driver for innovation. Journal of innovation and entrepreneurship, 1(1), 1-12.
- Carayannis, E. G., Grigoroudis, E., Campbell, D. F., Meissner, D., & Stamati, D. (2018). The ecosystem as helix: an exploratory theory-building study of regional co-opetitive entrepreneurial ecosystems as Quadruple/Quintuple Helix Innovation Models. R&D Management, 48(1), 148-162.
- Carayannis, E. G., & Morawska-Jancelewicz, J. (2022). The Futures of Europe: Society 5.0 and Industry 5.0 as Driving Forces of Future Universities. Journal of the Knowledge Economy, 1-27.
- Carayannis, E. G., Grigoroudis, E., Stamati, D., & Valvi, T. (2019). Social business model innovation: A quadruple/quintuple helix-based social innovation ecosystem. IEEE Transactions on Engineering Management, 68(1), 235-248.
- Carlsson, B., Jacobsson, S., Holmén, M., & Rickne, A. (2002). Innovation systems: analytical and methodological issues. Research policy, 31(2), 233-245.
- Chao, C. H., Wu, W. N., Tsai, T. H., Hsu, C. W., & Chen, M. C. (2018, July). Visualizing the intellectual structure and evolution of innovation systems research: a bibliometric analysis. In Proceedings of the 5th Multidisciplinary International Social Networks Conference (p. 27). ACM
- Chesbrough, H. W. (2006). The era of open innovation. Managing innovation and change, 127(3), 34-41.
- Christensen, J. F., Olesen, M. H., & Kjær, J. S. (2005). The industrial dynamics of Open Innovation-Evidence from the transformation of consumer electronics. Research policy, 34(10), 1533-1549.
- Coenen, L., & López, F. J. D. (2010). Comparing systems approaches to innovation and technological change for sustainable and competitive economies: an explorative study into conceptual commonalities, differences and complementarities. Journal of cleaner production, 18(12), 1149-1160.
- Coenen, L., Benneworth, P., & Truffer, B. (2012). Toward a spatial perspective on sustainability transitions.Research policy, 41(6), 968-979.

- Colombo, M. G., & Delmastro, M. (2002). How effective are technology incubators: Evidence from Italy. Research policy, 31(7), 1103-1122.
- Cooke, P. (1992). Regional innovation systems: competitive regulation in the new Europe. Geoforum, 23(3), 365-382.
- Cooke, P. (2005). Regionally asymmetric knowledge capabilities and open innovation: Exploring 'Globalisation 2'-A new model of industry organization. Research policy, 34(8), 1128-1149.
- Cooke, P., Uranga, M. G., & Etxebarria, G. (1997). Regional innovation systems: Institutional and organizational dimensions. Research policy, 26(4-5), 475-491.
- Cooke, P., Uranga, M. G., & Etxebarria, G. (1998). Regional systems of innovation: an evolutionary perspective. Environment and planning A, 30(9), 1563-1584.
- Dahlander, L., & Gann, D. M. (2010). How open is innovation? Research policy, 39(6), 699-709.
- Del Giudice, M., Carayannis, E. G., & Maggioni, V. (2017). Global knowledge intensive enterprises and international technology transfer: emerging perspectives from a quadruple helix environment. The Journal of Technology Transfer, 42(2), 229-235.
- Edquist, C. (2010). Systems of innovation perspectives and challenges. African Journal of Science, Technology, Innovation and Development, 2(3), 14-45.
- Etzkowitz, H., & Leydesdorff, L. (2000). The dynamics of innovation: from National Systems and "Mode 2" to a Triple Helix of university-industry-government relations. Research policy, 29(2), 109-123.
- Etzkowitz, H., de Mello, J. M. C., & Almeida, M. (2005). Towards "meta-innovation" in Brazil: The evolution of the incubator and the emergence of a triple helix. Research policy, 34(4), 411-424.
- Fagerberg, J., & Srholec, M. (2008). National innovation systems, capabilities and economic development. Research policy, 37(9), 1417-1435.
- Freeman, C. (1995). The 'National System of Innovation'in historical perspective. Cambridge Journal of economics, 19(1), 5-24.
- Freeman, C. (2002). Continental, national and sub-national innovation systems-complementarity and economic growth. Research policy, 31(2), 191-211.
- Fuenfschilling, L., & Truffer, B. (2014). The structuration of socio-technical regimes-Conceptual foundations from institutional theory. Research Policy, 43(4), 772-791.
- Furman, J. L., Porter, M. E., & Stern, S. (2002). The determinants of national innovative capacity. Research policy, 31(6), 899-933.
- Furrer, O., Thomas, H., & Goussevskaia, A. (2008). The structure and evolution of the strategic management field: A content analysis of 26 years of strategic management research. International Journal of Management Reviews, 10(1), 1-23.
- Geels, F. W. (2004). From sectoral systems of innovation to socio-technical systems: Insights about dynamics and change from sociology and institutional theory. Research policy, 33(6-7), 897-920.
- Geels, F. W. (2004). Understanding system innovations: a critical literature review and a conceptual synthesis. System innovation and the transition to sustainability: Theory, evidence and policy, 19-47.
- Geels, F. W. (2005). The dynamics of transitions in socio-technical systems: a multi-level analysis of the transition pathway from horse-drawn carriages to automobiles (1860-1930). Technology analysis & strategic management, 17(4), 445-476.
- Geels, F. W. (2010). Ontologies, socio-technical transitions (to sustainability), and the multi-level perspective. Research policy, 39(4), 495-510.
- Geels, F. W. (2012). A socio-technical analysis of low-carbon transitions: introducing the multi-level perspective into transport studies. Journal of transport geography, 24, 471-482.
- González-Loureiro, M., Dabic, M., & Furrer, O. (2015). A content and comparative analysis of strategic management research in the Baltic area. Baltic Journal of Management, 10(2), 243-266.
- Gonzalez-Loureiro, M., Kiessling, T., & Dabic, M. (2015). Acculturation and overseas assignments: A review and research agenda. International Journal of Intercultural Relations, 49, 239-250.
- Gonzalez-Loureiro, M., Sousa, M. J., & Pinto, H. (2017). Culture and innovation in SMEs: the intellectual structure of research for further inquiry. European Planning Studies, 25(11), 1908-1931.
- Grundel, I., & Dahlström, M. (2016). A quadruple and quintuple helix approach to regional innovation systems in the transformation to a forestry-based bioeconomy. Journal of the Knowledge Economy, 7(4), 963-983.
- Hekkert, M. P., Suurs, R. A., Negro, S. O., Kuhlmann, S., & Smits, R. E. (2007). Functions of innovation

systems: A new approach for analysing technological change. Technological forecasting and social change, 74(4), 413-432.

- Hjalager, A. M. (2010). A review of innovation research in tourism. Tourism management, 31(1), 1-12.
- Howaldt, J., & Schwarz, M. (2011). Social innovation-social challenges and future research fields. In Enabling Innovation (pp. 203-223). Springer, Berlin, Heidelberg.
- Huizingh, E. K. (2011). Open innovation: State of the art and future perspectives. Technovation, 31(1), 2-9.
- Jacobsson, S., & Johnson, A. (2000). The diffusion of renewable energy technology: an analytical framework and key issues for research. Energy policy, 28(9), 625-640.
- Kaufmann, A., & Tödtling, F. (2001). Science-industry interaction in the process of innovation: the importance of boundary-crossing between systems. Research policy, 30(5), 791-804.
- Köhler, J., Geels, F. W., Kern, F., Markard, J., Onsongo, E., Wieczorek, A., . . . & Fünfschilling, L. (2019). An agenda for sustainability transitions research: State of the art and future directions. Environmental Innovation and Societal Transitions, 31, 1-32.
- Liu, X., & White, S. (2001). Comparing innovation systems: a framework and application to China's transitional context. Research policy, 30(7), 1091-1114.
- López-Duarte, C., González-Loureiro, M., Vidal-Suárez, M. M., & González-Díaz, B. (2016). International strategic alliances and national culture: Mapping the field and developing a research agenda. Journal of World Business, 51(4), 511-524.
- López-Rubio, P., Roig-Tierno, N., & Mas-Verdú, F. (2021). Assessing the origins, evolution and prospects of national innovation systems. Journal of the Knowledge Economy, 1-24.
- López-Rubio, P., Roig-Tierno, N., & Mas-Tur, A. (2020). Regional innovation system research trends: toward knowledge management and entrepreneurial ecosystems. International Journal of Quality Innovation, 6(1), 1-16.
- López-Rubio, P., Roig-Tierno, N., & Mas-Tur, A. (2021). A research journey from national systems of innovation to national systems of entrepreneurship: Introducing the Sextuple Helix. International Journal of Innovation and Technology Management (IJITM), 18(08), 1-23.
- Lundvall, B. Å., Johnson, B., Andersen, E. S., & Dalum, B. (2002). National systems of production, innovation and competence building. Research policy, 31(2), 213-231.
- Malerba, F. (2002). Sectoral systems of innovation and production. Research policy, 31(2), 247-264.
- Markard, J., Raven, R., & Truffer, B. (2012). Sustainability transitions: An emerging field of research and its prospects. Research policy, 41(6), 955-967.
- Meyer-Krahmer, F., & Schmoch, U. (1998). Science-based technologies: university-industry interactions in four fields. Research policy, 27(8), 835-851.
- Miller, K., McAdam, R., Moffett, S., Alexander, A., & Puthusserry, P. (2016). Knowledge transfer in university quadruple helix ecosystems: an absorptive capacity perspective. R&D Management, 46(2), 383-399.
- Miller, K., McAdam, R., & McAdam, M. (2018). A systematic literature review of university technology transfer from a quadruple helix perspective: toward a research agenda. R&D Management, 48(1), 7-24.
- Morgan, K. (2004). The exaggerated death of geography: learning, proximity and territorial innovation systems. Journal of economic geography, 4(1), 3-21.
- Moulaert, F., & Sekia, F. (2003). Territorial innovation models: a critical survey. Regional studies, 37(3), 289- 302.
- Muller, E., & Zenker, A. (2001). Business services as actors of knowledge transformation: the role of KIBS in regional and national innovation systems. Research policy, 30(9), 1501-1516.
- Negro, S. O., Alkemade, F., & Hekkert, M. P. (2012). Why does renewable energy diffuse so slowly? A review of innovation system problems. Renewable and Sustainable Energy Reviews, 16(6), 3836-3846.
- Phillips, W., Lee, H., Ghobadian, A., O'Regan, N., & James, P. (2015). Social innovation and social entrepreneurship: A systematic review. Group & Organization Management, 40(3), 428-461.
- Pittaway, L., Robertson, M., Munir, K., Denyer, D., & Neely, A. (2004). Networking and innovation: a systematic review of the evidence. International journal of management reviews, 5(3-4), 137-168.
- Ribeiro, L., Ruiz, R., Bernardes, A., & Albuquerque, E. (2010). Matrices of science and technology interactions and patterns of structured growth: implications for development. Scientometrics, 83(1),

55-75.

- Rogge, K. S., & Reichardt, K. (2016). Policy mixes for sustainability transitions: An extended concept and framework for analysis. Research Policy, 45(8), 1620-1635.
- Rodríguez-Pose, A., & Crescenzi, R. (2008). Research and development, spillovers, innovation systems, and the genesis of regional growth in Europe. Regional studies, 42(1), 51-67.
- Schartinger, D., Rammer, C., & Fröhlich, J. (2006). Knowledge interactions between universities and industry in Austria: sectoral patterns and determinants. In Innovation, networks, and knowledge spillovers (pp. 135-166). Springer, Berlin, Heidelberg.
- Schmutzler, J., Suarez, M., Tsvetkova, A., & Faggian, A. (2017). Introduction. A context-specific two-way approach to the study of innovation systems in developing and transition countries. In Innovation in Developing and Transition Countries. Edward Elgar Publishing.
- Schumpeter, J. A. (1976). Capitalism, socialism and democracy (1942). J. Econ. Literature, 20, 1463.
- Smith, A., Voß, J. P., & Grin, J. (2010). Innovation studies and sustainability transitions: The allure of the multi- level perspective and its challenges. Research policy, 39(4), 435-448.
- Solleiro, J. L., & Castañón, R. (2005). Competitiveness and innovation systems: the challenges for Mexico's insertion in the global context. Technovation, 25(9), 1059-1070.
- Spigel, B. (2017). The relational organization of entrepreneurial ecosystems. Entrepreneurship Theory and Practice, 41(1), 49-72.
- Teece, D. J. (2010). Business models, business strategy and innovation. Long range planning, 43(2-3), 172- 194.
- Tödtling, F., & Trippl, M. (2005). One size fits all: Towards a differentiated regional innovation policy approach. Research policy, 34(8), 1203-1219.
- Toivanen, H., & Ponomariov, B. (2011). African regional innovation systems: bibliometric analysis of research collaboration patterns 2005-2009. Scientometrics, 88(2), 471-493.
- Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review. British journal of management, 14(3), 207-222.
- Van der Have, R. P., & Rubalcaba, L. (2016). Social innovation research: An emerging area of innovation studies?. Research Policy, 45(9), 1923-1935.
- Udoh, E., & Rhoades, J. (2006, April). Mining Documents in a Small Enterprise Using WordStat. In Third International Conference on Information Technology: New Generations (ITNG'06) (pp. 490-494). IEEE
- Uriona-Maldonado, M., dos Santos, R. N., & Varvakis, G. (2012). State of the art on the Systems of Innovation research: a bibliometrics study up to 2009. Scientometrics, 91(3), 977-996. Uyarra, E. (2010). What is evolutionary about 'regional systems of innovation'? Implications for regional policy. Journal of evolutionary economics, 20(1), 115-137.
- Uyarra, E., & Flanagan, K. (2010). From regional systems of innovation to regions as innovation policy spaces. Environment and Planning C: Government and Policy, 28(4), 681-695.
- Watt, A. (2015). QDA Miner 4.0. Qualitative Research Journal, 15(2), 250-251.
- Wolfe, D. A., & Gertler, M. S. (2004). Clusters from the inside and out: local dynamics and global linkages. Urban studies, 41(5-6), 1071-1093.
- Woolthuis, R. K., Lankhuizen, M., & Gilsing, V. (2005). A system failure framework for innovation policy design. Technovation, 25(6), 609-619.

Appendix

CATEGORY Mental construct	WORDS Family of related words
CLUSTER	Cluster, industrial cluster.
COLLABORATION	Innovation network, network.

CATEGORY Mental construct	WORDS Family of related words
DEVELOPMENT	Advanced economies, develop, developing countries, economic develop, economic growth, emergence, emerging, factors, economic geography, globalization, integrated, patterns, regional economics, regional level, sectoral systems, socio economic, underdevelopment, world, world development.
DYNAMIC	Dynamic, evolution, transformation, transition.
ENTERPRISE	Model, small and medium, business services, enterprise.
ENTREPRENEURSHIP	Creation, creative, entrepreneurial, entrepreneurship ecosystem, entrepreneurship, innovative, start up, venture capital, business services, enterprise, small and medium.
CAP_INN (Innovation Capacity)	innovation process, innovation, innovation activities, innovation and production, innovation capacity, innovation process, innovative capacity, learning process, market formation, potential, process, role, sources, structural, structure, innovative activities.
G__ AUSTRALIA AND OCEANIA	Australia, Micronesia, New Zealand, Papua_ new Guinea, Samoa.
G__AFRICA	Africa, Africa, Algeria, Angola, Benin, Botswana, Burkina Faso, Burundi, Cameroon, Cape Verde, Central African Republic, Chad, Comoros, democratic republic of the Congo, Djibouti, Egypt, equatorial Guinea, Eritrea, Eswatini, Ethiopia, Gabon, Ghana, Guinea, Guinea-Bissau, ivory coast, Kenya, Lesotho, Liberia, Libya, Madagascar, Malawi, Mali, Mauritania, Mauritius, morocco, Mozambique, Namibia, Niger, Nigeria, Rwanda, Senegal, Sierra Leone, Somalia, south Africa, south Africa, south Africa, south Sudan, Tanzania, the Gambia, Tunisia, Uganda, western Sahara, Zambia, Zimbabwe.
G__ASIA	Afghanistan, Armenia, Asian, Azerbaijan, Bahrein, Bangladesh, Brunei, Bhutan, Cambodia, china, Chinese, Cyprus, Georgia, India, Indonesia, Iran, Iraq ,Israel, Japan, Japanese, Jordan, Kazakhstan, Korea, Kuwait, Kyrgyzstan, Laos, Lebanon, Malaysia, Maldives, Mongolia, Myanmar, Nepal, north Korea, Oman, Pakistan, Palestine, Philippines, Qatar, Russia, Russian, Saudi Arabia, Singapore, south Korea, Sri Lanka, Syria, Taiwan, Taiwan, Taiwanese, Tajikistan, Thailand, Timor-Leste, turkey, Turkmenistan, United Arab Emirates, Uzbekistan, Vietnam, Yemen.

CATEGORY Mental construct	WORDS Family of related words
G__EUROPA	Albania, Andorra, Armenia, Austria, Azerbaijan, Belarus, Belgium, Bulgaria, Croatia, Cyprus, Czechia, Denmark, Denmark, Dutch, Estonia, EU, Europe, European, European, European countries, European union, Finland, France, France, Georgia, German, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Liechtenstein, Lithuania, Malta, Moldavan, Monaco, Montenegro, Netherlands, north Macedonia, Norway, Poland, Portugal, Romania, Russia, san Marino, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, turkey, UK, Ukraine, United kingdom.
G__LATAM	Argentina, Bolivia, Brazil, chile, Colombia, Costa Rica, Cuba, Dominican Republic, Ecuador, el Salvador, Guatemala, Honduras, México, Nicaragua, Panamá, Paraguay, Perú, Uruguay, Venezuela.
G—NORTHAMERICA	Canada, United States.
INDUSTRY	Agricultural, agriculture system, Agriculture, conservation agriculture, Biotechnology, Business service, companies, creative industries, energy, energy policy, energy system, firm, generation, industrial, industries, industry, manufacturing, manufacturing firm, nanotechnology, oil, patent, product, production research and development, sector, sectoral, service, small and medium, software, software industries, transformation transition.
KNOWLEDGE	Information, innovation studies, innovation theory, intensive, knowledge creation
Knowledge _network, knowledge transfer, knowledge, knowledge base, knowledge based economy, knowledge bases, knowledge creation, knowledge diffusion, knowledge economy, knowledge exchange, knowledge exchange, knowledge flows, knowledge intensive, knowledge interactions, knowledge management, knowledge networks, knowledge production, knowledge spillovers, knowledge transfer, regional knowledge, tacit knowledge, tacit knowledge, transfer	
MANAGEMENT	Adaptative management, administration, advantage, business model, challenge, competitive advantage, complex, cooperation, cost, decision, external, forecasting, future, goals, innovation management, management of technology, market, modelling, organization, organizational, organization, resource, specialization, stakeholders, strategic, strategic management, strategies, technology management
NATIONAL_INNOVATION	Nation, national, National innovation, National innovation system, nation level, national system of innovation

CATEGORY Mental construct	WORDS Family of related words
PERFORMANCE	Competition, competitive, competitive advantage, competitiveness, demand, efficiency, efficient, gap, impact, indicator, innovation performance, market, performance, productivity, results, role, success, successful, update
POLICY	Energy policy, innovation policy, Policy maker, Policy implication, policy instrument, policy making, politic, public, public policy, research policy, technology policy
REGIONAL_INNOVATION	Geographical proximity, local, local innovation, proximity, region, region development, regional innovation system, regional system, regional, regional advantage, regional development, regional innovation, regional innovation system, regional level, regional science, regional studies, regional system of innovation, territorial, region
SCIENCE	Research, research and development, science, science and technology, science based, scientific, researchers,
SUSTAINABILITY	Climate, climate change, eco innovation, ecological, ecosystem, environment, environmental, environmental innovation, environments, green, nature, renewable, renewable energy, renewable energy technologies, selection environments, sustain develop, sustain transit, sustainability transitions, sustainable, sustainable development,
SYSTEM	Function, innovation system, innovation system approach, system, system function, system of innovation, systemic, systems of innovation, technological innovation system
T SOCIAL_INNOVATION	Social, social innovation
T__OPEN_INNOVATION	Collaboration, Collaborative, collaborative innovation, open innovation, open
T__SOCIO_TECNOLOGICA	socio technical, socio technical system, socio technical transition
T__TRIPL HELICE	Actor network, actor strategies, actor, actor and institution, actors involved, firm, governance, government, helix, institution, institutional, institutional change, interaction, political, public sector, society, triple helix, triple helix model, universities, university industry, university industry government
TRANSFER	

Table A: Dictionary of Categories. Source: Own elaboration.

A Comparative Study of the Value of Asian NFT Art on OpenSea, with a Focus on Art Incorporating Asian Characters

Anela Ilijas, Yeungnam University, SOUTH KOREA (anelailijas@yu.ac.kr)

Abstract

The emergence of Non-Fungible Tokens (NFTs) has revolutionized the landscape of digital art, providing artists with new avenues for creation, distribution, and monetization. This research investigates the landscape of NFT art collections featuring Asian characters, specifically Japanese kanji and Korean hangeul, on the OpenSea platform. Theoretical frameworks drawn from digital art valuation, cultural representation in art, and blockchain technology guide the analysis. Findings reveal diverse manifestations of cultural identity and artistic innovation within the analyzed collections, reflecting the intersection of tradition and modernity in Asian art. Economically, prominent artists command substantial sales volumes, while others experience more modest transactions, underscoring the growing interest in NFT art within the global market. Culturally, NFT art collections serve as platforms for preserving and promoting traditional scripts and symbols, contributing to the global dissemination of Asian cultural heritage. Technologically, the adoption of blockchain technology empowers artists to assert ownership over their creations and establish new paradigms of value and authenticity. In conclusion, this research highlights the transformative potential of NFTs in reshaping the landscape of art and culture, emphasizing the need for continued exploration and collaboration to amplify cultural voices and expand artistic horizons in the digital age.

Keywords: NFT art, Asian digital art, kanji art, hangeul art, asian characters

1. Introduction

In 2021, non-fungible tokens (usually abbreviated as NFT) gained significant attention from the mainstream media, artists, celebrities, and investors. So-called NFT boom refers to a period of widespread interest and adoption of NFT that took place primarily during the first half of 2021. During this period, the NFT market experienced a surge in transaction volumes, sales, and overall market capitalization. While the intense hype and media attention around NFT may have subsided since the peak of the boom in 2021, the market for NFT remains active and evolving. Many art critics say that the NFT era is the art's digital revolution: we live in the time where digital art has gained an economical value, just like traditional art. Record-breaking sales and rise of cryptocurrencies' value attracted the attention of many artists. And as it is often the case in modern art, there are always some artists who include their national identity and traditions to their art. Analyzing the trends of the NFT art industry in Asia, we have noticed that Asian digital artists, particularly those from Japan and Korea, have been incorporating their language's letters in their NFT artworks, leveraging the recent popularity of East Asian pop culture. This trend reflects a cultural exchange and the artists' desire to express their identity and heritage through their art. In this study, we focus on the analysis of NFT art collections featuring Asian characters, specifically Japanese kanji and Korean hangeul, on the OpenSea platform. By examining the economic value, cultural representation, and artistic expression within these collections, we aim to shed

light on the intersection of technology, art, and culture in the digital realm. Through the analysis of the unique characteristics and cultural nuances of Asian NFT art collections, we hope to deepen the understanding of the evolving landscape of digital culture and its impact on global artistic expression.

2. Theoretical framework and literature review

The analysis of NFT art collections featuring Asian characters, particularly Japanese kanji and Korean hangeul, will be guided by a theoretical framework that integrates concepts from digital art valuation, cultural representation in art, and blockchain technology.

Digital Art Valuation:

Digital art valuation theory provides insights into the economic significance and market dynamics of NFT art collections. It encompasses discussions on factors influencing the value of digital artworks. For instance, the article on Binance (2021) suggests considering the following indicators when evaluating NFT art: utility, rarity, community size, potential of the NFTs, provenance, and personal taste. According to Bsteh (2021), the factors that determine the value of digital artworks on the secondary market include the price, the history of artists and the innovations they implemented. This may include various aspects such as the use of new technologies, algorithms, games, communication methods, and the overall brand of the artist. They also add that, in general, the aesthetic value and personal taste are significant factors when considering investments in NFT art. The increased transparency facilitated by blockchain technology and the inclusion of smart contracts are expected to enhance the perceived security associated with investing in NFT art (Bsteh, 2021). Dylan-Ennis (2021) claims that blockchain art is often valued for reasons that go beyond the aesthetics. For example, certain NFTs like Cryptopunks are highly sought after due to their age, just like valuable antique items in the blockchain world. Despite appearing as simple pixel art, the most expensive Cryptopunk was sold for a significant amount. Dylan-Ennis (2021) emphasizes that what makes Cryptopunks desirable is not just the artwork itself, but rather their metadata, such as their longevity on the blockchain, and to truly understand the value, one must look beyond the art and consider the medium. Russell (2022) emphasizes that while previous research on NFT has shown that the connection to prominent collectors and visibility on popular crypto platforms are the indicators of the expected price of an NFT, it is crucial to consider a variety of difficult-to-quantify factors. While measurable forms of influence hold significance, the ability of an artwork, whether associated with an NFT or not, to gain discursive enrichment is essential for its long-term value and recognition as historically and culturally significant. This enrichment should be independent of the opinions or motivations of individuals who stand to gain financially from its reevaluation (Russell, 2022). As noted by Kılıçaslan and Ekizler (2022), the concept of "digital scarcity" enabled by blockchain technology contributes to the perceived value of NFT art, shaping consumer preferences and investment behavior. By examining sales data and total volume metrics from platforms like OpenSea, we aim to elucidate patterns and trends in the valuation of Asian NFT art.

Cultural Representation in Art:

Theoretical perspectives on cultural representation in art elucidate how artworks reflect and shape cultural identities, narratives, and values. According to Hall (2015), cultural identity is a dynamic construct shaped by historical, social, and political contexts, manifested through symbols, languages, and artistic expressions. In the context of NFT art collections featuring Asian characters, the incorporation of kanji and hangeul symbols serves as a means of cultural representation and preservation. As highlighted by Economou (2015), digital art platforms like OpenSea facilitate the global dissemination of cultural heritage, enabling artists to share their unique cultural perspectives with diverse audiences worldwide. Saul (2020) explores the notion of cultural symbols in art and argues that they play a crucial role in communicating cultural meanings and identities. He suggests that cultural symbols serve as vehicles for expressing

collective experiences and traditions. In our analysis, we consider how Japanese and Korean artists utilize their respective scripts, kanji and hangeul, to express cultural heritage and identity. Through qualitative examination of artwork designs and artists' statements, we explore the ways in which Asian characters serve as symbols of cultural authenticity in the NFT art domain.

Blockchain Technology:

Theoretical framework on blockchain technology elucidates its transformative impact on various industries, including the art market. According to Tapscott and Tapscott (2016), blockchain enables decentralized and transparent transactions, ensuring the integrity and authenticity of digital assets. In the context of NFT art collections, blockchain technology facilitates the creation, distribution, and ownership of digital artworks, revolutionizing traditional notions of art ownership and provenance. As noted by Catalini and Gans (2020), blockchain-based platforms like Ethereum and Klaytn provide a secure and immutable infrastructure for trading NFTs, empowering artists to monetize their creations and engage with global audiences without intermediaries. Additionally, Kochetkova (2020) highlights the potential of blockchain technology to democratize access to the art market by eliminating intermediaries and reducing transaction costs. She argues that blockchain-based platforms enable direct peer-to-peer transactions, fostering greater inclusivity and transparency in the art ecosystem. Blockchain theory informs our analysis of NFT platforms like OpenSea and the implications of utilizing decentralized ledgers for art ownership and provenance. We will examine how blockchain facilitates transparent transactions and empowers artists to monetize their creations in novel ways.

3. Methodology

Our study relies on a mixed-method approach combining quantitative and qualitative data collection methods. Quantitative data, including total volume metrics, sales data, and pricing information, are gathered from OpenSea for each analyzed NFT art collection. Additionally, qualitative data, such as artists' statements, collection descriptions, and visual analyses of artworks, are collected to provide context and insights into artistic intentions and cultural representations. The collected data will undergo a systematic analysis to fulfill the objectives of our research. Quantitative analysis will involve examination of total volumes of the selected NFT art collections. Qualitative analysis will entail thematic coding of artists' statements and visual elements of artworks to identify recurring motifs, cultural themes, and aesthetic choices. Comparative analysis between collections will elucidate distinctions in artistic approaches, cultural interpretations, and economic outcomes. The findings from our analysis will be interpreted within the theoretical framework outlined above. This combined theoretical framework and methodology will facilitate a comprehensive examination of Asian NFT art collections on OpenSea, providing insights into their cultural significance, economic value, and technological implications.

4. Explaining NFT

Wang, et al. (2021) define non-fungible token (NFT) as a type of cryptocurrency, which was firstly proposed in Ethereum Improvement Proposals (EIP)-721 and further developed in EIP-1155. They mention that the reason why NFT differs from other cryptocurrencies is that NFT is a unique, non-fungible digital asset, unlike classical cryptocurrencies where all coins are equivalent and indistinguishable. For our research, it is important to explain NFT in the context of art. NFT art is digital art that is tokenized, meaning it is turned into an NFT, which is then recorded on a blockchain. What sets NFT art apart is its indivisibility, authenticity, and scarcity. Each NFT art piece is one-of-a-kind, verifiable as the original work, and its ownership and transaction history are permanently recorded on the blockchain, making it nearly impossible to duplicate or counterfeit. This fusion of technology and creativity has transformed the art world by allowing artists to sell their digital creations directly to collectors and providing a new way to estab-

lish provenance and value in the digital art market. Blockchain technology is foundational to the concept of Non-Fungible Tokens (NFTs). Firstly, it provides the decentralized infrastructure necessary for NFT creation, storage, and transactions. Each NFT is typically minted or created on a blockchain, which records the ownership and transaction history of that token in an immutable and transparent manner. This ensures the scarcity and authenticity of the digital asset. Secondly, blockchain technology enables the secure and transparent transfer of NFTs. The ownership of NFTs is recorded on the blockchain's ledger, and ownership can be transferred between users with cryptographic signatures, ensuring provable ownership and provenance. Caxton, et al. (2022) explain the blockchain as a secure database without any need for a centralized authority to update or maintain the data on it which makes blockchain's use-cases endless, and blockchain networks customisable. A blockchain is a decentralized and distributed ledger technology that records transactions across a network of computers. The key feature is decentralization, meaning no single entity controls the entire network, enhancing transparency and security. The most well-known NFT marketplaces like OpenSea predominantly operate on second-generation blockchains, including Ethereum, Polygon, and Klaytn (Prasad, et al., 2023). Ethereum, Klaytn, and Polygon are all distinct blockchain networks, each with unique characteristics. Ethereum, the first and most well-known of the three, is a decentralized platform primarily used for smart contracts and decentralized applications (DApps). It employs a Proof-of-Stake (PoS) consensus mechanism, transitioning from Proof-of-Work (PoW), to improve scalability and energy efficiency. Klaytn, on the other hand, is a blockchain developed by the South Korean tech giant, Kakao. It focuses on providing a user-friendly environment for businesses and developers to build DApps, with a particular emphasis on mass adoption. Klaytn employs a hybrid consensus mechanism, combining PoW and PoS elements for enhanced security and efficiency. Polygon, formerly Matic Network, is a layer 2 scaling solution for Ethereum, designed to tackle Ethereum's scalability issues. It allows faster and cheaper transactions on the Ethereum network by utilizing sidechains and PoS. Polygon aims to enhance the user experience for DApps and DeFi projects on Ethereum, providing a bridge between the Ethereum ecosystem and other blockchains.

5. Analysis of NFT art collections

暗号漢字 -CryptoKanji-

暗号漢字 -CryptoKanji- is an Ethereum-based NFT art collection fully created by famous Japanese NFT artist mera takeru. It is listed on OpenSea (https://opensea.io/collection/cryptokanji). Mera takeru is also known as a pioneer of crypto art in Japan, as he was one of the first Japanese artists who started creating and selling NFT art. He is highly recognized in the Japanese digital art community and has showcased his work in numerous exhibitions, both domestically and internationally. He is also known as an advocate for NFT and digital art, frequently participating in online seminars and talk sessions to promote new forms of art. 暗号漢字 -CryptoKanji- consisting of 139 art works (current state: June 11, 2023) represents different Japanese characters kanji. It's important to say that in the interview for Japanese web-portal Atarashii keizai (2021), mera takeru mentioned this art collection. During the interview, he said the following: „Japan is still a closed island country, so I think there are many authentic cultural elements that are still unknown to the world. As I've already said before, crypto art easily crosses the ocean, so I am conscious of incorporating these elements of Japanese culture into my art works. 暗号漢字 -CryptoKanji-, where each kanji is joined with an English word, is one of these works" (Atarashii keizai, 2021). While the statement that kanji, characters originated from ancient China, are the part of Japanese authentic culture might be controversial to some people, this is the artist's opinion and we can see from that statement that his 暗号漢字 -CryptoKanji- collection was made to promote kanji overseas (he is most likely referring to the West). In terms of design, all artworks in this collection feature a minimalist design with a clean white background and a bold black kanji letter placed prominently (refer to Figure 1).

Figure 1: 神. 暗号漢字 -CryptoKanji- collection by mera takeru (2020). (Source: OpenSea.io)

The simplicity of the composition allows the letter to take center stage, capturing attention. The contrast between the black letter and the white background adds to the overall clarity and readability of the artwork. At the moment of writing this paper, the total volume of 暗号漢字 -CryptoKanji-collection on the OpenSea is 3 ETH, which is approximately $5212,35. The significant total volume of this NFT collection can be attributed to the artist mera takeru's prominence and popularity within the NFT art industry. Checking the sales of this collection on the OpenSea, we noticed that most of the artworks of this collection were sold. The highest sale was the 0,1 WETH (approximately $173,89) purchase of the artwork titled 神 (kanji character meaning 'god'). The artwork titled 幻 (kanji character meaning 'phantom') was the second-highest sale, with a purchase of 0,099 WETH (approximately $172,15), while the third highest sale was 0,07 WETH (approximately $121,85) purchase of the artwork titled 空 (kanji character meaning 'sky').

Golden Kanji

Another NFT art collection incorporating Japanese kanji into its artworks is Golden Kanji. It is a Polygon-based NFT collection listed on OpenSea (https://opensea.io/collection/goldenkanji). This collection consists of 2136 artworks, each of them representing one kanji from the list of jōyō kanji (常用漢字) - list of kanji characters designated by the Japanese government as "commonly used" and officially recognized as literacy baseline in written Japanese. The creator of the Golden Kanji collection, Japanese artist Nekoraisu, describes this collection as follows: "Golden Kanji is Polygon based NFT collection of animated Golden Japanese Kanji. The first season of this collection consists of 2136 Jouyou Kanji officially established by the Japanese Ministry of Education. This project is dedicated to Japanese language learners around the world and the expression of gratitude to our ancestors that bring these awesome characters and are still used in today's world" (Golden Kanji, 2021). In terms of design, all artworks in the Golden Kanji collection are the animated images where the shimmering outlines of the kanji letters appear first, and then the gold-colored kanji letter appears in full (refer to Figure 2).

Figure 2: #0033 Golden Jouyou Kanji 侍. Golden Kanji collection by Nekoraisu (2021). (Source: OpenSea.io)

Each artwork shows a light up kanji letter with a black background. Each animated kanji letter is glowing brightly, illuminating the area around it and casting an abstract blur of light onto the surface behind it. At the moment of writing this paper, the total volume of the Golden Kanji collection on the OpenSea is 0,0192 ETH, which is approximately $33,36. Checking the sales of this collection on OpenSea, we noticed that only five artworks from this collection were sold. However, one of them, #0026 Golden Jouyou Kanji 何, was sold to the two users, which means that there were two sales for this artwork. The artworks that were sold are #0033 Golden Jouyou Kanji 侍 (kanji character meaning 'samurai'), #0011 Golden Jouyou Kanji 乙 (kanji character meaning 'duplicate' or 'strange'), #0004 Golden Jouyou Kanji 並 (kanji character meaning 'row' or 'line up'), #0573 Golden Jouyou Kanji 沖 (kanji character meaning 'open sea' or 'offing'), and #0026 Golden Jouyou Kanji 何 (kanji character meaning 'what'). All five artworks were sold at the same price of 0,0032 ETH (approximately $5,56).

Replica Kanji Flowers

The last NFT collection with Japanese kanji characters we want to mention in this paper is a Polygon-based NFT collection Replica Kanji Flowers, which is listed on OpenSea (https://opensea.io/collection/replica-kanji-flowers). This collection comprises a total of 2220 artworks. The author behind this idea is Kenta Suhara, Japanese Web3 developer and the CEO of Japanese IT company Pontech inc. He is well-known within the NFT community in Japan as he launches various NFT projects and participates in various blockchain technology events quite often. Firstly, Kenta Suhara launched the project Kanji Flowers at the website https://kanjiflowers.xyz, where he allowed the users to mint the artworks on Polygon blockchain platform for free. This project gained huge attention in the Japanese digital art community as all artworks were minted and taken within two hours from the release time. Then those newly created artworks were sold by the users who minted them at OpenSea. This collection has an interesting design as all artworks from this collection present one kanji character, which is circularly repeated (or replicated, as the name of the

collection suggests) in order to create a shape of a flower (refer to Figure 3).

Figure 3: 少. Replica Kanji Flowers collection by Kenta Suhara (2021). (Source: OpenSea.io)

The shape changes along with the size in correspondence with the user's mouse movement. While this idea is interesting, it is not new: the artist openly admits that he borrowed the concept and code from the project Cyberflowers done by peilingjiang. The background of each artwork is grayish white, while the main kanji character is colorful and all his duplicated versions (replicas) are black. At the moment of writing this paper, the total volume of this collection on the OpenSea is 0,8886 ETH, which is approximately $1561,97. Upon reviewing the sales data of the Replica Kanji Flowers collection on OpenSea, it is evident that the most significant transaction involved the acquisition of the artwork named 少 (kanji character meaning 'little') for a price of 0,2 ETH (approximately $351,56). The second highest sale price was 0,05 ETH (approximately $87,89) and four artworks were sold under this price. Those artwork are 乃 (kanji character meaning 'from'), 福 (kanji character meaning 'luck'), 詩 (kanji character meaning 'poetry'), and 化 (kanji character meaning 'change'). The third highest sale price was 0,032 ETH (approximately $56,25) and the artwork titled 水 (kanji character meaning 'water') was sold under this price. It is also interesting to add that one month after the release of the Replica Kanji Flowers, Kenta Suhara also launched the project Replica Jukugo Flowers, which was intended only for the owners of Replica Kanji Flowers artworks. In this project, Replica Kanji Flowers art owners could mint their jukugo NFT if they had at least two Kanji Flowers artworks. Jukugo (熟語) are Japanese kanji compound words: several kanji characters together can compose an idiom. The design of Replica Jukugo Flowers was identical to Replica Kanji Flowers, with the only difference that instead of one kanji and its numerous replicas, each artwork of this collection had two original kanji characters that were replicated in the flower shape.

Klminjeongeum NFT

Switching from Japanese kanji characters to Korean hangeul letters, there is an NFT collection that has to be mentioned. It is Klminjeongeum NFT, a Klaytn-based collection listed on OpenSea

(https://opensea.io/collection/klminjeongeum). Klminjeongeum NFT is a collection made by the NFTKLE team, a team of NFT artworks and technology enthusiasts based in Korea. Klminjeongeum NFT was their first project which they launched in September 2021. The creators of this collection define it as "the first & greatest hangul NFT that contains all Korean letters" (NFTKLE.co, 2021). The name of this collection comes from the Hunminjeongeum (훈민정음), the document describing hangeul - an entirely new script for Korean language invented by King Sejong. However, the first syllable from Hunminjeongeum was changed to Kl as Kl of Klaytn, the blockchain platform utilized for creating this particular NFT collection. Nevertheless, the creators claim that Klminjeongeum is not related to the content of Hunminjeongeum at all, but mention that Klminjeongeum embodies the fundamental concepts of yin-yang and the five elements found in oriental philosophy (NFTKLE.co, 2021). There are 11172 artworks in this collection. NFTKLE team explains that "399 characters without consonants are level 2, accounting for 3.5% of the total number. The remaining letters with consonants are level 1, and 10,773, which is 96.5% of the total number" (NFTKLE.co, 2021). In terms of design, all artworks in this collection feature minimalistic visuals (refer to Figure 4).

Figure 4: KMJ Hangul #1765. Klminjeongeum NFT collection by NFTKLE (2021). (Source: OpenSea.io)

The background of all artworks has a brownish color, while the foreground consists of a large black hangeul letter. The hangeul letter appears to be written in bold font, and it stands out against the plain brownish background. The letters are written in the traditional font. The design of these artworks showcases a composition that combines elements of traditional Korean aesthetics with contemporary minimalistic presentation. Klminjeongeum NFT is the only collection mentioned in this paper that is based on Klaytn. As we mentioned before, Klaytn is a blockchain developed by the South Korean Kakao Corporation. Unlike all other collections mentioned in this paper which use global blockains, Klminjeongeum NFT's creators use the blockchain developed in their own country, South Korea. Therefore, this collection is sold using the cryptocurrency WKLAY. WKLAY is based on Klaytn blockchain platform. It's not a surprise that an NFT project that claims to

promote Korean culture uses blockchain developed in Korea. At the moment of writing this paper, the total volume of the Klminjeongeum NFT collection on the OpenSea is 92133,6545 KLAY, which is approximately $13353,06. The highest sale price was 500 WKLAY (approximately $72,66) and thirty-two artworks in total were sold at this price. The second highest sale price was 400 WKLAY (approximately $58,12), and the third highest price was 299 WKLAY (approximately $43,45).

Hangul NFT

Hangul NFT is a Polygon-based NFT art collection created by GimbapTeam, a team of artists and engineers in Seoul creating Korean-inspired NFT collections. This collection was made specifically to commemorate Hangul Day (한글날), Korean national holiday which is celebrated annually on October 9th. It is listed on OpenSea (https://opensea.io/collection/hangul-nft). At the moment of writing this paper the collection contains 1009 artworks. In terms of design, all artworks in this collection showcase a simple, but vibrant and dynamic composition (refer to Figure 5).

Figure 5: Hangul #945. Hangul NFT collection by GimbapTeam (2021). (Source: OpenSea.io)

With colorful walls and floors as the background, each artwork features a central panel displaying a large, colorful hangeul letter. The panels cast shadows onto the floor, adding depth and dimension to the overall composition. Additionally, some of the panels imitate the wooden texture, further enhancing the visual appeal of the artworks. At the moment of writing this paper, the total volume of the Hangul NFT collection on the OpenSea is 0,0394 ETH, which is approximately $68,46. The highest sale was 0,002 ETH (approximately $3,48) purchase of the artwork titled Hangul #945 (artwork representing hangeul character ㅡ). All other sale prices were quite low, as the second highest price was only 0,0008 ETH (approximately $1,39) for the work Hangul #724 (artwork representing hangeul character ㅆ) and the third highest sale price was 0,0007 ETH (approximately $1,22) for the work Hangul #694 (artwork representing hangeul character ㅓ).

Hangul Art & Charism

Hangul Art & Charism is an NFT art collection fully created by Korean NFT artist HangulArt. It is an Ethereum-based collection listed on OpenSea (https://opensea.io/collection/hangul-art).

HangulArt is a new NFT artist who desires to share the beauty of hangeul letters in the form of art. In the description of his NFT collection, he says that Korean letters hangeul are art themselves, and this collection is made to explore the world of hangeul with the eyes of an artist. This collection is relatively small compared to other collections, as at the moment of writing this paper the collection contains only 70 artworks. Unlike all other collections mentioned previously, this one doesn't have a unified design and all works are unique. It can be also said that Hangul Art & Charism collection's works have much more complicated designs compared to art works in previously mentioned collections. Each artwork incorporates hangeul letters in the work in a unique and authentic way. For example, the work Flower drawn by 1 Million Flower(꽃, Korean Hangul Character) is an animated image portraying the white flower that is completely made by numerous tiny hangeul characters 꽃, which means 'flower' in Korean (refer to Figure 6).

Figure 6: Flower drawn by 1 Million Flower(꽃, Korean Hangul Character). Hangul Art & Charism collection by HangulArt (2021). (Source: OpenSea.io)

Next artwork from this collection worth mentioning is an artwork titled Yun DongJu's face written in his Seosi, Foreword (refer to Figure 7).

Figure 7: Yun DongJu's face written in his Seosi, Foreword. Hangul Art & Charism collection by HangulArt (2021). (Source: OpenSea.io)

In this work, artist created a portrait of famous Korean 20th century poet Yun Dong-ju, which is completely made from the lines of his most famous poem 서시 (Foreword). Another noteworthy artwork that deserves recognition is the piece titled 1M Char Pixels - Mona Lisa I Image : Mona Lisa in 1 Million characters of 11172 Hangul ㅁㄴㄹㅈ (refer to Figure 8).

Figure 8: 1M Char Pixels - Mona Lisa I Image : Mona Lisa in 1 Million characters of 11172 Hangul ㅁㄴㄹㅈ. Hangul Art & Charism collection by HangulArt (2021). (Source: OpenSea.io)

In this work HangulArt made a portrait of Mona Lisa using numerous hangeul characters ㅁ, ㄴ, ㄹ, ㅈ, which are the consonants taken from the Korean spelling of Mona Lisa - 모나리자. In conclusion, each artwork demonstrates the artist's creativity and skill in utilizing hangeul characters to create meaningful compositions. Unfortunately, not being a well-known artist, none of HangulArt's works was not sold as Hangul Art & Charism collection's total volume is 0 ETH.

6. Results

The analysis of NFT art collections featuring Asian characters, specifically Japanese kanji and Korean hangeul, on OpenSea yielded valuable insights into various aspects including economic value, cultural representation, and artistic expression. This section presents the key findings derived from the examination of the selected collections: 暗号漢字 -CryptoKanji-, Golden Kanji, Replica Kanji Flowers, Klminjeongeum NFT, Hangul NFT, and Hangul Art & Charism.

1) Economic Value:

暗号漢字 -CryptoKanji- collection, created by Japanese artist mera takeru, emerged as the most economically significant among the analyzed collections. With a total volume of 3 ETH (approximately $5212.35 USD), attributed to the artist's prominence and popularity within the NFT art industry, this collection experienced substantial sales activity, with most artworks being sold. Golden Kanji, comprising 2136 artworks representing jōyō kanji, exhibited a comparatively lower total volume of 0.0192 ETH (approximately $33.36 USD) on OpenSea. Despite the large number of artworks, sales activity was limited, with only five artworks being sold. Replica Kanji Flowers, a collection featuring circularly repeated kanji characters to create floral shapes, demonstrated

a total volume of 0.8886 ETH (approximately $1561.97 USD). Notably, the highest sale price was 0.2 ETH (approximately $351.56 USD) for the artwork named 少 (kanji character meaning 'little'). Klminjeongeum NFT, centered on Korean hangeul letters and utilizing the Klaytn blockchain, showcased a total volume of 92133.6545 KLAY (approximately $13353.06 USD). Significant sales were observed, with the highest price reaching 500 WKLAY (approximately $72.66 USD). Hangul NFT and Hangul Art & Charism collections, both dedicated to Korean hangeul, presented relatively lower total volumes of 0.0394 ETH (approximately $68.46 USD) and 0 ETH, respectively. Sales activity in these collections was modest, with minimal impact on total volume.

2) Cultural Representation and Artistic Expression:

暗号漢字 -CryptoKanji- collection by mera takeru and Golden Kanji by Nekoraisu demonstrated a focus on incorporating Asian characters into digital artworks as a means of cultural representation. Mera takeru's minimalist designs with bold black kanji letters aimed to promote Japanese culture overseas, while Nekoraisu's animated golden kanji letters celebrated the heritage of Japanese language learners. Replica Kanji Flowers by Kenta Suhara showcased an innovative approach to integrating kanji characters into floral designs, reflecting artistic creativity and cultural symbolism. The utilization of circularly repeated kanji characters to form flower shapes emphasized the fusion of traditional Japanese aesthetics with contemporary digital art techniques.Klminjeongeum NFT and Hangul NFT collections highlighted the significance of Korean hangeul letters as a cultural symbol and artistic medium. Klminjeongeum NFT, based on the Klaytn blockchain, emphasized the promotion of Korean culture through blockchain technology. Hangul NFT and Hangul Art & Charism collections celebrated the beauty of hangeul characters through vibrant visual compositions and creative interpretations.

3) Technological Implications:

The adoption of blockchain technology, particularly Ethereum and Klaytn platforms, facilitated the creation, distribution, and ownership of NFT art collections. The use of decentralized ledgers provided transparency, security, and authenticity to digital artworks, empowering artists to monetize their creations and engage with global audiences. The interoperability of NFT platforms like OpenSea enabled seamless trading of digital assets across different blockchain ecosystems, fostering a vibrant and interconnected NFT art market. The integration of Asian characters into NFT art collections transcended geographical boundaries, contributing to the global dissemination of cultural heritage and artistic innovation.

Overall, the analysis of Asian NFT art collections on OpenSea revealed diverse manifestations of cultural representation, artistic expression, and economic value within the rapidly evolving landscape of digital art and blockchain technology. These findings contribute to a deeper understanding of the intersection between art, culture, and technology in the digital age.

7. Conclusion

In conclusion, the analysis of NFT art collections featuring Asian characters, specifically Japanese kanji and Korean hangeul, on the OpenSea platform has revealed a rich tapestry of cultural representation, artistic expression, and economic value. Through a comprehensive examination of various collections such as 暗号漢字 -CryptoKanji-, Golden Kanji, Replica Kanji Flowers, Klminjeongeum NFT, Hangul NFT, and Hangul Art & Charism, we have uncovered the diverse ways in which artists leverage blockchain technology to showcase Asian cultural heritage and creativity. Economically, these collections have demonstrated varying degrees of success, with prominent artists like mera takeru commanding substantial sales volumes, while others experience more modest transactions. Nevertheless, the economic value of these collections goes beyond mere financial transactions, serving as a testament to the growing interest and investment in NFT art

within the global market. Culturally, Asian NFT art collections serve as a platform for preserving and promoting traditional scripts and symbols, such as kanji and hangeul, in the digital space. Artists like Nekoraisu and HangulArt have creatively incorporated these characters into their artworks, celebrating the beauty and complexity of Asian languages and cultures. Through vibrant visual compositions and innovative designs, these collections contribute to the global dissemination of Asian cultural heritage and foster cross-cultural appreciation and understanding. Technologically, the adoption of blockchain technology has revolutionized the art market, enabling artists to bypass traditional gatekeepers and engage directly with audiences worldwide. The transparency, security, and immutability provided by blockchain have empowered artists to assert ownership over their digital creations and establish new paradigms of value and authenticity in the digital art realm. In essence, the analysis of Asian NFT art collections on OpenSea underscores the transformative potential of technology in reshaping the landscape of art and culture. As we navigate the complexities of the digital age, it is imperative to recognize the profound impact of NFTs on artistic expression, cultural representation, and economic transactions. By embracing innovation and embracing diversity, we can harness the power of technology to create a more inclusive and vibrant artistic ecosystem for generations to come.

References

- Binance. (2021). 6 Key Indicators for NFT Collectors To Evaluate NFT Projects. Binance, [online] 10 November. Available at: https://www.binance.com/en/blog/nft/6-key-indicators-for-nft-collectors-to-evaluate-nft-projects-421499824684902985 [Accessed 17 March 2023].
- Bsteh, S. (2021). From Painting to Pixel: Understanding NFT artworks. [pdf] ResearchGate. Available at: https://www.researchgate.net/profile/Sheila-Bsteh/publication/351346278_From_Painting_to_Pixel_Understanding_NFT_artworks/links/609280ec92851c490fb7470e/From-Painting-to-Pixel-Understanding-NFT-artworks.pdf [Accessed 19 May 2023].
- Catalini, C., & Gans, J. S. (2020). Some simple economics of the blockchain. Communications of the ACM, 63(7), 80-90.
- Caxton, S. V. M., Naveen, K., Karthik, R., & Bama, S. S. (2022). User-Centered Evaluation and Design Suggestions for NFT Marketplaces. In: 2022 International Conference on Inventive Computation Technologies (ICICT). pp. 1214-1221.
- Dylan-Ennis, P. (2021). NFT Art: The Bizarre World Where Burning a Banksy Can Make It More Valuable. The Conversation, [online] 6 March. 2021. Available at: https://theconversation.com/nft-art-the-bizarre-world-where-burning-a-banksy-can-make-it-more-valuable-156605
- Economou, M. (2015). Heritage in the digital age. In A Companion to Heritage Studies (pp.215-228). John Wiley & Sons.
- Golden Kanji. (2021). What is Golden Kanji? Medium, [blog] 16 August. Available at: https://goldenkanji.medium.com/what-is-golden-kanji-f4b6b5cdf507 [Accessed 10 June 2021].
- Hall, S. (2015). Cultural Identity and Diaspora. In Colonial discourse and post-colonial theory (pp. 392-403). Routledge.
- Kılıçaslan, F., & Ekizler, H. (2022). Factors effecting purchase intention in blockchain and NFT (non-fungible token) technologies. Journal of Research in Business, 7(2), 604-623.
- Kochetkova, M. (2020). BLOCKCHAIN IN THE ART MARKET: Opportunities and Challenges [Bachelor's thesis, LAB University of Applied Sciences].
- NFTKLE.co. (2021). Klminjeongeum. [online] Available at: https://www.nftkle.co/klminjeongeum [Accessed 12 June 2023].
- OpenSea. (2023). [online] Available at: https://opensea.io/ [Accessed 12 June 2023].
- Prasad, C., Rao, B. S., Pujari, J. J., & Hema, C. (2023). Developing a Non-Fungible Token-Based Trade Marketplace Platform Using Web 3.0. In 2023 5th International Conference on Inventive Research in Computing Applications (ICIRCA). pp. 312-316.
- Russell, F. (2022). NFTs and Value. M/C Journal, 25(2). https://doi.org/10.5204/mcj.2863
- Saul, S. O. D. (2020). The Words We Can't Hear: Decoding the Language of Objects Through

the Eyes of Object-Oriented Ontology [Master's thesis, Colorado State University].

· Shitara, Y. (2021). Nihon no kuriputoāto no senku-sha mera takeru wa ima no NFT būmu o dō mite iru ka? (日本のクリプトアートの先駆者 mera takeru は今のNFTブームをどう見ているか). Atarashī keizai, [online] 20 April. Available at: https://www.neweconomy.jp/features/nft/110294 [Accessed 1 February 2022].

· Tapscott, D., & Tapscott, A. (2016). Blockchain revolution: how the technology behind bitcoin is changing money, business, and the world. Penguin.

· Wang, Q., Li, R., Wang, Q., & Chen, S. (2021). Non-fungible token (NFT): Overview, evaluation, opportunities and challenges. arXiv preprint arXiv:2105.07447.

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.

Key Words: Web3, decentralized Internet, ENS, blockchain, news website, Web feature analysis

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).
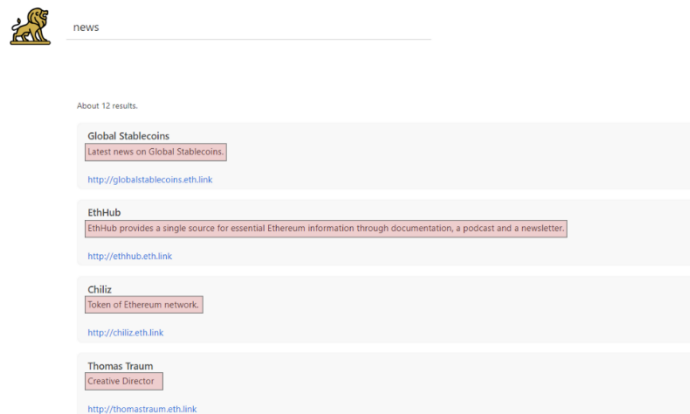

Figure 1: Introductory text on decentralized websites

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.
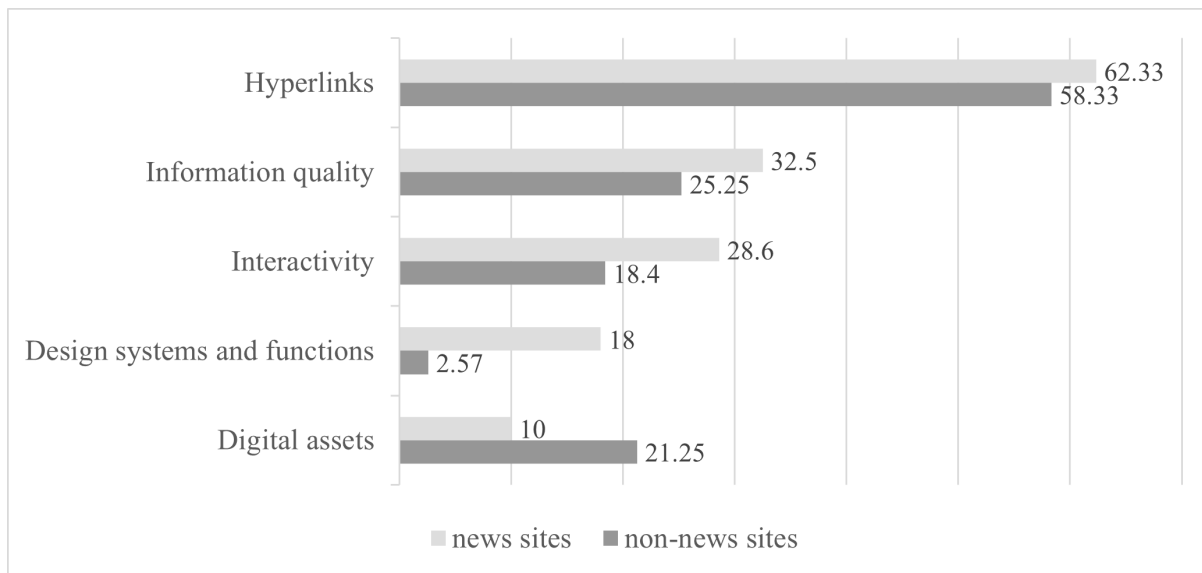

Figure 2: Comparing the characteristics of news and non-news sites across major categories

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.

Hi 👋

Welcome to my place on the web

My name is Alex Reyes and I'm a student majoring in Computer Science at Penn State University (class of 2021). If you're looking to get in touch, here's my email

Personal

A few things about me:

- I love to travel, explore new places, and snowboard
- I love reading. Here's my goodreads where you can see what I've been up to
- I also enjoy writing, and I'm getting a minor in World Literature
- I've loved tech my whole life, and it's something I'm super passionate about
- A fun fact about me is I used to live in Cape Verde, Africa

I like meeting new people! If you're reading this, send me a DM me on twitter

Professional

I will be @ Facebook as a TPM full time after graduation

A short history:

- I interned @ Facebook as a TPM intern after my junior year (Summer 2020)
- I interned @ Microsoft as a SWE intern after sophomore year (Summer 2019)
- I interned @ Microsoft as an Explore intern after freshman year (Summer 2018)
- I interned @ Listrak as a SWE intern after highschool (Summer 2017)

For more, connect with me on Linkedin
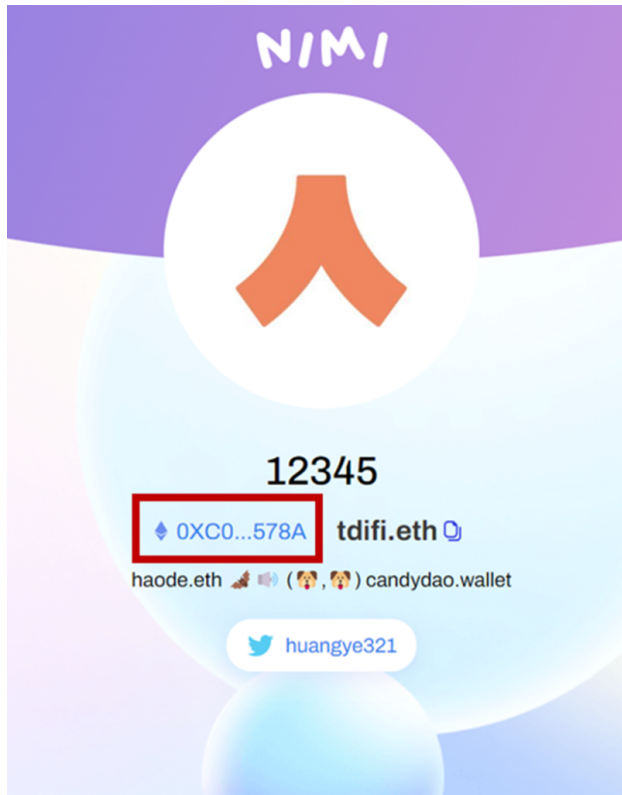
Figure 3: Single-page news site

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.
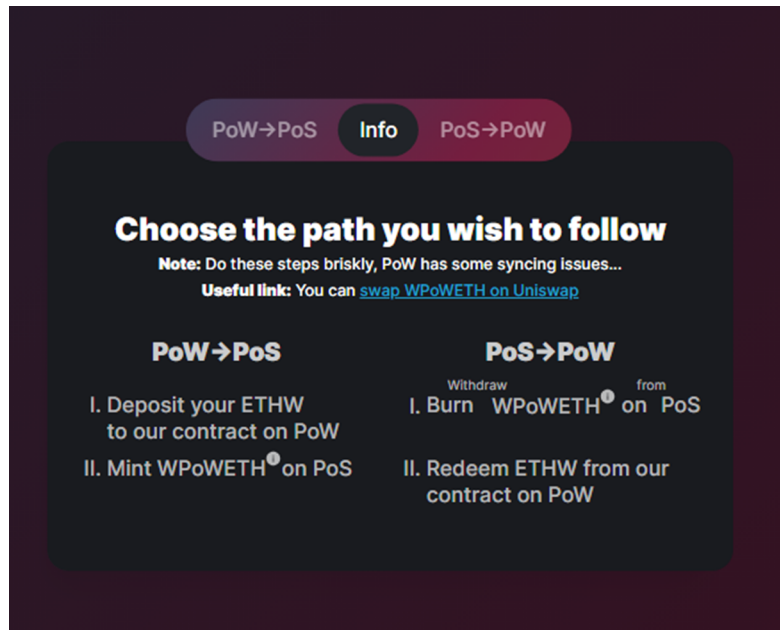
Figure 5: First screen of non-news site

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.
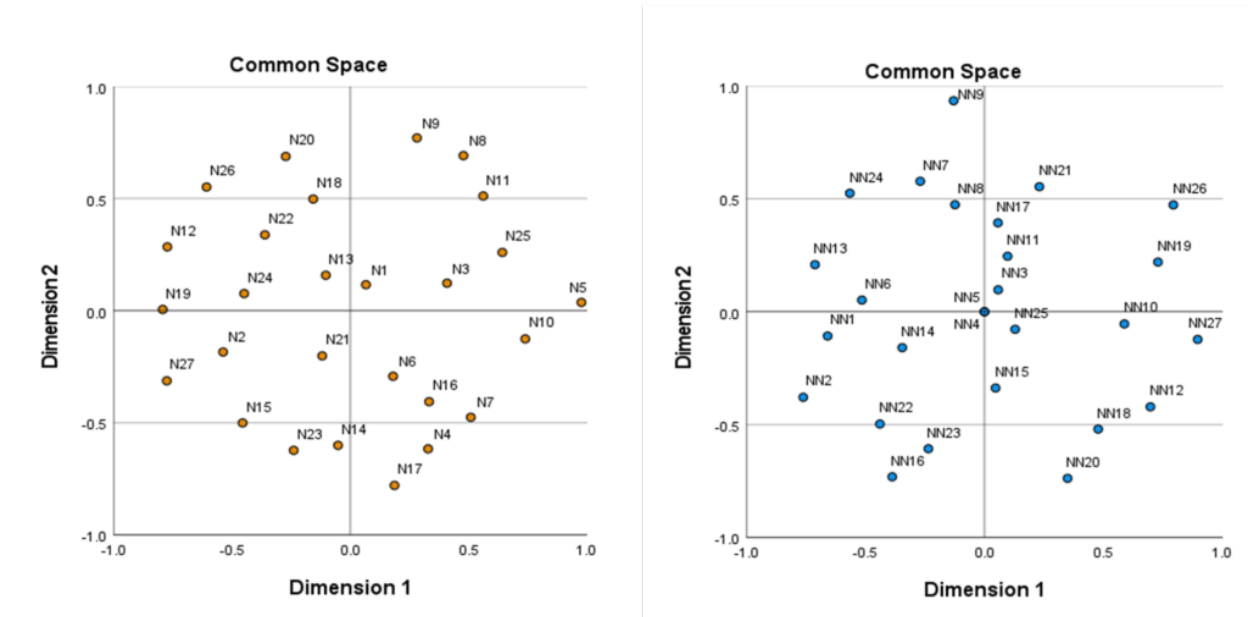
Figure 6: Results of multidimensional scaling analysis of news (left) and non-news sites (right)

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.
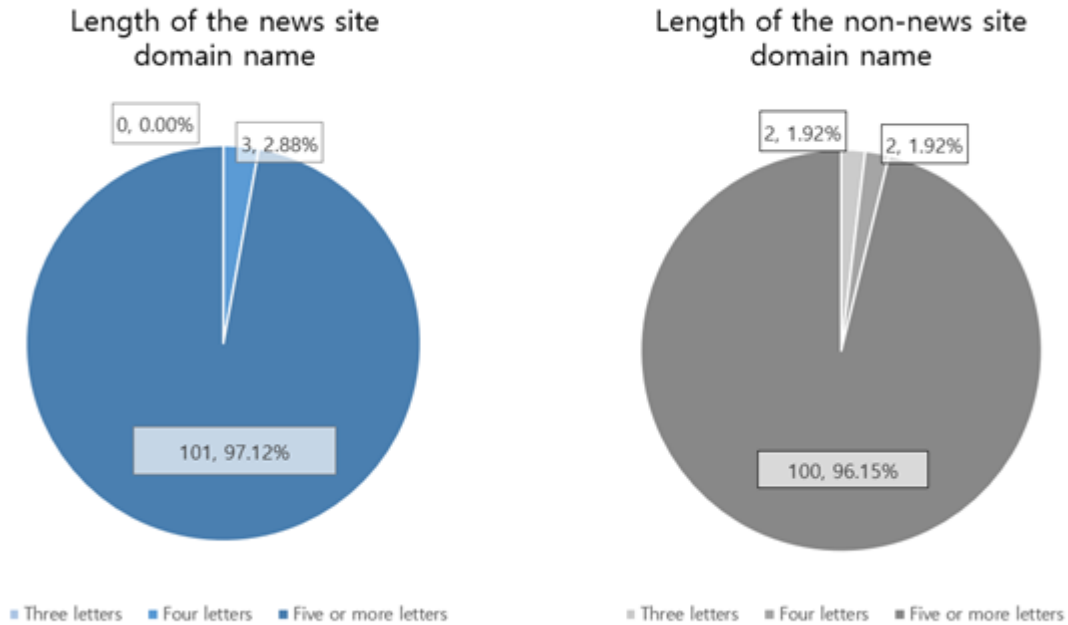
Length of the news site
domain name

0, 0.00%

3, 2.88%

101, 97.12%

■ Three letters ■ Four letters ■ Five or more letters

Length of the non-news site
domain name

2, 1.92%

2, 1.92%

100, 96.15%

■ Three letters ■ Four letters ■ Five or more letters

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

Rank	News	Degree centrality	Non-news	Degree centrality
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
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.
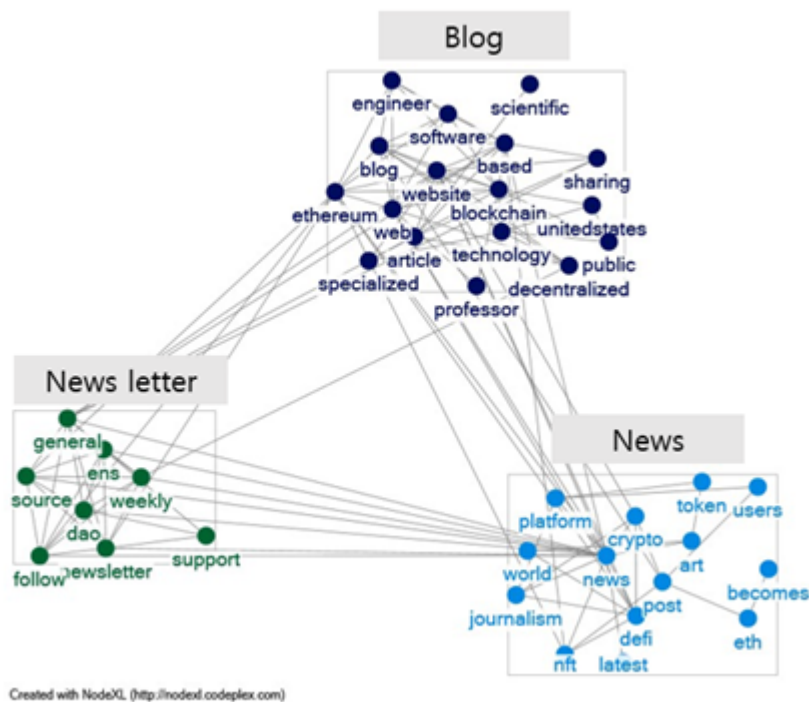

Figure 10: News sites cluster visualization

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."
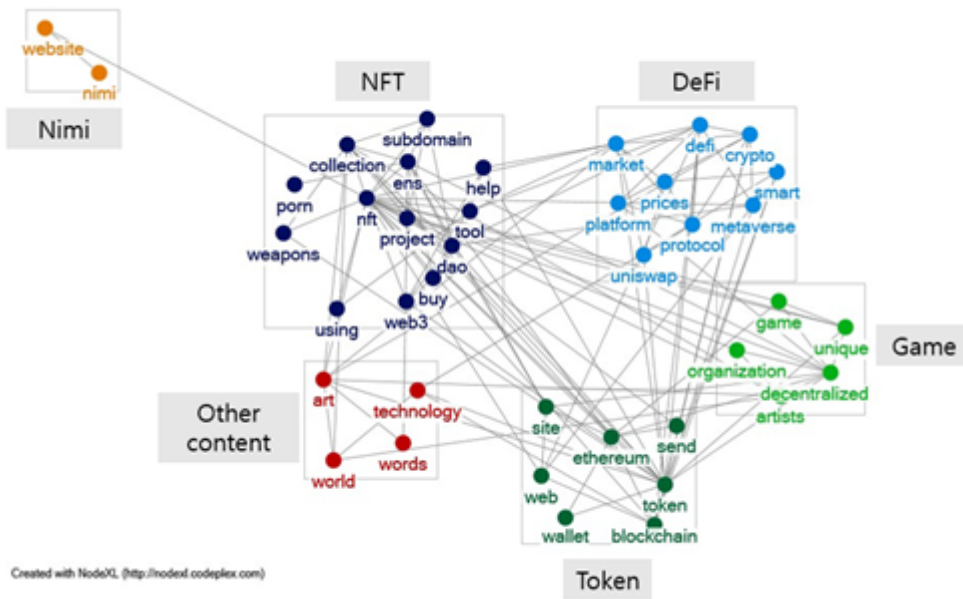
Figure 11: Non-news sites cluster visualization

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.