Scientometric Analysis of Papers on Ukrainian Heritage: History, Culture, and Literature

Iryna Balagura[1,2], Andriy Kryuchyn[1]

[1] Institute for Information Recording of National Academy of Sciences of Ukraine, 2, Mykoly Shpaka Street, Kyiv, 03113, Ukraine

[2] University of Nottingham, University Park, Nottingham, NG7 2RD, United Kingdom

Abstract
The present study consists of a comparative scientometric analysis of abstracts from a national and an international scientific database, and an analysis of information generated by artificial intelligence. The importance of research on topics regarding the national heritage of Ukraine rose during the last decade because of the necessity to preserve and defend the territory, identity and heritage. The aim of the study was to identify and compare the main keywords and corresponding co-word networks related to Ukrainian national heritage in Ukrainian scientific papers, international journals and on the Internet. The co-word and co-author networks were analyzed and compared, and the main clusters of keywords were distinguished. The semantic connections between the concepts *history, culture, literature* and *heritage* were shown. The most popular keywords in all the datasets were connected with national identity, heritage, culture and history. There was a significant increase in the number of papers in international journals during 2019-2022. Keywords with the highest degree of centrality in the co-word network using ChatGPT were Lviv, Easter traditions, Vasyl Stus, Ukrainian fashion and others which reflect the most common interests of wider society.

Keywords
Co-word network, national heritage, scientometrics, ChatGPT, Ukrainica naukova, Scopus

1. Introduction

Studying national heritage is essential for fostering cultural identity, preserving diversity, understanding history, and promoting social cohesion. It contributes to a sense of pride and appreciation of human culture, both on a national and global scale. Research regarding national heritage is of great importance for each country and plays a significant role in representing the country to the global community. World heritage was first described in the World Heritage Convention ratified by UNESCO in 1972 which defined cultural and natural heritage, drew up a World Heritage list and ratified nations to agree to cooperate in the protection of heritage [1]. However, the concept of heritage has evolved and changed through the years, influenced by political, economic, and social factors, such as the development of tourism and urbanization, reflecting the ongoing contributions of society [2]. Cultural heritage and the changing understanding of it can be described by scientometric research. Several scientometric studies have revealed hidden patterns and challenges, and have shown progress in research and collaboration in the area of cultural heritage [3-7].

Heritage consists of a wide range of elements, which can be categorized into Natural and Cultural heritage [8]. Cultural heritage includes tangible, intangible cultural heritage and cultural landscape heritage. Tangible heritage includes historical sites and monuments that hold historical, architectural and cultural significance, museums and archives (immovable heritage), arts and artefacts (movable cultural heritage). Intangible heritage includes traditions and customs, language and literature, folklore and oral traditions, performing arts and culinary traditions. Natural heritage includes geological, geochronological, and geomorphic heritage; biological, zoological and botanical heritage,

✉ balaguraira@gmail.com (I. Balagura); kryuchyn@gmail.com (A. Kryuchyn)

ⓘ 0000-0001-9627-2091 (I. Balagura); 0000-0002-5063-4146 (A. Kryuchyn)

and natural landscape including underwater cultural heritage (protected areas, landscapes and ecosystems).

The importance of studying, representing and preserving the national heritage has become of great importance for Ukraine in recent decades, for several reasons: reviving and integrating heritage into the EU and preservation during the current war. Consolidation of Ukrainian society particularly needs intangible cultural heritage [8]. Ukraine prioritizes European integration which includes cultural integration into the European Union (EU) space [9]. At the same time, the EU promotes a narrative of the common European past and shared cultural heritage [10]. Russia`s war has put at risk, destroyed, damaged and endangered Ukrainian heritage, which requires immediate intervention from the international cultural heritage community [11]. Given the current importance of research in the area of Ukrainian national heritage, scientometric analysis of papers regarding this topic could show progress and future routes in this direction. The aim of this paper is to discover the main research topics, scope and meaning of Ukrainian heritage in scientific papers and on the Internet using complex network analysis of the *Ukrainika naukova* abstract database, *Scopus* and *ChatGpt*. The research aimed to identify and compare the main keywords and corresponding co-word networks related to Ukrainian national heritage in Ukrainian scientific papers, international journals and on the Internet.

2. Methods and results

National heritage encompasses a wide range of tangible and intangible elements that collectively represent the cultural, historical, and natural identity of a nation. Therefore, to analyse the topic the keywords *history, culture* and *literature* were chosen and searched in *SCOPUS* and *Ukrainika naukova*. Co-word-author-journal networks were formed for each keyword. Times Series analysis of publication activity and main keywords showed that main research in these areas of Ukrainian national heritage gradually moved from Ukrainian to international space.

Keywords and co-word networks were also analysed using ChatGpt 3.5 and the methodology proposed by Lande et al [12].

The first such was conducted in the scientometric information system *Science metric library*, developed in the Institute for Information Recording of the National Academy of Sciences of Ukraine using the data from *Ukrainika naukova, Arxiv, Stack Exchange* and *CNKI* (China National Knowledge Infrastructure on the base of Manticore Search [13-15]. The information-analytical system was developed to investigate the records. The open-source database Manticoresearch was chosen as a base for the system to support SQL- and JSON-requests. The system interface provides options to use a range of records from scientific databases and Q&A websites for searching and obtaining analytical data. After searching by concept, analytical options include (the following): forming monthly time series of the frequency of concept usage, visualization of the wavelet transformation, charts of dynamics, and smoothing of data. The system also allows users to select keywords and form co-word networks using the improved *TD-IDF* method in Lande et al. [16], and also download the adjacency matrix for *Gephi*. A fuller picture of the chosen topic can be obtained by realizing the co-author network and the appropriate adjacency matrix for *Gephi.*

The search was implemented using the *Ukrainica naukova* database to obtain the time-serious of papers by year, the most common keywords and most productive authors, co-word and co-author matrixes and networks which could be further studied in other network software e.g. *Gephi. Ukrainica naukova* contains 825947 records with the retrospective to 1997 year for January of 2024, among which abstracts of books, journals, doctoral dissertations and conference papers. Most of the records are in Ukrainian, therefore the search was fulfilled with translation into Ukrainian using the same prompts *history* (історія), *culture* (культура) and *literature* (література). There were found 11695 records, the time series of the data presented in the fig. 1. The highest numbers of papers regarding the keyword in Ukrainian were published in 2000-2001 years.
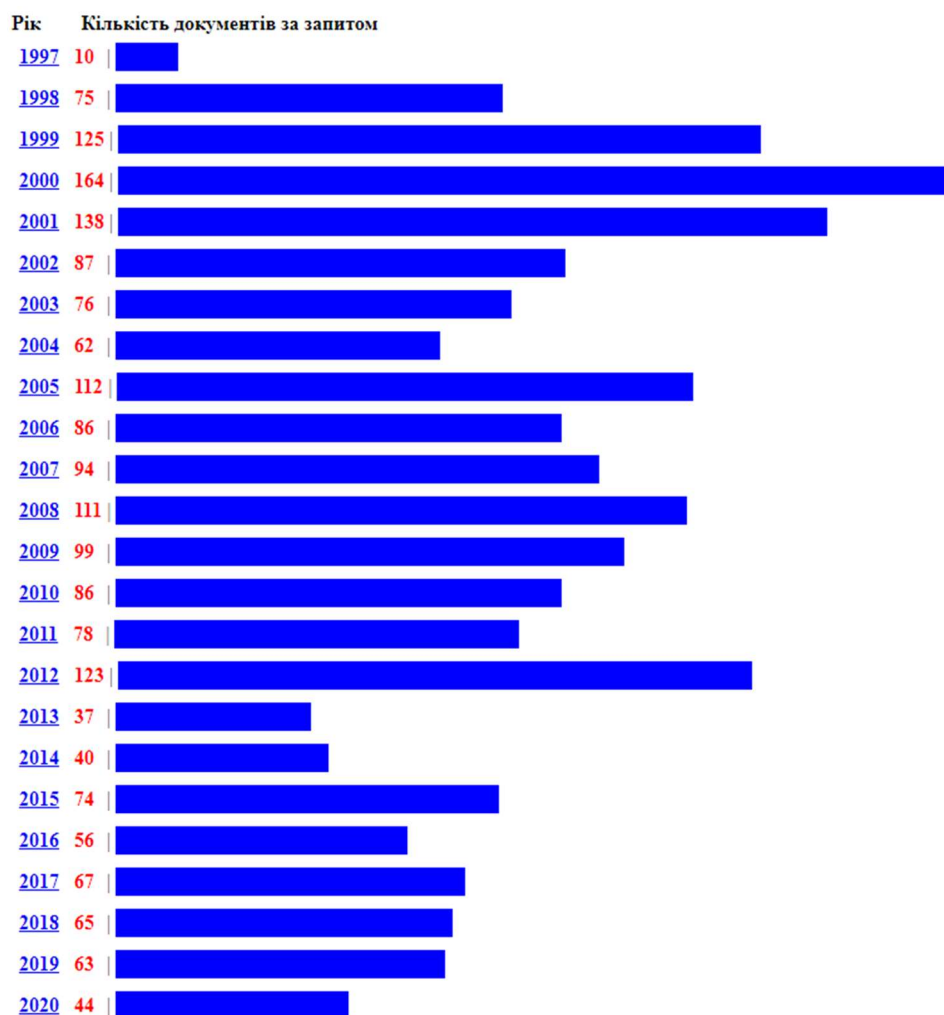
Figure 1: Time series of papers by the prompt *history* in the *Ukrainica naukova* database

The topic was analysed with co-word networks which are networks formed with the most common keywords which co-occur in the same text detected according to the frequency of their use in the abstract based on the *TF-IDF* algorithm [16-18]. The co-word network is presented in Fig. 2. Nodes of the network or keywords could be estimated by measuring centrality measures. The degree of centrality represents the number of links connected with the node, which could be interpreted as the most common words and terms [19]. The most common keywords and their degree of centrality are presented in Table 1.

Table 1

Most common keywords by the prompt *history* in *the Ukrainica naukova* database

Keywords	Degree
History	54
Historiography	9
History of science	9
Ukraine	8
Medicine	7
Numismatics	7
History	5
Kyiv	5
Education	5

Research history	5
History of development	5
Galicia	4
Department	4
History of multimedia	4
Introduction	4
History of Ukraine	4
History of science and	4
technology	4

The co-author and co-word networks use the *Ukraininika naukova* data for the keyword *history* obtained with *the Science metric library*. Both networks have a low graph density of about 0.03 according to the wideness of the term which applies to all disciplines, which explains the variety of terms in co-word networks (fig.2). Co-author networks have 30 connected components with an average degree of 3 (on average an author has three co-authors). History and other social sciences are areas of research where papers are written by fewer authors than in science, so the network is not dense.

Figure 2: Co-word network by the prompt *history* in the *Ukrainica naukova* database

To make a comparison of the research representation the search of the papers by keyword *history* also was done using the database *Scopus*, where the prompt was limited by the affiliation country Ukraine and subject area *Art and Humanities*: TITLE-ABS-KEY (history) AND (LIMIT-TO (AFFILCOUNTRY , "Ukraine")) AND (LIMIT-TO (SUBJAREA , "ARTS")). In *Scopus* were found 999

documents and most papers were published during 2019 – 2022 years, which could be explained by the government and scientific policy of Ukrainian integration into the EU. It is possible to check the evolution and main concepts which appear through the years in the papers (table 2). The keywords *culture, national heritage*, and *identity* were the most common topics, so appear about every year the topic of the war arises in 2022. The analysis of keywords was done using *Scincescape* [20].

Table 2
The most frequent keywords by year in the Scopus database by the keyword *history* in papers produced by scientists affiliated with Ukrainian institutions (made with *Sciencescape*)

2019	2020	2021	2022	2023
Ukraine (14)	Ukraine (13)	Ukraine (20)	Ukraine (15)	Ukraine (10)
Translation (7)	Culture (7)	Identity (9)	Culture (10)	Culture (4)
Culture (5)	Translation (7)	Culture (6)	Identity (5)	War(4)
National identity (4)	Barrow (5)	Poetry (5)	Cultural heritage (5)	Translation (3)
Transformation (4)	Identity (4)	Neolithic (4)	History (5)	Cultural heritage (3)
Eneolithic (3)	Neolithic (4)	Cultural heritage (4)	Japan (5)	Religion (3)
Eastern Europe (3)	Rusins (4)	Rusins (4)	Postmodernism (5)	Creativity (3)
Rusins (3)	Bronze age (4)	Archaeology (4)	Chronology (4)	Ukrainian Literature (3)
Bronze age (3)	National identity (4)	Ideology (4)	Bronze age (4)	
Concept (3)			Education (4)	

A co-word-author-journal network was also created which combines the most active authors, their main keywords and the journals, where most of the papers are published (fig. 3,table 3). The Papers are published mostly in Ukrainian historical journals such as *Eminak, Manuscript* and *Book Heritage of Ukraine, Shidnij svit* and others. The set of the keywords in the network (table 3) differ from the keywords (table 2) by the year and cumulatively describes the main topics of scientific research of authors affiliated with Ukraine.

Table 3
Co-word-journal- author networks description using *Scopus* database by the keyword *history* in papers produced by scientists affiliated with Ukrainian institutions (made with *Sciencescape*)

| Main authors (papers)	Main keywords (papers)	Main journals (papers)
Pavlenko S. (4)	Ukrainian (69)	Eminak (100)
Secundant S. (4)	History (31)	Manuscript and book heritage of Ukraine (70)
Shandra R. (4)	Historiography (20)	
Bazaluk O (3)	Russian empire (20)	Shidnij svit (49)
Danilets J. (3)	Crimea (17)	History of science and technology (42)
Grinchenko G. (3)	Education (15)	
Hanna D. (3)	Identity (13)	Sententiae (38)
Kapranov S. (3)	USSR (12)	Rusin (36)
Krupnyk I.(3)	Olbia (10)	Bylye gody(26)
Levchenko V (3)	History of archeology(9)	Bibliotekarz podlaski(35)
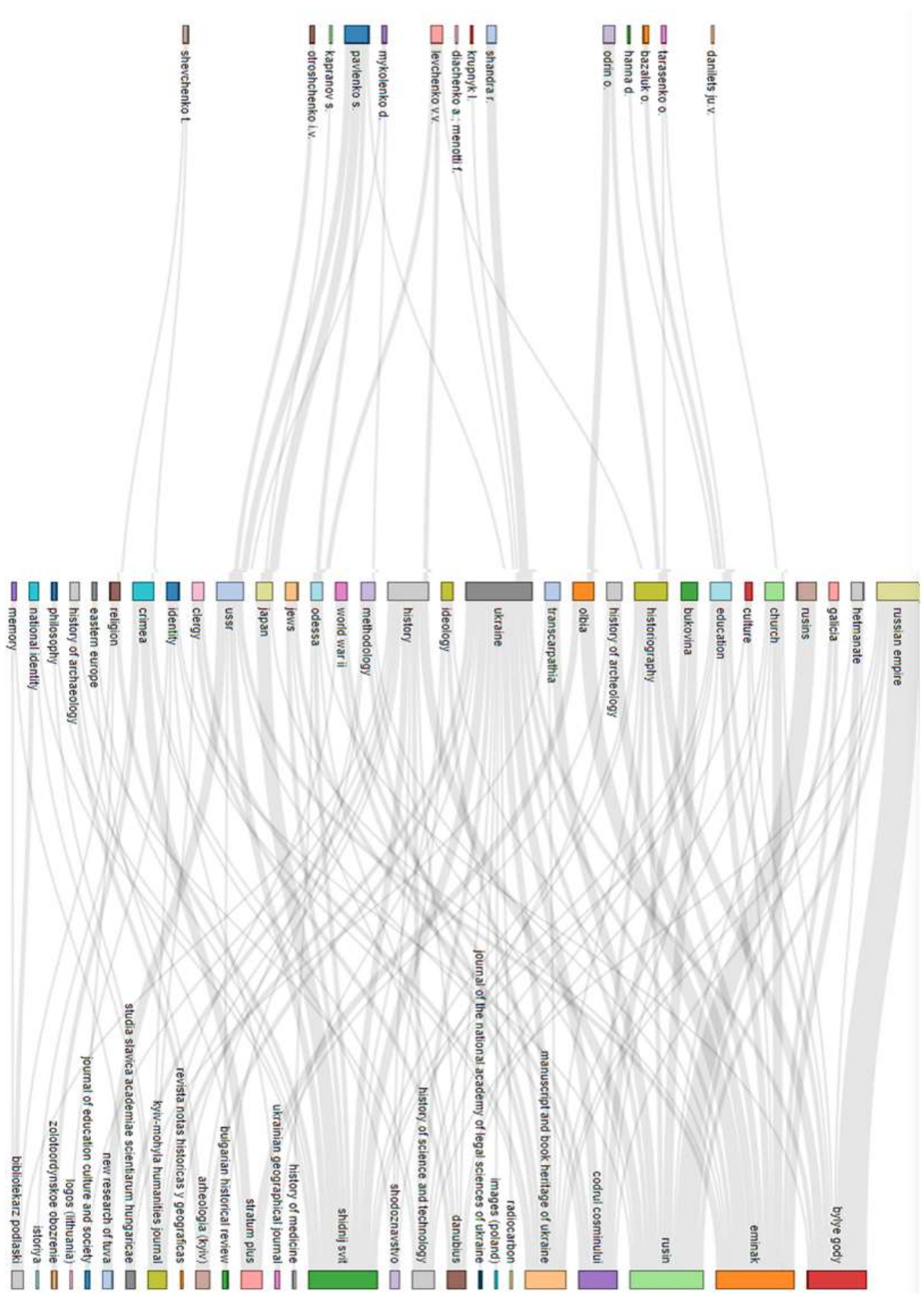
Figure 3: Co-word-journal- author networks using *Scopus* database by the keyword *history* in papers produced by scientists affiliated with Ukrainian institutions (made with *Sciencescape*).

Similar findings were done for the keyword *culture*, the number of papers in *Scopus* was 1642 in the section *Arts and Humanities.* The most common keywords in the co-word network using *VOSviewer* software are *Ukraine, culture, cultural heritage, identity, Eastern Europe, translation* and others, which recalls the dataset for the history keyword and shows the importance of the topic of national heritage and identity for Ukraine (fig 4.) [19]. The list of keywords also includes historical concepts such as *neolithic, chronology, archaeological evidence, archaeology, palaeolithic, bronze age, cultural history, settlement history*, and *historical geography,* which shows the closeness of these two areas.
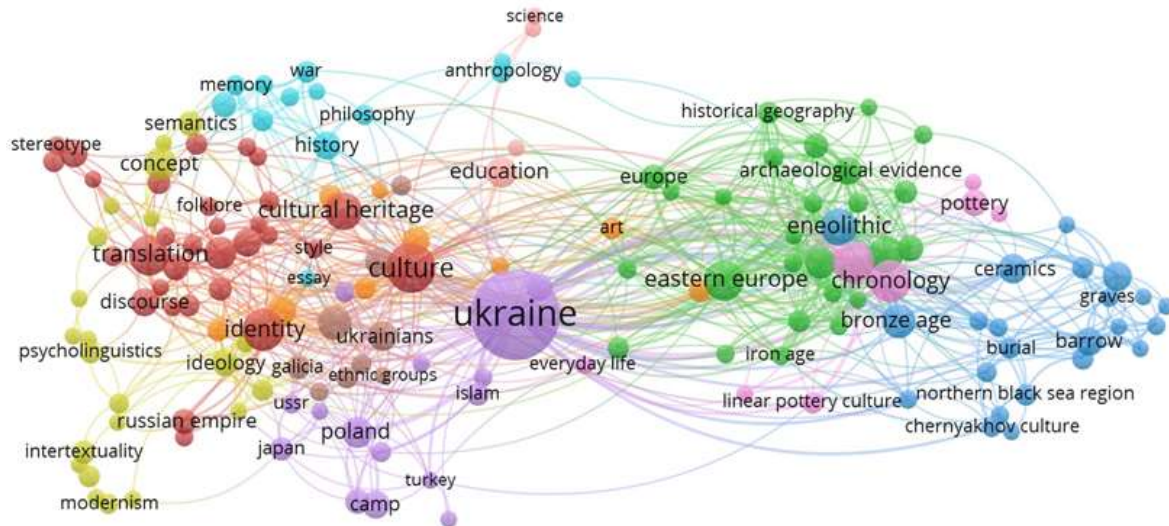

Figure 4: Co-word network by the keyword *culture* using *Scopus* database (made by VosViewer)

The co-word-author-journal network was done for the term *literature* using the prompt TITLE-ABS-KEY (litera*) AND (LIMIT-TO (AFFILCOUNTRY , "Ukraine")) AND (LIMIT-TO (SUBJAREA , "ARTS")) in the *Scopus* database, there were found 836 papers (table 4). The journals` list is similar to previous networks mostly includs Ukrainian journals, among them *Shidnij svit, Visnyk Universitetu imeni Alfreda Nobelya, Bibliotekarz podlaski, Manuscript and book heritage of Ukraine, Eminak, Rusin* and others. The main keywords are *Ukrainian literature, Ukraine, translation, Ukrainian language, intermediality, poetry*, and also *national identity, culture, history, historiography* which emphasize the connectivity to the history and culture and importance of research of the national heritage and identity.

Table 4

Co-word-journal- author networks description using *Scopus* database by the keyword *literature* in papers produced by scientists affiliated with Ukrainian institutions (made with *Sciencescape*)

Main authors (papers)	Main keywords (papers)	Main journals (papers)
Bezrukov A.(7)	Ukrainian literature (28)	Shidnij svit (38)
Bohovyk O. (7)	Translation (17)	Visnyk univeritetu Alfreda
Kolomiyets I. (5)	Ukrainian language (16)	nobelya.Seriya:filologichni
Petrova Y.(3)	Intermediality (13)	nauki(3)
Pukhonska O. (3)	Poetry (13)	Bibliotekarz podlaski(35)
Rudnytska N.(3)	Identity (12)	Manuscript and book heritage
Savchyn V.(3)	Intertextuality (11)	of Ukraine (32)
Shostak O. (3)	Genre (10)	Psycholinguistics (25)
Stepanova A. (3)	Text(10)	Eminak (22)
Vorobyova O (3)	Culture (9)	Rusin (19)

The three co-word networks which were obtained using data from *Scopus* and the co-word network using data from *Ukrainica naukova* are compared. The prompt for the Ukrainian database was not limited by the area of research, so the keywords in the networks are made of basic concepts which refer to history, literature and culture and connected disciplines. It is important to take into account that *Ukrainica naukova* consist of papers since 1998 and reflects the cumulative information. Co-word networks (fig.2-4) reflect mostly topics of the last five years.

Description and presentation of the national heritage for the society implemented through the Internet websites and social media. *Chat GPT* was used as an instrument for the comparison of the most popular topics regarding the national heritage of Ukraine on the Internet using the algorithm and technology presented by Lande et al in [20]. *Chat GPT* is a language model based on the Generative Pre-trained Architecture (GPT-3,5) and communicates with users using artificial intelligence (ChatGPT, 2023). To form a network of 500 nodes the following prompt was used iteratively: `Give me another 50 main pairs of linked concepts correspondent with Ukrainian heritage: history, culture, and literature in the format "concept 1; concept 2"` (table 5, fig.5).

Table 5

Most common keywords in co-word network by *Chat GPT*

Keywords	Frequency
TRADE	30
COSSACS	24
LVIV	11
TARAS SHEVCHENKO	9
HETMANATE	9
ARCHITECTURE	8
BANDURA	7
ODESA	7
BATTLE OF KRUTY	6
UKRAINIAN DIASPORA	6
IVAN FRANKO	6
UKRAINIAN FAIRY TALES	6
LESYA UKRAINKA	5
UKRAINIAN RIDDLES	5
SAINT SOFIA CATHEDRAL	5
HOLODOMOR	5
HETMAN IVAN MAZEPA	5
LVIV OPERA HOUSE	5
SALT TRADE ROUTES	5
UKRAINIAN KOBZAR	5
ORAL TRADITION	5
KYIV RUS	5
UKRAINIAN FOLK DANCES	5
BLACK SEA	5
KYIV ARSENAL	5

The most frequent terms in the network around which clusters were formed were *Taras Shevchenko, Kobzar, Ukrainian embroidery, Ivan Franko, Lviv`s intellectual life, Ukrainian riddles* and *oral traditions*. The biggest cluster consists of the topic of *Taras Shevchenko* and *the Ukrainian language*, the most common keywords in the whole network are *TRADE(30), Cossacks(24), Lviv(11),*

Hetmanate(9), Taras Shevchenko(9). These keywords show important figures in history and literature and national traditions.
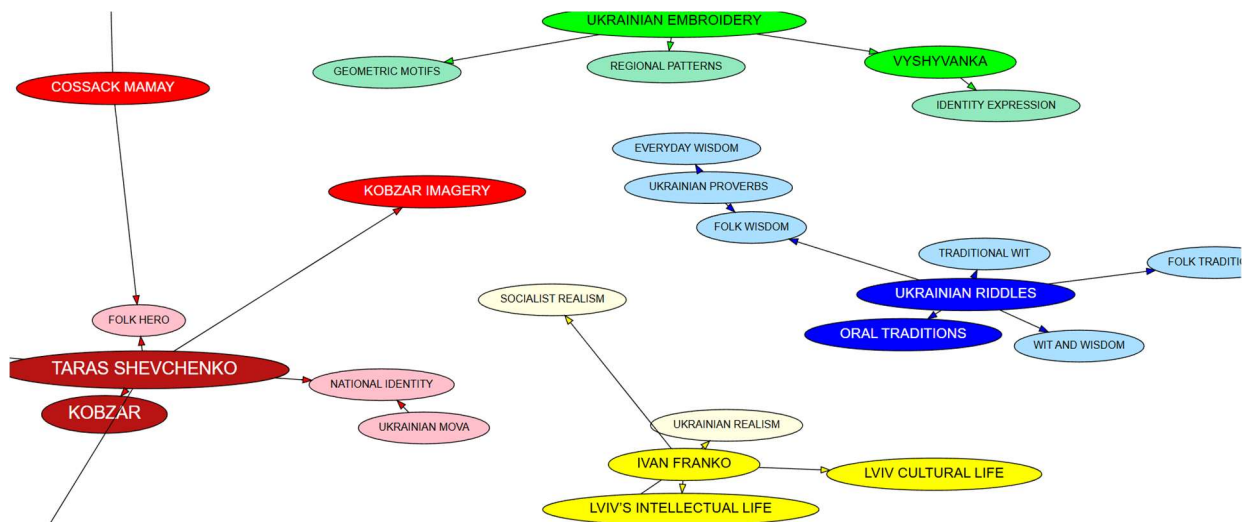

Figure 5: Part of the co-word network by ChatGPT

3. Conclusions

The topic of the national heritage of Ukraine was compared in the academic and non-academic data. Scientometric analysis of the research output of authors affiliated with Ukrainian institutions and the co-word networks analysis using *ChatGPT* were presented. The data for the analysis was gathered from the *Ukrainika naukova* and *Scopus* databases. Co-word network described the main topics of research and showed the strong connection between history, culture and literature and the importance of research in the directions of national identity and heritage. An increase of papers in the international database was also shown, which could be the result of the policy and intent of Ukraine of integration into the European Union space.

Co-author and co-word networks for the keyword history using the national and international databases revealed similar keywords and similar authors with high centrality measures. The low densities of the networks could be explained by the low level of co-authorships in the area of *Arts and Humanities.* The co-word network using *ChatGPT* is different and highlights the most important figures from history and literature, Ukrainian, which are present Ukrainian heritage on the Internet.

The methods and tools employed here could be used for many purposes, including the identification and description of scientific groups and research topics, the most communicative researchers, and main principles of science communication; estimating the level of inter-scientist cooperation, detecting actual topics and priorities, and possible science cooperation, and determining associative and entry relations for subject field model formation.

4. Acknowledgements

References

[1] UNESCO, the World Heritage Convention, 2024, URL: https://whc.unesco.org/en/convention/

[2] Yang Liu, Karine Dupre and Xin Jin, A systematic review of literature on contested heritage, Current Issues in Tourism, (24)4, pp. 442-465 2021 doi 10.1080/13683500.2020.1774516

[3] Mansuri, L., Udeaja, C., Trillo, C., Kwasi, G., Patel, D., Jha, K., Makore, CB and Gupta, S (2019). Scientometric analysis and mapping of digital technologies used in cultural heritage field. Association of Researchers in Construction Management, ARCOM 2019. Leeds 02 - 04 Sep 2019 pp. 255-264

[4] Vlase, I., Lähdesmäki, T. A bibliometric analysis of cultural heritage research in the humanities: The Web of Science as a tool of knowledge management. Humanit Soc Sci Commun 10, 84 (2023). https://doi.org/10.1057/s41599-023-01582-5

[5] (Mansuri et al., 2019, Vlase, 2023, Maldonado-Erazo et al, 2021, Basel, 2021).

[6] Maldonado-Erazo, C.P.; Álvarez-García, J.; Río-Rama, M.d.l.C.d.; Durán-Sánchez, A. Scientific Mapping on the Impact of Climate Change on Cultural and Natural Heritage: A Systematic Scientometric Analysis. Land 2021, 10, 76. https://doi.org/10.3390/land10010076

[7] Prados-Peña, María Belén, George Pavlidis, and Ana García-López. "New technologies for the conservation and preservation of cultural heritage through a bibliometric analysis." Journal of Cultural Heritage Management and Sustainable Development (2023).

[8] Chmil, H., Kuznietsova, I., Mishchenko, M., Oliynyk, O., & Demeshchenko, V. (2021). Intangible cultural heritage as a resource for consolidating modern Ukrainian society. Linguistics and Culture Review, 5(S4), 747-760. https://doi.org/10.21744/lingcure.v5nS4.1713

[9] Sun Hua, World Heritage Classification and Related Issues—A Case Study of the "Convention Concerning the Protection of the World Cultural and Natural Heritage", Procedia - Social and Behavioral Sciences, Volume 2, Issue 5, 2010, Pages 6954-6961, ISSN 1877-0428, https://doi.org/10.1016/j.sbspro.2010.05.048

[10] Shevchenko, M. O. "European Union experience in digitizing historical and cultural heritage: ways of implementation in Ukraine." (2020): 41.

[11] Lähdesmäki, Tuuli, et al. Creating and Governing Cultural Heritage in the European Union: The European Heritage Label. Taylor & Francis, 2020.

[12] Lande, Dmytro, Anatolii Feher, and Leonard Strashnoy. Cybersecurity in AI-Driven Casual Network Formation. Theoretical and Applied Cybersecurity 5.2 (2023).

[13] Science Metric library URL: https://bigsearch.space/cgi-bin/frontlib.pl

[14] Certificate of copyright registration for the work. Computer program of scientific potencial analysis Science Metric Library http://dwl.kiev.ua/art/AS/as119845/index.html

[15] Lande D.V., Kryuchyn A.A., Dobrovska S.V., Balagura I.V. Use of the «Library of Science Metrics» system for conducting science metric research. Data Rec., Storage & Processing. 2023. Vol. 25, No. 1. P. 32–42.

[16] Lande D., Dmytrenko O., Snarskii A. Transformations of texts into the complex network with applying visibility graphs algorithms (2018) CEUR Workshop Proceedings, 2318, pp. 95 - 106

[17] Lande, Dmytro, Dmytrenko, Oleh, Creating Directed Weighted Network of Terms Based on Analysis of Text Corpora, (2020) 2020 IEEE 2nd International Conference on System Analysis and Intelligent Computing, SAIC 2020, art. no. 9239182, Cited 6 times. DOI: 10.1109/SAIC51296.2020.9239182

[18] Lande D., Dmytrenko O., Radziievska O. Determining the directions of links in undirected networks of terms (2019) CEUR Workshop Proceedings, 2577, pp. 132 - 145

[19] Lande, Dmytro, et al. Link prediction of scientific collaboration networks based on information retrieval. World Wide Web 23 (2020): 2239-2257.

[20] ScienceScape, URL: https://medialab.github.io/sciencescape/

[21] VOSviewer Vizualization scientific landscapes URL: https://www.vosviewer.com/

[22] Lande, Dmitry, and Leonard Strashnoy. "GPT Semantic Networking." (2023).