International Journal of Knowledge Processing Studies
(KPS)

Homepage: http://kps.artahub.ir/

ORIGINAL RESEARCH ARTICLE

Proposing a Knowledge Organization Model for Iranian Website Content Based on Existing Metadata Frameworks

Seyedeh Sara Moosavi[1], Sepideh Ciruskabiri[2], Ahmadreza Varnaseri[3*]
[1] *PhD Candidate, Retrieval Information, Kharazmi University, Tehran, Iran.*
[2] *PhD Candidate, Knowledge Management, Tarbiat modares University, Tehran, Iran.*
[3] *PhD Candidate, Information Science and Knowledge_ Information Retrieval and Knowledge, Faculty of Management, University of Tehran, Tehran, Iran. ahmadrezavarnaseri@gmail.com*

ARTICLE INFO

Article History:
Received: 2022/05/29
Accepted: 2023/06/29
Published Online: 2022/06/29

Keywords:
Web Archiving
Metadata elements
Knowledge organization
Iranian websites

Number of Reference: 29
Number of Figures: 0
Number of Tables: 4

DOI:
http://dx.doi.org/10.22034/kps.2022.344828.1023

ABSTRACT

This study aims to create an integrated metadata system to organize information and content knowledge of Iranian websites to protect, represent, access, retrieve and provide a suitable metadata model for archiving such resources in the country. This applied research was done by a documentary (library) and a survey method. In the initial study, the world's leading patterns in web archiving were selected, then metadata elements were extracted from the official site of each template and refined and homogenized in the comparative table. The research tool, a questionnaire including the description of each metadata element and the importance of its existence in describing web content was provided to information science experts. Due to the indigenous needs, open tables were allocated to record them. To organize the content knowledge of Iranian websites, thirty-eight metadata elements were prioritized and evaluated by information science experts. knowledge organization of web content will strengthen the representation, facilitate retrieval and prevent content loss when it is based on the standard and integrated metadata templates while increasing user satisfaction. In this research, information science experts have recognized that the metadata elements of Title, Subject, Creator, and Format are more important.

►**Citation (APA):** Moosavi, S., Ciruskabiri, S. & Varnaseri, A. (2022). Proposing a Knowledge Organization Model for Iranian Website Content Based on Existing Metadata Frameworks. *International Journal of Knowledge Processing Studies, 2*(3): 47-57.

1. Introduction

The speed and ease of production and loading of content on the web, the exponential increase in the volume and short life cycle of data in this space, have made the processes of collection, refinement, organization, storage, management, and dissemination of such data, information and knowledge associated with it, seem vital. Websites are ephemeral and often considered at-risk born-digital content. New websites form constantly, URLs change, content changes and websites sometimes disappear entirely. Websites document current events, organizations, public reactions, government information, and cultural and scholarly information on a wide variety of topics. Materials that used to appear in print are increasingly published online (Program Web Archiving, 2022).

Fewer individuals and legal entities today can be found who do not have a virtual page on the Internet or have not uploaded content in this range. The high volume of resources, information, digital documents, and documents created by the development of cyberspace in the Internet environment, has led to fundamental changes in the methods for retrieving and representing information and knowledge, as well as the development of exploration tools for scientific research. These developments prompted information scientists to turn their attention to archiving websites. Archiving a website means preserving the content, data, and media for future reference. Using a dedicated service such as the Wayback Machine, you can view older versions of a website. On a technical level, crawlers take snapshots of a website, which constitutes the archive itself. You're able to access it using a simple calendar and view each iteration in a timeline format if you wish (Ravoof, 2022). Indeed, Web archives are an initiative to protect the unsustainable content of information and knowledge for the use of present and future generations.

The web is a vast and important information space, and over the last twenty years, the effort to preserve and document it has grown from the pioneering efforts of the Internet Archive to a community of archivists, librarians, and other practitioners who have contributed to hundreds of repositories (Wiedeman, 2019). Libraries, archives, and other memory collections have a long history of tackling the challenges of preserving unsustainable media. The longevity of a website is unstable and its maintenance is very challenging. These websites are frequently updated, expanded, or rebuilt. Even any site can go unnoticed or disappear without warning. Due to the constant change of websites, both in terms of content and updates and in terms of technologies used, it is necessary to use tools and standards in web archiving so that archive content remains readable and accessible over time. This is because a significant percentage of web content, if not stored timely, simply disappears. Therefore, to maintain the integrity and coherence of historical, cultural, and scientific documents, it is necessary to preserve web content as soon as possible (timely). Web archiving is the process of gathering up data recorded on the World Wide Web, storing it, ensuring the data are preserved in an archive, and making the collected data available for future research (Niu, 2012). The main purpose of creating a web archive is to protect and preserve valuable information and knowledge available to researchers and users based on their needs (Hakimzadeh, 2017). The Internet Archive and several national libraries-initiated web archiving practices in 1996. The International Web Archiving Workshop (IWAW), begun in 2001, has provided a platform to share experiences and exchange ideas. The later founding of the International Internet Preservation Consortium (IIPC), in 2003, has greatly facilitated international collaboration in developing standards and open-source tools for creating web archives (Niu, 2012). The attempts to build general Web archives based on ongoing 'deliberative and purposive preservation of Web material' took off in the mid-1990's only a few years after the spread of the Web protocols. The approaches differed concerning both the range of the materials collected and the criteria for

selection. In 1996, the private Internet Archive (archive.org) in the US took a 'generalized philanthropic' approach aiming to cover the whole Web. The same year, Kulturarw3 in Sweden and Pandora in Australia (both based within the national libraries) took a national domain perspective. The kind of material collected also differed as the Internet Archive and Kulturarw3 collected the widest possible set of materials, while the Pandora project focused on a selected set of sites considered the most valuable or authoritative sites.

The Library of Congress began archiving web content in 2000. The library's web archives comprise collections of archived web content selected by subject specialists and covering designated events and thematic topics. Content is selected according to Library of Congress collection policy statements and supplemental guidelines, and the collections are international in scope. The web archives are a part of a continuing effort by the library to evaluate, select, collect, catalog, provide access to, and preserve digital materials for future generations of researchers. (Library of Congress, 2022) More than 29,000 websites in 85 subject collections from 200 countries in more than 100 languages are available in the Library of Congress Web Archive (IIPC WAC 2022: #WHYWEBARCHIVING: PRESERVING INTERNET CONTENT FOR RESEARCH USE, 2022).

The early initiatives followed a growing range of national initiatives, especially in Europe. National libraries are predominant agencies covering national domains, except for the US, where the non-profit Internet Archive provides worldwide coverage and the Library of Congress maintains a huge selective archive. In addition, a range of selective archives is established at major universities. Only a few Web archives, if any in Near Middle East, Africa, and South America. Thus, Web archiving is mainly established in the northern hemisphere even if 'this ever-growing heritage may exist in any language, in any part of the world, and any area of human knowledge or expression' (UNESCO charter on the preservation of digital heritage, Article 1). According to the charter all sorts of digital heritage, born-digital heritage included should be 'protected and preserved for current and future generations' (Charter Article 1). Web archives belong to the category of 'born-digital cultural heritage' (materials created in digital form) but they differ from other kinds of born-digital materials because archived Web materials may include coded Internet links in their messages. Because of the global reach of the address system of possible destinations from any anchor and the indefinite number of possible instructions to be performed by any link on the live Web, Web archives become more complex than any formerly known set of data, except for the live Web as a whole.

A list of Web archiving initiatives can be found on Wikipedia. A range of the initiatives listed are also members of The International Internet Preservation Consortium (IIPC) established in 2003.

Today, two major general and philanthropic archive initiatives are the Internet Archive, established in 1996, and Common Crawl (commoncrawl.org) established in 2007. Since 2006, the Internet Archive has provided a subscription-based archive service, Archive-it (archive-it.org) allowing anybody to establish a tailored Web archive, which may also be incorporated into the Internet archive. The European Internet Memory Research, a commercial offspring of the Internet Memory Foundation, provided a similar service, archive-the-net, since 2011 (Finnemann, 2019).

In the archiving of collections on the web, several dimensions must be considered. To optimally manage digital and Internet archives, we should implement codified programs (Doroodi, 2009).

Like the management of many other kinds of information resources, the workflow of web archiving includes appraisal and selection, acquisition, organization and storage, description and access (Niu, 2012).

Metadata standards that are designed to facilitate the retrieval of Web archive content are rules and regulations to harmonize, integrate, and enhance the exchange of information, and communication, and improve interoperability. Therefore,

upgrading metadata standards in various fields will increase the quality of services and interactions in the information system (Ahmadi, et al., 2009).

In this regard, due to the unsustainable nature of information and knowledge available on the official websites of the country and the need to protect, maintain and access such information and knowledge, in the context of the national network and semantic web environment, the need for comprehensive archiving of important domestic websites has doubled. The lack of a common approach to the production and registration of metadata for web resources in Iran is one of the most important challenges faced by users and webmasters. According to studies conducted in Iran, including Hakimzadeh 's research (2017), Nooshinfard et al. (2012), and the feasibility of creating a web archive in the National Library and Archives of the Islamic Republic of Iran, 2010, there is no clear and codified metadata model for organizing the content of resources on the web to archive domestic websites. Accordingly, this study intends to study the existing patterns and study the indigenous needs from the perspective of information science experts, organize information and relevant knowledge in internal websites, to provide a metadata model for archiving such resources as an effective step to achieve integration in content archiving of Iranian websites. Therefore, the importance of this study is to improve the management and information organization and archived web content knowledge to enhance accessibility, retrieval, and maintaining integrity.

Research questions

1. What are the proposed metadata elements based on indigenous needs?

2. Is there a significant difference between prioritizing and the importance of metadata elements of web content archiving from the perspective of library and information science experts?

2. Literature Review

Hakimzadeh's (2017) research entitled "Web Archiving" attempts to describe the concept of web archiving, how to select and compile, organize, and protect, legal issues, and levels of access to web resources and provides solutions and suggestions for web archiving. In his research, he pointed to the need for scientific organization of web archive resources and stressed that the method of storage must be correct and scientific to recover quickly, accurately, and correctly. Therefore, Hakimzadeh emphasizes the need to use appropriate metadata to describe the elements of information that makes it easier to manage and retrieve the content of web resources.

Noushinfard et al. (2012) in a study entitled "Study of collection policies of national web archives of selected countries: providing a proposed list for national web archives" intends to study the collection policy of national web archives of selected countries, a proposed list for archiving collections Provide a national web. This research was conducted under a library study method and reviewed sources and texts and content analysis of national archival collection policies. The findings indicated that the criteria for collecting resources in the study web archives were different. Finally, a list of 8 factors, 32 criteria, and 107 markers was presented for the collection in the national web archives.

Shadanpour et al. (2010) in a research project entitled "Feasibility study of creating a web archive in the National Library and Archives of the Islamic Republic of Iran" actually examined the first stage of the "Iran Web Archive" project and assessed the possibilities in the organization and the country in different dimensions.

Shafiee Alavijeh et al. (2009) in a study entitled "Study of metadata elements in web pages resulting from a search in public search engines" A sample consisting of 90 web pages, which were selected by searching in public search engines such as Google, Yahoo and . . . In terms of the presence of metadata elements (Dublin Core and Hypertext sign language metadata) in these

pages and the existence of a significant relationship between the presence of metadata elements and the type of search engine were studied. The findings indicate that there is very little presence of metadata elements in web pages and there is no significant relationship between the presence of metadata elements and the type of search engine.

Taheri (2008) in his research entitled "Comparison of the efficiency of Dublin Core metadata design and Marc 21 metadata format in organizing the World Wide Web information resources" studied and review the Dublin core metadata designs and Marc 21 format, and the role of these projects in organizing resources He conducted the research using an analytical method, and the effectiveness of these schemes in indexing electronic information was compared. The results of the study showed that the Marc 21 format is more suitable for storing, processing, and exchanging web information.

Portoro and Jirart (2019) studied the PROMISE project in a study entitled "Behind the Scenes of Web Archiving: The Metadata of Extracted Websites." In this research, first, technical information about web archiving and descriptive metadata used for web archiving is presented, as well as various methods of web archiving. In the PROMISE project, which develops a stable web archiving service for Belgium, the State Archives and the Royal Library are working to provide a common model for descriptive metadata using the OCLC template. This project uses the Heritrix browser to record selected web content and its content is stored in the WARC file format, which is the dominant file format in the field of web archiving.

Costa et al. (2017) conducted two surveys in 2010 and 2014 in a study entitled "The Evolution of Web Archives" to examine web archiving projects and their evolution. In this study, several plans and models were studied to identify the challenges and opportunities for advancement in this field. The models are also discussed to define the strategy, estimate the resources and provide solutions for technology research and development. The results showed that during the past years, there has been significant growth in the number of projects and countries hosting these projects, the volume of data, and the number of contents retained. Based on two surveys conducted in this study, it was found that web archiving projects are typically hosted by developed countries but can be extended worldwide.

Gomez et al. (2011) in a study entitled "Review of Web Archiving Projects" has studied the current models of Web archiving in the world and the challenges ahead in this area. The results show that the number of web archiving projects has increased significantly since 2003, especially in developed countries. In this study, 42 web archiving projects and their hosts were identified worldwide and indicators such as archived data volume, archived file format, or the number of people were statistically analyzed.

Kim and Lee (2007) in a study entitled "Development of metadata elements for selective web archiving" have reviewed and analyzed several web archiving projects. In this research, he describes the stages of the web archiving process in the record life cycle (selection, download, archiving, quality control) and examines and analyzes various web archiving projects and archiving methods.

3. Methodology

This research is of the applied type and conducted by a library or documentary method and a survey. In the first step, a preliminary study was conducted on the available resources in the field of web archiving. In this section, the world's leading models are selected. Metadata elements for archiving web resources were then extracted from the official site of each template and its adaptive table prepared from the Online Computer Library Center (OCLC), the Library of Congress (LOC), and the European Broadcasting Union Metadata Set (EBU Core).

To extract the primary metadata elements for web archiving, the structure of web archive records in the desired patterns was examined. Common elements in a web

content organization were extracted and refined in all three models, and to further integrate, the content adaptation of each metadata element was studied in terms of content description and importance of that element. Common elements for web content organizing were extracted and refined in all three models, and to further integrate, the content adaptation of each metadata element was studied in terms of content description and importance of that element. Since reference templates collect and organize the website content in different countries of the world, the metadata elements in the structure of describing and arranging their content can be modeled and localized, and it is possible to define a suitable metadata model to organize web knowledge content, provided on internal sites. Table 1 provides a comparative comparison of the metadata elements of the patterns being studied.

Then, through a questionnaire including the description of each metadata element, the importance of metadata elements extracted from the comparative table in describing web resources, was questioned by information science experts. Due to the importance of indigenous needs in describing web resources and the possibility of metadata elements other than the experiences of other countries, open tables were allocated to declare and record this importance.

The data collection tool in this research is a researcher-made questionnaire that in two parts, polls the metadata elements provided from the existing models and suggests the required native elements.

This study population included the Iranian library and information scientists in metadata and due to the limited number of these people, sampling was not performed. Also, to ensure the validity of the research tool, a researcher-made questionnaire was provided to five experts and after ensuring the validity, it was distributed among the study population. Totally, 40 responses were received and finally, the collected data were analyzed by SPSS software, version 8. The results will be described.

Table 1. *Matching and comparing elements of metadata frameworks*

Elements	OCLC	EBU Core	LOC
Publisher		*	****
Creator	*	*	****
Contributor	*	*	****
Collector	*		
Title	*	*	****
Alternative Title			
Collection Title			
Subject	*	*	****
Description	*	*	
Source of description	*		
Function Descriptor			
Notes		*	****
Source		*	
Language	*	*	****
Relation	*	*	
Coverage		*	
Extent	*		
Part of			****
Scopes			****
Date	*	*	****
Date Captured			
Date Metadata Modified			
Date Validate			
Type		*	
Format		*	****
Additional Metadata Formats			****
Online Format			****
Genre/Form	*		****
Audience			
Identifier		*	
URL	*	*	****
Locations			****
Repository			****
Rights	*	*	****
Access Condition			****
Availability			
Mandate			
Harvest File			
Collecting Method			
Collecting Tool			

4. Findings

According to Table 1, data science professionals ranked the metadata elements in the following order in terms of importance and the need for web archiving:

Table 2. *Prioritize the importance of archiving metadata in terms of information science professionals*

Metadata prioritization	Frequency of Response along with the level of importance very much Much medium Low seldom					Percentage of response in order of importance very much much medium Low seldom					median
Subject	0	1	1	8	24	0	2.9%	2.9%	23.5%	70.6%	5
Title	0	0	4	6	24	0	0	11.8%	17.6%	70.6%	5
Creator	2	1	3	5	23	5.9%	2.9%	8.8%	14.7%	67.6%	5
Format	0	1	4	12	17	0	2.9%	11.8%	35.3%	50%	5
Source	0	3	7	5	19	0	8.8%	20.6%	14.7%	55.9%	5
publisher	1	2	5	11	15	2.9%	5.9%	14.7%	32.4%	44.1%	5
Alternative Title	0	2	5	16	11	0	5.9%	14.7%	47.1%	32.1%	4
Genre/Form	0	2	5	14	13	5.9%	0	14.7%	41.2%	38.2%	4
Language	1	1	7	7	17	2.9%	2.9%	20.6%	20.6%	50%	5
Date	2	2	6	7	17	5.9%	5.9%	17.6%	20.6%	50%	5
Relation	1	2	5	9	16	2.9%	5.9%	14.7%	26.5%	47.1%	5
Access Condition	1	1	9	9	14	2.9%	2.9%	26.5%	26.5%	41.2%	5
Availability	1	3	7	7	16	2.9%	8.8%	20.6%	20.6%	47.1%	5
Type	2	1	8	8	15	5.9%	2.9%	23.5%	23.5%	44.1%	5
Rights	0	3	12	4	15	0	8.8%	35.3%	11.8%	44.1%	5
Online Format	0	3	9	10	12	0	8.8%	26.5%	29.4%	35.3%	5
Source	1	2	8	12	11	2.9%	5.9%	23.5%	35.3%	32.4%	4
Coverage	1	2	8	12	11	2.9%	5.9%	23.5%	35.3%	32.4%	4
Collector	1	1	6	15	10	2.9%	2.9%	17.6%	44.1%	29.4%	4
Source of description	0	3	12	8	11	0	8.8%	35.3%	23.5%	32.4%	3
Identifier	1	5	7	8	13	2.9%	14.7%	20.6%	23.5%	38.2	5
Notes	0	3	9	16	6	0	8.8%	26.5%	47.1%	17.6%	4
Extent	0	2	12	13	7	0	5.9%	35.3%	38.2%	20.6%	4
Part of	1	2	10	13	8	2.9%	5.9%	29.4%	38.2%	23.5%	4
Mandate	1	5	8	9	11	2.9%	14.7%	23.5%	26.5%	32.4%	4
Function Descriptor	0	2	15	9	8	0	5.9%	44.1	26.5%	23.5%	3
Locations	0	4	13	9	8	0	11.8%	38.2%	26.5%	23.5%	3
Scopes	0	4	11	14	5	0	11.8%	32.4%	41.2%	14.7%	4
Contributor	2	5	10	7	10	5.9%	14.7%	29.4%	20.6%	29.4%	3
Collecting Method	2	3	12	9	8	5.9%	8.8%	35.3%	26.5%	23.5%	3
Repository	2	4	12	6	10	5.9%	11.8%	35.3%	17.6%	29.4%	3
Date Validate	1	4	14	9	6	2.9%	11.8%	41.2%	26.5%	17.6%	3
Additional Metadata Formats	0	3	14	11	5	0	8.8%	41.2%	32.4%	14.7%	3
Date Captured	1	4	15	8	6	2.9%	11.8%	44.1%	23.5%	17.6%	3
Collecting Tool	2	5	11	9	7	5.9%	14.7%	32.4%	26.5%	20.6%	3
Date Metadata Modified	2	5	11	9	7	5.9%	14.7%	32.4%	26.5%	20.6%	3
Harvest File	2	5	11	9	7	5.9%	14.7%	32.4%	26.5%	20.6%	3

Six metadata elements with first - to sixth priority include elements of subject, title, creator, format, source, and publisher, respectively, more important than other metadata elements. Subject metadata elements with (94%) and titles with (88.2%) have been proposed as the most important metadata elements in web archiving.

Seven other metadata elements as alternatives title, form, language, date, relation, access condition, and availability are more important than other metadata elements, and they have been proposed in the seventh to 13th priorities.

Six other metadata elements, including elements of type, rights, online format, description, coverage, and collector, are considered more important than other elements and suggested in priorities fourteen to nineteen.

In the following, according to experts, seven other components including the Source of description

Identifier, note, extent, part of, the mandate is considered as more important than other elements and was placed in the priorities of the twenty - twenty-sixth.

The other twelve metadata elements that were prioritized in the twenty-seventh to thirty-eighth include location, scopes, contributor, collecting method, repository, additional metadata formats, date captured, collecting tool, date metadata, modified, and harvest file. This group can be considered the least important metadata element for web archiving experts.

Table 3*. Metadata Elements Rank*

Metadata	Mean Rank
Title	27.78
Alternative Title	22.07
Creator	26.37
Collector	21.23
Publisher	23.48
Contributor	16.45
Subject	28.20
Type	21.45
Genre/Form	21.20
Format	24.95
Online Format	19.58
Additional Metadata Formats	14.55
Extent	16.85
Scopes	16.28
Part of	17.85
Locations	16.77
Metadata	Mean Rank
Coverage	19.60
Source	23.60
Description	20.80
Source of description	17.88
Function Descriptor	17.20
Note	17.80
Language	24.27
Date	22.67
Date Captured	13.98
Date Validate	15.35
Date Metadata Modified	12.93
Identifier	19.27
Collecting Method	15.85
Collecting Tool	14.75
Harvest File	12.77
Repository	15.73
Rights	20.47
Availability	22.35
Access Condition	20.40
Mandate	17.57
Audience	18.48
Relation	22.22

Table 4. *Test statistics*

N	30
Chi-Square	183.759
df	37
Asymp. Sig.	0.000
a. Friedman Test	

The data in Table 4, extracted by the Friedman test, confirm the results of the present study, and the information science experts have chosen the metadata elements of the subject, title creator, format, and source as the most important metadata elements for Web archiving. To validate the research hypotheses, according to sig (000) obtained in Table 1, which is less than the

significance level of 0.05, it is clear that there is a significant difference in the metadata selection. Also, by confirming the second hypothesis, we conclude that there is a significant difference between the rank of metadata elements according to information science experts and in terms of information science experts, these metadata elements do not have the same priority.

5. Conclusion

The web archiving project requires cooperation and convergence of different pieces of content and technical and legal domains that involves gathering, organizing, storing, and servicing available resources on the domestic websites. Due to the extensive changes that have taken place in the methods and tools of scientific research and the consequences that have followed, it is necessary to archive web content systematically as soon as possible to preserve and integrate historical, cultural, and scientific documents. The lack of a standard and integrated metadata template leads to loss of managerial power, lack of proper services, and not access to valuable knowledge resources organizing the content of websites based on standard and integrated metadata templates will strengthen retrieval and increase user satisfaction and will prevent the loss of website information content. In this regard, this study measured the importance of metadata elements for web archiving from the expert perspective, and finally, the elements of subject, title, creator, and format were identified as more important. This study is consistent with Hakimzadeh's research that the use of metadata on the web leads to better management and retrieval of web content. It is also consistent with the research of Shafiee Alavijeh et al, on the lack or absence of appropriate metadata elements on the web. Taheri's research in 2008 achieved a more appropriate exchange with brand format elements, which is compatible with this study due to the proper use of metadata elements.

Puerto Rico and Giarat Research 2019, Costa et al. 2017, Gomez et al., 2011; Kim & Lee's 2007 research highlighted the need and importance of using metadata elements in web archiving and emphasized that using appropriate metadata to describe information elements facilitates the management and retrieval of web resource content. Whereas one of the main requirements for optimal Web resource management is the creation of an appropriate and integrated metadata template in this area; the provision of essential elements to organize this type of resource can lead to increasing the access and the retrieving resource accuracy and improve interoperability and achieve integration

References

Ahmadi, N. S., Ojagh, F. S., Namvar, Z., Rahimnejhad, F., Zardari, S., Peymani, H., & Farizadeh, Z. (2009). *Metadata standards. Tehran*, Tehran, Tehran: Chapar. Retrieved 2022. (in Persian)

About the Internet Archive. (2021). Retrieved 5 11, 2022, from Internet Archive: https://archive.org/about/

Costa, M., Gomes, D., & Silva, M. J. (2016, 5 9). The evolution of web archiving. *International Journal on Digital Libraries*, 18(3), 191-205. doi:DOI 10.1007/s00799-016-0171-9

Doroodi, F. (2009). Web Archiving Book Review. Ganjineh Asnad, 2(74), 115-122. Retrieved 5 10, 2022, from http://ensani.ir/file/download/article/20120325113020-1011-102.pdf. (in Persian)

Di Pretoro, E., & Geeraert, F. (2019). Behind the scenes of web archiving: metadata of harvested websites. *Archives et Bibliothèques de Belgique-Archief-en Bibliotheekwezen in België*, 106, 63-74.

Emmanuel Di Pretoro, F. G. (2019, 5 9). Behind the Scenes of Web Archiving: Metadata of Harvested. 63-74. Retrieved 5 11, 2022, from https://hal.archives-ouvertes.fr/hal-02124714

Finnemann, N. O. (2019). Web archive. Knowledge Organization 49(1), 47-70. Retrieved 4 9, 2022, from ISKO: https://www.isko.org/cyclo/web_archive#7

Gomes, D., Miranda, J., & Costa, M. (2011). A Survey on Web Archiving Initiatives. Gradmann, S., Borri, F., Meghini, C.,

Schuldt, H. (eds) Research and Advanced Technology for Digital Libraries TPDL. doi:https://doi.org/10.1007/978-3-642-24469-8_41

Gomes, D., & Costa, M. (2014, 4). The Importance of Web Archives for Humanities. *International Journal of Humanities and Arts Computing*, 8(1), 106-123. doi:10.3366/ijhac.2014.0122

Hassanzadeh, Mohammad (2017) *Information storage and retrieval*. Tehran: Samt (in Persian)

Hakimzadeh, R. (2017). Web Archiving. National Archives Quarterly, 48-67. Retrieved 5 10, 2022, from magiran.com/p1959195. (in Persian)

IIPC WAC 2022: #WHYWEBARCHIVING: PRESERVING INTERNET CONTENT FOR RESEARCH USE. (2022, 6 21). Library of Congress. Retrieved 6 24, 2022, from https://www.youtube.com/watch?v=B26KMgI8iY8

Kim, J., & Lee, H. (2007). Development of Metadata Elements for Intensive Web Archiving. *Journal of Information Management*, 24(2), 143-160. doi:https://doi.org/10.3743/KOSIM.2007.24.2.143

Library of Congress. (2022). IIPC Web Archiving conference. Retrieved 6 24, 2022, from International Internet Preservation Consortium: https://netpreserve.org/ga2022/virtual-visit-to-the-library-of-congress/

Merriam Webster. (n.d.). Metadata. Retrieved 5 10, 2022, from Merriam Webster: https://www.merriam-webster.com/dictionary/metadata

Metadata. (2022). Retrieved 5 10, 2022, from EBU Technology & Innovation: https://tech.ebu.ch/metadata

New World Encyclopedia. (2018, 12 20). Online Computer Library Center. (New World Encyclopedia) Retrieved 5 10, 2022, from New World Encyclopedia: https://www.newworldencyclopedia.org/p/index.php?title=Online_Computer_Library_Center&oldid=1016746

Nooshin Fard, F., and Rezaei Sharifabadi, S., and Eshghabadi, M. (1391). Review of National Web Archive Collection Policies in Selected Countries: Provide Proposed Lists for National Web Archives. Document Treasure, 22 (1 (85 in a row)), 80-106. https://www.sid.ir/fa/journal/ViewPaper.aspx?id=173621

Niu, J. (2012). *An Overview of Web Archiving. School of Information Faculty Publications*, 308. Retrieved 5 10, 2022, from http://scholarcommons.usf.edu/si_facpub/308?utm_source=scholarcommons.usf.edu%2Fsi_facpub%2F308&utm_medium=PDF&utm_campaign=PDFCoverPages

Preservation Services. (2022). Retrieved 4 9, 2022, from Harvard Library: https://preservation.library.harvard.edu/web-archiving

Promoting the interests of public service media. (2022). Retrieved 5 10, 2022, from EBU: https://www.ebu.ch/about

Program Web Archiving. (2022). Retrieved 6 24, 2022, from Library of Congress: https://www.loc.gov/programs/web-archiving/about-this-program/

Ravoof, S. (2022). How to Archive a Website: Our Mammoth Guide to Saving Your Site? Retrieved 5 9, 2022, from Kinsta: https://kinsta.com/blog/archive-a-website/

Shafiei Alavicheh, S., Ghaebi, A., & Rezaee sharifabadi, S. (2009). Examine the metadata elements contained in web pages resulting from searches on public search engines. *Information Processing and Management (Information Science and Technology)*, 0. Retrieved 5 11, 2022, from https://www.sid.ir/fa/journal/ViewPaper.aspx?id=102949. (in Persian)

Shadanpoor, F. (2010). Iran Web Archive (Phase 1): Feasibility Study of Creating a Web Archive in the National Archives and Library Organization. Research project, Documents Organization and National Library of the Islamic Republic of Iran, National Library of the Islamic Republic of Iran. (in Persian)

Taheri, M. (2008). Comparing the efficiency of the Dublin Core Metadata Framework and the Marc 21 metadata format in

organizing the World Wide Web information resources. Library and Information Science, 43(3), 139-158. Retrieved 5 11, 2022, from http://lis.aqr-libjournal.ir/article_43734.html. (in Persian)

Web Archiving. (2022). Retrieved 4 9, 2022, from IIPC International Internet Preservation Consortium: https://netpreserve.org/web-archiving/about-archiving/

Web Archiving. (2022). Retrieved 4 9, 2022, from IIPC: https://us02web.zoom.us/webinar/register/WN_CB749fYYQ4OWG7R-V-v1XQ

Wiedeman, G. (2019). Describing Web Archives: A Computer Assisted Approach. Journal of Contemporary Archival Studies, 6. Retrieved 5 10, 2022, from https://elischolar.library.yale.edu/jcas/vol6/iss1/31?utm_source=elischolar.library.yale.edu%2Fjcas%2Fvol6%2Fiss1%2F31&utm_medium=PDF&utm_campaign=PDFCoverPages