Emerging Digital Heritage: Constructing the Selection Framework and Preservation Strategies for Born-Digital Cultural Heritage

Tianjiao Qi [1], Yinuo Zhou [2], Zizhuo Liu [3]

[1] School of Information Resource Management, Renmin university of China, China, qtjjoy@163.com
[2] School of Information Resource Management, Renmin university of China, China, zyn21@ruc.edu.cn
[3] School of Information Resource Management, Renmin university of China, China, 15153817939@163.com

Keywords: Born-digital Cultural Heritage, Cultural Heritage Conservation, Lifecycle Model, Digital Stewardship

Abstract

This paper explores the selection and preservation pathways for born-digital cultural heritage. The research first defines the conceptual connotation and common types of born-digital cultural heritage, proposing that it possesses core characteristics including multidimensional value, finite and closed nature, technology dependency, separability of information from carrier, multimedia integration, organic interconnectedness, and public interactivity. Building upon this foundation, a selection evaluation framework comprising 4 dimensions, 12 primary indicators, and 23 secondary indicators is constructed through the grounded theory methodology. Finally, based on this framework, tiered preservation strategies for born-digital cultural heritage are proposed.

1. Introduction

Since the advent of the information age, vast quantities of cultural resources and knowledge are digitally created, shared, and utilized. These born-digital resources possess significant value, with some capturing pivotal aspects of civilization's development (waag). As humanity's shared cultural assets, they constitute born-digital cultural heritage. The 2023 inclusion of "De Digitale Stad" (1994-2001) in UNESCO's Memory of the World Register, marking UNESCO's first recognition of such heritage(Jaillant, 2019) underscores its status. Despite the abundance of born-digital resources, few merit long-term preservation. Identifying content with potential heritage value requires subsequent archiving to become authenticated born-digital cultural heritage. Born-digital resources are inherently fragile(Lyman) dependent on ephemeral technologies and highly susceptible to loss through technological obsolescence, exemplified by the 44-day average webpage lifecycleand digital art.

Scholarship reflects growing interest. Theoretical work addresses preservation challenges and solutions, while applied research explores resource utilization, sharing, and related ethical concerns. Research demonstrates an interdisciplinary trend, predominantly led by information science, with contributions from cultural heritage and arts. However, three critical limitations persist. Firstly, non-information disciplines predominantly offer theoretical insights without practical solutions. Secondly, Frequent confusion among born-digital resources, archives and cultural heritage has led to indiscriminate preservation targeting. Thirdly, current discourse on preserving born-digital heritage remains predominantly theoretical, lacking practical approaches and institutional frameworks.

This study addresses these gaps. We first establish clear definitions of born-digital cultural heritage (concept, types, characteristics). Subsequently, relevant policies and cases are analyzed using grounded theory coding to develop a systematic selection and appraisal framework. Finally, based on this framework, preservation strategies for born-digital resources are proposed. This work establishes actionable pathways for born-digital cultural heritage selection and preservation.

2. The Concept of Born-Digital Cultural Heritage

2.1 The Concept and Types of Born-Digital Cultural Heritage

Digital cultural heritage comprises unique resources embodying human knowledge and expression, existing in two primary forms: resources created natively in digital format and those digitized from analog sources. Its content encompasses resources related to culture, education, science, and administration, as well as information from technical, legal, medical, and other domains. Based on form, it is categorized into born-Digital Cultural Heritage (resources created exclusively in digital form) and Digitized Cultural Heritage (resources converted into digital format from analog originals).
Current research and policy exhibit broad consensus on the core concept of born digital cultural heritage, with minor definitional variations. UNESCO's Charter on the Preservation of the Digital Heritage defines it as resources existing solely in digital form with no other format. The 2021 EU Proposal for a Council Recommendation on a common European data space for cultural heritage characterizes it as assets created digitally without analogous physical equivalents, such as digital art, virtual museums, or cultural content generated outside heritage institutions (e.g., social media or gaming industries). Scholars further define it as born-digital materials possessing enduring value that warrant preservation for future generations .

Regarding types, UNESCO's 2016 Guidelines for the Long-Term Preservation of Digital Heritage Selection identified born digital heritage as new expressions within the digital environment, including web pages, social media sites, private research databases, and online gaming environment. Revised in 2021, these guidelines expanded the types to include web pages, social media sites, private research databases, online gaming environments, VR/AR materials, digital art, and web archives, with appendices detailing social media, software source code, research data, and artificial intelligence as specific forms. Other projects classify types such as digitally photographed images, born-digital files, digital art, websites, applications, and games. Practically, the resource types currently most frequently

recognized as cultural heritage include web archives, video games and digital art.
Consequently, this paper defines Born-Digital Cultural Heritage as unique cultural resources imbued with human knowledge and expression, created natively in digital form with enduring preservation value, encompassing principal types such as web archives, video games, social media, digital art, research databases, VR/AR materials, software source code, and artificial intelligence assets.

2.2 The Characteristics of Born-Digital Cultural Heritage

As a novel heritage typology in the digital era, born-digital cultural heritage inherits core characteristics of traditional cultural heritage while developing unique attributes due to its formative methods and carrier modalities.

(1) Multidimensional Value
Born-digital cultural heritage stores high-density information with a complex value system. Aligning with the World Heritage Convention's requirement of "outstanding universal value" , it exhibits longitudinal significance through enduring intergenerational cultural transmission and support for historical research. Horizontally, it encompasses historical, artistic, scientific, socio-cultural, and economic values. Beyond the core "three values" of cultural heritage, it manifests cultural diversity and social cohesion, while its developmental potential enables economic value conversion (e.g., cultural tourism benefits).

(2) Finite and Closed Nature
Digital resources (websites, social media, databases) initially feature high dynamism with frequent updates. Their transition to cultural heritage necessitates closure — content stabilization and cessation of updates, transforming them into non-current resources. This finite, closed, and well-defined state constitutes the fundamental distinction from original digital resources. Projects like the AIDS Education Global Information System and Global Free Software were excluded from the Memory of the World Register due to ongoing updates, whereas the discontinued DDS project (2001) met closure criteria for inclusion.

(3) Technological Dependence and Informational Volatility
Born-digital heritage relies inherently on digital environments for creation, operation, and interpretation. Technological obsolescence risks permanent inaccessibility and memory fragmentation. While technology itself may hold value, format migration introduces volatility. New software and standards may alter or erase original information, compromising authenticity. High dynamism in digital resources necessitates preserving critical historical versions to ensure traceability.

(5) Contextual Integrity
Value resides not in isolated items but in composite files with intrinsic relationships. Preserving contextual background is essential to assess heritage as an organic whole. For web-based heritage, five hierarchical levels exist: web elements, single pages, websites, link collections, domain-level aggregate. (the minimal unit for historical validation).

(6) Public Interactivity
Digital carriers inherently connect born-digital heritage with public engagement. Common forms (websites, social media, games, digital art) feature real-time interactivity, enabling dynamic user participation. Heritage value emerges through co-evolution between public interaction and content, where user-generated contributions shape cultural diversity. Consequently, public engagement metrics become vital value indicators.

3. Grounded Coding Process

Current scholarly consensus on the connotation and value of born-digital cultural heritage is increasingly solidifying, yet systematic methodologies and actionable guidelines for selecting and preserving such heritage from extant digital resources remain critically underdeveloped. To address this gap, this paper employs grounded coding analysis of international cultural heritage conventions, inscribed born-digital heritage projects (e.g., UNESCO Memory of the World entries), and national or region preservation initiatives, aiming to establish a selection and appraisal framework.

3.1 Data Sources

To construct the selection and appraisal framework, this study will select relevant policies and cases as foundational references, conduct grounded theory coding, and distill key metrics for born-digital cultural heritage selection. To ensure the scientific validity of these metrics, high-impact texts (partially listed in Table 1) will be prioritized. The three documents A1-A3, officially issued by UNESCO, respectively target Documentary Heritage, Cultural and Natural Heritage, and Digital Heritage, all constituting programmatic document in the field of cultural heritage conservation. Born-digital cultural heritage lies at the intersection of documentary and digital heritage, closely related to policies A1 and A3. As a specialized heritage category, policy A2 also offers references for its selection. Cases B1-B9 are representative born-digital cultural heritage projects. Open, axial, and selective coding will be applied to these policy/case materials to establish a three-tiered selection framework for born-digital cultural heritage selection.

Name	Authority	Number
General guidelines of the MoW	UNESCO's programmatic document	A1
Convention Concerning the Protection of the World Cultural and Natural Heritage	UNESCO's programmatic document	A2
Guidelines for the Selection of Digital Heritage for Long-Term Preservation （2021）	UNESCO's programmatic document	A3
De Digitale Stad	The first born-digital heritage recognized by UNESCO	B1
PANDORA	Inscribed on the Australian National MoW Register	B2
Netarkivet	The Danish online archive project	B3
EU Web archive	The EU online archive project	B4
UK web archive	The UK online archive project	B5
Rhizome Artbase	Including the earliest Internet artworks	B6
AusStage	Inscribed on the Australian National Memory of the World Register	B7

AustLit	Inscribed on the Australian National Memory of the World Register	B8
Internet archive	The largest and most influential web archive project	B9

Table 1. Data Sources

3.2 Coding Process

As shown in Table 2, the first step is to conduct open coding of the policy or project content. Due to the large amount of content, this article only lists some of the initial concepts and their original statements. According to the principles and steps of open coding, the initial concepts are categorized to form 69 initial concepts and 27 sub-categories. The results of open coding are presented in Table 3. Then, the main axis coding is carried out to sort out the internal logic and hierarchical structure of these 27 sub-categories, determine semantic and structural relationships, and summarize to obtain 11 main categories (Table 3). Finally, selective coding is conducted on the main categories formed after the main axis coding to analyze the dimensions and adjust the hierarchical relationships. Eventually, 4 dimensions of "Significance", "Credibility", "Preservation Status", and "Regulatory Ethics" are extracted, as shown in Table 4.

Original Statements	Initial Concepts
Has significant historical, cultural, or artistic value (A3)	Historical, cultural, or artistic value
Has academic importance or research value (B5)	Research value
Provides unique testimony to a vanished civilization or cultural tradition (A2)	Unique testimony
Whether it has unique characteristics (A1)	Unique characteristics
Whether there is sufficient metadata for accessing or preserving digital heritage (A3)	Sufficient metadata
Files saved in sustainable formats (B1)	Sustainable file format
Whether this institution is the best place to preserve and access the resource (A3)	Institutional capacity
Whether the resource is rare or has numerous copies (A3)	Rare resource
Made by Danes, in Danish, or for a Danish audience (B3)	Denmark-related

Table 2. Examples of the Open Coding Process

Sub-category	Initial Concepts
Core Value	Historical or artistic value, artistic or technical value, historical imagery, artistic influence, research value
Cultural Value	Cultural significance, cultural value
Social Value	Social cohesion, social value, diversity of life, political influence, government websites
Uniqueness	Unique testimony, unique characteristics
Innovativeness	Novel digital ideas, critical content, innovative usage, technological experimentation, new technologies
Formation Status	Authoritative publications, provenance of works
Formation Time	Oldest, first generation
Historical Context	Historically significant themes, significant stages of human development
Page View	Active status
Search Index	Internet search volume
Public Engagement	Public attention, online discussions
Geographic Scope	Regional influence, online reputation, international contribution
Value Timeliness	Long-term value, potential utilization

Table 3. Results of Open Coding (Partial Results)

Main Category	Sub-category	Connotation
Content Significance	Core Value	The three core values of heritage: historical, artistic, and scientific
	Cultural Value	Beneficial role in cultural development
	Social Value	Social activities or spiritual content carried
	Innovativeness	Whether the design concept, theme, technology, or format is pioneering
Context Value	Formation Status	Social influence of the individual or organization creating the heritage
	Formation Time	Time when the heritage was initially created

Table 4. Results of Axial Coding

Dimension	Secondary metrics	Tertiary metrics
Significance	Content Significance	Core Value
		Cultural Value
		Social Value
		Uniqueness
		Innovativeness

	Context Value	Formation Status
		Formation Time
		Historical Context
	Carrier Influence	Page View
		Search Index
		Public Engagement
		Geographic Scope
	Value Timeliness	
Credibility	Integrity	Format Integrity
		Content Integrity
		Metadata Status
	Authenticity	Authenticity of Management
		Authenticity of Provenance
	Reusability	
Preservation Status	Institutional Capacity	Technical Maturity
		Institutional Resources
	Rarity	Version Quantity
		Version Originality
	Risk of Loss	Social Conditions
		Backup Status
Regulatory & Ethical	Privacy and Security	
	Institutional Mission Alignment	

Table 5. Results of Selective Coding

3.3 Analysis and Interpretation of the Framework

Through coding the aforementioned policies and projects, this study constructed a selection and appraisal framework consisting of 4 dimensions, 12 secondary metrics, and 23 tertiary indicators. The framework is divided into four dimensions: Significance, Credibility, Preservation Status, and Regulatory and Ethical. The Significance dimension includes Content Significance, Context Value, Carrier Influence, and Value Timeliness. Credibility can be evaluated based on Integrity, Authenticity, and Reusability. Preservation Status comprises Rarity and Risk of Loss. Regulatory and Ethical includes Privacy and Security, and Institutional Mission Alignment.

3.3.1 Significance

The significance of digital resources is key to assessing their eligibility as digital cultural heritage. Higher significance implies greater utilization value and a higher likelihood of transformation from born-digital resources to cultural heritage. Content Significance can be evaluated from five aspects: Core Value, Cultural Value, Social Value, Uniqueness, and Innovativeness. Cultural heritage values are diverse, encompassing core, academic, social, and other values. To be selected as cultural heritage, born-digital resources must possess historical, artistic, and scientific value. Born-digital heritage can serve as objective evidence for historical events and provide research materials for scholars, thus possessing scientific value. Additionally, born-digital heritage carries the spiritual culture of specific periods, relating to the emotional memory of certain communities. Some digital heritage may be linked to socio-political activities thus having social value. Cultural and social values are derivative values of born-digital cultural heritage. Uniqueness emphasizes whether the content, form, or other aspects of digital resources have irreplaceable and non-replicable characteristics, enabling them to serve as "witnesses" of historical eras. Innovativeness evaluates whether the design concepts, themes, technologies, formats, etc., of the resources have pioneered transformative innovations in knowledge production, emphasizing temporal primacy.

Context Value assesses the environmental information surrounding the creation of born-digital resources. Contextual information includes the social environment of resource creation, the activities generating the resources, and other related information Evaluating Context Value involves three metrics: Formation Status, Formation Time, and Historical Context. From an archive appraisal perspective, the higher the administrative or social status of the resource creator, the higher the resource value. Regarding formation time, since cultural heritage carries cultural attributes, older resources represent early cultural characteristics, possess more ancient cultural traits, and embody deeper cultural value, thus having a higher likelihood of becoming cultural heritage. Specific historical contexts shape cultural heritage, determine the interpretation of cultural connotations, form a dialogic relationship, and influence the value of cultural heritage.

Carrier Influence measures the external dissemination impact of born-digital resources, consisting of four tertiary indicators: Page View, Search Index, Public Engagement, and Geographic Scope. Due to the inseparability of born-digital resource content and carrier, carrier influence also reflects the value of cultural heritage. From an informetrics perspective, the page view and search index of born-digital resources in search engines reflect real-time attention, providing objective evidence of impact. For example, existing tools like Google Trends and Baidu Index can demonstrate the popularity of keywords in cyberspace. Public Engagement includes the frequency of news coverage related to the resources and the popularity of user comments, shares, and derivative creations.

This indicator empowers public participation in cultural heritage decision-making, democratizing heritage, and balances the tertiary indicators "Formation Status," countering elite cultural narratives. Geographic Scope reflects the regional influence of resource dissemination.

Furthermore, the Value Timeliness indicator helps filter resources with genuine long-term preservation value, excluding those with only short-term reference or utilization value. Cultural genes that withstand historical testing become the roots of civilizational inheritance. Over time, born-digital cultural heritage will acquire new meanings through historical changes, while resources with only short-term instrumental value lack sustainability and will gradually depreciate.

3.3.2 Credibility

Born-digital cultural heritage is a trustworthy electronic record preserved for the future. The Credibility dimension verifies the integrity, authenticity, and reusability of digital resources, ensuring that the resource content is genuine, the evidence chain is complete, and it is free from forgery or tampering, thereby guaranteeing the authority and historicity of cultural heritage Integrity can be assessed through Content Integrity and Metadata Status. Resources related to born-digital cultural heritage are diverse, including digital information and analog information. For example, the DDS project resources include not only early forum websites but also oral histories of early users and physical objects like disks. Good metadata quality preserves the context of resource creation, links resources to socio-historical contexts as an organic whole, and verifies the credibility of cultural heritage.

Authenticity occupies a core niche in the cultural ecosystem, determining the uniqueness of a cultural heritage. It is the core value of maintaining human cultural diversity and heritage conservation. Protecting the authenticity of emic culture should be prioritized in heritage conservation. Credibility assessment should be conducted throughout the entire lifecycle. Born-digital cultural heritage should possess authenticity from its inception, ensuring that the resource source and management process are reliable, free from forgery or tampering, and consistent with the original state. The Reusability indicator evaluates whether the resources support further activation and utilization, facilitating the construction of future-oriented smart data for cultural heritage and linking multi-carrier, diverse cultural connotations.

3.3.3 Preservation Status

The Preservation Status dimension consists of three secondary metrics: Institutional Capacity, Rarity, and Risk of Loss. Institutional Capacity is primarily evaluated based on Technical Maturity and Institutional Resources. Strictly speaking, this indicator is not for selecting born-digital cultural heritage itself but for identifying institutions capable of properly preserving cultural heritage, ensuring timely and appropriate protection of potential born-digital cultural heritage. Institutions with high maturity and abundant resources are more suitable for preserving cultural heritage resources; otherwise, the resources are better transferred to other institutions. Rarity includes two indicators: Version Quantity and Version Originality. Resources with no copies or few copies, as well as original versions, have higher rarity and correspondingly higher value. For example, the PANDORA project excludes compilations of non-original materials and easily accessible information; the UK Web Archive prioritizes preserving content published exclusively online. Risk of Loss includes two indicators: Social Conditions and Backup Status. Social Conditions primarily consider security, economic issues, etc., in the heritage's region, such as whether it is in a war zone or a low-income country. Cultural heritage in regions with poor social conditions and few backups should also be prioritized for preservation.

3.3.4 Regulatory and Ethical

The Regulatory and Ethical dimension includes two indicators: Privacy and Security, and Institutional Goal Alignment. Born-digital cultural heritage involves personal information of online users, especially research databases and social media resources, posing risks of sensitive information leakage. The compliance boundaries for preserving and using such data resources require deeper exploration. As cultural institutions, when selecting heritage, it is necessary to protect the public's right to know, solicit public opinions on disclosing personal data, and prioritize preserving publicly available and transparent content. On the other hand, Institutional Goal Alignment aims to select resources most suitable for preservation by the institution. Born-digital cultural heritage has the separability of information and carrier, scattered across different institutions. Cultural institutions need to define the scope of their cultural heritage resource protection. Resources of equal value related to their own country or nation should be prioritized. For example, national libraries typically collect cultural heritage within their national domain, while universities tend to collect web archives serving researchers or educational needs. Regional and organizational web archives are often collected for legal or record-keeping purposes, focusing on specific documents or websites

4. Classified Preservation Strategy for Born-Digital Resources Based on the Selection Framework

The recognition of cultural heritage requires historical validation, and its value assessment is diachronic. Born-digital cultural heritage evolves from born-digital resources, and the prominence of its cultural value gradually emerges through historical changes. However, the dynamic volatility of born-digital resources requires relevant institutions to have forward-looking preservation awareness, selecting resources with the potential to become cultural heritage for long-term preservation to prevent damage or loss of potential cultural heritage. Preserving born-digital resources requires multi-institutional collaboration, forming a joint effort from local to national and even international levels. Local cultural institutions mainly refer to local GLAM institutions responsible for specific resource management. National management agencies are national administrative departments, including central authorities like national archives administrations and ministries like national ministries of culture. International organizations are NGOs related to cultural heritage, with UNESCO as a typical representative. National management agencies and international organizations are primarily responsible for formulating policy standards, not specific preservation management. Based on the four dimensions of the selection and evaluation framework, this paper proposes a graded preservation strategy for born-digital resources at three levels — local cultural institutions, national management agencies, and international organizations — providing concrete implementation pathways for cultural heritage preservation.

4.1 Classified Preservation Strategy Based on Significance

Due to limited institutional resources, when preserving born-digital resources, they should be graded and classified based on Content Significance, Contextual Value, Carrier Influence, and Value Durability indicators. According to the Significance indicators, digital resources can be broadly divided into three levels: Priority Preservation Level, General Preservation Level, and Temporary Preservation Level. Resources with high content significance, high contextual value, high carrier influence, high value durability, and significant importance to the institution require priority preservation. The General Preservation Level consists of resources showing high significance currently but potentially shorter value durability. The Temporary Preservation Level includes resources with low significance, whose value will gradually disappear after a period of preservation. Institutions should expand the preservation scope as much as possible to avoid misjudging the value of cultural heritage.

Cultural heritage represents the cultural diversity of a nation or country. In Significance grading assessment, enhancing cultural

diversity requires balancing considerations of Creator Status and Public Engagement. Under the trend of decentralization, human-centered cultural narratives are increasingly prominent, and enhancing interaction between users and cultural institutions has become an important goal of cultural heritage preservation. Born-digital cultural heritage inherently possesses public interactivity, giving it a natural advantage in increasing public participation. Appropriately increasing the weight of public engagement in resource preservation and selection aligns with the intrinsic requirements of resource characteristics and is a beneficial way to build public-participatory heritage.

4.2 Classified Preservation Strategy Based on Credibility

The core of ensuring the authenticity of born-digital resources lies in guaranteeing integrity, authenticity, and data reusability. Ensuring the integrity of born-digital resources means that all materials closely related to the resources should be preserved to avoid omissions caused by subjective inference. Besides the digital resource content itself, digital management records, metadata, oral histories of relevant personnel, and associated physical materials related to the heritage need to be preserved intact, constructing an archival fonds in the digital world. For example, the DDS project preserved not only the DDS platform itself but also audiovisual materials, photos, and other digital heritage; artifact registration cards, metadata, and other heritage digital information; as well as physical materials like disk arrays and servers, serving as historical materials for web archaeology. Among these, Metadata Status plays a vital role in ensuring resource integrity, revealing the historical background or geographical context of resource creation, establishing connections between cultural heritages, and forming a linked cultural knowledge network, promoting open sharing of heritage resources. Simultaneously, standardized and distinctive metadata schemes should be developed based on the type and characteristics of cultural heritage to better describe resources. For instance, some scholars have proposed metadata structures for VR/AR context-aware services.

Regarding credibility assurance, full lifecycle management of born-digital resources should be implemented, with comprehensive control over resource creation, management, archiving, transfer, preservation, and access. Front-end control should be applied from the beginning of resource creation. During the resource acquisition stage, the content and attributes of the responsible parties should be accurately recorded and verified, and resource creator identity authentication should be conducted. After resource creation, any modifications or deletions must be fully recorded and periodically monitored. Technically, common methods to ensure the authenticity of digital resources include blockchain, digital signatures, metadata collection, identity authentication, firewalls, digital digest comparison, etc. The integrity and credibility of born-digital resources also ensure the data reusability of born-digital cultural heritage.

4.3 Classified Preservation Strategy Based on Preservation Status

To enable sustainable preservation of born-digital resources and minimize or prevent their loss or damage, strategies include cross-institutional preservation, developing reasonable backup strategies, and prioritizing high-risk resources. During born-digital resource preservation, if the original preserving institution lacks resources or technical capacity, misaligned with the resource's significance and preservation urgency, preservation work can be delegated to external institutions. Additionally, some institutions are establishing unified cultural resource preservation network platforms to promote cultural sharing, alleviating the burden on individual institutions to some extent, such as Europeana in Europe and NDIIPP in the US. On the other hand, preserving institutions need to develop backup strategies suitable for resource conditions, specifying requirements such as backup quantity, form, and frequency. Common backup methods include online backup, offline backup, and off-site backup, which should be used comprehensively to build a multi-layered backup system. Furthermore, resources with high rarity and high risk of loss should be prioritized for preservation, especially unique copies and resources in high-risk regions. For example, against the backdrop of the Russia-Ukraine war, the Saving Ukrainian Cultural Heritage Online (SUCHO) project urgently preserved born-digital resources like websites, online galleries, and meme walls of Ukrainian heritage institutions, effectively addressing the crisis of digital cultural heritage loss.

4.4 Classified Preservation Strategy Based on Regulatory and Ethical

Born-digital cultural heritage possesses public interactivity and carries a vast amount of public information. Therefore, institutions should incorporate considerations of personal privacy data when preserving resources, ensuring compliance with privacy regulations. For example, the UK Web Archive explicitly states that it does not collect online audio, video, private networks, emails, or social media records. When selecting preservation scope, DDS surveyed users for consent. On the other hand, for individual countries or regions, in cultural heritage preservation practice, institutions should first select resources aligned with their organizational mission to construct their national cultural memory. Resources of high significance but low alignment can be preserved through cross-institutional collaboration.

5. Conclusion

Born-digital cultural heritage serves as the physical carrier of cultural genes in the digital era. Its fragility and transience necessitate the selection and preservation of high-value resources from the vast digital landscape as a prerequisite for civilizational continuity. This paper delineates the concept and common types of born-digital cultural heritage, highlighting its core characteristics: multidimensional value, finiteness and closure, technological dependency, information-carrier separability, multimedia integration, organic interrelatedness, and public interactivity. Building on this foundation, the study selects relevant policies and cases, integrating archival science and informetrics theories to construct a selection and evaluation framework for born-digital cultural heritage. This framework comprehensively assesses resources across four dimensions — Importance, Authenticity, Preservation Status, and Regulatory Ethics — providing a clear pathway for nations to select born-digital resources.

Cultural heritage value assessment is diachronic, compounded by the fragile and volatile nature of born-digital resources. Preservation institutions should utilize this evaluation framework to identify resources with potential heritage value. Strategies include graded importance assessment, ensuring resource authenticity, achieving sustainable preservation, and establishing standardized protocols to guarantee long-term preservation and knowledge innovation.

Furthermore, comparing international and national policies and cases reveals differing perspectives: international organizations emphasize flexibility and inclusivity, while nations focus on cultural specificity and practicality. The proposed framework leans toward the former, offering adaptable extensibility and universality. Countries can incorporate localized indicators to preserve national-level cultural heritage. Future research will validate this framework through practical application and refinement.

References

Albaqawy, G.A., Alnaim, M.M., Bay, M.A., Touahmia, M., 2023. Assessment of Saudi Arabia's Classification and Selection Criteria for Heritage Sites: A Case Study of Barzan Heritage Area in Hail City. Sustainability, 15(2), 1015. doi.org/10.3390/su15021015

Alberts, G., Marc, W., Jansma, R., 2017. Archaeology of the Amsterdam digital city; why digital data are dynamic and should be treated accordingly. Internet Histories, 1(1-2), 146-159. doi.org/10.1080/24701475.2017.1309852

Baker, M.J., 2021. Foundations of the Age-Area Hypothesis. Humanities & Social Sciences Communications, 8(1), 311. doi.org/10.1057/s41599-021-00991-8

Barbuti, N., 2020. Creating Digital Cultural Heritage with Open Data: From FAIR to FAIR 5 principles. Communications in Computer and Information Science, 16th Italian Research Conference on Digital Libraries (IRCDL), Univ Bari Aldo Moro, Dept Comp Sci, Bari, ITALY.

Baron, R., 2016. Public folklore dialogism and critical heritage studies. International Journal of Heritage Studies, 22(8), 588-606. doi.org/10.1080/13527258.2016.1150320

China, I., 2015. Principles for the Conservation of Heritage Sites in China. Wen wu chu ban she.

Commision, E., 2021. Commission proposes a common European data space for cultural heritage. digital-strategy.ec.europa.eu/en/news/commission-proposes-common-european-data-space-cultural-heritage (2021-11-10)

Davis, S.B., Vane, O., Kräutli, F., 2021. Can I believe what I see? Data visualization and trust in the humanities. Interdisciplinary Science Reviews, 46(4), 522-546. doi.org/10.1080/03080188.2021.1872874

Filiciak, M., 2020. Playful Machines and Heritage: How to Prepare Future Cultural Histories?. Arts, 9(3), 82. doi.org/10.3390/arts9030082

Hawkins, A., 2022. Archives, linked data and the digital humanities: increasing access to digitised and born-digital archives via the semantic web. Archival Science, 22(3), 319-344. doi.org/10.1007/s10502-021-09381-0

IIPC. About archiving. netpreserve.org/web-archiving/about-archiving/

Jaillant, L., 2019. After the digital revolution: working with emails and born-digital records in literary and publishers ’ archives. Archives and Manuscripts, 47(3), 285-304. doi.org/10.1080/01576895.2019.1640555

Lee, J., Kim, J., Ahn, J., Woo, W., 2019. Remote Diagnosis of Architectural Heritage Based on 5W1H Model-Based Metadata in Virtual Reality. ISPRS International Journal of Geo-Information, 8(8), 339. doi.org/2220-9964/8/8/339

Lyman, P. Archiving the World Wide Web. clir.org/pubs/reports/pub106/web/#1

Ruan, J., McDonough, J.P., 2009. Preserving born-digital cultural heritage in virtual world.

Ruest, N., Fritz, S., Deschamps, R., Lin, J., Milligan, I., 2021. From archive to analysis: accessing web archives at scale through a cloud-based interface. International Journal of Digital Humanities, 2(1), 5-24. doi.org/10.1007/s42803-020-00029-6

Ruijgrok, E.C.M., 2006. The three economic values of cultural heritage: a case study in the Netherlands. Journal of Cultural Heritage, 7(3), 206-213. doi.org/10.1016/j.culher.2006.07.002

Ryan, M., Keating, D., Finegan, J., 2022. Managing and accessing web archives: Irish practitioners' perspectives. Ai & Society, 37(3), 975-984. doi.org/10.1007/s00146-021-01364-0

Sargent, S., 2016. Classical Horsemanship and the Dangers of the Emergent Intangible Cultural Heritage Authorised Discourse. International Journal of Intangible Heritage, 11, 35-53. doi.org/10.35638/ijih.2016.11.004

Schafer, V., Pailler, F., 2024. 'All your image are belong to us': heritagization, archiving and historicization of memes. Visual Communication, 23(3), 527-543. doi.org/10.1177/14703572231221030

Tzouganatou, A., 2022. Openness and privacy in born-digital archives: reflecting the role of AI development. Ai & Society, 37(3), 991-999. doi.org/10.1007/s00146-021-01361-3

UNESCO. Charter on the Preservation of Digital Heritage. unesco.org/en/legal-affairs/charter-preservation-digital-heritage (2003-10-15)

UNESCO, 2021a. General Guidelines of the Memory of the World (MoW) Programme. unesdoc.unesco.org/ark:/48223/pf0000378405

UNESCO, 2021b. Guidelines For the Selection Of Digital Heritage For Long-Term Preservation – 2nd Edition. ifla.org/wp-content/uploads/2019/05/assets/hq/topics/cultural-heritage/documents/persist-content-guidelines-zh.pdf (2021-09-14)

waag. De Digitale Stad nominated as Unesco World Heritage. waag.org/en/article/de-digitale-stad-nominated-unesco-world-heritage/ (2022-01-05)

waag, 2017. DIY Handboek voor Webarcheologie. waag.org/sites/waag/files/freeze-diy-handboek.pdf