CHAPTER-11

PRESERVING BORN-DIGITAL CONTENT

Sharad Maurya
Amity University Punjab, Mohali, India
Email: mauryasharad96@gmail.com

ABSTRACT

The exponential growth of born-digital content—items such as emails, social media updates, websites, digital photos and databases—needs immediate preservation measures to guarantee their long-term survival. Born-digital materials have distinct threats compared to physical records, which include technical obsolescence, file format incompatibility, media deterioration and data corruption. This article discusses the complex methods necessary to preserve the authenticity, integrity and usability of such content for future reference. Important strategies are the use of sustainable, open file formats; strong metadata standards for discoverability and the use of secure, redundant storage systems to avoid loss. Preservation techniques such as migration, emulation, and bitstream preservation tackle changing technological environments; while data integrity checks at regular intervals help avoid corruption. Legal and ethical frameworks such as copyright and privacy issues are important to ensure that access is balanced with protection. Organizational policies, global standards (e.g., OAIS, PREMIS) and inter-institutional cooperation improve preservation. Sophisticated tools, like Archivematica and LOCKSS, paired with new technologies like AI, provide scalable solutions. By anticipating these challenges, stakeholders can preserve cultural heritage, institutional memory and societal knowledge contained in born-digital material, ensuring its long-term value in an ever-changing digital environment.

Keyword: Born-digital content, Digital preservation, Emulation, Data Security, Metadata

E Book ISBN: 978-81-981814-6-6| DOI: https://doi.org/10.34256/vadlibs.25.11.139

INTRODUCTION

In the age of digital information, more and more information is generated and exists in digital form only—so-called born-digital content. These are emails, social media messages, websites, digital photos, databases and so on. In contrast to paper records, born-digital materials are by nature fragile, vulnerable to technological obsolescence, file corruption, and media degradation. Preserving such material is essential to the protection of cultural heritage, institutional memory and societal knowledge for the benefit of future generations.

Preservation of born-digital content requires a proactive, multi-faceted strategy to provide long-term accessibility and usability. It demands solutions to challenges like changing file formats, hardware and software dependency and the fast rate of technological change. Successful preservation methods include sustainable storage solutions, standardized metadata, data integrity checks and compliance with legal and ethical guidelines. Using strong policies and the application of specialized tools, institutions and individuals can reduce risk and preserve the context and authenticity of digital content. This chapter lays the groundwork for examining the fundamental principles, strategies and considerations of preserving born-digital content, pointing to its growing relevance in today's increasingly digital environment.

BORN DIGITAL CONTENT

Born-digital content is information that has been generated and only exists in a digital format, with no physical equivalent. These are:

- **Text and Documents:** Word documents, PDFs, emails and e-books.
- **Media:** Digital images, movies, voice recordings and animations.
- **Web Content:** Social media updates, blogs and web pages.
- **Data:** Databases, spreadsheets and datasets created by software.
- **Software and Games:** Programs, video games and interactive media.
- **Other:** Digital art, virtual reality content and Blockchain records.

In contrast to digitized materials, born-digital materials have their origin in digital spaces and are susceptible to technological obsolescence, file corruption and media degradation, which complicate their preservation.

Process of preservation of born digital content:

The preservation of born-digital content uses structured steps to make it accessible in the long term, in its original state and usable. The following is a brief description of the most important stages:

Identification and Selection

- Identify born-digital content (such as emails, websites, and databases) for preservation due to value, importance, or legal obligations.
- Determine selection criteria to prioritize content in accordance with cultural, historical or organizational significance.

Appraisal and Acquisition

- Assess content for feasibility of preservation, such as file types, dependencies and ethical implications.
- Purchase content by transfer, download or web harvesting, with proper documentation and chain of custody.

Pre-Ingest Preparation

- Scan content to learn about its structure, formats and dependencies (e.g., software or hardware).
- Establish an initial metadata (descriptive, technical) to record context and technical information.
- Virus check, duplicates or corrupted file check to maintain quality.

Ingest

- Migrate content to a preservation environment or repository (e.g., Archivematica, DSpace).
- Produce information (e.g., checksums) to confirm data integrity.
- Provide unique identifiers and standardize metadata employing schemas such as PREMIS.

Storage and Maintenance

- Deposit content in redundant, secure systems (e.g., cloud, local servers, and offline media) with geographic spread.
- Conduct periodic backups and media refreshment to avert degradation.
- Audit storage environments for security and stability.

Preservation Planning and Activity

Establish a preservation plan grounded in content requirements:

- **Migration:** Transfer files to sustainable file formats (e.g., PDF/A, TIFF) to prevent obsolescence.
- **Emulation:** Reconstitute original software/hardware environments for viewing.
- **Encapsulation:** Package content with its dependencies.
- **Bitstream Preservation:** Preserve exact replicas of original files.

Conduct routine format maintenance and integrity checks to guarantee usability.

Metadata Management

- Enhance and preserve metadata for discovery, context and technical needs.
- Modify metadata to indicate preservation activities or content changes.

Access Provision

a. Offer controlled access to preserved content via repositories or interfaces, honoring copyright and privacy.
b. Assure compatibility with contemporary systems or offer emulated environments for original rendering.

Monitoring and Auditing

- Regularly audit preserved content for integrity, accessibility and relevance.
- Update preservation plans to cover emerging technologies, risks or standards.

Collaboration and Compliance

- Comply with international standards (e.g., OAIS, ISO 14721) and legislation.
- Work together with institutions, archivists and technologists to exchange resources and best practices.

This cyclical process needs constant resources, expertise and adjustment to technological development. Tools such as LOCKSS,

BagIt, and digital forensics software automate workflows, while policies provide consistency.

SUGGESTIONS

Preserving born-digital material requires a thoughtful strategy to guarantee its availability and integrity in the future. Suggested practices are to use open and sustainable file formats such as PDF/A and TIFF to preclude obsolescence and to construct rich metadata using standards such as PREMIS for context and discovery. Secure, redundant storage on cloud, local and offline systems, combined with periodic backups to the 3-2-1 rule, protects against loss of data. Preservation methods like migration to new media, emulation of old environments and bitstream preservation ensure usability and authenticity. Workflows are simplified by tools like Archivematica and LOCKSS, while checksums and BagIt guarantee data integrity. Organizations must have well-defined preservation policies in accordance with standards like OAIS, provide resources for training and perform regular audits to track content health. Tackling legal and ethical concerns, such as copyright and privacy, assures compliance and working with networks like the Digital Preservation Coalition ensures collective solutions. Reducing obsolescence means capturing dependencies and investigating virtualization, with newer technologies like AI making scalability even better. Access testing on a regular basis and educating stakeholders reinforces further. By putting these measures in place, stakeholders can shield diverse born-digital assets, websites, datasets and more—from technological change and decay, saving cultural and institutional memory for generations to come in a rapidly changing digital environment.

CONCLUSION

Preservation of born-digital content is a crucial activity in protecting the digital heritage of our time. As emails, social media, websites and datasets continue to shape cultural, historical and institutional accounts, proactive approaches are essential in reversing threats such as technological obsolescence, data loss and media degradation. By embracing sustainable file formats, sound metadata practices, secure storage technologies and preservation methods like migration and emulation, stakeholders can guarantee long-term accessibility and authenticity. Legal and ethical frameworks, as well as compliance with international standards, enhance these efforts even further, while collaboration and creative tools maximize scalability. The process

requires continuous dedication, resources and flexibility to changing technologies. Finally, keeping born-digital information is not a technical hurdle alone but also a cultural mandate where future generations must be able to access and relate to the digital artifacts that transform our world presently.

REFERENCES

[1] Franklin, C. A., Usaf. (2011). Information Assurance (Ia) And Support To Computer Network Defense (Cnd). In Chairman Of The Joint Chiefs of Staff Instruction J-6 CJCSI 6510.01F. https://www.jcs.mil/Portals/36/Documents/Library/Instructions/6510_01.pdf

[2] Artefactual Systems Inc. (n.d.). DSpace exports, Documentation (Archivematica 1.11.2) , Archivematica: open-source digital preservation system. https://www.archivematica.org/en/docs/archivematica-1.11/user-manual/transfer/dspace/

Book Title "Advancing Library and Information Science: Innovations, Practices and Future Directions",
Edited By- Dr. Jatinder Kumar, Dr. Mariraj Vasudev Sedam.
A Book with CC-BY license at https://press.vyomhansjournals.com
Published by Vyom Hans Publications,
ISBN (Digital Download and Online): 978-81-981814-6-6
ISBN (Book): 978-81-981814-5-9
Year: April, 2025
DOI: https://doi.org/10.34256/vadlibs.25.11.139

HOW TO CITE
Sharad Maurya. (2025). Preserving Born-Digital Content. Ed. by Kumar, J. & Sedam, M. V. in "Advancing Library and Information Science: Innovations, Practices, and Future Directions" (pp. 139-144).Vyom Hans Publications. https://doi.org/10.34256/vadlibs.25.11.139;
E-ISBN: 978-81-981814-6-6
