

GLAM के डिजिटल डेटा के लिए आपदा प्रबंधन: एक अध्ययन Disaster Management Plan for Digital Data of GLAM: A Case Study

प्रतापानन्द झा¹ एवं प्रोफेसर ओम विकास²

Pratapanand Jha¹ (Research Scholar), Prof. Om Vikas²

¹Dept. of Computer Science and Engineering, Banasthali Vidyapith, Rajasthan

²Formerly, Director, ABV-Indian Inst. of IT & Management Gwalior; &

Counsellor (S&T), Indian Embassy, Japan; & Sr. Director, Ministry of Electronics & IT

pjha@ignca.nic.in, dr.omvikas@gmail.com

<https://doie.org/10.0524/VP.2024254982>

सारांश

GLAM दीर्घाओं (Galleries), पुस्तकालयों (Libraries), अभिलेखागारों (Archives) और संग्रहालयों (Museums) का संक्षिप्त रूप है, और GLAM संस्थाओं का मुख्य कार्य विभिन्न सांस्कृतिक विरासतों के संरक्षण के अलावा उनसे संबंधित ज्ञान को आम जनता तक पहुंचाना भी है। आज डिजिटल डेटा हमारी अधिकतर गतिविधियों का एक हिस्सा बन गया है। अपने सभी दैनिक कार्यों के लिए हम डिजिटल उपकरणों का प्रयोग करते हैं। लेखन, फोटोग्राफी, ऑडियोग्राफी, वीडियो रेकॉर्डिंग तथा एनीमेशन फिल्म आदि बनाने के लिए बाजार में उपलब्ध सभी उपकरण डिजिटल प्लेटफॉर्म पर ही काम करते हैं। इसके चलते हम प्रत्येक दिन बहुत बड़ी मात्रा में डिजिटल डेटा बना रहे हैं। इनके अलावा, विश्व भर में उपलब्ध सांस्कृतिक विरासतों को डिजिटल प्रारूप में परिवर्तित करने के लिए एक विशाल कार्यबल निरंतर कार्यरत है। विश्व आर्थिक मंच (World Economic Forum, WEF) के वर्ष 2023 रिपोर्ट के अनुसार यह अनुमान लगाया गया है कि वर्ष 2025 तक वैश्विक स्तर पर प्रत्येक दिन करीब 463 एक्सबाइट डिजिटल डेटा बनते रहेंगे। इनमें से अधिकांश डेटा, जैसे हवाई जहाज या रेलवे के टिकट, पार्सल रिकॉर्ड, बैंकिंग लेनदेन आदि की प्रासंगिकता अल्पकालिक हो सकती है लेकिन सांस्कृतिक विरासतों से संबंधित डेटा की प्रासंगिकता दीर्घकालिक ही होती है। सांस्कृतिक विरासतों का संरक्षण विभिन्न सांस्कृतिक संस्थाओं द्वारा किया जाता है, जिन्हें संयुक्त रूप से आजकल GLAM संस्थाएँ कहा जाता है। GLAM संस्थाओं द्वारा सांस्कृतिक डिजिटल डेटा को लोगों तक ऑनलाइन पहुँच बनाने के चलते उन्हें एक नई चुनौती का सामना करना पड़ता है, और वह है संरक्षित डिजिटल डेटा को साइबर अपराधों की पहुँच से बाहर रखना।

भारतीय आंकड़ों के अनुसार वर्ष 2020 में, देश भर में करीब 2.9 लाख से अधिक साइबर अपराध दर्ज किये गये थे। चूँकि साइबर अपराधों की संख्या में प्रत्येक वर्ष 15–20% की दर से वृद्धि हो रही है, इसलिए GLAM संस्थाओं द्वारा सांस्कृतिक डिजिटल डेटा का संरक्षण अत्यंत ही नाजुक मुद्दा होने के साथ साथ नितांत जरूरी उत्तरदायित्व भी बन गया है। साइबर अपराधों के अलावा सांस्कृतिक डिजिटल डेटा में हेरफेर हो जाने की संभावना कुछ संस्थागत गतिविधियों में जाने अनजाने चूक होने के चलते भी हो सकता है, जैसे कि मीडिया रिफ्रेशिंग, मीडिया माइग्रेशन, प्रारूप माइग्रेशन और प्लेटफॉर्म अपग्रेडेशन, इत्यादि। इस परिदृश्य में सभी GLAM संस्थाएँ अपने सांस्कृतिक डिजिटल डेटा को दीर्घकालीन अवधि तक संरक्षित रखने के लिए आज पूरी तरह से तैयार नहीं हैं। अतः, इन संस्थानों को अपने सांस्कृतिक डिजिटल डेटा की सुरक्षा के लिए 'ट्रस्टेड डिजिटल रिपॉजिटरी (टीडीआर) मानक जैसे किसी अन्य मानकों को जल्द से जल्द अपनाने

की जरूरत है।

Abstract

GLAM is an acronym for Galleries, Libraries, Archives, and Museums, which refers to the cultural institutions which are responsible for maintaining and preserving cultural heritage, and with a mission to provide its access to the society, in general. Today, digital data creation is a part of all our activities. We create most of our documents digitally. All the gadgets available in the market for creating texts, visual, audio, video, animation contents etc., work on digital platforms and generate digital contents. Hence, every day we generate large volumes of digital data which are born digital. In addition, a large workforce is employed across the world to convert most of the available analog data into digital form. As per the World Economic Forum Report of year 2023, it is estimated that by the year 2025 around 463 exabytes of data will be created every day, globally. Most of these data such as air/railway tickets, parcel records, banking transactions etc. may only have short-time relevance but a significant amount of the data, especially related with cultural heritage will have long-time relevance. And thus, to preserve the cultural history of a region, community, or a country the need to preserve such data is of utmost importance.

Further, any digital data which are offered for online for access by GLAM institutions are exposed to various types of cybercrimes. As per the national crime statistics of India, more than 2.9 lakh cyber security incidents were reported in the year 2020. As the volume of cybercrime is growing at the rate of 15-20% per year preserving digital data is a highly delicate but extremely important issue for the GLAM institutions. In addition to cybercrimes the digital data is also susceptible to damage by inappropriate internal operations like media refreshing, media migration, format

transformation, and platform upgradation. For such scenario, most of the GLAM institutions are not well equipped to maintain and preserve their digital data for long term duration. They need to adopt either the Trustworthy Digital Repository (TDR) standard or some similar standard to safeguard their precious digital data as early as possible.

मुख्य शब्द: आपदा प्रबंधन, अंकीय संरक्षण, विश्वसनीय अंकीय कोष, राष्ट्रीय सांस्कृतिक, दृश्य-श्रव्य अभिलेखागार

Keywords: Disaster Management, Digital Preservation, Trustworthy Digital Repository (TDR), National Cultural Audiovisual Archives (NCAA), GLAM

परिचय

आधुनिक जीवन के लगभग सभी पहलुओं को डिजिटल प्रौद्योगिकी ने बदल दिया है। हाल के दशकों में संचार, व्यापार, बाजार, मनोरंजन और बैंकिंग जैसे क्षेत्रों में क्रांतिकारी परिवर्तन आया है। अब कोई ऐसा उपकरण या मशीन मिलना दुर्लभ है जो किसी न किसी तरह के इलेक्ट्रॉनिक या डिजिटल प्रौद्योगिकी पर आधारित न हो। डिजिटल प्रौद्योगिकी के चलते उपकरण अधिक दक्ष, आकार में छोटे, वजन में कम तथा एक से अधिक कार्यों को सम्पन्न करने में सक्षम हो गये हैं। सूचनाओं को स्थानीय या दूरस्थ रूप से संगृहीत किया जा सकता है और उन्हें तत्काल एक-स्थान से दूसरे स्थान तक भेजा जा सकता है। यहां तक कि 'सूचना' शब्द अब केवल शब्दों और संख्याओं को संदर्भित नहीं करता अपितु इसका विस्तृत अर्थ अब फोटो, ऑडियो और वीडियो जैसे माध्यमों तक को समाविष्ट करता है।

हमारे सभी लेन-देन चाहे आधिकारिक हों या निजी, जैसे बैंकिंग, परिवहन, चिकित्सा, मनोरंजन या खरीददारी, ये सभी हम ऑनलाइन करने लगे हैं। इस कारण हम प्रत्येक दिन बहुत बड़ी मात्रा में डिजिटल डेटा बना रहे हैं। इसके अलावा वह विरासत के

प्राथमिक स्रोतों तक की पहुँच सांस्कृतिक और अन्य शोधकर्ताओं के लिये सुलभ बनाते हैं।

GLAM संस्थाओं द्वारा सांस्कृतिक डिजिटल डेटा को लोगों तक ऑनलाइन पहुँच बनाने का उत्तरदायित्व है। पर इसके कार्यान्वयन के चलते उन्हें एक नई चुनौती का सामना करना पड़ता है, और वह हैं संरक्षित डिजिटल डेटा को साइबर अपराधियों की पहुँच से बाहर रखना। विश्व आर्थिक मंच (WEF) ने अपनी वैश्विक सुरक्षा आउटलुक रिपोर्ट वर्ष 2023 में GLAM संगठनों के समक्ष उपस्थित चुनौतियों और साइबर अपराध के खतरे के खिलाफ खुद को सक्षम बनाने और तैयार रहने की तत्काल आवश्यकता पर बल दिया है। इस रिपोर्ट के अनुसार, 'जागरूकता और आवश्यक तैयारियाँ, संगठनों को नई डिजिटल प्रौद्योगिकी से होने वाले लाभ एवं साइबर अपराध जैसे घटनाओं के बीच संतुलन रखने में मदद करेगी।... विशेषज्ञों का मानना है कि वैश्विक भू-राजनैतिक अस्थिरता के चलते अगले दो वर्षों में किसी भयावह साइबर घटना घटित होने की संभावना अत्याधिक है।... प्रौद्योगिकियाँ आमतौर पर अनेक संगठनों के बीच साझा की जाती हैं, और जिसके परिणामस्वरूप इन संगठनों की निर्भरताएं और कमजोरियाँ लगभग एक जैसी ही होती हैं।'

ज्ञात/अज्ञात खतरों से अपने डेटा को सुरक्षित रखने के लिए, संस्थानों को एक आपदा प्रबंधन योजना (Disaster Management Plan) बनाने की आवश्यकता होती है। प्रस्तुत शोधपत्र में आपदा प्रबंधन योजना के बुनियादी पहलुओं, प्रकृति और डिजिटल डेटा से संबंधित मुद्दों को रेखांकित करने का प्रयास किया गया है।

डिजिटल डेटा के संरक्षण के लिए, ट्रस्टेड डिजिटल रिपॉजिटरी (Trusted Digital Repository - TDR) के रूप में मानक तैयार किया गया है, लेकिन आज तक दुनिया भर में केवल दो रिपॉजिटरी को ही ट्रस्टेड डिजिटल रिपॉजिटरी के रूप में ISO 16363:2012 के लिए प्रमाणित किया गया है।

संस्कृति मंत्रालय, भारत सरकार की 'राष्ट्रीय सांस्कृतिक दृश्य-श्रव्य अभिलेखागार (National Cultural Audiovisual Archive - NCA) परियोजना को नवंबर 2017 में विश्व का प्रथम ISO 16363:2012 प्रमाणित ट्रस्टेड डिजिटल रिपॉजिटरी बनने का सौभाग्य प्राप्त हुआ। यद्यपि टीडीआर प्रमाणन कोई अंतिम विकल्प नहीं है और इसकी विफलता की संभावना हमेशा बनी होती है, चाहे वह संगठनात्मक हो, वित्तीय हो, या कोई अन्य। NCA द्वारा अपनाई गई डेटा भंडारण और पुनरुप्राप्ति योजना के कुछ अंश को यहाँ प्रस्तुत किया गया है।

डिजिटल डेटा की प्रकृति एवं उसके संरक्षण की चुनौतियाँ:

तीव्र गति से हो रहे प्रौद्योगिकी विकास के कारण किसी खास प्रारूप में रखे डिजिटल डेटा पर प्रचलन में न बने रहने का भय सतत बना रहता है। इस खतरे के तीन निम्नलिखित कारण हैं—

- डिजिटल स्टोरेज मीडिया का प्राकृतिक ह्रास एवम् उनका अप्रचलन
- डिजिटल डेटा का मालिकाना प्रारूप (proprietary format)
- डेटा ऑपरेट करने के लिए आवश्यक हार्डवेयर और सॉफ्टवेयर का अप्रचलन/पुराना पड़ जाना।

वस्तुस्थिति:

- डिजिटल सामग्रियों के आपदा प्रबंधन योजना के अनेक सफल उदाहरण आज हमारे पास नहीं हैं।
- डिजिटल सूचना के निर्माण के पहले से ही उनके संरक्षण की प्रक्रिया शुरू करनी पड़ती है। संरक्षण का अर्थ है, उनकी पहुँच और उपयोगिता को लंबे समय तक के लिये सुनिश्चित करना। यह सिर्फ बिट-स्ट्रीम संरक्षण की बात नहीं है।
- डिजिटल संरक्षण की चुनौतियों में सिर्फ प्रौद्योगिकी ही नहीं बल्कि इनका संबंध संगठनात्मक ढाँचे से भी है।

– आज एक राष्ट्रीय नीति की आवश्यकता है, डेटा के डिजिटल संरक्षण और नीति की प्रभावी कार्यान्वयन प्रक्रिया के लिए।

डिजिटल डेटा लंबे समय में अनुपयोगी न हो जाए इसके लिए निम्नलिखित तकनीक प्रयोग में है:

- **रिफ्रेशिंग (Refreshing)** – रिफ्रेशिंग प्रक्रिया से डेटा को अपने वर्तमान भंडारण माध्यम से उसी प्रकार के, पर नये भंडारण माध्यम पर स्थानांतरण किया जाता है। इस प्रक्रिया से डेटा की बिटरेट में या उसके प्रारूप में कोई परिवर्तन या बदलाव नहीं किया जाता है। यह प्रक्रिया वर्तमान भंडारण माध्यम के प्राकृतिक ह्रास से डेटा को सुरक्षित रखती है।
- **माइग्रेशन (Migration)** – माइग्रेशन प्रक्रिया डेटा एवं एप्लीकेशन को एक नए सिस्टम (आधार संरचना) पर पुनर्सथापित करती है। माइग्रेशन दो प्रकार का होता है – (i) डेटा को एक फाइल प्रारूप से रूपांतरण कर दूसरे फाइल प्रारूप में स्थानांतरित करना, या (ii) एक ऑपरेटिंग सिस्टम से दूसरे पर आरोहित करना ताकि संसाधन का उपयोग निर्बाध बना रहे।
- **प्रतिकृति (Replication)** – एक या अधिक सिस्टम पर डेटा की डुप्लिकेट प्रतियां बनाना प्रतिकृति कहलाती है। जो डेटा केवल एक ही स्थान पर एकल प्रतिलिपि के रूप में मौजूद होती है, वह सॉफ्टवेयर या हार्डवेयर की विफलता, मानवीय भूल या चालाकी, तथा आग, बाढ़ जैसी पर्यावरणीय आपदाओं के प्रति अत्यधिक संवेदनशील होती है। यदि डिजिटल डेटा प्रतिलिपियाँ को कई स्थानों पर रखी जाए, तो उसके दीर्घकाल तक उपयोगी बने रहने की संभावना बनी रहती है।
- **अनुकरण (Emulation)** – अनुकरण किसी अप्रचलित प्रणाली की क्रियात्मकता (functionality) की प्रतिकृति है। इसका उद्देश्य

उस वातावरण को फिर से बनाना है जिस वातावरण में डिजिटल ऑब्जेक्ट्स को मूल रूप से बनाया गया था। इससे अप्रचलित प्रणाली से निर्मित डेटा को पुनः उपयोगी बनाया जा सकता है।

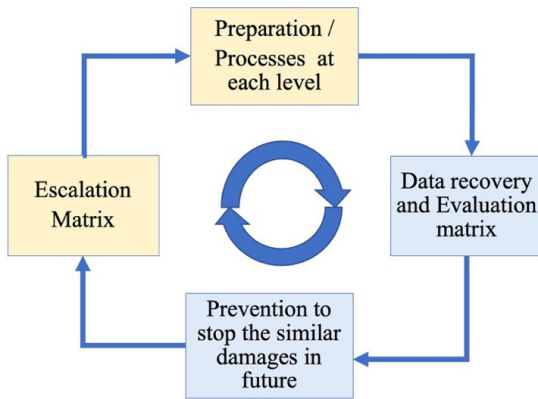
- **एनकैप्सुलेशन (Encapsulation)** – यह तकनीक सुनिश्चित करती है कि संरक्षित ऑब्जेक्ट्स का वर्णन डेटा के साथ रखना चाहिए। यह वस्तुतः सामग्री को, तथा उसे समझने के लिए आवश्यक सभी जानकारियों के, एक साथ जोड़ कर रखना है। डिजिटल ऑब्जेक्ट से जुड़ी फाइलों में सभी सूचनाओं तथा उसके घटकों के बीच के संबंधों का विवरण, "कंटेनर" या "रैपर" नामक संरचनाओं का उपयोग करके किया जाता है। जिसका उपयोग मशीन लर्निंग के माध्यम से 'एमुलेटर', 'व्यूअर' या 'कन्वर्टर्स' के विकास में किया जा सकता है।
- **मेटाडेटा एकीकरण (Metadata Integration)** – मेटाडेटा एक डिजिटल फाइल संबंधित डेटा है जिसमें फाइल के निर्माण, उपयोग, अधिकार, प्रतिबंध, संरक्षण इतिहास और प्रबंधन की जानकारियाँ शामिल हैं। डिजिटल फाइलों से जुड़े मेटाडेटा फाइल, उसके प्रारूप के अप्रचलन से प्रभावित हो सकता है। ASCII को मेटाडेटा के लिए सबसे टिकाऊ प्रारूप माना जाता है क्योंकि इसका प्रचलन व्यापक है। यह जानकारियों को बरकरार रखता है, पर उसके संरचना की जानकारी को नहीं जिसमें इसे प्रस्तुत किया गया है। उच्च क्रियात्मकता लिये एसजीएमएल (SGML) या एक्सएमएल (XML) का उपयोग किया जाता है।

डिजिटल ऑब्जेक्ट प्रबंधन तथा डेटा की निर्बाध उपलब्धता को सुनिश्चित करने के लिये, डिजिटल संरक्षण की प्रतिबद्धता को पूरा करने के लिए, तथा डेटा की अखंडता, प्रामाणिकता और गोपनीयता को सुनिश्चित करने के लिए उपरोक्त तकनीकियों को

कार्यप्रणाली के एक भाग के रूप में कार्यान्वित करने की आज जरूरत है।

आपदा प्रबंधन योजना

ज्ञात खतरों से डेटा की सुरक्षा के लिए तैयार की गई एक संस्थागत योजना, आपदा प्रबंधन योजना है। सभी महत्वपूर्ण डेटा को, चाहे वह ऐतिहासिक, सांस्कृतिक, वैज्ञानिक, आर्थिक या कोई अन्य हो, इन्हें भविष्य के संदर्भों के लिए सर्वथा संरक्षित किया जाना चाहिए। नियमित संरक्षण प्रक्रिया, डेटा समूह को एक खास तरह के खतरे से बचाने के लिये अपनाई जाती है। उनके अलावा प्राकृतिक आपदाओं, जैसे बाढ़, आगजनी, भूकंप आदि, तथा मानवीय भूलों और बदनीयती से भी डेटा की सुरक्षा के लिए भी संस्थान को ध्यान देने की आवश्यकता है। अतः, डेटा की तत्काल पुनरुप्राप्ति के लिए एक ऐसी योजना की आवश्यकता है जिसमें संस्थागत संरचना और प्रक्रियाएँ शामिल हो। जैसा कि ब्लॉक आरेख में नीचे दर्शाया गया है, ऐसी योजना में एस्केलेशन मैट्रिक्स, विभिन्न स्तरों पर आवश्यक प्रक्रियाओं की तैयारी, डेटा की पुनःप्राप्ति, और भविष्य में ज्ञात क्षतियों से रोकथाम के उपाय, जैसे घटक शामिल हैं:



चित्र 1. आपदा-प्रबंधन योजना का खंड आरेख

- I. एस्केलेशन मैट्रिक्स – एस्केलेशन मैट्रिक्स के माध्यम से खतरे या क्षति संबंधित जानकारियाँ अधिकारियों को दी जा सकती है। यह मुद्दों

के समाधान के लिए एक संरचित ढांचा प्रदान करता है। इसका लक्ष्य यह सुनिश्चित करना है कि सभी मुद्दों को समय पर, और अधिक प्रभावी तरीके से हल किया जा सके।

- II. विभिन्न स्तरों पर आवश्यक प्रक्रियाओं की तैयारी – डेटा को पुनःप्राप्त करने के लिये और नुकसान को कम करने के लिए संबंधित प्राधिकारी द्वारा कार्यवाई करने के लिए अपेक्षित प्रक्रियाओं को परिभाषित करना और उनकी तैयारी करना है।
- III. डेटा की पुनःप्राप्ति – डेटा की पुनःप्राप्ति के लिए प्राधिकरण द्वारा की गई कार्यवाइयाँ, और सुधारात्मक उपायों के लिए की गई प्रक्रियाओं का मूल्यांकन करते रहना है।
- IV. ज्ञात क्षतियों से रोकथाम के उपाय – भविष्य में ज्ञात क्षतियों से रोकथाम के लिए बनायी योजना में समयानुकूल बदलाव करते रहना अपेक्षित है।

उपरोक्त आरेख का उपयोग किसी भी प्रकार के डेटा के लिए किया जा सकता है, चाहे डेटा एनालॉग हो या डिजिटल। हालाँकि, विभिन्न प्रकार के डेटा और विभिन्न प्रकृति की आपदाओं के लिए तैयारियाँ, प्रतिक्रियाएँ और पुनःप्राप्ति के तरीके अलग-अलग होते हैं।

डिजिटल डेटा के उन मामलों में जहाँ डेटा अपनी प्रकृति के चलते, तेजी से होती प्रौद्योगिक अप्रचलन के कारण क्षति के प्रति अति संवेदनशील हो गए है, वहाँ संरक्षण के लिये आवश्यक अधोसंरचना, विशेषज्ञता और प्रक्रियाएं काफी भिन्न होती हैं, जैसा कि अगले अनुभाग में बताया गया है।

डिजिटल डेटा संरक्षण के प्रयास एवं उपलब्ध तकनीकियाँ:

नब्बे के दशक में डिजिटल संरक्षण की चुनौती से निपटने के लिए कई समानांतर प्रयास दुनिया भर में किए जा रहे थे। अंतरिक्ष डेटा सिस्टम (Consultative Committee for Space Data System - CCSDS), वाशिंगटन डीसी की सलाहकार समिति (जो विभिन्न

देशों की दस सदस्य एजेंसियों और प्रेक्षक के रूप में अन्य तेईस अतिरिक्त संगठनों से बनी है) ने एक संयुक्त गतिविधि शुरू की:

- ओपन अभिलेखीय सूचना प्रणाली (Open Archival Information System – OAIS) का एक संदर्भ मॉडल को विकसित करने के लिए, और इन मुद्दों पर मुख्य रूप से आम सहमति बनाने के लिए –
- दीर्घकालिक संरक्षण से संबंधित नियमों और अवधारणाओं की सामान्य रूपरेखा बनाना,
 - डिजिटल अभिलेख, उसके वातावरण और उसकी सूचनाओं का एक संदर्भ मॉडल तथा एक कार्यात्मक मॉडल बनाना, जिसके मानदण्ड पर किसी भी अन्य डिजिटल अभिलेखागारों की तुलना की जा सके,
 - एक डिजिटल अभिलेख की तकनीकी, संगठनात्मक संरचना, तथा उसकी प्रतिबद्धताओं की आवश्यकताओं को रेखांकित करना ताकि वे अपने ग्राहकों को, यानि नामित समुदाय को दीर्घकालिक संरक्षण सेवाएँ प्रदान कर सकें।

इसकी रिपोर्ट, जिसे 'ब्लूबुक' कहा जाता है, जनवरी 2002 में आई थी, उसे बाद में ISO मानक, ISO-14721, के रूप में अपनाया गया। उसका एक संशोधित और अद्यतन संस्करण वर्ष 2012 में ISO मानक, ISO-14721:2012, के रूप में प्रकाशित किया गया। इस मानक के अनुसार संदर्भ मॉडल की मुख्य संकल्पना एक ओपन अभिलेखीय सूचना प्रणाली (OAIS) की है। एक OAIS आधारित अभिलेखागार को कम से कम 6 निम्नलिखित उत्तरदायित्व निभाने पड़ते हैं, और वे हैं: अभिलेख का समावेशन (इनजेस्ट), अभिलेखीय भंडारण, डेटा प्रबंधन, दीर्घकालीन संरक्षण योजना, अभिलेखों की उपलब्धता तथा उनका प्रशासन।

ट्रस्टेड डिजिटल रिपॉजिटरी

किसी प्रणाली की विश्वसनीयता का माप है कि वह अपने उद्देश्यों और अपनी विशिष्टताओं

(specifications) के अनुरूप संचालन करने में कितना सक्षम है, यानी वह कितना कार्य अपने किये गए दावे के अनुरूप करता है। किसी भी डिजिटल रिपॉजिटरी की विश्वसनीयता का निर्धारण करने के लिये अनेक मुद्दों का आकलन जरूरी है, और वे हैं: डिजिटल डेटा की अखंडता, प्रामाणिकता, और उनकी गोपनीयता बनाये रखना; डेटा की उपलब्धता सुनिश्चित करने के लिए डिजिटल ऑब्जेक्ट्स का प्रबंधन; तकनीकी बुनियादी अधोसंरचना और डेटा सुरक्षा के उपाय; तथा डिजिटल संरक्षण की प्रतिबद्धता को पूरा करने के लिए उचित संस्थागत अधोसंरचना। संक्षेप में, एक ट्रस्टेड डिजिटल रिपॉजिटरी, डिजिटल डेटा से संबंधित सभी संभावित खतरों तथा अन्य जोखिमों को समझता है जैसे, मीडिया, सॉफ्टवेयर और हार्डवेयर की विफलता; संचार त्रुटियों और संचार नेटवर्क की विफलता; मीडिया, सॉफ्टवेयर और हार्डवेयर प्लेटफार्मों का अप्रचलन; मानवीय भूल, आंतरिक और बाहरी खतरे; संगठनात्मक विफलता आदि। एक डिजिटल रिपॉजिटरी स्वयम् को विश्वसनीय बनाने के लिये संगठनात्मक संरचना में बदलाव लाता है ताकि वह उपरोक्त जोखिमों का भलि भांति प्रबंधन कर सके।

तृतीय पक्ष द्वारा किसी डिजिटल रिपॉजिटरी का मूल्यांकन और उसका प्रमाणन उस डिजिटल रिपॉजिटरी और उसके उपयोगकर्ताओं के बीच एक उच्च स्तर का विश्वास सुनिश्चित करता है। एक डिजिटल रिपॉजिटरी, उसे दिये गए किसी भी डिजिटल दस्तावेज का संग्रह और संरक्षण करता है ताकि उसकी उपलब्धता भविष्य के लिए सुनिश्चित हो सके। एक डिजिटल भंडार का प्रमाणीकरण वस्तुतः इन सभी अवयवों/घटकों का प्रमाणीकरण है:

- (i) संगठन के अभिलेखीय कार्यक्रम,
- (ii) अभिलेखीय प्रक्रिया,
- (iii) संगृहीत डेटा, और

(iv) रिपॉजिटरी के प्रबंधन में कार्यरत सभी अभिलेखीय विशेषज्ञ। उपरोक्त सभी प्रयासों के बाद

सीआरएल (CRL) आरएलजी/एनएआरए RLC/NARA टास्क फोर्स ने (जिसके सह-अध्यक्ष सीआरएल (CRL) और 'ओसीएलसी (OCLC) हैं), फरवरी 2007 में ट्रस्टेड डिजिटल रिपॉजिटरी का पहला संस्करण प्रकाशित किया गया। बाद में, एक अंतर्राष्ट्रीय ऑडिट और प्रमाणन प्रक्रिया की दिशा में एक ऑडिट चेकलिस्ट का उपयोग करके एक मानदंड स्थापित किया गया। इसके प्रमाणन के लिए एक संशोधित और अद्यतन संस्करण वर्ष 2012 में एक मानक ISO 16363:2012 'Trustworthy Digital Repository: Establishing Audit and Certification Criteria towards an International Audit & Certification Process using Checklist for Audit and Certification' के रूप में प्रकाशित किया गया।

ISO/IEC 27000:2018 सूचना सुरक्षा प्रबंधन प्रणालियों (ISMS) का अवलोकन प्रदान करता है। यह आमतौर पर मानकों के आईएसएमएस परिवार में उपयोग किए जाने वाले नियम और परिभाषाएं भी प्रदान करता है। यह दस्तावेज सभी प्रकार और आकार के संगठन (जैसे वाणिज्यिक उद्यम, सरकारी एजेंसियां, गैर-लाभकारी संगठन) पर लागू होता है। संस्थागत मूल्यवान जानकारी को अच्छे व्यावसायिक कारणों से व्यवस्थित रूप से सुरक्षित रखने और वैध तरीके से उपयोग करने का प्रावधान ISO27K में है।

ISO/IEC 27000 पहली बार 2009 में प्रकाशित हुआ था और 2012, 2014, 2016 और 2018 में अपडेट किया गया था। वर्तमान 2018 पांचवां संस्करण ITTF से अंग्रेजी और फ्रेंच भाषा में उपलब्ध है। यह 2016 के चौथे संस्करण का एक छोटा सा संशोधन था जिसमें संक्षिप्तीकरण पर एक अनुभाग था, और आईएसओ/आईईसी 27004 के 2016 के पुनर्लेखन के बाद मेट्रिक्स-संबंधित परिभाषाओं का युक्तिकरण था। 27000 का छठा संस्करण का कार्य-प्रगति पर है। इसका संभावित शीर्षक "सूचना सुरक्षा, साइबर सुरक्षा और गोपनीयता सुरक्षा – सूचना सुरक्षा प्रबंधन प्रणाली – अवलोकन" है।

ऑडिट और प्रमाणन के लिए चेकलिस्ट में तीन खंड हैं: संगठनात्मक संरचना, डिजिटल ऑब्जेक्ट प्रबंधन, और प्रौद्योगिकी संरचना और सुरक्षा। एक विश्वसनीय डिजिटल रिपॉजिटरी को प्रलेखन, पारदर्शिता, पर्याप्ता और मापनीयता से संबंधित सभी मानदंडों पर खरा उतरना पड़ता है। दीर्घकालिक डिजिटल संरक्षण के लिए अपनी प्रतिबद्धता के विरुद्ध लेखा परीक्षकों द्वारा मूल्यांकन के लिये रिपॉजिटरी के सभी परिचालन चरणों के दस्तावेजों का होना अनिवार्य है।

संरक्षण और सुरक्षा संबंधित अपनी सभी प्रक्रिया में एक डिजिटल रिपॉजिटरी को पारदर्शी होने की जरूरत है ताकि सभी कथित खतरों और जोखिमों के खिलाफ अपनी विश्वसनीयता पर अधिकार जताया जा सके। एक रिपॉजिटरी को आवधिक निगरानी (periodic surveillance) ऑडिट से सालाना गुजरना पड़ता है, और तीन साल के बाद इन्हें पुनः प्रमाणन कराना पड़ता है। प्राइमरी ट्रस्टेड डिजिटल रिपोजिटरी ऑथराइजेशन बॉडी लिमिटेड (Primary Trusted Digital Repository Authorization Body Ltd. - PTAB), यूनाइटेड किंगडम एकमात्र निकाय है जिसे टीडीआर की ऑडिटिंग और प्रमाणित करने के लिए नेशनल एक्रिडिटेशन बोर्ड फॉर सर्टिफिकेशन बॉडी (National Accreditation Board for Certification Body - NABCB) द्वारा अनुमोदित किया गया है।

संस्कृति मंत्रालय, भारत सरकार की राष्ट्रीय सांस्कृतिक दृश्य-श्रव्य अभिलेखागार (National Cultural Audio-visual Archive – NCAA) परियोजना (<http://ncaa.gov.in/repository/>), जिसे इंदिरा गांधी राष्ट्रीय कला केंद्र 'http://ignca.gov.in' ने सी-डैक पुणे के तकनीकी सहयोग से कार्यान्वित किया है, उस अभिलेखागार को नवंबर 2017 में आईएसओ 16363:2012 प्रमाणित ट्रस्टेड डिजिटल रिपॉजिटरी बनने का सौभाग्य प्राप्त हुआ जो दुनिया का सबसे पहला प्रमाणित ट्रस्टेड डिजिटल रिपॉजिटरी है।

राष्ट्रीय सांस्कृतिक दृश्य-श्रव्य अभिलेखागार के लिए प्रयुक्त डेटा प्रबंधन

पिछले छह दशकों में बनाई गई भारत की सांस्कृतिक संपदा का एक बड़ा हिस्सा विभिन्न सरकारी, गैर-सरकारी संस्थानों और निजी संग्रहों में दृश्य-श्रव्य रूप में संगृहीत है। इनमें भारत के कई महान कलाकारों के दुर्लभ रचनाओं का भण्डार है। यह एक अमूल्य राष्ट्रीय धरोहर है जिसे सदैव संरक्षित रखने के साथ-साथ देश के नागरिकों के लिए उपलब्ध कराने का भी उत्तरदायित्व है। संगृहीत माध्यम की नाजुकता के साथ साथ जागरूकता की कमी एवम् उचित रखरखाव तथा व्यवस्थित और आधुनिक संरक्षण प्रौद्योगिकियों के अभाव में इन सामग्रियों का हमेशा के लिए खो जाने का खतरा सतत बना हुआ है। मुद्दे की गंभीरता को समझते हुए, संस्कृति मंत्रालय ने 12वीं पंचवर्षीय योजना में सांस्कृतिक दृश्य-श्रव्य सामग्रियों के संग्रह के लिए एक परियोजना का प्रस्ताव रखा गया जिसे राष्ट्रीय सांस्कृतिक दृश्य-श्रव्य अभिलेखागार (National Cultural Audio-visual Archive – NCAA) परियोजना के रूप में कार्यान्वयन करने के लिए इन्दिरा गाँधी राष्ट्रीय कला केन्द्र (Indira Gandhi National Center for Arts – IGNCA) को मंजूरी दे दी गई।

परियोजना के प्रारंभिक (पायलट) चरण में देश भर के कई सरकारी और गैर-सरकारी संस्थानों से अप्रकाशित और गैर-व्यावसायिक ऑडियो-वीडियो रिकॉर्डिंग का चुनाव किया गया है और इस 31,500 घंटे के ऑडियो-वीडियो रिकॉर्डिंग को डिजिटलाईज किया गया है। भारतीय सांस्कृतिक विरासत की इन रिकॉर्डिंग्स को जन-जन तक पहुंचाने के लिये इसे <http://ncaa.gov.in> वेबसाइट के माध्यम से ऑनलाइन उपलब्ध कराया गया है। मुख्यतः कलाकारों, शोधकर्ताओं, छात्रों और विद्वानों वृत्तचित्र तथा फिल्म निर्माताओं, फिल्म और टेलीविजन कलाकारों, लेखकों, GLAM (गैलरी, पुस्तकालय, अभिलेखागार और संग्रहालय) पेशेवरों एवम् संस्कृति पिपासुओं के लिए

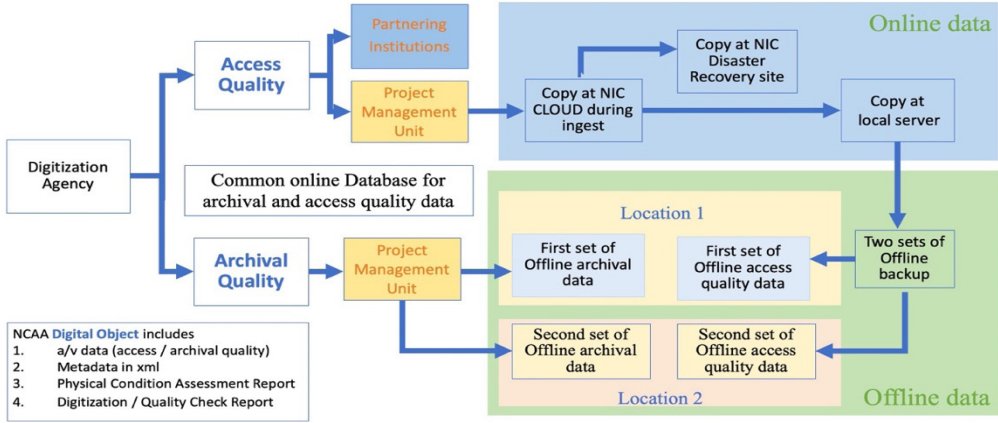
यह राष्ट्रीय सांस्कृतिक दृश्य-श्रव्य अभिलेखागार एक अमूल्य संसाधन के रूप में प्रतिष्ठित हुआ है।

राष्ट्रीय सांस्कृतिक दृश्य-श्रव्य अभिलेखागार की डिजिटल रिपॉजिटरी को सी-डैक, पुणे के डिजिटल संरक्षण उत्कृष्टता केंद्र (Centre for Excellence for Digital Preservation) के सहयोग से विकसित किया गया है जिसे बनाने में विश्व स्तर पर अपनाए जाने वाले ओपन आर्काइवल इंफॉर्मेशन सिस्टम (OAIS) संदर्भ मॉडल का अनुपालन किया गया है। इस रिपॉजिटरी के ऑनलाइन पहुँच को सहज बनाने के लिये आवश्यक आधुनिक संरचना राष्ट्रीय सूचना विज्ञान केन्द्र (National Informatics Centre - NIC) ने उपलब्ध कराया है। डिजिटाइजेशन और मेटाडेटा के मानक, इंटरनेशनल एसोसिएशन ऑफ साउंड एंड दृश्य-श्रव्य आर्काइव्स (International Association of Sound & Audio-visual Archives - IASA) और डबलिन कोर द्वारा निर्धारित मानकों के अनुरूप हैं।

राष्ट्रीय सांस्कृतिक दृश्य-श्रव्य अभिलेखागार का उद्देश्य भारत की मौखिक परंपराओं, प्रदर्शन कलाओं, आनुष्ठानिक और सांस्कृतिक प्रदर्शनों और भारतीय पारंपरिक ज्ञान के दृश्य-श्रव्य संग्रह की दीर्घकालिक उपलब्धता को सुनिश्चित करना है। जैसा कि ऊपर बताया गया है, इसे ISO प्रमाणित ट्रस्टेड डिजिटल रिपॉजिटरी के रूप में प्रमाणित किया गया है। ट्रस्टेड डिजिटल रिपॉजिटरी प्रमाणीकरण के अन्तर्गत इसकी डेटा संग्रहण नीति और आपदा रिकवरी योजना को ऑडिट प्रक्रिया के माध्यम से मंजूरी दी गई।

राष्ट्रीय सांस्कृतिक दृश्य-श्रव्य अभिलेखागार के संपूर्ण डेटा प्रबंधन की योजना को चित्र 2 में दर्शाया गया है। जैसा दर्शाया गया है, डिजिटल डेटा के संग्रहण में अभिलेखीय गुणवत्ता और उपलब्धता गुणवत्ता (संपीडित, छोटे आकार का – जो अभिलेखीय डेटा से बनाया गया) को ध्यान में रखा गया है। जहाँ अभिलेखीय गुणवत्ता आगे के शोध और संपादन के लिए जरूरी है, वहीं उपलब्धता गुणवत्ता ऑनलाइन पहुँच के लिए आवश्यक है।

NCAA Data Management Plan



चित्र 2. NCAA डेटा प्रबंधन योजना

प्रत्येक डिजिटल ऑब्जेक्ट में

- (i) ऑडियो-वीडियो डेटा (अभिलेखीय/उपलब्धता गुणवत्ता),
- (ii) मेटाडेटा एक्सएमएल में,
- (iii) भौतिक स्थिति का मूल्यांकन रिपोर्ट, और
- (iv) डिजिटलीकरण और गुणवत्ता जांच रिपोर्ट शा. मिल है।

ऑनलाइन डेटाबेस से अभिलेखीय डेटा और उपलब्धता गुणवत्ता की जरूरतें पूरी हो जाती हैं।

उपलब्धता गुणवत्ता वाले डेटा को दो समान हार्ड ड्राइव के सेट में रखा जाता है। पहला सेट उस भागीदार संस्थान के पास जमा कर दिया जाता है जिसके पास से यह मूल डेटा आया है। दूसरा सेट परियोजना प्रबंधन इकाई (Project Management Unit - PMU) के पास जमा कर दिया जाता है। परियोजना प्रबंधन इकाई में डेटा को कॉपी किया जाता है और इसे डेटाबेस के साथ समाहित कर दिया जाता है जो राष्ट्रीय सूचना विज्ञान केन्द्र के क्लाउड पर रखा है। जब इसे अधिकृत निदेशक द्वारा एप्लिकेशन के माध्यम से प्रकाशित कर दिया जाता है तो यह जनता के लिए उपलब्ध हो जाता

है। राष्ट्रीय सूचना विज्ञान केन्द्र के क्लाउड की एक डुप्लिकेट प्रति राष्ट्रीय सूचना विज्ञान केन्द्र के डिजास्टर रिकवरी साइट पर रखी जाती है। इसकी एक प्रति इन्दिरा गाँधी राष्ट्रीय कला केन्द्र के स्थानीय सर्वर पर भी रखी गई है, जिसे समय-समय पर राष्ट्रीय सूचना विज्ञान केन्द्र के क्लाउड से अद्यतन किया जाता है। उपलब्धता गुणवत्ता वाले डेटा के उपरोक्त तीन सेट और डिजिटाइजेशन एजेंसी से प्राप्त हार्ड ड्राइव के दो सेट के अलावा, इन्दिरा गाँधी राष्ट्रीय कला केन्द्र के स्थानीय सर्वर पर रखी डेटा से भंडारण के लिए दो ऑफलाइन डेटा सेट (Linear Tape-Open – LTO) एलटीओ टेप में बनाए जाते हैं। ये दो ऑफलाइन डेटा सेट अलग-अलग स्थानों पर जो विभिन्न भूकंपीय क्षेत्रों में हों, रख दिये जाते हैं जैसा चित्ररूप में दर्शाया गया है। अभिलेखागारों के भंडारों का वातावरण नियंत्रित रहता है ताकि टेप खराब न हो जाँय, और वहाँ अधिकारियों की पहुँच भी प्रतिबंधित रहती है।

चूँकि राष्ट्रीय सांस्कृतिक दृश्य-श्रव्य अभिलेखागार को विश्वसनीय डिजिटल रिपॉजिटरी प्रमाणन मिले अभी कुछ ही वर्ष बीते हैं जिसके चलते डेटा संरक्षण से जुड़ी कुछ गतिविधियों को, जैसे एलटीओ टेप को पुराने संस्करण से नए संस्करण में स्थानांतरित

करना, शुरू नहीं किया गया है। अभिलेखागार द्वारा अपनाई गई प्रक्रिया में डेटा पुनःप्राप्ति के लिए पर्याप्त समय हैं। इसलिए, डिजिटल डेटा की डिजास्टर रिकवरी के लिए एक रिपॉजिटरी को आज के समय में ट्रस्टेड डिजिटल रिपॉजिटरी के रूप में विकसित करना ही सबसे अच्छा समाधान है।

निष्कर्ष:

डिजिटल डेटा के नाजुक प्रकृति को ध्यान में रखते हुए उसकी रिपॉजिटरी लिये एक आपदा प्रबंधन योजना का होना अत्यन्त आवश्यक है ताकि उनका दीर्घकालिक संरक्षण किया जा सके। डिजिटल डेटा के दीर्घकालिक संरक्षण की दिशा में 'ट्रस्टेड डिजिटल रिपॉजिटरी' (टीडीआर) एक सुस्थापित प्रक्रिया है। प्रयास यह होना चाहिए कि सभी महत्वपूर्ण डेटा किसी न किसी प्रमाणित ट्रस्टेड डिजिटल रिपॉजिटरी में संरक्षित हो, जो किसी भी आपदा की स्थिति में अपने डेटा भंडार का दीर्घकालीन संरक्षण करने में सक्षम है। चूंकि किसी भी ट्रस्टेड डिजिटल रिपॉजिटरी के विफल होने की संभावना संस्थागत कारणों से हो ही सकती है, इसके चलते एक 'पूरक ट्रस्टेड डिजिटल रिपॉजिटरी' की अवधारणा विकसित की गई है जो डेटा संरक्षण की दृष्टि से मुख्य ट्रस्टेड डिजिटल रिपॉजिटरी का एक विश्वसनीय उत्तराधिकारी के रूप में कार्य कर सके। यह पूरक ट्रस्टेड डिजिटल रिपॉजिटरी मुख्य ट्रस्टेड डिजिटल रिपॉजिटरी के संपूर्ण डेटा की एक प्रतिलिपि अपने यहाँ रखती है। व्यवहारिक दृष्टिकोण से यह अवधारणा विकसित हुई है कि विभिन्न ट्रस्टेड डिजिटल रिपॉजिटरी एक-दूसरे के साथ एक नेटवर्क में जुड़े रहे और साथ साथ वे एक दूसरे के लिये पूरक ट्रस्टेड डिजिटल रिपॉजिटरी का भी रोल निभायें। ट्रस्टेड डिजिटल रिपॉजिटरियों का नेटवर्क ही आज के इतिहास को भविष्य में सुरक्षित रख सकेगा।

शोध पत्र में प्रयुक्त अंग्रेजी शब्दों की समानार्थक हिंदी शब्दावली

Alphabetically sorted terminology in English	वर्णमाला अनुक्रमित हिंदी शब्दावली
Access Quality	एक्सेस क्वालिटी / एक्सेस गुणवत्ता
Archival Quality	अभिलेखीय गुणवत्ता / आर्काईवल क्वालिटी
Digital Archive	डिजिटल आर्काईव्स / डिजिटल लेखागार
Digital Preservation	डिजिटल संरक्षण
Disaster Management	आपदा प्रबंधन
Emulation	एम्यूलेशन, अनुकरण
Encapsulation	एनकैप्सुलेशन, संपुटीकरण
GLAM - Gallery, Library, Archive and Museum	GLAM – दीर्घा, पुस्तकालय, अभिलेखागार एवं संग्रहालय के समूह को निर्देशित करता है, जो सांस्कृतिक विरासतों का संरक्षण व प्रसार करते हैं।
International Association of Sound and Audiovisual Archives (IASA)	इंटरनेशनल एसोसिएशन ऑफ साउंड एंड ऑडियोविजुअल आर्काइव्स (आईएएसए)
Metadata - Information about content	मेटाडेटा
Migration	माइग्रेशन, प्रवजन
National Cultural Audiovisual Archives (NCAA)	नेशनल कल्चरल दृश्य-श्रव्य आर्काइव्स
Offline data centre	ऑफलाइन डेटा सेंटर
Open Archive Information System (OAIS) Reference Model	ओपन-आर्काइव इन्फॉर्मेशन सिस्टम रेफरेंस मॉडल

Open format	ओपेन फॉर्मेट
Open Platform	ओपेन प्लेटफॉर्म
Refreshing	रिफ्रेशिंग
Replication	रिप्लिकेशन, प्रतिकृति
Standards for Digitization	डिजिटलीकरण का मानक
Trusted Digital Repository (TDR)	ट्रस्टेड डिजिटल रिपोजिटरी (टीडीआर)
Trustworthy Digital Repository (TDR)	ट्रस्टवर्दी डिजिटल रिपोजिटरी (टीडीआर)

संदर्भ:

1. Survey Report of World Economic Forum 2019. Available at: <https://www.weforum.org/agenda/2019/04/how-much-data-is-generated-each-day-cf4bddf29f/>
2. [https://en.wikipedia.org/wiki/GLAM_\(cultural_heritage\)](https://en.wikipedia.org/wiki/GLAM_(cultural_heritage))
3. The World Economic Forum (WEF), The Global Risks Report 2023, 18th Edition, https://www3.weforum.org/docs/WEF_Global_Risks_Report_2023.pdf
4. Goodman, Paul. 16 Advantages of Digital Technology. Available at: <https://turbofuture.com/computers/Advantages-of-Digital-Technology>
5. Sinha, Mukul K. Digital repository: audit & certification - study report on digital preservation in India; conducted by National Digital Preservation Program, Government of India. 2008.
6. Lavoie, Brian. The Open Archival Information System (OAIS) Reference model: introductory guide. 2nd ed. UK: Digital Preservation Coalition, 2014.
7. Gladney, Henry M. Preserving digital information. Heidelberg: Springer, 2007. ISBN 978-3540378877
8. The National Cultural Audiovisual Archives <http://ncaa.gov.in/repository/>
9. Standards for Trustworthy Digital Repositories, <http://www.iso16363.org>
10. https://www.cdac.in/index.aspx?id=mc_hc_national_digi_pre
11. <http://www.iso16363.org/first-repository-in-the-world-has-been-awarded-iso-16363-certification/>
12. <http://www.iso16363.org/iso-certification/certified-clients/united-states-government-publishing-office/>
13. Standards for information security management systems, <https://www.iso27001security.com/html/27000.html>