शोधामृत
(कला, मानविकी और सामाजिक विज्ञान की अंतर्राष्ट्रीय सहकर्मी समीक्षित, मूल्यांकित, त्रैमासिक शोध पत्रिका)

ISSN : 3048-9296 (Online)
3049-2890 (Print)
IIFS Impact Factor-5.375
Vol.-3; issue-3 (July-Sept.) 2026
Page No- 163-169
©2026 Shodhaamrit
DOI : 10.71037/Shodhaamrit.v3i3.22
https://shodhaamrit.gyanvividha.com

Author's :
डॉ. देवेन्द्र सिंह राजपूत
सहायक आचार्य, संस्कृत विभाग,
हरियाणा केन्द्रीय विश्वविद्यालय, महेंद्रगढ़.

Corresponding Author :
डॉ. देवेन्द्र सिंह राजपूत
सहायक आचार्य, संस्कृत विभाग,
हरियाणा केन्द्रीय विश्वविद्यालय, महेंद्रगढ़.

पांडुलिपियों के पुनरुद्धार में भाषा प्रौद्योगिकी एवं कृत्रिम बुद्धिमत्ता की भूमिका

शोधसारांश : संस्कृत पाण्डुलिपियां भारत की विशाल ज्ञानराशि, संस्कृति एवं विरासत की अमूल्य धरोहर हैं। इस बौद्धिक संपदा का एक बड़ा भाग भौतिक क्षरण, जलवायु परिवर्तन एवं उपेक्षा के कारण लगातार लुप्त हो रहा है। प्रस्तुत शोधपत्र संस्कृत पांडुलिपियों में निहित ज्ञान के रक्षित कर उसका पुनरुद्धार करने में भाषा प्रौद्योगिकी, विशेष रूप से कृत्रिम बुद्धिमत्ता की अपरिहार्य भूमिका का विश्लेषण करता है। पाण्डुलिपि संरक्षण के लिए प्राचीनकाल में प्रयुक्त होने वाली विधियों का विवेचन करते हुए डिजिटल मानवीकी के युग में प्रचलित मशीन लर्निंग, एआई एवं उन्नत प्रौद्योगिकी आधारित तकनीकों का विवेचन किया गया है। यह शोधपत्र मुख्य रूप से तीन तकनीकी आयामों पर केंद्रित है। प्रथम, डिजिटलीकरण और प्रतिबिम्ब संवर्धन तकनीकी, द्वितीय, स्वतः अक्षर अभिज्ञान (ओ.सी.आर. और तृतीय, प्राकृतिक भाषा प्रसंस्करण (NLP), जिसका उपयोग इन पांडुलिपियों के अनुवाद, व्याकरणिक विश्लेषण और विषय-आधारित वर्गीकरण के लिए किया जा रहा है। इस शोधपत्र में लेखक द्वारा विकसित विज्ञान एवं तकनीकी से सम्बद्ध प्राचीन संस्कृत पाण्डुलिपियों हेतु विकसित ऑनलाइन मैन्यूस्क्रिप्ट्स केटेलॉग सिस्टम (**OMACS)** - https://sanskrit.jnu.ac.in/omacs/index.jsp को भी प्रदर्शित किया गया है। इस शोध पत्र का उद्देश्य यह बतान है कि संस्कृत पांडुलिपियों का संरक्षण केवल अतीत के ज्ञान का संरक्षण नहीं है, बल्कि ए.आई. जैसी भविष्योन्मुखी तकनीकों के माध्यम से प्राचीन प्रज्ञा को आधुनिक विज्ञान के साथ जोड़ना भी है। यह प्रयास न केवल भारत की सांस्कृतिक संप्रभुता को सुदृढ़ करेगा बल्कि वैश्विक ज्ञान परंपरा में एक नया प्रतिमान स्थापित करेगा।

मुख्य शब्द : संस्कृत पांडुलिपियाँ, कृत्रिम बुद्धिमत्ता, डिजिटल संरक्षण, ओसीआर, एनएलपी, सांस्कृतिक विरासत, डेटा माइनिंग।

1. प्रस्तावना : भारत में ज्ञान विज्ञान की परंपरा अनवरत एवं अक्षुण्ण रूप से चली आ रही है। ज्ञान की यह धारा श्रुत परंपरा से आरम्भ होकर पाण्डुलिपियों से होती हुई ग्रन्थों के विविध लिपियों में लेखन, मुद्रण,

प्रिंटिंग से होकर आज डिजिटल माध्यमों के रूप में कृत्रिम बुद्धिमत्ता के दौर में आ पहुंची है। भारतीय मेधा ने सदियों से सम्पूर्ण विश्व को प्रभावित किया है। इसी कडी में ज्ञान के सशक्त संवाहक के रूप में पाण्डुलिपियां रही हैं। श्रुत परंपरा के बाद लेखन कला के विकास के साथ ही ग्रन्थों को एक पीढी से दूसरी पीढी को हस्तांतरित करने में पाण्डुलिपियों की महत्त्वपूर्ण भूमिका रही है। हमारे पूर्वजों द्वारा ज्ञान की अनुपम धरोहर को संरक्षित करने के परिणामस्वरूप पत्तों, वृक्षों की छालों पर लेखन को स्थायित्व प्रदान करने के उद्देश्य से पाण्डुलिपियों के रूप में ग्रन्थों के लेखन एवं संरक्षण का एक महत्त्वपूर्ण प्रयास किया। इन पाण्डुलिपियों के रूप में श्रुति, स्मृति, दर्शन, व्याकरण, साहित्य-काव्यशास्त्र, राजनीतिशास्त्र, कला, गणित, रसायन, आयुर्वेद, भौतिकी, ज्योतिषशास्त्र, औषधिशास्त्र इत्यादि ज्ञान के विविध अनुशासनों से सम्बन्धित ज्ञानराशि आज हमें प्राप्त होती है। भारत में अनेक संग्रहालयों एवं पुस्तकालयों में ताड़पत्र, भोजपत्र, लकड़ी, ताम्रपत्र, पत्थर, कागज एवं चमड़े पर लिखित पांडुलिपियों का विशाल संग्रह विद्यमान है, जिनका प्राकृतिक और जैविक क्षरण लगातार बढ़ रहा है। जिसके कारण प्रतिदिन अनेक पाण्डुलिपियां नष्ट हो रही हैं। वर्तमान समय में डिजिटल मानवीकी के क्षेत्र में हो रहे नवाचारों एवं नई तकनीकियों के निर्माण से ग्रन्थ एवं पाठ के संरक्षण की अनेक अत्याधुनिक तकनीकियों का विकास हो चुका है। भारतीय भाषाओं में प्रौद्योगिकी के विकास से सम्बन्धित अनेक संस्थाएं एवं अनके उपक्रम कार्यरत हैं। ऐसे समय में पाठ-संरक्षण तकनीकियों का उपयोग कर पाण्डुलिपियों के संरक्षण की महती आवाश्यकता है। ए. आई. आधारित उन्नत प्रौद्योगिकी ज्ञान के संरक्षण एवं बहुआयामी माध्यमों से प्रस्तुतिकरण में अत्यन्त सहायक है। प्रस्तुत शोधपत्र में संस्कृत पाण्डुलिपियों के पुनरुद्धार के समक्ष आ रही समस्याओं को चिह्नित करते हुए उनके डिजिटल माध्यमों से पुनरुद्धार हेतु प्रयुक्त की जाने वाली तकनीकों एवं कृत्रिम बुद्धिमत्ता आधारित प्रविधियों पर चर्चा परिचर्चा की गई है।

2. पाण्डुलिपि से आशय : "पाण्डुलिपि" शब्दः संस्कृत भाषा में "पाण्डु" (पाण्डु = पीला रंग, हल्का रंग) इस शब्द से तथा "लिपि" (लिखना) इन दो शब्दों के योग से बना है।

" पाण्डुवर्णे पत्रे लिखितं ग्रन्थरुपं लेख्यम्। "पाण्डुलेखिता लिपिः।"[1]

अर्थात् - पीले रंग के पत्र पर लिखना चाहिए। "***हस्तलिखितं पुस्तकम्। मुद्रितं न।"***[2] अर्थात् – जो पुस्तक हाथ से लिखित है, मुद्रित नहीं है वह पाण्डुलिपि है।

आधुनिकपरिभाषा - "पांडुलिपि वह हस्तलिखित दस्तावेज़ है जिसमें पाठ्य अथवा चित्रात्मक सामग्री निहित होती है, और जिसे पारंपरिक लिपियों का प्रयोग करते हुए ताड़पत्र, भोजपत्र, हस्तनिर्मित कागज या वस्त्र जैसी सामग्रियों पर लिखा जाता है।"[3]

"manuscriptum" शब्द मध्यकालीन लैटिन भाषा से उद्भूत है। पांडुलिपि का आशय हस्तलिखित पुस्तक अथवा किसी ग्रंथ की हस्तलिखित प्रतिलिपि से है। साहित्य एवं प्राचीन अभिलेखीय परंपरा में लेखन–सामग्रियों के संदर्भ-जैसे पाषाण, ताम्रपत्र, काष्ठफलकों, ताड़पत्रों, पत्तों, चर्मपत्रों तथा कागज के पृष्ठों-का उल्लेख व्यापक रूप से मिलता है। पारंपरिक रूप से मौखिक शिक्षण-पद्धति (गुरुकुल प्रणाली) ने दीर्घ अवधि तक संस्कृत-विद्या के संवहन और संरक्षण की भूमिका निभाई; इसके उपरांत विभिन्न लेखन–सामग्रियों के प्रयोग से लेखन–परंपरा का संगठित विकास प्रारंभ हुआ।

पांडुलिपि-लेखन के कौशल का उल्लेख अथर्ववेद तथा तैत्तिरीय संहिता जैसे वैदिक ग्रंथों में प्राप्त होता है। प्रारंभिक काल में इन ग्रंथों को विविध सामग्रियों—जैसे वर्तिका (छाल/वल्कल), ताड़पत्र, केले के पत्ते, काष्ठफलकों तथा ताम्रपत्र पर लिपिबद्ध किया जाता था। आज संस्कृत पांडुलिपियों की अनुमानित संख्या लगभग पाँच मिलियन मानी जाती है, जिनमें से लगभग 67 प्रतिशत पांडुलिपियाँ संस्कृत ग्रंथों की हैं (गौर एवं चक्रवर्ती के अनुसार)।

3. संस्कृत पांडुलिपियों के संरक्षण की आवश्यकता : संस्कृत पांडुलिपियों का संरक्षण भारतीय ज्ञान–परंपरा, सांस्कृतिक विरासत तथा बौद्धिक इतिहास के संरक्षण हेतु अत्यंत अनिवार्य है। पांडुलिपियाँ न केवल प्राचीन भारत की वैज्ञानिक, दार्शनिक एवं सांस्कृतिक चेतना की मूल वाहक हैं, अपितु ज्ञान के प्राथमिक स्रोत होने के कारण उनका संरक्षण राष्ट्रीय दायित्व भी है। संस्कृत पांडुलिपियाँ बहु–पीढ़ियों में संचित उस ज्ञान–परंपरा का प्रतिनिधित्व करती हैं, जिसे सदियों तक मौखिक और लिखित दोनों माध्यमों से संरक्षित किया गया। पांडुलिपियों के नाश का अर्थ केवल ग्रंथों का विनाश नहीं है बल्कि उन ज्ञान–तत्त्वों का लोप भी है, जो संस्कृति, दर्शन, कला, चिकित्सा, समाज–विज्ञान तथा वैज्ञानिक चिंतन के ऐतिहासिक विकास को समझने के लिए अत्यंत महत्वपूर्ण हैं।

वृहत्तर परिप्रेक्ष्य में देखा जाए तो पांडुलिपियाँ किसी भी राष्ट्र की बौद्धिक संपदा का साकार स्वरुप हैं। वैश्वीकरण के वर्तमान युग में जब विश्वभर के राष्ट्र बौद्धिक संपदा अधिकार के संरक्षण को प्राथमिकता दे रहे हैं, तब संस्कृत पांडुलिपियों का संरक्षण भारत के बौद्धिक स्वामित्व, वैज्ञानिक परंपरा तथा सांस्कृतिक अस्मिता की रक्षा के लिए अत्यंत आवश्यक हो जाता है।

4. पांडुलिपि संरक्षण की पारंपरिक–पद्धतियाँ : भारतीय परंपरा में पांडुलिपियों के संरक्षण के लिए विभिन्न परंपरागत उपायों का प्रयोग किया जाता रहा है। इनमें - धूप–स्वेद, वस्त्रावरण, लकड़ी के बक्सों में सुरक्षित रखना, भूमिगत कोषों में संग्रहण, भाद्रपद मास में विशेष धूप–प्रयोग तथा विजयदशमी के अवसर पर प्रदीपन जैसे उपाय प्रमुख रुप से प्रचलन में थे। ये पद्धतियाँ अल्प व्यय के साथ स्वास्थ्य की दृष्टि से सुरक्षित तथा स्थानीय परिस्थितियों के अनुरुप थीं।

5. पाण्डुलिपि संरक्षण की आधुनिक प्रौद्योगिकी आधारित पद्धतियाँ : प्राचीन पांडुलिपियाँ रखरखाव एवं अन्य प्राकृतिक कारणों से नष्ट होती जा रही हैं, अतः उनके स्थायी संरक्षण के उपायों में डिजिटल तकनीकी की महती भूमिका सिद्ध हो सकती है। वर्तमान समय में प्राचीन ग्रन्थों, पाठों एवं दस्तावेजों के संरक्षण हेतु प्रचलित अनेक तकनीकों एवं प्रौद्योगिकी में से पाण्डुलिपियों के संरक्षण हेतु निम्नलिखित प्रसिद्ध हैं -

5.1 उच्च गुणवत्ता वाला डिजिटल छायांकन - पांडुलिपि-पृष्ठ का ऐसा डिजिटल अभिलेखन जिसमें सूक्ष्मतम अक्षर-रूप, रेखांकन, कागज अथवा ताड़पत्र की बनावट, वर्ण–भिन्नता तथा भौतिक क्षरण के सभी सूक्ष्म संकेत अत्यधिक स्पष्टता और विस्तृत स्तर पर दृश्य हों। इसके लिए निम्नलिखित उपकरण प्रयोग में लिए जाते हैं -

- उच्च रेजोल्यूशन चित्रग्राहक (कैमरा), नियंत्रित प्रकाश, स्थिर सतह
- TIFF/RAW फॉर्मेट
- रंग–सुधार व स्पष्टता बढ़ाने हेतु सॉफ्टवेयर

इस हेतु अपनाए जाने वाली प्रक्रिया अधोलिखित है -

- पाण्डुलिपियों का पूर्व-सज्जीकरण — पाण्डुलिपि को धूल, गंदगी तथा सूक्ष्मजीवों से मुक्त करने हेतु प्रारम्भिक सफाई एवं कीटनाशन किया जाता है।
- स्थिरीकरण — पाण्डुलिपि को एक समतल एवं स्थिर सतह पर इस प्रकार स्थापित किया जाता है कि पृष्ठों का कोण अथवा स्थिति परिवर्तित न हो।
- प्रकाश व्यवस्था — छायांकन के लिए समान, नियंत्रित तथा कोण-परिवर्तन-रहित प्रकाश स्रोत का उपयोग किया जाता है।
- छायाचित्रण — उच्च-गुणवत्तायुक्त इमेजिंग उपकरणों द्वारा पाण्डुलिपि के प्रत्येक पृष्ठ का उच्च-रिजॉल्यूशन छायाचित्र लिया जाता है।
- डिजिटल फाइलों का संग्रहण — प्राप्त चित्रों को प्रायः RAW अथवा TIFF जैसे अवसंपीडित (lossless) प्रारूपों में संग्रहित किया जाता है।

- प्रसंस्करण — छवियों का रंग-संशोधन, स्पष्टता-वृद्धि, विकृति-निवारण तथा आवश्यक सीमा-समायोजन किया जाता है।
- प्रतिलिपियों का सृजन — संग्रहण, अभिलेखीकरण अथवा क्लाउड एकीकरण के लिए डिजिटल प्रतिलिपियाँ तैयार की जाती हैं।

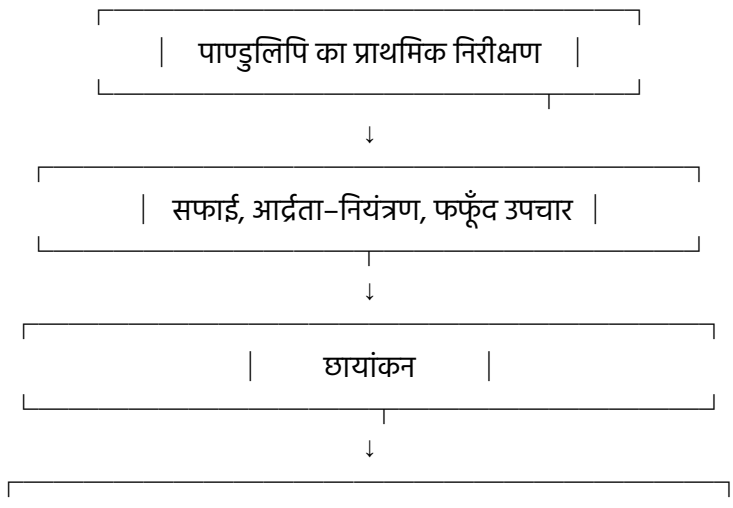
उच्च-रिजॉल्यूशन छवियों में छोटे अक्षर एवं चित्र-विवरण स्पष्ट रूप दिखाई देते हैं। बहु-वर्णीय प्रतिबिंबन के प्रयोग से मैले, कटे-फटे, प्रकाश में अस्पष्ट वर्ण भी स्पष्ट दिखाई देते हैं। परिशोधित चित्रों के विश्लेषण के लिए विविध डिजिटल उपकरणों एवं शोध-सॉफ्टवेयरों का उपयोग किया जाता है। ब्रिटिश पुस्तकालय सहित अनेक अंतरराष्ट्रीय अभिलेखागार प्राचीन हस्तलिखित पाण्डुलिपियों के ऐसे उच्च-रिजॉल्यूशन डिजिटल प्रतिरूप उपलब्ध कराते हैं।

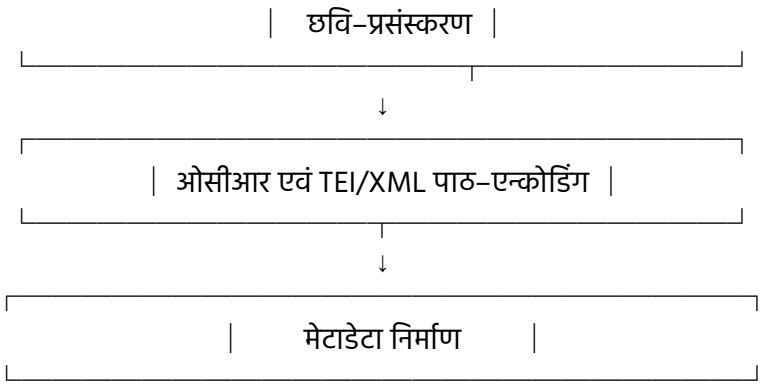
5.2 बहु-वर्णीय प्रतिबिंबन - बहु-वर्णीय छायाचित्रण (मल्टीस्पेक्ट्रल इमेजिंग) में पाण्डुलिपि के पृष्ठों को विभिन्न प्रकाश-तरंगदैर्घ्यों (यू.वी., आई.आर. तथा दृश्य प्रकाश) में चित्रित किया जाता है। इस प्रक्रिया से धुंधले, लुप्त, फफूँदग्रस्त या अस्पष्ट वर्ण भी स्पष्ट रूप से दिखाई देते हैं। बहु-वर्णीय छायाचित्रण द्वारा विशिष्ट प्रकाश स्रोत और उन्नत छवि-प्रसंस्करण सॉफ्टवेयर द्वारा पाण्डुलिपि का विभिन्न तरंगदैर्घ्यों में क्रमिक छायांकन किया जाता है। तत्पश्चात् प्राप्त छवियों का विश्लेषण, तुलनात्मक अध्ययन और डिजिटल संयोजन किया जाता है जिससे क्षतिग्रस्त लेख पुनः पठनीय हो जाते हैं।

5.3 स्वतः अक्षर पहचान - डिजिटल छवियों से पाठ को स्वतः पढ़ने हेतु स्वतः अक्षर-पहचान तन्त्र (**ओसीआर)** तकनीक का बृहद् स्तर पर प्रयोग किया जा रहा है। संस्कृत के प्राचीन लिपि-रूपों को पठनीय बनाने के लिए इस तकनीक के माध्यम से अनेक प्रयास किए गए हैं। इस दिशा में 'गूगल इंडिक ओसीआर' परियोजना सक्रिय है।[4] इसके अतिरिक्त आई. आई.आई.टी. हैदराबाद में भी संस्कृत पाण्डुलिपियों के लिए ओसीआर प्रणालियों के विकास से संबंधित विशेष अध्ययन संचालित हो रहे हैं। यह परियोजना विशेष रूप से प्राचीन हस्तलिखित लिपियों मुख्यतः देवनागरी एवं अन्य भारतीय लिपियों को यथासंभव सटीक रूप से मशीन-पठनीय बनाने का लक्ष्य रखती है।[5]

5.4 डिजिटल संग्रहालय बनाना – भारतीय राष्ट्रीय डिजिटल पुस्तकालय[6], इन्दिरा गांधी राष्ट्रिय कला केन्द्र का कलानिधि संभाग[7], राष्ट्रीय पुस्तकालय आदि संस्थाओं में पांडुलिपियों का डिजिटल कैटलॉग तैयार किया जा रहा है।

6. डिजिटल प्रसंस्करण की चरणबद्ध प्रक्रिया : इस प्रक्रिया को निम्नलिखित फ्लोचार्ट द्वारा क्रमबद्ध रूप से समझा जा सकता है -





7. भारत में संस्कृत पाण्डुलिपियों के डिजिटल संरक्षण हेतु प्रमुख प्रयास

- **भारतीय पुरातत्व सर्वेक्षण (ASI)** द्वारा *National Mission on Monuments and Antiquities (NMMA)* के अंतर्गत 1,18,359 पुरावशेषों एवं कलावस्तुओं का डिजिटल अभिलेखीकरण किया गया है। यह पहल भारतीय विरासत-सामग्री के दीर्घकालिक डिजिटल संरक्षण की दिशा में एक महत्त्वपूर्ण योगदान है। (https://nmma.nic.in/nmma/indexAction.do and https://www.pib.gov.in/PressReleasePage.aspx?PRID=2157465)
- **ज्ञान भारतम् मिशन** के अंतर्गत भारतीय पाण्डुलिपि-संपदा के डिजिटलीकरण एवं संरक्षण को प्रमुख लक्ष्य बनाकर अब तक *3,50,000 से अधिक पाण्डुलिपियों* का डिजिटलीकरण 37 सार्वजनिक एवं निजी संस्थानों के सहयोग से सम्पन्न किया गया है। यह मिशन राष्ट्रीय स्तर पर पाण्डुलिपि-धरोहर के एकीकृत डिजिटलीकरण के लिए महत्वपूर्ण मॉडल प्रस्तुत करता है। (https://gyanbharatam.com/)
- **संस्कृति मंत्रालय** द्वारा **सी-डैक, पुणे** के तकनीकी सहयोग से *"**जतन**"* नामक सॉफ्टवेयर का उपयोग संग्रहालयीय वस्तुओं के डिजिटलीकरण और डिजिटल अभिगम्यता हेतु किया जा रहा है। यह सॉफ्टवेयर बहुविध संग्रहालय संग्रहीत वस्तुओं के मानकीकृत डिजिटल दस्तावेज़ीकरण के लिए विकसित किया गया है। (https://www.cdac.in/index.aspx?id=mc_hc_virtual_museums)
- **इंदिरा गांधी राष्ट्रीय कला केंद्र (IGNCA)** के पास संस्कृत सहित अनेक भारतीय भाषाओं की पाण्डुलिपियों के विशाल डिजिटल भंडार उपलब्ध हैं। इसके अतिरिक्त, संस्थान द्वारा विस्तृत ऑनलाइन सूची-ग्रन्थ (catalogues) भी प्रदान किए जाते हैं, जो शोधकर्ताओं हेतु महत्वपूर्ण संदर्भ-सामग्री हैं। (https://www.ignca.gov.in/)
- **राष्ट्रीय अभिलेखागार (National Archives of India)** ने *विज्ञान एवं प्रौद्योगिकी विभाग* तथा *नेशनल इंस्टीट्यूट ऑफ एडवांस्ड स्टडीज़ (NIAS), बेंगलुरु* के सहयोग से *रामायण, भगवद्गीता तथा महाभारत* की दुर्लभ पाण्डुलिपियों के डिजिटलीकरण का एक महत्त्वपूर्ण परियोजना-कार्य प्रारंभ किया है। यह भारतीय महाकाव्य परंपरा के संरक्षण हेतु एक विशेष प्रयास है।
- **राष्ट्रीय पुस्तकालय, कोलकाता** ने भारत की बौद्धिक एवं सांस्कृतिक विरासत के संरक्षण के उद्देश्य से दुर्लभ पाण्डुलिपियों के संग्रहण, संरक्षण एवं डिजिटलीकरण का कार्यक्रम प्रारम्भ किया है। *"Down the Memory Lane"* नामक डिजिटलीकरण पहल के माध्यम से इन अमूल्य दस्तावेजों का दीर्घकालिक डिजिटल संरक्षण किया जा रहा है।

8. संरक्षण-प्रौद्योगिकियों का विश्लेषण : वर्तमान में प्रचलित डिजिटल संरक्षण प्रौद्योगिकियों का व्यवस्थित मूल्यांकन यह संकेत करता है कि संस्कृत पाण्डुलिपियों के संरक्षण एवं डिजिटलीकरण के संदर्भ में दक्षता, शुद्धता (accuracy), तथा प्रयोज्यता के स्तर पर पर्याप्त विविधताएँ विद्यमान हैं। ए.आई. आधारित प्रौद्योगिकियों

में PLM-SegFormer तकनीकी सीमेन्टिक सेगमेंटेशन कार्यप्रणाली पर आधारित है जो कि IGNCA में पायलट प्रोजेक्ट के तौर पर प्रयोग में ली जा रही है। ज्ञान-सेतु ए.आई. चैलेंज के अन्तर्गत एन्सेम्बल पद्धति के प्रयोग से पाण्डुलिपियों के डिजिटल संरक्षण के साथ कृत्रिम मेधा आधारित तकनीकी से पठन एवं प्रसंस्करण का कार्य किया जा रहा है एवं यह प्रौद्योगिकी अभी विकास के चरण में है।

डिजिटल संरक्षण में ओ.सी.आर. एवं ए.आई. प्रणालियों की सटीकता मुख्यतः लिपि की जटिलता, पाण्डुलिपि की भौतिक स्थिति तथा प्रशिक्षण-डेटा की गुणवत्ता पर निर्भर करती है। देवनागरी के संयुक्ताक्षर, मात्राएँ एवं पुरानी वर्तनी ओ.सी.आर. की शुद्धता को लगभग 10–15%, जबकि क्षतिग्रस्त या धुंधली पाण्डुलिपियाँ इसे 15–30% तक प्रभावित करती हैं। इसके विपरीत, डोमेन-विशिष्ट प्रशिक्षण डेटा ए.आई. मॉडलों की शुद्धता में लगभग 8–12% सुधार कर सकता है।

9. डिजिटल संरक्षण से सम्बंधित चुनौतियाँ

- लिपि की विविधता
- ओसीआर की कम सटीकता
- विशेषज्ञता की कमी
- धन और तकनीकी अवसंरचना का अभाव
- मानकीकरण का अभाव

10. निष्कर्ष : शोध से स्पष्ट है कि संस्कृत पाण्डुलिपियों का डिजिटल संरक्षण सांस्कृतिक दायित्व के साथ वैज्ञानिक आवश्यकता भी है। डिजिटल मानविकी, बहु-वर्णीय प्रतिबिम्बन, ए.आई.-आधारित ओ.सी.आर., सेमान्टिक मेटाडेटा तथा सहयोगी मंचों का समन्वित उपयोग अधिक प्रभावी सिद्ध होता है। स्वचालित प्रणालियों की तुलना में मानव–ए.आई. मिश्रित मॉडल अधिक शुद्धता प्रदान करते हैं। यद्यपि विभिन्न राष्ट्रीय पहलों से व्यापक डिजिटलीकरण हुआ है, फिर भी मानकीकरण एवं तकनीकी अवसंरचना की चुनौतियाँ बनी हुई हैं। अतः उच्च गुणवत्ता, लागत-प्रबंधन और अंतर-संचालन के लिए मानव–ए.आई मिश्रित कार्यप्रवाह तथा मानकीकृत मेटाडेटा को अपनाना आवश्यक है।

संदर्भ सूची :

1. *शब्दकल्पद्रुमः,* खण्डः 5, पृष्ठ - 306
2. *वाचस्पत्यम्,* खण्डः 5
3. National Mission for Manuscripts, Govt. of India – www.namami.gov.in
4. GitHub Repository: https://github.com/google/indic_ocr
5. IIIT Hyderabad OCR Research Portal: https://research.iiit.ac.in/projects/
6. https://ndl.iitkgp.ac.in/
7. https://ignca.gov.in/online-digital-resources/manuscripts/

संदर्भग्रन्थ सूची :

1. Gaur, R. C. (2009). Preservation and access to Indian Mss. *AAMES Newsletter, 6* (2).
2. Jha, G. N. (2009, December). E-learning & multimedia for Sanskrit. Paper presented at the Workshop on *E-learning & Multimedia for Education*, Special Centre for Sanskrit Studies, New Delhi.
3. Jha, P. (2007). Digitization and documentation of Mss and other cultural heritage at the IGNCA. *Kriti Rakshana, 8*. National Mission for Mss.
4. Krishna, A., Majumder, B. P., Bhat, R. S., & Goyal, P. (2018). Upcycle your ओसीआर: Reusing ओसीआरs for post-ओसीआर text correction in romanised Sanskrit. In

Proceedings of the 22nd Conference on Computational Natural Language Learning (CoNLL 2018) (pp. 345–355). Association for Computational Linguistics. https://doi.org/10.18653/v1/K18-1034

5. *Kriti Rakshana: A bi-monthly publication of the National Mission for Mss.* National Mission for Mss.
6. Ministry of Culture, Government of India. (2023, December 7). *National Mission for Mss documented 5.2 million Mss and conserved 90 million folios across the country* [Press release]. Press Information Bureau, Delhi. Retrieved from https://www.pib.gov.in/PressReleseDetए.आई.lm.aspx?PRID=1983483
7. Panda, G. P. (2007). Kritisampada: The national database of Mss. *Kriti Rakshana, 8*. National Mission for Mss.
8. Rajput, D. S. (2015). *Multimedia technology for Sanskrit heritage texts: A case of Śrīmadbhagavadgītā*. Lambert Academic Publishing.
9. Rajput, D. S. (2018). *Creating an online catalogue system for science and technology Mss of Sanskrit: Issues and challenges* (Doctoral dissertation, Jawaharlal Nehru University, School of Sanskrit and Indic Studies).
10. Tripathi, D. S. (Ed.). (n.d.). *Guidelines for digitization of archival material.* National Mission for Mss. https://namami.gov.in/sites/default/files/digitization.pdf
11. Tyler, N. (2025, May 5). *ओसीआर options – Sanskrit and Tech with Tyler (Kalpataru Diaries blog)*. Retrieved from https://www.tylerneill.info/blog-kalpataru-diaries/ओसीआर-options
12. https://doi.org/10.1186/s40494-023-01125-w

Web Resources :

1. Archaeological Survey of India - National Mission on Monuments and Antiquities. https://nmma.nic.in/nmma/indexAction.do
2. Digital Library of India. http://www.new.dli.ernet.in
3. Gyan Bharatam Mission Portal. https://gyanbharatam.com/
4. Indira Gandhi National Centre for the Arts - Digital Library. https://www.ignca.gov.in/
5. Khuda Bakhsh Oriental Public Library - Online Catalogue. http://kblibrary.bih.nic.in/
6. Ministry of Culture, Government of India - Mss Division. https://indiaculture.nic.in/Mss
7. National Archives of India. https://www.nationalarchives.gov.in
8. National Mission for Mss. https://www.namami.gov.in
9. Online Mss Catalogue System (OMACS). http://sanskrit.jnu.ac.in/omacs
10. Press Information Bureau, Government of India. (2024). *Digital Documentation of Antiquities*. https://www.pib.gov.in/PressReleasePage.aspx?PRID=2157465

•