

हिन्दी AI पुस्तकालय संस्करण

एआई प्राचीन लिपियों को पढ़ता है

एआई से पुनर्जीवित प्राचीन
अभिलेख

अधिवक्ता किम
क्यूंग-जिन

kimkj.com · AI पुस्तकालय · 2026

एआई प्राचीन लपियों को पढ़ता है

एआई से पुनर्जीवित प्राचीन अभिलेख

अधिवक्ता कमि क्यूंग-जनि

आग में जलकर कोयला बने ग्रंथ-चर्मपत्र, कीचड़ में दबे लकड़ी के टुकड़े और टूटकर बखिरी मट्टि की पट्टिकाएँ लंबे समय तक ऐसी वस्तुएँ थीं जिन्हें पढ़ा नहीं जा सकता था। यह पुस्तक उस सन्नाटे को तोड़ने वाली बारह घटनाओं की पड़ताल करती है। इसमें ज्वालामुखी की राख में दबे हर्कुलेनियम पपीरस को बना खोले पढ़ने वाला आभासी अनावरण, X-ray से पुनर्जीवित एन-गेदी की लैव्यवस्था, लखावट के आधार पर लपिकों की संख्या तय करने वाला यशायाह ग्रंथ-चर्मपत्र का विश्लेषण, दसियों हजार टुकड़ों को जोड़ने वाला डिजिटल गोद, गलिगमेश महाकाव्य का पुनर्निर्माण, अक्कादी भाषा के बही-खातों का अनुवाद, ओरेकल बोन लपिका 3D आभासी संयोजन, शलिला की काष्ठ पट्टिकाओं की घसीट लपिका वाचन, Linear A, सधु लपि और Rongorongo का अध्ययन करने वाला संगणकीय भाषावर्ज्ञान, टूटे हुए यूनानी शलिलेखों के छूटे अंश भरने वाले भाषा मॉडल, और मूल प्रतनिष्ठ होने पर भी सुरक्षित रहने वाले बहुवर्णिक इमेजिंग अभिलेखागार तक के विषय शामिल हैं। प्रत्येक अध्याय को घटना फाइल, वैज्ञानिक जाँच दल की पड़ताल और सामने आए सच के क्रम में व्यवस्थित किया गया है। बिना किसी पूर्व ज्ञान के भी इसे आसानी से पढ़ा जा सके, इसलिए पारभाषिक शब्दों की सरल व्याख्या की गई है, और प्रत्येक उप-अध्याय में जाँचे-परखे संदर्भ जोड़े गए हैं। परिशिष्ट में ऐसे सार्वजनिक स्रोत लिके एकत्र किए गए हैं जिन्हें पाठक घर बैठे सीधे खोलकर देख सकते हैं। यह पुस्तक NotebookLM नोट 'एआई आधारित प्राचीन भाषा एवं प्राचीन दस्तावेज़ वसिकेतन शोध संग्रह (2024-2026)' में संकलित 251 सामग्रियों पर आधारित है।

सार्वजनिक शोध सामग्री अभिलेखागार

यह पुस्तक एआई द्वारा संकलित और तैयार की गई है। इसमें दुनिया भर की प्रमुख भाषाओं में उपलब्ध सामग्री को एकत्र करके संश्लेषित किया गया है। NotebookLM में प्रश्न पूछने की भाषा पर कोई प्रतिबंध नहीं है, भले ही स्रोत सामग्री किसी भी भाषा में लिखी गई हो। पाठक इस पुस्तक को तैयार करने में इस्तेमाल किए गए उसी शोध अभिलेखागार से अपनी भाषा में प्रश्न पूछ सकते हैं। भले ही कुछ सामग्री कोरियाई, अंग्रेज़ी या अन्य भाषाओं में हो, तब भी यह सुविधा उपलब्ध है। NotebookLM में सेटिंग्स खोलकर अपनी मनपसंद आउटपुट भाषा चुनने पर सभी उत्तर और परिणाम उसी भाषा में प्राप्त होते हैं।

इस पुस्तक को तैयार करने में इस्तेमाल किया गया सार्वजनिक NotebookLM अभिलेखागार (एआई आधारित प्राचीन भाषा एवं प्राचीन दस्तावेज़ वसिकेतन शोध संग्रह 2024-2026, 251 सामग्रियों):

<https://notebook.google.com/notebook/cf32a6a1-66cc-473d-99d8-ed5a51596eb3>

इस पुस्तक का मूल कोरियाई संस्करण kimkj.com के एआई पुस्तकालय में उपलब्ध है:

<https://kimkj.com/ai-library/?mod=document&uid=12641>

वषिय-सूची

प्रत्येक अध्याय के शीर्षक पर क्लिक करके संबंधित पाठ पर जाया जा सकता है। शुरुआत से पढ़ने के लिए अध्याय 1 चुनें।

अध्याय 1. ज्वालामुखी की राख में समाया दार्शनिक का पुस्तकालय: हर्कुलेनियम के कार्बनीकृत ग्रंथ-चर्मपत्र • 4

अध्याय 2. जली हुई मंजूषा से निकले वचन: एन-गेदी की मुहरबंद लैव्यवस्था • 11

अध्याय 3. यशायाह चर्मपत्र में छपी दो छायाएँ: मुंशी प्रमाणीकरण • 18

अध्याय 4. मट्टी और चर्मपत्र के हजारों टुकड़ों की पहली जोड़ने वाला डिजिटल गोद • 25

अध्याय 5. मट्टी की पट्टिकाओं के टूटे टुकड़ों को जोड़ना: गलिगमेश महाकाव्य का संगणकीय पुनर्निर्माण • 33

अध्याय 6. प्राचीन मेसोपोटामिया की रसीदे और व्यापारियों की फुसफुसाहट: AI अनुवाद • 41

अध्याय 7. कछुए के उदर-कवच में सोया साम्राज्य का भविष्य: ओरेकल बोन लपि और 3D आभासी संयोजन • 49

अध्याय 8. कीचड़ के गड्ढे से पुनर्जीवित सुलेख: सलिला की काष्ठ पट्टिकाएँ और कर्सवि लपि पहचान एआई • 57

अध्याय 9. एजियन सागर की खोई पहली: लीनयिर ए और ज्यामितीय ग्रिड • 65

अध्याय 10. सूर्य चक्र और पशु मुहरों का कूट तोड़ना: सधु लपि और रोगोरोन्गो की ध्वनियाँ • 73

अध्याय 11. एआई ने टूटे हुए ग्रीक शिलालेखों को फरि से पढ़ा: फाउंडेशन मॉडल और Ithaca • 81

अध्याय 12. कभी न जलने वाला आभासी पुस्तकालय: अत-उच्च-रिज़ॉल्यूशन बहुवर्णक पुरालेख और भविष्य • 89

परशिष्ट. मेरी अपनी AI पुरातत्व यात्रा: आम जनता के लिए वैश्विक अन्वेषण पोर्टल • 97

अध्याय 1. ज्वालामुखी की राख में समाया दार्शनिक का पुस्तकालय: हर्कुलेनियम के कार्बनीकृत ग्रंथ-चर्मपत्र

1.1. घटना की फाइल

जसि दनि ज्वालामुखी ने पुस्तकालय को ढक दिया

ईस्वी 79 में दक्षिणी इटली का वसुवियस ज्वालामुखी फट पड़ा। ज्वालामुखी ने एक साथ गर्म गैसें और राख बाहर निकाली। पहाड़ के नीचे समुद्र तट पर हर्कुलेनियम नाम का एक शहर था। यह रोम के धनी लोगों का पसंदीदा ग्रीष्मकालीन वशिराम स्थल था। यह शहर भी एक ही दनि में ज़मीन के नीचे दफ़न हो गया[1]।

शहर के बाहर एक बड़ी हवेली स्थिति थी। आज वदिवान इसे पपीरस वलिा कहते हैं। हवेली का मालिक संभवतः लूसियस कैल्पर्नियस पसिो कैसोनिस था। वह जूलियस सीज़र का ससुर था। इस घर में यूनानी दार्शनिक फिलोडेमस आते-जाते थे। फिलोडेमस एपिक्यूरस संप्रदाय के वदिवान थे। हवेली के एक हस्तिसे में ग्रंथ-चर्मपत्रों को सहेज कर रखा गया एक पुस्तकालय था। यहाँ उनके लखे दस्तावेज़ असामान्य रूप से अधिक मिले हैं, इसलिए इसे उनसे जुड़ा ग्रंथालय माना जाता है[1]।

पपीरस प्राचीन भूमध्यसागरीय दुनिया का कागज़ है। इसे मसिर में नील नदी के किनारे उगने वाले पौधे से बनाया जाता है। तने के भीतरी गूदे को पतली पट्टियों में काटकर आड़ा बछिया जाता है। उसके ऊपर फिर से खड़ी दशा में एक और परत रखी जाती है। पानी छड़िक कर दबाने पर पौधे का रस गोंद का काम करता है। सूखने पर यह कागज़ की एक शीट बन जाता है। इसलिए पपीरस की सतह पर आड़े और खड़े रेशों का परस्पर जाल रह जाता है। यह बनावट आगे चलकर एक महत्वपूर्ण सुराग बनती है[7]।

उस दनि पुस्तकालय में गर्म गैसें घुस आईं। ग्रंथ-चर्मपत्र बनिा लपटों के ही सीधे भुन गए। कागज़ का काम करने वाला पपीरस काले कोयले में बदल गया। इस तरह जलकर कोयला बनने को कार्बनीकरण कहा जाता है। ग्रंथ-चर्मपत्र सकिड़ गए और उनकी सतह फूल गई[2]। उनके ऊपर 20 मीटर से अधिक ज्वालामुखी तलछट जमा हो गई। ज्वालामुखी तलछट राख और पत्थरों के चूर्ण के जमने से बनी सख्त परत होती है[1]।

ज़मीन के नीचे गहराई में दब जाने से उन तक हवा नहीं पहुँची। यदहिवा न पहुँचे, तो पौधों के रेशे आसानी से नहीं सड़ते। आपदा ने पुरावशेषों की रक्षा करने वाले ढक्कन का काम किया। ग्रंथ-चर्मपत्र इसी तरह 2,000 वर्षों तक टक़े रहे[2]।

भूमगित सुरंग से निकली काली छड़े

1738 में नेपल्स के शाही परिवार ने इस क्षेत्र की खुदाई शुरू की। 1750 में एक कुआँ खोद रहे मज़दूर को रंग-बरिगा मोज़ेक फ़र्श मिला। इस तरह हवेली का स्थान सामने आया। उत्खनन का ज़मिमा कार्ल वेबर ने संभाला। वह स्विट्ज़रलैंड के सैन्य इंजीनियर थे[1]।

वेबर ने सख्त ज्वालामुखी तलछट में सीधे नीचे की ओर एक सुरंग खोदी। ऐसी सुरंग को ऊर्ध्वाधर शाफ़्ट कहा जाता है। फिर उन्होंने भूलभुलैया की तरह संकरी सुरंगें खोदीं। सुरंग के भीतर अंधेरा था और हवा दूषित थी। दीवारें ढहने का खतरा भी हमेशा बना रहता था। ऐसी परस्थितियों में भी वेबर ने हवेली का फ़्लोर प्लान बड़ी बारीकी से बनाया। उस मानचित्र की बदौलत आज हम जान सकते हैं कि किस कमरे से क्या मिला था[1]।

1752 में मज़दूरों ने काली छड़ियाँ जैसी वस्तुएं बाहर निकालना शुरू किया। शुरू में उन्होंने समझा कि यह लकड़ी का कोयला है। बाद में पता चला कि वे पपीरस के ग्रंथ-चर्मपत्र थे। इस हवेली से मिले ग्रंथ-चर्मपत्रों की संख्या लगभग 1,100 बताई जाती है। यह प्राचीन दुनिया के किसी पुस्तकालय के पूर्ण रूप से बचे रहने का एकमात्र ज्ञात उदाहरण है[1]।

इस पुस्तकालय में फिलोडेमस के लेख असाधारण रूप से अधिक थे। वहाँ अंतिम प्रतियाँ, प्रतिलिपियाँ और आरंभिक मसौदे साथ-साथ रखे थे। वदिवानों का मानना है कि उन्होंने इसी घर में रचनाएँ लिखीं और संकलित की थीं[1]। कठिनाई से खोले गए ग्रंथ-चर्मपत्रों में दर्शन, वक्तव्यशास्त्र और नीतिशास्त्र पर लेख थे। वक्तव्यशास्त्र वाणी और लेखन से दूसरों को सहमत करने की वदिया है। जो ग्रंथ-चर्मपत्र नहीं खोले जा सके, उनमें क्या लिखा था, यह कोई नहीं जानता था[2]।

ज़बरदस्ती खोलने की कोशिश में नष्ट हुए ग्रंथ-चर्मपत्र

वदिवान उनके भीतर का पाठ पढ़ना चाहते थे। समस्या यह थी कि छूते ही ग्रंथ-चर्मपत्र बखिर जाते थे[2]। 1753 में पादरी एंटोनियो पियाजियो ने एक मशीन बनाई। यह धागे और वज़न की मदद से ग्रंथ-चर्मपत्र को धीरे-धीरे खींचकर खोलने वाला उपकरण था[3]। इस मशीन ने कुछ ग्रंथों को बचा तो लिया, लेकिन इसने भीतरी परतों को फाड़ भी दिया[2]।

19वीं सदी में ब्रिटिश रसायनशास्त्री हम्फ्री डेवी आगे आए। उन्होंने ग्रंथ-चर्मपत्रों पर क्लोरीन गैस, आयोडीन वाष्प और अम्लों का प्रयोग किया। यह रसायनों की मदद से परतों को अलग करने की विधि थी। परिणाम अच्छे नहीं रहे। रसायनों ने पपीरस को कमजोर बनाकर नष्ट कर दिया[3]।

भौतिकी रूप से खोलने के अधिकांश प्रयासों ने पुरावशेषों को नुकसान पहुँचाया। अक्षरों वाले टुकड़े ज़मीन पर बखिर कर गायब भी हो गए। आखिरकार विद्वत् समुदाय ने उन्हें खोलने के प्रयास रोक दिए। आज भी सैकड़ों ग्रंथ-चर्मपत्र बनी खुले बचे हुए हैं[2]। ये पुरावशेष नेपल्स, पेरिस और ऑक्सफ़ोर्ड के संग्रह संस्थानों में बँटे हुए रखे हैं[4]।

X-ray को रोकने वाली कार्बन की दीवार

वैज्ञानिकों ने उन्हें बना खोले पढ़ने का रास्ता खोजा। सबसे पहले जसि तकनीक का विचार आया, वह कंप्यूटेड टोमोग्राफी थी। इसे आमतौर पर CT कहा जाता है। CT वस्तु को काटे बना उसके भीतरी घनत्व के अंतर को मापने वाला उपकरण है। अधिक घनत्व वाला पदार्थ X-ray को अधिक रोकता है। इसलिए क्रॉस-सेक्शनल तस्वीरों में वह चमकीला दिखाई देता है[5]।

धातु युक्त स्याही CT छवियों में सफेद चमकती है। मध्यकालीन यूरोप में लौह युक्त स्याही का व्यापक उपयोग होता था। इज़राइल के एन-गेदी से मिला जला हुआ ग्रंथ-चर्मपत्र भी ऐसा ही था। यह पुरावशेष प्राचीन इज़राइल की हबिरू पांडुलिपि है। स्याही में धातु मेली होने के कारण तस्वीर में निशान रह गए। इसलिए 2016 में इसे बना खोले ही पढ़ लिया गया[5]।

हर्कुलेनियम के लिपिकों की स्थिति अलग थी। लिपिक वह व्यक्ति होता था जो हाथ से दस्तावेज़ों की नकल तैयार करता था। वे जली हुई लकड़ी की कालखि में पानी और राल मिलाकर स्याही बनाते थे। इस स्याही का मुख्य घटक कार्बन है। लेकिन ज्वालामुखी की गर्मी से झुलसा पपीरस भी कार्बन बन चुका था। यानी कागज़ और लिखावट दोनों एक ही तत्व बन गए थे[6]।

यदि घनत्व समान हो, तो X-ray दोनों स्थानों से एक समान रूप से गुजरती है। क्रॉस-सेक्शनल तस्वीरों में कोई सीमा रेखा दिखाई नहीं देती। विद्वानों ने इस बाधा को कार्बन स्याही की समस्या कहा। 2015 में इतालवी और फ्रांसीसी शोधकर्ताओं ने एक अन्य विधि आजमाई। यह फेज़-कंट्रास्ट इमेजिंग नामक तकनीक थी। यह इस गुण का उपयोग करती है कि जब X-ray किसी पदार्थ की सीमा से गुजरती है, तो उसकी दिशा थोड़ी मुड़ जाती है। घनत्व समान होने पर भी सीमा रेखा पर संकेत बदल जाता है। शोधकर्ताओं ने इस विधि से मुड़े हुए ग्रंथ-चर्मपत्र के भीतर कुछ अक्षरों की रूपरेखा पकड़ ली। लेकिन यह वाक्यों को पढ़ने के स्तर तक पर्याप्त नहीं था[6]। लंबे समय तक लोगों का मानना था कि इस दीवार को पार नहीं किया जा सकता[2]।

संदर्भ सामग्री

- [1] Herculaneum Society. The Villa of the Papyri. University of Oxford. <https://www.herculaneum.ox.ac.uk/index.php/the-villa-of-the-papyri/> (3 सितंबर 2026 को अभिगमिति)
- [2] Vesuvius Challenge. FAQ. scrollprize.org. <https://scrollprize.org/faq> (3 सितंबर 2026 को अभिगमिति)
- [3] Villa of the Papyri. Wikipedia. https://en.wikipedia.org/wiki/Villa_of_the_Papyri (3 सितंबर 2026 को अभिगमिति)
- [4] University of Kentucky. (2026, 6월 25일). The day the Herculaneum scrolls began speaking again. UKNow. <https://uknow.uky.edu/research/day-herculaneum-scrolls-began-speaking-again> (3 सितंबर 2026 को अभिगमिति)
- [5] Seales, W. B., Parker, C. S., Segal, M., Tov, E., Shor, P., & Porath, Y. (2016). From damage to discovery via virtual unwrapping: Reading the scroll from En-Gedi. Science Advances, 2(9), e1601247. <https://doi.org/10.1126/sciadv.1601247>
- [6] Mocella, V., Brun, E., Ferrero, C., & Delattre, D. (2015). Revealing letters in rolled Herculaneum papyri by X-ray phase-contrast imaging. Nature Communications, 6, 5895. <https://doi.org/10.1038/ncomms6895>
- [7] Papyrus. Encyclopaedia Britannica. <https://www.britannica.com/topic/papyrus-writing-material> (3 सितंबर 2026 को अभिगमिति)

1.2. वैज्ञानिकी अनुवेषण दल की खोज 1: आभासी प्याज के छलिके उतारना

बना खोले पढ़ने का विचार

अमेरिका के केंटकी विश्वविद्यालय के प्रोफेसर ब्रेंट सील्स ने एक अलग रास्ता चुना। उनका विचार पुरावशेष को खोलने के बजाय कंप्यूटर के भीतर फैलाने का था। वास्तविक वस्तु को वैसा ही रहने देकर केवल 3D छवियाँ प्राप्त की जाती हैं। इसके बाद स्क्रीन पर एक-एक करके परतों को अलग किया जाता है। यह प्याज के छलिके उतारने की तरह अंदर की ओर जाने की विधि है[1]।

शोध दल ने इस वधिका आभासी अनावरण नाम दिया। अंग्रेजी में इसे वर्चुअल अनरैपिंग कहते हैं। 2016 में यह वधि पहली बार सफल हुई। इसका वधि पछिले भाग में देखा गया एन-गेदी ग्रंथ-चर्मपत्र था। यह पुरावशेष 1970 में एक प्राचीन सनिगॉग के खंडहर से मिला था। बाहरी रूप से यह कोयले का एक काला ढेला मात्र था। जब शोध दल ने इसे स्क्रीन पर खोला, तो हब्रू अक्षर प्रकट हुए। यह पुराने नियम की लैव्यवस्था का प्रारंभिक भाग था[1]।

आभासी अनावरण पाँच चरणों में पूरा होता है। 3D तस्वीर लेना, परतों को वभाजति करना, स्याही के मान लगाना, समतल फैलाना और आपस में जोड़ना। एन-गेदी पांडुलिपि में यह क्रम पूरी तरह काम कर गया। क्योंकि तस्वीर में स्याही पहले से ही दिखाई दे रही थी[1]।

हर्कुलेनियम के ग्रंथ-चर्मपत्र की स्थिति एन-गेदी पांडुलिपि से भी बदतर थी। ज्वालामुखी की गर्मी से दबकर उसका क्रॉस-सेक्शन वकृत हो गया था। गोल आकार बगिड़ गया था और परतें आपस में चपिक गई थी। इसके अलावा, तस्वीरों में स्याही दिखाई नहीं दे रही थी। पाँच चरणों में से दूसरे और तीसरे चरण ने एक साथ बाधा खड़ी कर दी[2]।

इसलिए हर्कुलेनियम के मामले में क्रम बदल दिया गया। पहले परतों को अलग करके समतल फैलाया जाता है। स्याही ढूँढने का काम उसके बाद अलग से किया जाता है। खंड 1.3 इसी स्याही की खोज की कहानी है[2]।

कण त्वरक द्वारा ली गई 3D तस्वीरें

पहला चरण ग्रंथ-चर्मपत्र के पूरे भीतरी हिस्से की तस्वीर लेना है[1]। अस्पताल का CT आमतौर पर लगभग 1 मिलीमीटर तक का अंतर ही पहचान पाता है। पपीरस की परतें इससे कहीं अधिक पतली होती हैं। एक परत की मोटाई बाल की मोटाई से भी कम होती है। अस्पताल के उपकरणों से परतें आपस में धुंधली और चपिकी हुई दिखती हैं[3]।

शोध दल कण त्वरक के पास गया। वे पहले ब्रिटन के डायमंड लाइट सोर्स गए[3]। बाद में फ्रांस के ग्रैनोबल स्थिति यूरोपीय सिक्रोटरॉन विकिरण सुविधा का भी उपयोग किया गया[4]। ये सुविधाएं अत्यंत चमकदार और सीधी दिशा में जाने वाली X-ray उत्पन्न करती हैं। प्रकाश जतिना तीव्र होगा, उतने ही कम समय में स्पष्ट तस्वीर प्राप्त होती है[3]।

शोध दल ने ग्रंथ-चर्मपत्र को एक पात्र में सीधा खड़ा करके मजबूती से स्थिर किया। क्योंकि स्कैनिंग के दौरान थोड़ा सा भी हिलने पर तस्वीर धुंधली हो जाती है। इस प्रकार प्राप्त डेटा का रजिऑल्यूशन प्रतियॉक्सेल लगभग 8 माइक्रोमीटर था। वॉक्सेल गहराई के मान वाले 3D पक्सेल को कहते हैं। इसे चित्र के एक बटु को छोटे घन में बदलने जैसा समझा जा सकता है[3]।

डेटा की मात्रा अत्यधिक थी। एक ग्रंथ-चर्मपत्र को स्कैन करने पर लगभग 300 टेराबाइट डेटा एकत्र होता है। यह हजारों स्मार्टफोन की भंडारण क्षमता को भरने जतिना बड़ा आकार है। इस डेटा को संभालना अपने आप में एक नई चुनौती बन गया[4]।

फ्रांस के ग्रैनोबल की सुविधा को चौथी पीढ़ी का प्रकाश स्रोत कहा जाता है। 2020 में इसके नए उपकरण चालू हुए। यूरोप के 21 देशों ने 15 करोड़ यूरो खर्च करके इसके प्रदर्शन को उन्नत किया। यह सुविधा मूल रूप से पदार्थ और जीवन की परिघटनाओं का अध्ययन करने के लिए बनाई गई थी। शुरू में किसी को नहीं पता था कि इसका उपयोग प्राचीन ग्रंथ-चर्मपत्रों को पढ़ने के लिए किया जाएगा[4]।

परतों को अलग करने वाला वॉल्यूमेट्रिक वभाजन

3D तस्वीर मिलने के बाद अगला चरण परतों को अलग करना होता है। इस प्रक्रिया को वॉल्यूमेट्रिक वभाजन कहा जाता है। यह 3D छवि में से परतों को एक-एक करके अलग करने का कार्य है। कंप्यूटर प्रत्येक क्रॉस-सेक्शनल तस्वीर में कागज की सतह के मार्ग का अनुसरण करता है। उन रेखाओं को जोड़ने पर कागज की एक शीट का 3D मॉडल बन जाता है। इस मॉडल को मेश कहा जाता है[1]।

शुरू में लोग माउस से एक-एक करके सीमाओं पर बटु लगाते थे[2]। एक ग्रंथ-चर्मपत्र में नहिंति कागज की लंबाई 1 मीटर से अधिक होती है[4]। इसे हाथ से ट्रेस करना बहुत लंबा और सघन कार्य था। इसके अलावा, परतें आपस में चपिकी होने के कारण सीमाएं अस्पष्ट थीं[2]।

प्रतियोगिता के प्रतियोगियों द्वारा स्वचालित उपकरण पेश करने से गतिबढ़ गई। स्विट्जरलैंड के जूलियन शलिगिर द्वारा बनाया गया प्रोग्राम इसका प्रमुख उदाहरण है। यह प्रोग्राम आस-पास की क्रॉस-सेक्शनल छवियों के पैटर्न का अनुसरण करते हुए परतों को जोड़ता है। इसने उस काम को कुछ ही घंटों में समेट दिया जिसमें इंसानों को कई दिन लग जाते थे[1]।

यह कार्य इतना बड़ा था कि कोई एक प्रयोगशाला इसे अकेले नहीं कर सकती थी। आयोजकों ने स्कैनिंग डेटा को इंटरनेट पर पूरी तरह खुला रखा। प्रतियोगियों ने अपने बनाए प्रोग्रामों का सोर्स कोड सार्वजनिक कर दिया। दूसरों के बनाए उपकरणों पर अगले व्यक्ति ने नई सुविधाएं जोड़ी। पढ़ा गया पाठ और कोड आज भी कोई भी डाउनलोड कर सकता है[5]।

समतल फैलाना और आपस में जोड़ना

परतों को अलग करने के बाद भी कागज का मॉडल गोल मुड़ा हुआ रहता है। इस स्थिति में अक्षर मुड़े हुए दिखाई देते हैं। इसलिए मॉडल को सपाट फैलाने की गणना की जाती है। इस चरण को समतलीकरण कहा जाता है[1]।

कंप्यूटर मॉडल के शीर्षों को छोटी मोतियों के रूप में बदलता है। यह माना जाता है कि मोतियाँ आपस में स्पर्श से जुड़ी हैं। इसके बाद इस जाल को फर्श पर सपाट बछिने की गणना बार-बार की जाती है। स्पर्श के खचाव या सकिड़न को न्यूनतम स्तर पर समायोजित किया जाता है। इससे अक्षरों का आकार अधिक वक्रित नहीं होता[1]।

अंतिम चरण आपस में जोड़ना है। ग्रंथ-चर्मपत्र एक बार में पूरा नहीं खुलता। कंप्यूटर इसे कई क्षेत्रों में विभाजित करके अलग-अलग फैलाता है। फैलाए गए क्षेत्रों को फिर से एक शीट के रूप में मलाना पड़ता है। इस समय पपीरस के विशिष्ट जालीदार रेशदार पैटर्न का उपयोग किया जाता है। यदि दो क्षेत्रों के रेशों का पैटर्न मेल खाता है, तो इसका अर्थ है कि वे मूल रूप से एक-दूसरे के बगल में थे। इस मलाना कार्य के पूरा होने के बाद आभासी अनावरण संपन्न होता है[1]।

ये पाँच चरण एक बार में पूरे नहीं होते। यदि किसी चरण में कोई त्रुटि हो जाए, तो उसके बाद के सभी चरण बगिड़ जाते हैं। यदि पिरतों को गलत तरीके से विभाजित किया जाए, तो दूसरे पन्नों के अक्षर आपस में मेल जाते हैं। इसलिए शोधकर्ता एक ही ग्रंथ-चर्मपत्र को बार-बार संसाधित करते हैं। वे परिणामों को आँखों से जाँचते हैं, सेटिंग्स बदलते हैं और प्रक्रिया को फिर से चलाते हैं[5]।

संदर्भ सामग्री

- [1] Seales, W. B., Parker, C. S., Segal, M., Tov, E., Shor, P., & Porath, Y. (2016). From damage to discovery via virtual unwrapping: Reading the scroll from En-Gedi. *Science Advances*, 2(9), e1601247. <https://doi.org/10.1126/sciadv.1601247>
- [2] Vesuvius Challenge. (2024, 2 मई 5 मई). Vesuvius Challenge 2023 Grand Prize awarded: we can read the scrolls! scrollprize.org. <https://scrollprize.org/grandprize> (3 सितंबर 2026 को अभिगमिति)
- [3] Vesuvius Challenge. FAQ. [scrollprize.org](https://scrollprize.org/faq). <https://scrollprize.org/faq> (3 सितंबर 2026 को अभिगमिति)
- [4] University of Kentucky. (2026, 6 मई 25 मई). The day the Herculaneum scrolls began speaking again. UKNow. <https://uknow.uky.edu/research/day-herculaneum-scrolls-began-speaking-again> (3 सितंबर 2026 को अभिगमिति)
- [5] Vesuvius Challenge. (2026, 6 मई 25 मई). An entire Herculaneum scroll has been read for the first time. [scrollprize.org](https://scrollprize.org/firstscroll). <https://scrollprize.org/firstscroll> (3 सितंबर 2026 को अभिगमिति)

1.3. वैज्ञानिकी अनुवेषण दल की खोज 2: स्याही के सूक्ष्म फगिरप्रति की खोज

कार्बन पर लिखा कार्बन

आभासी अनावरण समाप्त होने के बाद भी स्क्रीन खाली थी। कागज़ तो फैल गया था, लेकिन अक्षर दिखाई नहीं दे रहे थे[1]। इसका कारण पहले देखी गई कार्बन की समस्या थी। पपीरस भी कार्बन था और स्याही भी कार्बन थी। X-ray दोनों में भेद नहीं कर पाती[2]।

प्रोफेसर सील्स के शोध दल ने एक परिकल्पना तैयार की। उनका मानना था कि ऐसा नहीं हो सकता कि स्याही ने कोई नशान ही न छोड़ा हो[1]। स्याही तरल रूप में कागज़ के संपर्क में आती है। तरल रेशों के बीच रसि जाता है। और सूखने के साथ ही सतह को थोड़ा खींचता है[3]।

इससे कागज़ की सतह का आकार सूक्ष्म रूप से बदल जाता है। घनत्व भले ही समान हो, सतह के उभार और ढलान बदल जाते हैं। शोध दल ने इन सूक्ष्म उभारों को ढूँढ़ने का निर्णय लिया[3]। यह आँखों से न देखिने जितना छोटा था, इसलिए इसे मशीन पर छोड़ना पड़ा[1]।

यहाँ इस्तेमाल किया गया साधन मशीन लर्नगि था। यह ऐसा तरीका नहीं है जिसमें इंसान एक-एक करके नियम लिख कर दे। इसमें सही उत्तरों वाले ढेर सारे उदाहरण दिखाए जाते हैं और उसे खुद नियम खोजने दिए जाते हैं। यह उसी तरह है जैसे हज़ारों बल्लियों की तस्वीरें दिखाने पर वह बल्लियों को पहचानना सीख जाता है। स्याही की खोज में, स्याही मौजूद होने और न होने वाले स्थानों को दर्शाने वाले नमूने दिखाए जाते हैं[1]।

क्रैकल नामक सुराग

मार्च 2023 में वसुवियस चैलेज नामक एक प्रतियोगिता आयोजित की गई[1]। इसे प्रोफेसर सील्स, नैट फ्राइडमैन और डैनयिल ग्रॉस ने मलिकर शुरू किया था[4]। आयोजकों ने स्कैनगि डेटा को इंटरनेट पर मुफ्त उपलब्ध करा दिया। उन्होंने पुरस्कार राशिरिखकर दुनिया भर के प्रोग्रामरों को एकत्रित किया[1]।

पहले कुछ महीनों तक किसी को भी स्याही नहीं मिली। उस वर्ष की गर्मियों में किसी हैडमर नामक एक प्रतिभागी को एक सुराग मिला। उन्होंने खोली गई सतह के डेटा को काफी देर तक सीधे आँखों से देखा। तभी उन्हें एक असामान्य पैटर्न दिखाई दिया। यह बारीक दरारों के फटने का नशान था[1]।

सूखे पड़े खेत की ज़मीन की कल्पना की जा सकती है। सूखी मट्टि जाल की तरह फट जाती है। स्याही लगे स्थानों पर भी ऐसी ही बारीक दरारें बन गईं। प्रतिभागियों ने इस पैटर्न को 'क्रैकल' कहा। हैडमर को इस खोज के लिए पहला स्याही पुरस्कार मिला[1]।

हैडमर ने खोज के तुरंत बाद ऑनलाइन चैट रूम में इसकी सूचना दी। इस साझाकरण ने प्रतियोगिता की दशा बदल दी। अन्य प्रतभागियों ने तुरंत इस पैटर्न को प्रशिक्षण सामग्री बना लिया[1]।

केवल छोटी खड़की के माध्यम से दिखाना

अमेरिकी विश्वविद्यालय के छात्र ल्यूक फ़ैरेटर ने सबसे पहले कदम उठाया। उन्होंने मशीन लर्निंग मॉडल को क्रैकल पैटर्न सिखाया। मॉडल संभावना के साथ उत्तर देता है कि स्याही कहाँ है। जैसे-जैसे प्रशिक्षण आगे बढ़ा, उत्तर अधिक सटीक होते गए[1]।

समस्या यह थी कि प्रशिक्षण डेटा बहुत कम था। जब डेटा कम होता है, तो मशीन उत्तर रट लेती है। इसे ओवरफिटिंग कहा जाता है। यह परीक्षा के प्रश्नों को समझे बिना केवल उत्तर रटने वाले छात्र जैसा है[1]।

प्रतभागियों ने डेटा को घुमाकर और उलटकर उसकी मात्रा बढ़ाई। छवि को घुमाने पर भी बारीक दरारों के पैटर्न की प्रकृति वैसी ही रहती है। इसकी बदौलत मशीन ने केवल पैटर्न की ही विशेषताओं को सीखा[1]।

इनपुट वडिों के आकार में भी सुधार किया गया। शुरुआत में एक बार में बड़ा क्षेत्र दिखाया गया था। व्यापक रूप से देखने पर पूरे अक्षर का आकार नज़र में आ जाता है। तब मशीन ने मनगढ़ंत अक्षर बनाना शुरू कर दिया। जहाँ स्याही नहीं थी, वहाँ उसने संभावित स्ट्रोक जोड़ दिए। इसे एआई का मतभ्रम कहा जाता है[1]।

शोधकर्ताओं ने वडिों का आकार घटाकर 64 गुणा 64 पिकसेल कर दिया। उस आकार में अक्षर का रूप प्रकट नहीं होता है। मशीन को केवल बारीक दरारों की बनावट देखकर नरिणय लेना होता है। मनगढ़ंत नरिमाण में भारी कमी आई। इसके बजाय स्क्रीन पर वास्तविक ब्रश के निशान स्पष्ट रूप से दिखाई देने लगे[1]।

इस गणना के लिए ज़मिंदार मॉडल दो प्रकारों को जोड़कर बनाया गया था। एक 3D कनवोल्यूशनल न्यूरोल नेटवर्क है। यह छोटे घन के भीतर के पैटर्न को स्कैन करके विशेषताओं को निकालने वाली संरचना है। दूसरा ट्रांसफार्मर है। यह दूर स्थिति संकेतों के बीच के संबंध की जाँच करने वाली संरचना है। दोनों संरचनाओं का एक साथ उपयोग करने पर पतली परतों के भीतर के धुंधले पैटर्न भी छूटते नहीं हैं[1]।

मशीन द्वारा प्रस्तुत स्क्रीन को तुरंत सही उत्तर के रूप में स्वीकार नहीं किया जाता है। पुरालेख विशेषज्ञ उस स्क्रीन को अलग से पढ़ते हैं। वे जाँचते हैं कि ग्रीक शब्दों का क्रम व्याकरण के अनुसार सही है या नहीं। वे यह भी पुष्टि करते हैं कि आगामी वाक्यों का अर्थ स्पष्ट है या नहीं। यदि मशीन और मानव का पठन मेल नहीं खाता, तो उस भाग की फरि से गणना की जाती है[1]।

केवल सतह के उभारों से

इसके बाद भी स्याही खोजने की विधि में नरितर सुधार होता रहा। प्रतभागि यूसुफ़ नादेर ने पूर्व-प्रशिक्षण पद्धतिका उपयोग किया। यह वह तरीका है जिसमें बिना उत्तर-कुंजी वाला चर्मपत्र डेटा पहले बड़ी मात्रा में दिखाया जाता है। मशीन पपीरस के रेशों की बनावट को खुद सीखती है। उसके बाद बहुत कम उत्तर-कुंजी के साथ केवल स्याही खोजने का प्रशिक्षण दिया जाता है[1]।

मार्च 2026 में सतह की वक्रता पर ही केंद्रित एक शोध सामने आया। इसे जयिर्जयिों एंजेलोटी, फेडेरिका निकोलार्डी, पॉल हैडरसन और ब्रेट सील्स ने मलिकर लिखा था। शोध दल ने पहले से खुले हुए पपीरस नमूनों की सतह को ऑप्टिकल उपकरणों से मापा। उन्होंने केवल ऊँचाई और नचाई वाले डेटा से मशीन को प्रशिक्षित किया। इसमें चमक या घनत्व की कोई जानकारी शामिल नहीं की गई थी[3]।

फरि भी मशीन ने स्याही वाले स्थानों की पहचान कर ली। इसका अर्थ है कि केवल सतह के उतार-चढ़ाव से ही स्याही खोजी जा सकती है। शोध दल ने यह भी गणना की कि माप कतिना सघन होना चाहिए। यह संख्या अगली इमेजिंग के रजिऑल्यूशन का मानक बनती है। यह शोधपत्र उसी वर्ष साइंटिफिक रिपोर्ट्स पत्रिका में प्रकाशित हुआ था[3]।

फरि भी स्याही खोजना अभी पूर्ण तकनीक नहीं है। जहाँ स्याही हल्की बची है, वे खंड अभी भी धुंधले दिखाई देते हैं। जहाँ परते आपस में चपिकी हुई हैं, वहाँ इमेजिंग डेटा ही अस्पष्ट होता है। कभी-कभी दो अलग-अलग पन्ने जुड़कर एक हो जाते हैं। एक चर्मपत्र पर अच्छा काम करने वाला मॉडल दूसरे चर्मपत्र पर नष्टप्रभावी भी हो सकता है। इसलिए इंसान बीच-बीच में आँखों से जाँच करते हैं और सुधार करते हैं। शोध आज भी एक-एक कदम आगे बढ़ रहा है[5]।

संदर्भ सामग्री

[1] Vesuvius Challenge. (2024, 2월 5일). Vesuvius Challenge 2023 Grand Prize awarded: we can read the scrolls! scrollprize.org. <https://scrollprize.org/grandprize> (3 सितंबर 2026 को अभिगमिति)

[2] Mocella, V., Brun, E., Ferrero, C., & Delattre, D. (2015). Revealing letters in rolled Herculaneum papyri by X-ray phase-contrast imaging. Nature Communications, 6, 5895. <https://doi.org/10.1038/ncomms6895>

- [3] Angelotti, G., Nicolardi, F., Henderson, P., & Seales, W. B. (2026). Ink detection from surface topography of the Herculaneum papyri. Scientific Reports. <https://doi.org/10.1038/s41598-026-58467-1>
- [4] University of Kentucky. (2026, 6월 25일). The day the Herculaneum scrolls began speaking again. UKNow. <https://uknow.uky.edu/research/day-herculaneum-scrolls-began-speaking-again> (3 सितंबर 2026 को अभिगमिति)
- [5] Vesuvius Challenge. (2026, 7월 10일). Open problems: why reading every Herculaneum scroll is still a challenge. scrollprize.org. https://scrollprize.org/2026_open_problems (3 सितंबर 2026 को अभिगमिति)

1.4. सामने आया सच

2023 में पुनः प्राप्त पहला शब्द

पहली सफलता केवल एक शब्द थी। इसकी घोषणा 12 अक्टूबर 2023 को हुई थी। इसके नायक अमेरिकी विश्वविद्यालय के छात्र ल्यूक फ़ैरेटर थे। उस समय उनकी आयु 21 वर्ष थी। उन्होंने जो शब्द पढ़ा, वह पोरफ़रिस था। यह बैंगनी रंग का अर्थ देने वाला एक ग्रीक शब्द है। प्राचीन साहित्य में यह सामान्यतः प्रयुक्त होने वाला शब्द नहीं है। फ़ैरेटर को प्रथम अक्षर पुरस्कार के रूप में 40,000 डॉलर मिले। दूसरे स्थान पर रहे यूसुफ़ नादेर को 10,000 डॉलर मिले। यह बनिा खोले किसी चर्मपत्र से शब्द पढ़ने का पहला उदाहरण था[7]।

5 फ़रवरी 2024 को प्रतियोगिता के परिणामों की घोषणा की गई। ग्रैंड प्रक्सिस् तीन लोगों ने संयुक्त रूप से जीता। वे मसिर के यूसुफ़ नादेर, अमेरिका के ल्यूक फ़ैरेटर और स्वट्ज़रलैंड के जूलियन शलिगिर थे। पुरस्कार राशि 700,000 डॉलर थी[1]।

इनके द्वारा पढ़ा गया अवशेष पेरसि अकादमी का चर्मपत्र है। इसका नाम PHerc. Paris. 4 है। तीनों ने इस चर्मपत्र से 2,000 से अधिक ग्रीक अक्षरों को पुनः प्राप्त किया। यह पंद्रह कॉलमों का पाठ था। कॉलम चर्मपत्र पर लंबवत रूप से विभाजित पाठ का एक खंड होता है। यह पूरे चर्मपत्र के लगभग 5 प्रतिशत के बराबर है[1]।

मूल्यांकन के मानदंड बहुत कड़े थे। उन्हें 140 अक्षरों वाले चार अंश प्रस्तुत करने थे। प्रत्येक अंश में 85 प्रतिशत से अधिक अक्षरों को पढ़ने पर ही उत्तीर्ण माना जाता था। पुरालेख विशेषज्ञों ने पहचान गोपनीय रखकर इनकी समीक्षा की[1]।

वषिय आनंद पर एक चर्चा थी। इसमें यह परखा गया कि क्या दुर्लभ वस्तुएँ सामान्य वस्तुओं की तुलना में नश्चिती रूप से अधिक आनंद देती हैं। मूल पाठ में भोजन का उदाहरण देने वाला एक अंश है। इसका आशय यह है कि केवल दुर्लभ होने के कारण ही कोई चीज़ अधिक आनंददायक नहीं मान ली जाती। यह वाक्य एपिक्यूरोस दर्शन के विचार को सीधे दर्शाता है[1]।

शीर्षक पुनः प्राप्त करने वाला चर्मपत्र

अगली उपलब्ध शीर्षक थी। 5 मई 2025 को प्रथम शीर्षक पुरस्कार दिया गया। इसके विजेता मार्सेल रोथ और महिा नोवाक थे। पुरस्कार राशि 60,000 डॉलर थी[2]।

संबंधित अवशेष ऑक्सफ़ोर्ड की बोडलियन लाइब्रेरी का PHerc. 172 था। स्क्रीन पर दिखाई देने वाला पाठ फिलोडेमस की कृतदुरुगुणों के वषिय में था। इसे पुस्तक 1 के रूप में पढ़ा गया। नेपल्स के फेडेरिको द्वितीय विश्वविद्यालय की फेडेरिका निकोलार्डी के शोध दल ने इसकी पुष्टि की। यह लपेटे हुए चर्मपत्र के शीर्षक को बनिा खोले पढ़ने का पहला मामला है[2]।

इसके बाद इस चर्मपत्र से 70 से अधिक कॉलम पढ़े गए। यह लालच, घमंड और चापलूसी जैसे स्वभावगत दोषों पर केंद्रित पाठ है। अब चर्चा के प्रवाह को आरंभ से समझा जा सकता है[3]।

शुरू से अंत तक पूरी पढ़ी गई पहली पुस्तक

25 जून 2026 को और भी बड़ा समाचार आया। यह घोषणा थी कि एक पूरे चर्मपत्र को शुरू से अंत तक पढ़ लिया गया है। यह अवशेष नेपल्स नेशनल लाइब्रेरी का PHerc. 1667 है। प्रतियोगिता के प्रतिभागी इसे चर्मपत्र संख्या 4 कहते हैं[4]।

स्कैनिंग फ़्रांस की यूरोपियन साइंटिफ़िक रीइजिशन फ़ाउंडेशन में की गई थी। इसके बाद आभासी अनावरण और स्याही की खोज के चरण पूरे किए गए। पुनः प्राप्त पाठ की लंबाई लगभग 1.4 मीटर थी। इसमें लगभग बाईस कॉलम थे। यह पहली बार था जब किसी हर्कुलेनियम चर्मपत्र को पूरी तरह से पढ़ा गया था[4]।

इसकी सामग्री नैतिकता और मानवीय चरित्र से संबंधित एक ग्रीक नबिंध है[3]। पाठ में स्टोइक दार्शनिक क्रसिपिस के भतीजे का उल्लेख मिलता है। उसका नाम एरिस्टोक्रयोन है[4]। इस पाठ का लेखन काल ईसा पूर्व तीसरी शताब्दी के अंत से दूसरी शताब्दी के आसपास माना जाता है। इसके लेखक कौन हैं, यह अभी तय नहीं हुआ है[3]।

शोध परिणाम एक शोधपत्र के रूप में भी प्रकाशित किए गए। जयिर्जियो एंजेलोटी, स्टीफन पार्सनस और फेडेरिका निकोलार्डी सहित 27 लोग इसके सह-लेखक बने[5]। डेटा, कोड और पढ़ा गया पाठ भी साथ में जारी किया गया[4]।

शोध दल ने सत्यापन प्रक्रिया का भी खुलासा किया। उन्होंने 2023 में पढ़े गए PHerc. Paris. 4 को उच्च रजिऑल्यूशन पर फरि से स्कैन किया। नए डेटा से प्राप्त पाठ पछिले परिणामों से पूरी तरह मेल खाता था। यह इस बात का प्रमाण है कि यह पाठ मशीन द्वारा गढ़ा नहीं गया था। इसका यह भी अर्थ है कि यही वधि अन्य चर्मपत्रों पर भी लागू की जा सकती है[4]।

अभी भी शेष 600 ग्रंथ

उसी घोषणा में एक अन्य समाचार भी शामिल था। PHerc. 139 नामक चर्मपत्र का शीर्षक पढ़ा गया। यह फलिोडेमस की कृतदिवताओं के वषिय में की आठवीं पुस्तक थी। उससे पहले तक इस कृतकी केवल पहली पुस्तक ही ज्ञात थी। पहली बार यह तथ्य सामने आया कि यह रचना आठ से अधिक पुस्तकों तक वसित थी[3]।

अभी भी 600 से अधिक ऐसे चर्मपत्र हैं जिन्हें पढ़ा नहीं जा सका है[3]। प्रतियोगिता ने वर्ष 2027 के लिए ग्रैंड प्रिक्स की घोषणा की है। कुल पुरस्कार राशि 1,000,000 डॉलर है। प्रथम स्थान प्राप्त करने वाले को 800,000 डॉलर मिलेंगे। लक्ष्य 13 निर्दिष्ट चर्मपत्र है। प्रत्येक कॉलम में बचे हुए अक्षरों का 70 प्रतिशत से अधिक भाग पढ़ना आवश्यक है। इसकी अंतिम तिथि 25 जून 2027 है[6]।

इस कहानी में अभी भी कई रक्ति स्थान हैं। अक्षरों को पढ़ लेने के बाद भी उनके अर्थ को समझना एक अलग काम है। पढ़े गए पाठ के एक कॉलम को अकादमिक रूप से परिष्कृत करने में भी लंबा समय लगता है। ऐसे चर्मपत्र भी बचे हैं जिनके लेखक का पता नहीं चला है[3]।

फरि भी दशिया स्पष्ट हो गई है। जो पुस्तक 2,000 वर्षों से कोई नहीं खोल सका था, वह स्क्रीन पर खुल गई। जलकर नष्ट हो चुकी समझी जाने वाली लाइब्रेरी धीरे-धीरे लौट रही है[4]।

यह काम करने वाले केवल एक व्यक्ति नहीं हैं। स्कैनिंग का काम करने वाले भौतिक विज्ञानी और परतों को अलग करने वाले प्रोग्रामर थे। प्राचीन ग्रीक भाषा को पढ़ने वाले पुरालेखवेत्ता भी थे। प्रतियोगिता में भाग लेने वाले छात्रों और कर्मचारियों ने भी हाथ बँटाया। 2,000 वर्षों बाद पुस्तकों का खुलना इन लोगों का संयुक्त परिणाम है। शेष चर्मपत्र भी इसी तरह खोले जा सकते हैं[4]।

संदर्भ सामग्री

- [1] Vesuvius Challenge. (2024, 2월 5일). Vesuvius Challenge 2023 Grand Prize awarded: we can read the scrolls! scrollprize.org. <https://scrollprize.org/grandprize> (3 सितंबर 2026 को अभिगमिति)
- [2] Vesuvius Challenge. (2025, 5월 5일). \$60,000 First Title Prize awarded! scrollprize.substack.com. <https://scrollprize.substack.com/p/60000-first-title-prize-awarded> (3 सितंबर 2026 को अभिगमिति)
- [3] University of Kentucky. (2026, 6월 25일). The day the Herculeum scrolls began speaking again. UKNow. <https://uknow.uky.edu/research/day-herculeum-scrolls-began-speaking-again> (3 सितंबर 2026 को अभिगमिति)
- [4] Vesuvius Challenge. (2026, 6월 25일). An entire Herculeum scroll has been read for the first time. scrollprize.org. <https://scrollprize.org/firstscroll> (3 सितंबर 2026 को अभिगमिति)
- [5] Angelotti, G., Parsons, S., Nicolardi, F., Nader, Y., Johnson, S., Josey, D., et al. (2026). Complete virtual unwrapping and reading of a rolled Herculeum papyrus. arXiv:2606.29085. <https://arxiv.org/abs/2606.29085> (3 सितंबर 2026 को अभिगमिति)
- [6] Vesuvius Challenge. Open Prizes. scrollprize.org. <https://scrollprize.org/prizes> (3 सितंबर 2026 को अभिगमिति)
- [7] Vesuvius Challenge. (2023, 10월 12일). First word discovered in unopened Herculeum scroll by 21yo computer science student. scrollprize.org. <https://scrollprize.org/firstletters> (3 सितंबर 2026 को अभिगमिति)

अध्याय 2. जली हुई मंजूषा से नकिले वचन: एन-गेदी की मुहरबंद लैव्यव्यवस्था

2.1. केस फाइल

मृत सागर के पश्चिमि का नखलसितान

एन-गेदी मृत सागर के पश्चिमि तट पर स्थिति एक नखलसितान है। इसे एनगेदी भी लिखते हैं। इसके चारों ओर सूखा रेगसितान है, लेकिन यहाँ झरने का पानी फूटता है। पानी होने के कारण लोग यहाँ इकट्ठा हुए। लोग हजारों वर्षों तक इस जलस्रोत के किनारे गाँव बनाकर रहे। पुराने नियम की 1 शमूल पुस्तक में भी इस भूमिका नाम आता है। यह वह कथा है जिसमें दाऊद राजा शाऊल से बचकर एन-गेदी के बीहड़ में छुपा था।

पुरातात्विक सर्वेक्षण के अनुसार, इस पहाड़ी पर ईसा पूर्व 8वीं शताब्दी के उत्तरार्ध से लोग रहते थे। और ईस्वी 600 के आसपास लगी एक भीषण आग से यह गाँव नष्ट हो गया[1]। यह ऐसी भूमि है जहाँ तेरह सौ से अधिक वर्षों का इतिहास मिट्टी में परतों के रूप में दबा हुआ है।

मृत सागर के आसपास बारिश कम होती है और हवा शुष्क रहती है। ऐसी जलवायु पुरानी वस्तुओं को लंबे समय तक सुरक्षित रखने में मदद करती है। लकड़ी, चमड़ा और कपड़ा अन्य स्थानों की तुलना में कम सड़ते हैं। मृत सागर पांडुलिपियाँ इसी क्षेत्र से मिलने का कारण भी यही है।

1970 में यहाँ एक बड़ा उत्खनन शुरू हुआ। इस उत्खनन का नेतृत्व डैन बाराग और एहूद नेतज़र ने किया था। ये दोनों यरुशलम के हब्रू विश्वविद्यालय के पुरातत्व संस्थान से थे। योसेफ पोरथ ने 5 मई 1970 को इस स्क्रॉल को खोदकर निकाला[1]। उत्खनन दल को एक प्राचीन यहूदी आराधनालय का स्थल मिला। आराधनालय वह इमारत है जहाँ यहूदी समुदाय प्रार्थना करने और धर्मग्रंथ पढ़ने के लिए इकट्ठा होता था। इसके फर्श पर अलग-अलग रंगों के छोटे पत्थरों को जोड़कर बनाया गया मोज़ेक बछियाँ हुआ था।

आराधनालय के भीतर पवित्र मंजूषा रखने का एक स्थान होता है। पवित्र मंजूषा वह संदूक है जिसमें पवित्र धर्मग्रंथ के चर्मपत्र रखे जाते हैं। यहूदी समुदाय आराधनालय में इस संदूक को सबसे पवित्र स्थान मानता है। उत्खनन दल ने उस पवित्र मंजूषा से जले हुए मलबे को हटाया। मलबे के बीच से एक काला पडि निकला[1]।

पवित्र मंजूषा से इसका मिलना अपने आप में एक महत्वपूर्ण सुराग है। आराधनालय की पवित्र मंजूषा में कोई भी साधारण दस्तावेज़ नहीं रखा जाता। उसमें वही धर्मग्रंथ-चर्मपत्र रखे जाते हैं जिनसे समुदाय पढ़ता और सुरक्षित रखता था। उस काले पडि की पहचान का अनुमान लगाने का आधार यही था।

प्राचीन धर्मग्रंथ की पांडुलिपियाँ अपने आप में बहुमूल्य सामग्री हैं। प्रटिगि प्रेस न होने के दौर में लोग हाथ से लिखकर इनकी नकल करते थे। हर बार नकल करते समय अक्षरों के बदलने की संभावना रहती है। जब कोई पुरानी पांडुलिपि मिलती है, तो यह पुष्टि की जा सकती है कि मूल पाठ किस प्रकार आगे बढ़ा। यही कारण था कि विद्वान पवित्र मंजूषा से निकले उस काले पडि पर इतना ध्यान दे रहे थे।

आग से बचा काला पडि

आग अभिलिखों को नष्ट करने वाली शक्ति है। कागज़ और चमड़ा दोनों ही आग में जलकर राख हो जाते हैं। लेकिन एन-गेदी में स्थिति थोड़ी अलग थी। मलबे के नीचे दबा हुआ चर्मपत्र ऑक्सीजन की कमी वाली स्थिति में धीरे-धीरे झुलस गया। वह राख बनकर बखिरने के बजाय कोयले की तरह सख्त हो गया। इस तरह कोयले की तरह सख्त होने के बदलाव को कार्बनीकरण कहते हैं। हवा की कमी वाली आग में जलने पर ऐसा होता है।

यह सदिधांत कोयले की भट्टी में लकड़ी से कोयला बनने की प्रक्रिया जैसा ही है। लकड़ी को सीधे जलाने पर केवल राख बचती है। यदि उसे मिट्टी से ढँककर हवा रोक दी जाए और तपाया जाए, तो काला कोयला बचता है। कार्बनीकृत ग्रंथ-चर्मपत्र भी इसी तरह बना था।

कार्बनीकृत पडि ढही हुई इमारत की राख और मिट्टी के नीचे दब गया। हवा न मिलने के कारण उसका सड़ना भी रुक गया। चमड़ा स्वाभाविक रूप से जलदी सड़ने वाली सामग्री है। यदि वह ज़मीन पर ऐसे ही पड़ा रहता, तो बहुत पहले नष्ट हो गया होता। जिस आग ने पुरावशेष को नष्ट किया, वही उसे बचाए रखने का कारण भी बनी।

उत्खनन दल द्वारा निकाला गया पडि पहली नज़र में कोयले का ढेला लगता था। लेकिन टूटे हुए पार्श्व भाग पर गोल लपिटी हुई परतें बची हुई थीं। पुरातत्वविदों ने इसे मुड़ा हुआ चमड़े का चर्मपत्र माना[1]। इस पुरावशेष को उस स्थान का नाम दिया गया जहाँ यह मिला था। इस प्रकार इसे एन-गेदी ग्रंथ-चर्मपत्र नाम मिला।

यह ग्रंथ-चर्मपत्र जानवरों की खाल को तैयार करके बनाया गया पार्चमेंट है[1]। इटली के हर्कुलेनियम से मिला ग्रंथ-चर्मपत्र पपीरस से बना था। पपीरस पौधों के तनों को दबाकर चपिकाकर बनाया गया वानस्पतिक कागज है। दोनों पुरावशेष अपनी मूल सामग्री से ही भिन्न हैं। यह अंतर बाद में इन्हें पढ़ने की वधितय करने वाला मोड़ साबित हुआ।

पार्चमेंट तैयार करने में बहुत मेहनत लगती है। जानवरों की खाल से बाल और चर्बी हटाकर उसे धोया जाता है। फिर उसे एक ढाँचे पर कसकर खींचकर सुखाया जाता है। इस प्रकार तैयार की गई पतली सतह पर अक्षर लिखे जाते हैं। यह कागज की तुलना में अधिक मजबूत और टिकाऊ होता था, लेकिन बहुत महंगा था।

इसके मलिन की स्थिति भी असामान्य थी। प्राचीन चर्मपत्र आमतौर पर छोटे-छोटे टुकड़ों में मिलते हैं। 1947 से कुमरान की गुफाओं से मिली मृत सागर पांडुलिपियाँ भी अधिकतर टुकड़ों के रूप में ही हैं। एन-गेदी का पुरावशेष लपिटे हुए आकार को बनाए रखते हुए मिला था। इसका अर्थ था कि इसके भीतर कई परतें अभी भी सुरक्षित हैं।

वह पुरावशेष जसिं छुआ नहीं जा सकता था

पाठ पढ़ने के लिए चर्मपत्र को खोलना पड़ता है। लेकिन इस पुरावशेष को खोला नहीं जा सकता था। अत्यधिक गर्मी के कारण चमड़ा सिकुड़ गया था और परतें आपस में चपिक गई थीं। बाहर से यह पत्थर की तरह सख्त था और अंदर से बहुत भुरभुरा था। उंगलियों से छूने मात्र से इसके चूर्ण बन जाने जैसी स्थिति थी[1]।

उपाय बिल्कुल नहीं थे, ऐसा भी नहीं था। चाकू से परतों को अलग करने का एक तरीका था। रसायनों का लेप लगाकर चमड़े को नरम करने का तरीका भी था। ये दोनों ही तरीके पुरावशेष को नष्ट कर देते। एक बार नष्ट हुए ग्रंथ-चर्मपत्र को दोबारा ठीक नहीं किया जा सकता।

केवल बाहर से देखकर यह भी पता नहीं लगाया जा सकता था कि इसमें कोई अक्षर है या नहीं। कार्बनीकृत सतह पूरी तरह काली और बनी किसी निशान के थी। स्याही भी काली थी और चमड़ा भी काला था, इसलिए आँखों से भेद करना संभव नहीं था। अंदर कुछ लिखा हुआ है, यह भी परतों के आकार से लगाया गया केवल एक अनुमान था।

सांस्कृतिक वरिसत पर काम करने वाले लोग इस बात को हमेशा सबसे आगे रखते हैं। पुरावशेष दुनिया में केवल एक ही होता है और एक भी असफलता बर्दाश्त नहीं की जा सकती। इसलिए संरक्षण विज्ञान में बना छुएँ जाँच करने का सदिधांत है। इस प्रकार की जाँच पद्धतिको गैर-आक्रामक परीक्षण कहते हैं। आक्रामक का अर्थ भीतर तक चीर-फाड़ करना है और गैर-आक्रामक इसका विपरीत है। एन-गेदी ग्रंथ-चर्मपत्र ऐसा पुरावशेष था जिसके लिए केवल गैर-आक्रामक परीक्षण की ही अनुमति थी।

45 वर्षों का मौन

यह ग्रंथ-चर्मपत्र इज़राइल के पुरावशेष भंडार में चला गया। बाद में इज़राइल पुरावशेष प्राधिकरण ने इस पुरावशेष की देखभाल की ज़िम्मेदारी ली[1]। बक्से के भीतर इंतज़ार करना ही उस समय की एकमात्र संरक्षण वधि थी। पुरावशेष भंडार वह स्थान होता है जहाँ तापमान और आर्द्रता को स्थिर रखा जाता है। हवा, रोशनी और कीड़ों से बचाकर पुरावशेष को और अधिक नुकसान पहुँचने से रोका जाता है। जसि पुरावशेष को पढ़ा न जा सके, उसकी स्थितिको सुरक्षित रखने का कार्य तब भी जारी रहता है।

मृत सागर पांडुलिपियों के संरक्षण का दायित्व संभालने वाली पनीना शोर ने इसके प्रबंधन का कार्यभार संभाला। बाद में इस ग्रंथ-चर्मपत्र को पढ़ने वाले शोध में भी उनका नाम शामिल हुआ[1]।

यह इंतज़ार 1970 से 2015 तक जारी रहा[1][2]। 45 वर्षों तक कोई भी इसके भीतर के अक्षरों को नहीं देख सका। बाहरी सतह पूरी तरह काली होने के कारण यह जानना भी कठिन था कि इसमें अक्षर हैं भी या नहीं। विद्वानों के बीच यह बात भी कही जाने लगी थी कि यह एक ऐसा पुरावशेष है जसिं कभी पढ़ा नहीं जा सकेगा।

इस इंतज़ार का एक कारण था। 1970 की तकनीक से इस पडि को संभालना संभव नहीं था। किसी वस्तु के आंतरिक भाग को सूक्ष्म परतों में विभाजित करके स्कैन करने वाले उपकरण भी आम नहीं थे। यदि स्कैन कर भी लिया जाता, तो उस भारी डेटा की गणना करने के लिए कंप्यूटर उपलब्ध नहीं थे। पुरावशेष को वैसा ही छोड़ दिया गया और तकनीक के विकसित होने का इंतज़ार किया गया।

इस दौरान दुनिया भर में इसी तरह की चित्ताँ बनी रही। इटली के हर्कुलेनियम के ग्रंथ-चर्मपत्र भी इसी दीवार के सामने अटके हुए थे। उन्हें भौतिक रूप से खोलने के पुराने प्रयासों ने पुरावशेषों को केवल नष्ट ही किया था। दोनों पुरावशेषों का स्थान और सामग्री अलग थी, लेकिन समस्या एक ही थी। बना छुएँ आंतरिक भाग को पढ़ने की एक वधिकी आवश्यकता थी।

इस मौन को तोड़ने वाली कोई नई शारीरिक तकनीक नहीं थी। वह वस्तु को काटे बिना उसके भीतर झाँकने वाली इमेजिंग और कंप्यूटर गणना थी। सील्स शोध दल लंबे समय से बना छुएँ प्राचीन ग्रंथों को पढ़ने के तरीकों पर शोध कर रहा था। इस शोध को नेशनल साइंस फाउंडेशन और

Google का समर्थन प्राप्त था[2]। उनके द्वारा लंबे समय से विकसित पद्धतिका परीक्षण करने के लिए एक पुरावशेष की आवश्यकता थी, और एन-गेदी ग्रंथ-चर्मपत्र इसका लक्ष्य बना।

2015 में अमेरिका के केंटकी विश्वविद्यालय के ब्रैट सील्स शोध दल ने यह काम अपने हाथ में लिया। शोध के परिणाम 21 सितंबर 2016 को शोध पत्रिका 'साइंस एडवांसेज' (Science Advances) में प्रकाशित हुए[1][2]।

अब केवल एक ही प्रश्न शेष था। बर्ना छुए भीतर के अक्षरों को कैसे पढ़ा गया? इसका उत्तर इमेजिंग उपकरण और गणना के क्रम में नहिति है।

संदर्भ सामग्री

[1] Seales, W. B., Parker, C. S., Segal, M., Tov, E., Shor, P., & Porath, Y. (2016). From damage to discovery via virtual unwrapping: Reading the scroll from En-Gedi. Science Advances, 2(9), e1601247. <https://doi.org/10.1126/sciadv.1601247>

[2] University of Kentucky Research. (2016, 9월 22일). Virtually Unwrapping the En-Gedi Scroll. research.uky.edu. <https://research.uky.edu/news/virtually-unwrapping-en-gedi-scroll> (3 सितंबर 2026 को अभिगमिति)

2.2. फॉरेंसिक टीम की पड़ताल: X-ray डिजिटल शोधन

बर्ना काटे भीतर देखने वाली इमेजिंग

शोध दल द्वारा चुना गया पहला उपकरण माइक्रो-CT था। CT का अर्थ कंप्यूटेड टोमोग्राफी है। यह अस्पताल में शरीर के अंदर देखने के लिए उपयोग किए जाने वाले स्कैन के समान ही सिद्धांत पर काम करता है। विभिन्न कोणों से X-ray करिणें छोड़कर वस्तु के भीतर के घनत्व को मापा जाता है। कंप्यूटर उन मानों को एकत्र करके एक त्रि-आयामी नक्शा तैयार करता है। माइक्रो-CT एक ऐसा उपकरण है जो इसे अत्यंत सूक्ष्म पैमाने पर करता है।

एन-गेदी ग्रंथ-चर्मपत्र के स्कैन के लिए Bruker SkyScan 1176 नामक उपकरण का उपयोग किया गया था। स्कैनिंग का कार्य इजराइल की Merkel Technologies ने संभाला था। इस कंपनी ने स्कैनिंग डेटा इजराइल पुरावशेष प्राधिकरण को दान कर दिया। स्कैनिंग का पैमाना प्रतिग्रंथ 18 माइक्रोमीटर था[1]। 1 माइक्रोमीटर 1 मिलीमीटर का 1000वाँ हिस्सा होता है। यह मानव बाल की मोटाई के कुछ हिस्सों के बराबर आकार है।

इस पैमाने के आकार का एक कारण है। परतों को अलग-अलग देखने के लिए चमड़े की एक परत की तुलना में बहुत अधिक सूक्ष्मता से मापना पड़ता है। यदि पैमाना परत की मोटाई से बड़ा हो, तो दो परतें आपस में मलिकर एक पड़ि बन जाएँगी। स्कैनिंग का पैमाना उस सीमा को निर्धारित करने वाली शर्त है कि किस स्तर की परतों को पढ़ा जा सकता है।

स्कैनिंग से प्राप्त डेटा केवल एक तस्वीर नहीं है। यह अत्यंत छोटे घनाकार कणों से घनी तरह बना एक त्रि-आयामी पड़ि है। इन कणों को वोक्सेल कहा जाता है। यदि किसी तस्वीर का पक्सेल एक समतल कण है, तो वोक्सेल एक त्रि-आयामी कण है। प्रत्येक कण में उस स्थान के घनत्व को दर्शाने वाली एक संख्या दर्ज होती है।

इस चरण में पुरावशेष को बर्ना छुए ही कंप्यूटर के भीतर स्थानांतरित कर दिया गया। अब से नपिटा जाने वाला वषिय चमड़ा नहीं, बल्कि संख्याएँ थीं। संख्याओं को चाहे जतिना भी छुआ जाए, वे टूटती नहीं हैं। यदि गणना गलत हो जाए, तो शुरुआत से फेरि से किया जा सकता है। गैर-आक्रमक परीक्षण की शक्त यही नहिति है।

धातु-मश्रिति स्याही द्वारा छोड़े गए चमकदार बड़ि

त्रि-आयामी डेटा प्राप्त कर लेने मात्र से अक्षर दिखाई नहीं देते। अक्षर दिखने के लिए स्याही और पृष्ठभूमि के घनत्व में अंतर होना चाहिए। यह शर्त हर्कुलेनियम के ग्रंथ-चर्मपत्रों में वफिल हो गई थी। वहाँ की स्याही लकड़ी के कोयले के चूरे से बनाई गई थी और जला हुआ पपीरस भी कार्बन का एक पड़ि था। दोनों का घनत्व समान होने के कारण X-ray से उनमें आसानी से अंतर नहीं किया जा सका।

एन-गेदी ग्रंथ-चर्मपत्र की स्थिति अनुकूल थी। शोध दल ने स्कैन डेटा में पाया कि जहाँ अक्षर स्थिति थे, वहाँ का घनत्व आसपास की तुलना में अधिक था। इसलिए उन्होंने माना कि इस स्याही में लोहा या सीसा जैसी धातुएँ मली हुई होगी। हालाँकि, बाहरी सतह पर कोई खुला भाग न होने के कारण वे इसके घटकों का सीधे विश्लेषण नहीं कर सके। शोध दल ने इस सीमा का अपने शोधपत्र में स्पष्ट उल्लेख किया[1]।

धातु X-ray करिणों को अधिक अवशोषित करती है। अधिक अवशोषण वाले क्षेत्र स्क्रीन पर चमकीले दिखाई देते हैं। जला हुआ चमड़ा X-ray को कम अवशोषित करता है, इसलिए वह गहरा दिखाई देता है। इस प्रकार अक्षरों और पृष्ठभूमि के बीच चमक का अंतर उत्पन्न होता

है। यह अंतर कंप्यूटर द्वारा अक्षरों को खोजने का पहला सुराग बना। एन-गेदी ग्रंथ-चर्मपत्र को हरकुलेनियम से पहले पढ़े जाने का कारण भी यही है।

हालाँकि, इस स्याही के रासायनिक घटकों की अभी तक पुष्टि नहीं हुई है। शोध दल ने केवल यह अवलोकन प्रस्तुत किया कि इसका घनत्व अधिक था और उसी के आधार पर अनुमान लगाया [1]। प्राचीन मुंशी ने इसमें क्या मलिया था, यह नमूना प्राप्त होने पर ही जाना जा सकता है। जो बात प्रमाणित नहीं है, उसे प्रमाणित तथ्य की तरह न मानना ही इस क्षेत्र का सदिधांत है।

आभासी अनावरण के चार चरण और धब्बों को हटाना

चमकदार बटु खोज लेने मात्र से पाठ पढ़ा नहीं जा सकता। अक्षर गोल मुड़ी हुई परतों पर बखिरे हुए हैं। इन परतों को कंप्यूटर के भीतर समतल करना होगा। इस प्रक्रिया को आभासी अनावरण कहा जाता है। सील्स शोध दल ने आभासी अनावरण को चार चरणों में विभाजित किया [1]।

पहला चरण परतों की सतह की पहचान करना है। कंप्यूटर त्रि-आयामी डेटा को पतली परतों में काटकर चमड़े की परत के निर्देशांक तय करता है। इस प्रकार एकत्र किए गए बटुओं को जोड़कर एक पतली शीट के आकार की सतह बनाई जाती है। यदि परतें आपस में चपिकी हों, तो इस चरण में भटकाव हो जाता है। यदि दो परतों को एक परत समझकर पढ़ा जाए, तो अक्षर भी आपस में मलि जाएंगे।

दूसरा चरण उस सतह पर चमक के मान लागू करना है। सतह जहाँ से गुजरती है, उन निर्देशांकों के स्कैन मानों को पढ़कर सतह पर उकेरा जाता है। उच्च घनत्व वाली स्याही इस समय चमकीले स्ट्रोक के रूप में उभरती है।

तीसरा चरण सफ़ाई हुई सतह को समतल करना है। किसी त्रि-आयामी आकार को समतल तल पर दबाने से उसका रूप खिंचा जाता है या विकृति हो जाता है। यह ठीक वैसी ही समस्या है जैसे ग्लोब को कागज़ के नक्शे पर फैलाने पर ध्रुवीय क्षेत्र विसृत हो जाते हैं। गणना इस विकृति को यथासंभव कम करते हुए सतह को समतल करती है।

चौथा चरण अलग-अलग समतल की गई कई सतहों को आपस में जोड़ना है। सतहों की सीमाओं को मिलाकर उन्हें एक बड़ी समग्र तस्वीर में संयोजित किया जाता है [1]।

इन चार चरणों में पहला चरण सबसे कठिन है। यदि परत की पहचान में एक बार भी चूक हो जाए, तो बाद के तीनों चरण गलत हो जाते हैं। गलत तरीके से समतल की गई छवि में ऐसे अक्षर भी बन सकते हैं जो वास्तव में मौजूद नहीं थे। यही कारण है कि शोधकर्ता इस चरण की मैन्युअल रूप से पुष्टि करते हैं।

इन चार चरणों से गुजरने के बाद वह काला बेलन एक समतल पृष्ठ बन जाता है। शोध दल ने इस विधि से चमड़े की पाँच परतों को समतल किया। खोली गई सतह का क्षेत्रफल लगभग 94 वर्ग सेंटीमीटर आंका गया था [1]। यह नोटबुक के एक पन्ने से थोड़ा छोटा आकार है।

समतल की गई छवि में भी समस्याएँ बनी रहती हैं। चमड़े की मोटाई जगह-जगह अलग होती है और सलिवर्टे भी बाकी रहती हैं। स्कैनिंग प्रक्रिया के दौरान उत्पन्न हुआ नॉइज़ भी इसमें मलि जाता है। इसलिए कुछ हद से बहुत अधिक चमकीले और कुछ बहुत गहरे होते हैं। जब अक्षर और धब्बे आपस में मलि जाते हैं, तो पढ़ना कठिन हो जाता है।

शोध दल ने गणना के माध्यम से इन धब्बों को हटा दिया। सबसे पहले, आसपास के पक्सेल की औसत चमक को मापकर पृष्ठभूमि का एक ब्राइटनेस मैप बनाया जाता है। यह मैप केवल व्यापक रूप से फैली चमक के प्रवाह को दर्ज करता है। इसके बाद मूल छवि को इस पृष्ठभूमि मैप से विभाजित किया जाता है। इस विभाजन के बाद व्यापक रूप से फैला चमक का अंतर समाप्त हो जाता है। केवल संकीर्ण दायरे में केंद्रित स्याही के संकेत ही स्पष्ट रूप से बचते हैं। इस प्रक्रिया से गुजरी छवि में हबिबू अक्षर स्पष्ट रूप से उभर आए [1]।

2026 में भी जारी चुनौती

आभासी अनावरण का उपयोग बाद में अन्य पुरावशेषों पर भी किया गया। 25 जून 2026 को Vesuvius Challenge ने एक समाचार साझा किया। इसमें बताया गया कि एक हरकुलेनियम ग्रंथ-चर्मपत्र को आदि से अंत तक पढ़ लिया गया है। पुरावशेष संख्या PHerc. 1667 है। इमेजिंग फ्रांस के ग्रैनोबल स्थिति यूरोपीय सकिरोट्रॉन विकिरण सुविधा में की गई थी। खोला गया यूनानी पाठ लगभग 1.4 मीटर लंबा और बाइस कॉलम का है। यह ईसा पूर्व दूसरी शताब्दी की नीतिशास्त्र की रचना प्रमाणित हुई [2]।

सकिरोट्रॉन एक ऐसा उपकरण है जो तीव्र गति से चलने वाले इलेक्ट्रॉनों को चुंबकीय क्षेत्र की मदद से मोड़ता है। जब इलेक्ट्रॉन मुड़ते हैं, तो अत्यंत शक्तिशाली X-ray निकलती है। इस प्रकाश से स्कैन करने पर घनत्व का मामूली अंतर भी स्पष्ट हो जाता है। यह ऐसी स्पष्टता है जिसे प्रयोगशाला के इमेजिंग उपकरणों से प्राप्त करना कठिन है। यह एन-गेदी में प्रयुक्त उपकरण से एक स्तर उन्नत साधन है।

इस उपलब्धि में इमेजिंग पद्धति का सुधार शामिल था। नई इमेजिंग पद्धति में सामग्री के भीतर कार्बन स्याही सीधे दिखाई दी। घनत्व में कम अंतर के कारण जो स्याही दिखाई नहीं दे रही थी, उसे पकड़ लिया गया। एन-गेदी में स्वाभाविक रूप से प्राप्त चमक के अंतर को यहाँ इमेजिंग तकनीक से उत्पन्न किया गया। बताया गया कि शेष लगभग 600 ग्रंथ-चर्मपत्रों पर भी यही विधि लागू की जा सकती है [3]।

चुनौतियाँ अभी बाकी हैं। 10 जुलाई 2026 को प्रकाशित एक सारांश में शेष समस्याओं को छह भागों में गणित किया गया है। परतों के दबे हुए हिससों में X-ray बखिर जाती है और स्क्रीन धुंधली हो जाती है। परतें खोजने की गणना छेद कर देती है या दो परतों को आपस में चपिका देती है। वोक्सेल डेटा में यह दर्ज नहीं होता कि कौन सी परत किस परत से जुड़ती है। मनुष्यों द्वारा तैयार किए गए प्रशिक्षण लेबल कण स्तर पर सटीक नहीं होते। एक ग्रंथ-चर्मपत्र से सीखा मॉडल दूसरे ग्रंथ-चर्मपत्र पर ठीक से काम नहीं करता। स्याही परत की किस गहराई पर स्थिति है, यह भी स्पष्ट नहीं है[4]।

समाधान भी साथ में प्रस्तुत किए गए। यदि सही लेबल पर्याप्त न हों, तो डेटा से ही नियम सीखने का एक तरीका मौजूद है। स्वयं सीखने वाले मॉडल इसके प्रमुख विकल्प माने जाते हैं। परतें खोजने की गणना में भी नए न्यूरल नेटवर्क ढांचे का उपयोग किया जा रहा है[4]।

एक नया सुराग भी सामने आया। 29 मार्च 2026 को सार्वजनिक किए गए एक अध्ययन में इस तथ्य का सहारा लिया गया कि स्याही सतह के आकार को बदल देती है। शोधकर्ताओं ने पहले से खुले हुए पपीरस की सतह की बनावट को सटीकता से मापा और मशीन लर्निंग को सिखाया। इसके परिणामस्वरूप केवल सतह की बनावट से ही अक्षरों के स्थान की पहचान की जा सकी। इस अध्ययन में यह भी जांचा गया कि इस बनावट को पढ़ने के लिए किस स्तर के इमेजिंग पैमाने की आवश्यकता है[5]। यह एन-गेदी में घनत्व से मल्लि सुराग को दूसरे गुण में फरि से खोजने वाला अध्ययन है।

संदर्भ सामग्री

- [1] Seales, W. B., Parker, C. S., Segal, M., Tov, E., Shor, P., & Porath, Y. (2016). From damage to discovery via virtual unwrapping: Reading the scroll from En-Gedi. *Science Advances*, 2(9), e1601247. <https://doi.org/10.1126/sciadv.1601247>
- [2] Vesuvius Challenge. (2026, 6월 25일). An entire Herculaneum scroll has been read for the first time. [scrollprize.org](https://scrollprize.org/firstscroll). <https://scrollprize.org/firstscroll> (3 सितंबर 2026 को अभिगमिति)
- [3] ASNT. (2026, 8월 13일). X-Ray Imaging and AI Recover First Complete Text from Sealed Ancient Scroll. *Materials Evaluation*. <https://www.asnt.org/me/26/8/x-ray-ai-recover-text-sealed-ancient-scroll> (3 सितंबर 2026 को अभिगमिति)
- [4] Vesuvius Challenge. (2026, 7월 10일). Open Problems: Why Reading Every Herculaneum Scroll Is Still a Challenge. [scrollprize.org](https://scrollprize.org/2026_open_problems). https://scrollprize.org/2026_open_problems (3 सितंबर 2026 को अभिगमिति)
- [5] Angelotti, G., Nicolardi, F., Henderson, P., & Seales, W. B. (2026, 3월 29일). Ink Detection from Surface Topography of the Herculaneum Papyri. *arXiv:2603.27698*. <https://arxiv.org/abs/2603.27698> (3 सितंबर 2026 को अभिगमिति)

2.3. सामने आया सच

स्क्रीन पर उभरी लैव्यव्यवस्था

कंप्यूटर स्क्रीन पर उभरे अक्षर इब्रानी लिपि के थे। शोधकर्ताओं ने चमड़े की पाँच परतें खोलीं और उसके भीतर से दो कॉलम का पाठ निकाला। यह मूल पाठ पुराने नियम की तीसरी पुस्तक लैव्यव्यवस्था का था। पढ़ा गया भाग लैव्यव्यवस्था का अध्याय 1 और अध्याय 2 है[1]।

संयुक्त चित्र में ग्रंथ-चर्मपत्र के एक पृष्ठ की पैंतीस पंक्तियाँ शामिल थीं। इनमें से अठारह पंक्तियों में अक्षर सुरक्षित थे। शेष सत्रह पंक्तियों का पुनर्निर्माण विद्वानों ने किया। पुनर्निर्माण का अर्थ यह नहीं है कि कंप्यूटर मगदंत पाठ बनाता है। यह बचे हुए नशानों और ज्ञात पाठ का मिलाप करके छूटे हुए स्थानों को भरने का कार्य है। एक पंक्ति में अक्षरों और खाली स्थानों की संख्या पैंतीस से चौतीस थी[1]।

शोधपत्र में बची हुई पंक्तियों और पुनर्निर्मित पंक्तियों को अलग-अलग स्पष्ट करना महत्वपूर्ण है। ऐसा इसलिए है ताकि पाठक पहचान सकें कि किस सीमा तक तथ्य प्रमाणित हैं। यदि बची हुई सामग्री और भरी गई सामग्री को मिलाकर प्रस्तुत किया जाए, तो बाद के शोधकर्ता इसकी पुष्टि नहीं कर सकते। इसके कारण अन्य विद्वान उसी डेटा के आधार पर दोबारा समीक्षा कर सकें।

लैव्यव्यवस्था का अध्याय 1 होमबला के नियमों की व्याख्या करता है। होमबला वह बला है जिसमें वेदी पर भेंट को पूरी तरह जलाकर अर्पित किया जाता है। अध्याय 2 अन्नबला के नियमों से संबंधित है। इसमें महीन आटे में तेल मलाने और लोबान रखने की विधि बताई गई है। आग में जले हुए पवित्र संदूक से आग में जलाकर अर्पित की जाने वाली बला के नियम निकले हैं।

पुनर्निर्मित पाठ में लैव्यव्यवस्था का पहला श्लोक भी शामिल है[1]। यह वह वाक्य है जिसमें यहोवा ने मल्लिपाले तम्बू से मूसा को बुलाकर बात की थी। पुस्तक का आरंभिक भाग बचे होने का अर्थ है कि यह ग्रंथ-चर्मपत्र लैव्यव्यवस्था की पूरी पुस्तक की शुरुआत था।

दो कॉलम होने का भी एक विशेष अर्थ है। प्राचीन ग्रंथ-चर्मपत्रों में पाठ को लंबवत लंबे खानों में विभाजित करके लिखा जाता था। एक खाना भरने के बाद अगले खाने में जाया जाता था। दो कॉलम निकलने का अर्थ है कि ऐसे दो खाने लगातार सामने आए हैं।

केवल अक्षर ही नहीं पढ़े गए। लिपिकी द्वारा कलम दबाकर लिखे गए स्ट्रोक की मोटाई भी स्क्रीन पर दर्ज हुई। यह इस बात का प्रमाण है कि यह मनुष्य के हाथ से उतारी गई प्रतिलिपि है।

एक हजार वर्ष के अंतराल को भरना

शोधकर्ताओं ने ग्रंथ-चर्मपत्र की आयु भी मापी। रेडियोकार्बन डेटिंग से पता चला कि यह तीसरी से चौथी शताब्दी ईस्वी का है[1]। यह वधि पुरावशेष में बचे कार्बन के परिवर्तन को मापकर आयु निर्धारित करती है। यह परिणाम प्राचीन बाइबल पांडुलिपि शोध के एक लंबे अंतराल को भरता है।

पहले यह समझें कि यह अंतराल क्या है। कुमरान की गुफाओं से मिली मृत सागर पांडुलिपियों की बाइबल प्रतियाँ तीसरी शताब्दी ईसा पूर्व से पहली शताब्दी ईस्वी के बीच की हैं। इसके विपरीत, पूर्ण रूप में बची मानक इब्रानी बाइबल पांडुलिपियाँ लगभग 1000 ईस्वी की हैं। इन दोनों समूहों के बीच लगभग एक हजार वर्ष का अंतराल है। इस अवधि के दौरान पाठ को कैसे प्रतिलिपिकिया गया, यह बताने वाली पांडुलिपियाँ बहुत कम थीं।

अंतराल जितना बड़ा होता है, संदेह भी उतना ही बढ़ता है। ऐसा इसलिए है क्योंकि यह जांचने का कोई साधन नहीं था कि मध्यकालीन पांडुलिपियों ने प्राचीन पाठ को क्यों उतारा था या नहीं। विद्वानों के बीच यह विचार भी था कि मध्यकालीन प्रतिलिपिकारों ने पाठ को परिष्कृत किया होगा।

एन-गेदी ग्रंथ-चर्मपत्र इसी अंतराल के ठीक मध्य से सामने आया। और पढ़ा गया पाठ मध्यकालीन इब्रानी बाइबल के व्यंजन पाठ के पूरी तरह समान था[1]। इब्रानी भाषा में केवल व्यंजन लिखे जाते हैं और स्वर नहीं लिखे जाते। केवल उन व्यंजनों से लिखे गए पाठ को व्यंजन पाठ कहा जाता है। इस मध्यकालीन पाठ को मेसोराटिक पाठ कहा जाता है। यह वह मानक पाठ है जिसे यहूदी प्रतिलिपिकारों ने नियम बनाकर सुरक्षित रखा था। अनुच्छेदों के विभाजन का स्थान भी समान था[2]। यह परिणाम दर्शाता है कि सदियों के दौरान पाठ को बिना किसी बदलाव के उतारा गया था।

यह माना जाता है कि यहूदी प्रतिलिपिकार बाइबल की नकल करते समय नियमों का कड़ाई से पालन करते थे। अक्षरों को गिनकर पुष्टि करने की वधि भी ऐसे ही नियमों में से एक थी। एन-गेदी ग्रंथ-चर्मपत्र इस बात का ठोस भौतिक प्रमाण बन गया कि उन नियमों का वास्तव में पालन किया गया था। एक ही पांडुलिपि ने इस पुरानी बहस में एक महत्वपूर्ण साक्ष्य जोड़ दिया।

इस पांडुलिपि के साथ एक और कीर्तमान जुड़ गया। अब तक पवित्र संदूक से निकली पंचग्रंथ की पांडुलिपियों में यह सबसे प्राचीन है[1]। पंचग्रंथ पुराने नियम की पहली पाँच पुस्तकों का सामूहिक नाम है।

प्रमाणित तथ्य और शेष प्रश्न

जो स्पष्ट हो चुका है और जिस पर अभी चर्चा चल रही है, उन्हें अलग करके देखना चाहिए।

जो स्पष्ट हो चुका है, वह इस प्रकार है। पढ़ा गया पाठ लैव्यव्यवस्था का अध्याय 1 और अध्याय 2 है। कार्बन डेटिंग का मान तीसरी से चौथी शताब्दी ईस्वी की ओर संकेत करता है। व्यंजन पाठ मध्यकालीन मानक पाठ के समान है। स्याही का घनत्व चमड़े से अधिक होने के कारण इसमें धातु के घटक मिले होने की संभावना दिखती है[1]।

शेष समस्याएँ भी स्पष्ट हैं। स्याही के घटकों का प्रत्यक्ष विश्लेषण नहीं किया जा सका[1]। ऐसा इसलिए हुआ क्योंकि बाहर कोई खुली सतह न होने के कारण नमूना प्राप्त नहीं किया जा सका। हस्तलेख के आधार पर काल निर्धारण को अलग तरह से देखने वाला दृष्टिकोण भी मौजूद है। पुरालेखविद Ada Yardeni ने इसे पहली शताब्दी ईस्वी के अंत या दूसरी शताब्दी की शुरुआत का माना था। पुरालेखविद प्राचीन हस्तलेख की तुलना करके काल निर्धारित करने वाले विद्वान होते हैं। इसके विपरीत Drew Longacre ने तर्क दिया कि रेडियोकार्बन डेटिंग अधिक विश्वसनीय है[2]। इन दोनों विचारों पर अभी कोई अंतिम निर्णय नहीं निकला है।

रखरखाव की अवधि भी एक अनसुलझा प्रश्न है। ग्रंथ-चर्मपत्र तीसरी से चौथी शताब्दी ईस्वी में बनाया गया था। आराधनालय लगभग 600 ईस्वी में आग से नष्ट हो गया था[1]। यदि ऐसा है, तो यह पांडुलिपि दो सौ से अधिक वर्षों तक पवित्र संदूक में रखी रही थी। समुदाय ने इस पुरानी पांडुलिपि को इतने समय तक क्यों बनाए रखा, यह स्पष्ट नहीं हो पाया है। केवल यह अनुमान लगाया जाता है कि वे वरिष्ठतम में मिली पांडुलिपि को बहुत मूल्यवान मानते होंगे।

पढ़ा गया दायरा भी ग्रंथ-चर्मपत्र का केवल एक अंश है। जो परते खोली गई, वे बाहर की पाँच परतें थीं[1]। अंदर के गहरे हिस्से में परतें और अधिक मजबूती से दबी हुई हैं। दबे हुए हिस्सों में परतों को अलग न कर पाने की समस्या 2026 में भी बनी हुई है। हर्कुलेनियम ग्रंथ-चर्मपत्रों पर काम करने वाले शोधकर्ता आज भी इस चुनौती से जूझ रहे हैं[3]।

भीतरी भाग को पढ़ने के लिए दो सुधार एक साथ होने चाहिए। परतों को अधिक बारीक रूप से देखने के लिए सूक्ष्म इमेजिंग पैमाने की आवश्यकता है। आपस में चपिकी परतों को अलग करने वाली गणना में भी सुधार होना चाहिए। जब ये दोनों शर्तें पूरी हो जाएंगी, तो एन-गेदी ग्रंथ-चर्मपत्र को फरि से स्कैनिंग टेबल पर लाया जा सकता है। यह इसलिए संभव है क्योंकि पुरावशेष आज भी सुरक्षित है।

काल निर्धारण में सहायता करने वाली वधियाँ भी बढ़ रही हैं। 2025 में एक ऐसा मॉडल प्रस्तुत किया गया जिसने प्राचीन इब्रानी पांडुलिपियों के हस्तलेख को मशीन लर्निंग के माध्यम से सीखकर काल का अनुमान लगाया। इस मॉडल ने रेडियोकार्बन डेटिंग मान और हस्तलेख डेटा दोनों से सीखा। यह वधिसिद्दिक के ऐसे सूक्ष्म घुमाव को भी माप लेती है जिससे मानव आँख से पकड़ना कठिन होता है[4]। जब ऐसे साधन वकिसति हो जाएँगे, तो एन-गेदी ग्रंथ-चर्मपत्र के काल संबंधी विवाद की भी पुनः समीक्षा की जा सकेगी।

आभासी अनावरण द्वारा खोला गया मार्ग केवल बाइबलि पांडुलिपियों तक सीमित नहीं है। मुड़े हुए और मुहरबंद प्राचीन पत्रों पर भी यही वधि अपनाई गई। आसानी से नष्ट होने वाली पुस्तकों और दस्तावेजों के लिए भी इमेजिंग और गणना का उपयोग किया जा रहा है। स्पर्श किए बिना पढ़ने का सिद्धांत सामग्रियों को संभालने का बुनियादी नियम बनता जा रहा है।

कंप्यूटर द्वारा प्रस्तुत चित्र अपने आप में अंतमि उत्तर नहीं है। 2016 के शोध में कंप्यूटर का कार्य परतों को खोलकर चित्र बनाने तक ही सीमित था। उस चित्र से अक्षरों को पढ़ना और उनका अर्थ निकालना मनुष्य का काम था। इसलिए पुरालेखविदों को एक-एक अक्षर की फरि से जांच करनी पड़ती है। इस घटना से केवल एक पाठ ही प्राप्त नहीं हुआ। हाथ लगाने पर नष्ट हो जाने वाले पुरावशेषों से निपटने का एक नया तरीका सामने आया। यह क्रम है इमेजिंग से डेटा प्राप्त करना, गणना से परतों को खोलना और अंत में मनुष्य द्वारा समीक्षा करना। यह क्रम आगे आने वाले कई पुरावशेषों पर भी उसी रूप में लागू होता है।

एन-गेदी से पढ़ा गया लैव्यव्यवस्था का पाठ भी उसी समीक्षा से गुजरकर अकादमिक रिकॉर्ड का हिस्सा बना। 45 वर्षों की चुप्पी को तोड़ना मशीन और मनुष्य का संयुक्त प्रयास था। इमेजिंग उपकरण ने घनत्व मापा और गणना ने परतों को खोला। फरि मनुष्य ने अक्षरों को पढ़ा और उनका अर्थ समझा। अगले अध्याय में अक्षरों के स्वरूप की स्वयं जांच करने की चर्चा जारी रहेगी।

संदर्भ सामग्री

- [1] Seales, W. B., Parker, C. S., Segal, M., Tov, E., Shor, P., & Porath, Y. (2016). From damage to discovery via virtual unwrapping: Reading the scroll from En-Gedi. *Science Advances*, 2(9), e1601247. <https://doi.org/10.1126/sciadv.1601247>
- [2] En-Gedi Scroll. (2026, 9 मई 3 मई). Wikipedia. https://en.wikipedia.org/wiki/En-Gedi_Scroll (3 सितंबर 2026 को अभिगमिति)
- [3] Vesuvius Challenge. (2026, 7 मई 10 मई). Open Problems: Why Reading Every Herculaneum Scroll Is Still a Challenge. [scrollprize.org](https://scrollprize.org/2026_open_problems). https://scrollprize.org/2026_open_problems (3 सितंबर 2026 को अभिगमिति)
- [4] Popović, M., Dhali, M. A., Schomaker, L., van der Plicht, J., Rasmussen, K. L., La Nasa, J., Degano, I., Colombini, M. P., & Tigheelaar, E. (2025). Dating ancient manuscripts using radiocarbon and AI-based writing style analysis. *PLOS ONE*, 20(6), e0323185. <https://doi.org/10.1371/journal.pone.0323185>

अध्याय 3. यशायाह चर्मपत्र में छपि दो छायाएँ: मुंशी प्रमाणीकरण

3.1. केस फाइल

गुफा से मलि चमड़े का चर्मपत्र

1947 में मृत सागर के पश्चिम में स्थिति कुमरान क्षेत्र की एक गुफा में प्राचीन चमड़े का चर्मपत्र मिला। बकरियाँ चराने वाले बंदू लोगों ने चट्टानी गुफा के भीतर से मर्तबान और चर्मपत्र खोज निकाले। इन पुरावशेषों को मृत सागर पांडुलिपियों कहा जाता है। इसके बाद कई गुफाओं में उत्खनन जारी रहा। अब तक एकत्र की गई मृत सागर पांडुलिपियों को लगभग 930 प्रतियों में व्यवस्थिति किय गया है। हाथ से नकल करके लिखी गई पुस्तक को प्रतिया पांडुलिपि कहते हैं। अलग-अलग टुकड़ों में गनि तो चमड़े और पपीरस के टुकड़े कई हजार से अधिक हैं[1]। इनके लिखे जाने का समय तीसरी शताब्दी ईसा पूर्व से पहली शताब्दी ईसवी के बीच माना जाता है[2]।

मृत सागर अत्यधिक खारे पानी की झील है और इसके चारों ओर सूखा रेगस्तान है। यहाँ बारिश लगभग न के बराबर होती है और हवा शुष्क रहती है। गुफा के भीतर का तापमान भी बहुत अधिक नहीं बदलता। चमड़े और पपीरस के बनि सड़े बचे रहने का कारण यही है। इतने बड़े पैमाने पर प्राचीन दस्तावेजों का संग्रह अन्यत्र कहीं मुश्कलि से ही बचा है।

इन पुरावशेषों में एक ऐसा चर्मपत्र है जसिने वदिवानों का ध्यान खींचा। यह महान यशायाह चर्मपत्र है। अकादमिकि जगत में इसके लिए 1QIsaa प्रतीक का उपयोग किय जाता है। इसका अर्थ है कुमरान की पहली गुफा से मली यशायाह की पुस्तक की पहली प्रतिया यह ईसा पूर्व दूसरी शताब्दी में पशु के चमड़े पर हब्रू भाषा में लिखी गई प्रतिया है। इसकी लंबाई लगभग 7.3 मीटर है। इसमें पुराने नयिम की यशायाह की पुस्तक की लगभग पूरी सामग्री मौजूद है। यह पाठ लंबवत 54 स्तंभों में वभिजति हैं[2]। इस स्तंभ को तकनीकी शब्दावली में कॉलम कहा जाता है। मृत सागर की गुफाओं से मली बाइबलि की प्रतियों में इसकी संरक्षण स्थिति सबसे अच्छी है।

यह चर्मपत्र चमड़े के कई टुकड़ों को धागे से आपस में जोड़कर बनाया गया था। एक टुकड़े पर कई कॉलम लिखे जाते हैं और फरि अगले टुकड़े पर बढ़ा जाता है। पूरा लिखने के बाद चर्मपत्र को गोल लपेटकर रखा जाता है। पढ़ते समय एक सरि खोला जाता है और दूसरा सरि लपेटा जाता है। 'durumari' (स्कुरॉल) नाम यही से आया है।

इस प्रतिका मूल्य इसकी प्राचीनता से आता है। मृत सागर पांडुलिपियों में आज सुरक्षिति बची हब्रू बाइबलि की प्रतियों में सबसे पुरानी प्रतियाँ शामिल हैं[3]। महान यशायाह चर्मपत्र भी उनमें से एक है। इससे यह तुलना करने की सामग्री मलि गई कबाइबलि का मूल पाठ सदियों से कैसे आगे बढ़ा। लेकिन इस चर्मपत्र को तैयार करने वाले व्यक्तिके बारे में कुछ भी ज्ञात नहीं था।

कसिने लिखा, इसे लेकर बंटे वदिवान

चर्मपत्र को पढ़ना और यह पता लगाना कि इसे कसिने लिखा, दोनों अलग-अलग बातें हैं। अक्षर पूरी तरह पढ़े जा सकते हैं, लेकिन इस प्रतिया में नकल करने वाले का नाम नहीं है। ग्रंथ की नकल करने का काम करने वाले व्यक्तिको मुंशी कहा जाता है। इस चर्मपत्र पर मुंशी के हस्ताक्षर भी नहीं हैं और कोई तारीख भी नहीं है[2]। इसलिए वदिवानों को केवल लिखावट के रूप को देखकर ही लिखने वालों की संख्या का अनुमान लगाना पड़ा।

इस मुद्दे को लेकर वदिवानों में दो राय बन गई। एक पक्ष का मानना था कि एक ही मुंशी ने शुरू से अंत तक इसे अकेले लिखा था। उनका आधार यह था कि पूरे चर्मपत्र में लिखावट की बनावट एक समान है। दूसरे पक्ष का मानना था कि इसे दो या दो से अधिक लोगों ने मलिकर लिखा था। उनका आधार यह था कि मध्य भाग को पार करते ही स्ट्रोक का आकार थोड़ा-थोड़ा बदलने लगता है।

दोनों दावों के सामने एक कठनि समस्या थी। एक ही व्यक्तिकी लिखावट भी शुरुआत और अंत में बलिकुल एक जैसी नहीं होती। हाथ थक जाने पर स्ट्रोक ढीले हो जाते हैं और कलम बदलने पर मोटाई बदल जाती है। इस स्वाभाविकि वचिलन के वसितार को जाने बनि यह तय नहीं किय जा सकता कि लेखक बदला है या नहीं। दोनों ही दावे आँखों से देखे गए प्रभाव पर नरिभर थे। कौन सा पक्ष सही है, यह तय करने का कोई पैमाना नहीं था।

माना जाता है कि कुमरान में एक ऐसा समुदाय था जो चर्मपत्रों को बनाता और उनकी रक्षा करता था। लोगों ने चर्मपत्रों को मर्तबानों में रखकर गुफाओं में छपि दिया था। ऐसा लगता है कि युद्ध निकट आने पर उन्होंने दस्तावेजों की रक्षा करने का प्रयास किय था। छपिने वालों के नाम भी शेष नहीं हैं। चर्मपत्र को रचने वाले हाथों की खोज करना उन लोगों के करीब पहुँचने का एक रास्ता भी है।

यह विवाद केवल लिखावट की चर्चा तक सीमिति नहीं है। यदकिई लोगों ने मलिकर प्रतिया बनाई थी, तो इसका अर्थ है कि मुंशियों का एक संगठित समूह साथ मलिकर काम करता था। यह इस बात को उजागर करने से जुड़ता है कि बाइबलि का मूल पाठ कनि लोगों के हाथों से होकर आज तक पहुँचा[2]। यही कारण है कि मृत सागर पांडुलिपियों के शोधकर्ता लंबे समय तक इस सवाल से जुड़े रहे।

आँखों द्वारा की जाने वाली जाँच की सीमाएँ

प्राचीन लिपियों का अध्ययन करने वाले विषय को पुरालेखविद्या कहा जाता है। इसमें वदिवान आवर्धक लेंस लेकर अक्षरों के आकार की जाँच करते हैं। वे अक्षरों की ऊँचाई, झुकाव और स्ट्रोक की मोटाई को मापते हैं तथा उन्हें चित्र के रूप में उतारते हैं। कई प्रतियों की लिखावट को कालक्रमानुसार व्यवस्थिति करके वे बदलाव के प्रवाह को भी पकड़ते हैं। लंबे समय तक यही तरीका प्रमाणीकरण का एकमात्र साधन था।

इस पद्धति में एक कमजोरी है। यह बहुत हद तक वदिवान की दृष्टि और अनुभव पर निर्भर करती है। एक ही अक्षर को देखकर भी अलग-अलग वदिवान अक्सर अलग-अलग नष्टिर्ण नकिलते थे[2]। मानव आँख बहुत छोटे संख्यात्मक अंतरों को नहीं माप सकती। सरकंडे की कलम के दबाव की शक्त या स्याही के फैलने की मोटाई का आँखों से अंदाज़ा लगाना कठिन है। स्मृतिकी भी अपनी सीमा होती है। कई मीटर की दूरी पर स्थिति दो अक्षरों को दमिाग में एक साथ रखकर तुलना करना आसान नहीं है। अंततः केवल मतभेद ही बार-बार दोहराए जाते रहे।

एक और बाधा उस युग की शक्तिषण पद्धति थी। मुंशी बनने की इच्छा रखने वाले व्यक्तिको अपने गुरु की लिखावट की हूबहू नकल करने का प्रशक्तिषण दिया जाता था। उन्हें सिखाया जाता था कि वे अपनी व्यक्तिकी शैली को प्रकट न करें। परिणामस्वरूप एक ही पाठशाला से निकले मुंशियों की लिखावट देखने में लगभग एक जैसी होती थी। मानव वदिवानों के लिए यह एक घने कोहरे जैसा था। चर्मपत्र के अगले और पछिले हिस्से में आया बारीक अंतर किसी एक व्यक्तिकी शारीरिक स्थितिकी बदलाव था या लेखक का बदलना, यह तय कर पाना असंभव था[2]।

यदि प्रमाणीकरण गलत होता है, तो उसके बाद के नष्टिर्ण भी गलत हो जाते हैं। यदि मुंशी एक ही व्यक्ती था, तो एक प्रतिये एक ही व्यक्ती का समय लगा था। यदि दो लोग थे, तो इसका अर्थ है कि काम बाँटकर करने वाला कोई संगठन मौजूद था। प्रतितैयार करने वाले स्थान का आकार और स्वरूप बदल जाता है। लिखावट के केवल एक प्रमाणीकरण पर कई नष्टिर्ण टकें हुए थे।

कक्षा में एक ही सुलेख प्रतमिान को देखकर इमला लिखने वाले बच्चों को याद करें, तो स्थितिलगभग वैसी ही है। सबके द्वारा साफ-साफ लिखने की कोशिश की गई कॉपियों में बाहर से कोई अंतर नहीं दिखता। फिर भी स्ट्रोक मोड़ने की आदत हर बच्चे की थोड़ी अलग होती है। समस्या यह थी कि उस अंतर को मापने वाला कोई पैमाना नहीं था।

स्थिति तब बदली जब पांडुलिपिकी तस्वीरें डिजिटल रूप से सार्वजनिक की गईं। वास्तविक चर्मपत्र प्रकाश और नमी के प्रतिसंवेदनशील होते हैं, इसलिए उन्हें बार-बार खोला नहीं जा सकता। तस्वीरें होने पर पुरावशेष को छुए बिना जितनी बार चाहें उतनी बार देखा जा सकता है। इज़राइल पुरावशेष प्राधिकरण ने Google के साथ मिलकर मृत सागर पांडुलिपियों का एक डिजिटल संग्रह बनाया। उन्होंने MegaVision नामक उपकरण से विभिन्न तरंग दैर्ध्य के प्रकाश डालकर पुरावशेषों की तस्वीरें लीं। मानव आँख से न देखने वाले नकिट-अवरक्त प्रकाश का उपयोग करके धुंधले पड़ चुके अक्षरों को पुनर्जीवित किया गया। उनका लक्ष्य वास्तविक वस्तु को सीधे देखने के समान छवि गुणवत्ता प्राप्त करना था[1]।

कंप्यूटर द्वारा पढ़ी जा सकने वाली उच्च-रिज़ॉल्यूशन तस्वीरें एकत्र होने पर नए उपकरणों के आने का मार्ग प्रशस्त हुआ। तस्वीर संख्याओं का एक समूह होती है। स्क्रीन बनाने वाले प्रत्येक छोटे बटु को पक्सेल कहा जाता है। प्रत्येक पक्सेल का एक चमक मान होता है। इसका अर्थ है कि स्ट्रोक की मोटाई और कोण को संख्याओं में मापा जा सकता है। आँखों से कए जाने वाले मूल्यांकन को गणना में बदलने की तैयारी पूरी हो चुकी थी। यहाँ से एआई जासूस आगे आता है।

संदर्भ सामग्री

- [1] Israel Antiquities Authority. (n.d.). The Digital Library. The Leon Levy Dead Sea Scrolls Digital Library. <https://www.deadseascrolls.org.il/about-the-project/the-digital-library> (3 सितंबर 2026 को अभिगमिति)
- [2] Popović, M., Dhali, M. A., & Schomaker, L. (2021). Artificial intelligence based writer identification generates new evidence for the unknown scribes of the Dead Sea Scrolls exemplified by the Great Isaiah Scroll (1QIsaa). PLOS ONE, 16(4), e0249769. <https://doi.org/10.1371/journal.pone.0249769>
- [3] University of Groningen. (2021, 4월 21일). Cracking the code of the Dead Sea Scrolls. <https://www.rug.nl/news/2021/04/cracking-the-code-of-the-dead-sea-scrolls?lang=en> (3 सितंबर 2026 को अभिगमिति)

3.2. वैज्ञानिक जाँच दल की पड़ताल: लिखावट में छूटे बायोमेट्रिक फिंगरप्रिंट

जाँच दल का गठन

नीदरलैंड्स के ग्रोनगन विश्वविद्यालय के शोध दल ने इस मामले को संभाला। पुरालेखविद् म्लादेन पोपोविच ने दल का नेतृत्व किया। एआई शोधकर्ता लैम्बर्ट शोमेकर और मारूफ धारी उनके साथ जुड़े[1]। उन्होंने 2017 में कंप्यूटर द्वारा मृत सागर पांडुलिपियों की लिखावट को मापने की विधिसबसे पहले प्रकाशित की[2]। उस विधिको परीक्षित करके महान यशायाह चर्मपत्र पर लागू करने के परिणाम 2021 में सामने आए[1, 5]।

शोध दल का उद्देश्य स्पष्ट था। वह मानव व्यक्तिपरिक मूल्यांकन को संख्याओं में बदलना था। पहले एक ही व्यक्ति की लिखावट में होने वाले स्वाभाविक वचिलन के वसितार को मापा जाता है। उसके बाद यह देखा जाता है कि चर्मपत्र के अगले और पछिले हिस्से का अंतर उस वचिलन से बड़ा है या नहीं। यदि अंतर बड़ा हो, तो इसका अर्थ है कि इसे किसी अन्य व्यक्ति ने लिखा था।

2017 में वकिसति विशेषता गणना विधिको उपयोग प्रतियों की आपस में तुलना करने के लिए किया गया था[2]। शोध दल ने उस गणना विधिको विशेष रूप से केवल महान यशायाह चर्मपत्र पर दोबारा लागू किया। एक ही प्रतियों के भीतर मुंशी के बदलने का स्थान खोजना अलग-अलग प्रतियों की तुलना करने से अधिक कठिन है। क्योंकि लिखावट का अंतर बहुत सूक्ष्म होता है।

इस दृष्टिकोण का एक बड़ा लाभ है। गणना की प्रक्रिया दर्ज होने पर कोई भी उसी तस्वीर से वही उत्तर प्राप्त कर सकता है। विद्वान की व्यक्तिगत अंतरदृष्टि पर निर्भर नहीं रहना पड़ता। यदि परिणाम गलत हो, तो यह दोबारा जाँचा जा सकता है कि किस चरण में गलती हुई। प्रमाणीकरण के परिणामों पर फिर से विवाद होने पर भी बहस का आधार संख्याओं के रूप में बना रहता है।

दाग हटाने का पहला चरण

पहला काम तस्वीर में से केवल स्याही को बाकी रखना है। इस प्रक्रिया को बाइनराइजेशन (द्विआधारीकरण) कहा जाता है। यह हर पिकसेल को स्याही और पृष्ठभूमि में से किसी एक में अलग करने की प्रक्रिया है। यह तस्वीर को काले अक्षरों और चमकदार पृष्ठभूमि में विभाजित करने का काम है।

इस चरण की आवश्यकता का एक कारण है। चर्मपत्र लगभग दो हजार वर्षों तक गुफा में रहा था। चमड़े की सतह पर दाग और फफूंद के निशान पड़ गए थे तथा बारीक दरारे भी आ गई थी। कंप्यूटर की दृष्टि में ऐसे निशान और स्याही के स्ट्रोक दोनों ही गहरे रंग के पिकसेल हैं। यदि पृष्ठभूमि को वैसा ही छोड़ दिया जाए, तो मशीन लिखावट के बजाय चमड़े की बनावट को सीख लेगी। तब वह मुंशी के बजाय चमड़े के प्रकारों में भेद करने लगेगी और प्रमाणीकरण गलत हो जाएगा[1]।

शोध दल ने इस कार्य के लिए एक अलग तंत्रिका नेटवर्क बनाया। उसका नाम BiNet है। तंत्रिका नेटवर्क मानव मस्तिष्क की तंत्रिका कोशिकाओं के जुड़ाव की नकल करके बनाई गई एक गणना संरचना है। BiNet हाथ से स्याही और पृष्ठभूमि को चिह्नित किए गए नमूनों को देखकर निर्णय के मापदंड सीखता है। प्रशिक्षण पूरा करने के बाद BiNet नई तस्वीरों में से दाग-धब्बों को हटाकर केवल स्ट्रोक छोड़ता है।

यहाँ एक नियम का पालन किया जाना चाहिए। मूल स्ट्रोक की मोटाई और कनारों के आकार से छेड़छाड़ नहीं की जानी चाहिए। क्योंकि मुंशी की आदतें इन्हीं कनारों में छिपी होती हैं। दाग मटाने के चक्कर में यदि स्ट्रोक को भी चिकना कर दिया जाए, तो सबूत नष्ट हो जाता है। BiNet को स्याही के निशान वैसा ही बनाए रखने के लिए डिज़ाइन किया गया था[1]। यह तकनीक सबसे पहले 2019 में एक सार्वजनिक मसौदे के रूप में सामने आई थी[3]। इसके डिज़ाइन और प्रदर्शन को संकलित करने वाला शोधपत्र 2026 में अंतरराष्ट्रीय दस्तावेज़ पहचान अकादमिकी शोध-संग्रह में प्रकाशित हुआ था[4]।

पूरी लिखावट की आदतों को मापने वाला पैमाना

साफ़ किए गए अक्षरों से दो प्रकार की विशेषताएँ निकाली जाती हैं। पहली बनावट से संबंधित विशेषता है और दूसरी आकार से संबंधित विशेषता है।

बनावट से संबंधित विशेषता को Hinge (हजि) कहा जाता है। Hinge दरवाज़े के कब्जे को कहते हैं। जहाँ भी स्ट्रोक मुड़ता है, वहाँ दो दिशाओं के कोण मापे जाते हैं। उन कोण युग्मों की आवृत्ति की एक तालिका बनाई जाती है। यह तालिका लिखते समय कलाई के घूमने की आदत को संख्याओं में दर्ज करती है। प्रत्येक अक्षर को अलग-अलग पहचाने बिना भी इसकी गणना की जा सकती है[1]।

आकार से संबंधित विशेषता को Fraglet (फ्रेगलेट) कहा जाता है। यह स्ट्रोक के छोटे कटे हुए टुकड़ों को संदर्भित करने वाला शब्द है। स्ट्रोक को काटकर इकट्ठा किया जाता है और एक सूची बनाई जाती है। शोध दल ने कई प्रतियों से निकाले गए टुकड़ों से एक मानक सूची बनाई। जब नई लिखावट आती है, तो यह गिना जाता है कि वह टुकड़ा सूची के किस घटक से मिला-जुलता है। मनुष्य द्वारा अक्षर सीखते समय सीखे गए स्ट्रोक का मूल रूप यहाँ प्रकट होता है[1]।

इन दोनों वशिषताओं में एक समानता है। वह यह कि अक्षरों का अर्थ जाने बना भी इनकी गणना की जा सकती है। हबिरू न पढ़ पाने वाला कंप्यूटर भी स्ट्रोक के कोण और घुमावदार आकार को माप सकता है। क्षतग्रिस्त अक्षर मलि होने पर भी बचे हुए स्ट्रोक से मान निकाले जाते हैं। प्राचीन दस्तावेजों से नपिते समय यह बहुत उपयोगी गुण है।

ये दोनों वशिषताएँ अलग-अलग चीजों को देखती हैं। Hinge हाथ की गति को देखता है और Fraglet सीखी गई लिखावट की शैली को देखता है। शोध दल ने दोनों का अलग-अलग भी उपयोग किया और मिलाकर भी देखा। यह सुनिश्चित करने के लिए कि क्या अलग-अलग पैमाने एक ही उत्तर देते हैं[1]।

अलेफ़ के एक अक्षर में छूटे नशान

पछिली दो वशिषताएँ पूरी लिखावट को मिलाकर समग्र रूप से मापती हैं। अब किसी एक अक्षर को बड़ा करके गहराई से देखने की बारी है।

शोध दल ने हबिरू वर्णमाला के एक अक्षर को अलग से चुनकर उसकी सूक्ष्म जाँच की। वह अक्षर अलेफ़ है। अलेफ़ हबिरू वर्णमाला का पहला अक्षर है और चर्मपत्र में सबसे अधिक बार आता है। इसकी आकृति एक विकर्ण रेखा और उससे जुड़े दो छोटे स्ट्रोक जैसी होती है।

अलेफ़ को चुनने का एक कारण है। नमूने जतिने अधिक होंगे, सांख्यिकी उतनी ही पुख्ता होगी। अक्षर जतिनी बार आएगा, आगे और पीछे के हिससों से उसे उतना ही समान रूप से एकत्र किया जा सकता है। पूरे चर्मपत्र में अलेफ़ के 5,011 अक्षर हैं। शोध दल ने कंप्यूटर द्वारा उनमें से 758 को चुना और विश्लेषण में उपयोग किया[1]। केवल उन्हीं अक्षरों को छांट लिया जानिके स्ट्रोक टूटे नहीं थे और पूरी तरह सुरक्षित थे।

चुने गए अलेफ़ को दो समूहों में विभाजित किया गया। चर्मपत्र के अगले भाग से मलि अक्षर और पछिले भाग से मलि अक्षर। प्रत्येक समूह की छवियों को एक ही स्थान पर परत-दर-परत रखकर औसत आकार बनाया गया। यह अर्धपारदर्शी कागज़ की कई शीटों को एक के ऊपर एक रखकर ऊपर से देखने जैसा है। अक्सर दोहराए जाने वाले स्ट्रोक गहरे हो जाते हैं और व्यक्ति-दर-व्यक्ति भिन्न स्ट्रोक धुंधले पड़ जाते हैं। दोनों औसत चित्रों को एक-दूसरे से घटाने पर केवल अंतर ही शेष बचता है। वह अंतर किस स्ट्रोक पर केंद्रित है, इसे रंगों के माध्यम से देखा जा सकता है[1]।

दो समूहों में बंटे बटु

संख्याओं में परिवर्तित वशिषताएँ सैकड़ों मानों की एक लंबी सूची होती हैं। मनुष्य आँखों से ऐसी सूची की तुलना नहीं कर सकता। इसलिए शोध दल ने सूची को दो मानों में घटाकर समतल पर बटुओं के रूप में अंकित किया। समान मान वाले अक्षर पास-पास आ जाएँ और भिन्न मान वाले दूर हो जाएँ, इस तरह उन्हें संक्षिप्त किया गया। यह बटुओं की स्थितिको धीरे-धीरे स्थानांतरित करने की गणना को दोहराकर तैयार किया गया चित्र है।

54 स्तंभों में से जनि स्तंभों की तस्वीरें पूरी तरह सुरक्षित थीं, उन सभी को इस समतल पर रखा गया। स्तंभ संख्या 16 और 46 को छोड़ दिया गया क्योंकि उनकी उपयोगी तस्वीरें उपलब्ध नहीं थीं। जैसे ही बटु अंकित हुए, दो समूह उभर कर सामने आए। आगे के स्तंभों ने एक समूह बनाया और पीछे के स्तंभों ने दूसरा समूह बनाया[1]। दोनों समूहों के बीच एक स्पष्ट अंतर दिखाई दे रहा था।

शोध दल ने जांच की कि यह विभाजन महज एक संयोग था या नहीं। उन्होंने कार्डी-वर्ग परीक्षण और टी-परीक्षण जैसी सांख्यिकीय विधियों का उपयोग किया। कार्डी-वर्ग परीक्षण यह जांचने की गणना है कि अंतर संयोगवश आया है या नहीं। टी-परीक्षण दो समूहों के औसत अंतर को आंकने की गणना है। दोनों विधियाँ संयोगवश ऐसा पैटर्न उभरने की संभावना निकालती हैं। यदि विह संभावना बहुत कम हो, तो इसे संयोग नहीं माना जाता। केवल एक व्यक्ति की लिखावट में होने वाले स्वाभाविक उतार-चढ़ाव से इतना बड़ा अंतर नहीं आया था[1]।

कई मापदंडों का एक ही उत्तर देना भी महत्वपूर्ण है। हजि से मापने पर भी और फ्रैगलेट से मापने पर भी उसी जगह सीमा तय हुई। अलेफ़ की औसत छवि में भी वही जगह विभाजित हुई। यह सीमा कॉलम 27 और कॉलम 28 के बीच स्थित थी[1]। यदि अलग-अलग तरीके एक ही जगह की ओर इशारा करते हैं, तो वह सीमा तस्वीर की किसी खामी के कारण नहीं बनी थी। इंसानी आंख से जो रेखा छूट गई थी, मशीन ने उसे खींचकर दिखा दिया।

शोध दल ने इसके विपरीत संभावना पर भी विचार किया। यह परिणाम तस्वीर की चमक या आगे-पीछे के चर्मपत्र की स्थिति में अंतर के कारण भी हो सकता था। लेकिन बाइनरीकरण द्वारा पृष्ठभूमि हटाने के बाद भी वह सीमा रेखा वैसी ही बनी रही[1]। इसका अर्थ है कि इसका कारण चर्मपत्र नहीं, बल्कि स्वयं लिखावट थी।

संदर्भ सामग्री

[1] Popović, M., Dhali, M. A., & Schomaker, L. (2021). Artificial intelligence based writer identification generates new evidence for the unknown scribes of the Dead Sea Scrolls exemplified by the Great Isaiah Scroll (1QIsaa). PLOS ONE, 16(4), e0249769. <https://doi.org/10.1371/journal.pone.0249769>

- [2] Dhali, M. A., He, S., Popović, M., Tigchelaar, E., & Schomaker, L. (2017). A digital palaeographic approach towards writer identification in the Dead Sea Scrolls. *Proceedings of the 6th International Conference on Pattern Recognition Applications and Methods*, 693-702. <https://doi.org/10.5220/0006249706930702>
- [3] Dhali, M. A., de Wit, J. W., & Schomaker, L. (2019). BiNet: Degraded-manuscript binarization in diverse document textures and layouts using deep encoder-decoder networks. *arXiv:1911.07930*. <https://arxiv.org/abs/1911.07930> (3 सितंबर 2026 को अभिगमिति)
- [4] Dhali, M. A., de Wit, J. W., & Schomaker, L. (2026). BiNet: A deep encoder-decoder network for binarizing degraded ancient manuscripts. In *Document Analysis and Recognition - ICDAR 2025 Workshops (Lecture Notes in Computer Science*, pp. 166-193). Springer Nature Switzerland. https://doi.org/10.1007/978-3-032-09371-4_11
- [5] University of Groningen. (2021, 4월 21일). Cracking the code of the Dead Sea Scrolls. [rug.nl](https://www.rug.nl/news/2021/04/cracking-the-code-of-the-dead-sea-scrolls?lang=en). <https://www.rug.nl/news/2021/04/cracking-the-code-of-the-dead-sea-scrolls?lang=en> (3 सितंबर 2026 को अभिगमिति)

3.3. सामने आया सच

कॉलम 27 पर बदला हाथ

वश्लेषण के परिणामों ने बहु-प्रतलिपिकार सदिधांत का समर्थन किया। महान यशायाह ग्रंथ-चर्मपत्र को दो लोगों ने बांटकर लिखा था। पहले प्रतलिपिकार ने कॉलम 1 से कॉलम 27 तक लिखा। दूसरे प्रतलिपिकार ने कॉलम 28 से कॉलम 54 तक लिखना जारी रखा[1]।

यह सीमा अन्य प्रमाणों से भी मेल खाती है। कॉलम 27 का नचिला हस्सा तीन पंक्तियों की ऊंचाई जतिना खाली है। कॉलम 28 चमड़े की एक नई जोड़ी गई शीट से शुरू होता है। वह स्थान यशायाह की पुस्तक का नया अध्याय शुरू होने की जगह भी है[1]। काम आगे सौपने के लिए यह एक उपयुक्त पड़ाव था। कंप्यूटर द्वारा खोजी गई सांख्यिकीय सीमा और ग्रंथ-चर्मपत्र की भौतिक सीमा एक ही स्थान पर थी।

कॉलम 27 के बाद प्रतलिपिकार बदल गया होगा, यह अनुमान पहले भी था। यह राय चमड़े की शीट जुड़ने के स्थान और मूल पाठ के विभाजन को देखकर सामने आई थी। कति यह अकादमिकी जगत में आम सहमत नहीं बन सकी थी। एआई वश्लेषण ने इस अनुमान को सांख्यिकीय आंकड़े प्रदान कर दिए[2]।

इस खोज का बहुत बड़ा महत्व है। बाइबल की एक पांडुलिपि को कई लोगों ने मलिकर तैयार किया था, इसका सांख्यिकीय प्रमाण सामने आया[1]। दोनों प्रतलिपिकारों ने न तो अपना नाम छोड़ा और न ही कोई हस्ताक्षर। उन्होंने बस अपने हाथों के संचालन की आदतें छोड़ी थीं। 2,000 साल बाद की मशीन ने उन आदतों को पढ़ लिया। पोपोविच ने कहा कि भले ही हम उनके नाम कभी न जान पाएं, लेकिन उनकी लिखावट के जरिए ऐसा लगता है मानो हम उनसे हाथ मिला रहे हों[2]।

एक ही वदियालय में सीखे दो लोग

दोनों प्रतलिपिकारों की लिखावट अब भी एक-दूसरे से बहुत मिलती-जुलती है। मशीन द्वारा मापने के बाद भी अंतर बहुत सूक्ष्म था। शोध दल ने इस समानता को ही एक अन्य प्रमाण के रूप में देखा। इसका अर्थ है कि दोनों की लेखन शैली एक-दूसरे के बहुत नजदीक थी[1]। शोध दल का मानना था कि दोनों प्रतलिपिकारों ने एक ही जगह प्रशिक्षण लिया था या वे एक ही परंपरा से संबंध रखते थे[2]।

यहाँ प्राचीन प्रतलिपिकार वदियालयों का दृश्य उभरता है। प्रशिक्षु एक-एक स्ट्रोक के स्तर पर मानक लिपि शैली सीखते थे। अपनी व्यक्तिगत शैली को दबाकर ही कोई एक अच्छा प्रतलिपिकार बन पाता था। उसी कड़े अनुशासन की बदौलत पूरे ग्रंथ-चर्मपत्र में लिखावट एक समान बनी रही।

फरि भी शरीर की स्वाभाविक आदतों को मटिया नहीं जा सकता था। नरकट की कलम पकड़ने का कोण हर व्यक्ति का अलग होता है। हाथ चलाने की गति और दबाव भी भिन्न होता है। यह अंतर स्ट्रोक के मुड़ने के कोण और उसके कनारों की बनावट में रह जाता है। इंसानी आंख इन सूक्ष्म अंतरों को एक साथ देखकर नरिणय नहीं कर सकती। कंप्यूटर हजारों मानों को संकलित कर औसत और वचिलन की गणना करता है। अनुशासन के पीछे छिपे व्यक्तिगत नशान इस तरह उजागर हुए[1]।

यह परिणाम दिखाता है कि पांडुलिपि तैयार करना एक सामूहिक कार्य था। एक लंबा ग्रंथ-चर्मपत्र बनाने में कई महीने लग जाते हैं। किसी एक ही व्यक्ति के अंत तक इस पर लगे रहने का कोई कारण नहीं था। यहीं से काम बांटने और एक-दूसरे को सौपने की संभावना निकलती है[1]। बाइबल का यह पाठ ऐसे ही हाथों-हाथ गुजरते हुए आज तक पहुंचा है।

जो काम मशीन अभी नहीं कर सकती

इस उपलब्धि का अर्थ यह नहीं है कि एआई सर्वशक्तिमान है। उसी शोध दल द्वारा वकिसति अगला उपकरण इसकी सीमाओं को दर्शाता है।

शोध दल ने 2025 में Enoch नामक काल-नरिधारण मॉडल पेश किया। Enoch लखावट के आकार को देखकर पांडुलिपि लिखे जाने के समय का अनुमान लगाने वाला मॉडल है। रेडियोकार्बन काल-नरिधारण प्राचीन वस्तुओं में बचे कार्बन से उनकी आयु मापने की वधि है। Enoch ने उस वधि से आयु नरिधारित किए गए 24 नमूनों से सीखा। अज्ञात काल की 135 पांडुलिपियों पर इसे लागू करने के परिणाम भी साथ में प्रकाशित किए गए। इसमें भवषियवाणी की त्रुटि औसतन 27.9 वर्ष से 30.7 वर्ष के बीच रही। पुरालेखवदों द्वारा इन अनुमानों की समीक्षा करने पर 79 प्रतिशत को उचित पाया गया[3]।

शेष 21 प्रतिशत पर विवाद खड़ा हो गया। पुरालेखवदि क्रिस्टोफर रोलस्टन ने एक उदाहरण देकर इसका खंडन किया। Enoch ने दानियेल की पुस्तक वाली 4Q114 पांडुलिपि को 230 ईसा पूर्व से 160 ईसा पूर्व के बीच का माना था। रोलस्टन का मानना था कि दानियेल के अध्याय 7 से 12 एक वशिष्ट घटना की ओर संकेत करते हैं। वह घटना 167 ईसा पूर्व में एंटोकोस IV द्वारा यरुशलम मंदिर को अपवर्तित किया जाना है। वह पाठ उस घटना से पहले नहीं लिखा जा सकता था। किंतु Enoch द्वारा सुझाई गई सीमा अधिकांशतः 167 ईसा पूर्व से पहले की है। रोलस्टन ने बताया कि Enoch का काल-नरिधारण बहुत पहले का है। उन्होंने कहा कि Enoch को टूलबॉक्स के कई उपकरणों में से बस एक होना चाहिए[4]।

यह विवाद एक अहम बात सिखाता है। मशीन केवल अक्षरों की बनावट देखती है। वह पाठ में दर्ज ऐतिहासिक घटनाओं या वाक्यों के अर्थ नहीं समझ सकती। इसलिए मशीन का उत्तर अंतिम निष्कर्ष नहीं, बल्कि एक संभावना मात्र है। संभावनाओं को सीमित करने के बाद इतिहासकारों को संदर्भ से मिलाकर करके नरिणय लेना होता है। किस अक्षर को मापना है, यह तय करने वाले भी पुरालेखवदि ही थे। अलेफ को चुनने का फैसला प्राचीन पांडुलिपियों के दीर्घकालिक अध्ययन के अनुभव से आया था।

प्रशिक्षण सामग्री की सीमा सिख्या भी एक बाधा है। Enoch ने ज्ञात काल वाले केवल 24 नमूनों से सीखा था[3]। प्राचीन अवशेषों के नमूने बढ़ाना मुश्किल होता है। काल-नरिधारण के लिए पुरावशेष का एक छोटा हिस्सा काटकर जलाना पड़ता है। सीमा नमूनों पर प्रशिक्षित मॉडल सीखे गए दायरे से बाहर जाते ही डगमगाने लगता है। प्रतिलिपिकार की पहचान की स्थिति भी ऐसी ही है। बार-बार आने वाले अक्षरों के नमूने तो पर्याप्त हैं, लेकिन दुर्लभ अक्षर गणि-चुने ही होते हैं। जहाँ नमूने कम हों, वहाँ मशीन का नरिणय कमजोर पड़ जाता है।

इसलिए शोध दल भी एआई को एकमात्र नरिणायक नहीं मानता। चमड़े के जुड़ने का स्थान, मूल पाठ का वभाजन और वदिवानों के संचित विचार एक साथ रखे जाते हैं। जब कई प्रमाण एक ही बटु की ओर इशारा करते हैं, तभी कोई निष्कर्ष निकलता है। महान यशायाह ग्रंथ-चर्मपत्र का उत्तर भी इसी तरह मिला था। मशीन ने केवल मनुष्य द्वारा उठाए गए सवालों का आंकड़ों में उत्तर दिया था।

अगली जांच का लक्ष्य

महान यशायाह ग्रंथ-चर्मपत्र का अध्ययन एक और बड़ी परियोजना में बदल गया। यूरोपीय अनुसंधान परिषद ने 23 जून 2026 को पोपोविच को एडवांस्ड ग्रांट देने का फैसला किया। अनुसंधान अनुदान 25 लाख यूरो है और इसकी अवधि 5 वर्ष है[5]। नई परियोजना का नाम Tracing Scribes and Scrolls है[6]। इसका अर्थ है प्रतिलिपिकारों और ग्रंथ-चर्मपत्रों के पदचिह्नों का पीछा करना।

इज़राइल पुरातत्व प्राधिकरण ने 30 जून 2026 को इस परियोजना में शामिल होने की घोषणा की। पीसा विश्वविद्यालय, नेपल्स फेडेरिको द्वितीय विश्वविद्यालय, दक्षिणी डेनमार्क विश्वविद्यालय और ल्यूवेन विश्वविद्यालय भी इसमें साथ हैं। बर्लिन और ट्यूरिन के मस्तिष्क संग्रहालय भी इसमें शामिल हुए[6]।

इस बार का सवाल 'किसने लिखा' से हटकर 'कहाँ बनाया गया' पर आ गया है। शोध दल मृत सागर पांडुलिपियों के लगभग 250 नमूनों की जांच करता है। वे चर्मपत्र, पपीरस और स्याही की रासायनिक संरचना को मापते हैं। उस संरचना की तुलना प्राचीन मस्तिष्क के पपीरस की संरचना से की जाती है। सामग्री कहाँ से आई, यह जानकर पांडुलिपिबिन्ने के स्थान का दायरा सीमित किया जा सकता है। एआई इस रासायनिक डेटा को व्यवस्थित करने और लखावट में छिपी नयिमों को खोजने का कार्य करता है[7]। पोपोविच ने कहा कि प्रयोगशाला विश्लेषण और प्राचीन लखावट के अध्ययन को जोड़कर हम उन सवालों तक पहुँच सकते हैं जहाँ पहले नहीं पहुँचा जा सका था[6]।

यह वधि केवल महान यशायाह ग्रंथ-चर्मपत्र तक सीमित नहीं है। मृत सागर पांडुलिपियों में ऐसी कई प्रतियाँ मौजूद हैं जिनके प्रतिलिपिकार अब भी अज्ञात हैं। एक ही हाथ से लिखी गई पांडुलिपियों को जोड़कर किसी एक प्रतिलिपिकार के कार्यों की सूची बनाई जा सकती है। ऐसी सूचियाँ जमा होने पर यह समझा जा सकता है कि प्रतिलिपिकार आपस में कैसे जुड़े थे[1]। दो व्यक्तियों को खोजने वाला यह शोध उस यात्रा का प्रारंभिक बटु बना।

इस शोध की एक और उपलब्धि स्वयं इसकी पद्धति है। तस्वीरों से खामियाँ दूर करना, स्त्रोक को संख्याओं में बदलना और सांख्यिकी से परीक्षण करने का क्रम सुव्यवस्थित हो गया। अन्य भाषाओं के प्राचीन दस्तावेजों पर भी यही प्रक्रिया लागू की जा सकती है। इस पुस्तक के आगामी अध्यायों में मिलने वाले मट्टी की पट्टिकाओं और काष्ठ पट्टिकाओं के अध्ययन भी इसी राह पर चलते हैं।

महान यशायाह ग्रंथ-चर्मपत्र से शुरू हुई जांच अभी समाप्त नहीं हुई है। दो व्यक्तियों की परछाई खोजने वाला उपकरण अब उस स्थान की तलाश में निकल पड़ा है जहाँ यह ग्रंथ-चर्मपत्र बनाया गया था।

प्राचीन दस्तावेजों को पढ़ना केवल अक्षरों की पहचान पर खत्म नहीं होता। यह भी पूछना आवश्यक है कि उन अक्षरों को किसने और कहाँ लिखा था। महान यशायाह ग्रंथ-चर्मपत्र 'इसे कतिने लोगो ने लिखा' के सवाल का आंकड़ों में जवाब देने वाला पहला उदाहरण है। ग्रंथ-चर्मपत्र 2,000 वर्षों तक मौन रहा। मनुष्य अब लिखावट के सूक्ष्म कंपन से वह उत्तर पढ़ रहा है।

संदर्भ सामग्री

- [1] Popović, M., Dhali, M. A., & Schomaker, L. (2021). Artificial intelligence based writer identification generates new evidence for the unknown scribes of the Dead Sea Scrolls exemplified by the Great Isaiah Scroll (1QIsaa). PLOS ONE, 16(4), e0249769. <https://doi.org/10.1371/journal.pone.0249769>
- [2] University of Groningen. (2021, 4월 21일). Cracking the code of the Dead Sea Scrolls. rug.nl. <https://www.rug.nl/news/2021/04/cracking-the-code-of-the-dead-sea-scrolls?lang=en> (3 सितंबर 2026 को अभिगमिति)
- [3] Popović, M., Dhali, M. A., Schomaker, L., van der Plicht, J., Rasmussen, K. L., La Nasa, J., Degano, I., Colombini, M. P., & Tigchelaar, E. (2025). Dating ancient manuscripts using radiocarbon and AI-based writing style analysis. PLOS ONE, 20(6), e0323185. <https://doi.org/10.1371/journal.pone.0323185>
- [4] Steinmeyer, N. (2025, 6월 20일). Can AI date the Dead Sea Scrolls? Bible History Daily, Biblical Archaeology Society. <https://www.biblicalarchaeology.org/daily/biblical-artifacts/dead-sea-scrolls/can-ai-date-the-dead-sea-scrolls/> (3 सितंबर 2026 को अभिगमिति)
- [5] University of Groningen. (2026, 6월 23일). Mladen Popović awarded ERC Advanced Grant to trace the origins of the Dead Sea Scrolls. rug.nl. <https://www.rug.nl/about-ug/latest-news/news/archief2026/nieuwsberichten/mladen-popovi-awarded-erc-advanced-grant-to-trace-the-origins-of-the-dead-sea-scrolls?lang=en> (3 सितंबर 2026 को अभिगमिति)
- [6] The Jerusalem Post. (2026, 6월 30일). New AI-powered research project aims to uncover the origins of the Dead Sea Scrolls. jpost.com. <https://www.jpost.com/archaeology/article-900938> (3 सितंबर 2026 को अभिगमिति)
- [7] Archaeology News Online Magazine. (2026, 6월). Where the Dead Sea Scrolls come from: AI and chemical analysis trace origins. archaeologymag.com. <https://archaeologymag.com/2026/06/dead-sea-scrolls-origins-studied-with-ai-and-chemical-analysis/> (3 सितंबर 2026 को अभिगमिति)

अध्याय 4. मट्टी और चर्मपत्र के हजारों टुकड़ों की पहली जोड़ने वाला डिजिटल गोद

4.1. केस फाइल

बकरी ढूँढ़ते समय मलिा घड़ा

1947 के वसंत में मृत सागर की पश्चिमी चट्टानों पर एक घटना घटी। बदलू लड़के अपनी भागी हुई बकरी ढूँढ़ रहे थे। लड़कों ने चट्टानों के बीच एक संकरा गुफा-द्वार देखा। गुफा के भीतर मट्टी के लंबे घड़े रखे थे। घड़ों के अंदर चमड़े पर लिखे ग्रंथ-चर्मपत्र रखे थे[1]।

ये ग्रंथ-चर्मपत्र ही मृत सागर पांडुलिपियाँ हैं। मृत सागर पांडुलिपियाँ ईसा पूर्व तीसरी शताब्दी के आसपास से ईसा की पहली शताब्दी के बीच लिखे गए यहूदी दस्तावेज़ हैं। उसके बाद के नौ वर्षों में पास की दस और गुफाएँ मिलीं। ग्रंथ-चर्मपत्र मलिन वाली गुफाएँ कुल ग्यारह हैं। यहाँ से मलि दस्तावेज़ 900 से अधिक हैं। ये हबिरू, अरामी और ग्रीक भाषा में लिखे गए थे[1]। अधिकांश भेड़ या बकरी की खाल से बने पतले चर्मपत्र पर लिखे गए थे। पपीरस नामक पौधे के तने से बने कागज़ पर लिखी प्रतियाँ भी हैं[7]।

इन ग्रंथ-चर्मपत्रों में हबिरू बाइबल की सबसे पुरानी प्रतियाँ शामिल थीं। समुदाय के नयिमों और कैलेंडर से संबंधित दस्तावेज़ भी मलि[1]। सूखे रेगिस्तान की गुफाओं ने इन शब्दों को 2,000 वर्षों तक सुरक्षित रखा।

गुफाओं के पास कुमरान नामक एक प्राचीन बस्ती के अवशेष हैं। अधिकांश वदिवानों का मानना है कि यहाँ रहने वाले यहूदी समुदाय ने ये ग्रंथ-चर्मपत्र छोड़े थे। वे धर्मग्रंथ की नकल करते थे और अपने नयिम लिखते थे। माना जाता है कि रोमन सेना के आगे बढ़ने पर उन्होंने इन ग्रंथ-चर्मपत्रों को गुफाओं में छिपा दिया था।

गुफा 4 में टुकड़ों का ढेर

अलग-अलग गुफाओं में पुरावशेषों की स्थिति बहुत भिन्न थी। गुफा 1 के ग्रंथ-चर्मपत्र घड़ों के अंदर काफी हद तक सुरक्षित रहे। गुफा 4 की स्थिति अलग थी। इस गुफा के ग्रंथ-चर्मपत्र घड़ों में रखे बनि ज़मीन पर ढेर लगे पड़े थे। केवल इसी एक गुफा से लगभग 500 ग्रंथ-चर्मपत्रों के हजारों टुकड़े निकले[1]।

गुफा की नमी ने चमड़े को धीरे-धीरे खराब कर दिया। कीड़ों और जानवरों ने चमड़े को कुतर दिया। सदयिँ बीतने के साथ ग्रंथ-चर्मपत्र छोटे-छोटे टुकड़ों में बखिर गए। पीछे केवल नाखून जतिने छोटे टुकड़ों का ढेर बचा था।

इज़राइल पुरावशेष प्राधिकरण द्वारा संरक्षित टुकड़ों की संख्या 25,000 से अधिक है[2]। अधिकांश टुकड़ों पर यह नहीं लिखा है कि वे किस ग्रंथ-चर्मपत्र से अलग हुए हैं। किसी एक टुकड़े पर केवल दो या तीन अक्षर ही बचे होना आम बात है। वदिवान इन्हें बड़ी काँच की प्लेटों पर फैलाते थे और आँखों से मलिन करते थे।

मलिन के लिए कुछ ही सुराग उपलब्ध थे। वे देखते थे कि अक्षरों की बनावट मलिती है या नहीं। वे देखते थे कि चमड़े का रंग एक जैसा है या नहीं। वे टूटे हुए कनारों को सटाकर देखते थे कि वे आपस में जुड़ते हैं या नहीं। इस वधिसे एक बार में केवल दो या तीन टुकड़ों की ही तुलना की जा सकती है। टुकड़ों की संख्या हजारों में पहुँचने पर तुलना किए जाने वाले जोड़ों की संख्या बेकाबू होकर बढ़ जाती है।

हाथ से पहली जोड़ने में बहुत समय लगा। मानव आँख एक साथ हजारों टुकड़ों के कनारों की तुलना नहीं कर सकती। इसके अलावा, टुकड़ों के कनारे पहले ही घसि चुके थे और सकिड़ गए थे। वदिवानों द्वारा दशकों तक मेहनत करने के बाद भी कई टुकड़े अपनी सही जगह नहीं पा सके।

दस्तावेज़ों को पहले पढ़ने का अधिकार केवल कुछ ही वदिवानों को दिया जाना भी एक समस्या थी। बाहरी शोधकर्ताओं के लिए तस्वीरें देखना भी मुश्किल था[1]। सतिंबर 1991 में अमेरिका के हंटिंगटन पुस्तकालय ने इस पाबंदी को हटा दिया। पुस्तकालय ने अपने पास मौजूद फोटो फ़िल्मों को बनि किसी शर्त के सार्वजनिक कर दिया[3]। तस्वीरें उपलब्ध होने के बाद कंप्यूटर से टुकड़ों को जोड़ने के प्रयास शुरू हुए।

पुरावशेषों का कई देशों में बखिर जाना भी एक बाधा थी। खोज के शुरुआती दौर में कुछ ग्रंथ-चर्मपत्र बेथलहम के पुरावशेष व्यापारियों के ज़रिए बकि गए थे[1]। उत्खनन का खर्च उठाने वाले देशों के संस्थानों को भी टुकड़े बाँट दिए गए थे। इस वजह से एक ही ग्रंथ-चर्मपत्र के टुकड़े अलग-अलग शहरों में बाँट गए। दो टुकड़े आपस में जुड़ते हैं या नहीं, यह देखने के लिए व्यक्तियों को हवाई यात्रा करनी पड़ती थी।

बाज़ार में आए नकली टुकड़े

मृत सागर पांडुलिपियों के टुकड़ों की कीमत बढ़ने के साथ एक और समस्या पैदा हुई। 2002 के आसपास से पुरावशेष बाज़ार में मृत सागर पांडुलिपियों के टुकड़े बताकर वस्तुएँ बेची जाने लगीं। बाइबल के श्लोक लिखे लगभग 70 टुकड़े सामने आए[4]। वदिवानों ने इन वस्तुओं को "पोस्ट-2002 टुकड़े" कहा। इन वस्तुओं के साथ ऐसा कोई रकिर्ड नहीं था जो बताता कि वे किस गुफा से और कब मलि थे[6]।

वाशिंगटन, अमेरिका के बाइबल संग्रहालय ने भी ऐसे 16 टुकड़े खरीदकर प्रदर्शित किए। कई विद्वानों को संदेह हुआ कि अक्षरों की बनावट अस्वाभाविक है। संग्रहालय ने कलाकृत जालसाजी जाँच कंपनी को गहन परीक्षण का काम सौंपा। इस जाँच का नेतृत्व अन्वेषक कोलेट लोल ने किया[4]।

मार्च 2020 में परणाम सामने आया। सभी 16 टुकड़े आधुनिक समय में बनाई गई जालसाजी निकले[4][5]। जाँच दल ने 3D माइक्रोस्कोप से सतह की जाँच की। सामग्री असली मृत सागर पांडुलिपियों का पतला चर्मपत्र नहीं थी। कम से कम 15 टुकड़े मोटे और खुरदरे चमड़े के थे। स्याही पहले से मौजूद दरारों के भीतर जमा थी[4]। यह इस बात का प्रमाण था कि चिमड़ा फटने के बाद उस पर लिखा गया था।

इस घटना ने पूरे बाइबल पुरातत्व जगत को हिलाकर रख दिया। जॉर्ज वाशिंगटन विश्वविद्यालय के प्रोफेसर क्रिस्टोफर रोलस्टन का कहना है कि आज भी बाज़ार में नकली टुकड़े आ रहे हैं। वे अक्सर ऐसे लोगों से मिलते हैं जो अपनी वस्तुओं को असली मानते हैं, लेकिन उन्होंने स्पष्ट किया कि लिगभग सभी नकली हैं। उत्खनन रिकॉर्ड वाले पुरावशेषों और बाज़ार से खरीदी गई वस्तुओं के बीच का अंतर यहाँ स्पष्ट हो गया[6]।

टूटे हुए टुकड़ों को जोड़ना और नकली वस्तुओं को अलग करना, दोनों की जड़ें एक ही हैं। दोनों में उन सूक्ष्म भौतिक नशानों को पढ़ना पड़ता है जिन्हें मानवीय आँख नहीं पकड़ सकती। एआई इस काम में शामिल हो गया है। जून 2026 में यूरोपीय अनुसंधान परिषद ने नया शोध अनुदान आवंटित किया। मृत सागर पांडुलिपियों की उत्पत्ति का पता लगाने वाले 5 साल के शोध के लिए 25 लाख यूरो दिए गए। शोध परियोजना का नाम "लेखक और ग्रंथ-चर्मपत्र की खोज" है। यह शोध सामग्री के घटक विश्लेषण, एआई और प्राचीन हस्तलेख अध्ययन का एक साथ उपयोग करता है[7]।

इस अध्याय में हम दो मामलों पर नज़र डालेंगे। पहला है बखिरे हुए टुकड़ों को फिर से जोड़ना। दूसरा है नकली वस्तुओं को अलग करना। ये दोनों काम एक ही सामग्री को लेकर अलग-अलग सवाल पूछते हैं। पहला सवाल है कि क्या ये दोनों मूल रूप से एक ही थे। दूसरा सवाल है कि क्या यह वास्तव में प्राचीन है।

संदर्भ सामग्री

- [1] Israel Antiquities Authority. (n.d.). Discovery and publication. The Leon Levy Dead Sea Scrolls Digital Library. <https://www.deadseascrolls.org.il/learn-about-the-scrolls/discovery-and-publication> (3 सितंबर 2026 को अभिगमिति)
- [2] Israel Antiquities Authority. (n.d.). The digital library. The Leon Levy Dead Sea Scrolls Digital Library. <https://www.deadseascrolls.org.il/about-the-project/the-digital-library> (3 सितंबर 2026 को अभिगमिति)
- [3] Deseret News. (1991, 9월 22일). Research library will give scholars access to Dead Sea Scroll photos. Deseret News. <https://www.deseret.com/1991/9/22/18942605/research-library-will-give-scholars-access-to-dead-sea-scroll-photos/> (3 सितंबर 2026 को अभिगमिति)
- [4] Greshko, M. (2020, 3월 13일). "Dead Sea Scrolls" at the Museum of the Bible are all forgeries. National Geographic. <https://www.nationalgeographic.com/history/article/museum-of-the-bible-dead-sea-scrolls-forgeries> (3 सितंबर 2026 को अभिगमिति)
- [5] Solly, M. (2020, 3월 16일). All of the Museum of the Bible's Dead Sea Scrolls are fake, report finds. Smithsonian Magazine. <https://www.smithsonianmag.com/smart-news/all-museum-bibles-dead-sea-scrolls-are-fake-report-finds-180974425/> (3 सितंबर 2026 को अभिगमिति)
- [6] Govier, G. (2026, 4월 8일). New Dead Sea Scrolls exhibit is the real deal. Christianity Today. <https://www.christianitytoday.com/2026/04/new-dead-sea-scrolls-exhibit-is-the-real-deal/> (3 सितंबर 2026 को अभिगमिति)
- [7] The Jerusalem Post. (2026, 6월 30일). New AI-powered research project aims to uncover the origins of the Dead Sea Scrolls. The Jerusalem Post. <https://www.jpost.com/archaeology/article-900938> (3 सितंबर 2026 को अभिगमिति)

4.2. फोरेसिक टीम की पड़ताल 1: ज्यामितीय एज एफनिटी और भौतिक फगिरप्रटि

फटे हुए कनारों को संख्याओं में बदलना

टूटे हुए टुकड़ों को जोड़ने का पहला सुराग कनारों का आकार होता है। कागज़ को हाथ से फाड़ने पर आरी के दाँतों जैसी टेढ़ी-मेढ़ी रेखा बनती है। पास-पास के दो टुकड़ों की रेखाएँ आपस में जुड़ती हैं। इंसान अंदाज़े से इन रेखाओं की तुलना करता है। कंप्यूटर इस रेखा को संख्याओं में बदल देता है।

कंप्यूटर कनारों के साथ-साथ मुड़ाव को मापता है। मुड़ाव को क्रम से रखने पर तरंग के आकार का एक संख्या अनुक्रम बनता है। जब दो टुकड़ों के संख्या अनुक्रम एक-दूसरे के विपरीत रखकर सटीक बैठते हैं, तो वे जुड़ने के उम्मीदवार बन जाते हैं। इस विधिको एज एफनिटी गणना कहा जाता है। अपनी भाषा में इसे कनारा जुड़ाव स्कोर कहते हैं।

इस स्कोर को अपनी भाषा में कनिरा जुड़ाव स्कोर कहा जा सकता है। यह वह संख्यात्मक मान है जो दर्शाता है कति टुकड़ों के कनिरा आपस में कतिने सटीक बैठते हैं।

केवल कनिरों को देखने की वधिकी स्पष्ट सीमाएँ हैं। प्राचीन टुकड़ों के कनिरा पहले ही घसि चुके होते हैं। सूखकर सकिड़ने के कारण कई जगह मूल आकार खो चुका होता है। इसलिए शोधकर्ता कनिरों के अलावा अन्य सुरागों को भी साथ देखते हैं।

सामग्री द्वारा छोड़ा गया फगिरप्रति

केवल कनिरा ही पर्याप्त नहीं है। टुकड़ों के कनिरा अक्सर पहले से ही घसि होते हैं। इसलिए शोधकर्ता स्वयं सामग्री पर ध्यान देते हैं।

पपीरस पौधे के तने को पतले रेशों में काटकर और उन्हें आड़ा-तरिखा रखकर बनाया जाता है। इसलिए सतह पर तने के रेशों के पैटर्न की बनावट रह जाती है। मूल रूप से एक ही पन्ने के रहे दो टुकड़ों में यह बनावट आपस में जुड़ती है[1]।

इजराइल के हाइफा विश्वविद्यालय में रॉय एबटिबुल की शोध टीम ने इन रेशों की बनावट को पढ़ा। शोध टीम ने टुकड़े के कनिरों के पास के हसिसे को छोटे-छोटे चौकोर टुकड़ों में काटा। तंत्रिका नेटवर्क ने दो चौकोर हसिसों को देखकर यह स्कोर दिया कि रेशे आपस में जुड़ते हैं या नहीं। उन स्कोरों को मिलाकर पूरे टुकड़े के जोड़े का निर्धारण किया गया[1]।

परीक्षण के लिए खराब स्थिति वाले मृत सागर पांडुलिपियों के टुकड़ों के समूह को चुना गया था। चूँकि सही उत्तर अज्ञात था, इसलिए कई गलत उम्मीदवार भी सामने आए। फिर भी, जो जोड़े मेल नहीं खा सकते थे, उनमें से 98 प्रतिशत को हटा दिया गया[1]। इसान द्वारा जाँचे जाने वाले काम की मात्रा घटकर 50वाँ हसिसा रह गई।

इस परिणाम का अर्थ स्पष्ट है। कंप्यूटर ने जोड़ों की पुष्टि नहीं की। कंप्यूटर ने उन जोड़ों को हटा दिया जिन्हें देखने की ज़रूरत नहीं थी। टुकड़ों को जोड़ने में एआई की भूमिका आम तौर पर ऐसी ही छँटाई की होती है।

फ्रांस के बोर्डो में एंटोनी परीन की शोध टीम ने इसके लिए सयामी तंत्रिका नेटवर्क का उपयोग किया। सयामी तंत्रिका नेटवर्क दो छवियों को साथ रखकर यह तय करता है कि वे एक ही स्रोत से हैं या नहीं। शोध टीम ने इस मॉडल का नाम PapySNet रखा। मॉडल ने पपीरस के लगभग 500 टुकड़ों से प्रशिक्षण लिया। उसने अक्षरों को नहीं, बल्कि केवल कागज़ के रेशों को देखकर जोड़ों का चयन किया। वास्तविक डेटा में उसने 79 प्रतिशत जोड़ों का सही मिलान किया[2]।

चर्मपत्र पर भी इसी तरह के फगिरप्रति होते हैं। चमड़े पर बालों के रोमछिद्रों के नशान रह जाते हैं। एक ही जानवर की खाल से निकले टुकड़ों में इन नशानों का क्रम जुड़ा रहता है। कीड़ों द्वारा ग्रंथ-चर्मपत्र में किए गए छेद भी एक सुराग हैं। लपिटे हुए ग्रंथ-चर्मपत्र को खोलने पर कीड़ों के छेद एक कतार में बार-बार दिखाई देते हैं। ग्रंथ-चर्मपत्र की अंदरूनी परतें छोटी और बाहरी परतें लंबी होती हैं। इसलिए बाहर की ओर जाने पर छेदों के बीच की दूरी थोड़ी-थोड़ी बढ़ती जाती है। इस दूरी को मापकर यह अनुमान लगाया जा सकता है कि टुकड़ा किस परत का था।

ऐसे भौतिक फगिरप्रति से निपटने के लिए सबसे पहले तस्वीरों को व्यवस्थित करना ज़रूरी है। इजराइल पुरावशेष प्राधिकरण की मृत सागर पांडुलिपियों की तस्वीरों में पैमाना, रंग चार्ट और संख्या पट्टिकाएँ साथ दिखाई देती हैं। पृष्ठभूमिकाली होने के कारण वह स्याही के साथ भ्रमति करती है। ब्रॉनसन ब्राउन-डेवोस्ट की शोध टीम ने इन बाधाओं को पहले हटाने के लिए चार चरणों वाली प्रसंस्करण वधिकिविकिसति की। शोध टीम ने सही उत्तरों से युक्त डेटासेट भी सार्वजनिक किया[3]।

सही उत्तरों वाले डेटासेट को सार्वजनिक करने का बहुत बड़ा महत्व है। अन्य शोधकर्ता उसी डेटा से अपनी वधियों का परीक्षण कर सकते हैं। दोनों वधियों में से कौन-सी बेहतर है, इसकी तुलना भी संख्याओं में की जा सकती है। प्राचीन दस्तावेज़ अनुसंधान इस तरह परस्पर सत्यापन वाले क्षेत्र में बदल रहा है।

मिट्टी की पट्टिकाओं, हड्डियों और कागज़ तक फैली तकनीक

यही विचार अन्य पुरावशेषों पर भी लागू हुआ। जर्मनी के म्यूनिख स्थित लुडविग मैक्समिलियन विश्वविद्यालय में प्रोफेसर एनरिक जिमिनेज है। वे इलेक्ट्रॉनिक बेबीलोनियन लाइब्रेरी (eBL) परियोजना का नेतृत्व करते हैं। इस परियोजना का टुकड़ा मिलान उपकरण Fragmentarium है[4]।

मेसोपोटामिया के लोग गीली मिट्टी पर नरकट की कलम दबाकर कीलाकार लिपिलिखिते थे। ये मिट्टी की पट्टिकाएँ आसानी से टूट जाती हैं। शोध टीम ने ब्रिटिश संग्रहालय और इराक संग्रहालय से मिट्टी की पट्टिकाओं की तस्वीरें लीं। 2018 से लगभग 22,000 टुकड़ों को डेटा के रूप में संसाधित किया गया[5]। कंप्यूटर टुकड़ों पर बचे अक्षरों से तैयार टेक्स्ट स्ट्रिंग्स की आपस में तुलना करता है। इस दौरान जीवविज्ञान में जीन अनुक्रमों की तुलना के लिए उपयोग की जाने वाली वधि उधार ली गई। ऐसा इसलिए क्योंकि प्राचीन लेखक क्षेत्र के अनुसार वर्तनी अलग-अलग लिखिते थे[4]।

उपलब्धियों संख्याओं में स्पष्ट हुई। इस परियोजना ने केवल 5 वर्षों में लगभग 1,500 नए जुड़ाव खोज निकाले। पछिल्ले 150 वर्षों में वदिवानों द्वारा खोजे गए जुड़ावों की संख्या लगभग 5,000 थी[4]।

मट्टी की पट्टिकाओं में एक और सुराग होता है जो चर्मपत्र में नहीं मिलता। वह है टूटी हुई सतह का 3D आकार। मट्टी की पट्टियों को आम तौर पर धूप में सुखाया जाता था और केवल कुछ को आग में पकाया जाता था। कठोर होने के दौरान उनका भीतरी हिस्सा असमान रूप से सख्त हो जाता है। इसलिए हर टूटी हुई सतह पर अलग-अलग घुमाव बनते हैं। इस सतह को 3D में मापकर आपस में जुड़ने वाले जोड़े खोजे जा सकते हैं।

चीन में ओरेकल बोन के टुकड़ों पर इसी तकनीक का उपयोग किया जाता है। ओरेकल बोन शांग राजवंश के लोगों द्वारा भविष्यवाणी करते समय कछुए के नचिले कवच और मवेशियों की हड्डियों पर उकेरे गए रिकॉर्ड हैं। हेनान विश्वविद्यालय में झांग चोंगशेंग की शोध टीम ने ओरेकल बोन जुड़ाव डेटासेट तैयार किया। शोध टीम के जुड़ाव एल्गोरिदम ने शीर्ष 10 में 98.39 प्रतिशत सटीकता हासिल की[6]। झांग ज्ञान की शोध टीम ने भी ओरेकल बोन के टुकड़ों की तस्वीरों से डीप रीजॉइनिंग मॉडल को प्रशिक्षित किया। इस शोध टीम ने भी डेटासेट को सार्वजनिक किया[7]।

शीर्ष 10 सटीकता का अर्थ वसतिार से समझने की आवश्यकता है। मशीन केवल एक सही उत्तर नहीं देती। वह उच्चतम संभावना वाले दस उम्मीदवारों को क्रम से प्रस्तुत करती है। उन दस में सही उत्तर होने की संभावना 98.39 प्रतिशत है[6]। अंतिम निर्णय इंसान करता है।

डेटा को एक जगह इकट्ठा करने का काम भी साथ-साथ हुआ। अन्यांग नॉर्मल यूनिवर्सिटी 'यनिची वेनयुआन' नामक ओरेकल बोन डेटा प्लेटफॉर्म संचालित करती है। अप्रैल 2023 तक इस पर ओरेकल बोन की 2,30,000 से अधिक तस्वीरें अपलोड की जा चुकी हैं। इसके साथ ही ओरेकल बोन लपिसे संबंधित 152 पुस्तकें और लगभग 3,000 शोध पत्र भी संकलित किए गए हैं[8]।

दुनहुआंग की मोगाओ गुफाओं से मिले बौद्ध धर्मग्रंथों के टुकड़े भी विभिन्न देशों में बखिरे हुए हैं। चिन मगिकुन की शोध टीम ने 2026 में एक नई वर्धिपेश की। यह वर्धिइंसान द्वारा सही उत्तर दिए बनिा भी खुद सीखती है। सीखने की इस पद्धति को स्व-पर्यवेक्षण शक्ति कहा जाता है। शोध टीम ने पूरी प्रतियों और ज्ञात जोड़ों वाले टुकड़ों को एक साथ एकत्र किया। जिन टुकड़ों का कोई जोड़ा ज्ञात नहीं था, उन्हें भी डेटा में शामिल किया गया। कंप्यूटर ने इस डेटा से खुद ही अभ्यास प्रश्न तैयार किए। शोध टीम ने कनिारों की दशा को शामिल करने वाली एक नई एन्कोडिंग वर्धि तैयार की। इसके बाद सियामी तंत्रिका नेटवर्क से दो टुकड़ों की समानता की डिग्री को स्कोर दिया गया[9]।

मृत सागर पांडुलिपियों के मामले में, यूरोपीय और इजराइली शोधकर्ताओं ने मिलकर एक ऑनलाइन कार्यस्थल बनाया। इसका नाम Scripta Qumranica Electronica है। इस कार्यस्थल पर शोधकर्ता स्क्रीन पर टुकड़ों की तस्वीरों को हलिकर और घुमाकर व्यवस्थित करते हैं। व्यवस्थित करने का यह परिणाम अन्य शोधकर्ताओं के साथ साझा किया जा सकता है[10]। टुकड़ों को जोड़ना एक व्यक्तिके दमिग के काम से बदलकर कई लोगों द्वारा मिलकर समीक्षा किए जाने वाले काम में बदल गया।

संदर्भ सामग्री

- [1] Abitbol, R., Shimshoni, I., & Ben-Dov, J. (2021). Machine learning based assembly of fragments of ancient papyrus. *Journal on Computing and Cultural Heritage*, 14(3), 1-21. <https://doi.org/10.1145/3460961>
- [2] Pirrone, A., Beurton-Aimar, M., & Journet, N. (2019). Papy-S-Net: A Siamese network to match papyrus fragments. *Proceedings of the 5th International Workshop on Historical Document Imaging and Processing*, 78-83. <https://doi.org/10.1145/3352631.3352646>
- [3] Brown-deVost, B., Kurar-Barakat, B., & Dershowitz, N. (2024). Segmenting Dead Sea Scroll fragments for a scientific image set. *arXiv*. <https://arxiv.org/abs/2406.15692>
- [4] Jiménez, E. (2023, 5 मई 25 मई). AI is helping us read ancient Mesopotamian literature. *The Conversation*. <https://theconversation.com/ai-is-helping-us-read-ancient-mesopotamian-literature-205091> (3 सितंबर 2026 को अभिगमति)
- [5] Phys.org. (2023, 2 मई 2 मई). Researcher uses AI to make texts that are thousands of years old readable. *Phys.org*. <https://phys.org/news/2023-02-ai-texts-thousands-years-readable.html> (3 सितंबर 2026 को अभिगमति)
- [6] Zhang, C., Zong, R., Cao, S., Men, Y., & Mo, B. (2020). AI-powered oracle bone inscriptions recognition and fragments rejoining. *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence*, 5309-5311. <https://doi.org/10.24963/ijcai.2020/779>
- [7] Zhang, Z., Zhang, H., Guo, A., Jiao, Q., Liu, Y., Li, B., & Gao, F. (2025). Deep rejoining model and dataset of oracle bone fragment images. *npj Heritage Science*, 13(1), 66. <https://doi.org/10.1038/s40494-025-01651-9>
- [8] 王者. (2023, 4 मई 8 मई). "殷契文渊"数据平台为甲骨文研究提供助力. *人民日报*. https://news.hangzhou.com.cn/gnxw/content/2023-04/08/content_8509027.htm (3 सितंबर 2026 को अभिगमति)
- [9] Chen, M., Xiang, Y., & Zheng, Y. (2026). A self-supervised learning approach for pairwise matching of ancient Dunhuang fragments. *Journal on Computing and Cultural Heritage*, 19(1), 1-19. <https://doi.org/10.1145/3776642>

4.3. फोरेंसिक टीम की पड़ताल 2: जालसाजी जासूस एआई

16 नकली टुकड़ों को पकड़ने का तरीका

बाइबल संग्रहालय द्वारा खरीदे गए मृत सागर पांडुलिपियों के सभी 16 टुकड़े नकली थे[1][2]। जांच दल ने केवल आंखों से अक्षरों को ही नहीं देखा। उन्होंने भौतिक और रासायनिक विधियों से सामग्री तथा स्याही की बारीकी से जांच की। इस प्रमाणीकरण में एआई का उपयोग नहीं किया गया था। इंसानों ने माइक्रोस्कोप और घटक विश्लेषण के जरिए इसे पकड़ा था[1]।

बाइबल संग्रहालय का प्रमाणीकरण एक बार में समाप्त नहीं हुआ। अक्टूबर 2018 में जर्मनी के संघीय सामग्री अनुसंधान एवं परीक्षण संस्थान ने पहले पांच टुकड़ों की जांच की। ये पांच टुकड़े प्राचीन पुरावशेषों की विशेषताओं से मेल नहीं खाते थे, इसलिए उन्हें प्रदर्शनी से हटा दिया गया[5]। इसके बाद शेष सभी नमूनों की जांच का परिणाम 2020 में सामने आया[1][2]।

पहला प्रमाण आधार सामग्री थी। असली मृत सागर पांडुलिपियाँ पतले फैलाए गए चर्मपत्र पर लिखी गई थीं। बाइबल संग्रहालय के टुकड़ों में से कम से कम 15 चमड़े के थे। यह चमड़ा चर्मपत्र की तुलना में अधिक मोटा और खुरदरा है[1]।

दूसरा प्रमाण स्याही का स्थान था। जांच दल ने 3D माइक्रोस्कोप से सतह की जांच की। दरारों के भीतर स्याही जमा थी। फटे हुए कनारों पर स्याही के नीचे की ओर बहने के निशान दिखाई दिए[1]।

यह निशान नरिणायक क्यों है, यह इसके क्रम पर विचार करने से समझा जा सकता है। असली पांडुलिपि में पहले समतल चर्मपत्र पर लिखा गया था। उसके बाद लंबे समय में चमड़ा सूखने के साथ फट गया। इसलिए असली पांडुलिपि में दरारों पर अक्षरों के स्ट्रोक टूट जाते हैं। नकली में यह क्रम उल्टा है। जालसाज ने पहले से फटे हुए पुराने चमड़े को हासिल किया और उस पर स्याही लगाई। स्याही खुली दरारों में बहकर अंदर चली गई[1]।

तीसरा प्रमाण सतह पर लगाया गया पदार्थ था। जांच दल को टुकड़ों के एम्बर रंग के तरल में डुबोए जाने के निशान मिले। यह तरल जानवरों की खाल से बना गोद प्रतीत होता है। जालसाज ने असली मृत सागर पांडुलिपियों की सतह की चमक की नकल करने की कोशिश की थी[1]।

तीनों प्रमाण अलग-अलग दशाओं से एक ही उत्तर की ओर संकेत करते थे। केवल एक प्रमाण संयोग हो सकता है। जब तीनों एक साथ मिलते हैं, तो यह संयोग नहीं होता। पुरावशेषों के प्रमाणीकरण में विभिन्न प्रकार के प्रमाण एक साथ जुटाने का यही कारण है।

हाथ द्वारा छोड़े गए निशान

नकली बनाने वाला व्यक्ति असली तस्वीर देखकर अक्षरों की नकल करता है। यह नकल अपने पीछे निशान छोड़ जाती है।

जब कोई व्यक्ति परिचित लिखावट में लिखता है, तो हाथ बना रुकें बहता है। जब वह किसी अपरिचित अक्षर के आकार की नकल करता है, तो स्थिति अलग होती है। वह स्ट्रोक खींचते समय रुकता है और अगले स्ट्रोक की स्थिति की पुष्टि करता है। ब्रश की नोक की गति धीमी हो जाती है और रेखा सूक्ष्म रूप से कांपती है। बड़ा करके देखने पर यह अंतर स्ट्रोक की बाहरी रेखा के रूप में दिखाई देता है।

कंप्यूटर इस अंतर को संख्यात्मक रूप से पकड़ लेता है। जून 2026 में ब्रिटन के ब्रैडफोर्ड विश्वविद्यालय में हसन उगाइल के शोध दल ने संबंधित शोध प्रकाशित किया। शोध दल ने एक ऐसा एआई तैयार किया जो यह तय करता है कि पुरानी रेखाचित्र कृतियाँ असली हैं या नहीं। यह मॉडल चित्रों से निकाले गए केवल पांच अंकों को देखता है। यह पैटर्न के दोहराव की मात्रा और चमक मान के बखिराव की सीमा को मापता है। यह हल्के और गहरे क्षेत्रों के अंतर को भी मापता है। इसमें बनावट की एकरूपता और रेखाओं के टहनियों की तरह शाखति होने की सीमा भी जोड़ी जाती है[3]।

शोध दल ने दस चित्रकारों के वास्तविक रेखाचित्रों से मॉडल को प्रशिक्षित किया। इन कृतियों का काल लगभग 1510 से 1917 तक फैला हुआ है। डेटा सार्वजनिक संग्रह उपलब्ध कराने वाले छह कला संग्रहालयों से एकत्र किया गया था। इनमें मेट्रोपॉलिटन कला संग्रहालय और एशमोलयिन संग्रहालय शामिल हैं[3]।

शोध पत्र में 900 नरिणयों के साथ परीक्षण के परिणाम प्रकाशित किए गए। संतुलित सटीकता 87.6 प्रतिशत थी। असली को असली के रूप में स्वीकार करने का अनुपात 77.8 प्रतिशत था। किसी अन्य की कृति को उस चित्रकार की कृति मान लेने की त्रुटि का अनुपात 2.6 प्रतिशत था[3]।

इस शोध में एक और उल्लेखनीय परणाम है। व्यापक रूप से उपयोग किए जाने वाले बड़े डीप लर्निंग मॉडल ने इस काम में अपेक्षाकृत खराब प्रदर्शन किया। बड़े मॉडल अक्सर असली कृतियों को नकली बताकर खारज कर देते थे। जब डेटा कम होता है, तो समस्या के अनुकूल तैयार किया गया छोटा मॉडल बेहतर साबित हुआ[3]।

यह परणाम प्राचीन पांडुलिपियों के शोध के लिए भी उपयोगी संदर्भ है। बची हुई प्राचीन पांडुलिपियों की संख्या बहुत कम है। बड़े मॉडलों को तृप्त करने लायक पर्याप्त डेटा नहीं है। इसका अर्थ यह है कि पहले से यह तय करने वाला डिज़ाइन बेहतर है कि क्या देखना है।

लखावट मापने वाली मशीन

यही तकनीक वास्तविक पुरावशेषों के अध्ययन में भी उपयोग की जाती है। नीदरलैंड के ग़रोनगिन विश्वविद्यालय में म्लादेन पोपोविक के शोध दल ने एआई के साथ मृत सागर पांडुलिपियों की लखावट का विश्लेषण किया। शोध दल ने न्यूरोल नेटवर्क से केवल स्याही के नशानों को निकाला। इसके बाद उन्होंने अक्षरों के घुमाव और झुकाव को संख्यात्मक मानों में बदल दिया[4]।

शोध दल ने कार्बन डेटिंग पूरी कर चुकी 24 पांडुलिपियों से मॉडल को प्रशिक्षित किया। मॉडल का नाम Enoch है। Enoch ने अज्ञात काल की 135 पांडुलिपियों की आयु का अनुमान लगाया। औसत त्रुटि लगभग 28 से 31 वर्ष थी। कई पांडुलिपियाँ अब तक ज्ञात समय से अधिक प्राचीन पाई गई[4]।

यह तकनीक प्रामाणिकता के निर्धारण में भी मदद करती है। वास्तविक लिपिकों की लखावट में प्रत्येक युग में नश्चित सांख्यिकीय विशेषताएँ होती हैं। नकली इस वितरण से बाहर हो जाता है। मानवीय आँख को लग सकता है कि दोनों लखावटें समान हैं, लेकिन मशीन इस अंतर को संख्याओं में प्रस्तुत करती है।

हालाँकि यह सहायता केवल एक अतिरिक्त संदर्भ सामग्री प्राप्त करने के स्तर तक है। Enoch के शोध में एक सावधानी बरतने वाली बात है। यह मॉडल काल का अनुमान लगाने के लिए प्रशिक्षित किया गया था। इसे असली और नकली के बीच भेद करने के लिए प्रशिक्षित नहीं किया गया था। केवल इसलिए कि काल का अनुमान अजीब आता है, उसे तुरंत नकली नहीं कहा जा सकता। मशीन का परणाम केवल उसी कार्य के दायरे में अर्थ रखता है जो उसने सीखा है।

मशीन जसिकी जगह नहीं ले सकती

मशीन के निर्णय की भी सीमाएँ हैं। बाइबल संग्रहालय के टुकड़ों की प्रामाणिकता का अंतिम निर्णय अकेले एआई ने नहीं किया था। माइक्रोस्कोप अवलोकन, रासायनिक परीक्षण और ग्रंथवैज्ञानिक ज्ञान का एक साथ उपयोग किया गया था[1][2]।

स्रोत रिकॉर्ड का महत्व यथावत बना रहता है। स्रोत रिकॉर्ड वह दस्तावेज है जिसमें दर्ज होता है कि पुरावशेष कहाँ से मिला था। प्रोफेसर रोलस्टन का कहना है कि जिन पुरावशेषों के उत्खनन रिकॉर्ड हैं और जो बाज़ार से खरीदे गए हैं, उन्हें अलग करके देखा जाना चाहिए। उत्खनन से मिले ग्रंथ-चर्मपत्रों में यह दर्ज रहता है कि वे किस गुफा से और कब मिले थे। बाज़ार में आए टुकड़ों में वह रिकॉर्ड नहीं होता[5]।

मशीन उन सूक्ष्म संकेतों को पढ़ती है जिन्हें इंसान छोड़ देते हैं। इंसान यह निर्णय करता है कि उन संकेतों का क्या अर्थ है। नकली की पहचान का काम इन दोनों के एक साथ होने पर ही पूरा होता है।

संदर्भ सामग्री

- [1] Greshko, M. (2020, 3월 13일). "Dead Sea Scrolls" at the Museum of the Bible are all forgeries. National Geographic. <https://www.nationalgeographic.com/history/article/museum-of-the-bible-dead-sea-scrolls-forgeries> (3 सितंबर 2026 को अभिगमिति)
- [2] Solly, M. (2020, 3월 16일). All of the Museum of the Bible's Dead Sea Scrolls are fake, report finds. Smithsonian Magazine. <https://www.smithsonianmag.com/smart-news/all-museum-bibles-dead-sea-scrolls-are-fake-report-finds-180974425/> (3 सितंबर 2026 को अभिगमिति)
- [3] Ugail, H., Ritch-Frel, J., Matuzava, I., & Stork, D. G. (2026). Verification of historical sketches via one-class learning on compact feature representations. PLOS One, 21(6), e0344796. <https://doi.org/10.1371/journal.pone.0344796>
- [4] Popović, M., Dhali, M. A., Schomaker, L., van der Plicht, J., Lund Rasmussen, K., La Nasa, J., Degano, I., Perla Colombini, M., & Tigchelaar, E. (2025). Dating ancient manuscripts using radiocarbon and AI-based writing style analysis. PLOS One, 20(6), e0323185. <https://doi.org/10.1371/journal.pone.0323185>
- [5] Museum of the Bible. (2018, 10월 22일). Museum of the Bible releases research findings on fragments in its Dead Sea Scrolls collection. Museum of the Bible. <https://www.museumofthebible.org/newsroom/museum-of-the-bible-releases-research-findings-on-fragments-in-its-dead-sea-scrolls-collection> (3 सितंबर 2026 को अभिगमिति)

4.4. सामने आया सच

पुष्ट संयोजन और संभावित संयोजन को अलग करने वाली रेखा

यदि कंप्यूटर यह भी कहता है कि दो टुकड़े आपस में जुड़ते हैं, तो बात वहीं खत्म नहीं होती। मशीन केवल संभावना बताती है। शोधकर्ता पुष्ट संयोजनों और संभावित संयोजनों के बीच कड़ाई से भेद करते हैं।

पुष्ट संयोजन को तभी मान्यता दी जाती है जब कई प्रकार के प्रमाण एक साथ मेल खाते हों। सबसे पहले, टूटी हुई सतहों का आकार बर्तन किसी अंतर के आपस में फिट होना चाहिए। इसके बाद, सामग्री की बनावट जुड़ी होनी चाहिए। पपीरस में तने की धारियाँ जुड़ी होनी चाहिए[1]। चर्मपत्र में रोमछद्मों के निशान जुड़े होने चाहिए। उसके बाद, सीमा के आर-पार जाने वाले अक्षर के स्ट्रोक स्वाभाविक रूप से जुड़े होने चाहिए। अंत में, जोड़े गए पाठ को उस भाषा के व्याकरण के अनुकूल होना चाहिए।

संभावित संयोजन वह स्थिति है जहाँ इन शर्तों में से केवल कुछ ही पूरी होती हैं। कीड़ों द्वारा किए गए छेदों के बीच के अंतर से टुकड़ों की स्थिति का अनुमान लगाने का परिणाम इसमें शामिल है। स्याही के घटक समान होने के कारण उन्हें एक ग्रंथ-चर्मपत्र समूह में रखने का परिणाम भी संभावित संयोजनों में आता है। ऐसे परिणाम शोध का प्रारंभिक बिंदु बनते हैं, लेकिन निष्कर्ष नहीं होते। क्रम यह है कि कंप्यूटर द्वारा छांटे गए संभावित संयोजनों की पुष्टि इंसान करता है[1]।

यदि इस भेद को न बनाए रखा जाए तो यह खतरनाक हो सकता है। केवल उच्च संभावना के आधार पर दो टुकड़ों को जोड़ने से ऐसा वाक्य बन जाता है जो पहले असततत्व में ही नहीं था। यदि उस वाक्य को उद्धृत किया जाता है, तो त्रुटि अगले शोध तक फैल जाती है। इसीलिए डेटा रिकॉर्ड में पुष्ट और संभावित संयोजनों को अलग-अलग दर्ज किया जाता है।

मशीन द्वारा दी गई संख्याएं भी इसी भेद को मानकर चलती हैं। केवल छँटाई अनुपात अधिक होने से संयोजन की पुष्टि नहीं हो जाती[1]। शीर्ष उम्मीदवारों में सही उत्तर होने की संभावना के संबंध में भी यही बात लागू होती है[2]। दोनों ही स्थितियों में अंतिम निर्णय इंसान ही करता है।

पुनर्जीवित पाठ

इस तरह वास्तविक प्राचीन ग्रंथ पुनर्जीवित हुए। Fragmentarium ने 2018 से मट्टी की पट्टिकाओं के लगभग 22,000 टुकड़ों पर काम किया। इस प्रोग्राम ने कई नए संयोजन और नई पांडुलिपियाँ खोजीं। शोध दल ने फ़रवरी 2023 में Fragmentarium को आम जनता के लिए खोल दिया। गलिंगमेश महाकाव्य का एक डिजिटल संस्करण भी इसके साथ जारी किया गया[3]।

पुनर्जीवित किए गए पाठों में गलिंगमेश महाकाव्य शामिल है। गलिंगमेश महाकाव्य मेसोपोटामिया के एक नायक की कहानी है। नवंबर 2022 में प्रोग्राम ने एक टुकड़े की पहचान की। यह टुकड़ा लगभग 130 ईसा पूर्व में प्रतलिपिकिए गए संस्करण से संबंधित था। यह गलिंगमेश की अब तक पहचानी गई पांडुलिपियों में काफी बाद के काल की है। इसका अर्थ यह है कि उस काल तक भी लोग इस कहानी की प्रतलिपि बना रहे थे[3]।

कीलाकार लिपि की वर्तनी हर क्षेत्र में अलग थी। ऐसा इसलिए था क्योंकि लिपिकी अपनी स्थानीय बोलियों के अनुसार अक्षरों को बदलकर लिखते थे। जब इंसान आंखों से खोजते थे, तो यह अंतर एक बड़ी बाधा था। शोध दल ने जीन अनुक्रमण की तुलना करने की वधिकी अपनाकर इस बाधा को पार किया। थोड़ी भिन्न पाठ्य-श्रृंखलाओं की भी परस्पर तुलना करना संभव हो गया[4]।

मृत सागर पांडुलिपियों के क्षेत्र में भी डेटा का आधार वसित हुआ। इज़राइल पुरावशेष प्राधिकरण ने 25,000 से अधिक टुकड़ों की उच्च-रिज़ॉल्यूशन तस्वीरें खींचकर उन्हें सार्वजनिक किया[5]। शोधकर्ता इन तस्वीरों पर टुकड़ों को खसिकाकर व्यवस्थित करते हैं[6]। पैमाने और रंग चार्ट जैसी बाधाओं को स्वचालित रूप से हटाने की प्रसंस्करण वधि भी जारी की गई[7]।

2026 में खुला अगला द्वार

जून 2026 में शुरू हुआ "लिपिकों और ग्रंथ-चर्मपत्रों की खोज" शोध अगला द्वार खोलता है। यह शोध टुकड़ों को जोड़ने से एक कदम आगे का प्रश्न उठाता है। यह पूछता है कि मृत सागर पांडुलिपियों को कहाँ और किसने लिखा था[8]।

शोध दल चर्मपत्र, पपीरस और स्याही के लगभग 250 नमूनों की जांच कर रहा है। नमूना परीक्षण के लिए उपयोग किए जाने वाले छोटे प्रतदिश को कहते हैं। इसमें एआई लिखावट विश्लेषण और प्राचीन पुरालेख अध्ययन को जोड़ा गया है। इज़राइल पुरावशेष प्राधिकरण, पीसा विश्वविद्यालय और नेपल्स विश्वविद्यालय इसमें एक साथ काम कर रहे हैं। दक्षिणी डेनमार्क विश्वविद्यालय और बेलजियम का ल्यूवेन विश्वविद्यालय भी इसमें भाग ले रहे हैं। इसकी तुलना बर्लिन और ट्यूरिन के संग्रहालयों में रखे मसिर के पपीरस से भी की जाती है[8]।

यदि नमूनों के घटक समान हों, तो उनके एक ही कार्यशाला से आने की संभावना बढ़ जाती है। इस जानकारी का उपयोग टुकड़ों को जोड़ने में भी किया जाता है। यदि एक ही स्रोत की सामग्री को पहले एक साथ समूहीकृत कर दिया जाए, तो तुलना किए जाने वाले जोड़ों की संख्या घट जाती है।

बचे हुए टुकड़े और शेष कार्य

उपलब्धियाँ हासिल होने के बावजूद, ऐसे कई टुकड़े हैं जिनमें अभी तक अपना सही स्थान नहीं मिला है। पुरावशेष प्राधकिरण के भंडारण बक्सों में ऐसे टुकड़े बचे हुए हैं जिनके बारे में यह ज्ञात नहीं है कि वे किस ग्रंथ-चर्मपत्र से आए हैं[8]। मट्टी की पट्टिकाओं में भी ऐसे कई संयोजन हैं जो अभी तक नहीं मिले हैं[4]।

उपकरणों का अंतर भी एक समस्या बना हुआ है। 3D स्कैनर से पुरावशेष को मापने के लिए पुरावशेष को उपकरण के पास ले जाना पड़ता है। कई संग्रहालय ऐसे हैं जहाँ से पुरावशेषों को बाहर ले जाना कठिन होता है। इसलिए कुछ पुरावशेषों का केवल फोटोग्राफिक डेटा उपलब्ध है और 3D डेटा नहीं है। ऐसे पुरावशेषों के लिए टूटी हुई सतहों की तुलना करने की वधिकी उपयोग नहीं किया जा सकता है।

प्रौद्योगिकी द्वारा छोड़ी गई कुछ चुनौतियाँ भी हैं। टुकड़ों की तस्वीरें अलग-अलग देशों में अलग-अलग उपकरणों से ली गई थीं। यदि प्रकाश और पृष्ठभूमि भिन्न हों, तो वही एल्गोरिदम अलग परिणाम देता है। इसलिए शोधकर्ताओं ने सही उत्तरो के साथ एक सार्वजनिक डेटासेट तैयार किया। ऐसा इसलिए किया गया ताकि समान पैमाने पर एक-दूसरे के परिणामों की तुलना की जा सके[7]। दुनहुआंग के टुकड़ों पर शोध ने भी इस दिशा में रुख किया है कि इंसान द्वारा सही उत्तर लेबल किए बनिा भी सीखा जा सके[9]।

यह तकनीक इंसानों को बाहर नहीं निकालती है। मशीन तुलना किए जाने वाले जोड़ों की संख्या कम करती है और इंसान अर्थ का निर्णय करता है। पुनर्जीवित वाक्य उस काल की भाषा-शैली से मेल खाता है या नहीं, यह इंसान के पढ़ने पर ही पता चलता है। चूँकि दोनों की भूमिकाएँ आपस में टकराती नहीं हैं, इसलिए वे मलिकर काम कर सकते हैं।

सार्वजनिक पढ़ें भी बदल गई है। पहले केवल कुछ ही विद्वान मूल तस्वीरें देख पाते थे। अब मट्टी की पट्टियों का डेटा और मृत सागर पांडुलिपियों की तस्वीरें इंटरनेट पर उपलब्ध हैं[4][5]। छात्र भी स्क्रीन पर टुकड़ों को खसिकाकर देख सकते हैं[6]।

डिजिटल गौद टुकड़ों को भौतिक रूप से नहीं चपिकाता है। पुरावशेष बक्सों के अंदर सुरक्षित रहता है। केवल कंप्यूटर स्क्रीन पर ही ग्रंथ-चर्मपत्र दोबारा जुड़ता है। इसलिए यदि निर्णय गलत हो जाए, तो इसे कभी भी पहले जैसा किया जा सकता है। 2,000 वर्षों से बखिरा हुआ पाठ इस तरह धीरे-धीरे वाक्यों के रूप में वापस लौट रहा है।

संदर्भ सामग्री

- [1] Abitbol, R., Shimshoni, I., & Ben-Dov, J. (2021). Machine learning based assembly of fragments of ancient papyrus. *Journal on Computing and Cultural Heritage*, 14(3), 1-21. <https://doi.org/10.1145/3460961>
- [2] Zhang, C., Zong, R., Cao, S., Men, Y., & Mo, B. (2020). AI-powered oracle bone inscriptions recognition and fragments rejoining. *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence*, 5309-5311. <https://doi.org/10.24963/ijcai.2020/779>
- [3] Phys.org. (2023, 2११ 2११). Researcher uses AI to make texts that are thousands of years old readable. Phys.org. <https://phys.org/news/2023-02-ai-texts-thousands-years-readable.html> (3 सतिबर 2026 को अभगिमति)
- [4] Jiménez, E. (2023, 5११ 25११). AI is helping us read ancient Mesopotamian literature. *The Conversation*. <https://theconversation.com/ai-is-helping-us-read-ancient-mesopotamian-literature-205091> (3 सतिबर 2026 को अभगिमति)
- [5] Israel Antiquities Authority. (n.d.). The digital library. The Leon Levy Dead Sea Scrolls Digital Library. <https://www.deadseascrolls.org.il/about-the-project/the-digital-library> (3 सतिबर 2026 को अभगिमति)
- [6] Scripta Qumranica Electronica. (n.d.). Scripta Qumranica Electronica. University of Haifa. <https://megillot.haifa.ac.il/index.php/en/research/scripta-qumranica-electronica> (3 सतिबर 2026 को अभगिमति)
- [7] Brown-deVost, B., Kurar-Barakat, B., & Dershowitz, N. (2024). Segmenting Dead Sea Scroll fragments for a scientific image set. *arXiv*. <https://arxiv.org/abs/2406.15692>
- [8] The Jerusalem Post. (2026, 6११ 30११). New AI-powered research project aims to uncover the origins of the Dead Sea Scrolls. *The Jerusalem Post*. <https://www.jpost.com/archaeology/article-900938> (3 सतिबर 2026 को अभगिमति)
- [9] Chen, M., Xiang, Y., & Zheng, Y. (2026). A self-supervised learning approach for pairwise matching of ancient Dunhuang fragments. *Journal on Computing and Cultural Heritage*, 19(1), 1-19. <https://doi.org/10.1145/3776642>

अध्याय 5. मट्टी की पट्टिकाओं के टूटे टुकड़ों को जोड़ना: गलिगमेश महाकाव्य का संगणकीय पुनर्निर्माण

5.1. केस फाइल

4,000 वर्षों को पार कर आई कहानी

गलिगमेश महाकाव्य आज बची हुई कथा-साहित्य की कृतियों में सबसे पुराने ग्रंथों में गिना जाता है। इससे भी पहले के सुमेरियाई गीत भी मिलते हैं। इसका मुख्य पात्र मेसोपोटामिया के उरुक शहर पर शासन करने वाला राजा गलिगमेश है[1]। वह अपने मतिर एन्कडू के साथ सुदूर देवदार के जंगल में जाता है। वे दोनों उस जंगल की रक्षा करने वाले राक्षस हुम्बाबा से लड़ते हैं[2]। बाद में एन्कडू की मृत्यु हो जाती है। तब गलिगमेश अमरता का मार्ग खोजने निकल पड़ता है[1]।

इस कहानी को कागज़ पर नहीं, बल्कि मट्टी की पट्टिकाओं पर लिखा गया था। प्राचीन लपिकि नरकट को तरिखा काटकर नुकीली तीली बनाते थे। इस तीली को सटाइलस कहा जाता है। सटाइलस से गीली मट्टी को दबाने पर कील के आकार का निशान बन जाता है। यह रेखा खींचकर नहीं, बल्कि दबाकर बनाए गए अक्षर हैं। मट्टी सूखने पर वे निशान वैसे ही जम जाते हैं। इसलिए इस लपिकि कीलाकार लपिकि कहा जाता है। लपिकि बनने के इच्छुक छात्र लिखावट के अभ्यास के लिए इस महाकाव्य की नकल करते थे। इसी कारण मध्य पूर्व के कई पुरातात्विक स्थलों से इसकी प्रतियों के टुकड़े मिलते हैं। मसिर के अमारना स्थल से भी इस महाकाव्य के टुकड़े मिले हैं। इसका अर्थ है कि यह कहानी मेसोपोटामिया से बहुत दूर तक फैल चुकी थी[1]।

इसके मुख्य रूप से दो संस्करण मिलते हैं। पहला संस्करण दूसरी सहस्राब्दी ईसा पूर्व की शुरुआत में लिखा गया प्राचीन बेबीलोनियन संस्करण है। दूसरा बाद में संकलित किया गया मानक बेबीलोनियन संस्करण है। मानक संस्करण बारह मट्टी की पट्टिकाओं से मिलकर बना है। इनमें से ग्यारह पट्टिकाएँ मुख्य कथा हैं। बारहवीं पट्टिका को बाद में जोड़ा गया परिशिष्ट माना जाता है। आज हम जो अनुवाद पढ़ते हैं, उसका आधार यही मानक संस्करण है। इसके रचित स्थानों को प्राचीन बेबीलोनियन संस्करण और अन्य टुकड़ों से भरा जाता है। मानक संस्करण को परिष्कृत करने वाले व्यक्ति सिनि-लेकी-उन्ननिनी (Sin-leqi-unninni) नामक लपिकि माने जाते हैं। उन्होंने बखिरे हुए पुराने गीतों को एक लंबी कृति में परिया। उन्होंने प्रत्येक पट्टिका की सामग्री तय की और पंक्तियों की संख्या का तालमेल बिठाया। फिर भी इस मानक संस्करण में भी कई खाली स्थान हैं। शुरू से अंत तक पूरी तरह सुरक्षित बची हुई एक भी प्रतिलिपि नहीं है[1]।

कहानी की वषिय-वस्तु आज पढ़ने पर भी अपरिचित नहीं लगती। गलिगमेश शक्तिशाली था, लेकिन अपनी प्रजा को सताने वाला राजा था। देवताओं ने उसका सामना करने के लिए एन्कडू को बनाकर भेजा। दोनों पहले लड़े, फिर मतिर बन गए। उन्होंने मिलकर हुम्बाबा को हराया और स्वर्ग के बैल को भी मार गिराया। इसके बदले में एन्कडू बीमार पड़कर मर जाता है। मतिर की मृत्यु देखकर गलिगमेश को अहसास होता है कि एक दिन उसे भी मरना है। ग्यारहवीं पट्टिका में महाजलप्रलय से बचे व्यक्ति के पास जाकर उसकी कहानी सुनने का दृश्य शामिल है[1]।

यह कहानी 2,000 से अधिक वर्षों तक वसिष्ठ रही। कीलाकार लपिकि पढ़ने वाले लोग लुप्त हो गए थे। 19वीं शताब्दी में वदिवानों द्वारा कीलाकार लपिकि को फिर से पढ़े जाने के बाद यह महाकाव्य दुनिया के सामने लौटा। नीनवे से लाए गए मट्टी की पट्टिकाओं के बक्सों को खंगालते समय एक शोधकर्ता ने जलप्रलय की कहानी पहचानी। यहीं से इसकी शुरुआत हुई। उसके बाद के 150 वर्षों में वदिवानों ने बखिरी हुई प्रतियों को एकत्र कर मूल पाठ को धीरे-धीरे पूरा किया है[1]।

नीनवे और 612 ईसा पूर्व की आग

मानक संस्करण की अधिकांश प्रतियाँ एक प्राचीन पुस्तकालय से मिलीं। यह 7वीं शताब्दी ईसा पूर्व में असीरिया के राजा अशूरबानपिल द्वारा नीनवे के राजमहल में स्थापित किया गया ग्रंथागार था। इस पुस्तकालय ने ज्योतिष-ग्रंथों, चिकित्सा-पुस्तकों, शब्दकोशों और साहित्यिक कृतियों को व्यापक रूप से एकत्र किया। राजा ने अपने साम्राज्य के कोने-कोने से मट्टी की पट्टिकाओं को इकट्ठा करवाया। उसने अपने अधिकारियों को पत्र भेजकर अनुपलब्ध ग्रंथों को ढूँढ़ लाने का आदेश दिया। अशूरबानपिल ऐसा राजा था जो कीलाकार लपिकि स्वयं पढ़ना जानता था[3]।

नीनवे 612 ईसा पूर्व के आसपास भीषण आग की चपेट में आ गया। राजमहल ढह गया और पुस्तकालय भी जलकर खाक हो गया। इस आग ने पुरावशेषों के साथ एक साथ दो काम किए। केवल छाया में सुखाई गई मट्टी की पट्टिकाएँ लपटों के बीच पककर कठोर हो गईं[4]। पकी हुई मट्टी की पट्टिकाओं को भट्ठों से निकाली ईंटों की तरह दीर्घजीवी स्वरूप मिल गया। यदि उसी आग में पपीरस या चमड़ा होता, तो वह राख बनकर नष्ट हो गया होता।

लेकिन ढहती इमारतों के भारी मलबे ने उन कठोर पट्टिकाओं के टुकड़े-टुकड़े कर दिए। आज ब्रिटिश संग्रहालय में सुरक्षित नीनवे से प्राप्त मट्टी की पट्टियों और टुकड़ों की संख्या 30,000 से अधिक है [3][4]। ये वे पुरावशेष हैं जिन्हें 19वीं शताब्दी के मध्य में उत्खनन दल ने मलबे के ढेरों से निकाला और जहाजों में लादकर लाया था [3]। एक ही कहानी दर्जनों टुकड़ों में बखिरकर हमारे पास पहुँची है।

टुकड़ों का आकार अलग-अलग है। कुछ हथेली जितने बड़े हैं, तो कुछ अंगूठे के नाखून जितने छोटे हैं। कई मामलों में एक टुकड़े पर केवल कुछ ही पंक्तियाँ बची होती हैं। वे चंद पंक्तियाँ महाकाव्य की बारह पट्टिकाओं में कहाँ बैठती हैं, यह कोई नहीं बताता। अक्षरों को पढ़ना और उनका सही स्थान खोजना दोनों पूरी तरह अलग कार्य हैं।

जब कोई टुकड़ा अपनी सही जगह पाता है, तो उसका अर्थ भी बदल सकता है। आगे-पीछे का संदर्भ जुड़ने पर ही पता चलता है कि वह किसकी बात थी और कौन सा दृश्य था। अलग होने पर जो अस्पष्ट शब्दों की सूची मात्र था, वह एक अर्थपूर्ण वाक्य बन जाता है। टुकड़ों का एक सही जोड़ पूरी कहानी के किसी दृश्य को जीवंत कर देता है।

3,50,000 पंक्तियों को हाथ से नहीं जोड़ा जा सकता

टूटे टुकड़ों को जोड़ने का काम लंबे समय तक मानव हाथों और स्मृतिपर निर्भर रहा। वदिवान संग्रह कक्ष से एक टुकड़ा उठाते हैं। फिर वे अन्य टुकड़ों के चित्रों और हाथ से बनाए गए रेखाचित्रों को याद करते हुए उनकी जोड़ी ढूँढते हैं। जब यह नष्टिर्ष नकिलता है कि दो टुकड़े एक ही पट्टिका के हैं, तो इसे जोड़ कहा जाता है।

यह पद्धति मानव स्मृति की क्षमता के कारण सीमिति हो जाती है। असीरियावदिया के आरंभ के बाद से 150 वर्षों में इस तरह जोड़े गए अंश लगभग 5,000 हैं [5]। बचे हुए टुकड़ों की संख्या इससे कहीं अधिक है। अकेले नीनवे से माली 30,000 से अधिक मट्टी की पट्टिकाएँ ब्रिटिश संग्रहालय के पास हैं [3]। अन्य संग्रहों को जोड़ दिया जाए तो यह संख्या बहुत बढ़ जाती है। उनमें से कई तो प्राचीन काल के बाद से कभी पढ़े ही नहीं गए हैं [5]। दुनिया भर के संग्रहालयों में बची कीलाकार पट्टिकाओं की कुल संख्या लगभग 5,00,000 आंकी गई है [6]।

काम के कठिने होने का एक और कारण भी है। टुकड़े कई देशों में बखिरे हुए हैं। ब्रिटिश संग्रहालय, बर्लिन, येल विश्वविद्यालय और फिलाडेल्फिया के पेन संग्रहालय के पास अपने-अपने संग्रह हैं [7]। दो टुकड़ों को सीधे मिलाकर देखने के लिए दो महाद्वीपों के बीच यात्रा करनी पड़ती है।

कीलाकार लिपि स्वयं भी बहुत जटिल और श्रमसाध्य है। लिपिक एक ही शब्द को थोड़ा अलग-अलग तरह से लिखते थे। क्षेत्र और काल के अनुसार वर्तनी बदलती रहती है [5]। इंसान आँखों से दो टुकड़ों के अक्षरों की तुलना करके भी यह आसानी से नहीं पहचान सकता कि वे एक ही रचना के हैं।

इसके अलावा, एक ही कीलाकार चिह्न के कई उच्चारण हो सकते हैं। संदर्भ के अनुसार एक ही नशान का अलग-अलग अर्थ और ध्वनि-मूल्य होता है। ऐसी लाखों पंक्तियों को सामने रखकर जोड़े खोजना मानवीय आँखों के वश की बात नहीं है।

इसलिए जर्मनी के म्यूनिख स्थित लुडविग मैक्सिमिलियन विश्वविद्यालय के प्रोफेसर एनरिक ज़िमेंनेज (Enrique Jiménez) ने एक अलग रास्ता चुना। वर्ष 2018 में उन्होंने बेबीलोनियन साहित्य को यथासंभव पूर्ण रूप में पुनर्निर्मित करने की परियोजना शुरू की। इसका उद्देश्य दुनिया भर में बखिरे मट्टी की पट्टियों के पाठ-पाठनों को एक जगह एकत्र करना है [5]। इस परियोजना का नाम इलेक्ट्रॉनिक बेबीलोनियन साहित्य पुस्तकालय है, जिसे संक्षेप में eBL कहा जाता है। eBL द्वारा टुकड़ों को जोड़ने के लिए बनाया गया डिजिटल कार्यस्थल Fragmentarium है। 2024 तक इस कार्यस्थल पर लगभग 25,000 मट्टी की पट्टिकाओं के पाठ संकलित हो चुके थे। पंक्तियों की संख्या के अनुसार यह 3,50,000 से अधिक पंक्तियाँ हैं [7]।

इस काम की तात्कालिकता का कारण समय है। मट्टी की पट्टिकाएँ आग में पकी होने पर भी अमर नहीं हैं। लवण और नमी सतह को फुला देते हैं और अक्षरों को मट्टि देते हैं। संग्रहालयों के संग्रह कक्षों में पढ़े जाने की प्रतीक्षा कर रही पट्टिकाएँ आज भी लाखों में हैं [5]। उनमें कतिनी कहानियाँ छपी हैं, यह कोई नहीं जानता।

कीलाकार लिपि पढ़ने वाले वदिवानों की संख्या भी एक समस्या है। दुनिया में असीरियावदिया के शोधकर्ता बहुत कम हैं। बची हुई पट्टिकाओं में से केवल एक छोटे हसिसे का ही अनुवाद हो पाया है [6]। एक व्यक्ति जीवन भर में जितनी पट्टिकाएँ पढ़ सकता है, उसकी एक सीमा होती है। मशीनों की सहायता लेने का कारण यही है।

संदर्भ सामग्री

[1] George, A. R. (2003). *The Babylonian Gilgamesh Epic: Introduction, Critical Edition and Cuneiform Texts*. Oxford University Press.

[2] Al-Rawi, F. N. H., & George, A. R. (2014). Back to the Cedar Forest: The beginning and end of Tablet V of the Standard Babylonian Epic of Gilgamesh. *Journal of Cuneiform Studies*, 66, 69-90. <https://doi.org/10.5615/jcunestud.66.2014.0069>

- [3] The Ashurbanipal Library Project. (n.d.). What is the Library? Oracc, University of Pennsylvania. <https://oracc.museum.upenn.edu/asbp/whatisthelibrary/index.html> (3 सितंबर 2026 को अभिगमिति)
- [4] British Museum. (n.d.). What was Ashurbanipal's Library? britishmuseum.org. <https://www.britishmuseum.org/research/projects/what-was-ashurbanipals-library> (3 सितंबर 2026 को अभिगमिति)
- [5] Jiménez, E. (2023, 5월 25일). AI is helping us read ancient Mesopotamian literature. The Conversation. <https://theconversation.com/ai-is-helping-us-read-ancient-mesopotamian-literature-205091> (3 सितंबर 2026 को अभिगमिति)
- [6] Wang, Z. (2026). TabletCraft: Bridging a 4,000-year cultural gap with bidirectional Akkadian NMT and cuneiform rendering. arXiv:2608.02609. <https://arxiv.org/abs/2608.02609> (3 सितंबर 2026 को अभिगमिति)
- [7] Cobanoglu, Y., Laasonen, J., Simonjetz, F., Khait, I., Cohen, S., Földi, Z., Härtinen, A., Heinrich, A., Mitto, T., Rozzi, G., Sáenz, L., & Jiménez, E. (2024). Transliterated cuneiform tablets of the electronic Babylonian Library platform. Journal of Open Humanities Data, 10, 19. <https://doi.org/10.5334/johd.148>

5.2. फॉरसकि दल की पड़ताल: Clay Edge डिजिटल संलयन और Fragmentarium

अक्षरों को वर्णमाला में लपियंतरति करने का पहला चरण

Fragmentarium का प्रारंभिक बटु अक्षर है। शोधकर्ता मट्टी की पट्टियों की तस्वीरों को देखकर कीलाकार लपि को रोमन अक्षरों में लपियंतरति करते हैं[1]। इस कार्य को लपियंतरण कहा जाता है। यह Hanja को Hangul ध्वनियों में लिखने जैसा है। यह प्रत्येक कीलाकार चहिन को उसके ध्वनि-मूल्य के अनुसार वर्णमाला में बदलने का कार्य है। जो भाग टूट जाने के कारण पढ़े नहीं जा सकते, उन्हें रक्ति स्थान के प्रतीक से दर्शाया जाता है।

इस प्रकार तैयार किया गया लपियंतरति पाठ ऐसा स्ट्रिंग बन जाता है जिसे कंप्यूटर संसाधित कर सकता है। eBL ने मई 2018 से अगस्त 2023 तक इन लपियंतरति पाठों का संचय किया। परिणामस्वरूप लगभग 25,000 मट्टी की पट्टिकाएँ और 3,50,000 से अधिक पंक्तियाँ खोज योग्य डेटा बन गई। eBL द्वारा प्रबंधित पट्टिका अभिलेखों की सूची 2,60,000 से अधिक है। ब्रिटिश संग्रहालय, येल बेबीलोनियन संग्रह और पेन संग्रहालय की वस्तुएँ इसमें शामिल हैं[1]।

लपियंतरण का काम इंसान करते हैं। क्योंकि मशीनें तस्वीरों से अक्षरों को पूरी तरह सटीक रूप से नहीं निकाल पाती हैं। eBL शोध दल ने पाँच वर्षों से अधिक समय तक इस लपियंतरण कार्य को जारी रखा। इसका परिणाम सार्वजनिक डेटा के रूप में जारी किया गया, जिसे कोई भी डाउनलोड कर सकता है[1]। मशीन से काम कराने के लिए पहले इंसान को डेटा तैयार करना होता है।

आनुवंशिक परीक्षण से मलिता-जुलता स्ट्रिंग मलिन

मुख्य एल्गोरिदम एक अप्रत्याशित क्षेत्र से आया। यह स्ट्रिंग संरेखण की वही वधि है जिसका उपयोग जीववैज्ञानी जीन अनुक्रमों की तुलना के लिए करते हैं[2]। जीन A, T, G, C इन चार अक्षरों की लंबी श्रृंखला होती है। जब दो जीवों के जीनों की तुलना की जाती है, तो अक्षर थोड़े भिन्न होने पर भी उन्हें एक ही स्थिति में रखकर देखा जाता है।

कीलाकार लपि में भी यही समस्या उत्पन्न होती है। क्योंकि अलग-अलग लपि एक ही शब्द को अलग तरह से लिखते थे। प्रोफेसर जमिनेज ने इस वर्तनी-परिवर्तन को जीन के उत्परिवर्तन जैसा माना। इसलिए उन्होंने जीववैज्ञान के संरेखण एल्गोरिदम को कीलाकार लपि के अनुसार संशोधित करके लागू किया[2]।

कंप्यूटर वास्तव में अक्षरों के छोटे-छोटे टुकड़ों की गिनती करता है। लपियंतरति पाठ को कुछ-कुछ अक्षरों में काटे गए टुकड़ों को n-gram कहा जाता है। यह प्रणाली दोनों मट्टी की पट्टिकाओं में समान n-gram की संख्या की गणना करती है। लंबे समान टुकड़े साझा करने वाली जोड़ी को अधिक अंक मिलते हैं। उच्च अंक पाने वाली जोड़ियाँ मानव शोधकर्ता के सामने उम्मीदवार के रूप में प्रस्तुत की जाती हैं।

Fragmentarium वेब स्क्रीन पर खुलता है। दुनिया के किसी भी देश के शोधकर्ता लॉग इन करके समान डेटा देख सकते हैं। लंदन के टुकड़े और म्यूनिख के टुकड़े को एक ही स्क्रीन पर साथ रखकर तुलना की जा सकती है[8]। अब सामान बाँधकर वमिन में बैठने की कोई आवश्यकता नहीं रह गई है।

यह वधि विवर्तनी में भिन्नता होने पर भी प्रभावी रहती है। यदि कोई एक शब्द अलग लिखा हो, तब भी आस-पास के टुकड़े मेल खाते हैं तो अंक मिलते हैं। eBL ने इस मलिन कार्यक्रम को एक सार्वजनिक रपिजिटरी में उपलब्ध कराया है। अन्य शोधकर्ता भी इस टूल को डाउनलोड करके उपयोग कर सकते हैं[6]।

कंप्यूटर अक्षरों को पढ़ने में भी सहायता करता है। एक ही कीलाकार चहिन के कई उच्चारण होना बहुत आम बात है। eBL द्वारा प्रशिक्षित मॉडल ने इन कई अर्थों में से सही का चयन करने में 98 प्रतिशत की सटीकता प्राप्त की[2]।

यह कार्य क्यों आवश्यक है, इसे एक उदाहरण से आसानी से समझा जा सकता है। एक ही कीलाकार नशान को किसी वाक्य में एक ध्वनि के रूप में पढ़ा जाता है। किसी अन्य वाक्य में उसे पूरे शब्द को दर्शाने वाले प्रतीक के रूप में पढ़ा जाता है। यदि इसे गलत पढ़ा जाए, तो मलिन के अंक गड़बड़ा जाते हैं। पाठ-पाठन को सही करना ही टुकड़ों के जोड़ की सटीकता को बढ़ाना है।

उपलब्धियाँ आँकड़ों में स्पष्ट दिखाई दीं। असीरियावदिया में 150 वर्षों के दौरान हाथ से जोड़े गए टुकड़ों की संख्या लगभग 5,000 थी। eBL ने केवल पाँच वर्षों में लगभग 1,500 नए जोड़ और खोज निकाले[2]। केवल लपियंतरण डेटा संग्रह तैयार करने की प्रक्रिया में ही लगभग 1,250 जोड़ सामने आए[1]। जो सीधे आपस में नहीं जुड़ते, लेकिन एक ही रचना का हिस्सा हैं, ऐसे भी हजारों टुकड़ों की पहचान हुई[2]।

टूटे हुए हिस्सों को भरने वाले भाषा मॉडल

टुकड़ों को जोड़ने के बाद भी पाठ में कई जगह खाली स्थान रह जाते हैं। इन रिक्त स्थानों को लैक्युना (lacuna) कहा जाता है। यहाँ से भाषा मॉडल अपनी भूमिका निभाते हैं।

वर्ष 2023 में तेल अवीव विश्वविद्यालय और एरियल विश्वविद्यालय के शोधकर्ताओं ने अक्कादियन भाषा का अनुवाद मॉडल प्रस्तुत किया। यह मॉडल कीलाकार लपिके चहिन डालने पर या वर्णमाला के लपियंतरण पाठ डालने पर, दोनों को अंग्रेजी में अनुवादित करता है। इसने दोनों प्रकार के इनपुट को संभाला। इसने अंग्रेजी से वापस अक्कादियन में अनुवाद करने का कार्य नहीं किया। प्रशिक्षण के लिए उपयोग किया गया डेटा पहले से प्रकाशित मट्टी की पट्टियों के पाठ थे[3]।

वर्ष 2026 में वपिरीत दशिया में भी काम करने वाला एक टूल सामने आया। TabletCraft नाम का यह एक ओपन सोर्स सॉफ्टवेयर है। यह टूल अक्कादियन से अंग्रेजी और अंग्रेजी से अक्कादियन, दोनों दशियाओं में अनुवाद करता है। इसका प्रशिक्षण डेटा 1,16,000 वाक्य-युग्म हैं। कीलाकार चहिन रूपांतरण तालिका में 14,000 से अधिक मलिन शामिल हैं। डेवलपर ने कहा कि बिची हुई 5,00,000 मट्टी की पट्टियों को पढ़ने के लिए ऐसे साधन की आवश्यकता थी[4]।

ऐसे मॉडल रिक्त स्थान के आगे-पीछे के संदर्भ को देखकर संभावित शब्दों को प्रायिकता के आधार पर प्रस्तुत करते हैं। बेबीलोनियन कविति में समान संरचना वाली दो पंक्तियों को समानांतर रखने का नियम होता है। मॉडल इस प्रारूप को भी एक सुराग के रूप में उपयोग करता है। मॉडल जो प्रस्तुत करता है, वह अंतिम उत्तर नहीं बल्कि संभावित विकल्पों की सूची है।

लखावट की शैली तक को पढ़ने वाला 2026 का टूल

18 मई 2026 को जर्मनी के बुरजबर्ग विश्वविद्यालय ने एक नया टूल जारी किया। इसे मैज वजिजान एवं साहित्य अकादमी और डॉरटमुंड तकनीकी विश्वविद्यालय ने संयुक्त रूप से विकसित किया। इसका नाम Paleographikum है। यह टूल मट्टी की पट्टियों की तस्वीरों से प्रत्येक कीलाकार चहिन के स्वरूप को सीखता है। फिर वह पूरे संग्रह में से उसी तरह दबाकर लिखे गए चहिनों को खोज निकालता है[5]।

यह उपकरण 70,000 तस्वीरों पर काम करता है, और इसमें 50 लाख से अधिक प्रतीक दर्ज हैं। पहले पाँच टुकड़ों की लखावट की तुलना करने में तीन दिन लगते थे। इस उपकरण से यह काम 5 मिनट में पूरा हो जाता है। जिस तरह हर व्यक्तिकी लखावट अलग होती है, उसी तरह हर लपिके के कील दबाने का कोण और अंतराल अलग होता है। उस अंतर को सूत्र मानकर एक ही पट्टिका के टुकड़ों को अलग छाँटा जा सकता है[5]।

इस उपकरण द्वारा विश्लेषित हतिती कीलाकार लपिके में 375 प्रतीक हैं। अब तक ज्ञात हतिती मट्टी की पट्टिकाओं के टुकड़े 30,000 तक हैं। इस शोध का नेतृत्व करने वाले प्रोफेसर डेनयिल श्वैमर ने बताया कि इस उपकरण से हजारों घंटे बचाए जा सकते हैं[5]।

शोध दल का अगला लक्ष्य व्यक्तिगत लपिके की पहचान करना है। इसका अर्थ है एक ही व्यक्ति द्वारा लिखी गई सभी मट्टी की पट्टिकाओं को खोजना। यदि यह संभव हो जाता है, तो यह रेखांकित किया जा सकता है कि एक लपिके ने कहाँ काम किया और वह कहाँ गया। इस प्रकार 3,500 वर्ष पहले के लोगों का व्यावसायिक जीवन आँकड़ों के माध्यम से सामने आता है[5]।

मट्टी के टूटे कनिारों को जोड़ने का दूसरा तरीका

अक्षरों के बजाय आकार के आधार पर टुकड़ों को जोड़ने का शोध भी लंबे समय से चल रहा है। टूटी हुई मट्टी की पट्टिका पर टूटी हुई सतह रह जाती है। इस सतह को वभिग सतह कहा जाता है। यदि दो टुकड़ों की वभिग सतहें आपस में जुड़ जाती हैं, तो वे एक ही पट्टिका से आए होते हैं।

इस दशिया में आगे बढ़ने वाली प्रमुख परियोजना 'आभासी कीलाकार मट्टी की पट्टिका पुनर्निर्माण परियोजना' है। बर्लिन के कील विश्वविद्यालय की सैड्रा वूली इस परियोजना का नेतृत्व करती हैं। टमि कॉलनिस् और एर्लेड गेहलकेन भी उनके साथ हैं। शोधकर्ता टुकड़ों को 3D स्कैनर से स्कैन करते हैं। फिर कंप्यूटर यह तय करता है कि दोनों टुकड़ों की टूटी हुई सतहें आपस में मेल खाती हैं या नहीं। पुराने टुकड़ों के कनिारे घसि होते हैं। इस घसि को भी गणना में शामिल किया जाता है[9]।

इस पद्धत की एक उपलब्ध महाजलप्रलय की कहानी में सामने आई। वह 'अत्रहासीस' नामक जलप्रलय महाकाव्य की तीसरी पट्टिका का टुकड़ा है। एक टुकड़ा लंदन में था और दूसरा जनिवा में था। दोनों स्थान लगभग 1,000 किलोमीटर दूर हैं। वास्तविक वस्तुओं को आमने-सामने न रख पाने के कारण इनके जुड़ाव की पुष्टि नहीं हो पा रही थी। जब 3D डेटा को एक के ऊपर एक रखा गया, तो दोनों टुकड़े आपस में जुड़ गए[9]। यह वर्ष 2014 में एक शोध पत्र के रूप में प्रकाशित हुई थी[10]।

3D डेटा से जुड़ा शोध आज भी जारी है। मार्च 2026 में एक शोध पत्र प्रकाशित हुआ। यह मट्टी की पट्टिका के 3D पॉइंट क्लाउड डेटा से पुरावशेषों की जानकारी को वर्गीकृत करने वाले तंत्रिका नेटवर्क पर आधारित है। पॉइंट क्लाउड वह डेटा है जो किसी वस्तु की सतह को अनगिनत बंदियों के निर्देशांकों के रूप में दर्ज करता है। 3D स्कैनर से मट्टी की पट्टिका को स्कैन करने पर ऐसा डेटा प्राप्त होता है। सतह की बनावट और मोटाई संख्याओं के रूप में दर्ज हो जाती है। इस शोध ने पछिले ट्रांसफॉर्मर-आधारित मॉडल से बेहतर परिणाम दिए[7]।

हालाँकि वर्तमान में फ्रैग्मेंटैरियम द्वारा वास्तव में उपयोग किया जाने वाला मुख्य साधन 3D आकार नहीं है। वह तस्वीरें और लपियंतरण हैं, यानी अक्षर और चित्र[1][2]। टूटी सतहों को जोड़ना प्रत्येक टुकड़े को अलग-अलग स्कैन करने की माँग करता है, जिसमें बहुत मेहनत लगती है। ये दोनों रास्ते एक-दूसरे का स्थान लेने के बजाय साथ-साथ चलते हैं।

तस्वीरों में से प्रतीकों को खोजने का काम भी तेज़ी से आगे बढ़ा है। जून 2026 में eBL के शोधकर्ताओं ने एक नया शोध पत्र प्रकाशित किया। यह मट्टी की पट्टिका की तस्वीरों में कीलाकार प्रतीकों के स्थान का स्वचालित रूप से पता लगाने की वधि है। इसमें 87,000 से अधिक टुकड़ों की तस्वीरों को शामिल किया गया। खोजे गए प्रतीक 29 लाख तक पहुँचते हैं। इसने पछिली वधियों की तुलना में काफी बेहतर प्रदर्शन किया[11]।

संदर्भ सामग्री

- [1] Cobanoglu, Y., Laasonen, J., Simonjetz, F., Khait, I., Cohen, S., Földi, Z., Härtinen, A., Heinrich, A., Mitto, T., Rozzi, G., Sáenz, L., & Jiménez, E. (2024). Transliterated cuneiform tablets of the electronic Babylonian Library platform. *Journal of Open Humanities Data*, 10, 19. <https://doi.org/10.5334/johd.148>
- [2] Jiménez, E. (2023, 5월 25일). AI is helping us read ancient Mesopotamian literature. *The Conversation*. <https://theconversation.com/ai-is-helping-us-read-ancient-mesopotamian-literature-205091> (3 सितंबर 2026 को अभिगमति)
- [3] Gutherz, G., Gordin, S., Sáenz, L., Levy, O., & Berant, J. (2023). Translating Akkadian to English with neural machine translation. *PNAS Nexus*, 2(5), pgad096. <https://doi.org/10.1093/pnasnexus/pgad096>
- [4] Wang, Z. (2026). TabletCraft: Bridging a 4,000-year cultural gap with bidirectional Akkadian NMT and cuneiform rendering. *arXiv:2608.02609*. <https://arxiv.org/abs/2608.02609> (3 सितंबर 2026 को अभिगमति)
- [5] Julius-Maximilians-Universität Würzburg. (2026, 5월 18일). AI reads cuneiform: A milestone for Ancient Near Eastern Studies. *EurekaAlert!*. <https://www.eurekaalert.org/news-releases/1128630> (3 सितंबर 2026 को अभिगमति)
- [6] Electronic Babylonian Literature. (n.d.). ngram-matcher [소프트웨어]. *GitHub*. <https://github.com/ElectronicBabylonianLiterature/ngram-matcher> (3 सितंबर 2026 को अभिगमति)
- [7] Hagelskjær, F. (2026). A novel network for classification of cuneiform tablet metadata. *arXiv:2603.03892*. <https://arxiv.org/abs/2603.03892> (3 सितंबर 2026 को अभिगमति)
- [8] electronic Babylonian Library. (n.d.). Fragmentarium. *ebli.mu.de*. <https://www.ebli.mu.de/fragmentarium> (3 सितंबर 2026 को अभिगमति)
- [9] International Association for Assyriology. (2022, 10월 22일). In the spotlight: The "Virtual Cuneiform Tablet Reconstruction Project". *iaassyriology.com*. <https://iaassyriology.com/in-the-spotlight-the-vctr-project/> (3 सितंबर 2026 को अभिगमति)
- [10] Collins, T., Woolley, S. I., Hernandez Munoz, L., Lewis, A., Ch'ng, E., & Gehlken, E. (2014). Computer-assisted reconstruction of virtual fragmented cuneiform tablets. In *Proceedings of the 2014 International Conference on Virtual Systems and Multimedia* (pp. 70-77). *IEEE*. <https://doi.org/10.1109/VSMM.2014.7136691>
- [11] Che, W., Garcés Arias, E., Niaz, A., Bender, A., & Jiménez, E. (2026). Automated sign detection across the Electronic Babylonian Library: A large-scale dataset and end-to-end cuneiform OCR pipeline. *arXiv:2606.22608*. <https://arxiv.org/abs/2606.22608> (3 सितंबर 2026 को अभिगमति)

5.3. उजागर हुआ सत्य

गलिगमेश महाकाव्य में लौटी पंक्तियाँ

फ्रैगमेंटेरियम ने गलिगमेश महाकाव्य के 20 नए टुकड़ों की खोज की। इसके कारण 100 से अधिक पंक्तियों का मूल पाठ अपने स्थान पर लौट आया। कई लुप्त दृश्य भी पुनः जीवति हो गए। एन्कड्डि द्वारा गलिगमेश को हुम्बाबा की हत्या न करने के लिए समझाने का दृश्य उनमें से एक है। दोनों द्वारा हुम्बाबा को मारने के बाद नपिपुर जाने का दृश्य भी नया जुड़ा। नपिपुर देवता एनललि की पूजा का धार्मिक केंद्र था[1]।

यह वर्धित कतिनी तेज़ है, इसका एक उदाहरण 2025 में सामने आया। बगदाद विश्वविद्यालय के अनमार फ़ाधलि और प्रोफ़ेसर जमिनेज़ ने बेबीलोन की प्रशंसा में लिखी गई एक लंबी कविता को पुनः जीवति किया। यह कविता 1,000 से अधिक वर्षों से वसिमत थी। शोध दल ने eBL प्लेटफ़ॉर्म के माध्यम से उसी कविता से संबंधित 30 प्रतियों को ढूँढ़ निकाला। यदि इसे हाथों से किया जाता, तो दशकों लग जाते। बखिरी हुई प्रतियों को एकत्र करने पर 250 पंक्तियों की इस कविता का अधिकांश भाग फिर से जुड़ गया[3]।

यह कविता बेबीलोन शहर के एक दनि को दर्शाती है। इसमें नदी, नगर के द्वार और लोगों के दैनिक कार्य दिखाई देते हैं। इसमें पुजारियों द्वारा किए जाने वाले कार्यों के नए अभिलेख भी शामिल हैं। 30 प्रतियों का मिलाया यह दर्शाता है कि यह कविता व्यापक रूप से पढ़ी जाती थी। पूरी संभावना है कि यह विद्यालयों में छात्रों द्वारा नकल करके लिखा जाने वाला पाठ रहा होगा[3]।

इस कार्य में लगा समय इस पद्धति की शक्ति को दर्शाता है। यदि 30 प्रतियों को हाथों से खोजना होता, तो दुनिया भर के कई संग्रहालयों की सूचियों को एक-एक करके खंगालना पड़ता। eBL प्लेटफ़ॉर्म पर यही काम केवल एक खोज से पूरा हो जाता है[3]।

यह वर्धित केवल गलिगमेश महाकाव्य में ही उपयोग नहीं की जाती है। बेबीलोनिया के शकुन ग्रंथों, चकित्सा ग्रंथों, शब्दकोशों और प्रार्थनाओं पर भी यही तरीका लागू होता है। eBL द्वारा खोजे गए जुड़ाव कई प्रकार के ग्रंथों में फैले हुए हैं[5]।

मानव हाथों द्वारा पहले सुलझाया गया देवदार का जंगल

मशीन ने सब कुछ नहीं सुलझाया है। गलिगमेश महाकाव्य में व्यापक रूप से ज्ञात नई खोजें मनुष्य के हाथों से ही हुई हैं।

2011 में इराक के सुलेमानिया संग्रहालय ने एक मट्टी की पट्टिका खरीदी। फारूक अल-रावी (Al-Rawi) और एंड्रयू जॉर्ज ने इस पट्टिका को पढ़ा। दोनों ने 2014 में अपने परिणाम प्रकाशित किए। यह पट्टिका मानक संस्करण की 5वीं पट्टिका के शुरुआती और अंतिम भाग को समेटे हुए थी। यहाँ से लगभग 20 पंक्तियों का नया मूल पाठ सामने आया[4]।

नया पाठ देवदार के जंगल को ध्वनियों के माध्यम से चित्रित करता है। बंदर ज़ोर-ज़ोर से आवाज़ें निकालते हैं और झीगुर समूह गान करते हैं। तरह-तरह के पक्षी चहचहाते हैं। ये ध्वनियाँ जंगल के रक्षक हुम्बाबा को प्रतदिनि आनंदति करती थीं। पहली बार जंगल को देखकर दोनों नायकों के दंग रह जाने का दृश्य भी इसमें आता है। हुम्बाबा को मारने के बाद गलिगमेश द्वारा पेड़ों को काटने पर पछताने वाला अंश भी है[4]। मेसोपोटामिया के मैदानों में बंदरों के निवास वाले जंगल नहीं हैं। इसलिए यह वर्णन कहाँ से आया, यह विद्वानों के लिए एक पहली बना हुआ है।

इस पट्टिका के मिलने का स्रोत साफ़-सुथरा नहीं था। संग्रहालय ने इसे एक पुरावशेष व्यापारी से खरीदा था। यह युद्ध के बाद इराक से लूटी गई प्राचीन वस्तुओं को बखिरने से रोकने का एक प्रयास था। यह किस पुरातात्विक स्थल से निकली थी, इसका कोई रिकॉर्ड नहीं है[4]। ऐसी पुरावशेष वस्तुओं के लिए उनके दूसरे टुकड़ों को ढूँढ़ने के सुराग बहुत कम होते हैं।

यह खोज कृत्रिम बुद्धिमत्ता (एआई) के बनिा भी संभव हुई थी। लेकिन भविष्य में ऐसी खोजें कतिनी बार की जा सकेंगी, यह उपकरणों पर निर्भर करता है। सुलेमानिया पट्टिका के दुनिया के सामने आने के लिए संयोग और मानवीय दृष्टि की आवश्यकता थी। eBL उस संयोग को एक खोज में बदलने का प्रयास कर रहा है[5]।

गोलाकार सन्दूक की कहानी

महाजलप्रलय की कहानी की स्थिति भी ऐसी ही है। यदि गलिगमेश महाकाव्य की 5वीं पट्टिका में देवदार का जंगल शामिल है, तो 11वीं पट्टिका में भीषण जलप्रलय का वर्णन है[6]। नीचे उल्लिखित मट्टी की पट्टिका गलिगमेश महाकाव्य की प्रतिनिधी है। यह उसी जलप्रलय परंपरा को दर्शाने वाली और भी प्राचीन पट्टिका है[8]।

ब्रिटिश संग्रहालय के डॉ. इरविंग फकिल ने 2009 में एक मट्टी की पट्टिका की जाँच की। यह एक नज़ी संग्रहकर्ता द्वारा संग्रहालय में लाई गई पट्टिका थी। उन्होंने माना कि यह लगभग 1750 ईसा पूर्व लिखी गई थी। इसका आकार लगभग 11.5 सेटीमीटर चौड़ा और 6 सेटीमीटर लंबा है। इस पर 60 पंक्तियाँ लिखी हैं[8]। इसकी सामग्री में ईश्वर द्वारा एक व्यक्ति को नाव बनाने का निर्देश देने की कहानी है। इसमें जानवरों को जोड़ों में चढ़ाने का भी उल्लेख है[7]।

वर्षीय रूप से ध्यान खींचने वाली बात नाव का आकार है। यह नाव कोई लंबा बक्सा नहीं, बल्कि एक गोल नाव है। यह नदियों में इस्तेमाल होने वाली 'कोराकल' नामक टोकरी जैसी गोल नाव के समान है[7]। पट्टिका पर रस्सियों और लकड़ी के ढाँचों का विवरण है। पानी के रिसाव को रोकने के लिए लगाए जाने वाले बटुमेन की मात्रा भी दी गई है। बटुमेन एक काला और चपिचपि प्राकृतिक तेल है। डॉ. फकिल ने इस पाठ-वाचन को 2014 में एक पुस्तक के रूप में प्रकाशित किया[8]।

इस मट्टी की पट्टिका को कृत्रिम बुद्धिमत्ता (एआई) ने नहीं पढ़ा। इसे एक इंसान ने पंक्ति-दर-पंक्ति पढ़ा। एआई की भूमिका किसी अन्य स्थान पर है। वह दुनिया भर में बखिरी जलप्रलय की कहानियों की प्रतियों को एक जगह एकत्र करने और आपस में मेल खाती पंक्तियों को खोजने का काम है[5]।

मशीन का काम और इंसान का काम

यह क्षेत्र परणामों को तीन स्तरों में विभाजित करके देखता है।

पहला स्तर भौतिक जुड़ाव है। यह वह स्थिति है जहाँ दो टुकड़ों की टूटी हुई सतहें वास्तव में आपस में मेल खाती हैं। इस स्थिति में यह पुष्टि हो जाती है कि दोनों टुकड़े एक ही मट्टी की पट्टिका से आए हैं।

दूसरा स्तर साहित्यिक जुड़ाव है। यह तब होता है जब दो टुकड़े सीधे एक-दूसरे को नहीं छूते, लेकिन यह स्पष्ट हो जाता है कि वे एक ही रचना के अलग-अलग हिस्से हैं। फ्रैगमेंटेरियम द्वारा खोजे जाने वाले अधिकांश परणाम इसी श्रेणी में आते हैं[5]। इस प्रकार एक ही कृति से संबंधित हजारों टुकड़ों की पुष्टि की गई है[2]।

इन दोनों स्तरों की प्रकृतिक भिन्नता है। भौतिक जुड़ाव जब आँखों से सत्यापित हो जाता है, तो वह निश्चित तथ्य बन जाता है। साहित्यिक जुड़ाव संभावनाओं पर निर्भर करता है, इसलिए नया टुकड़ा सामने आने पर इसमें बदलाव हो सकता है। eBL इन दोनों स्तरों को अलग-अलग दर्ज करता है[5]।

तीसरा स्तर व्याख्या का है। यह विचार करने का काम है कि पुनर्जीवित मूल पाठ का क्या अर्थ है। मशीन यह नहीं जानती कि देवदार के जंगल में बंदरों की आवाज़ किस स्मृति से आई है। गोल नाव बाइबिल के चौकोर सन्दूक में कैसे बदल गई, मशीन इसका भी उत्तर नहीं दे सकती। यह प्रश्न प्राचीन ग्रंथों के शोधकर्ताओं और इतिहासकारों के हिस्से का है।

मशीन द्वारा सुझाए गए उम्मीदवार मानवीय समीक्षा के बाद ही आधिकारिक रिकॉर्ड बनते हैं। विद्वान उम्मीदवारों की सूची खोलते हैं और दोनों टुकड़ों की तस्वीरों को साथ-साथ देखते हैं। वे अक्षरों के आकार, पंक्तियों के बीच की दूरी और पट्टिका की मोटाई की जाँच करने के बाद जुड़ाव की पुष्टि करते हैं[5]। मशीन का काम तो अनगिनत जोड़ों में से केवल कुछ विचारणीय जोड़ों को छोटकर देने तक ही सीमित है[2]।

जून 2026 में प्रकाशित एक समीक्षात्मक लेख भी यही सीमा रेखा खींचता है। वर्तमान कृत्रिम बुद्धिमत्ता कीलाकार लिपि को अपने आप नहीं पढ़ सकती। यह मानव अनुसंधान में सहायता करने वाला एक उपकरण मात्र है[9]। वुर्जबर्ग विश्वविद्यालय का पैलियोग्राफ़िक भी ऐसा ही है। यह लिखावट की तुलना में लगने वाले तीन दिनों के समय को घटाकर केवल 5 मिनट कर देता है। क्या तुलना करनी है और परणामों का मूल्यांकन कैसे करना है, यह तय करने वाला इंसान ही है[10]।

जो बदला है, वह गत है। 150 वर्षों में जो जुड़ाव 5,000 मामले थे, वे पछिले पाँच वर्षों में 1,500 बढ़ गए[2]। गलिंगमेश जसि अमर जीवन की तलाश में भटका था, वह अंततः उसकी कहानी को मिला गया। मट्टी में टूटी हुई अवस्था में 2,600 वर्षों तक प्रतीक्षा करने वाले वाक्य अब एक-एक करके अपने स्थान पर लौट रहे हैं।

अभी बहुत काम बाकी है। दुनिया की 5 लाख मट्टी की पट्टिकाओं में से केवल कुछ ही पढ़ी जा सकी हैं[11]। eBL के लिप्यंतरण डेटा में शामिल पट्टिकाएं 25,000 हैं[5]। हित्ति भाषा के 30,000 टुकड़ों को अभी-अभी तस्वीरों के रूप में व्यवस्थित किया गया है[10]। टूटी हुई सतहों को जोड़ने का 3D कार्य अभी शुरुआती चरण में है। अगली पीढ़ी के छात्र इस सूची को आगे बढ़ाएंगे। उस समय गलिंगमेश महाकाव्य के रक्ति स्थान आज की तुलना में कम हो चुके होंगे। ननिवे की आग से बचे टुकड़े अभी पूरी तरह पढ़े नहीं गए हैं। उन्हें पढ़ने वाले उपकरण हर साल बढ़ रहे हैं।

संदर्भ सामग्री

[1] Ofgang, E. (2024, 8월 12일). Piecing together an ancient epic was slow work. Until AI got involved. The New York Times. https://edition.pagesuite.com/tribune/article_popover.aspx?guid=235bb893-7f2d-41cd-9857-139c7d1e8355 (3 सितंबर 2026 को अभिगमिति)

[2] Jiménez, E. (2023, 5월 25일). AI is helping us read ancient Mesopotamian literature. The Conversation. <https://theconversation.com/ai-is-helping-us-read-ancient-mesopotamian-literature-205091> (3 सितंबर 2026 को अभिगमिति)

- [3] Fadhil, A. A., & Jiménez, E. (2025). Literary texts from the Sippar Library V: A hymn in praise of Babylon and the Babylonians. *Iraq*, 1-58. <https://doi.org/10.1017/irq.2024.23>
- [4] Al-Rawi, F. N. H., & George, A. R. (2014). Back to the Cedar Forest: The beginning and end of Tablet V of the Standard Babylonian Epic of Gilgameš. *Journal of Cuneiform Studies*, 66, 69-90. <https://doi.org/10.5615/jcunestud.66.2014.0069>
- [5] Cobanoglu, Y., Laasonen, J., Simonjetz, F., Khait, I., Cohen, S., Földi, Z., Härtinen, A., Heinrich, A., Mitto, T., Rozzi, G., Sáenz, L., & Jiménez, E. (2024). Transliterated cuneiform tablets of the electronic Babylonian Library platform. *Journal of Open Humanities Data*, 10, 19. <https://doi.org/10.5334/johd.148>
- [6] George, A. R. (2003). *The Babylonian Gilgamesh Epic: Introduction, Critical Edition and Cuneiform Texts*. Oxford University Press.
- [7] PBS NewsHour. (2014, 1월 24일). New evidence suggests Noah's ark may have been round. [pbs.org. https://www.pbs.org/newshour/world/new-evidence-suggests-noahs-ark-may-have-been-round](https://www.pbs.org/newshour/world/new-evidence-suggests-noahs-ark-may-have-been-round) (3 सितंबर 2026 को अभिगमति)
- [8] Finkel, I. (2014). *The Ark Before Noah: Decoding the Story of the Flood*. Hodder & Stoughton.
- [9] Centre for the Study of Manuscript Cultures. (2026, 6월 14일). Can artificial intelligence decipher cuneiform tablets? Universität Hamburg. <https://www.csmc.uni-hamburg.de/publications/mesopotamia/2026-06-14.html> (3 सितंबर 2026 को अभिगमति)
- [10] Julius-Maximilians-Universität Würzburg. (2026, 5월 18일). AI reads cuneiform: A milestone for Ancient Near Eastern Studies. *EurekAlert!*. <https://www.eurekalert.org/news-releases/1128630> (3 सितंबर 2026 को अभिगमति)
- [11] Wang, Z. (2026). TabletCraft: Bridging a 4,000-year cultural gap with bidirectional Akkadian NMT and cuneiform rendering. *arXiv:2608.02609*. <https://arxiv.org/abs/2608.02609> (3 सितंबर 2026 को अभिगमति)

अध्याय 6. प्राचीन मेसोपोटामिया की रसीदे और व्यापारियों की फुसफुसाहट: AI अनुवाद

6.1. केस फाइल

मिट्टी पर उकेरी गई रसीदे

प्राचीन मेसोपोटामिया के लोग गीली मिट्टी की पट्टिकाओं पर लिखते थे। वे नरकट के सरि को तराशकर लेखनी बनाते थे। लेखनी बना स्याही के केवल नशान बनाने वाला उपकरण है। वे इस लेखनी को मिट्टी पर दबाते थे और तुरंत उठा लेते थे[1]। दबा हुआ नशान कील जैसे तकिये आकार का बन जाता था। इस लिपिका नाम कीलाकार लिपि है। अकादमिक जगत में इसे Seolhyeongmunja भी कहा जाता है। Seolhyeong का अर्थ कील का आकार होता है।

मिट्टी सूखने पर पत्थर जैसी सख्त हो जाती है। कागज़ या चमड़े के विपरीत, यह समय बीतने पर भी सड़ती नहीं है। आग लगने पर भी यह राख नहीं बनती। बल्कि तेज़ गर्मी मिट्टी को ईंट की तरह ठोस बना देती है। युद्ध में जब शहर जले, तब मिट्टी की पट्टिकाएँ और मज़बूत हो गईं[3]। उसी बदौलत 4 हजार साल पुराने बही-खाते आज तक बचे रहे।

मिट्टी की कई पट्टिकाएँ हथेली के बराबर हैं। यह जब में रखकर ले जाने लायक आकार है। लिपिकी गीली मिट्टी को गूँधकर पट्टिका बनाते थे। लिपिकी दूसरों के लिए लिखने और पढ़कर सुनाने वाले लोग थे। लिखना पूरा होने पर वे इसे धूप में सुखाते थे। इन्हें जानबूझकर भट्टी में पकाना दुर्लभ था। आज संग्रहालयों में इन्हें पकाना संरक्षण के लिए उठाया गया कदम है[2]। यदि कुछ सुधारना होता, तो वे सतह को रगड़कर मटि देते थे।

अब तक खोजी गई मिट्टी की पट्टिकाओं की संख्या लगभग 5,00,000 आंकी गई है[4]। ज़मीन के नीचे अभी भी ऐसी कई पट्टिकाएँ दबी हैं जिनसे निकाला नहीं गया है। यहाँ लिखे शब्द केवल राजा की प्रशंसा तक सीमिति नहीं हैं। अनाज की कुछ बोरियाँ मलिन की रसीदे हैं। चाँदी उधार देने और ब्याज तय करने वाले अनुबंध हैं। दूर गए पत्रों को पत्नी द्वारा भेजा गया पत्र भी है[6]।

एक रसीद छोटी होती है। पहले वस्तु का नाम और मात्रा आती है। उसके बाद देने वाले और पाने वाले का नाम आता है। अंत में तारीख और गवाह का नाम लिखा जाता है। यह क्रम आज दुकानों से मलिन वाली रसीद जैसा ही है। ऐसे एक दस्तावेज़ की सामग्री बहुत सीमिति होती है। लेकिन जब हजारों पट्टिकाएँ एकत्र होती हैं, तो बात बदल जाती है। पता चलता है कि किस वर्ष अनाज की कीमतें बढ़ीं। यह भी दिखाई देता है कि किस शहर ने क्या खरीदा और बेचा।

अनूदति न हुआ अधिकांश हस्सि

समस्या यह है कि इन लेखों को शायद ही पढ़ा गया है। मिट्टी की पट्टिकाओं में से केवल एक छोटा हस्सि ही आधुनिक भाषाओं में अनूदति हुआ है[5]। बाकी पट्टिकाएँ संग्रहालयों के बक्सों में सो रही हैं।

हर कोई लिपिकी नहीं बन सकता था। लिपिकी बनने के लिए स्कूल में लंबे समय तक पढ़ना पड़ता था। छात्र मिट्टी की पट्टिकाओं पर प्रतीकों को बार-बार लिखते थे। गलत लिखी गई पट्टिका को पानी में घोलकर फिर से बनाया जाता था। बची हुई अभ्यास पट्टिकाएँ भी अवशेषों के रूप में खोदी जाती हैं। ये अभ्यास पट्टिकाएँ उस युग की कक्षा को दर्शाती हैं।

लिखना-पढ़ना जानने वाले लोग बहुत कम थे। ज्यादातर लोग लिपिकी से अपने लिए लिखवाते थे। पत्र मलिन पर लिपिकी ही उसे पढ़कर सुनाता था। पत्रों के वाक्यों का एक औपचारिक और वनिम्र ढाँचा होने का यही कारण था। बोलने वाला और लिखने वाला व्यक्ति अलग-अलग होते थे।

मिट्टी की एक पट्टिका को पढ़ने के लिए भी कई चरणों से गुजरना पड़ता है। सबसे पहले एक-एक प्रतीक को पहचाना जाता है। इसके बाद तय किया जाता है कि वह प्रतीक किस ध्वनिका प्रतिनिधित्व करता है। फिर शब्दों को अलग किया जाता है और उनके अर्थ खोजे जाते हैं। अंत में वाक्यों को जोड़कर अनुवाद किया जाता है। यदि प्रारूप जाना-पहचाना हो, तो काम जल्दी समाप्त हो जाता है। अपरचित दस्तावेज़ों में लंबा समय लगता है।

कारण यह है कि इसे पढ़ने वाले लोगों की कमी है। कीलाकार लिपिका अध्ययन करने वाले विषय को असीरियोलॉजी (Assyriology) कहा जाता है। एक विद्वान को तैयार करने में दशकों लग जाते हैं। क्योंकि प्राचीन भाषा और प्राचीन इतिहास दोनों को साथ सीखना पड़ता है। इसलिए दुनिया में इस लिपिकी पढ़ने वाले लोग बहुत ही कम हैं[5]।

नकिलने वाली नई मिट्टी की पट्टिकाओं की संख्या विद्वानों के पढ़ने की गति से कहीं अधिक है। विशेषज्ञों की संख्या नहीं बढ़ रही, इसलिए केवल बक्से जमा होते जा रहे हैं। जून 2026 में ब्रिटन के यॉर्क विश्वविद्यालय ने एक नया शोध प्रोजेक्ट शुरू किया। इस प्रोजेक्ट का नाम

"कीलाकार वरिसत तक सार्वभौमिक पहुँच की ओर" है। योजना कीलाकार सामग्री का अरबी में भी अनुवाद करने की है ताकि कोई भी इसे देख सके। यह प्रोजेक्ट 2029 तक 3 वर्षों तक जारी रहेगा[7]।

सामग्री को एक स्थान पर एकत्र करने का कार्य भी साथ-साथ चल रहा है। कीलाकार डिजिटल लाइब्रेरी प्रोजेक्ट ऐसी ही एक पहल है। यह दुनिया भर में बखिरी मट्टी की पट्टिकाओं की जानकारी एकत्र कर इंटरनेट पर सार्वजनिक करता है। तस्वीरें, लिप्यंतरण और अनुवाद एक साथ रखे जाते हैं। लिप्यंतरण कीलाकार प्रतीकों की ध्वनियों को वर्णमाला में लिखा गया रूप है। अब तक 4,00,000 से अधिक कलाकृतियों को सूचीबद्ध किया जा चुका है[4]। डेटा इकट्ठा होने पर ही मशीनों को सीखने की सामग्री मिलती है।

मट्टी की पट्टियाँ आज कई देशों में बखिरी हुई हैं। इराक और तुर्किये के संग्रहालयों में इनकी बड़ी संख्या मौजूद है। लंदन, पेरिस और बर्लिन के संग्रहालयों के पास भी इनका बड़ा हस्सा है। ये अमेरिका और जापान के विश्वविद्यालयों में भी बँटी हुई हैं। यद्यपि किसी एक दस्तावेज़ का अगला और पछिला हस्सा अलग-अलग देशों में हो, तो उन्हें जोड़ना कठिन होता है। इंटरनेट पर तस्वीरें एक साथ अपलोड होने के बाद ही उनका मिलान किया जा सकता है।

राजा के अभिलेखों के पीछे छूटे बही-खाते

रोज़मर्रा के दस्तावेज़ों के पीछे छूट जाने के पीछे एक पुराना कारण है। 19वीं और 20वीं सदी की शुरुआत में उत्खनन दलों ने सबसे पहले राजमहलों और मंदिरों की खुदाई की। विशाल शिलालेखों और सोने के आभूषणों ने लोगों का ध्यान खींचा। विद्वानों ने भी शाही शिलालेखों और कानून संहिताओं को पहले पढ़ा। उन्होंने अपना समय सबसे पहले मथिकों और महाकाव्यों के अनुवाद में लगाया।

इस बीच व्यापारियों की रसीदें और पत्र प्राथमिकता सूची में पीछे छूट गए। अनाज वितरण की सूचियों को मामूली काम के रिकॉर्ड माना गया। एक-एक करके देखने पर इनमें कोई विशेष सार नहीं दिखता था। विद्वानों ने इन उबाऊ दस्तावेज़ों को हाथ से लिखने का काम टाल दिया। इस तरह आम लोगों की आवाज़ लंबे समय तक दबी रही।

इस पक्षपात के कारण बहुत कुछ खो गया। कीमतों में किस तरह उतार-चढ़ाव आया, यह जानना कठिन हो गया। महिलाएँ क्या काम करती थीं, यह छुप गया। बच्चों और सेवकों के नाम केवल सूचियों में ही समिट कर रह गए। राजा के शिलालेख आत्मप्रशंसा लिखते हैं। रसीदें अपनी बड़ाई नहीं करती। वे केवल लेन-देन के तथ्य दर्ज करती हैं। इसलिए रसीदें कभी-कभी अधिक ईमानदार ऐतिहासिक स्रोत बन जाती हैं।

एक ही कीलाकार चहिन के कई अर्थ होते हैं

मशीनों से पढ़ने की कोशिश करने पर भी कीलाकार लिपि आसान नहीं है। क्योंकि एक ही प्रतीक कई भूमिकाएँ निभाता है[8]। किसी स्थान पर यह ध्वनिदर्शाने वाला शब्दांश प्रतीक बन जाता है। किसी अन्य स्थान पर यह पूरे शब्द का अर्थ व्यक्त करता है। और कुछ स्थानों पर इसका कोई उच्चारण नहीं होता। यह केवल यह बताने के लिए प्रयोग होता है कि अगला शब्द पेड़ है या देवता। ऐसे चहिनों को निर्धारक (डिटरमिनेटिव) कहा जाता है।

कीलाकार लिपि केवल एक भाषा लिखने वाली लिपि नहीं थी। सुमेरियन और अक्कादी भाषाएँ एक ही लिपि का उपयोग करती थीं। बाद में हित्ति और एलामी भाषाओं ने भी इस लिपि को उधार लिया। एक ही प्रतीक हर भाषा में अलग ध्वनि निकालता है। पढ़ने वाले को सबसे पहले दस्तावेज़ की भाषा तय करनी पड़ती है।

पेड़ का अर्थ देने वाला सुमेरियन प्रतीक gish इसका अच्छा उदाहरण है। फर्नीचर के नाम के आगे आने पर यह संकेत देता है कि सामग्री लकड़ी है। अकेले आने पर यह स्वयं लकड़ी शब्द बन जाता है। अन्य स्थानों पर यह केवल gish ध्वनिकी प्रतिनिधित्व करता है। पढ़ने वाले को आगे-पीछे के संदर्भ को देखकर इसकी भूमिका चुननी पड़ती है।

संख्याएँ लिखने का तरीका भी आज से भिन्न था। मेसोपोटामिया के लोग 60 को एक इकाई मानते थे। घड़ी के 60 मिनट और 60 सेकंड यही से आए हैं। रसीदों में मात्रा भी इसी पद्धति से लिखी जाती थी। अनाज और चाँदी के लिए अलग-अलग इकाइयों का प्रयोग होता था। एक ही प्रतीक अनाज होने पर और चाँदी होने पर अलग मान रखता है। मशीन को इस इकाई के अंतर को भी साथ सीखना होता है।

शब्दों को जोड़कर लिखना भी एक बड़ी बाधा है। प्राचीन लिपिकी शब्दों के बीच जगह नहीं छोड़ते थे। कंप्यूटर के लिए यह पता लगाना मुश्किल होता है कि शब्द कहाँ समाप्त होता है[8]। मट्टी की पट्टिकाओं का टूटना और घिस जाना भी एक समस्या है। कीलाकार निशान हल्के पड़ जाँ, तो मानवीय आँखों से भी भेद करना कठिन हो जाता है। अक्कादी भाषा में शब्दों के रूपों का बार-बार बदलना भी जटिल है। एक ही मूल रूप में कई प्रत्यय जुड़ते हैं। कर्ता है या कर्म, इसके अनुसार अंतिम ध्वनि बदल जाती है। काल और पुरुष के अनुसार यह फरि बदल जाती है। शब्दकोश में खोजने के लिए पहले मूल रूप का पता लगाना पड़ता है।

मट्टी की पट्टिकाओं के टूटने के कारण भी कई हैं। वे ज़मीन के नीचे दबती हैं और पानी में भीगती हैं। उत्खनन के दौरान उनके कोने टूटकर अलग भी हो जाते हैं। पुराने समय में कलाकृतियों के व्यापार के कारण टूटी हुई पट्टिकाएँ कई देशों में बखिर गईं। ज़मीन से निकालने के बाद नमक का रसिब अक्षरों को ढक भी देता है[2]। विद्वान अक्षरों के गायब होने से बने रिक्त स्थान को gyeolsonbu (lacuna) कहते हैं।

बाधाओं की इन अनेक परतों के कारण लाखों बही-खाते अब तक मौन रहे हैं।

संदर्भ सामग्री

- [1] Truman State University. (2026). How it works. Truman State University Cuneiforms. <https://exhibits.truman.edu/cuneiform/how-it-works/> (3 सितंबर 2026 को अभिगमिति)
- [2] Guinan, A., Oller, G., & Ormsby, D. (1976). Nippur rebaked: The conservation of cuneiform tablets. Expedition, 18(3). <https://www.penn.museum/sites/expedition/nippur-rebaked/> (3 सितंबर 2026 को अभिगमिति)
- [3] Bouscaren, D. (2023). Assyrian women of letters. Archaeology, 76(6). <https://archaeology.org/issues/november-december-2023/features/assyrian-women-of-letters/> (3 सितंबर 2026 को अभिगमिति)
- [4] Cuneiform Digital Library Initiative. (2026). About the CDLI. [cdli.earth](https://cdli.earth/about). <https://cdli.earth/about> (3 सितंबर 2026 को अभिगमिति)
- [5] Gutherz, G., Gordin, S., Sáenz, L., Levy, O., & Berant, J. (2023). Translating Akkadian to English with neural machine translation. PNAS Nexus, 2(5), pgad096. <https://doi.org/10.1093/pnasnexus/pgad096>
- [6] Michel, C. (2020). Women of Assur and Kanesh: Texts from the Archives of Assyrian Merchants (Writings from the Ancient World 42). SBL Press.
- [7] Cuneiform Digital Library Initiative. (2026). Research Programme "Towards Universal Access to Cuneiform Heritage" 2026-2029. [cdli.earth](https://cdli.earth/postings/239). <https://cdli.earth/postings/239> (3 सितंबर 2026 को अभिगमिति)
- [8] Gordin, S., Gutherz, G., Elazary, A., Romach, A., Jiménez, E., Berant, J., & Cohen, Y. (2020). Reading Akkadian cuneiform using natural language processing. PLoS ONE, 15(10), e0240511. <https://doi.org/10.1371/journal.pone.0240511>

6.2. वैज्ञानिक जाँच दल की पड़ताल: अक्कादी AI अनुवाद

अनुवाद के दो रास्ते तैयार करना

2023 में तेल अवीव विश्वविद्यालय के शोधकर्ताओं ने सबसे पहले यह राह खोली। गैरि गुथरज़ (Gai Gutherz) और उनके साथियों ने अक्कादी भाषा का अंग्रेज़ी में अनुवाद करने वाला मॉडल बनाया[1]। अक्कादी भाषा मेसोपोटामिया में लंबे समय तक बोली जाने वाली भाषा थी। यह कीलाकार लिपि में लिखी गई थी और आज इसे बोलने वाला कोई नहीं है।

शोध दल ने अनुवाद के दो रास्ते तैयार किए। पहले रास्ते का नाम C2E है। यह कंप्यूटर यूनिकोड में लिखे कीलाकार प्रतीकों को दर्ज़ करके अंग्रेज़ी प्राप्त करने का मार्ग है। यह सीधे मट्टी की पट्टिका की तस्वीर पढ़ने की वधि नहीं है। वद्वानों के लिप्यंतरण की आवश्यकता न होने के कारण यह चरण छोटा है। दूसरे रास्ते का नाम T2E है। यह वह तरीका है जिसमें वद्वान द्वारा कीलाकार प्रतीकों की ध्वनिको वर्णमाला में लिखे गए पाठ को इनपुट दिया जाता है। वर्णमाला के इस रूपांतरण को लिप्यंतरण कहा जाता है। लिप्यंतरण के चरण में प्रतीकों का अर्थ एक बार व्यवस्थित हो जाता है। इसलिए T2E अनुवाद अधिक सटीक होता है[1]।

लिप्यंतरण तैयार करना भी आसान काम नहीं है। वद्वान प्रतीकों को देखकर उसकी ध्वनितय करते हैं। यदि कोई ध्वनियों हो, तो वे संदर्भ देखकर चुनते हैं। यह चुनाव ही मानवीय नरिणय है। एक बार नरिणय हो जाने के बाद मशीन के लिए उस पाठ को संभालना आसान हो जाता है। T2E का स्कोर अधिक होने का कारण यही है।

अनुवादक की आंतरिक संरचना दो भागों में बँटी है। पहला भाग मूल पाठ को संख्याओं के समूह में बदलता है। दूसरा भाग उन संख्याओं को देखकर एक-एक अंग्रेज़ी शब्द चुनता है। पहले भाग को एन्कोडर और दूसरे भाग को डिकोडर कहा जाता है। एन्कोडर शब्दों के अर्थ और उनके क्रम दोनों को समाहित करता है। डिकोडर पहले चुने गए शब्दों को देखते हुए अगले शब्द का नरिधारण करता है।

तस्वीरों से सीधे प्रतीकों को पढ़ने का शोध अलग से जारी है। जून 2026 में इलेक्ट्रॉनिक बेबीलोनियन लाइब्रेरी के शोधकर्ताओं ने परणाम प्रस्तुत किए। मशीन ने 87,000 से अधिक टूटी हुई मट्टी की पट्टिकाओं की तस्वीरों को स्कैन किया। पहचाने गए प्रतीक चिह्नों की संख्या लगभग 29 लाख के करीब है। पछिली वधियों की तुलना में पहचान की सटीकता 37 प्रतिशत तक बढ़ गई। लेकिन यह अत्यधिक टूटी हुई पट्टिकाओं और अव्यवस्थित पंक्तियों की स्थिति में अभी भी कमज़ोर है[6]।

8,000 दस्तावेज़ों से सीखने वाली मशीन

मशीन उदाहरण वाक्यों को देखकर अनुवाद सीखती है। शोधकर्ताओं ने Oracc नामक एक सार्वजनिक डेटाबेस से दस्तावेज़ एकत्र किए। Oracc कीलाकार साहित्य को व्यवस्थित और सार्वजनिक करने वाला एक ऑनलाइन संग्रह है। यहाँ से एकत्र किए गए दस्तावेज़ों की संख्या 8,056 थी। वाक्यों की संख्या 50,000 से अधिक थी[1]।

एकत्र किए गए दस्तावेज़ वभिन्न प्रकार के थे। राजा की उपलब्धियों को उकेरने वाले शिलालेख लगभग 3,000 थे। प्रशासनिक पत्र लगभग 2,000 थे। इनमें न्यायिक रिकॉर्ड और ज्योतिषि रिपोर्टें भी शामिल थीं[1]। प्रत्येक दस्तावेज़ के साथ वद्वानों द्वारा किया गया अंग्रेज़ी अनुवाद संलग्न था। मशीन ने इन जोड़ियों को हजारों बार देखकर अपने आप नियमों को खोज निकाला।

मशीन के सीखने की प्रक्रिया परीक्षा की तैयारी जैसी होती है। पहले वह प्रश्न और उत्तर दोनों को साथ देखती है। फिर उत्तर छपाकर स्वयं हल करने का प्रयास करती है। गलती होने पर वह देखती है कि गलती कहाँ हुई। वह इस समीक्षा को लाखों बार दोहराती है। मनुष्यों की तुलना में उसका अंतर केवल गति और नरितरता का है।

मशीन शब्दों को संपूर्ण रूप में संसाधित नहीं करती। वह शब्दों को छोटे-छोटे टुकड़ों में वभिजति करके संभालती है। इन टुकड़ों को टोकन कहा जाता है। दुर्लभ शब्दों को भी टुकड़ों में बाँटकर संभाला जा सकता है। कई प्रत्ययों वाली अक्कादी भाषा के लिए यह वधिसिटीक बैठती है।

अनुवाद की गुणवत्ता को BLEU नामक स्कोर से मापा जाता है। यह तरीका मापता है कि मशीन का अनुवाद मानव अनुवाद से कतिना मेल खाता है। यह मेल खाने वाले शब्दों के समूहों के अनुपात के आधार पर स्कोर देता है। इसे 0 से 100 तक के अंकों में दर्शाया जाता है। 30 अंक से ऊपर होने पर अक्सर माना जाता है कि अर्थ स्पष्ट हो रहा है। हालाँकि यह कोई पूर्ण पैमाना नहीं है। दस्तावेज़ के प्रकार और वाक्य की लंबाई के अनुसार व्याख्या बदल जाती है।

शोधकर्ताओं ने कनवलशनल न्यूट्रल नेटवर्क (CNN) नामक संरचना को चुना। यह एक ऐसी संरचना है जो निकटतम सूचनाओं को एक साथ बाँधकर सीखती है। क्योंकि डेटा कम होने पर इसका सीखना तेज़ और स्थिर होता है। सत्यापन के परिणामों में C2E को 36.52 अंक मिले। T2E का स्कोर 37.47 के साथ और भी अधिक था। पहले इसे माल होने वाली स्वचालित खोज वधियों केवल 27.09 और 23.51 अंक ही प्राप्त कर सकी थी। छोटे वाक्यों और निर्धारित प्रारूप वाले दस्तावेज़ों में इसके परिणाम बहुत अच्छे रहे[1]।

डेटा को वभिजति करने का तरीका भी महत्वपूर्ण है। सीखने वाले हस्से और परीक्षण वाले हस्से को अलग-अलग रखा जाता है। यदि परीक्षण वाले दस्तावेज़ पहले ही दिखा दिए जाएँ, तो स्कोर निरर्थक हो जाता है। यह मानवीय परीक्षा में प्रश्नों को पहले से जानने जैसा है। शोध दल ने इस अंतर को बनाए रखते हुए स्कोर निकाला।

एक ही मूल पाठ का अनुवाद भी हर व्यक्ति के अनुसार भिन्न होता है। इसलिए केवल स्कोर देखकर अच्छा या बुरा तय नहीं किया जाता। स्कोर के अलावा मानवीय मूल्यांकन का भी उपयोग किया जाता है। अनुवाद वद्वानों को दिखाया जाता है और उनसे ग़रेडिग करवाई जाती है। उनसे अलग से पूछा जाता है कि क्या अर्थ समझ में आ रहा है। यह भी अलग से पूछा जाता है कि क्या कहीं कोई त्रुटि है। वद्वानों द्वारा प्रयोग योग्य माने गए अनुवाद T2E में लगभग 44 प्रतिशत थे। C2E में यह लगभग 38 प्रतिशत था[1]। कभी-कभी स्कोर और मानवीय मूल्यांकन में अंतर आ जाता है। ऐसे समय में मनुष्य के निर्णय को प्राथमिकता दी जाती है।

अक्कादी भाषा का अनुवाद अल्प-संसाधन (low-resource) कार्यों के अंतर्गत आता है। अल्प-संसाधन का अर्थ है सीखने के लिए उदाहरण वाक्यों की कमी। अंग्रेज़ी और फ़्रेंच अनुवादक करोड़ों वाक्य सीखते हैं। अक्कादी भाषा के युग्मति वाक्य उसके हजारों हस्से के बराबर भी नहीं हैं। डेटा कम होने पर बड़े मॉडल उलटे भटक जाते हैं।

2025 में कई शोध टीमों के बीच एक साझा प्रतियोगिता भी आयोजित की गई। कार्य शब्दों के मूल रूप खोजना और लुप्त अक्षरों का अनुमान लगाना था। सामग्री इलेक्ट्रॉनिक बेबीलोनियन लाइब्रेरी और ARCHIBAB से ली गई थी। एक ही डेटा पर परिणामों को मापने से तुरंत स्पष्ट हो गया कि कौन सी वधि बेहतर है[2]।

मूल रूप खोजना अनुवाद का प्राथमिक आधार कार्य है। अकादमिक जगत में इस कार्य को लेमाटाइजेशन (lemmatization) कहा जाता है। यह प्रत्यय जुड़कर बढ़े हुए शब्दों को उनके शब्दकोश रूप में लौटाने का कार्य है। यह प्राथमिक कार्य पूरा होने के बाद ही शब्द-वार खोज और सांख्यिकी संभव हो पाती है।

वश्वसनीय लगने वाले झूठ को पकड़ना

मशीन अनुवाद की एक खतरनाक आदत होती है। बनिा किसी आधार के भी यह सच लगने वाले वाक्य गढ़ लेती है। इस आदत को मतभ्रम कहा जाता है। शोधकर्ताओं ने मतभ्रम को दो भागों में बाँटकर देखा। एक वह स्थिति है जहाँ मूल पाठ को देखकर भी उसका गलत अर्थ निकाला जाता है। दूसरी वह स्थिति है जहाँ मूल पाठ की अनदेखी करके कोई अलग ही कहानी बना दी जाती है[1]।

टूटी हुई मट्टी की पट्टिकाओं में मतभ्रम अधिक होता है। ऐसा इसलिए है क्योंकि अक्षर मट्टि से बने खाली स्थानों को मशीन अपनी कल्पना से भर देती है। इसलिए अनुवाद की दोबारा जाँच वद्वानों को अवश्य करनी चाहिए। मशीन द्वारा तैयार किया गया वाक्य केवल पहला मसौदा है, कोई अंतिम निर्णय नहीं।

मतभ्रम को कम करने के कई तरीके हैं। टूटे हुए हसिंसों को रक्ति स्थान के चहिन से स्पष्ट रूप से दर्शाया जाता है। मशीन को यह उत्तर देने दिया जाता है कि वह उस अंश को नहीं जानती। एक ही वाक्य का कई बार अनुवाद भी कराया जाता है। यदि उत्तर हर बार बदलता है, तो उस पर विश्वास नहीं किया जाता। वदिवान इन संकेतों को देखकर यह तय करते हैं कि किस हसिंसे की दोबारा समीक्षा करनी है।

कीलाकार प्रतीकों को वर्णमाला में बदलना और शब्दों को अलग करना अनुवाद से पहले का काम है। उसी शोध धारा में यह काम भी मशीन को सौंपकर देखा गया। वदिविशी LSTM नामक तंत्रिका नेटवर्क का प्रदर्शन सबसे आगे रहा। इसकी सटीकता अधिकतम 97 प्रतिशत तक पहुँच गई[3]।

दोनों दशाओं में खुला द्वार

अब तक के अनुवादक केवल एक ही दशा में चलते थे। वे केवल प्राचीन लेखों को आधुनिक अंग्रेज़ी में बदलते थे। मई 2026 में अमेरिका के दक्षिणी कैलिफ़ोर्निया विश्वविद्यालय के वांग झाओहुई ने इसकी उल्टी दशा का मार्ग खोला। उनके बनाए उपकरण का नाम TabletCraft है[4]।

TabletCraft अंग्रेज़ी को अक्कादी भाषा में भी अनुवादित करता है। यह अनुवाद के परिणाम को कीलाकार प्रतीकों में उकेरकर मट्टी की पट्टिका का चित्र भी तैयार कर देता है। इसके प्रशिक्षण में दोनों दशाओं के 1,16,000 वाक्य-युग्मों का उपयोग किया गया। इसमें 14,000 से अधिक प्रतीकों की मलिन सूची थी। परीक्षण में अक्कादी से अंग्रेज़ी का स्कोर 49.1 अंक रहा। अंग्रेज़ी से अक्कादी में 48.5 अंक मल्लि[4]।

इस उपकरण का उपयोग संग्रहालय के अनुभव उपकरणों और विश्वविद्यालय की कक्षाओं में किया गया। इसके निर्माता द्वारा स्पष्ट की गई सीमाएँ भी हैं। मशीन द्वारा गढ़ी गई अक्कादी भाषा को असली प्राचीन अभिलेख मानकर उद्धृत नहीं करना चाहिए। यह उपकरण नव-असीरियाई काल के शाही शिलालेखों पर सबसे अच्छा काम करता है[5]।

दोनों दशाओं के अनुवाद की अपनी उपयोगिता है। वदिवार्थी अपने वाक्यों को कीलाकार प्रतीकों में बदलकर देख सकते हैं। प्रतीकों को स्वयं बनाने से लिपिकी बनावट स्पष्ट समझ आती है। संग्रहालयों में दर्शक अपना नाम उकेर कर देखते हैं। प्राचीन लिपिकेवल देखने की वस्तु से बदलकर हाथ से आजमाने वाला साधन बन जाती है।

उपर्युक्त अध्ययनों में डेटा और प्रोग्राम दोनों को सार्वजनिक किया गया[1][5]। इसे सार्वजनिक करने से अन्य दल भी वही परीक्षण दोहरा सकते हैं। कोई गलती होने पर वह जल्दी पकड़ में आ जाती है। प्राचीन भाषाओं के अध्ययन में बहुत कम लोग जुड़े हैं। ऐसे में खुली जानकारी ही प्रगत की गति करती है।

अनुवादक का उपयोग करने का तरीका भी व्यवस्थित हो रहा है। वदिवान स्क्रीन पर मशीन का पहला मसौदा खोलते हैं। वे मूल पाठ और अनुवाद को आमने-सामने रखकर मिलाते हैं। संदेहास्पद शब्दों पर नशान लगाते हैं। जब ऐसे नशान एकत्र हो जाते हैं, तो वे अगली बार के शक्तिषण की सामग्री बन जाते हैं।

संदर्भ सामग्री

- [1] Gutherz, G., Gordin, S., Sáenz, L., Levy, O., & Berant, J. (2023). Translating Akkadian to English with neural machine translation. PNAS Nexus, 2(5), pgad096. <https://doi.org/10.1093/pnasnexus/pgad096>
- [2] Gordin, S., Sahala, A., Spencer, S., & Klein, S. (2025). EvaCun 2025 shared task: Lemmatization and token prediction in Akkadian and Sumerian using LLMs. In Proceedings of the Second Workshop on Ancient Language Processing (pp. 242-250). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2025.alp-1.33>
- [3] Gordin, S., Gutherz, G., Elazary, A., Romach, A., Jiménez, E., Berant, J., & Cohen, Y. (2020). Reading Akkadian cuneiform using natural language processing. PLoS ONE, 15(10), e0240511. <https://doi.org/10.1371/journal.pone.0240511>
- [4] Wang, Z. (2026, 5월 14일). TabletCraft: Bridging a 4,000-year cultural gap with bidirectional Akkadian NMT and cuneiform rendering. arXiv:2608.02609 (2026년 5월 14일 제출, 2026년 8월 공개). <https://arxiv.org/abs/2608.02609> (3 सितंबर 2026 को अभिगमिति)
- [5] Buyukyildirim, K. (2026, 8월 11일). TabletCraft opens a two-way AI translation channel for archaeologists. Arkeonews. <https://arkeonews.net/tabletcraft-opens-a-two-way-ai-translation-channel-for-archaeologists/> (3 सितंबर 2026 को अभिगमिति)
- [6] Che, W., Garces Arias, E., Niaz, A., Bender, A., & Jimenez, E. (2026, 6월 21일). Automated sign detection across the Electronic Babylonian Library: A large-scale dataset and end-to-end cuneiform OCR pipeline. arXiv:2606.22608. <https://arxiv.org/abs/2606.22608> (3 सितंबर 2026 को अभिगमिति)

6.3. उजागर हुआ सच

कानेश के व्यापारियों के पत्र

मध्य तुर्किय में काइसेरी के पास कुलतेपे नाम का एक टीला है। 4,000 वर्ष पूर्व इस स्थान का नाम कानेश था। असीरियाई व्यापारियों ने यहाँ अपना व्यापारिक केंद्र बनाया था। व्यापारियों के उस केंद्र समुदाय को कारुम कहा जाता था। कारुम का मूल अर्थ नदी किनारे का घाट होता है। माल उतारने की जगह आगे चलकर व्यापारियों का इलाका बन गई।

कानेश का कारुम असीरियाई व्यापारियों का एक स्वायत्त संगठन था। वे आपस में नियम तय करते थे और अपने विवाद नपिताते थे। वे स्थानीय राजा को कर चुकाते थे और बदले में सुरक्षा पाते थे। सामान गधों की पीठ पर लादकर भेजा जाता था। अशूर से कानेश तक का रास्ता कई हफ्तों का लंबा सफर था।

1948 से जारी खुदाई में मट्टी की पट्टियों का ढेर मिला गया। अब तक निकले दस्तावेजों की संख्या लगभग 23,500 है। यह आँकड़ा खुदाई का नेतृत्व कर रहे फ़िक्री कुलाकोउलू ने साझा किया है[1]। इनकी वषिय-वस्तु किसी राजा का गुणगान नहीं है। इनमें कर्ज़ के कागज़ात और माल-चालान शामिल हैं। माल-चालान भेजे गए सामान की सूची होती है। इनमें अदालती फ़ैसले और पारिवारिक चिट्ठियाँ भी हैं।

चिट्ठियों में तकाज़े और शक़ायते वैसी की वैसी बची हुई हैं। पूछा जाता है कि पैसा कब भेजा जाएगा। सामान देर से आने पर उलाहना दिया जाता है। अपना ख़्याल रखने की शुभकामनाएँ भेजी जाती हैं। बेटे को मन लगाकर पढ़ने की सीख भी दी जाती है। 4,000 साल पहले के इंसानों की चिंताएँ इन वाक्यों में आज भी मौजूद हैं।

इन दस्तावेजों से प्राचीन अर्थव्यवस्था का ढाँचा उजागर हुआ। व्यापारी अशूर से वस्त्र और टनि लाते थे। अनातोलिया में वे इनके बदले चाँदी और सोना प्राप्त करते थे। कर और चुंगी कैसे दी जाए, यह भी लिखा हुआ था। विवाद होने पर क्या नरिणय दिया गया, इसका विवरण भी दर्ज था।

काँसा बनाने के लिए टनि अनिवार्य धातु है। काँसा ताँबे और टनि को मिलाकर बनता है। अनातोलिया में टनि मलिना मुश्किल था। टनि को बहुत दूर पूरब से ढोकर लाना पड़ता था[8]। यही कमी लंबे व्यापारिक मार्ग बनने की असली ताक़त थी।

अशूर में रह गई महिलाओं की भूमिका भी उजागर हुई। व्यापारियों की पत्नियाँ और बेटियाँ घर पर कपड़े बुनकर भेजती थीं[2][8]। वे उत्पादन की मात्रा तय करती थीं और कीमतों पर मोलभाव करती थीं। अपने पतियों को लिखी चिट्ठियों में वे बाज़ार भाव का हिसाब भी रखती थीं। सेसिल मशिल ने ऐसे 300 से अधिक दस्तावेज़ एकत्र करके उनका अंग्रेज़ी में अनुवाद किया। मशिल फ़्रांसीसी राष्ट्रीय वैज्ञानिक अनुसंधान केंद्र और हैम्बर्ग विश्वविद्यालय में कार्यरत हैं[2]।

लफ़ाफ़ा खोले बना पढ़ना

प्राचीन व्यापारी महत्वपूर्ण मट्टी की पट्टिकाओं को मट्टी के लफ़ाफ़े में डालकर सील कर देते थे। भीतर का लेख पढ़ने के लिए लफ़ाफ़े को तोड़ना पड़ता था। क़ति लफ़ाफ़े पर भी मुहरों की छाप होती है। लफ़ाफ़ा तोड़ने पर वह जानकारी नष्ट हो जाती है।

लफ़ाफ़ा जालसाज़ी रोकने की एक व्यवस्था थी। अंदर की पट्टिका पर अनुबंध की बातें लिखी जाती थीं। उस पर मट्टी की परत चढ़ाकर लफ़ाफ़ा बनाया जाता था। लफ़ाफ़े के ऊपर भी वही बात दोबारा लिखी जाती थी। गवाह उस पर बेलनाकार मुहर घुमाकर छाप लगाते थे। यदि बाहर और भीतर का विवरण अलग होता, तो बदलाव की बात तुरंत पकड़ में आ जाती थी।

लफ़ाफ़े को नष्ट किए बिना भीतर का पाठ पढ़ने के तरीके की ज़रूरत थी। जून 2026 में काइसेरी पुरातत्व संग्रहालय ने इसका समाधान दिखाया। लफ़ाफ़े में बंद एक मट्टी की पट्टिका को बिना तोड़े सफलतापूर्वक पढ़ लिया गया। यह फ़्रांसीसी राष्ट्रीय वैज्ञानिक अनुसंधान केंद्र और हैम्बर्ग विश्वविद्यालय का संयुक्त कार्य था। किसी वस्तु को काटे बिना भीतर झाँकने वाली CT स्कैन जैसी तकनीक का इस्तेमाल किया गया। अंदर गेहूँ और जौ के सौदे का अनुबंध लिखा था। कुलाकोउलू ने बताया कि यह स्कैनिंग जौ का एक दाना तक दिखा देती है[1]। इस प्रकार प्राप्त अक्षरों की छवियाँ अनुवादक के लिए इनपुट सामग्री बनती हैं। जब पढ़े जाने वाले दस्तावेज़ बढ़ेंगे, तभी मशीन को भी सीखने के लिए सामग्री मिलेगी।

50,000 डॉलर की अनुवाद प्रतियोगिता

एक अनुवाद प्रतियोगिता भी आयोजित हुई। Deep Past Initiative नामक संगठन ने इस प्रतियोगिता का आयोजन किया[3]। यह प्रतियोगिता एआई प्रतियोगिता मंच Kaggle पर आयोजित की गई[4]। कार्य कानेश के व्यापारिक दस्तावेजों का अंग्रेज़ी में अनुवाद करना था[3]। प्रतियोगिता 16 दिसंबर 2025 को शुरू हुई। आवेदन की अंतिम तिथि 23 मार्च 2026 थी। पुरस्कार राशि 50,000 डॉलर थी। संचालन का खर्च एक वित्तीय कंपनी ने वहन किया[7]।

प्रतियोगिता का डेटा कोई भी डाउनलोड कर सके, इसके लिए सार्वजनिक किया गया था[4]। खुला डेटा प्रतियोगिता समाप्त होने के बाद भी उपलब्ध रहता है। अगले शोधकर्ता उसी डेटा पर आगे का काम शुरू करते हैं। इस व्यवस्था में प्रत्येक प्रतियोगिता अपने पीछे उपयोगी डेटा छोड़ जाती है।

कुलतेपे के दस्तावेजों में से भी कई का अनुवाद होना अभी बाकी है[3]। खुदाई आज भी हर वर्ष जारी है[1]। जब नई पट्टिकाएँ मिलती हैं, तो वे सूची में जुड़ जाती हैं। यदि पिढ़ने वाले लोग नहीं बढ़ेंगे, तो केवल सूची लंबी होती जाएगी। यही पर मशीन की मदद की ज़रूरत पड़ती है।

यह स्थिति केवल एक ही स्थल तक सीमिति नहीं है। विश्व के कई संग्रहालयों में ऐसे दस्तावेजों के ढेर लगे हैं। वहाँ भी ऐसे बक्से रखे हैं जनिहें अब तक खोला और पढ़ा नहीं गया है। तयशुदा प्रारूप वाले दस्तावेज हर जगह बहुतायत में मिलते हैं। अनाज के खाते और कर रजिस्टर इसी तरह के हैं। जो तरीका कानेश में सफल रहा, उसे दूसरे अभिलेखागारों में भी आजमाया जा सकता है।

प्रतियोगिता के विवरण में दस्तावेजों में उल्लिखित वास्तविक व्यक्तियों के नामों को आगे रखा गया। इनमें अशूर-नादा और अशूर-इद्दी जैसे व्यापारियों के नाम थे। इश्तार-लामास्सी और एन्नम-अशूर के नाम भी साथ जोड़े गए[3]। इस प्रकार 4,000 वर्ष पहले के एक व्यापारी परिवार के नाम प्रतियोगिता का मुख्य विषय बन गए।

प्रतियोगिता के तरीके के कई फायदे हैं। केवल कुछ प्रयोगशालाएँ ही नहीं, बल्कि दुनिया भर के डेवलपर्स एक ही समस्या को हल करने में जुट जाते हैं। एक ही पैमाने से नतीजों को नापा जाता है, जिससे तुरंत पता चल जाता है कि कौन-सा तरीका श्रेष्ठ है। पुरस्कार राशिलोगों को एक साथ जोड़ने की ताकत बनती है।

दस्तावेजों का यह समूह मशीन अनुवाद के लिए बेहद अनुकूल सामग्री है। इनका प्रारूप निश्चित है और वाक्य छोटे हैं। एक जैसी अभिव्यक्तियाँ कई दस्तावेजों में बार-बार आती हैं। मशीन दोहराए जाने वाले नियमों को बहुत अच्छे से पकड़ लेती है। विद्वान पहले ही बहुत से दस्तावेजों का अनुवाद कर चुके हैं। वे अनुवाद मशीन के लिए अध्ययन सामग्री बन जाते हैं।

जो मशीन नहीं कर सकती

उपलब्धियाँ भले ही बढ़ी हों, लेकिन सामने खड़ी दीवारें भी स्पष्ट हैं। पहली रुकावट डेटा की कमी है। आधुनिक भाषाओं के अनुवादक करोड़ों वाक्यों का अध्ययन करके बड़े होते हैं। अक्कादी भाषा में अनुवादित वाक्यों के केवल कुछ ही हजार जोड़े उपलब्ध हैं[5]। जब डेटा कम होता है, तो दुर्लभ स्थानों और लोगों के नामों में भूल होना स्वाभाविक है।

दूसरी रुकावट प्रतीकों का बहु-अर्थी होना है। एक ही प्रतीक ध्वनिका सूचक है, पूरा शब्द है या कोई वर्गीकरण चिह्न, इसे समझना कठिन होता है[6]। यदि आगे-पीछे का संदर्भ बहुत छोटा हो, तो मशीन भी उलझ जाती है।

तीसरी रुकावट मशीन के फैसले करने का तरीका है। मशीन इस बात का हिसाब लगाती है कि किस शब्द के आने की संभावना सबसे अधिक है। वह यह नहीं परख सकती कि वाक्य ऐतिहासिक रूप से सही है या नहीं। इसलिए अंतिम चरण में विद्वान की जाँच बेहद ज़रूरी होती है[5]। जब मशीन और इंसान मिलकर काम करते हैं, तभी अनुवाद भरोसेमंद बनता है।

काम बाँटकर करने का यह तरीका अब स्थापित हो रहा है। मशीन हजारों पट्टिकाओं को सरसरी तौर पर देखकर उनका सामान्य अर्थ निकालती है। विद्वान उनमें से महत्वपूर्ण दस्तावेज छोट लेते हैं। केवल चुने हुए दस्तावेजों को ही वे स्वयं बारीकी से दोबारा पढ़ते हैं। इस तरह पढ़े जाने वाले दस्तावेजों की संख्या में भारी बढ़ोतरी होती है।

अनुवाद पर स्पष्ट लेबल लगाया जाना ज़रूरी है। कौन-सा अंश मशीन ने तैयार किया, यह दर्ज किया जाता है। जहाँ इंसान ने सुधार किया, उसे भी चिह्नित किया जाता है। तभी बाद में आने वाले शोधकर्ता साक्ष्यों की दोबारा पुष्टि कर सकते हैं। ऐतिहासिक अनुसंधान साक्ष्यों की पड़ताल से ही शुरू होता है।

अनुवाद में नाम सबसे जटिल हिसा होते हैं। व्यक्तियों और स्थानों के नामों का अर्थ अनुवादित नहीं किया जा सकता। उनके उच्चारण को ज्यों का त्यों लिप्यंतरित करना होता है। लेकिन एक ही नाम अलग-अलग दस्तावेजों में भिन्न रूप में भी लिखा मिलता है। मशीन नए नाम को किसी असंगत शब्द में बदल सकती है। इसलिए नामों की एक अलग सूची बनाकर मशीन को सखाई जाती है।

फिर भी दशिया बिल्कुल स्पष्ट है। पहले एक विद्वान अपने पूरे जीवन में केवल कुछ सौ पट्टिकाएँ ही पढ़ पाता था। अब मशीन पहले कच्चा मसौदा तैयार कर देती है। विद्वान उस मसौदे को सुधारने में अपना समय लगाते हैं। बक्सों में सोई 4,000 साल पुरानी रसीदे अपनी बारी की प्रतीक्षा कर रही हैं। अब राजाओं की नहीं, आम लोगों की आवाज़ें सुनाई देने लगी हैं।

कानेश के व्यापारियों ने यह नहीं सोचा था कि उनकी चिट्ठियाँ इतने लंबे समय तक बची रहेगी। उन्होंने तो केवल उस दनि का हिसाब-किताब लिखा था। वही हिसाब आज एक युग की अर्थव्यवस्था को दिखाता है। जो शेष रह जाता है, वह आत्मप्रशंसा नहीं बल्कि अभिलेख है।

संदर्भ सामग्री

- [1] Türkiye Today. (2026, 6월 13일). 4,000-year-old clay tablet sealed in mud envelope goes on display in central Türkiye. Türkiye Today. <https://www.turkiyetoday.com/culture/4000-year-old-clay-tablet-sealed-in-mud-envelope-goes-on-display-in-central-turkiye-3221841> (3 सितंबर 2026 को अभिगमिति)
- [2] Michel, C. (2020). *Women of Assur and Kanesh: Texts from the Archives of Assyrian Merchants (Writings from the Ancient World 42)*. SBL Press.
- [3] Deep Past Initiative. (2026). Deep Past Challenge: Introduction. [deeppast.org. https://www.deeppast.org/challenge/intro/](https://www.deeppast.org/challenge/intro/) (3 सितंबर 2026 को अभिगमिति)
- [4] Kaggle. (2026). Deep Past Challenge: Translate Akkadian to English. [kaggle.com. https://www.kaggle.com/competitions/deep-past-initiative-machine-translation](https://www.kaggle.com/competitions/deep-past-initiative-machine-translation) (3 सितंबर 2026 को अभिगमिति)
- [5] Guthertz, G., Gordin, S., Sáenz, L., Levy, O., & Berant, J. (2023). Translating Akkadian to English with neural machine translation. *PNAS Nexus*, 2(5), pgad096. <https://doi.org/10.1093/pnasnexus/pgad096>
- [6] Gordin, S., Guthertz, G., Elazary, A., Romach, A., Jiménez, E., Berant, J., & Cohen, Y. (2020). Reading Akkadian cuneiform using natural language processing. *PLoS ONE*, 15(10), e0240511. <https://doi.org/10.1371/journal.pone.0240511>
- [7] Deep Past Initiative. (2025, 12월 16일). Deep Past Initiative partners with Kaggle on prize challenge to translate 4,000-year-old tablets. EIN Presswire. <https://www.einpresswire.com/article/875885564/deep-past-initiative-partners-with-kaggle-on-prize-challenge-to-translate-4-000-year-old-tablets> (3 सितंबर 2026 को अभिगमिति)
- [8] Bouscaren, D. (2023). Assyrian women of letters. *Archaeology*, 76(6). <https://archaeology.org/issues/november-december-2023/features/assyrian-women-of-letters/> (3 सितंबर 2026 को अभिगमिति)

अध्याय 7. कछुए के उदर-कवच में सोया साम्राज्य का भवष्य: ओरेकल बोन लपि और 3D आभासी संयोजन

7.1. केस फाइल

दवा की दुकान में मल्लि 3,000 वर्ष पुराने अक्षर

1899 में चंगि राजवंश के वदिवान वांग यरिंग ने एक प्राचीन अक्षर को पहचाना। इस घटना से जुड़ी एक पुरानी कहानी प्रचलित है। कहानी यह है कि वांग यरिंग बीमार पड़े और उन्हें 'ड्रैगन की हड्डी' (लोगगु) नाम की औषधि मिली। ड्रैगन की हड्डी प्राचीन जानवरों की हड्डियों को पीसकर बनाई गई चूर्ण औषधि थी। प्राचीन लोग इन हड्डियों को ड्रैगन की हड्डियाँ मानते थे। कहा जाता है कि उन्हें जो हड्डी मिली, वह अभी पसी नहीं थी। उन्होंने उसकी सतह पर एक अनोखा पैटर्न देखा[1]।

यह कहानी प्रमाणित तथ्य नहीं बन पाई है। चीनी अकादमिक जगत का मानना है कि इसका समर्थन करने वाला कोई अभिलेख नहीं है। प्राचीन लपि विशेषज्ञ ली शुएकनि ने 2007 के एक लेख में इस कहानी का खंडन किया। उनका कहना था कि औषधि के रूप में बेची जाने वाली हड्डियाँ पहले ही टूट चुकी होती थीं, इसलिए उन पर अक्षर देखना संभव नहीं था। वास्तव में एक पुरावस्तु व्यापारी ने आन्यांग से नकिली हड्डियों को खरीदकर पहुँचाया था। माना जाता है कि वे हड्डियाँ पुरालेखविद् वांग यरिंग के पास पहुँचीं और उन्होंने उनका मूल्यांकन किया[2]।

वांग यरिंग का अगले ही वर्ष नधिन हो गया, इसलिए वे हड्डियों के स्रोत का पता नहीं लगा सके। उनके मतिर लयू ई ने 1903 में इन अक्षरों को संकलित कर पहली पुस्तक प्रकाशित की। 1908 में वदिवान लुओ झेनयु ने हड्डियों के स्रोत का पता लगाया[1]। वे हड्डियाँ हेनान प्रांत के आन्यांग के खेतों से निकल रही थीं।

आन्यांग की वही भूमि यिनक्सू है। यिनक्सू शांग राजवंश की अंतिम राजधानी का स्थल था। शांग राजवंश ने 13वीं शताब्दी ईसा पूर्व से 11वीं शताब्दी ईसा पूर्व तक यहाँ शासन किया था[3]। औपचारिक उत्खनन 1928 में शुरू हुआ। तब तक किसानों द्वारा खोदी गई हड्डियाँ औषधि के रूप में बिकती रही। कभी-कभी अक्षरों वाली सतह को खुरचकर बेच दिया जाता था[1]। 3,000 वर्षों तक बचा रहा यह अभिलेख औषधि बनकर समाप्त होने ही वाला था।

वह देश जहाँ राजा देवताओं से प्रश्न पूछता था

यिनक्सू की हड्डियों पर उकेरे गए अक्षरों को ओरेकल बोन लपि कहा जाता है। पूर्वी एशिया में ओरेकल बोन लपि से पुराना कोई लिखित अभिलेख ज्ञात नहीं है[3]।

शांग राजवंश के लोग किसी बड़े काम से पहले देवताओं से पूछते थे। वे पूछते थे कि क्या बारिश होगी। वे पूछते थे कि क्या फसल अच्छी होगी। वे पूछते थे कि क्या युद्ध में जाना चाहिए। वे रानी के बच्चे को जन्म देने के दिन के बारे में भी पूछते थे। देवताओं की इच्छा जानने के इस अनुष्ठान को भवष्यवाणी अनुष्ठान कहा जाता है। भवष्यवाणी अनुष्ठान राजपरिवार द्वारा सीधे प्रबंधित किया जाने वाला एक राष्ट्रीय कार्य था[3]।

शांग राजवंश की रानी फू हाओ सेना का नेतृत्व करने वाली सेनापति भी थीं। फू हाओ का नाम ओरेकल बोन लपि के अभिलेखों में कई बार आता है। 1976 में यिनक्सू में फू हाओ का मकबरा पूरी तरह सुरक्षित स्थिति में खोजा गया[4]। मकबरे में मल्लि कांस्य और जेड की वस्तुएँ उनकी प्रतीक्षा को दर्शाती थीं। इस प्रकार ओरेकल बोन लपि मकबरे के साथ मलिकर प्राचीन लोगों के जीवन को उजागर करती है।

ओरेकल बोन लपि में लिखे प्रश्न केवल राजकीय मामलों तक सीमित नहीं थे। उनमें राजा के दाँत दर्द की चिंता करने वाले प्रश्न भी थे। उनमें यह पूछने वाले प्रश्न भी थे कि क्या शक्तिार पर जाना चाहिए। इसलिए ये हड्डियाँ शांग राजवंश के लोगों के दैनिक जीवन को दर्शाने वाले अभिलेख बन जाती हैं[3]।

आग से तपी हड्डियाँ, चाकू से उकेरे गए अक्षर

शांग राजवंश के लोग भवष्यवाणी अनुष्ठान के लिए दो सामग्रियों का उपयोग करते थे। एक कछुए का चपटा उदर-कवच था। दूसरा बैल की चौड़ी कंधे की हड्डी थी। इन दोनों सामग्रियों को मिलाकर 'ओरेकल बोन' (कवच-अस्थि) कहा जाता है[3]।

तैयारी की प्रक्रिया बहुत व्यवस्थित थी। लपिकि हड्डियों की बाहरी सतह को चिकना बनाते थे। फिर वे पछिली सतह पर खाँचों की कतारें तराशते थे। तराशी गई खाँचों में वे लाल-गर्म छड़ लगाते थे। गर्मी पाकर हड्डी की बाहरी सतह चटक जाती थी। राजा और ज्योतिषी दरारों की दिशा और कोण की जाँच करते थे। और वे उन दरारों को देवताओं के उत्तर के रूप में पढ़ते थे[3]।

पढ़ने का काम पूरा होने पर लपिकि चाकू उठाते थे। वे भवषियवाणी अनुष्ठान की प्रक्रिया को हड्डी की सतह पर एक-एक रेखा करके उकेरते थे। यही उकेरे गए लेख ओरेकल बोन लपि हैं।

अभिलेख एक नश्चिती कर्म का पालन करते थे। पहले भवषियवाणी की तथि और प्रभारी का नाम लिखा जाता था। इसके बाद देवताओं से पूछा गया प्रश्न लिखा जाता था। फिर राजा द्वारा दिया गया शुभ-अशुभ का नरिणय लिखा जाता था। अंत में वास्तव में जो घटित हुआ, उसे लिखा जाता था। यह चार भागों को क्रमानुसार भरने की वधि थी। हालाँकि चारों भाग पूरी तरह सुरक्षित हों, ऐसे अभिलेख दुर्लभ हैं। हड्डी टूटने के कारण एक या दो भाग गायब भी हो गए थे। कुछ मामलों में लपिकि ने शुरू से ही खाली जगह छोड़ दी थी[3]। फिर भी लिखने का कर्म हमेशा सुसंगत था।

यह नियमिता बाद में बहुत मददगार साबित होती है। क्योंकि प्रारूप नश्चिती होने पर छूटे हुए स्थानों का दायरा सीमिति किया जा सकता है। यदि पहले भाग में तथि है, तो अगले भाग में प्रश्न आता है। कंप्यूटर भी इस कर्म को सीखकर छूटे हुए अक्षरों का अनुमान लगाता है[3]।

तथि लिखने की वधि भी नश्चिती थी। शांग राजवंश दस और बारह प्रतीकों के संयोजन से दनि गनिता था। यह चक्र हर साठ दनि में एक चक्र पूरा करता था। हड्डी पर तथि के केवल दो प्रतीक बचे हों, तब भी कर्म की गणना की जा सकती है। टूटे हुए अभिलेखों की अवधि को संकीर्ण करने का सुराग यही से मिलता है[3]।

1,50,000 टुकड़ों में बखिरा अभिलेख

शांग राजवंश के पतन के बाद यनिक्सू मटिटी में दब गया। उस लंबे कालखंड में हड्डियाँ अक्षुण्ण नहीं रह सकीं। मटिटी के भार ने पतले उदर-कवच को दबाकर टुकड़े-टुकड़े कर दिया। पानी और सूक्ष्मजीवों ने हड्डियों की बनावट को नष्ट कर दिया। मूल रूप से एक पूरा कवच दर्जनो टुकड़ों में वभाजित हो गया।

अब तक ज्ञात ओरेकल बोन के टुकड़ों की संख्या लगभग 1,50,000 है[5]। ये पुरावशेष एक ही स्थान पर एकत्रित नहीं हैं। 19वीं सदी के अंत में खोजे जाने के बाद इनका एक बड़ा हसिसा पुरावस्तु बाज़ार में चला गया[1]। इसलिए एक ही हड्डी से टूटे हुए टुकड़े अलग-अलग देशों के संग्रहालयों में पहुँच गए[5]। कुछ पुरावशेष चीन के आन्यांग और बीजिंग में हैं। कुछ पुरावशेष यूरोप और दक्षिण कोरिया के संस्थानों के पास भी हैं[6]।

बखिरे हुए टुकड़ों को जोड़कर फिर से एक मूल रूप देने के कार्य को संयोजन कहा जाता है। संयोजन का कार्य लंबे समय तक केवल मानवीय आँखों पर ही नरिभर रहा। वदिवान आवरधक लेस से हजारों श्वेत-श्याम रगड़-छापों की जाँच करते थे। रगड़-छाप हड्डी पर कागज़ रखकर और स्याही थपथपाकर तैयार की गई प्रतिलिपि होती है। यह सकि के पर कागज़ रखकर पेसलि रगड़कर उसका पैटर्न उभारने जैसा है। वदिवान हड्डी की बनावट और टूटे हुए कनिारों के घुमावों की तुलना करते थे। वे यह भी एक-एक करके देखते थे कि किकी हुई रेखाओं के सरि आपस में जुड़ते हैं या नहीं[5]।

यह तरीका धीमा और कठिन था। एक जोड़ा खोजने में कभी-कभी कई साल लग जाते थे। जनि टुकड़ों की जोड़ी नहीं मिली, वे बक्सों में ही पड़े रहे। जोड़ी न मिलने पर वाक्य भी अधूरे ही रह जाते हैं। ऐसे अभिलेखों का ढेर लग गया जनिमें केवल प्रश्न बचा था और परिणाम गायब था।

कार्य के पैमाने पर वचिार करने पर कठिनाई और अधकि स्पष्ट हो जाती है। यदि 1,50,000 टुकड़े हों, तो दो चुनने के संभावति संयोजनों की संख्या 10 अरब जोड़ों से भी अधकि होती है। यदि कोई व्यक्ति प्रतिदिन सौ जोड़ों की तुलना करे, तब भी जीवनभर में उन्हें पूरा नहीं देख सकता। इसके अतिरिक्त, टुकड़े कई देशों में बखिरे हुए हैं[5]। अक्सर तो दो टुकड़ों को साथ रखकर देखने का अवसर तक नहीं मिल पाता था।

3,000 अक्षर अब भी मौन हैं

टुकड़ों को पूरी तरह जोड़ लेने के बाद भी समस्या बनी रहती है। क्योंकि स्वयं अक्षरों को पढ़ा नहीं जा सकता। ओरेकल बोन लपि में कुल पहचाने गए अलग-अलग अक्षरों की संख्या लगभग 4,500 है। कति अकादमिक जगत केवल लगभग 1,500 अक्षरों के अर्थ पर ही सहमत हो पाया है। शेष लगभग 3,000 अक्षरों का अर्थ अभी तक अज्ञात है[7]।

पढ़ने में कठिनाई के कई कारण हैं। चाकू के नशान घसि जाने से रूपरेखा धुंधली हो गई। हड्डी टूटने से अक्षर का आधा हसिसा गायब हो गया। हर लपिकि की चाकू पकड़ने की आदत अलग थी, इसलिए उन्होंने एक ही अक्षर को अलग-अलग तरह से उकेरा। फलस्वरूप समान अर्थ वाले कति भिन्न रूप वाले कई अक्षर बन गए। ऐसे अक्षरों को भिन्न-रूपी अक्षर कहा जाता है[3]। आधुनिकि चीनी शब्दकोश इन 3,000 वर्ष पुराने रूपों की व्याख्या नहीं कर सकते।

यह मौन शांग राजवंश के इतिहास में भी रकिता छोड़ देता है। यदि अनुष्ठानों के नाम न पढ़े जा सकें, तो अनुष्ठानों के प्रकार नहीं गनि जा सकते। यदि स्थानों के नाम न पढ़े जा सकें, तो राज्य की सीमाओं का नक्शा नहीं बनाया जा सकता। यदि व्यक्तियों के नाम न पढ़े जा सकें, तो यह नहीं जाना जा सकता कि किसने क्या ज़िम्मेदारी संभाली थी। एक अक्षर का मौन इतिहास के एक पूरे दृश्य के मौन में बदल जाता है।

अज्ञात अर्थ वाले 3,000 अक्षर और अपनी जोड़ी खो चुके हजारों टुकड़े शेष रह गए। इस मौन को तोड़ने के लिए एआई ने यनिक्सू में प्रवेश किया।

संदर्भ सामग्री

- [1] Mark, E. (2016, 2월 26일). Oracle bones. World History Encyclopedia. https://www.worldhistory.org/Oracle_Bones/ (3 सितंबर 2026 को अभिगमिति)
- [2] 姚小鸥. (2020, 7월 16일). 谁是甲骨文的最早发现者. 澎湃新闻. https://www.thepaper.cn/newsDetail_forward_8297008 (3 सितंबर 2026 को अभिगमिति)
- [3] Chen, Z., Hua, W., Li, J., Zhu, Y., Zhi, X., Liu, Z., Chen, T., Zhang, W., & Zhai, G. (2026). Oracle bone inscriptions information processing: a comprehensive survey. npj Heritage Science, 14, 220. <https://doi.org/10.1038/s40494-026-02511-w> (3 सितंबर 2026 को अभिगमिति)
- [4] Smarthistory. War and sacrifice: The tomb of Fu Hao. Smarthistory. <https://smarthistory.org/tomb-of-fu-hao/> (3 सितंबर 2026 को अभिगमिति)
- [5] Zhang, C., Wang, B., Chen, K., Zong, R., Mo, B., Men, Y., Alpanidis, G., Chen, S., & Zhang, X. (2022). Data-driven oracle bone rejoining: A dataset and practical self-supervised learning scheme. In Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (pp. 4482-4492). ACM. <https://doi.org/10.1145/3534678.3539050>
- [6] 梁文艳. (2026, 5월 26일). 以科技赋能, 让甲骨文"活"起来. 中国青年报. https://zqb.cyol.com/pc/content/202605/26/content_426168.html (3 सितंबर 2026 को अभिगमिति)
- [7] Liu, Y., Guan, H., Wang, P., Wang, X., Wan, J., Zhang, K., Zheng, H., Liu, X., Kuang, Z., Yang, H., Li, B., Liu, Y., Jin, L., & Bai, X. (2026). AlphaOracle: Oracle bone script decipherment via human-workflow-inspired deep learning. The Innovation, 7(11), 101462. <https://doi.org/10.1016/j.xinn.2026.101462>

7.2. वैज्ञानिक जाँच दल की पड़ताल: ओरेकल बोन समरूपति मल्टीमॉडल मॉडल और 3D आभासी संयोजन

बखिरी हुई सामग्री को एक स्क्रीन पर समेटना

जाँच सामग्री एकत्र करने से शुरू होती है। ओरेकल बोन अनुसंधान में भी ऐसा ही था। आन्यांग नॉर्मल यूनिवर्सिटी और चीनी सामाजिक विज्ञान अकादमी के शोधकर्ताओं ने 2019 में 'यनिकी वेनयुआन' नामक डेटा प्लेटफॉर्म शुरू किया। यनिकी वेनयुआन ओरेकल बोन की रगड़-छापों, अक्षरों की जानकारी और शोध-पत्रों को एक जगह एकत्र कर मुफ्त उपलब्ध कराता है[1]।

यह प्लेटफॉर्म केवल तस्वीरें ही अपलोड नहीं करता। यह ओरेकल बोन के टुकड़ों की उच्च-रिज़ॉल्यूशन तस्वीरें लेता है और उनके 3D मॉडल बनाकर अपलोड करता है। 3D मॉडल किसी वस्तु को विभिन्न दशाओं से स्कैन करके कंप्यूटर में त्रि-आयामी रूप में पुनर्जीवित करने वाला डेटा है। स्क्रीन घुमाने पर हड्डी का अग्र भाग, पृष्ठ भाग और टूटी हुई कनारे वाली सतह देखी जा सकती है। शोधकर्ता मल्टीमीटर स्तर की सटीकता के साथ ये मॉडल बनाते हैं। चाकू से उकड़े गए उथले नशानों की गहराई भी इस डेटा में दर्ज होती है[1]।

वर्द्धित चले गए ओरेकल बोन पुरावशेषों को वापस पाने का काम भी इस प्लेटफॉर्म से जुड़ा। शोधकर्ताओं ने वास्तविक वस्तु के स्थान पर 3D डेटा लाने का तरीका चुना। जर्मनी, दक्षिण कोरिया और फ्रांस जैसे देशों के संग्रहालयों के साथ सहयोग करके उन्होंने 900 से अधिक पुरावशेषों का डेटा एकत्र किया[1]। वास्तविक वस्तुएँ मूल संग्रहालयों में ही रही, केवल उनका स्वरूप आन्यांग लौट आया।

अक्टूबर 2025 में इस प्लेटफॉर्म में 'यनिकी शगिझी' नामक एक उपकरण जोड़ा गया। यह उपकरण अक्षरों को पहचानता है, रगड़-छापों की तुलना करता है और संबंधित साहित्य खोजकर देता है[1]। डेटा का भंडारण करने वाला गोदाम अब प्रश्नों के उत्तर देने वाले सहायक में बदल गया।

डेटा कितना संचित हुआ है, इसकी पुष्टि आँकड़ों से भी होती है। 2026 में प्रकाशित एक व्यापक समीक्षा शोध-पत्र ओरेकल बोन से संबंधित 36 प्रकार के सार्वजनिक डेटासेट प्रस्तुत करता है। इनमें अक्षर पहचान के लिए डेटासेट भी हैं और संयोजन अनुसंधान के लिए भी[2]।

विभिन्न प्रकार के डेटा का एक साथ उपयोग करने की पद्धतिको मल्टीमॉडल कहा जाता है। ओरेकल बोन अनुसंधान में इसका अर्थ तीन प्रकार के डेटा को एक साथ देखना है। पहला स्याही से ली गई श्वेत-श्याम रगड़-छाप है। दूसरा रंग और बनावट को दर्शाने वाली रंगीन तस्वीर है। और तीसरा गहराई की जानकारी से युक्त 3D स्कैन डेटा है। रगड़-छाप में रेखाएँ स्पष्ट होती हैं, कति गहराई गायब हो जाती है। 3D डेटा गहराई को दर्शाता है, कति रेखाएँ धुंधली दिखाई देती हैं। तीनों को एक साथ रखकर देखने से वे एक-दूसरे की कमियों को पूरा कर देते हैं[2]।

अक्षर पढ़ने के मानवीय क्रम का अनुकरण करना

2026 में जारी किया गया AlphaOracle इस डेटा पर काम करने वाला एआई है। यह मॉडल हुआझोग वजिज्ञान एवं प्रौद्योगिकी विश्वविद्यालय तथा आन्यांग नॉर्मल यूनिवर्सिटी आदि में मलिकर वकिसति किया। इसके डिजाइन का सिद्धांत मानव वदिवानों की कार्य-प्रक्रिया का हबहू पालन करना है[3]।

मानव वदिवान चार चरणों से गुजरते हैं। पहले वे रगड़-छाप की जाँच करके रेखाओं को व्यवस्थिति करते हैं। इसके बाद वे अक्षर की बनावट का विश्लेषण करते हैं। फिर वे उन अन्य वाक्यों को खोजते हैं जहाँ वही अक्षर आता है। अंत में वे प्राचीन साहित्य के ज्ञान से जाँचते हैं कि अर्थ सही बैठता है या नहीं। AlphaOracle इन चारों चरणों के लिए अलग-अलग घटकों में वभिजति है[3]।

इस संरचना का एक कारण है। वदिवान उस मशीन पर आसानी से वशिवास नहीं कर सकते जो केवल सीधे उत्तर दे देती है। चरणों को वभिजति करने पर यह देखा जा सकता है कि किस चरण में नषिकर्ष अलग हुआ। वदिवान उस चरण पर वापस जाकर मशीन के साथ अपने वचारों का मलिन कर सकते हैं[3]।

इस मॉडल द्वारा पढ़े गए डेटा की मात्रा एक व्यक्ति द्वारा जीवनभर में देखे जा सकने वाले डेटा से अधिक है। इसने 73,883 रगड़-छापों को संसाधित किया। साथ ही 45,364 स्कैन डेटा और 23,755 शोध-पत्रों का भी विश्लेषण किया। शोधकर्ताओं ने 86 वशिषज्जों से इस उपकरण का उपयोग करवाया। परणामस्वरूप एक अक्षर का विश्लेषण करने में लगने वाला समय 64 प्रतशित घट गया। 79 प्रतशित प्रतभिगयिों ने इस उपकरण को उपयोगी बताया[3]।

अक्षरों को घटकों में वभिजति करके देखना

अन्य शोधकर्ताओं ने अक्षरों को पूरे रूप में याद न करने का तरीका चुना। ओरेकल बोन लपि के अक्षर बार-बार दोहराए जाने वाले चतिरात्मक घटकों से बने हैं[4]। हाथ, पैर और बर्तन जैसे आकारों के घटक कई अलग-अलग अक्षरों में प्रयुक्त होते हैं। घटकों का अर्थ समझकर पहली बार देखे गए अक्षर का भी अनुमान लगाया जा सकता है।

इसी वचार के साथ अप्रैल 2026 में OB-Radix डेटासेट प्रकाशित किया गया। OB-Radix में 934 अलग-अलग अक्षर शामिल हैं। उन अक्षरों को बनाने वाले घटक चतिर 1,853 हैं। घटकों के प्रकार 478 हैं। प्रत्येक घटक के साथ वशिषज्जों द्वारा सत्यापित व्याख्या जुड़ी हुई है[4]।

एआई किसी अपरचिति अक्षर को देखते ही पहले उसे घटकों में तोड़ता है। इसके बाद वह जाँचता है कि घटक किस संबंध में स्थिति है। दो घटक ऊपर-नीचे रखे हैं या अगल-बगल, इससे उनका अर्थ बदल जाता है। मॉडल इस संबंध की भी जाँच करके अर्थ का सुझाव देता है[4]।

उसी वर्ष जून में OracleAnalyzer नामक एक मॉडल भी सामने आया। यह मॉडल 3 अरब पैरामीटर वाले एक छोटे भाषा मॉडल को ओरेकल बोन डेटा के साथ पुनः प्रशिक्षित करके बनाया गया था। शोधकर्ताओं ने नए ओरेकल बोन तर्क डेटा और प्राथमिकता डेटा तैयार करके जोड़े। प्राथमिकता डेटा वह डेटा है जिसमें मनुष्यों द्वारा बेहतर चुने गए उत्तरों को एकत्रित किया जाता है। परणामस्वरूप इस छोटे मॉडल ने बहुत बड़े मॉडलों की तुलना में बेहतर प्रदर्शन किया[5]। इसका अर्थ है कि यदि डेटा का चयन सही हो, तो केवल आकार के बल पर निर्भर रहने की आवश्यकता नहीं होती।

धुंधली रेखाओं को पुनर्निर्मित करने वाला मॉडल

पढ़ने से पहले अक्षरों को दिखाई देने योग्य बनाने पर भी शोध चल रहा है। 2024 में प्रस्तुत किया गया OBSD ऐसा ही एक उपकरण है। यह उपकरण नई छवियाँ उत्पन्न करने वाले डिफ्यूज़न मॉडल का उपयोग करता है[6]। डिफ्यूज़न मॉडल वह तकनीक है जो धुंधले धब्बों से धीरे-धीरे एक स्पष्ट छवि तैयार करती है।

OBSD घिसी-पट्टी और धुंधली ओरेकल बोन अक्षर छवियों को इनपुट के रूप में लेता है। और यह छवियों के रूप में सुझाव देता है कि वह अक्षर किस आधुनिक चीनी अक्षर से संबंधित है। इस शोध को 2024 के एक अंतरराष्ट्रीय सम्मेलन में सर्वश्रेष्ठ शोध-पत्र के रूप में चुना गया[6]। अक्षर पढ़ने को छविनिर्माण की समस्या में बदलने के वचार को व्यापक मान्यता मिली।

टूटे हुए टुकड़ों के कनारों को मलिता कंप्यूटर

अक्षर पढ़ने के साथ-साथ टुकड़ों को जोड़ने पर भी शोध आगे बढ़ा। 2022 में जारी किया गया OB-Rejoin संयोजन अनुसंधान के लिए एक डेटासेट है। इसमें वास्तविक ओरेकल बोन की 998 रगड़-छापें शामिल हैं। उसी शोध दल ने S3-Net नामक एक टुकड़ा-जोड़क तंत्रिका नेटवर्क भी प्रस्तुत किया[7]।

यह न्यूरोल नेटवर्क गणना करता है कि टुकड़ों के कनारों के घुमाव आपस में कतिने मिलते-जुलते हैं। जब दो कनारे आपस में सटीक बैठते हैं, तो स्कोर बढ़ जाता है। मशीन सबसे अधिक स्कोर के क्रम में उम्मीदवारों की सूची बनाती है। उसी शोध दल ने 2020 में एक पछिला शोध भी प्रकाशित किया था। उस समय शीर्ष दस उम्मीदवारों में सही उत्तर शामिल होने का अनुपात 98.39 प्रतिशत था[11]।

पहेली जोड़ते समय इंसान चित्र से पहले टुकड़ों का आकार देखता है। मशीन भी यही तरीका अपनाती है। बस अंतर यह है कि मशीन एक दिन में लाखों बार तुलना कर सकती है।

मशीन जो देती है वह सीधा सही उत्तर नहीं, बल्कि संभावित उम्मीदवारों की सूची है। कई बार सूची में सबसे ऊपर का विकल्प गलत भी होता है। इसलिए शोध दल शीर्ष दस में सही उत्तर होने को अपने प्रदर्शन का पैमाना मानता है[7]। किसी इंसान के लिए केवल दस विकल्प जाँचना और एक लाख विकल्प जाँचना पूरी तरह अलग बात है।

कैपटिल नॉर्मल विश्वविद्यालय के ओरेकल बोन लिपि शोध केंद्र ने हेनान विश्वविद्यालय के साथ मिलकर टुकड़ों को जोड़ने का एक प्रोग्राम बनाया। इसका नाम Zhui Da Da है और यह 2020 से इस्तेमाल हो रहा है। यह प्रोग्राम 1,80,000 स्याही छापों (रबगिंस) की एक साथ तुलना करके जोड़ीदार टुकड़े खोजता है। शोध केंद्र ने इस तरीके से ओरेकल बोन के लगभग 1,500 टुकड़ों के रिकॉर्ड व्यवस्थित किए। पहले एक टुकड़े की जोड़ी खोजने में आठ से दस साल तक लग जाते थे। अब कुछ ही दिनों में संभावित उम्मीदवार सामने आ जाते हैं[8]।

वाक्य के स्तर तक पहुँचता परीक्षण

एक अक्षर पढ़ना और पूरे वाक्य को समझना दो अलग बातें हैं। जून 2026 के अंत में जारी S-OBI इस अंतर को मापने वाला एक परीक्षण सेट है। शोधकर्ताओं ने विशेषज्ञों द्वारा सत्यापित 95 मूल स्याही छापों (रबगिंस) से प्रश्न तैयार किए। इसमें कुल 695 प्रश्न हैं और इनके उत्तर वाक्य के स्तर पर देने होते हैं। इसके बाद कई बड़े एआई मॉडलों को ये प्रश्न हल करने के लिए दिए गए। परीक्षा के परिणामों में विभिन्न मॉडलों के बीच प्रदर्शन का अंतर साफ दिखाई दिया। इससे यह भी पता चला कि पूरे वाक्य के प्रवाह को समझना अब भी एक कठिन काम है[9]।

इस परीक्षण ने एक और तथ्य सामने रखा। सामान्य एआई ओरेकल बोन लिपि की समस्याओं को ठीक से हल नहीं कर पाता है[9]। दुनिया भर के पाठ से सीखने वाले मॉडल भी 3,000 साल पुराने वाक्यों के सामने लड़खड़ा जाते हैं। इसका कारण यह है कि ये अक्षर और व्याकरण उन्होंने कभी पढ़े ही नहीं हैं। इसलिए केवल ओरेकल बोन के लिए मॉडलों को फरि से प्रशिक्षित करने का शोध जारी है[5]।

टूटे हुए कनारों से आगे बढ़कर संदर्भ को जोड़ने वाला शोध भी सामने आया है। एक शोध दल ने यह तय करने का कार्य तैयार किया कि क्या दो वाक्य मूल रूप से एक ही हड्डी पर मौजूद थे। उन्होंने OBID-ACR नाम से एक डेटासेट जारी किया। साथ ही उन्होंने अक्षरों के आकार की जानकारी और वाक्य की जानकारी को एक साथ इस्तेमाल करने का तरीका सुझाया[10]। इस तरह टुकड़ों को वास्तव में जोड़े बना भी संदर्भ को जोड़कर देखने का मार्ग प्रशस्त हुआ।

संदर्भ सामग्री

- [1] 梁文艳. (2026, 5월 26일). 以科技赋能, 让甲骨文"活"起来. 中国青年报. https://zqb.cyol.com/pc/content/202605/26/content_426168.html (3 सितंबर 2026 को अभिगमति)
- [2] Chen, Z., Hua, W., Li, J., Zhu, Y., Zhi, X., Liu, Z., Chen, T., Zhang, W., & Zhai, G. (2026). Oracle bone inscriptions information processing: a comprehensive survey. npj Heritage Science, 14, 220. <https://doi.org/10.1038/s40494-026-02511-w> (3 सितंबर 2026 को अभिगमति)
- [3] Liu, Y., Guan, H., Wang, P., Wang, X., Wan, J., Zhang, K., Zheng, H., Liu, X., Kuang, Z., Yang, H., Li, B., Liu, Y., Jin, L., & Bai, X. (2026). AlphaOracle: Oracle bone script decipherment via human-workflow-inspired deep learning. The Innovation, 7(11), 101462. <https://doi.org/10.1016/j.xinn.2026.101462>
- [4] Zhang, J., Li, R., Pang, H., Xia, D., Zhu, Z., Zhang, Q., Li, C., & Yang, X. (2026, 4월 8일). Specializing large models for oracle bone script interpretation via component-grounded multimodal knowledge augmentation. arXiv:2604.06711. <https://arxiv.org/abs/2604.06711> (3 सितंबर 2026 को अभिगमति)
- [5] Song, Z., Wang, Y., Ma, Z., Yu, Z., Wang, T., Zhang, J., Wang, T., & Yu, K. (2026, 6월 24일). OracleAnalyser: Analysing implicit semantics of oracle bone scripts through MLLMs with post-training. arXiv:2606.25906. <https://arxiv.org/abs/2606.25906> (3 सितंबर 2026 को अभिगमति)
- [6] Guan, H., Yang, H., Wang, X., Han, S., Liu, Y., Jin, L., Bai, X., & Liu, Y. (2024). Deciphering oracle bone language with diffusion models. In Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Association for Computational Linguistics. <https://aclanthology.org/2024.acl-long.831/> (3 सितंबर 2026 को अभिगमति)

- [7] Zhang, C., Wang, B., Chen, K., Zong, R., Mo, B., Men, Y., Almpandis, G., Chen, S., & Zhang, X. (2022). Data-driven oracle bone rejoining: A dataset and practical self-supervised learning scheme. In Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (pp. 4482-4492). ACM. <https://doi.org/10.1145/3534678.3539050>
- [8] 北京日报. (2026, 8월 28일). 高校团队"唤醒"三千年前文明碎片. 新浪教育. <https://edu.sina.com.cn/l/2026-08-28/doc-inipwhrw0020843.shtml> (3 सितंबर 2026 को अभिगमिति)
- [9] Li, Z., Chen, Z., Chen, T., & Zhai, G. (2026, 6월 30일). Beyond single character: Evaluating MLLMs for sentence-level oracle bone inscription understanding. arXiv:2606.31169. <https://arxiv.org/abs/2606.31169> (3 सितंबर 2026 को अभिगमिति)
- [10] Zhang, H., Wang, T., Zhang, Z., Li, B., Sun, H., Hou, C., Wang, N., Yu, Y., Jiao, Q., Xiong, J., & Liu, Y. (2026). A multi-modal dataset and method for bone-level association prediction in oracle bone inscriptions. npj Heritage Science, 14, 19. <https://doi.org/10.1038/s40494-025-02282-w> (3 सितंबर 2026 को अभिगमिति)
- [11] Zhang, C., Zong, R., Cao, S., Men, Y., & Mo, B. (2020). AI-powered oracle bone inscriptions recognition and fragments rejoining. In Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence (IJCAI-20), 5309-5311. <https://doi.org/10.24963/ijcai.2020/779>

7.3. सामने आया सच

एक अक्षर बदलते ही नक्शा बदल गया

AlphaOracle की पहली उपलब्धि केवल एक अक्षर थी। शोधकर्ताओं ने लंबे समय से विवाद का विषय रहे एक अक्षर को इस प्रणाली में डाला। यह वही अक्षर है जो बाद में चीनी संप्रतीक 'लाओ (勞)' में बदल गया[1]। इस प्रणाली ने स्ट्रोक की बनावट और घटकों की संरचना को व्यवस्थित किया। फिर उन अन्य वाक्यों को खोजकर सामने रखा जहाँ यह अक्षर आता था। इसके बाद प्राचीन ग्रंथों के ज्ञान से तुलना करके इसके अर्थ का दायरा सीमित किया गया।

प्रणाली ने नष्टिर्क्ष निकाला कि यह अक्षर किसी स्थान का नाम या किसी वंश (कबीले) का नाम था। वंश का अर्थ है एक ही पूर्वज से जुड़े रश्तेदारों का कुल। यह उस पुरानी व्याख्या से बिल्कुल अलग परिणाम था, जिसमें इस अक्षर को वाक्य को सजाने वाला सहायक शब्द माना जाता था। शोधकर्ताओं ने स्पष्ट किया कि इस व्याख्या की प्राचीन लिपिविशेषज्ञों द्वारा वधिवित समीक्षा की गई थी[1]।

एक अक्षर का अर्थ बदलते ही वे सारे वाक्य बदल जाते हैं जिनमें वह अक्षर आता है। यद्विह किसी व्यक्तिका नाम हो, तो वह वाक्य किसी व्यक्तिकी नज्जी कहानी बन जाता है। यद्विह किसी स्थान का नाम हो, तो वही वाक्य स्थानीय प्रशासन का ऐतिहासिक रिकॉर्ड बन जाता है। फिर शांग राजवंश उस भूमिपर कैसे शासन करता था, इसे भी दोबारा समझना पड़ता है। शोध दल ने रिपोर्ट दी कि यह परिणाम शांग साम्राज्य के प्रशासन और समाज को देखने के नज़रिए को पूरी तरह बदल देता है[1]।

इस शोध में सबसे महत्वपूर्ण बात केवल उत्तर नहीं है। असली बात यह है कि प्रणाली ने उत्तर के साथ-साथ उसका ठोस आधार भी प्रस्तुत किया। AlphaOracle यह स्पष्ट दिखाता है कि उसने कनि वाक्यों और कनि घटकों के आधार पर वह नष्टिर्क्ष निकाला[1]। विद्वान उस तर्क का परीक्षण करके तय कर सकते हैं कि वे उससे सहमत हैं या नहीं। इंसान द्वारा जाँचे जा सकने वाले प्रारूप में उत्तर देना ही इस शोध का मुख्य केंद्र था।

सालों का मलिन कार्य घटकर कुछ दनों का रह गया

टुकड़ों को जोड़ने के क्षेत्र में भी परिणाम सामने आए। यह ऊपर उल्लिखित Zhui Da Da प्रोग्राम की ही उपलब्धि है। इस प्रोग्राम ने 300 से अधिक ऐसे जोड़े खोज निकाले जिनमें विद्वान लंबे समय से नहीं पकड़ पाए थे[2]। यह सौ से अधिक वर्षों से इंसानी हाथों द्वारा खोजे गए जोड़ों में जुड़ा एक नया आँकड़ा है।

इंसान की अपनी भूमिका अब भी वैसी ही बनी हुई है। विद्वान मशीन द्वारा दी गई सूची के प्रत्येक विकल्प की जाँच करते हैं। यह सत्यापन वास्तविक अवशेषों के प्रत्यक्ष मलिन के साथ पूरा होता है।

अन्यांग नॉर्मल विश्वविद्यालय की प्रयोगशाला ने भी अपनी उपलब्धियाँ प्रस्तुत कीं। मई 2026 तक नए जोड़े गए ओरेकल बोन के मामले 155 थे। उसी प्रयोगशाला द्वारा पहचाने गए दोहरे रिकॉर्ड्स के मामले लगभग 7,000 तक पहुँचते हैं[3]। कई बार एक ही हड्डी अलग-अलग पुरातात्विक कैटलॉग में दर्ज हो जाती थी। तब एक ही हड्डी दो अलग-अलग अवशेषों के रूप में गिनी जाती थी। नए जोड़े गए टुकड़ों की तुलना में ऐसे दोहरेपन को छोटने की संख्या बहुत अधिक है। इसका सीधा अर्थ है कि मशीन का कार्य केवल नई खोजें करना ही नहीं, बल्कि पुराने ऐतिहासिक अभिलेखों की जाँच करना भी है।

जब टुकड़े आपस में जुड़ते हैं, तो वाक्य लंबे हो जाते हैं। अक्सर एक टुकड़े पर केवल प्रश्न रह जाता है और उसका परिणाम दूसरे टुकड़े पर होता है। जब दोनों टुकड़े मिलते हैं, तभी प्रश्न पूछने की तथि और वास्तव में घटी घटना एक ही सूत्र में जुड़ पाती है। तब यह कई छोटे-छोटे टुकड़ों के बजाय एक संपूर्ण घटना का वस्तुतः वर्णन बन जाता है।

यह अंतर सीधे इतिहास लेखन को प्रभावित करता है। टूटे हुए अधूरे अभिलेखों को देखने पर घटनाएँ बखिरी हुई प्रतीत होती हैं। जुड़े हुए अभिलेखों को देखने पर उनका सही कालक्रम सामने आ जाता है। इससे यह पता चलता है कि राजा ने कब प्रश्न पूछा था और उसे उत्तर कब मिला था।

दोहराव छॉटकर अलग किए गए 7,000 रिकॉर्ड भी उतना ही महत्व रखते हैं[3]। एक ही हड्डी को दो बार गनिने से प्राथमिक डेटा की संख्या ही गलत हो जाती है। शोधकर्ता एक ही वाक्य को दो अलग-अलग रिकॉर्ड समझकर उद्धृत कर बैठते हैं। मशीन ने बड़े पैमाने पर स्याही छापों (रबगिंस) की तुलना करके इस दोहरापन को साफ कर दिया। यह कोई दिखावटी खोज भले न हो, लेकिन बुनियादी ऐतिहासिक साक्ष्यों को ठीक करने का अत्यंत महत्वपूर्ण काम है।

समुद्र पार बखिरी हड्डियाँ स्क्रीन पर मिलती हैं

टुकड़ों को जोड़ने की राह में सबसे बड़ी बाधा दूरी थी। जब एक ही हड्डी के टूटे टुकड़े अलग-अलग देशों में होते, तो उन्हें पास रखकर देखना असंभव था। केवल तस्वीरें भेजने से उनकी मोटाई और कनारों के घुमाव का पता नहीं चल पाता था। 3D डेटा ने इस बाधा को बहुत कम कर दिया।

ऊपर बताए गए 900 से अधिक विदेशी अवशेष इस बाधा को पार करने का सीधा प्रमाण हैं[3]। मूल वस्तुएँ अपने स्थान पर ही सुरक्षित रहती हैं। केवल उनकी सतह के आकार और बनावट का डेटा यहाँ लाया गया है।

अन्यांग का कोई शोधकर्ता अब अपनी स्क्रीन पर उन टुकड़ों को चारों तरफ घुमाकर देख सकता है। बीजिंग के किसी टुकड़े और यूरोप में रखे किसी टुकड़े को एक ही स्क्रीन पर साथ रखकर उनके कनारों को आपस में मिलाया जा सकता है।

सामग्रियों को सार्वजनिक करने का तरीका भी बदल चुका है। पहले पुरावशेषों को देखने के लिए उन्हें सहेजने वाले संस्थानों में व्यक्तिगत रूप से जाना पड़ता था। केवल अनुमत मिलने में ही कई-कई महीने लग जाते थे। अब तो कोई छात्र भी घर बैठे ऑरेकल बोन का डेटा खोल सकता है[3]। जब सामग्री का अध्ययन करने वालों की संख्या बढ़ती है, तो उसे परखने वाली आँखें भी बढ़ जाती हैं। आँखें बढ़ने से गलत जोड़े गए टुकड़ों को पकड़ने के अवसर भी बढ़ जाते हैं।

जो काम मशीन अभी नहीं कर सकती

उपलब्धियों के साथ-साथ सीमाएँ भी स्पष्ट हो गईं। जून 2026 के अंत में जारी S-OB1 परीक्षण ने इन सीमाओं को आँकड़ों के साथ सिद्ध किया। कई बड़े एआई मॉडलों ने ऑरेकल बोन लिपि के वाक्यों वाले एक जैसे प्रश्न हल किए। किसी एक अक्षर को पहचानने वाले प्रश्नों की तुलना में पूरे वाक्य को समझने वाले प्रश्नों में उनका प्रदर्शन गिर गया। इसका कारण भी सामने आया। मॉडल स्पष्ट रूप से बचे हुए अक्षरों को भी शुरू में ही गलत पढ़ लेते थे। उसी त्रुटि के साथ आगे का अनुमान लगाने के कारण उनका अंतिम उत्तर गलत हो गया[4]। 3,000 वर्ष पुराने वाक्यों के प्रवाह को समझना अभी भी काफी हद तक इंसानों का ही काम है।

जिस बात से सबसे अधिक सतर्क रहने की आवश्यकता है, वह है मनगढ़ंत उत्तर गढ़ना (हैलुसिनेशन)। एआई संभावनाओं के आधार पर उत्तर चुनता है। ऑरेकल बोन लिपि जैसे सीमित अध्ययन सामग्री वाले क्षेत्रों में, वह ऐसे काल्पनिक अक्षर भी गढ़ सकता है जो देखने में तो स्वाभाविक लगें पर वास्तव में कभी थे ही नहीं। शांग राजवंश की भाषा की तुलना के लिए उस दौर की पड़ोसी भाषाओं के स्रोत भी पर्याप्त नहीं हैं[5]। इसलिए शब्दकोश देखकर उत्तर की पुष्टि करना बेहद कठिन है।

डेटा का एकतरफा झुकाव भी एक समस्या बना हुआ है। अब तक एकतरफे ऑरेकल बोन सामग्री किसी खास समय और विशिष्ट उत्खनन स्थलों पर ही केंद्रित है। ऐसे असंतुलित डेटा से प्रशिक्षित मशीन जब किसी नए क्षेत्र के ऑरेकल बोन देखती है, तो उससे गलती होने की पूरी संभावना रहती है। 2026 में प्रकाशित एक व्यापक समीक्षा शोध-पत्र ने भी डेटा की गुणवत्ता और संतुलन को भविष्य की बड़ी चुनौती माना है[5]।

आँकड़ों के हिसाब से देखें तो अभी लंबा सफर तय करना बाकी है। जिन अक्षरों का अर्थ अज्ञात है, उनकी संख्या अब भी लगभग 3,000 है। AlphaOracle ने तर्कों के साथ जिस अक्षर की गुत्थी सुलझाई, वह इनमें से केवल एक था[1]। एक अकेले अक्षर को समझने में लगी मेहनत को देखते हुए बचा हुआ कार्य बहुत बड़ा है। फिर भी, यद्विधि पूरी तरह स्थापित हो जाए, तो अगले अक्षरों का समाधान थोड़ा तेजी से हो सकेगा।

इंसान और मशीन के बीच बँटा काम

इसलिए वर्तमान शोध पद्धति में मनुष्य और मशीन को साथ मलिकर काम करने का स्वरूप तय हो चुका है। मशीन संभावित उम्मीदवारों का व्यापक दायरा प्रस्तुत करती है और उसके तर्कों को व्यवस्थित करती है। वहीं इंसान ऐतिहासिक ग्रंथों और वास्तविक अवशेषों से उन वकिल्पो की पुष्टि करता है।

ऊपर देखा गया विश्लेषण के समय में संक्षेप भी इसी श्रम-वभाजन का फल है। मशीन ने मानवीय नर्णय की जगह नहीं ली, बल्कि उसने मनुष्य के परीक्षण के लिए आवश्यक सामग्रियों को पहले ही छाँटकर रख दिया। विशेषज्ञों द्वारा इस उपकरण को बेहद उपयोगी बताए जाने का मुख्य कारण भी यही है[1]।

टुकड़ों को जोड़ने की प्रक्रिया में भी यही प्रणाली दृढ़ हो चुकी है। पहले जसि तुलना में वर्षों लग जाते थे, उसे प्रोग्राम कुछ ही दिनों में समाप्त कर देता है[2]। शोधकर्ता उन संभावित वकिल्पो को लेकर संग्रहालय के भंडार कक्ष में जाते हैं। वे वास्तविक हड्डियों को आमने-सामने रखकर देखते हैं कि क्या उनके खंडित किनारे आपस में जुड़ते हैं। अंतमि नर्णय मनुष्य की उंगलियों के स्पर्श से ही तय होता है।

इस कार्य-वभाजन में एक अनविर्य शर्त जुड़ी है। वह यह कि मशीन द्वारा दिए गए उत्तर का आधार मनुष्य को साफ दिखाई देना चाहिए। यदि आधार ही स्पष्ट न हो, तो वदिवानों के पास सत्यापन का कोई उपाय नहीं बचता। और जसि उत्तर की पुष्टि हो सके, उसे इतिहास लेखन में स्थान नहीं दिया जा सकता। AlphaOracle द्वारा अपनी प्रक्रिया को वभिन्न चरणों में बाँटने का मुख्य कारण भी यही है[1]।

3,000 वर्ष पहले शांग राजवंश का राजा हड्डियों की दरारों को देखकर भवष्य पूछता था। आज का शोधकर्ता उसी हड्डी के टूटे हुए किनारों को देखकर अतीत के बारे में प्रश्न पूछता है। प्रश्नों की दशा भले ही उलट हो, लेकिन जसि दरार को वे देख रहे हैं, वह आज भी वही है।

संदर्भ सामग्री

- [1] Liu, Y., Guan, H., Wang, P., Wang, X., Wan, J., Zhang, K., Zheng, H., Liu, X., Kuang, Z., Yang, H., Li, B., Liu, Y., Jin, L., & Bai, X. (2026). AlphaOracle: Oracle bone script decipherment via human-workflow-inspired deep learning. *The Innovation*, 7(11), 101462. <https://doi.org/10.1016/j.xinn.2026.101462>
- [2] 北京日报. (2026, 8월 28일). 高校团队"唤醒"三千年前文明碎片. 新浪教育. <https://edu.sina.com.cn/l/2026-08-28/doc-inipwhrw0020843.shtml> (3 सितंबर 2026 को अभिगमिति)
- [3] 梁文艳. (2026, 5월 26일). 以科技赋能, 让甲骨文"活"起来. 中国青年报. https://zqb.cyol.com/pc/content/202605/26/content_426168.html (3 सितंबर 2026 को अभिगमिति)
- [4] Li, Z., Chen, Z., Chen, T., & Zhai, G. (2026, 7월 1일). Beyond single character: Evaluating MLLMs for sentence-level oracle bone inscription understanding. *arXiv:2606.31169*. <https://arxiv.org/abs/2606.31169> (3 सितंबर 2026 को अभिगमिति)
- [5] Chen, Z., Hua, W., Li, J., Zhu, Y., Zhi, X., Liu, Z., Chen, T., Zhang, W., & Zhai, G. (2026). Oracle bone inscriptions information processing: a comprehensive survey. *npj Heritage Science*, 14, 220. <https://doi.org/10.1038/s40494-026-02511-w> (3 सितंबर 2026 को अभिगमिति)

अध्याय 8. कीचड़ के गड्ढे से पुनर्जीवित सुलेख: सलिला की काष्ठ पट्टिकाएँ और कर्सवि लपि पहचान एआई

8.1. केस फाइल

लकड़ी पर लिखने का युग

एक समय था जब कागज़ दुर्लभ था। प्राचीन पूर्वी एशिया के लोग लकड़ी को पतला तराश कर उस पर लिखते थे। अक्षर लिखी इन लकड़ी की पट्टियों को काष्ठ पट्टिकाएँ कहा जाता है। काष्ठ पट्टिकाएँ आमतौर पर दो से तीन उँगलियों जितनी चौड़ी होती हैं। इनमें से कई की लंबाई लगभग एक बालशित जितनी होती है। कुछ लकड़ियों को खंभे की तरह भी तराशा जाता था। ऐसी काष्ठ पट्टिकाओं के चारों तरफ लिखा जा सकता है। कई सतहों पर लिखे जाने के कारण इन्हें बहुमुखी काष्ठ पट्टिका कहा जाता है।

काष्ठ पट्टिकाओं के कई उपयोग थे। एक उपयोग सामान पर बाँधने वाले टैग के रूप में था। प्रांतों से राजधानी अनाज या नमकीन मछली भेजते समय काष्ठ पट्टिका को डोरी से बाँध दिया जाता था। उस पर भेजने वाले क्षेत्र का नाम, वस्तु का नाम और मात्रा लिखी होती थी। दूसरा उपयोग सरकारी कार्यालयों के दस्तावेज़ के रूप में था। अधिकारियों के नाम, पद, प्राप्त वस्तु और तारीख दर्ज की हुई काष्ठ पट्टिकाएँ मिलती हैं^[1]। लिखावट सीखने के लिए एक ही चीनी अक्षर को बार-बार लिखकर अभ्यास की गई काष्ठ पट्टिकाएँ भी हैं।

लकड़ी कागज़ की तुलना में कठोर थी, इसलिए इसका दोबारा उपयोग किया जा सकता था। उपयोग समाप्त होने पर काष्ठ पट्टिका की ऊपरी सतह को चाकू से पतला छील दिया जाता था। नई सतह सामने आने पर उस पर फिर से लिखा जाता था। इस तरह छिली गई पतली लकड़ी की कतरने भी पुरावशेष स्थलों से साथ मिलती हैं। इन कतरनों को sakseol कहा जाता है। इसका अर्थ लिखावट मटिने के लिए छिली गई लकड़ी की कतरन है। बुयेओ ग्वानबुक-री अवशेष स्थल से मिली 329 काष्ठ पट्टिकाओं को देखने पर यह समझा जा सकता है। उनमें से 247 वस्तुएँ छीलन ही थीं। इसका अर्थ है कि दस्तावेज़ बनाने और सुधारने की व्यवस्था पास में ही मौजूद थी^[1]।

इतिहासकारों द्वारा काष्ठ पट्टिकाओं को बहुमूल्य मानने का एक कारण है। इतिहास की पुस्तकें आमतौर पर घटनाएँ बीत जाने के बाद लिखी जाती हैं। उनमें लेखक का दृष्टिकोण और राजवंश की परिस्थितियाँ शामिल होती हैं। काष्ठ पट्टिकाएँ उस दिने के काम को उसी दिने दर्ज करने वाले दस्तावेज़ के समान हैं। ये भावी पीढ़ी को दिखाने के लिए सजाकर लिखे गए लेख नहीं हैं। कसिने, कब और कतिना अनाज भेजा, इसका विवरण ज्यों का त्यों बचा रहा। घटना के समय बने ऐसे अभिलेखों को प्राथमिक स्रोत कहा जाता है। काष्ठ पट्टिकाओं पर भी गलत लिखे गए अक्षर और सुधारे गए स्थान मिलते हैं। इसलिए शोधकर्ता कई नमूनों को एक साथ रखकर उनकी तुलना करते हैं।

काष्ठ पट्टिकाएँ कोरियाई प्रायद्वीप और जापानी द्वीपसमूह दोनों जगहों से मिलती हैं। ये दोनों ही चीनी वर्णों से दस्तावेज़ बनाने वाले समाज थे। इसके बावजूद उन्होंने अपनी मातृभाषा लिखने के उपाय खोजे। सलिला के लोगों ने चीनी अक्षरों के अर्थ और ध्वनिको मिलाकर अपनी भाषा के शब्द-क्रम को लिखा। इस लेखन पद्धतिको idu और hyangchal कहा जाता है। जापानी लोगों ने भी चीनी अक्षरों की ध्वनिका सहारा लेकर अपनी भाषा लिखी। काष्ठ पट्टिकाएँ इस बात का प्रत्यक्ष प्रमाण हैं कि यह प्रयास कब शुरू हुआ था। इसका कारण यह है कि कागज़ के दस्तावेज़ अब लगभग नहीं बचे हैं। यही वजह है कि दोनों देशों की काष्ठ पट्टिकाओं की तुलना करने वाले शोध नरितर जारी हैं।

कीचड़ द्वारा सुरक्षित एक हजार वर्ष

काष्ठ पट्टिकाएँ मिलने के स्थान प्रायः नश्चित होते हैं। वे महल के पास तालाब के तल में मिलती हैं। वे कल्लि की दीवार के भीतर जलभराव वाली नचिली दलदली भूमि में मिलती हैं। वे नालों को भरने वाली कीचड़ की परत में मिलती हैं। लोगों द्वारा उपयोग के बाद फेंकी गई लकड़ी पानी में डूबकर सीधे मट्टी में दब गई।

पानी के भीतर कीचड़ में हवा लगभग नहीं पहुँच पाती। लकड़ी को सड़ाने वाले मुख्य कारक हवा में साँस लेने वाले सूक्ष्मजीव होते हैं। हवा बंद होने पर ये सूक्ष्मजीव काम नहीं कर पाते। बिना हवा के जीवित रहने वाले सूक्ष्मजीव हालांकि बिचे रहते हैं। फिर भी लकड़ी के सड़ने की गति बहुत धीमी हो जाती है। इसी कारण एक हजार से अधिक वर्षों तक लकड़ी का आकार बना रहा। जसि मट्टी में चमड़ा और कागज़ बिना नशान मटि जाते हैं, वहाँ भी काष्ठ पट्टिकाएँ सुरक्षित रही। कब्रों ने नहीं, बल्कि कचरे के गड्ढों ने इन अभिलेखों को सुरक्षित रखा।

इसके बदले एक अन्य परिवर्तन हुआ। लकड़ी के भीतर पानी रसि गया और उसने कोशिका भित्तियों को भर दिया। मट्टी का लौह तत्व और कई खनजि भी साथ में अंदर चले गए। लकड़ी पूरी तरह जलमग्न अवस्था में आ गई। ऐसी लकड़ी को जल-संतृप्त काष्ठ कहा जाता है।

जल-संतृप्त काष्ठ को बाहर निकालकर साधारण तरीके से नहीं सुखाया जा सकता। पानी निकलते ही यह तेज़ी से सिकुड़ती और चटक जाती है। उत्खनन स्थल पर काष्ठ पट्टिकाओं को तुरंत पानी के टब में डुबोने का यही कारण है। संरक्षण प्रयोगशाला में लकड़ी के भीतर के पानी को रसायनों से धीरे-धीरे बदला जाता है। इस प्रक्रिया में ही कई महीनों से कई वर्षों का समय लग जाता है। पढ़ने का काम इससे पहले ही पूरा कर लेना चाहिए। रासायनिक उपचार के दौरान सतह का रंग बदल सकता है।

काली लकड़ी पर काले अक्षर

खनजिों के समाने से लकड़ी का रंग गहरा हो जाता है। कई लकड़ियाँ भूरे से बदलकर बलिकूल काली हो जाती हैं। लकड़ी के काले पड़ने की इस परिघटना को कजलीकरण कहा जाता है। समस्या लिखने की सामग्री से जुड़ी है। प्राचीन लोग चीड़ की कालखि को सरेस में घोलकर पारंपरिक स्याही बनाते थे। स्याही का मुख्य तत्व कार्बन होता है और इसका रंग काला होता है।

काली पड़ चुकी लकड़ी पर काली स्याही से लिखे अक्षर मौजूद हैं। तेज़ रोशनी डालने पर भी दोनों में कोई अंतर नहीं दिखता। इसके अलावा पानी में सदियों रहने के कारण स्याही के कई कण धुल चुके हैं। लकड़ी के रेशों की गहरी खड़ी धारियाँ ब्रश के नशानों से मलि जाती हैं। मट्टी और सड़े हुए तत्व ब्रश की लकीरों के खाँचों को भर देते हैं। मानवीय आँखों से यह जानना भी मुश्कलि हो जाता है कि वहाँ कोई अक्षर है भी या नहीं।

पढ़ना कठनि होने पर वदिवानों के सदिधांत भी बँट जाते हैं। एक ही काष्ठ पट्टिका को लेकर अलग-अलग वदिवान भनिन चीनी अक्षर पढ़ लेते हैं। एक अक्षर बदलने से वाक्य का अर्थ बदल जाता है और इतहिास की व्याख्या भी बदल जाती है। इसलए काष्ठ पट्टिका शोधकर्ता लंबे समय से एक नई आँख की तलाश में थे। उन्हें ऐसे उपकरण की ज़रूरत थी जो मानवीय आँखों से छूट जाने वाली बारीकियों को पकड़ सके।

पढ़े जाने योग्य काष्ठ पट्टिकाओं की संख्या लगातार बढ़ रही है। ग्योगनाम के हामआन सेओगसान पर्वत दुर्ग से सलिला की काष्ठ पट्टिकाएँ लगातार मलिती रही हैं। अगस्त 2025 में भी दो नई अभलिखति काष्ठ पट्टिकाओं की पुष्टि हुई[2]। अभलिख का अर्थ पुरावशेष पर उकरे या लिखे गए अक्षरों से है। चुंगनाम के बुयेओ ग्वानबुक-री पुरावशेष स्थल से बाएकजे के साबी काल की काष्ठ पट्टिकाएँ बड़ी संख्या में नकिली[1]। ग्योगगी के यांगजू दायमोसान दुर्ग से 5वीं सदी की मानी जाने वाली काष्ठ पट्टिका मलि है। यदयिह सच है, तो यह अब तक ज्ञात बाएकजे लिखति सामग्रियों में सबसे प्राचीन दौर की होगी[3]।

जापान में भी स्थिति सिमान है। नारा राष्ट्रीय सांस्कृतिक संपदा अनुसंधान संस्थान वभिनिन स्थानों से प्राप्त काष्ठ पट्टिकाओं की जानकारी एकत्र कर प्रकाशति करता है। इस डेटाबेस का नाम Mokkanko है[4]। यहाँ केवल काष्ठ पट्टिका पर लिखे अक्षरों से ही खोज नहीं होती। इसके उपयोग, काल, सामग्री और प्राप्त स्थल के आधार पर भी खोज की जा सकती है[5]। पढ़ने के लिए इस तरह लकड़ियों का ढेर लगा हुआ है।

संदर्भ सामग्री

- [1] 파이낸셜뉴스. (2026, 2월 5일). 부여 관북리 '백제 피리' 첫 확인...목간 329점 무더기 발견. 파이낸셜뉴스. <https://www.fnnews.com/news/202602051448564101> (3 सतिंबर 2026 को अभगिमति)
- [2] 경남일보. (2025, 8월 7일). 함안 성산산성서 새로운 목간 2점 명문 확인. 경남일보. <https://www.gnnews.co.kr/news/articleView.html?idxno=616028> (3 सतिंबर 2026 को अभगिमति)
- [3] 뉴시스. (2025, 11월 20일). 1500년 전 '백제 문자' 깨어났다...양주 대모산성서 5세기 목간 출토. 뉴시스. https://www.newsis.com/view/NISX20251120_0003410107 (3 सतिंबर 2026 को अभगिमति)
- [4] 奈良文化財研究所. 木簡庫(MOKKAN-KO). <https://mokkanko.nabunken.go.jp/ja/> (3 सतिंबर 2026 को अभगिमति)
- [5] カレントアウェアネス・ポータル. (2018). 奈良文化財研究所、木簡に関する情報を検索するシステム「木簡庫」(MOKKAN-KO)を公開. 国立国会図書館. <https://current.ndl.go.jp/car/35714> (3 सतिंबर 2026 को अभगिमति)

8.2. वैज्ञानिक जांच दल की खोज 1: डिजिटल स्याही आवर्धक

अदृश्य प्रकाश का उपयोग

मानव आँखें प्रकाश के एक सीमति दायरे को ही देख पाती हैं। इंद्रधनुष के लाल रंग के बाहर भी प्रकाश होता है। इस प्रकाश को अवरक्त (इन्फ्रारेड) प्रकाश कहा जाता है। इसका तरंगदैर्घ्य लंबा होने से यह आँखों को नहीं दिखता, लेकिन कैमरे से इसकी तस्वीर ली जा सकती है। रमिोट कंट्रोल जो संकेत टेलीवज़िन को भेजता है, वह भी इन्फ्रारेड ही होता है।

काष्ठ पट्टिकाओं को पढ़ने के लिए इन्फ्रारेड का उपयोग करने का कारण सामग्री का गुण है। स्याही के कार्बन कण इन्फ्रारेड को भली-भाँति अवशोषति करते हैं। काली पड़ चुकी लकड़ी के रेशे बहुत अधिक इन्फ्रारेड परावर्तति करते हैं। मानवीय आँखों को एक जैसे काले दिखने वाले दोनों पदार्थ इन्फ्रारेड कैमरे के सामने अलग हो जाते हैं। अवशोषति करने वाला भाग गहरा आता है और परावर्तति करने वाला भाग चमकीला आता है। जो ब्रश के नशान मटि चुके माने गए थे, वे तस्वीर पर उभर आते हैं।

मुख्य रूप से प्रयुक्त प्रकाश लगभग 800 नैनोमीटर तरंगदैर्घ्य वाला नकिट-अवरक्त प्रकाश है। नैनोमीटर 1 मीटर के 1 अरबवें हिस्से की लंबाई होती है। मनुष्य को दिखने वाला लाल प्रकाश 700 नैनोमीटर के पास समाप्त हो जाता है। इसका अर्थ है कि उसके ठीक बाहर के

प्रकाश का उपयोग किया जाता है। छायांकन बाहरी रोशनी को रोके गए कमरे में किया जाता है। प्रकाश को एकतरफा किए बिना समान रूप से फैलाकर डाला जाता है। यदि केवल एक तरफ से रोशनी दी जाए, तो लकड़ी के रेशों की खाइयों में लंबी परछाइयाँ बन जाती हैं।

इन्फ्रारेड फोटोग्राफी काष्ठ पट्टिका अनुसंधान की बुनियादी प्रक्रिया बन चुकी है। उत्खनन रपिर्टों में छपने वाली काष्ठ पट्टिकाओं की तस्वीरों के बगल में प्रायः इन्फ्रारेड तस्वीरें भी साथ रखी जाती हैं। नारा राष्ट्रीय सांस्कृतिक संपदा अनुसंधान संस्थान ने इन्फ्रारेड कैमरे से काष्ठ पट्टिका की अक्षर-छवियों को स्पष्ट करने वाला उपकरण काफी पहले ही विकसित कर लिया था[1]।

इन्फ्रारेड से हाइपरस्पेक्ट्रल की ओर

सिर्फ एक इन्फ्रारेड तस्वीर से सभी अक्षर स्पष्ट नहीं हो पाते। दाग-धब्बों के स्थान और लकड़ी के रेशों की स्थितियों के अनुसार सबसे साफ दिखने वाला तरंगदैर्घ्य अलग होता है। इसीलिए सामने आई तकनीक हाइपरस्पेक्ट्रल इमेजिंग है।

सामान्य तस्वीर केवल लाल, हरे और नीले तीन मान ही दर्ज करती है। हाइपरस्पेक्ट्रल इमेजिंग इससे अलग है। यह एक दृश्य को दर्जनों से लेकर सैकड़ों संकीर्ण तरंगदैर्घ्य खंडों में बाँटकर चित्रित करती है। आप मान सकते हैं कि प्रकाश को इंद्रधनुषी रंगों में बारीक बाँटकर अलग-अलग खींचा जाता है। स्क्रीन के प्रत्येक बिंदु के लिए प्रत्येक तरंगदैर्घ्य का चमक मान दर्ज होता है। इसलिए एक तस्वीर की जगह तस्वीरों का एक मोटा बंडल बन जाता है।

इन मानों की तुलना करके स्याही लगे स्थान और मट्टी के दाग वाले स्थान में अंतर किया जा सकता है। ऐसा इसलिए है क्योंकि स्याही और मट्टी के लिए प्रत्येक तरंगदैर्घ्य पर चमक बदलने का तरीका अलग होता है। कुछ अक्षर 900 नैनोमीटर के आसपास स्पष्ट दिखते हैं। कुछ अक्षर अन्य खंडों में उभरते हैं। शोधकर्ता तरंगदैर्घ्य खंडों को चुन-चुनकर सबसे पठनीय संयोजन खोजते हैं। कंप्यूटर कई खंडों की स्वचालित रूप से तुलना करके संभावित विकल्पों को भी छांटता है। मनुष्य द्वारा एक-एक तस्वीर को आँखों से जाँचने के काम को गणना द्वारा आसान बना दिया गया है।

राष्ट्रीय गया सांस्कृतिक वारिसत अनुसंधान संस्थान ने 2025 में पहली बार इस विधि का प्रयोग किया। उन्होंने हामआन सेओगसान पर्वत दुर्ग की काष्ठ पट्टिकाओं को पढ़ने के लिए हाइपरस्पेक्ट्रल इमेजिंग को अपनाया। संस्थान ने बताया कि यह तरीका पारंपरिक इन्फ्रारेड फोटोग्राफी से अधिक सटीक है। यह उन अक्षरों को भी स्पष्ट कर सकता है जिन्हें नग्न आँखों या सामान्य फोटोग्राफी से पहचानना कठिन होता है[2]।

अक्षरों के आकार से अक्षर ढूँढना

तस्वीर साफ होने के बाद भी अगली समस्या बनी रहती है। यह तय करना होता है कि ब्रश का धुंधला निशान कौन सा चीनी अक्षर है। कर्सिव लिपि में घसीटकर लिखे गए अक्षरों का आकार बहुत बदल जाता है। कई अक्षरों में तो केवल कुछ ही लकीरें बची रह जाती हैं। चीनी अक्षरों की संख्या हजारों में होने के कारण किसी इंसान द्वारा शब्दकोश पलटने की भी एक सीमा होती है।

नारा राष्ट्रीय सांस्कृतिक संपदा अनुसंधान संस्थान और टोक्यो विश्वविद्यालय के ऐतिहासिक सामग्री संकलन संस्थान ने इसमें मदद के लिए एक प्रणाली बनाई। इसका नाम Mojizo है। उपयोगकर्ता जिस अक्षर को पढ़ना चाहता है, उसकी छविको काटकर अपलोड करता है। यह प्रणाली काष्ठ पट्टिकाओं और प्राचीन अभिलेखों से संकलित अक्षर-छवियों को खंगालती है। उनमें से समान आकार वाले अक्षरों को ढूँढ़कर दिखाती है[3]। इंसान उन विकल्पों को देखकर निर्णय लेता है। यह सीधे उत्तर देने वाली मशीन नहीं, बल्कि मिलिते-जुलते उदाहरण सामने लाने वाला एक खोज उपकरण है।

ऐसे उपकरणों ने काम करने के क्रम को बदल दिया है। पहले विद्वान अपनी याददाश्त और शब्दकोश के सहारे संभावित अक्षरों का अनुमान लगाते थे। अब मशीन हजारों उदाहरणों में से मिलिते-जुलते अक्षरों को पहले ही चुन देती है। इंसान उन सीमित विकल्पों में से संदर्भ के अनुसार सही अक्षर चुनता है। इससे पढ़ने में लगने वाला समय बहुत कम हो गया है।

मशीन द्वारा खींची गई लकीरों को अलग करना

आजकल का जनरेटिव एआई धुंधली तस्वीरों को स्पष्ट बना देता है। यह टूटी हुई रेखाओं को जोड़ देता है और खाली जगहों को भर देता है। काष्ठ पट्टिका की तस्वीरों पर इस तकनीक का प्रयोग करने से परिणाम देखने में बहुत सुंदर लगते हैं। यही पर एक खतरा छुपा है।

मशीन द्वारा भरी गई लकीर लकड़ी पर वास्तव में बची स्याही नहीं होती। यह सीखी गई हजारों लिखावटों के औसत से निकाला गया एक अनुमान मात्र है। यदि लकड़ी के रेशे गोल आकार में बचे हों, तो मशीन उसे गोल लकीर मान सकती है। ऐसा होने पर जो अक्षर वास्तव में था ही नहीं, वह भी सच जैसा बनकर उभर आता है। इससे ऐतिहासिक स्रोत दूषित हो जाते हैं और इतिहास की व्याख्या गलत हो जाती है।

इसलिए शोध के क्षेत्र में एक सख्त नियम है। मूल तस्वीर और सुधारे गए परिणाम को हमेशा साथ-साथ रखा जाता है। वास्तव में दिखने वाले स्याही के निशानों और मशीन द्वारा बनाए गए पक्सिल को अलग-अलग चिह्नित किया जाता है। मशीन के उत्तर को नष्ट नहीं, बल्कि

संभावित विकल्पों की सूची माना जाता है। अंतिम निर्णय मनुष्य ही करता है। पठन का प्रारूप प्रस्तुत करते समय यह भी स्पष्ट किया जाता है कि किस उपकरण से और कनि परस्थितियों में तस्वीर ली गई थी।

सरकारी स्तर पर भी तैयारियाँ चल रही हैं। राष्ट्रीय विज्ञान एवं प्रौद्योगिकी सलाहकार परिषद ने 16 अप्रैल 2026 को राष्ट्रीय वरिसत प्रशासन की कार्यान्वयन योजना को मंजूरी दी। यह 2026 से 2030 तक राष्ट्रीय वरिसत की जाँच और जीर्णोद्धार में एआई, रोबोट और सेसर को शामिल करने की योजना है। इस वर्ष का नविश आकार 14 अरब डॉन है। इसमें से 4.4 अरब डॉन बुद्धिमान उन्नत संरक्षण तकनीक के विकास के लिए आवंटित किए गए हैं[4]। राष्ट्रीय गमिहे संग्रहालय ने 23 जून 2026 को एक शैक्षणिक संगोष्ठी का आयोजन किया। इसका विषय गया वरिसत अनुसंधान पर वैज्ञानिक दृष्टिकोण था। इसमें गैर-वनिशकारी विश्लेषण, डिजिटल पुनरुद्धार और 3D X-ray माइक्रोस्कोपी तकनीक पर सात प्रस्तुतियाँ दी गईं[5]। नग्न आँखों से न देखने वाले पुरावशेषों के भीतर झाँकने के तरीके अब काष्ठ पट्टिकाओं से परे भी फैल रहे हैं।

संदर्भ सामग्री

- [1] 奈良文化財研究所. 赤外線カメラを利用した木簡文字画像鮮明化システムの開発. 学術機関リポジトリデータベース. <https://irdb.nii.ac.jp/01144/0004944317> (3 सितंबर 2026 को अभिगमिति)
- [2] 경남일보. (2025, 8월 7일). 함안 성산산성서 새로운 목간 2점 명문 확인. 경남일보. <https://www.gnnews.co.kr/news/articleView.html?idxno=616028> (3 सितंबर 2026 को अभिगमिति)
- [3] 奈良文化財研究所, 東京大学史料編纂所. 木簡・くずし字解読システム MOJIZO. <https://aimojizo.nabunken.go.jp/> (3 सितंबर 2026 को अभिगमिति)
- [4] 뉴시스. (2026, 4월 16일). "AI, 블록체인으로 국가유산 보존관리"...과기자문회의, 유산청 시행계획 의결. 뉴시스. https://www.newsis.com/view/NISX20260416_0003594216 (3 सितंबर 2026 को अभिगमिति)
- [5] 경남일보. (2026, 6월 17일). 국립김해박물관, 23일 '가야학술제전' 심포지엄 개최. 경남일보. <https://www.gnnews.co.kr/news/articleView.html?idxno=639314> (3 सितंबर 2026 को अभिगमिति)

8.3. वैज्ञानिक जांच दल की खोज 2: KuroNet और Miwo

जुड़े हुए अक्षरों को अलग नहीं किया जा सकता

जापान में एदो काल तक लिखे गए लाखों पुराने दस्तावेज़ मौजूद हैं। ये दस्तावेज़ Kuzushiji नामक कर्सिव लिपि में लिखे गए थे। Kuzushiji 8वीं सदी से एक हजार से अधिक वर्षों तक इस्तेमाल की गई कर्सिव लिपि है। वर्ष 1900 में जापान ने अपनी लेखन प्रणाली बदल दी। उसके बाद से स्कूलों में यह लिपि सिखाना बंद कर दिया गया। इसलिए आज के अधिकांश जापानी 150 साल पहले की किताबें भी नहीं पढ़ पाते[1]।

इसे पढ़ना कठिन होने का कारण सिर्फ अक्षरों की बनावट नहीं है। ब्रश को कागज़ से उठाए बना कई अक्षरों को लगातार लिखने की एक शैली होती है। पछिले अक्षर की अंतिम लकीर और अगले अक्षर की पहली लकीर आपस में जुड़ जाती है। ऐसा होने पर अक्षरों के बीच कोई सीमा रेखा नहीं बचती। पूरी पंक्ति ही एक लंबी नरितर रेखा जैसी दिखाई देती है।

कंप्यूटर अक्षर पहचान प्रणाली लंबे समय तक एक ही क्रम में काम करती रही। पहले वह प्रत्येक अक्षर को एक चौकोर बॉक्स में काटती है। उसके बाद कटे हुए बॉक्स की छविको देखकर पहचानती है कि वह कौन सा अक्षर है। यह तरीका केवल तभी काम करता है जब सीमा रेखा मौजूद हो। लगातार जुड़ी हुई कर्सिव लिखावट के सामने यह प्रक्रिया पहले ही कदम पर विफल हो गई। यदि पहला चरण गलत हो जाए तो बाद का परिणाम भी गलत होता है, इसलिए सटीकता बढ़ाने का कोई रास्ता नहीं था।

अक्षरों को अलग किए बना पढ़ने वाला KuroNet

जापान के सेंटर फॉर ओपन डेटा इन द ह्यूमैनिटीज के शोधकर्ताओं ने एक नया रास्ता खोला। थारनि क्लानुवात, एलेक्स लैम्ब और असानोबु कितामोतो ने मलिकर एक मॉडल विकसित किया। इसका नाम KuroNet है। इसे 2019 के अंतरराष्ट्रीय दस्तावेज़ विश्लेषण एवं पहचान सम्मेलन में पहली बार प्रस्तुत किया गया था[1]। वर्ष 2020 में इसकी संरचना को और सुधारकर इसके परिणामों को एक शोध पत्रिका में प्रकाशित किया गया[2]।

KuroNet अक्षरों को पहले से अलग नहीं करता है। यह प्राचीन दस्तावेज़ के पूरे पृष्ठ को एक साथ इनपुट के रूप में लेता है। मॉडल की मुख्य संरचना U-Net है। U-Net छविको धीरे-धीरे छोटा करके उसकी विशेषताएँ निकालता है। इसके बाद यह उसे फिर से मूल आकार में बढ़ा करने वाला न्यूरल नेटवर्क है। छोटा करने वाला भाग और बढ़ा करने वाला भाग सममति होते हैं, इसलिए इसका आकार अंग्रेज़ी अक्षर U जैसा बनता है। छोटा करते समय यह ब्रश के स्ट्रोक की मोटाई और दिशा सीखता है। बढ़ा करते समय यह उस जानकारी को उसके मूल स्थान पर वापस पहुँचाता है।

यह न्यूरोल नेटवर्क पूरे पृष्ठ पर एक साथ दो बातों का अनुमान लगाता है। पहली बात यह है कि अक्षर का केंद्र कहाँ है। दूसरी बात यह है कि उस स्थान पर कौन सा अक्षर है। इसका परिणाम प्रायिकता मानचित्र के रूप में आता है। यह स्क्रीन पर जगह-जगह अक्षर होने की संभावना दर्शाने वाला चित्र है। उच्च प्रायिकता वाले स्थानों का अनुसरण करने पर अक्षर की स्थिति और पहचान एक साथ तय हो जाती है। अक्षरों को काटने का कोई चरण नहीं होता, इसलिए आपस में जुड़े हुए अक्षरों में भी यह प्रणाली लड़खड़ाती नहीं है।

प्रशिक्षण के लिए जापान के राष्ट्रीय जापानी साहित्य अनुसंधान संस्थान द्वारा तैयार की गई सामग्री का उपयोग किया गया। इसका नाम 'जापानी शास्त्रीय ग्रंथ कुजुशिजी डेटासेट' (Kuzushiji Dataset) है। इसमें 10 लाख से अधिक अक्षरों के नमूने शामिल हैं[2]। उसी शोध दल ने मशीन लर्निंग शोधकर्ताओं के आसान उपयोग के लिए इसका एक संक्षिप्त संस्करण भी जारी किया। वह हस्तलिखित अंकों के डेटा की तर्ज पर बनाया गया Kuzushiji-MNIST था। उन्होंने अक्षरों की संख्या बढ़ाने वाला संस्करण और केवल कांजी (चीनी वर्ण) एकत्र करने वाला संस्करण भी साथ जारी किया[3]। इसकी बदौलत जापानी शास्त्रीय साहित्य से अपरिचित शोधकर्ता भी इस चुनौती पर काम करने के लिए आगे आ सके।

KuroNet ने अक्टूबर 2022 में अपनी ज़िम्मेदारी आगे सौंप दी। पहचान मॉडल का नाम बदलकर Ruri हो गया[4]। Ruri वस्तु पहचान तकनीक को घसीट लपि के अनुकूल बनाकर तैयार किया गया मॉडल है। अक्षरों के पीछे का रंग या पैटर्न जटिल होने पर भी यह पहले से बेहतर पढ़ता है। इसने उसी डेटासेट को और समृद्ध करके प्रशिक्षण लिया और पहले न पढ़े जा सकने वाले अक्षरों को भी पढ़ लेता है[5]।

हाथ में आया Miwo

सर्वर पर चलने वाले मॉडल को मोबाइल फोन में डालना अगला कार्य था। इस तरह सामने आया ऐप Miwo है। इसे पहली बार 30 अगस्त 2021 को सार्वजनिक किया गया था[5]। जापानी में इसका अंकन *みゝ* है और अंग्रेजी में Miwo है[6]।

इसके उपयोग का तरीका सरल है। प्राचीन दस्तावेज़ की तस्वीर मोबाइल फोन के कैमरे से ली जाती है। कुछ ही देर में स्क्रीन पर आधुनिक अक्षरों में बदला हुआ परिणाम दिखाई देता है। स्लाइडर की मदद से मूल दस्तावेज़ और परिणाम को ऊपर-नीचे रखकर तुलना की जा सकती है। इसमें यह भी दिखाता है कि मूल रूप से वह कौन सी hentaigana थी। hentaigana पुरानी हरिगाना का रूप है, जिसका उपयोग अब नहीं किया जाता। परिणाम को टेक्स्ट फाइल के रूप में निर्यात करने की सुविधा भी है। खींची गई तस्वीरें सर्वर पर नहीं छोड़ी जाती हैं। यह Android और iPhone दोनों पर मुफ्त में उपलब्ध है[6]।

26 अक्टूबर 2022 को आए संस्करण 1.1 से पहचान मॉडल को Ruri में बदल दिया गया[5]। उसी महीने इस ऐप को 2022 का Good Design Award मिला[7]। उपयोग के आँकड़े लगातार बढ़ते गए। जारी होने के बाद से 3 सितंबर 2026 तक पहचानी गई छवियों की संख्या 3,523,373 दर्ज की गई[6]। इसका अर्थ है कि लोग पुस्तकालयों, मंदिरों और स्थानीय अभिलेखागारों में पुराने दस्तावेज़ों की तस्वीरें खींच रहे हैं।

नागरिक मलिकर पढ़ते हैं

मशीन पहले पढ़े और इंसान उसमें सुधार करे, यह तरीका अन्य जगहों पर भी काम आता है। 'Minna de Honkoku' नाम की एक ऑनलाइन लपियांतरण परियोजना है। प्राचीन लिखावट को आधुनिक अक्षरों में उतारने का काम लपियांतरण कहलाता है। इसका जापानी अंकन *みんなの転写* है और इसका अर्थ 'सब मलिकर लिखना' है। इसे 2017 में क्योटो विश्वविद्यालय के प्राचीन भूकंप अध्ययन समूह ने शुरू किया था। वर्तमान में इसका संचालन राष्ट्रीय इतिहास एवं लोक-संस्कृति संग्रहालय और टोक्यो विश्वविद्यालय का भूकंप अनुसंधान संस्थान मलिकर करते हैं[8]।

भागीदारी का तरीका सबके लिए खुला है। इंटरनेट ब्राउज़र के ज़रिए कोई भी व्यक्ति पुराने दस्तावेज़ों को लपिबिद्ध कर सकता है। स्क्रीन पर एआई द्वारा पढ़ा गया प्रारंभिक परिणाम दिखाई देता है। इंसान को केवल गलत अक्षरों को ठीक करना होता है। यह वेबसाइट मानविकी ओपन डेटा संयुक्त उपयोग केंद्र की कुजुशिजी पहचान सेवा का उपयोग करती है। अब तक 10,800 से अधिक प्रतिभागियों ने 4.8 करोड़ से अधिक अक्षरों को लपिबिद्ध किया है[8]।

लपिबिद्ध किए गए पाठ केवल यँ ही जमा नहीं होते। इंसान द्वारा सुधारे गए सही उत्तर दोबारा प्रशिक्षण सामग्री बन जाते हैं। प्रशिक्षण सामग्री बढ़ने से मॉडल बेहतर होता है। मॉडल बेहतर होने पर प्रारंभिक परिणाम अधिक सटीक आता है और इंसानी मेहनत कम लगती है। एक बार यह चक्र शुरू हो जाए तो काम में गति आ जाती है।

पढ़े गए दस्तावेज़ों का उपयोग भी व्यापक है। इस परियोजना की शुरुआत भूकंप शोधकर्ताओं ने की थी। पुरानी डायरियों और सरकारी अभिलेखों में भूकंप, सुनामी और अकाल की तारीखें तथा नुकसान दर्ज हैं। घसीट लपि की वजह से जो रिकॉर्ड बंद पड़े थे, उनके डेटा में बदलते ही अतीत की आपदाओं के चक्र की गणना संभव हो जाती है। सदियों पुराना ब्रश-सुलेख आज आपदा रोकथाम की सामग्री बन जाता है।

यहाँ तक की बात कागज़ पर लिखे जापानी प्राचीन दस्तावेज़ों की थी। KuroNet, Ruri और Miwo सभी कागज़ पर लिखी कुजुशिजी को पढ़ने वाले मॉडल हैं। कीचड़ में दबी काष्ठ पट्टिकाओं को इन मॉडलों द्वारा स्वचालित रूप से पढ़े जाने का कोई उदाहरण अब तक नहीं मिला है। काष्ठ पट्टिकाओं पर अक्षरों की संख्या कम होती है, आधार लकड़ी का होता है और स्याही काफी मटि चुकी होती है। प्रशिक्षण में उपयोग होने वाले सही डेटा का भंडार भी कागज़ी दस्तावेज़ों जितना नहीं बना है। यदि कोरिया और जापान की काष्ठ पट्टिकाओं का डेटा और एकत्र हो जाए, तो स्थिति बदल सकती है। फ़िलहाल यह केवल एक आशा है, कोई ठोस उपलब्धि नहीं।

संदर्भ सामग्री

- [1] Clanuwat, T., Lamb, A., & Kitamoto, A. (2019). KuroNet: Pre-modern Japanese Kuzushiji character recognition with deep learning. 2019 International Conference on Document Analysis and Recognition (ICDAR), 607-614. <https://doi.org/10.1109/ICDAR.2019.00103>
- [2] Lamb, A., Clanuwat, T., & Kitamoto, A. (2020). KuroNet: Regularized residual U-Nets for end-to-end Kuzushiji character recognition. SN Computer Science, 1, 177. <https://doi.org/10.1007/s42979-020-00186-z>
- [3] Clanuwat, T., Bober-Irizar, M., Kitamoto, A., Lamb, A., Yamamoto, K., & Ha, D. (2018). Deep learning for classical Japanese literature. arXiv. <https://arxiv.org/abs/1812.01718>
- [4] ROIS-DS 人文学オープンデータ共同利用センター. KuroNetくずし字認識サービス(AI OCR). CODH. <https://codh.rois.ac.jp/kuronet/> (3 सितंबर 2026 को अभिगमिति)
- [5] ROIS-DS 人文学オープンデータ共同利用センター. みを(miwo)とは? CODH. <https://codh.rois.ac.jp/miwo/about/> (3 सितंबर 2026 को अभिगमिति)
- [6] ROIS-DS 人文学オープンデータ共同利用センター. みを(miwo): AIくずし字認識アプリ. CODH. <https://codh.rois.ac.jp/miwo/> (3 सितंबर 2026 को अभिगमिति)
- [7] 国立情報学研究所. (2022, 10월 26일). 世界初のAIくずし字認識アプリ「みを(miwo)」が2022年度グッドデザイン賞を受賞. 国立情報学研究所. <https://www.nii.ac.jp/news/release/2022/1026.html> (3 सितंबर 2026 को अभिगमिति)
- [8] みんなで翻刻. みんなで翻刻について. <https://honkoku.org/> (3 सितंबर 2026 को अभिगमिति)

8.4. सामने आया सच

Wolji की काष्ठ पट्टिकाओं से पता चला राजघराने का कामकाज

ग्योगजू का Donggung और Wolji एकीकृत सलि राजवंश का उप-महल और तालाब है। 1975 में Wolji की खुदाई में पहली बार काष्ठ पट्टिकाएँ मिलीं। इसके बाद भी अन्वेषण जारी रहा और उनकी संख्या बढ़ती गई। राष्ट्रीय ग्योगजू संग्रहालय यहाँ से मलि अवशेषों को Wolji दीर्घा में प्रदर्शित करता है[1]।

Wolji की काष्ठ पट्टिकाएँ राजघराने के दैनिक जीवन की झलक दिखाती हैं। कुछ काष्ठ पट्टिकाएँ मर्तबानों पर बाँधे जाने वाले लेबल (टैग) थीं। लेबल वाली काष्ठ पट्टियों के सरिँ पर दोनों तरफ खाँचे कटे होते थे। उन खाँचों पर रस्सी लपेटकर उन्हें सामान से बाँधा जाता था। इनका काम आज की कूरियर पर्ची जैसा ही था। Wolji से मिली काष्ठ पट्टियों पर nokgap और ong जैसे शब्द लिखे हैं। nokgap हरिण के मांस से बना कण्वित व्यंजन है और ong का अर्थ मर्तबान है। इसका मतलब है कि सलि राजघराना अचार के प्रकार और उसे तैयार करने का समय लिखकर प्रबंधन करता था[2]।

ऐसे रिकॉर्ड इतिहास की कतिबों में नहीं मिलते। Samguk Sagi में राजाओं, युद्धों और प्रशासनिक व्यवस्थाओं को दर्ज किया गया था। रसोई में क्या सामान पहुँचा, यह उसमें नहीं लिखा था। काष्ठ पट्टिकाएँ इस कमी को पूरा करती हैं। कीचड़ में दबा लकड़ी का एक टुकड़ा राजमहल का बहीखाता बन जाता है।

कोरियाई काष्ठ पट्टिका सोसायटी और राष्ट्रीय ग्योगजू संग्रहालय ने 3 और 4 अप्रैल 2026 को एक अंतरराष्ट्रीय सम्मेलन आयोजित किया। इसका विषय 'सलि की राजधानी ग्योगजू का परविश और लिखित जीवन' था। यह सम्मेलन ग्योगजू हवाबेक कन्वेंशन सेंटर में दो दिनों तक चला। इसमें Wolseong खाई और Donggung तथा Wolji से मिली सामग्रियों को एक साथ रखा गया। यह काष्ठ पट्टिकाओं, शिलालेखों और स्याही के अभिलेखों के ज़रिए सलि की राजधानी के स्वरूप को दोबारा संजोने का प्रयास था। इसमें कोरिया और जापान के शोधकर्ताओं ने मलिकर प्रस्तुतियाँ दीं[3]।

Seongsan कलि की काष्ठ पट्टिका में सज़ा की रपौर्त

हामान का Seongsan कलि सलि द्वारा छठी शताब्दी में नर्मित एक पहाड़ी कलि है। कलि की दीवार के भीतरी दलदल से लगातार काष्ठ पट्टिकाएँ मलिती रहीं। यहाँ से प्राप्त स्याही वाली काष्ठ पट्टिकाओं की संख्या अगस्त 2025 तक 247 दर्ज की गई। यह कोरिया में खोजी गई प्राचीन काष्ठ पट्टिकाओं के आधे से भी अधिक हैं[4]।

18वें उत्खनन सर्वेक्षण में नई पढ़ी गई दो पट्टिकाएँ चीड़ की लकड़ी से बनी थीं। इनमें से एक चारों सतहों पर लिखी गई बहुफलकीय काष्ठ पट्टिका है। दूसरी दोनों तरफ लिखी गई द्विपक्षीय काष्ठ पट्टिका है[4]। बहुफलकीय पट्टिका के चारों में से तीन फलों पर प्रशासनिक विवरण दर्ज थे। यह किसी व्यक्ति को दी गई सज़ा का रिकॉर्ड है[5]। इनके बनने का समय छठी शताब्दी का मध्य माना जाता है[4]।

यह पाठोद्धार ऊपर बताई गई हाइपरस्पेक्ट्रल इमेजिंग से प्राप्त हुआ परिणाम है। आँखों और सामान्य तस्वीरों से जो अक्षर नहीं पहचाने जा सकते थे, वे पुनर्जीवित हो गए[4]। एक हजार से अधिक वर्षों तक कीचड़ में दबा रहा सलि के नचिले दर्जे के अधिकारी का प्रतविदन वाक्यों के रूप में वापस लौट आया। Seongsan कलि की काष्ठ पट्टिकाओं पर अनाज भेजने वाले कई क्षेत्रों के नाम भी दर्ज हैं। यह सामग्री दिखाती है कि छठी शताब्दी में सलि ने स्थानीय क्षेत्रों को कैसे बाँटकर शासन किया था। इससे यह भी समझा जा सकता है कि कौन सा रास्ते से कलि तक पहुँचा था।

Baekje राजमहल के दस्तावेज़ों का ढेर

Buyeo का Gwanbuk-ri स्थल Baekje के साबी काल के राजमहल का स्थान माना जाता है। राष्ट्रीय Buyeo सांस्कृतिक वरिसत अनुसंधान संस्थान ने 2024 और 2025 में 16वाँ उत्खनन सर्वेक्षण किया। यहाँ से 329 काष्ठ पट्टिकाएँ और उनसे संबंधित अवशेष मिले। इनमें लिखे हुए अक्षरों वाली काष्ठ पट्टिकाएँ 82 और लकड़ी की छीलन (sakseol) 247 हैं[6]। यह देश के किसी एक स्थल से मिली सबसे अधिक संख्या है और साबी काल की काष्ठ पट्टिकाओं में शुरुआती दौर की हैं[7]।

काष्ठ पट्टिकाओं में तथियाँ लिखी पट्टिकाएँ भी शामिल हैं। इनमें gyeongsinnyeong और gyehaenyeon नाम के वर्ष दर्ज हैं। इनमें क्रमशः 540 और 543 माना जाता है। इसका अर्थ है कि यह Baekje द्वारा राजधानी को Ungjin से Sabi ले जाने के ठीक बाद का रिकॉर्ड है। इसमें अधिकारियों के कार्मिक रिकॉर्ड, राजकीय बहीखाते और पदों के नाम भी पाए गए[7]। बाँस से बना Baekje का सुपरि वाद्ययंत्र भी पहली बार वास्तविक रूप में सामने आया[8]।

22 और 23 जुलाई 2026 को एक अंतरराष्ट्रीय सम्मेलन आयोजित किया गया। इसे राष्ट्रीय Buyeo सांस्कृतिक वरिसत अनुसंधान संस्थान, कोरियाई काष्ठ पट्टिका सोसायटी और Baekje सोसायटी ने संयुक्त रूप से आयोजित किया। इसका विषय 'Buyeo Gwanbuk-ri स्थल के वाद्ययंत्र और काष्ठ पट्टिकाएँ' था। यह सम्मेलन बसान के Haeundae-gu में हुआ। इसकी शुरुआत चुंगनाम राष्ट्रीय विश्वविद्यालय के मानद प्रोफेसर पार्क सून-बाल के मुख्य भाषण से हुई। तीन सत्रों में बारह शोधपत्र प्रस्तुत किए गए। कोरिया, चीन और जापान के शोधकर्ताओं ने मिलकर Gwanbuk-ri की काष्ठ पट्टिकाओं के अर्थ की समीक्षा की[8]।

शिलालेखों के सामने सावधानी

काष्ठ पट्टिकाओं पर काम करने वाला तरीका सभी प्राचीन अक्षरों पर लागू नहीं होता। जब आधार पत्थर में बदल जाता है, तो स्थिति बदल जाती है। Gwanggaeto शिलालेख इसका एक उदाहरण है। 414 में स्थापित इस शिलालेख पर Goguryeo का इतिहास उत्कीर्ण है। लंबे समय तक धूप-बारिश झेलने से शिलालेख की सतह पर अक्षर घिस चुके हैं। यहाँ स्याही को नहीं, बल्कि पत्थरों में खुदे हुए खाँचों को पढ़ना होता है।

इसलिए शोधकर्ता छाप प्रतियाँ (rubblings) की तुलना करते हैं। मूल पाषाण छाप वह प्रतियाँ हैं जो शिलालेख पर चूना पोतने से पहले ली गई थी। मूल पत्थर पर कागज़ रखकर सीधे थपथपा कर यह छाप तैयार की गई थी। अब तक ऐसी केवल दस के करीब प्रतियों की पुष्टि हुई है। 19वीं सदी के उत्तरार्ध में एक अलग रूप सामने आया। अक्षरों को स्पष्ट दिखाने के लिए शिलालेख पर चूना लगाकर ली गई 'चूना छाप' तैयार की गई। ssang-gu-gamuk-bon नाम की कई प्रतियाँ भी बनाई गई। इनमें शिलालेख पर कागज़ रखकर अक्षरों के चारों ओर रेखाएँ खींची जाती थीं। फिर उसके खाली स्थान को गाढ़ी स्याही से भरकर ऐसा दिखाया जाता था जैसे वह असली छाप हो। यह थपथपा कर ली गई छाप नहीं, बल्कि निकल करके बनाई गई प्रतिलिपि है। इस प्रक्रिया में अक्षर गलत उतर सकते हैं[9]। इसी कारण एक ही अक्षर अलग-अलग छाप प्रतियों में अलग दिखाई देने लगा।

विवाद के केंद्र में sinmyongyeon-jo का वाक्य है। इस वाक्य की व्याख्या प्राचीन कोरिया-जापान संबंधों के इतिहास से जुड़ने के कारण लंबे समय तक बहस का विषय बनी रही। मुख्य विवाद यह था कि क्या किसी इंसान के हाथ से अक्षरों को बदला गया था[10]। मूल अक्षरों को जानने के लिए प्रारंभिक मूल पाषाण छाप और विभिन्न चूना छापों की तुलना करनी होती है। इसके लिए प्रत्येक छाप को परखने का मानक तय करना पहला कदम है[9]।

ऐसी बातें प्रचलित हैं कि 3D मापन और एआई ने इस विवाद का समाधान कर दिया है। लेकिन जाँच करने पर अक्सर वह शोध ही नहीं मिलता जसि आधार बताया जाता है। काष्ठ पट्टिकाओं पर कारगर रही वधि शिलालेखों पर भी काम करेगी, यह उम्मीद वास्तविकी शोध परिणामों से मेल नहीं खाती। ऊपर बताई गई तकनीकों केवल अदृश्य चीजों को दृश्यमान बनाती हैं। जो सामने आया है उसका अर्थ तय करना इसके बाद का कार्य है।

इसलिए निष्कर्ष दो स्तरों वाला है। इन्फ्रारेड और हाइपरस्पेक्ट्रल इमेजिंग ने सलि और Baekje की काष्ठ पट्टिकाओं से गायब हो चुकी स्याही को जीवित किया। बिना विभाजन वाले पहचान मॉडल ने जापानी कागज़ी प्राचीन दस्तावेज़ों की न कटने वाली घसीट लपिको पढ़ा। दोनों तकनीकों के विषय अलग हैं, लेकिन उनकी दशा एक है। इंसानी आँख से छूटे स्ट्रोक को मशीन पहले पहचान लेती है। इसके साथ ही, यदि मशीन द्वारा उकड़े गए स्ट्रोक और मूल में बचे स्ट्रोक के बीच भेद करने का अनुशासन न हो, तो वही तकनीक ऐसा इतिहास रच देगी जो कभी था ही नहीं। सुलेख को जगाने वाली मशीन है, लेकिन उस लिखावट की ज़िम्मेदारी उठाने वाला इंसान है।

संदर्भ सामग्री

- [1] 국립경주박물관. 월지판. 국립경주박물관. <https://gyeongju.museum.go.kr/kor/html/sub02/0201.html> (3 सितंबर 2026 को अभिगमति)
- [2] 국사편찬위원회. 목간. 우리역사넷. https://contents.history.go.kr/mobile/tt/view.do?levelId=tt_b36 (3 सितंबर 2026 को अभिगमति)
- [3] 경북일보. (2026, 3월 31일). 1500년 전 신라의 목소리 깨어난다...경주서 국제 목간 학술대회. 경북일보. <https://www.kyongbuk.co.kr/news/articleView.html?idxno=4068614> (3 सितंबर 2026 को अभिगमति)
- [4] 경남일보. (2025, 8월 7일). 함안 성산산성서 새로운 목간 2점 명문 확인. 경남일보. <https://www.gnnews.co.kr/news/articleView.html?idxno=616028> (3 सितंबर 2026 को अभिगमति)
- [5] 경남도민일보. (2025, 8월 7일). 누군가를 처벌한 신라 행정 보고, 함안 성산산성 목간서 확인. 경남도민일보. <https://www.idomin.com/news/articleView.html?idxno=943461> (3 सितंबर 2026 को अभिगमति)
- [6] 파이낸셜뉴스. (2026, 2월 5일). 부여 관북리 '백제 피리' 첫 확인..목간 329점 무더기 발견. 파이낸셜뉴스. <https://www.fnnews.com/news/202602051448564101> (3 सितंबर 2026 को अभिगमति)
- [7] 경향신문. (2026, 2월 17일). 부여 관북리서 발견된 목간 329점, '백제 피리'만큼이나 중요한 이유. 경향신문. <https://www.khan.co.kr/article/202602171427001> (3 सितंबर 2026 को अभिगमति)
- [8] 이데일리. (2026, 7월 21일). 백제 왕궁터 출토 목간 329점·횡적 발굴 성과, 국제학술대회서 조명. 이데일리. <https://edaily.co.kr/News/Read?mediaCodeNo=257&newsId=03276726645516816> (3 सितंबर 2026 को अभिगमति)
- [9] 국립중앙박물관. 광개토대왕비 탁본. 국립중앙박물관 큐레이터 추천 소장품. <https://www.museum.go.kr/MUSEUM/contents/M0501000000.do?schM=view&relicRecommendId=254462> (3 सितंबर 2026 को अभिगमति)
- [10] 국사편찬위원회. 광개토대왕릉비. 우리역사넷. https://contents.history.go.kr/mobile/kc/view.do?levelId=kc_r100300 (3 सितंबर 2026 को अभिगमति)

अध्याय 9. एजयिन सागर की खोई पहली: लीनयिर ए और ज्यामतीय ग्रडि

9.1. केस फाइल

क्रीट द्वीप के महल से नकिली तीन लपियिँ

एजयिन सागर के दक्षिण में क्रीट द्वीप है। कांस्य युग में इस द्वीप पर मनीअन सभ्यता का विकास हुआ। मनीअन लोग नावे बनाकर भूमध्य सागर में आते-जाते थे। वे जैतून का तेल, शराब और अनाज ढोते थे। महल इन चीज़ों को इकट्ठा करता था और बाद में लोगों में बाँट देता था।

1900 के आसपास ब्रिटिश पुरातत्वविद् आर्थर इवांस ने क्नोसोस महल के स्थल की खोई की। उन्होंने ज़मीन के अंदर से अक्षरों से उकेरी गई मट्टी की पट्टिकाएँ खोजीं। वे अक्षर केवल एक प्रकार के नहीं थे। इवांस ने मली लपियिँ को तीन वर्गों में बाँटा। वे क्रीट की चतुरलपि, लीनयिर ए और लीनयिर बी हैं[1]।

मनीअन सभ्यता एक हजार से अधिक वर्षों तक चली। फिर कांस्य युग के उत्तरार्ध में इसमें बड़ा बदलाव आया। उत्तर में स्थिति सेटोरनी द्वीप पर ज्वालामुखी का फटना उनमें से एक था। इस वसिफोट का समय ईसा पूर्व 17वीं सदी से 16वीं सदी के बीच माना जाता है। सटीक काल पर आज भी विवाद जारी है[7]। माना जाता है कि ज्वालामुखी की राख क्रीट द्वीप तक उड़कर आई थी और सुनामी तटों से टकराई थी। ईसा पूर्व 1450 के आसपास माइसीनयिन यूनानियों ने इस द्वीप पर अधिकार कर लिया[1]। ज्वालामुखी और माइसीनयिन शासन के बीच सौ साल से अधिक का समय बीता था। इन दोनों घटनाओं को सीधे कारण और परिणाम के रूप में जोड़ना कठिन है। उसके बाद मनीअन लोगों की भाषा अभिलेखों से लुप्त हो गई।

तीनों लपियिँ के उपयोग का समय अलग-अलग है। लीनयिर ए का उपयोग लगभग ईसा पूर्व 1750 से ईसा पूर्व 1450 तक किया गया था। लीनयिर बी इसके बाद सामने आई। यह वह लपि है जिसका उपयोग माइसीनयिन यूनानियों के क्रीट द्वीप पर अधिकार के बाद हुआ। दोनों लपियिँ में कई ऐसे चिह्न समान हैं जिनका आकार मलिता-जुलता है[1]।

मट्टी की पट्टिकाओं का बचे रहना एक संयोग था। मनीअन लोग मट्टी की पट्टिकाओं को आग में नहीं पकाते थे। वे गीली मट्टी पर अक्षर उकेर कर सुखाते थे, फिर लगभग एक साल तक इस्तेमाल करके फेंक देते थे। लेकिन महल में आग लगने से मट्टी की पट्टिकाएँ चीनी मट्टी के बर्तनों की तरह पक गईं। केवल आग में पकी पट्टिकाएँ ही 3,500 वर्षों तक बची रहीं। आज जो सामग्री बची है, वह आग की ही देन है।

इवांस द्वारा वर्गीकृत तीन लपियिँ में से क्रीट की चतुरलपि सबसे पहले सामने आई। यह ऐसी लपि है जिसमें मानव सरि या बर्तन का आकार साफ दिखाई देता है। लीनयिर ए उस चित्र को रेखाओं में समेटे हुए रूप के नकिट है। दोनों लपियिँ लगभग दो सौ वर्षों तक एक ही काल में साथ-साथ उपयोग की गईं। ऐसा नहीं था कि एक लपि के लुप्त होने के बाद दूसरी लपि बनी। कुछ विद्वानों का विचार है कि लीनयिर ए क्रीट की चतुरलपि से प्रेरित होकर बनाई गई थी। फिर भी यह निश्चित रूप से नहीं कहा जा सकता कि तीनों लपियिँ एक सीधी रेखा में विकसित हुईं। यह स्पष्ट है कि उनके बीच कई चिह्न समान हैं, लेकिन उनके विकास की कड़ी अभी तक स्पष्ट नहीं हुई है[1]।

लीनयिर बी को पढ़ लिया गया

लीनयिर बी को 1950 के दशक में पढ़ लिया गया। यह उपलब्ध ब्रिटिश वास्तुकार माइकल वेटर्सि ने हासिल की। उन्होंने साबित किया कि लीनयिर बी में दर्ज भाषा यूनानी भाषा का ही एक प्राचीन रूप थी। शास्त्रीय विद्वान जॉन चैडविक ने उनके इस शोध में मदद की[1]।

वेटर्सि ने एक तालिका बनाकर शुरुआत की। इस तालिका में क्षैतजि पंक्तियों में व्यंजन और लंबवत स्तंभों में स्वर रखे गए थे[8]। ऐसी तालिका को शब्दांश ग्रडि कहा जाता है। यह एक ऐसा ढांचा था जिसमें प्रत्येक शब्दांश के बैठने के लिए पहले से खाने बनाए गए थे। वेटर्सि ने सांख्यिकी की मदद से यह दायरा संकरा किया कि कौन-सा चिह्न किस खाने में आएगा[1]।

उन्होंने इन तय खानों में क्नोसोस और फैस्टोस जैसे स्थानों के नाम बठिआए। जैसे ही स्थानों के नाम सटीक बैठे, बाकी चिह्न भी एक के बाद एक सुलझते चले गए। विद्वानों ने इसी विधिको लीनयिर ए पर भी आजमाकर देखा। उन्होंने लीनयिर बी के ध्वनिमानों को मलिते-जुलते चिह्नों पर सीधे लागू किया[3]। ध्वनितो नकिल आई। लेकिन वह ध्वनिकौन-सी भाषा थी, यह कोई नहीं समझ सका[1]।

ध्वनिमान लागू करने के बाद भी अर्थ सामने न आना एक अजीब स्थिति है। ऐसा इसलिए है क्योंकि अक्षरों को पढ़ने और भाषा को समझने में अंतर होता है। हंगुल लपिजानने वाला व्यक्ति किसी अपरचित विदेशी भाषा को भी हंगुल में लिखे जाने पर पढ़ सकता है। वह उसका उच्चारण तो कर सकता है, लेकिन अर्थ नहीं समझ पाता। लीनयिर ए का अध्ययन करने वाले विद्वानों की स्थिति भी ठीक ऐसी ही है।

दो भाषाओं में लिखा कोई पत्थर नहीं है

प्राचीन लिपियों को पढ़ते समय विद्वान प्रायः समानांतर दस्तावेज़ खोजते हैं। यह ऐसा दस्तावेज़ होता है जिसमें एक ही सामग्री दो भाषाओं में अलग-बगल लिखी होती है। ऐसे दस्तावेज़ को द्विभाषी शिलालेख कहा जाता है। मसिर की चित्रलिपि रोसेटा स्टोन की बदीलत सुलझ पाई थी। रोसेटा स्टोन पर एक ही राजाज्जा तीन लिपियों में उकेरी गई थी। उनमें से एक लिपि प्राचीन यूनानी भाषा की थी, जिसे विद्वान पहले से ही पढ़ना जानते थे।

मेसोपोटामिया की कीलाकार लिपिका मामला भी कुछ ऐसा ही था। ईरान की बेहसितून चट्टान पर तीन भाषाओं में उकेरा गया शिलालेख मौजूद था। प्रारंभ में इस शिलालेख की तीनों भाषाएँ पढ़ी नहीं जा सकती थीं। उनमें से प्राचीन फ़ारसी को सबसे पहले सुलझाया गया। ऐसा इसलिए हुआ क्योंकि इसमें चिन्ह कम थे और शब्दों के बीच अंतराल था। विद्वानों ने पहले सुलझाई गई प्राचीन फ़ारसी के आधार पर बाकी भाषाओं को पढ़ लिया।

क्रीट द्वीप से ऐसा कोई दस्तावेज़ नहीं मिला। ऐसा कोई अवशेष नहीं है जहाँ लीनियर ए के बगल में उसका यूनानी अनुवाद दर्ज हो। मनीअन लोगों द्वारा छोड़े गए लेख केवल उनकी अपनी ही भाषा में लिखे हैं। तुलना के लिए कोई दूसरा पाठ न होने से उसका अर्थ समझने का रास्ता बंद हो जाता है[1]।

सुराग बन सके ऐसा कोई चिन्ह भी दिखाई नहीं देता। मसिर की चित्रलिपि को पढ़ने वाले शांपोलियों के पास कारतूश मौजूद था। यह राजा के नाम के चारों ओर खींचा गया एक लंबा घेरा था। यह पहले से स्पष्ट था कि इस घेरे के भीतर लिखे अक्षर किसी व्यक्तिका नाम हैं। शांपोलियों ने यहाँ से कुछ अक्षरों के ध्वनिमान निकाले और बाकी को सुलझा लिया। लीनियर ए की मट्टी की पट्टिकाओं पर ऐसा कोई घेरा नहीं है[1]।

एक और समस्या भी है। वह यह की लीनियर ए में दर्ज भाषा की जड़ों का कोई पता नहीं है। विद्वान इस भाषा को मनीअन भाषा कहते हैं। मनीअन भाषा किस भाषा-परिवार की है, यह आज तक स्पष्ट नहीं हो सका है। कुछ विद्वान इसे भारोपीय भाषा-परिवार से जोड़ते हैं, तो कुछ पश्चिमी एशियाई मूल का मानते हैं। इनमें से अपेक्षाकृत बेहतर प्रमाण सामी भाषा परिवार और लाइसियाई भाषा के पक्ष में माने जाते हैं। लाइसियाई भाषा आज के तुर्किये के दक्षिण-पश्चिम में बोली जाने वाली एक प्राचीन भाषा थी[1]। लेकिन किसी भी पक्ष को विद्वानों की सर्वसम्मति नहीं मिल सकी है।

लीनियर ए केवल क्रीट द्वीप पर ही नहीं पाई जाती। यह एजियन सागर के विभिन्न द्वीपों और यूनान की मुख्य भूमि पर भी खोजी गई है[1]। ज्वालामुखी में दबे सेंटोरीनी द्वीप के अवशेषों से भी यह मिली है। मनीअन लोगों के समुद्री व्यापारिक मार्गों के साथ-साथ यह लिपि भी फैलती गई। यह इस बात का प्रमाण है कि किस युग में भूमध्य सागर कितने व्यापक स्तर पर आपस में जुड़ा हुआ था।

विद्वानों द्वारा लीनियर ए पर इतना ध्यान केंद्रित करने का कारण भी यही है। लीनियर ए यूरोपीय धरती पर मिली सबसे प्राचीन लिपियों में गिनी जाती है। मनीअन लोगों द्वारा अपनी भाषा में छोड़ा गया एकमात्र रिकॉर्ड यही है। यदि लीनियर ए सुलझ जाती है, तो यह स्पष्ट हो जाएगा कि वे क्या मानते थे और कनि चीज़ों का व्यापार करते थे। यूनानी पौराणिक कथाओं में दर्ज मनीटोर की कथा की जड़ें भी यहीं छपी हैं।

1,400 संक्षिप्त बही-खाते

बची हुई सामग्री की मात्रा भी पर्याप्त नहीं है। अभिलेख का अर्थ है पुरावशेषों पर उकेरा या चित्रित किया गया लेख। अब तक पहचाने गए लीनियर ए के अभिलेखों की संख्या लगभग 1,370 है। उनमें लिखे कुल चिन्हों को मिला दिया जाए, तो भी वे लगभग 7,400 ही बैठते हैं[1]। मसिर या मेसोपोटामिया के अभिलेखों की तुलना में यह मात्रा नगण्य है।

लेखों की लंबाई भी बहुत छोटी है। अधिकांश मट्टी की पट्टिकाएँ गोदाम के बही-खाते हैं। उनमें बस इतना दर्ज है कि क्या सामान कितनी संख्या में आया और किस कतिना दिया गया। वहाँ केवल व्यक्तियों के नाम, वस्तुओं के नाम और संख्याओं की कतारें हैं। वास्तविक वाक्य न के बराबर हैं। वाक्य न होने पर व्याकरण को समझने के साधन भी घट जाते हैं।

चिन्हों की संख्या गनिते समय श्रेणियों को अलग करना पड़ता है। शब्दांश दर्शाने वाले 97 चिन्ह हैं। वस्तु का पूरा नाम दर्शाने वाले लगभग 50 चिन्ह हैं। संख्याएँ लिखने के 4 चिन्ह और भिन्न लिखने के 17 चिन्ह हैं। दो चिन्हों को मिलाकर बने लगभग 150 संयोजनों की भी जानकारी है[6]। इन सभी को मिलाकर मानक चिन्हों की सूची 300 तक पहुँचती है[2]। इनमें से बार-बार आने वाले चिन्ह बहुत थोड़े हैं। बाकी चिन्ह बस कुछ ही बार आकर रह जाते हैं। कंप्यूटर से सांख्यिकी करवाने के लिए एक ही चिन्ह का बार-बार आना आवश्यक है। जो चिन्ह गनि-चुने मौकों पर ही आते हैं, सांख्यिकी उन्हें ठीक से नहीं पकड़ पाती।

इसलिए शोधकर्ताओं ने डेटा को नए सरि से तैयार करने का नरिणय लिया। यह हाथ से बनाई गई आकृतियों को कंप्यूटर द्वारा प्रयोग किए जा सकने वाले प्रारूप में बदलने का काम है[2]। 2026 में रोशनी की मदद से उकेरे गए नशानों को फरि से पढ़ने की एक विधि भी सामने आई[4]। यह वह तरीका है जिससे आँखों से छूट जाने वाले उथले खाँचे भी साफ़ देखने लगते हैं[5]। एजियन सागर की खामोशी को तोड़ने का प्रयास इस तरह डेटा को व्यवस्थित करने के साथ फरि से शुरू हुआ।

संदर्भ सामग्री

- [1] Braović, M., Krstinić, D., Štula, M., & Ivanda, A. (2024). A systematic review of computational approaches to deciphering Bronze Age Aegean and Cypriot scripts. *Computational Linguistics*, 50(2), 725-779. https://doi.org/10.1162/coli_a_00514
- [2] Salgarella, E., & Castellan, S. (2021). SigLA: The Signs of Linear A. A palaeographical database. In *Grapholinguistics in the 21st Century 2020: Proceedings (Grapholinguistics and Its Applications, Vol. 5, pp. 945-962)*. Fluxus Editions. <https://sigla.phis.me/> (3 सितंबर 2026 को अभिगमिति)
- [3] Younger, J. G. (n.d.). Linear A texts and inscriptions in phonetic transcription. University of Kansas. <http://people.ku.edu/~jyounger/LinearA/> (3 सितंबर 2026 को अभिगमिति)
- [4] Giorgi, L., & Lopez, S. (2026). Aegean Bronze Age writing systems through RTI: New potential and perspectives. *Digital Applications in Archaeology and Cultural Heritage*, 41, e00508. <https://www.sciencedirect.com/science/article/abs/pii/S2212054826000159> (3 सितंबर 2026 को अभिगमिति)
- [5] GreekReporter. (2026, 3월 14일). New technology reveals hidden secrets of 3,500-year-old Bronze Age Aegean scripts. [greekreporter.com. https://greekreporter.com/2026/03/14/new-technology-reveals-hidden-secrets-aegean-bronze-age-scripts/](https://greekreporter.com/2026/03/14/new-technology-reveals-hidden-secrets-aegean-bronze-age-scripts/) (3 सितंबर 2026 को अभिगमिति)
- [6] Del Frio, M. (2022). Linear A. Mnamon: Ancient writing systems in the Mediterranean. *Scuola Normale Superiore*. <https://doi.org/10.25429/sns.it/lettere/mnamon022> (3 सितंबर 2026 को अभिगमिति)
- [7] Pearson, C. L., Brewer, P. W., Brown, D., Heaton, T. J., Hodgins, G. W. L., Jull, A. J. T., Lange, T., & Salzer, M. W. (2018). Annual radiocarbon record indicates 16th century BCE date for the Thera eruption. *Science Advances*, 4(8), eaar8241. <https://doi.org/10.1126/sciadv.aar8241>
- [8] Stavrakakis, E. (2024). Ventris' Grids. *Bulletin of the Institute of Classical Studies*, 67(1), 86-103. <https://doi.org/10.1093/bics/qbae033>

9.2. वैज्ञानिक जांच दल की पड़ताल: अपर्यवेक्षति ध्वन्यात्मक ग्रडि मलिन

शब्दकोश के बना जोड़ी खोजने वाली मशीन

अनुवादक बनाने के लिए आमतौर पर समानांतर वाक्यों की जरूरत होती है। इसमें कोरियाई और अंग्रेजी वाक्यों को आमने-सामने रखकर मशीन को दिखाया जाता है। ऐसी सामग्री को समानांतर कॉर्पस कहा जाता है। लीनयिर ए के लिए कोई समानांतर कॉर्पस उपलब्ध नहीं है[3]।

इसलिए शोधकर्ताओं ने बना उत्तर कुंजी के सीखने का तरीका चुना। यह पहले से सही उत्तर दिए बिना केवल सांख्यिकी की सहायता से नयिम खोजने का तरीका है। इस वधिको अपर्यवेक्षति शिक्षण कहा जाता है। मशीन यह गिनती करती है कि चिन्ह कसि क्रम में एक-दूसरे के साथ आते हैं। फिर वह अन्य लिपि प्रणालियों के आंकड़ों से तुलना करके उनकी जोड़ी तलाशती है।

2019 में अमेरिका के मैसाचुसेट्स इंस्टीट्यूट ऑफ टेक्नोलॉजी के शोधकर्ताओं ने यह तरीका पेश किया। यह जियामिगि लुओ, युआन काओ और रेजिनी बारजलि द्वारा मलिकर बनाया गया एक मॉडल है। इन्होंने दो लिपि प्रणालियों को जोड़ने के कार्य को 'न्यूनतम लागत प्रवाह समस्या' में बदल दिया। यह चिन्हों के बीच जलमार्ग बनाकर पानी के बहने की लागत तय करने की विधि है। जिस जलमार्ग में सबसे कम लागत आती है, उसके सही जोड़ी होने की संभावना सबसे अधिक होती है[1]।

शोधकर्ताओं ने इस मॉडल को युगारटिक लिपि पर लागू करके देखा। इसने पछिले सर्वश्रेष्ठ रकिर्ड को 5 प्रतिशत अंक से पीछे छोड़ दिया। उन्होंने इसे लीनयिर बी पर भी आजमाया। इसने यूनानी भाषा के समान मूल वाले 67.3 प्रतिशत शब्दों का सही मलिन किया[1]। अंतर केवल इतना था कि लीनयिर बी का उत्तर पहले से ज्ञात था। शोधकर्ता पहले से जानते थे कि वह लिपि यूनानी भाषा को दर्शाती थी।

शोध दल द्वारा लीनयिर बी को परीक्षण के लिए चुनने का एक ठोस कारण था। मॉडल के प्रदर्शन को तभी परखा जा सकता है जब समस्या का उत्तर ज्ञात हो[1]। यदि इसे सीधे किसी अनसुलझी समस्या पर लागू किया जाए, तो परिणाम सही है या नहीं, यह जानने का कोई रास्ता नहीं रहता। इसलिए पहले से पढ़ी जा चुकी लिपि से क्षमता नापी जाती है और फिर अपठित लिपि की ओर कदम बढ़ाया जाता है[3]।

मॉडल जो परिणाम देता है, वह कोई पूरा अनुवाद नहीं होता। वह प्रत्येक चिन्ह के लिए संभावित ध्वनिमानों के विकल्पों की एक सूची होती है[1]। मनुष्य इस सूची को लेकर पुरावशेषों से मलिन करता है और अनुपयुक्त विकल्पों को छोटता है। कंप्यूटर का काम मनुष्य के देखने योग्य कार्यभार को घटाना है[3]।

ध्वनियों के स्वरूप को नरिदेशांको में बदलना

लीनयिर ए की स्थिति भिन्न है। इसकी तुलना किस भाषा से की जाए, पहले तो यही स्पष्ट नहीं है[3]। 2021 में उसी प्रयोगशाला ने इस परिस्थिति से निपटने वाला एक मॉडल प्रस्तुत किया। इसे लुओ, फ्रेडरिक हार्टमैन और एनरिको सैटस ने बारज़लि तथा काओ के साथ मलिकर बनाया था[2]।

यह मॉडल एक साथ दो कठिनाइयों से निपटता है। पहली कठिनाई यह है कि शब्दों की सीमाएँ स्पष्ट नहीं हैं। प्राचीन लेखों में अक्सर शब्दों के बीच खाली जगह नहीं छोड़ी जाती थी। दूसरी कठिनाई यह है कि इसकी किसी सजातीय भाषा का निर्धारण नहीं हुआ है[2]।

शोधकर्ताओं ने ध्वनियों की विशेषताओं को संख्यात्मक नरिदेशांको में बदल दिया। अंतर्राष्ट्रीय ध्वन्यात्मक वर्णमाला (IPA) के आधार पर उन्होंने व्यंजनों और स्वरों की विशेषताओं को अक्षरों पर स्थापित किया। होठों से निकलने वाली ध्वनियों के नरिदेशांक आपस में निकट होते हैं। जीभ की जड़ से निकलने वाली ध्वनियों के नरिदेशांक उनसे दूर होते हैं। इन नरिदेशांको की मदद से मशीन द्वारा बेमेल जोड़ियाँ बनाने की संभावना बहुत कम हो जाती है[2]।

शोधकर्ताओं ने इस मॉडल का परीक्षण गोथिक और युगारटिक लिपि पर किया। इसके बाद उन्होंने इसे अभी तक अनसुलझी इबेरियन भाषा पर आजमाया। इबेरियन लिपि के चिन्हों के ध्वनिमान काफी हद तक ज्ञात हैं। जो बात नहीं सुलझी है, वह यह है कि उन ध्वनियों में कौन-सी भाषा दर्ज की गई थी। शोध दल को ऐसा कोई प्रमाण नहीं मिला कि इबेरियन भाषा बास्क भाषा की सजातीय है। यह परिणाम अकादमिक जगत की मुख्यधारा की राय से मेल खाता था। उन्होंने यह भी प्रदर्शित किया कि किसी भाषा के संभावित सजातीय विकल्पों की तुलना किस प्रकार की जाती है[2]। यही विधि लीनयिर ए के शोध में भी पूरी तरह लागू की जा सकती है[3]।

खाली खानों से मलिने वाले नयिम

मशीन केवल चिन्हों के आकार की ही तुलना नहीं करती। चिन्ह अपने पड़ोसी चिन्हों के साथ किस प्रकार जुड़ते हैं, इस प्रवृत्ति की भी गणना की जाती है। कुछ चिन्ह केवल शब्दों के आरंभ में आते हैं। कुछ चिन्ह केवल अंत में ही लगते हैं। कभी-कभी दो चिन्ह कभी भी एक साथ पास-पास नहीं आते[3]।

साथ आने की इस आवृत्ति को सह-उपस्थिति सांख्यिकी कहा जाता है। सह-उपस्थिति के आंकड़े जुटाने से तालिका के खाली स्थान सामने आ जाते हैं। एक ही क्षैतिज पंक्ति में रखे गए चिन्हों में समान व्यंजन होने की संभावना अधिक होती है। एक ही लंबवत स्तंभ में रखे गए चिन्हों में समान स्वर होने की संभावना अधिक होती है[8]। जो कार्य वेटर्सि ने अपने हाथों से किया था, कंप्यूटर उसी प्रक्रिया को कहीं अधिक तेजी से दोहराता है[3]।

कंप्यूटर इस गणना को मनुष्य की तुलना में बहुत तेजी से समाप्त कर देता है। यदि 300 चिन्ह हों, तो जोड़ियों की संख्या हजारों में पहुँच जाती है। कागज़ पर लिखकर गिनना किसी मनुष्य के लिए अत्यंत भारी काम है। वेटर्सि ने इस कार्य को कई वर्षों के परिश्रम से पूरा किया था[3]।

यहाँ तक का काम ध्वनियों की संरचना तैयार करने का है। संरचना बन जाने पर भी यह स्पष्ट नहीं होता कि उन ध्वनियों का अर्थ क्या है। ग्रंडि को भरना और भाषा को समझना दो पूरी तरह अलग बातें हैं[3]।

अक्षरों के आकार को डेटा में बदलना

मशीन से सांख्यिकी करवाने के लिए डेटा मशीन-पठनीय प्रारूप में होना चाहिए। कागज़ की कतिबों में छपी तस्वीरों मात्र से गणना नहीं कराई जा सकती। इस काम को SigLA नामक डेटा भंडार ने संभाला है। इसे कैम्ब्रिज की एस्तेर सालगारेला और रेनेस के समीर कास्टेलान ने मलिकर तैयार किया था[4]।

SigLA में 300 मानक चिह्न और हाथ से उतारे गए लगभग 400 अभिलेख शामिल हैं। इन अभिलेखों में 3,000 से अधिक व्यक्तिगत चिह्न देखे जा सकते हैं। डेटा और चित्र किसी के भी उपयोग के लिए सार्वजनिक रूप से उपलब्ध हैं। किसी एक चिह्न के स्रोत को अलग करके तुलना भी की जा सकती है[4]।

मार्च 2026 में सामने आए एक अध्ययन ने डेटा के दायरे को और विस्तृत किया। लवनीया जोर्गी और सारा लोपेज़ ने एजियन लिपि पर परावर्तन रूपांतरण इमेजिंग तकनीक लागू की। यह एक ही पुरावशेष पर विभिन्न दशाओं से प्रकाश डालकर कई तस्वीरें लेने की विधि है। इन तस्वीरों को मलिन पर सतह के सूक्ष्म उभार उभरकर सामने आते हैं। यहाँ तक कि यह भी जांचा जा सकता है कि उत्कीर्णक ने कौन सा स्ट्रोक पहले खींचा था[7]।

जब ऐसा डेटा इकट्ठा होता है, तो यह अन्य लिपियों के लिए भी उपयोगी हो जाता है। 2022 में बोलोग्ना विश्वविद्यालय के शोधकर्ताओं ने साइप्रो-मनीअन लिपि का पुनर्मूल्यांकन किया। यह मकिले कोरात्सा, फेबियो तामबुरिनी, मगिुल वालेरियो और सलिविया फेरारा द्वारा किया गया संयुक्त शोध था। उन्होंने पूर्व ज्ञान के बिना न्यूरल नेटवर्क का उपयोग करके चिह्नों के आकारों की तुलना की। इसके परिणामस्वरूप

उन्होंने नषिकर्ष नकाला कइसे तीन समूहों में वभिजति करने वाले मौजूदा वर्गीकरण की समीक्षा की जानी चाहिए। आंकड़ों के अनुसार इसे एक ही लिपि प्रणाली के रूप में देखना अधिक उचित है[5]।

कंप्यूटर द्वारा दिया गया परिणाम तुरंत शोध-पत्र नहीं बन जाता। पुरातत्वविद् और भाषाविद् इसकी दोबारा समीक्षा करने की प्रक्रिया से गुजरते हैं[3]। वे मलिन करते हैं कि कौन सी मट्टी की पट्टिका किस कमरे से मली थी। वे स्ट्रोक की आदतों की तुलना करके यह देखते हैं कि क्या यह एक ही लिपिकी द्वारा लिखा गया था[7]। इस समीक्षा को पार करने के बाद ही इसे अकादमिकी सामग्री में शामिल किया जाता है[3]।

एजियन सागर की तीनों लिपियों को एक साथ संभालने वाले शोध भी जारी हैं। यह क्रीट की चतुरलिपि, लीनियर ए और साइप्रो-मनीअन लिपि को एक साथ रखने की पद्धति है। तीनों लिपियों की तुलना करने पर ऐसे नियम सामने आते हैं जो केवल एक लिपि को देखने पर नहीं दिखते थे। कंप्यूटर इस प्रकार की व्यापक तुलना में अपनी शक्ति दिखाता है[3]।

वह काम जो मशीनें नहीं कर सकती

इन तरीकों ने लीनियर ए का अर्थ स्पष्ट नहीं किया है। पार करने के लिए अभी भी तीन बाधाएं बची हैं।

पहली बात तुलना के लिए समकक्ष का होना है। अपठित लिपि और ज्ञात लिपि के बीच कोई संबंध होने पर ही अनसुपरवाइज्ड अलाइनमेंट प्रभावी होता है[1]। मनीअन भाषा की कोई जीवित संबंधी भाषा नहीं बची है[3]। निर्देशांक तो बन जाते हैं, लेकिन उन निर्देशांकों को स्थिर करने के लिए कोई संदर्भ बटु नहीं है।

दूसरी बात डेटा की मात्रा है। 1,370 अभिलेख मशीन लर्निंग में उपयोग के लिए बहुत कम हैं[3]। डेटा कम होने पर बड़े न्यूरल नेटवर्क प्रशंसनीय भ्रम गढ़ लेते हैं। इस प्रकार की त्रुटि को मतभ्रम (हैलुसिनेशन) कहा जाता है।

तीसरी बात सत्यापन है। जुलाई 2026 में The Conversation में प्रकाशित एक लेख में इस समस्या को रेखांकित किया गया। यह डबलनि सटी विश्वविद्यालय की जेन एडकनिस् द्वारा लिखा गया लेख है। उन्होंने लिखा कि सांख्यिकी द्वारा पैटर्न मिलाना गैर-मौजूद अर्थ उत्पन्न नहीं कर सकता। उनका कहना है कि ज्ञात भाषा या द्विभाषी दस्तावेज जैसी बुनियाद होने पर ही संयोग और वास्तविकी नियम के बीच अंतर किया जा सकता है। मशीन केवल संभावित विकल्पों को कम करती है, और सत्यता परखने का काम मनुष्यों के जम्मे रहता है[6]।

संदर्भ सामग्री

- [1] Luo, J., Cao, Y., & Barzilay, R. (2019). Neural decipherment via minimum-cost flow: From Ugaritic to Linear B. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Association for Computational Linguistics. <https://aclanthology.org/P19-1303/> (3 सितंबर 2026 को अभिगमिति)
- [2] Luo, J., Hartmann, F., Santus, E., Barzilay, R., & Cao, Y. (2021). Deciphering undersegmented ancient scripts using phonetic prior. Transactions of the Association for Computational Linguistics, 9, 69-81. https://doi.org/10.1162/tacl_a_00354
- [3] Braović, M., Krstinić, D., Štula, M., & Ivanda, A. (2024). A systematic review of computational approaches to deciphering Bronze Age Aegean and Cypriot scripts. Computational Linguistics, 50(2), 725-779. https://doi.org/10.1162/coli_a_00514
- [4] Salgarella, E., & Castellan, S. (2021). SigLA: The Signs of Linear A. A palaeographical database. In Grapholinguistics in the 21st Century 2020: Proceedings (Grapholinguistics and Its Applications, Vol. 5, pp. 945-962). Fluxus Editions. <https://sigla.phis.me/> (3 सितंबर 2026 को अभिगमिति)
- [5] Corazza, M., Tamburini, F., Valério, M., & Ferrara, S. (2022). Unsupervised deep learning supports reclassification of Bronze age cypriot writing system. PLOS ONE, 17(7), e0269544. <https://doi.org/10.1371/journal.pone.0269544>
- [6] Adkins, J. (2026, 7월 27일). Cracking the code: can AI help us decipher ancient languages? The Conversation. <https://theconversation.com/cracking-the-code-can-ai-help-us-decipher-ancient-languages-288238> (3 सितंबर 2026 को अभिगमिति)
- [7] Giorgi, L., & Lopez, S. (2026). Aegean Bronze Age writing systems through RTI: New potential and perspectives. Digital Applications in Archaeology and Cultural Heritage, 41, e00508. <https://www.sciencedirect.com/science/article/abs/pii/S2212054826000159> (3 सितंबर 2026 को अभिगमिति)
- [8] Stavrakakis, E. (2024). Ventris' Grids. Bulletin of the Institute of Classical Studies, 67(1), 86-103. <https://doi.org/10.1093/bics/qbae033>

9.3. सामने आया सच

संख्याएँ और भग्नि सुलझ गए

लीनियर ए में सबसे पहले सुलझने वाला हस्सा संख्याओं की गणना पद्धति थी। मनीअन लोग दस के समूहों में गनने वाली दशमलव प्रणाली का उपयोग करते थे। एक लंबवत रेखा 1 है और एक कर्षित रेखा 10 है। एक वृत्त 100 है, और करिणो वाला वृत्त 1000 है[3]। इस नियम का अनुमान केवल चित्रों को देखकर भी लगाया जा सकता था।

कठिन काम भग्नि के चिह्न थे। मट्टी की पट्टिकाओं पर सत्रह भग्नि चिह्न मिलते हैं[8]। दो चिह्नों को मिलाकर लिखे गए चिह्नों को भी गनि तो यह संख्या लगभग तीस हो जाती है[9]। लंबे समय तक यह नहीं सुलझ सका कि कौन सा चिह्न किस मान को दर्शाता है।

2021 में बोलोग्ना विश्वविद्यालय के शोधकर्ताओं ने इस समस्या पर काम किया। यह मकिले कोरात्सा, सलिविया फेरारा, बारबरा मोटेकी, फैबियो ताम्बुरिनी और मगुएल वालेरियो का संयुक्त शोध था। उन्होंने चिह्नों के आकार का परीक्षण करने की पारंपरिक विधि और कंप्यूटर गणना को एक साथ लागू किया[1]।

शोधकर्ताओं ने सबसे पहले उन नियमों को व्यवस्थित किया जिनका पालन पट्टिकाओं पर चिह्न करते थे। उन्होंने गनि कि कौन से चिह्न एक साथ आते हैं और कौन से चिह्न एक साथ नहीं आते हैं। फिर उन्होंने कंप्यूटर से मानों के संयोजनों की जांच की। संभावित विकल्प लगभग 40 लाख थे। सांख्यिकीय परीक्षणों और अन्य सभ्यताओं की भग्नि प्रणालियों की तुलना के माध्यम से उन्होंने इन विकल्पों को सीमित किया[1]।

अंतिम परिणाम इस प्रकार था। मनीअन भग्नि प्रणाली में सबसे छोटा मान $1/60$ था। इस प्रणाली से 60 को हर मानने वाले अधिकांश मान लिखे जा सकते थे। $1/10$ को दर्शाने वाला चिह्न लीनियर बी में भी वैसा ही जारी रहा। लीनियर बी में इस चिह्न का उपयोग सूखे अनाज को मापने की इकाई के रूप में किया गया था। दोनों लिपियों के अभिलेखों के बीच लगभग 250 वर्षों का अंतर है। उस समयांतराल के बावजूद यह गणना पद्धति जीवित रही[1]।

जब गणना पद्धति सुलझ गई, तो बहीखाते की संरचना भी दिखाई देने लगी। पट्टिकाओं में मदों के चिह्नों और उनकी मात्रा को जोड़कर लिखने का नियम अपनाया गया था। यही प्रारूप कई पुरातात्विक स्थलों पर दोहराया गया है। इसका अर्थ है कि क्रीट द्वीप के महल समान लेखांकन नियमों का साझा उपयोग कर रहे थे[3]। अर्थ न जानने पर भी दस्तावेज़ का स्वरूप इस प्रकार प्रकट होता है।

स्थानों के नाम तो पढ़े जाते हैं, लेकिन अर्थ नहीं सुलझा

ध्वनियों के संबंध में भी कुछ उपलब्धियाँ मिली हैं। अमेरिका के कंसास विश्वविद्यालय के जॉन यंगर ने लीनियर ए अभिलेखों के ध्वन्यात्मक प्रतिलिखनों की एक सूची सार्वजनिक की है। यह एक ऐसा प्रतिलिखन है जिसमें लीनियर बी में समान आकार वाले चिह्नों को समान ध्वनिमान देकर देखा गया है[2]।

इस प्रतिलिखन में कई ऐसे शब्द मिलते हैं जो क्रीट द्वीप के प्राचीन स्थान-नाम प्रतीत होते हैं। ये क्नोसोस और फेस्टोस जैसे नाम हैं जो बाद के युगों में भी बने रहे[2]। स्थान-नामों का मेल खाना यह दर्शाता है कि ध्वनिमान की परिकल्पना पूरी तरह गलत नहीं थी[3]।

ध्वनिमान की परिकल्पना की सीमाएं भी स्पष्ट हैं। भले ही दोनों लिपियों में आकार समान दिखे, लेकिन इसकी कोई गारंटी नहीं है कि ध्वनियाँ भी समान थीं। लीनियर बी बनाने वाले लोगों ने दूसरों की लिपि उधार लेकर अपनी भाषा लिखी थी। उधार लेने की प्रक्रिया में ध्वनियों में बदलाव आ गया होगा[3]। इसलिए प्रतिलिखन का उपयोग केवल अनंतिम पठन के रूप में ही किया जाता है[2]।

लेकिन सीमा यहीं तक है। स्थान-नामों को छोड़ दिया जाए, तो शेष शब्दों के अर्थ आपस में नहीं जुड़ते। भले ही ध्वन्यात्मक रूप से पढ़ा जा सके, पर यह ज्ञात नहीं होता कि इसका क्या अर्थ है। ध्वनि और अर्थ के बीच अभी कोई सेतु नहीं बना है[3]।

2026 की हलचल

2026 की गर्मियों में लीनियर ए को पढ़ लिए जाने की घोषणा हुई। यह एआई इंजीनियर टॉम डमिनी द्वारा किया गया दावा था। उन्होंने औपचारिक रूप से भाषाविज्ञान का अध्ययन नहीं किया है[4]।

डमिनी ने बताया कि उन्होंने Python स्क्रिप्ट लिखकर दो डेटा स्रोतों की पड़ताल की। ये हाथ से उतारी गई सामग्री को संकलित करने वाला GORILA और पहले देखा गया SigLA है। अपनी पहली घोषणा में उन्होंने 40 चिह्नों को ध्वनिमान दिए और 408 शब्दों की एक सूची तैयार की[4]। अगस्त में 41 पृष्ठों का एक पत्र जारी करते हुए उन्होंने इस सूची को 508 प्रवर्णितियों तक बढ़ा दिया। इसमें 42 चिह्नों का पठन और 443 अभिलेखों के अनुवाद शामिल थे। उन्होंने दावा किया कि मनीअन भाषा सेमिटिक भाषा परिवार से संबंधित है, जिसमें हब्रू और अरामी शामिल हैं[5]।

अकादमिक जगत ने इस दावे को स्वीकार नहीं किया। क्योंकि कोई भी सहकर्मी-समीक्षित शोध-पत्र प्रकाशित नहीं हुआ था[4]। बताया गया कि रेटगर्स विश्वविद्यालय और कैम्ब्रिज विश्वविद्यालय में इसकी समीक्षा चल रही है। लेकिन अभी तक कोई नबिकरण नहीं निकला है[5]। लीनियर ए को पढ़ लेने के दावे पछिले सौ वर्षों में कई बार किए जा चुके हैं। इसलिए शोधकर्ता नए दावों के प्रतिसावधानी बरतते हैं[6]।

उसी वर्ष अगस्त में एक अन्य व्याख्या भी सामने आई। यह माउंट जुक्लास के शिखर पर बने पवित्र स्थल से मली पत्थर की मेज के अभिलेख से संबंधित थी। यह देवताओं को मदरि अर्पित करने वाली वेदी थी। प्रस्ताव यह था कि इस पत्थर पर खुदे दो नाम बाद के ग्रीक देवताओं से जुड़े हैं। यह प्रस्ताव भी मात्र एक परकिल्पना बनकर रह गया है। क्योंकि लीनियर ए अभी तक नश्विचि रूप से हल नहीं हो पाई है[7]।

यह घटना संपूर्ण एआई अनुसंधान के लिए भी एक सबक छोड़ती है। मशीनें उन जगहों पर नयिमों को अच्छी तरह खोज लेती हैं जहाँ नयिम मौजूद होते हैं। भनिं चहिनो जैसे डेटा में परणाम मलिं, जहाँ तारकिक संगतकी आवश्यकता थी[1]। दूसरी ओर, ऐसी जगहें भी हैं जहाँ केवल नयिमों के सहारे नहीं पहुँचा जा सकता। कौन सी ध्वनिकिसि अर्थ से जुड़ी है, यह सामाजिक सहमतिका वषिय है। जब उस समझौते को जानने वाले सभी लोग लुप्त हो जाते हैं, तो केवल गणना द्वारा उसे पुनर्जीवित करने का मार्ग संकरा हो जाता है[6]।

शेष डेटा को खुला रखना

शोध के तौर-तरीके भी बदल गए हैं। पहले अभिलेखों के चित्र केवल महंगे कैटलॉग में ही प्रकाशित होते थे। केवल वही लोग आंकड़े तैयार कर सकते थे जो वे दुर्लभ पुस्तकें प्राप्त कर पाते थे। अब डेटा इंटरनेट पर सार्वजनिक रूप से उपलब्ध है। एक छात्र भी इसे डाउनलोड कर सकता है और चहिनो की गणना कर सकता है[3]।

डेटा खुला होने पर सत्यापन भी तीव्र हो जाता है। यदि कोई नया पठन प्रस्तावित करता है, तो दूसरा व्यक्ति उसी डेटा से उसकी पुष्टि कर सकता है[6]। 2026 की गर्मियों के हंगामे में भी सत्यापन प्रक्रिया इसी प्रकार आगे बढ़ी[4]। दावा करने वाला पक्ष और समीक्षा करने वाला पक्ष दोनों एक ही डेटा को देख रहे थे[5]।

डेटा सार्वजनिक होने के कारण शोध की बाधाएं भी कम हो गई हैं। विश्वविद्यालय से असंबद्ध व्यक्ति भी चहिनो के आंकड़े निकाल सकता है। नए दृष्टिकोण बढ़ने पर छूटे हुए नयिम भी पकड़ में आ जाते हैं। इसके विपरीत, असत्यापित दावे भी बढ़ जाते हैं। पारदर्शिता और सत्यापन को एक साथ चलना चाहिए[6]।

अभी तक न पढ़ी जा सकी लिपि

लीनियर ए के संदर्भ में एआई ने चार प्रमुख कार्य किए हैं। मशीन ने बखिरे हुए डेटा को एक जगह एकत्रित किया। इसने चहिनो के आकार और स्ट्रोक को संख्याओं में बदलकर उनकी तुलना करना संभव बनाया[3]। भनिं चहिनो के मान को कंप्यूटर गणना ने मनुष्यों के पठन में सहायता करते हुए सीमिति किया। इसने लाखों विकल्पों वाली समस्या को उस सीमा तक घटा दिया जिसकी समीक्षा मनुष्य कर सके[1]।

मशीन जो नहीं कर सकी, वह भी स्पष्ट है। मनिंअन किस प्रकार की भाषा थी, यह अभी भी ज्ञात नहीं हुआ है। रोसेटा स्टोन की भूमिका नभाने वाला कोई दस्तावेज या तो अभी भी जमीन के नीचे है या अस्तित्व में ही नहीं है[3]। ध्वनियों का प्रतिलिखन करने पर भी यह स्पष्ट नहीं होता कि वे शब्द किस ओर संकेत करते हैं[6]।

साइप्रो-मनिंअन लिपिका उदाहरण इस अंतर को भली-भांति दर्शाता है। नयूरल नेटवर्क ने यह परणाम दिया कि मौजूदा वर्गीकरण डगमगा रहा है[10]। लेकिन इसका अर्थ यह नहीं है कि उस लिपिको पढ़ लिया गया है। मशीन ने जो व्यवस्थित किया था, वह अक्षरों की संरचना थी, न कि शब्दों का अर्थ।

आगे दो ही रास्ते दिखाई देते हैं। एक रास्ता नए पुरावशेषों का मलिना है। दो भाषाओं में लिखा एक भी दस्तावेज पूरी स्थितिको बदल सकता है। दूसरा रास्ता डेटा को और अधिक बारीकी से परिष्कृत करना है। यदि एक-एक स्ट्रोक तक दर्ज कर लिया जाए, तो भविष्य के शोध उसी आधार पर शुरू होंगे[3]।

लीनियर ए आज भी एक अपठित लिपि है। मशीन ने उसके सामने दरवाजा नहीं खोला, बल्कि ताले की रूपरेखा खींच दी। दरवाजा खोलना नए पुरावशेषों और मानवीय नरिणय पर नरिभर करता है[6]। एजियन सागर की मट्टी की पट्टिकाएं अभी भी अपनी बात नहीं कह रही हैं।

संदर्भ सामग्री

- [1] Corazza, M., Ferrara, S., Montecchi, B., Tamburini, F., & Valério, M. (2021). The mathematical values of fraction signs in the Linear A script: A computational, statistical and typological approach. *Journal of Archaeological Science*, 125, 105214. <https://doi.org/10.1016/j.jas.2020.105214>
- [2] Younger, J. G. (n.d.). *Linear A texts and inscriptions in phonetic transcription*. University of Kansas. <http://people.ku.edu/~jyounger/LinearA/> (3 सितंबर 2026 को अभिगमिति)

- [3] Braović, M., Krstinić, D., Štula, M., & Ivanda, A. (2024). A systematic review of computational approaches to deciphering Bronze Age Aegean and Cypriot scripts. *Computational Linguistics*, 50(2), 725-779. https://doi.org/10.1162/coli_a_00514
- [4] GreekReporter. (2026, 7월 30일). AI engineer claims to have deciphered Linear A. [greekreporter.com](https://greekreporter.com/2026/07/30/ai-engineer-claim-decipher-linear-a/). <https://greekreporter.com/2026/07/30/ai-engineer-claim-decipher-linear-a/> (3 सितंबर 2026 को अभगिमति)
- [5] GreekReporter. (2026, 8월 13일). AI engineer claims to have deciphered Linear A, identifies Minoan language as Semitic. [greekreporter.com](https://greekreporter.com/2026/08/13/ai-engineer-claim-linear-a-minoans-semitic-language/). <https://greekreporter.com/2026/08/13/ai-engineer-claim-linear-a-minoans-semitic-language/> (3 सितंबर 2026 को अभगिमति)
- [6] Adkins, J. (2026, 7월 27일). Cracking the code: can AI help us decipher ancient languages? *The Conversation*. <https://theconversation.com/cracking-the-code-can-ai-help-us-decipher-ancient-languages-288238> (3 सितंबर 2026 को अभगिमति)
- [7] GreekReporter. (2026, 8월 14일). Linear A inscription points to Minoan origins of Greek gods Dionysos and Demeter. [greekreporter.com](https://greekreporter.com/2026/08/14/minoan-linear-a-inscription-greek-gods-dionysos-demeter/). <https://greekreporter.com/2026/08/14/minoan-linear-a-inscription-greek-gods-dionysos-demeter/> (3 सितंबर 2026 को अभगिमति)
- [8] Del Frio, M. (2022). Linear A. *Mnamon: Ancient writing systems in the Mediterranean*. Scuola Normale Superiore. <https://doi.org/10.25429/sns.it/lettere/mnamon022> (3 सितंबर 2026 को अभगिमति)
- [9] Salgarella, E., & Castellan, S. (2021). SigLA: The Signs of Linear A. A palaeographical database. In *Grapholinguistics in the 21st Century 2020: Proceedings (Grapholinguistics and Its Applications, Vol. 5, pp. 945-962)*. Fluxus Editions. <https://sigla.phis.me/> (3 सितंबर 2026 को अभगिमति)
- [10] Corazza, M., Tamburini, F., Valerio, M., & Ferrara, S. (2022). Unsupervised deep learning supports reclassification of Bronze age cypriot writing system. *PLOS ONE*, 17(7), e0269544. <https://doi.org/10.1371/journal.pone.0269544>

अध्याय 10. सूर्य चक्र और पशु मुहरों का कूट तोड़ना: सधु लपि और रोगोरोन्गो की ध्वनियाँ

10.1. केस फाइल

मुहर पर खुदे चार अक्षर

सधु सभ्यता लगभग 2600 ईसा पूर्व से लगभग 1900 ईसा पूर्व तक जारी रही[1]। इसका क्षेत्र आज का पाकिस्तान और उत्तर-पश्चिमी भारत है। शहरों के आकार और संख्या को देखते हुए यह कांस्य युग की प्रमुख सभ्यताओं में से एक है। कांस्य युग वह प्राचीन काल है जब लोहे की तुलना में कांस्य का अधिक उपयोग होता था। माना जाता है कि इसके शहर और गाँव 10 लाख वर्ग किलोमीटर से अधिक क्षेत्र में फैले हुए थे[2]।

1920 के दशक में हड़प्पा और मोहनजोदड़ो में बड़े पैमाने पर उत्खनन शुरू हुआ[1]। इस कार्य का संचालन ब्रिटिश पुरातत्वविद् जॉन मार्शल ने किया। वह भारतीय पुरातत्व सर्वेक्षण के प्रमुख थे। हड़प्पा स्थल की खुदाई 1921 में दयाराम साहनी ने की थी। मोहनजोदड़ो स्थल का कार्य 1922 में राखालदास बनर्जी ने संभाला। मार्शल ने इन उत्खननों को जोड़कर दुनिया के सामने एक नई सभ्यता के अस्तित्व की घोषणा की[3]।

जमीन के नीचे से ईंटों की सीधी सड़के और जल निकासी व्यवस्था मिली। और हथेली से भी छोटी कई चौकोर मुहरें बड़ी संख्या में मिली[1]।

केवल नगरों की बनावट से ही बहुत कुछ जाना जा सकता है। सड़के शतरंज के बोर्ड की तरह सीधी फैली थी और हर घर से नालियाँ जुड़ी थीं। ईंटों को एक निश्चित आकार में पकाया गया था। वजन मापने के बाटों के भी मानक थे। इसका अर्थ है कि यह योजना बनाने और नियमों का पालन करने वाला समाज था। लेकिन उनके छोड़े गए लेख पढ़े नहीं जा सके हैं।

मुहरें सेलखड़ी नामक मुलायम पत्थर से बनाई गई थीं। सेलखड़ी ऐसा पत्थर है जिस पर नाखून से भी खरोच लग जाती है। मुहर के नचिले हिस्से में पशुओं के चित्र उकेरे गए हैं। इनमें एक सींग वाला जानवर, कूबड़ वाला बैल और हाथी जैसे चित्र शामिल हैं। चित्र के ऊपर खाली जगह में प्रतीकों की एक पंक्ति बनी हुई है[1]।

मुहर दबाकर छाप लगाने वाली वस्तु है। गीली मट्टी पर दबाने से चित्र और प्रतीक ज्यों के त्यों उस पर छप जाते हैं। सामान बांधने वाली रस्सी पर गीली मट्टी लगाकर मुहर दबाने से वह सील बन जाती है। सधु सभ्यता के शहरों में मुहर लगी मट्टी की सीले भी मिली हैं[1]।

वद्वान इन प्रतीकों के समूह को सधु लपि या हड़प्पा लपिकहते हैं[1]। नाम में लपि शब्द आता है, लेकिन इसे अभी तक पढ़ा नहीं जा सका है। न इसकी ध्वनिपिता है और न इसका अर्थ। इसकी खोज हुए सौ साल से अधिक बीत चुके हैं, फिर भी स्थिति यही है[2]।

'पढ़ नहीं पाना' कहने के दो परस्पर जुड़े अर्थ हैं। पहला यह कि प्रतीकों की ध्वनि अज्ञात है। दूसरा यह कि प्रतीकों का अर्थ अज्ञात है। वर्तमान में दोनों ही अज्ञात हैं।

समस्या लंबाई की है। एक मुहर पर उकेरे गए प्रतीक आमतौर पर केवल चार या पाँच ही होते हैं[1]। एक फलक पर लिखा सबसे लंबा अभिलेख भी सत्रह अक्षरों से अधिक नहीं है। तीन फलकों पर अलग-अलग उकेरे गए सभी प्रतीकों को मिलाकर भी केवल छब्बीस अक्षर बनते हैं[4]। यह मात्रा कक्षा की कॉपी की एक पंक्ति को भी नहीं भर सकती।

मुहरों या बर्तनों पर उकेरे गए ऐसे लेखों को अभिलेख कहा जाता है। अब तक एकत्र किए गए सधु अभिलेख लगभग 3,000 हैं[5]। संख्या कम नहीं है, लेकिन प्रत्येक अभिलेख बहुत छोटा है। केवल छोटे लेखों को जोड़कर वाक्य की संरचना समझना कठिन है। क्योंकि किरता और क्रिया कैसे व्यवस्थित होते हैं, यह देखने का अवसर नहीं मिलता।

समुद्र पार काष्ठ पट्टिकाओं के प्रतीक

दक्षिणी प्रशांत महासागर के मध्य में रापा नुई नामक एक द्वीप है। हम ईस्टर द्वीप नाम से अधिक परिचित हैं। मानव आबादी वाली नकिटतम भूमि 2,000 किलोमीटर से भी अधिक दूर है। यह विशाल पत्थर की मानव प्रतमाओं मोआई के लिए भी जाना जाता है। इस द्वीप से भी अज्ञात अर्थ वाले प्रतीकों से लिखी लकड़ी की पट्टिकाएँ मिलीं। इस अभिलेख को आज रोगोरोन्गो कहा जाता है[6]।

ये प्रतीक बर्श से लिखे गए अक्षर नहीं हैं। नुकीले ओब्सिडियन के टुकड़ों या हड्डियों के औजारों से लकड़ी को खोदकर उकेरा गया था। उकेरी गई आकृतियाँ मनुष्य, पक्षी और मछली जैसी दिखती हैं[6]। हाथ-पैर उठाए मनुष्य की आकृति अक्सर दिखाई देती है।

पढ़ने की दिशा भी अनोखी है। पढ़ने वाला व्यक्ति पट्टिका के नचिले बाएँ कोने से शुरू करता है। एक पंक्ति पढ़ने के बाद पट्टिका को 180 अंश घुमाया जाता है। ऊपर-नीचे उलटी स्थिति में अगली पंक्ति पढ़ी जाती है। प्रत्येक पंक्ति पर पट्टिका को पलटने की इस पद्धति को रिवर्स

बॉउस्ट्रोफडॉन कहा जाता है[6]। सरल शब्दों में, यह पलटकर पढ़ने की वधि है। प्राचीन यूनान में भी प्रत्येक पंक्ति में दशा बदलने वाली लेखन शैली थी। वहाँ पट्टिका को घुमाए बना केवल अक्षरों की दशा बदली जाती थी।

यह वधि शोधकर्ताओं की एक और परीक्षा लेती है। पहले यह तय करना होता है कि कौन सी पंक्ति पहली है, तभी क्रम बनता है[6]। यदि क्रम गलत चुना जाए तो प्रतीकों का पूरा क्रम ही उलट जाता है। कंप्यूटर में डेटा डालने से पहले यह नरिणय मनुष्य को ही लेना पड़ता है।

बचे हुए अवशेष बहुत कम हैं। दुनिया भर के संग्रहालयों और नज्जी संग्रहकर्ताओं के पास बखिरी हुई वास्तविक कलाकृतियाँ केवल 26 हैं। ये सभी रापा नुई द्वीप से बाहर हैं[6]। उन पर उक्रे गए कुल प्रतीकों की संख्या लगभग 15,000 अक्षर ही हैं[7]। यह मात्रा एक पुस्तक के बराबर भी नहीं है।

अवशेष इतने कम होने का एक कारण है। 19वीं सदी में द्वीप की जनसंख्या में भारी गरिवट आई और कई लकड़ी की पट्टिकाएँ नष्ट हो गईं। बची हुई पट्टिकाएँ यूरोप के संग्रहालयों और मठों में बखिर गईं। उनमें से चार पट्टिकाएँ रोम की एक धार्मिक मण्डली में सुरक्षित रखी गई हैं[8]।

ये लकड़ी की पट्टिकाएँ कब बनाई गईं, यह भी लंबे समय से विवाद का विषय था। संदेह था कि यूरोपीय लोगों के द्वीप पर आने के बाद उनकी नकल करके इन्हें बनाया गया था। 2024 में सलिविया फेरारा के शोध दल ने रोम की चार लकड़ी की पट्टिकाओं की आयु मापी[8]। उन्होंने लकड़ी में शेष कार्बन की मात्रा मापकर आयु निर्धारित करने की वधि का उपयोग किया। इस वधि को रेडियोकार्बन डेटिंग कहा जाता है।

चार में से तीन पट्टिकाएँ 19वीं सदी की लकड़ी की थीं। शेष एक पट्टिका बहुत पुरानी लकड़ी की नकिली। यह यूरोपीय जहाजों के द्वीप पर पहुँचने के 1720 के दशक से भी 200 से अधिक वर्ष पहले की है[8]। इससे संभावना बढ़ गई कि रापा नुई के लोगों ने स्वयं अपनी लिपि विकसित की थी।

शोध दल ने एक शर्त भी जोड़ी। पेड़ काटने का समय प्रतीकों को उक्रेने के समय के समान नहीं होता। लकड़ी की आयु केवल यह बताती है कि उस पर नक्काशी बाद में की गई थी। क्योंकि पुरानी लकड़ी को बाद में तराश कर उपयोग किया गया हो सकता है। यह बात भी ध्यान देने योग्य है कि इतनी पुरानी तिथि वाली केवल एक ही पट्टिका मिली है[8]।

क्या यह वास्तव में लिपि है

इन दो प्रतीक प्रणालियों को लेकर विद्वानों में लंबा विवाद रहा। विवाद कूटवाचन से शुरू नहीं हुआ। पहले इस बात पर बहस हुई कि क्या ये वास्तव में बोली जाने वाली भाषा को दर्ज करने वाली लिपि हैं[1][9]।

2004 में स्टीव फार्मर, रचर्ड स्प्रोएट और माइकल वटिजेल ने संयुक्त रूप से एक शोध पत्र प्रकाशित किया। उन्होंने तर्क दिया कि संधि प्रतीक कोई लिपि नहीं है। उन्होंने आधार दिया कि अभिलेख बहुत छोटे हैं और कोई भी लंबा दस्तावेज़ मौजूद नहीं है। यानी ये प्रतीक कुल चहिन, संपत्तिके स्वामित्व के नशान या ताबीज के अधिक नकिट हैं[9]।

दूसरी ओर के विद्वानों ने सांख्यिकी से उत्तर खोजने का प्रस्ताव रखा। उनका कहना था कि प्रतीकों की आवृत्ति और उनके आगे-पीछे आने के नयिमों की गणना की जाए[5]। मानव भाषा में बोलने की आदत जैसी सांख्यिकीय प्रवृत्तियाँ होती हैं। यह दृष्टिकोण यह मापने का प्रयास करता है कि क्या वही प्रवृत्तियाँ इन प्रतीकों में भी मौजूद हैं।

इस वधि का लाभ यह है कि अर्थ जाने बना भी इसका उपयोग किया जा सकता है। प्रतीकों को संख्याओं में बदलकर केवल उनके क्रम की गणना करनी होती है। यह वह काम है जिसमें कंप्यूटर बहुत कुशल है।

कूटवाचन के मार्ग में तीन दीवारें

दोनों ही प्रतीक प्रणालियों का अभी तक कूटवाचन नहीं हो सका है। राह रोकने वाली तीन दीवारें हैं।

तुलना करने के लिए कोई पाठ नहीं है। मसिर की चतुरलिपि रोसेटा स्टोन की बदौलत सुलझाई गई थी। रोसेटा स्टोन पर वही सामग्री ग्रीक भाषा में भी लिखी हुई थी। फ्रांसीसी विद्वान शांपोलियों ने ज्ञात और अज्ञात लिपियों को आमने-सामने रखकर उनका मलान किया। एक ही सामग्री को दो अलग-अलग भाषाओं में समानांतर रूप से लिखने वाले पाठ को द्विभाषी पाठ कहा जाता है। संधि स्थलों और रापा नुई द्वीप से ऐसा कोई अभिलेख नहीं मिला है[2][6]।

यह भी ज्ञात नहीं है कि किस भाषा को लिखा गया था। प्रतीकों द्वारा दर्ज की गई वास्तविक भाषा को अंतरनिहित भाषा कहा जाता है। संधि लिपि की अंतरनिहित भाषा को लेकर द्रवड़ि भाषा परिवार और भारत-आर्य भाषा परिवार के सिद्धांत आमने-सामने हैं। यह दावा भी किया जाता है कि यह किसी भी ज्ञात भाषा परिवार से संबंधित नहीं है[2]।

अंतिम बाधा डेटा की मात्रा है। आज का एआई विशाल मात्रा में पाठ्य सामग्री सीखकर बढ़ा होता है। लेकिन यहाँ केवल कुछ हज़ार सधु मुहरे और 26 लकड़ी की पट्टिकाएँ ही उपलब्ध हैं[1][6]। डेटा कम होने पर एआई विश्वसनीय लगने वाले झूठ गढ़ने लगता है। इस परिघटना को मतभ्रम कहा जाता है।

ये तीनों बाधाएँ एक-दूसरे से जुड़ी हुई हैं। यदि अंतर्राष्ट्रीय भाषा ज्ञात हो तो ध्वनिभूल्यों का मिलान किया जा सकता है। यदि दिवभाषी पाठ उपलब्ध हो तो अंतर्राष्ट्रीय भाषा का पता लगाया जा सकता है। यदि डेटा प्रचुर मात्रा में हो तो सांख्यिकी से सुराग मिल सकता है। क्योंकि इनमें से एक भी नहीं सुलझ पा रहा है, इसलिए बाकी भी वहीं के वहीं रुके हैं।

5 जनवरी 2025 को भारत के तमिलनाडु राज्य ने एक पुरस्कार की घोषणा की। यह घोषणा चेन्नई में आयोजित सधु सभ्यता की खोज के 100 वर्ष पूरे होने पर आयोजित अंतरराष्ट्रीय सम्मेलन के उद्घाटन समारोह में की गई थी[10]। इसके अनुसार जो व्यक्ति पुरातत्व विशेषज्ञों द्वारा स्वीकार्य रूप से सधु लिपि का कूटवाचन करेगा, उसे 10 लाख डॉलर दिए जाएँगे[10][11]। सितंबर 2026 तक कोई ज्ञात विजेता नहीं है।

संदर्भ सामग्री

- [1] Rao, R. P. N., Yadav, N., Vahia, M. N., Joglekar, H., Adhikari, R., & Mahadevan, I. (2009). Entropic evidence for linguistic structure in the Indus script. *Science*, 324(5931), 1165. <https://doi.org/10.1126/science.1170391>
- [2] Khan, M. A., Khurshid, S. K., & Aslam, M. (2025). Modern algorithms for ancient scripts: A review of AI-based techniques in Indus civilization research. *Cultural and Historical Heritage: Preservation, Presentation, Digitalization*, 11(2), 11-21. <https://doi.org/10.55630/KINJ.2025.110201>
- [3] Grand Valley State University. (n.d.). Archaeologist of the month: John Marshall. [gvsu.edu. https://www.gvsu.edu/archaeology/module-spotlight-view.htm?entryId=A748C17B-A353-F70E-D7422D41F7889433](https://www.gvsu.edu/archaeology/module-spotlight-view.htm?entryId=A748C17B-A353-F70E-D7422D41F7889433) (3 सितंबर 2026 को अभिगमति)
- [4] Farmer, S. (n.d.). Claims concerning the longest Indus "inscription". [safarmer.com. https://safarmer.com/indus-longestinscription/](https://safarmer.com/indus-longestinscription/) (3 सितंबर 2026 को अभिगमति)
- [5] Yadav, N., Joglekar, H., Rao, R. P. N., Vahia, M. N., Adhikari, R., & Mahadevan, I. (2010). Statistical analysis of the Indus script using n-grams. *PLoS ONE*, 5(3), e9506. <https://doi.org/10.1371/journal.pone.0009506>
- [6] Australian Museum. (2026, 5 मई 28 मई). Kohau Rongorongo: Talking tablet. [australian.museum. https://australian.museum/learn/cultures/pasifika-collections/rapa-nui-collections/kohau-rongorongo-talking-tablet/](https://australian.museum/learn/cultures/pasifika-collections/rapa-nui-collections/kohau-rongorongo-talking-tablet/) (3 सितंबर 2026 को अभिगमति)
- [7] Ross, L. (2026, 3 मई 31 मई). Reading Rongorongo (Version 2). *Zenodo*. <https://doi.org/10.5281/zenodo.19362491> (3 सितंबर 2026 को अभिगमति)
- [8] Ferrara, S., Tassoni, L., Kromer, B., Wacker, L., Friedrich, M., Tonini, F., Lastilla, L., Ravanelli, R., & Talamo, S. (2024). The invention of writing on Rapa Nui (Easter Island): New radiocarbon dates on the Rongorongo script. *Scientific Reports*, 14, 2794. <https://doi.org/10.1038/s41598-024-53063-7>
- [9] Farmer, S., Sproat, R., & Witzel, M. (2004). The collapse of the Indus-script thesis: The myth of a literate Harappan civilization. *Electronic Journal of Vedic Studies*, 11(2), 19-57. <https://doi.org/10.11588/ejvs.2004.2.620>
- [10] The Tribune. (2025, 1 मई 6 मई). Stalin announces \$1 mn prize for deciphering Indus Valley script. [tribuneindia.com. https://www.tribuneindia.com/news/india/stalin-announces-1-mn-prize-for-deciphering-indus-valley-script/](https://www.tribuneindia.com/news/india/stalin-announces-1-mn-prize-for-deciphering-indus-valley-script/) (3 सितंबर 2026 को अभिगमति)
- [11] Smithsonian Magazine. (2025, 1 मई 30 मई). Officials are offering \$1 million to anyone who can decode this ancient script. [smithsonianmag.com. https://www.smithsonianmag.com/smart-news/officials-are-offering-1-million-to-anyone-who-can-decode-this-ancient-script-180985922/](https://www.smithsonianmag.com/smart-news/officials-are-offering-1-million-to-anyone-who-can-decode-this-ancient-script-180985922/) (3 सितंबर 2026 को अभिगमति)

10.2. वैज्ञानिक जाँच दल की पड़ताल: सूचना एंट्रॉपी मापक

अक्षरों के फ़िरफ़ट्टि मापने का पैमाना

सधु मुहरे और रापा नुई की लकड़ी की पट्टिकाएँ कभी एक-दूसरे से नहीं मिलीं। फिर भी दोनों स्थलों के शोधकर्ताओं ने एक ही उपकरण चुना। क्योंकि अज्ञात अर्थ वाले प्रतीकों की पंक्तियों के सामने वे जो कर सकते थे, वह लगभग एक जैसा था।

अपठित लिपि के सामने शोधकर्ताओं ने एक अलग प्रश्न चुना। इसका अर्थ क्या है, यह पूछने के बजाय उन्होंने पूछा कि यह किस प्रकार व्यवस्थित है। क्योंकि अर्थ जाने बिना भी क्रम की गणना की जा सकती है।

इस समय उपयोग किया जाने वाला उपकरण सूचना एंट्रॉपी है। एंट्रॉपी यह मापने वाला मान है कि आगे क्या आएगा, इसका अनुमान लगाना कतिना कठिन है। यदि मान अधिक है, तो इसका अर्थ है कि आगे का अनुमान लगाना कठिन है। यदि मान कम है, तो इसका अर्थ है कि आगे क्या आएगा, यह लगभग तय है।

यह अवधारणा संचार इंजीनियरिंग से आई है। यह टेलीफोन लाइन से सिग्नल भेजते समय कतिनी सूचना प्रसारित होती है, इसे मापने के लिए बनाया गया पैमाना है। यह अर्थ की परवाह किए बिना केवल क्रम को देखकर गणना करता है। यही कारण है कि इसे अपठित लिपि पर भी लागू किया जा सकता है।

कक्षा में बैठने की सीटों के उदाहरण से इसे समझना आसान है। यदि लॉटरी निकालकर बैठा जाए, तो पता नहीं होता कि बगल में कौन बैठेगा। ऐसी स्थिति में एंट्रॉपी अधिक होती है। यदि शील नंबर के क्रम में बैठा जाए, तो बगल में बैठने वाला व्यक्ति पहले से तय होता है। ऐसी स्थिति में एंट्रॉपी कम होती है।

मानव भाषा इन दोनों चरम सीमाओं के बीच स्थिति है। व्याकरण क्रम को बांधता है, लेकिन कहने की बात स्वतंत्र रूप से चुनी जाती है। इसलिए मानव भाषा की एंट्रॉपी मध्य क्षेत्र में आती है। इस मध्य क्षेत्र को पैमाने के रूप में उपयोग करना शोध का शुरुआती बिंदु है।

कोरियाई भाषा में भी इसे देखा जा सकता है। "hakgyo-e" के बाद आने वाले शब्द कई हो सकते हैं। ganda, gatda, danyeowatda सभी संभव हैं। लेकिन कोई भी मनमाना शब्द नहीं आ सकता। विकल्प होना लेकिन सीमिति दायरे में होना, भाषा की यही विशेषता है।

संक्षिप्त प्रतीकों को मापना

2009 में राजेश राव और नशिा यादव सहित शोधकर्ताओं के दल ने इस पैमाने को संक्षिप्त प्रतीकों पर लागू किया। उन्होंने गणना की कि पिछले प्रतीक की जानकारी होने पर अगले प्रतीक की अनिश्चितता कतिनी होती है [1]। पूर्ववर्ती प्रतीक के ज्ञात होने पर अगले प्रतीक की प्रत्याशा को मापने वाले इस मान को संक्षिप्त एंट्रॉपी कहा जाता है।

तुलना के लिए अन्य विषय भी तैयार किए गए। एक ओर अंग्रेजी, संस्कृत और सुमेरियन जैसी मानव भाषाओं को रखा गया। दूसरी ओर गैर-मानव भाषा प्रणालियों को रखा गया। इनमें जैविक डीएनए अनुक्रम, कंप्यूटर प्रोग्राम कोड और यादृच्छिक रूप से उत्पन्न प्रतीक शृंखलाएँ शामिल थीं [1]।

परिणाम यह निकला कि संक्षिप्त प्रतीक मानव भाषाओं के समूह के निकट पाए गए। वे यादृच्छिक प्रतीक शृंखलाओं की तुलना में अधिक स्पष्ट नियमों वाले थे और कंप्यूटर कोड की तुलना में अधिक स्वतंत्र थे। शोध दल ने इस परिणाम को Science पत्रिका में प्रकाशित किया [1]।

उसी शोध दल ने एक संभाव्यता मॉडल भी प्रस्तुत किया। यह एक ऐसी विधि है जिसमें पिछले प्रतीक के निर्धारित होने पर अगले प्रतीक के आने की संभावना की तालिका बनाई जाती है। केवल पूर्ववर्ती प्रतीक को देखकर अगले की गणना करने वाले इस मॉडल को मार्कोव मॉडल कहा जाता है। तालिका को देखने पर पता चलता है कि कौन से प्रतीक पाठ की शुरुआत में अधिक आते हैं। और कौन से प्रतीक पाठ के अंत में अधिक आते हैं, यह भी स्पष्ट होता है [2]।

उसी शोध दल ने अगले वर्ष प्रतीकों की आवृत्ति की भी जाँच की। उन्होंने लगभग 3,000 अभिलेख एकत्र किए और उनमें से अच्छी स्थिति वाले 1,548 अभिलेख चुने। बहुत कम प्रतीकों का उपयोग अधिकांश स्थानों पर हुआ था और बाकी प्रतीक बहुत कम दिखाई दिए। यह स्वरूप मानव भाषा में हमेशा देखने वाला वितरण है [3]।

इस वितरण का एक नाम है। इसे ज़िफ़ी का नियम कहा जाता है। यह वह नियम है जिसके अनुसार अक्सर उपयोग होने वाले कुछ शब्द पाठ का अधिकांश भाग बनाते हैं। कोरियाई भाषा में भी eun, neun, i, ga जैसे शब्द शीर्ष स्थान पर होते हैं। इसके विपरीत, केवल एक बार उपयोग होकर समाप्त होने वाले शब्द बहुत अधिक होते हैं। संक्षिप्त प्रतीकों में भी यही स्वरूप दिखाया [3]।

उनके स्थान के क्रम में भी नियम सामने आए। 23 प्रतीकों ने पाठ के अंत का 80 प्रतिशत भाग घेरा। इसके विपरीत, पाठ की शुरुआत का 80 प्रतिशत भरने के लिए 82 प्रतीकों की आवश्यकता थी [3]। इसका अर्थ है कि अंत में आने वाले प्रतीक बहुत अधिक सीमिति थे। यह कोरियाई भाषा में वाक्य के अंत में परसर्गों या प्रत्ययों के एकत्रित होने के स्वरूप जैसा है।

शोध दल ने पूर्ववर्ती प्रतीक से अगले प्रतीक का अनुमान लगाने वाला एक मॉडल भी बनाया [2]। इस मॉडल ने घसिकर अपठनीय हो चुके प्रतीकों को लगभग 75 प्रतिशत संभावना के साथ पुनर्निर्मित किया [3]। यह ऐसा परिणाम है जो नियम होने पर ही संभव है।

'पुनर्निर्मित करने' शब्द का गलत अर्थ नहीं निकाला जाना चाहिए। मॉडल केवल प्रतीक के आकार को ही भरता है। वह प्रतीक कौन सी ध्वनि था, यह अभी भी अज्ञात है। उसने पहली का एक छूटा हुआ टुकड़ा भरा है, पूरी तस्वीर पढ़कर नहीं सुनाई है।

खंडन और प्रत-खंडन

इस मॉडल की सीमाएँ भी हैं। क्योंकि यह केवल पछिले एक या दो प्रतीकों को देखकर अगले का अनुमान लगाने की वधि है [2]। मानव वाक्यों में बहुत दूर स्थिति शब्दों के बीच भी संगति होती है। यदि डेटा छोटा हो तो वह इस दूरस्थ संगति को नहीं सीख सकता।

इस परिणाम का तुरंत खंडन किया गया। रचिर्ड स्प्रोट ने 2010 में एक आलोचनात्मक लेख प्रकाशित किया। उनका तर्क था कि केवल एंटरपी मान समान होने के तथ्य से ही इसे लपि घोषित नहीं किया जा सकता। क्योंकि गैर-मानव भाषा प्रतीक प्रणालियाँ भी समान मान उत्पन्न कर सकती हैं [4]।

राव की शोध टीम ने उसी वर्ष उसी शोध पत्रिका में उत्तर दिया। उन्होंने स्पष्ट किया कि उन्होंने सधु प्रतीकों के लपि होने को सिद्ध नहीं किया है। उनकी व्याख्या थी कि उन्होंने केवल सांख्यिकी से इसके लपि होने की संभावना का समर्थन किया है [5]।

स्प्रॉट ने 2014 में अपनी आलोचना को एक कदम और आगे बढ़ाया। उन्होंने गैर-भाषाई प्रतीकों के कई डेटा सेट एकत्र किए। इनमें मेसोपोटामिया के देवताओं के चिह्न, उत्तरी अमेरिका के कुलचिह्न (टोटेम) खंभे और खलहिन की दीवारों पर बने तारों के चित्र शामिल थे। उन्होंने मौसम के पूर्वानुमान में इस्तेमाल होने वाले चित्र भी जोड़े। उन्होंने इस डेटा और सधु प्रतीकों की तुलना एक ही पैमाने से की [6]।

उनकी बनाई गई नरिणय प्रक्रिया ने सधु प्रतीकों को गैर-भाषाई वर्ग में रखा। स्प्रॉट ने लिखा कि केवल सांख्यिकी पर्याप्त नहीं है। उनका तर्क है कि इसके लिए या तो सफल पाठोद्धार होना चाहिए या उस संस्कृति द्वारा लेखन के उपयोग का कोई अलग पुरातात्विक प्रमाण मलिन चाहिए [6]। यह विवाद अभी समाप्त नहीं हुआ है।

अप्रैल 2026 में आया नया विश्लेषण भी इसी प्रश्न के सामने खड़ा है। आशीष नायर ने सधु घाटी के 1,916 अभिलेखों को लेकर परीक्षण की रूपरेखा फिर से तैयार की। उन्होंने गैर-भाषाई प्रतीक प्रणालियों की नकल करते हुए कंप्यूटर से कृत्रिम डेटा बनाया। उन्होंने वास्तविक सधु डेटा और कृत्रिम डेटा के सांख्यिकीय आँकड़ों की आमने-सामने तुलना की [7]।

सधु डेटा किसी भी कृत्रिम डेटा से पूरी तरह मेल नहीं खाया। यह दो कृत्रिम डेटा सेटों के ठीक बीच में रहा। वास्तव में मौजूद कोई भी गैर-भाषाई प्रतीक प्रणाली सधु के सांख्यिकीय आँकड़ों को हूबहू नहीं दोहरा सकती। नायर ने लिखा कि यह लपि है या नहीं, इसका नरिणय अभी टाल देना चाहिए। यह लेख शोध पत्रिका की समीक्षा से पहले इंटरनेट पर प्रकाशित प्रीप्रिंट है [7]।

प्रतीकों को पहचानने वाली आँख

सांख्यिकी से पहले एक काम करना होता है। वह है पत्थरों पर उकेरे गए धुंधले निशानों को प्रतीकों के रूप में उतारना। यह कार्य लंबे समय तक मानव की आँखों और हाथों पर निर्भर रहा।

निशानों को उतारना सोच से कहीं अधिक कठिन है। मुहरे 4,000 वर्षों में घिस चुकी हैं। एक ही प्रतीक भी उकेरने वाले के अनुसार थोड़ा भिन्न दिखता है। इसे दो प्रतीक माना जाए या एक, इस पर भी राय बँट जाती है। यदि यह नरिणय बदलता है, तो इसके बाद के सांख्यिकीय आँकड़े भी बदल जाते हैं।

2017 में सतीश पलनयिपन और रोनोजॉय अधिकारी ने यह काम कंप्यूटर को सौंपा। उन्होंने एक ऐसी कार्यप्रणाली बनाई जिसमें मुहर की तस्वीर डालने पर प्रतीकों की सूची निकल आती है [8]।

2025 में वैष्णवी दीक्षित की शोध टीम ने इस सलिसलि को आगे बढ़ाया। उन्होंने मुहरों की तस्वीरों से प्रतीकों को पहचानने वाले और जानवरों के चित्रों को पहचानने वाले अलग-अलग तंत्रिका नेटवर्क बनाए। उन्होंने डेटा को पाँच भागों में बाँटकर बारी-बारी से परीक्षण किया। दोनों तंत्रिका नेटवर्कों ने औसतन लगभग 85 प्रतिशत की सटीकता दिखाई [9]।

इस समस्या को कम करने के लिए नरिणय के मानदंडों को सार्वजनिक करना और उनकी आपस में तुलना करना आवश्यक है। 2025 में मुज़फ़्फ़र अली ख़ान की शोध टीम ने इस कार्य को संकलित किया। उन्होंने 2009 से 2025 तक सधु लपि से जुड़े एआई अध्ययनों को एकत्र कर प्रकारों में विभाजित किया। उन्होंने बार-बार दोहराई जाने वाली गलतियों को भी साथ में दर्ज किया [10]।

2025 में अतुल शर्मा और शुभाजीत रॉय चौधरी ने एक उपकरण प्रस्तुत किया। इसका नाम AI-EPIGRAPHY है। जब कोई उपयोगकर्ता प्रतीकों का क्रम डालता है, तो यह उच्च संभावना वाले संभावित अर्थ प्रस्तुत करता है। यह उपकरण आवृत्ति विश्लेषण और शृंखला प्रायिकता गणना को जोड़ता है [11]।

मई 2026 में डेटाबेस के आधार को ही फिर से तैयार करने के प्रयास की घोषणा की गई। बीटा मेज़ेसी की शोध टीम ने दुर्लभ और अपठित लपियों को एक स्थान पर एकत्र करने वाला डेटाबेस तैयार किया। इसमें उच्च-रिज़ॉल्यूशन तस्वीरें, प्रतलिखित प्रतीक, टपिपणियाँ और स्रोत की जानकारी एक ही प्रारूप में रखी गई। उनका मानना था कि बिखरे हुए डेटा को एकत्र करना पाठोद्धार से पहले का कार्य है [12]।

संदर्भ सामग्री

- [1] Rao, R. P. N., Yadav, N., Vahia, M. N., Joglekar, H., Adhikari, R., & Mahadevan, I. (2009). Entropic evidence for linguistic structure in the Indus script. *Science*, 324(5931), 1165. <https://doi.org/10.1126/science.1170391>
- [2] Rao, R. P. N., Yadav, N., Vahia, M. N., Joglekar, H., Adhikari, R., & Mahadevan, I. (2009). A Markov model of the Indus script. *Proceedings of the National Academy of Sciences*, 106(33), 13685-13690. <https://doi.org/10.1073/pnas.0906237106>
- [3] Yadav, N., Joglekar, H., Rao, R. P. N., Vahia, M. N., Adhikari, R., & Mahadevan, I. (2010). Statistical analysis of the Indus script using n-grams. *PLoS ONE*, 5(3), e9506. <https://doi.org/10.1371/journal.pone.0009506>
- [4] Sproat, R. (2010). Ancient symbols, computational linguistics, and the reviewing practices of the general science journals. *Computational Linguistics*, 36(3), 585-594. https://doi.org/10.1162/coli_a_00011
- [5] Rao, R. P. N., Yadav, N., Vahia, M. N., Joglekar, H., Adhikari, R., & Mahadevan, I. (2010). Entropy, the Indus script, and language: A reply to R. Sproat. *Computational Linguistics*, 36(4), 795-805. https://doi.org/10.1162/coli_c_00030
- [6] Sproat, R. (2014). A statistical comparison of written language and nonlinguistic symbol systems. *Language*, 90(2), 457-481. <https://doi.org/10.1353/lan.2014.0031>
- [7] Nair, A. (2026, 4월 20일). How non-linguistic is the Indus sign system? A synthetic-baseline scorecard. [arXiv:2604.17828](https://arxiv.org/abs/2604.17828). <https://arxiv.org/abs/2604.17828> (3 सितंबर 2026 को अभिगमिति)
- [8] Palaniappan, S., & Adhikari, R. (2017). Deep learning the Indus script. [arXiv:1702.00523](https://arxiv.org/abs/1702.00523). <https://arxiv.org/abs/1702.00523>
- [9] Dixit, V., Hussain, N., Basak, S., Atturu, D., Mitra, D., & Bhattacharya, U. (2025). Deep learning in archiving Indus script and motif information. *Journal of Computer Applications in Archaeology*, 8(1), 156-169. <https://doi.org/10.5334/jcaa.175>
- [10] Khan, M. A., Khurshid, S. K., & Aslam, M. (2025). Modern algorithms for ancient scripts: A review of AI-based techniques in Indus civilization research. *Cultural and Historical Heritage: Preservation, Presentation, Digitalization*, 11(2), 11-21. <https://doi.org/10.55630/KINJ.2025.110201>
- [11] Sharma, A., & Roy Chowdhury, S. (2025). AI-EPIGRAPHY: An interactive tool for computational decipherment of the Indus Valley script. *Proceedings of IndiaHCI '25*. <https://doi.org/10.1145/3768633.3770145>
- [12] Megyesi, B., Rattenborg, R., Lang, B., Waldispuhl, M., & Heder, M. (2026). Building a corpus and database for rare and undeciphered scripts. *Proceedings of the Fourth Workshop on Language Technologies for Historical and Ancient Languages (LT4HALA 2026) @ LREC 2026*, 184-196. ELRA Language Resources Association. <https://aclanthology.org/2026.lt4hala-1.17/> (3 सितंबर 2026 को अभिगमिति)

10.3. उजागर हुआ सत्य

चंद्रमा की गणना करने वाली काष्ठ पट्टिका

रापा नुई की काष्ठ पट्टिकाओं में से एक का नाम मामारी है। इस काष्ठ पट्टिका में एक ऐसा भाग है जो अन्य भागों से भिन्न दिखता है। इसमें अर्धचंद्र के आकार के प्रतीक लगातार बने हुए हैं[1]।

जर्मन विद्वान थॉमस बारथेल ने 20वीं सदी के मध्य में इस भाग पर ध्यान दिया। उन्होंने इसे चंद्रमा की तथियों का अभिलेख माना[1]। 1990 में ज़ाक गी ने डेटा को पुनः व्यवस्थिति करके वही नक्षत्र प्रस्तुत किया[2]।

2011 में पॉल हॉर्ली ने इस व्याख्या को और सुदृढ़ किया। मामारी काष्ठ पट्टिका का पंचांग तीस रातों से बना है। यह संख्या चंद्रमा के वास्तविक चक्र से मेल खाती है। यह पोलनिशिया के कई द्वीपों के पंचांगों से भी मेल खाती है[1]।

हॉर्ली ने ताहिती और न्यूजीलैंड के पंचांगों की भी साथ रखकर तुलना की। प्रत्येक रात को दिया जाने वाला नाम हर द्वीप में मलिता-जुलता था। रापा नुई का पंचांग अन्य द्वीपों की तुलना में लगभग तीन दिनों पीछे था। ऐसा प्रतीत होता है कि यह अंतर द्वीपों के दूर-अलग-थलग रहने के दौरान आया[1]।

प्रतीकों का वन्यास भी नियमों का पालन करता था। अर्धचंद्र प्रतीकों के बीच-बीच में अन्य प्रतीक आकर उन्हें आठ समूहों में बाँटते हैं। अंतिम दो अर्धचंद्र उन रातों को दर्शाते हैं जब चाँद दिखाई नहीं देता[1]।

यही क्रम काष्ठ पट्टिका के बाहर भी मलिता है। द्वीप के शैलचित्रों में कछुए के चित्र और अर्धचंद्र एक साथ उकेरे गए हैं। वहाँ भी अंतिम दो अर्धचंद्र कछुए की पीठ के खोल के भीतर स्थिति हैं। यह काष्ठ पट्टिका द्वारा चाँद-वहीन रातों को अलग दर्शाने की विधि के समान है[1]। काष्ठ पट्टिका और चट्टान दोनों एक ही गणना पद्धतिसाझा करते थे।

यह व्याख्या रोगोरोन्गो के अध्ययनों में व्यापक रूप से स्वीकार की जाती है। फरि भी, जो पढ़ा गया है वह केवल तथियों का क्रम है। हम यह नहीं जानते कि रापा नुई के लोग उन रातों को क्या पुकारते थे।

एआई द्वारा दिए गए उत्तर और उनका भार

चाँद की गणना और सूर्य की गणना आपस में जुड़ी हुई हैं। सूर्य ग्रहण तब होता है जब सूर्य और चंद्रमा एक सीध में आते हैं। इसलिए पंचांग को पढ़ लेने का दावा तुरंत पूरे आकाश को पढ़ लेने के दावे में बदल जाता है।

मार्च 2026 में लोगन रॉस ने एक प्रीप्रिंट शोधपत्र प्रकाशित किया। उन्होंने अपने नज्जी कंप्यूटर पर भाषा मॉडल चलाकर रोगोरोन्गो का विश्लेषण किया। उन्होंने 89 प्रतीकों का एक शब्दकोश बनाया और दावा किया कि इससे शेष बचे अभिलेखों का लगभग 97 प्रतिशत हसिंसा शामिल हो गया। उन्होंने यह दावा भी किया कि भामारी काष्ठ पट्टिका सूर्य ग्रहण और चंद्र ग्रहण की भविष्यवाणी करने वाली तालिका है[3]।

इस दावे का अभी तक अकादमिक जगत द्वारा सत्यापन नहीं किया गया है। रॉस का लेख किसी शोध पत्रिका की समीक्षा से पारित शोधपत्र नहीं है। यह इंटरनेट रीपोजिटरी पर पहले अपलोड किया गया प्रीप्रिंट है[3]। प्रीप्रिंट कोई भी अपलोड कर सकता है। इसकी सामग्री सही है या नहीं, इसकी बाद में अन्य शोधकर्ताओं द्वारा जाँच की जानी चाहिए।

ऐसे दावों पर विचार करते समय परीक्षण की वधिको भी देखना चाहिए। रोगोरोन्गो के शेष अभिलेख केवल 26 वस्तुएँ और लगभग 15,000 अक्षर हैं[3]। जहाँ केवल इतना ही डेटा उपलब्ध हो, वहाँ संयोगवश परिणाम मेल खा जाना आसान होता है। यदि 89 प्रतीकों को इच्छानुसार ध्वन्यात्मक मान दे दिए जाएँ, तो कोई भी वाक्य बनाया जा सकता है। इसलिए अकादमिक जगत सबसे पहले यह पूछता है कि क्या कोई अन्य व्यक्ति उसी वधि से वही परिणाम प्राप्त कर सकता है।

वास्तविक पाठोद्धार के लिए एक परीक्षा से गुजरना आवश्यक होता है। वह यह कि पाठोद्धार प्रस्ताव तैयार करते समय उपयोग न की गई किसी अन्य काष्ठ पट्टिका को लाकर पढ़ा जाए। वहाँ भी अर्थपूर्ण वाक्य निकलना चाहिए। माइसीनियाई ग्रीक को लिखने वाली लीनियर बी उस परीक्षा को उत्तीर्ण करने का एक उदाहरण है[4]। रोगोरोन्गो और सधु लपि अभी तक उस सीमा को पार नहीं कर पाई है।

मशीन क्या कर सकती है और क्या नहीं

अब तक एआई ने इन दोनों लिपियों में जो कार्य किया है, वह स्पष्ट है।

मशीन धुंधले निशानों में से प्रतीकों को पहचानती है[5]। जिस काम को मनुष्य को आँखों से मिलाने में महीनों लगते, उसे वह कुछ ही घंटों में पूरा कर देती है। घसि जाने के कारण पढ़ने में कठिनी प्रतीकों को संदर्भ के आधार पर पुनर्जीवित भी करती है[6]। यह बखिरे हुए पुरावशेषों की तस्वीरों और अभिलेखों को एक स्थान पर एकत्र करके खोजने योग्य बनाती है[7]।

डेटा एकत्र करने का कार्य आसानी से दिखाई नहीं देता। लेकिन इस कार्य के बिना इसके बाद की गणना भी संभव नहीं है। यदि तस्वीरों की चमक और कोण अलग-अलग हों, तो प्रतीकों की तुलना नहीं की जा सकती। यदि प्रतिलिखन के नियम हर शोधकर्ता के लिए अलग हों, तो सांख्यिकी भी गलत हो जाती है। 2026 में आया डेटाबेस उन मानकों को एक समान करने का एक प्रयास है[7]।

सांख्यिकी के क्षेत्र में भी उपलब्धियाँ रही हैं। कई अध्ययनों ने इस तथ्य की पुष्टि की है कि सधु प्रतीक कोई यादृच्छिक क्रम नहीं हैं[6][8]। व्यवस्था में नियम होने का निष्कर्ष बार-बार दोहराया गया है। फरि भी, इसका अर्थ यह नहीं है कि यह लिपि ही है। नियम का होना और मानव वाणी को लिपिबद्ध करना दोनों अलग बातें हैं।

लेकिन जो कार्य मशीन नहीं कर सकती, वह चुनौती कहीं अधिक बड़ी बनी हुई है। सधु लिपि के प्रत्येक प्रतीक का ध्वन्यात्मक मान अभी भी अज्ञात है[9]। रोगोरोन्गो प्रतीकों का ध्वन्यात्मक मान भी अज्ञात है[3]। दोनों ही लिपियों में अभी तक ऐसा कोई शब्द नहीं है जिसका अर्थ निश्चित हो चुका हो। सधु मुहरो पर लिखा गया नाम किसी व्यक्तिका है या किसी वस्तु का, यह भी तय नहीं हुआ है[9]।

इसका कारण डेटा की प्रकृति में निहित है। यदि द्विभाषी पाठ उपलब्ध हो, तो पाठोद्धार बहुत आसान हो जाता है। न तो सधु स्थलों पर और न ही रापा नुई द्वीप पर ऐसा कोई पाठ मिलता है[9]। ऐसा नहीं है कि द्विभाषी पाठ के बिना कोई लिपिसुलझी न हो। लीनियर बी इसी तरह सुलझी थी। कति उस समय अंतर्राष्ट्रि भाषा का अनुमान लगाने के संकेत मौजूद थे और डेटा भी पर्याप्त था[4]। सधु लिपि की अंतर्राष्ट्रि भाषा कौन सी है, यह भी तय नहीं हुआ है[9]। रोगोरोन्गो के शेष अभिलेख बहुत ही कम हैं[3]।

अप्रैल 2026 के विश्लेषण ने इस स्थिति को आँकड़ों के माध्यम से पुनः दर्शाया। सधु घाटी के 1,916 अभिलेख न तो मानव भाषा के सांख्यिकी से पूरी तरह मेल खाते हैं और न ही गैर-भाषाई प्रतीकों के सांख्यिकी से[10]। इसका अर्थ है कि सौ से अधिक वर्षों से चला आ रहा प्रश्न अभी भी अनुत्तरित है।

प्रतीक्षा के पक्ष में खड़े होकर

अपठित लिपियों का अध्ययन करने वाला शोध दो प्रकार के दृष्टिकोण की माँग करता है। एक है नए उपकरणों को अंत तक आजमाने का दृष्टिकोण। दूसरा है सामने आए परिणामों पर संदेह करने का दृष्टिकोण।

एआई पहला काम बहुत अच्छी तरह करता है। वह गणना करता है, तुलना करता है और संभावित विकल्पों की सूची प्रस्तुत करता है। बाद वाला काम मनुष्य का है। उसे यह नर्णय लेना होता है कि उन विकल्पों में से उस युग में किसके होने की संभावना थी। उस नर्णय में पुरातात्विक स्थल की स्थिति, कब्रों से मिले अवशेष और बची हुई मौखिक परंपराएँ भी शामिल होती हैं।

रापा नुई के पंचांग की व्याख्या इसका एक उदाहरण है। अर्धचंद्र के प्रतीकों की गिनती करना एक गणना है। उन तीस रातों की तुलना पोलिनेशिया के कई द्वीपों के पंचांगों से करना ऐतिहासिक सत्यापन है [1]। जब ये दोनों कार्य आपस में जुड़े, तब व्याख्या को स्वीकार्यता मिली।

सधु मुहरो पर बने जानवरों के चित्र भी ऐसा ही डेटा है। एक सींग वाले जानवर का क्या अर्थ था, यह हम नहीं जानते। मामारी काष्ठ पट्टिका की तीस रातों का भी यही हाल है। क्रम तो पता चल गया है, लेकिन उन रातों में लोग क्या करते थे, यह अज्ञात है।

यह स्वीकार करने का भी अपना महत्व है कि हम नहीं जानते। कहाँ तक स्पष्ट हो चुका है और कहाँ से अनुमान शुरू होता है, इसकी सीमा रेखा खींचने पर ही अगला शोध शुरू होता है। यदि इस रेखा को धुंधला कर दिया जाए, तो मनगढ़ंत कहानी सत्य का रूप धारण कर लेती है। अपठित लिपियों के शोध में यह सीमा रेखा खींचना परिणाम निकालने जतिना ही महत्वपूर्ण है।

इस अध्याय में चर्चा दोनों लिपियों कभी एक-दूसरे से नहीं मिली। इनका समय भी 3,000 से अधिक वर्षों के अंतराल पर है और इनका स्थान भी पृथ्वी के विपरीत छोरों पर है। फरि भी, शोधकर्ताओं के सामने आने वाली कठिनाइयाँ लगभग एक जैसी हैं: संक्षिप्त डेटा, अनुपलब्ध दृष्टिपाठ और अज्ञात अंतर्राष्ट्रि भाषा। यही कारण है कि एक ही उपकरण का उपयोग दोनों स्थानों पर किया जा रहा है।

तमलिनाडु राज्य द्वारा घोषित 10 लाख डॉलर का पुरस्कार भी इसी स्थिति को दर्शाता है [11]। पुरस्कार घोषित हुए एक वर्ष से अधिक बीत जाने के बाद भी कोई विजेता सामने नहीं आया है। संगणक तेज़ हो गए हैं, परंतु चाबी अभी भी नहीं मिली है।

संदर्भ सामग्री

- [1] Horley, P. (2011). Lunar calendar in rongorongo texts and rock art of Easter Island. *Journal de la Societe des Oceanistes*, 132, 17-38. <https://doi.org/10.4000/jso.6314>
- [2] Guy, J. B. M. (1990). The lunar calendar of Tablet Mamari. *Journal de la Societe des Oceanistes*, 91(2), 135-149. <https://doi.org/10.3406/jso.1990.2882>
- [3] Ross, L. (2026, 3 मई 31 मई). Reading Rongorongo (Version 2). Zenodo. <https://doi.org/10.5281/zenodo.19362491> (3 सितंबर 2026 को अभिगमति)
- [4] Chadwick, J. (1958). *The decipherment of Linear B*. Cambridge University Press.
- [5] Palaniappan, S., & Adhikari, R. (2017). Deep learning the Indus script. *arXiv:1702.00523*. <https://arxiv.org/abs/1702.00523>
- [6] Yadav, N., Joglekar, H., Rao, R. P. N., Vahia, M. N., Adhikari, R., & Mahadevan, I. (2010). Statistical analysis of the Indus script using n-grams. *PLoS ONE*, 5(3), e9506. <https://doi.org/10.1371/journal.pone.0009506>
- [7] Megyesi, B., Rattenborg, R., Lang, B., Waldispühl, M., & Heder, M. (2026). Building a corpus and database for rare and undeciphered scripts. *Proceedings of the Fourth Workshop on Language Technologies for Historical and Ancient Languages (LT4HALA 2026) @ LREC 2026*, 184-196. ELRA Language Resources Association. <https://aclanthology.org/2026.lt4hala-1.17/> (3 सितंबर 2026 को अभिगमति)
- [8] Rao, R. P. N., Yadav, N., Vahia, M. N., Joglekar, H., Adhikari, R., & Mahadevan, I. (2009). Entropic evidence for linguistic structure in the Indus script. *Science*, 324(5931), 1165. <https://doi.org/10.1126/science.1170391>
- [9] Khan, M. A., Khurshid, S. K., & Aslam, M. (2025). Modern algorithms for ancient scripts: A review of AI-based techniques in Indus civilization research. *Cultural and Historical Heritage: Preservation, Presentation, Digitalization*, 11(2), 11-21. <https://doi.org/10.55630/KINJ.2025.110201>
- [10] Nair, A. (2026, 4 मई 20 मई). How non-linguistic is the Indus sign system? A synthetic-baseline scorecard. *arXiv:2604.17828*. <https://arxiv.org/abs/2604.17828> (3 सितंबर 2026 को अभिगमति)
- [11] The Tribune. (2025, 1 मई 6 मई). Stalin announces \$1 mn prize for deciphering Indus Valley script. *tribuneindia.com*. <https://www.tribuneindia.com/news/india/stalin-announces-1-mn-prize-for-deciphering-indus-valley-script/> (3 सितंबर 2026 को अभिगमति)

अध्याय 11. एआई ने टूटे हुए ग्रीक शिलालेखों को फरि से पढ़ा: फाउंडेशन मॉडल और Ithaca

11.1. घटना फ़ाइल

पत्थर पर उकेरे गए देश के अभिलेख

प्राचीन ग्रीस के लोग महत्वपूर्ण बातों को पत्थर पर उकेरते थे। जब वे कोई नया कानून बनाते थे, तो उसे पत्थर के खंभों पर उकेरकर सार्वजनिक चौराहों पर स्थापित करते थे। कर संग्रह के बही-खाते भी पत्थर पर दर्ज किए जाते थे। मंदिरों को समर्पित वस्तुओं की सूची भी पत्थर पर ही छोड़ी जाती थी। पत्थर या कांस्य पर उकेरे गए ऐसे लेखों को शिलालेख कहा जाता है। शिलालेखों को एकत्र करने और पढ़ने के अध्ययन को पुरालेखशास्त्र कहा जाता है। इसका अर्थ पत्थर या धातु पर उकेरे गए लेखों का अध्ययन करने वाला शास्त्र है[4]।

शिलालेख आमतौर पर उस काल के लोगों द्वारा सीधे छोड़े गए अभिलेख होते हैं। बाद के लोगों द्वारा प्रतिलिपिबिनाते समय किए गए संशोधनों के नशान इनमें बहुत कम मिलते हैं। इसलिए इतिहासकारों के लिए ये मूल्यवान सामग्री हैं। हालांकि, वे हमेशा अपने मूल स्वरूप में नहीं रहते। कुछ ऐसे शिलालेख भी हैं जहां बाद में अक्षरों को मटिया गया या दोबारा उकेरा गया। पत्थर को घिसकर उस पर दोबारा लिखने के उदाहरण भी मिलते हैं। एथेंस ने किस शहर के साथ कब हाथ मलिया, यह भी शिलालेख ही बताते हैं[4]।

शिलालेख कई लोगों के देखने के लिए लिखे गए लेख होते हैं। इन्हें चौराहों या मंदिरों के सामने खड़ा किया जाता था ताकि आने-जाने वाले लोग इनमें पढ़ सकें। इनका उपयोग स्कूल के गलियारों में लगे सूचना-पत्रों जैसा ही है। इसलिए शिलालेखों में उस दनि के अधिकारी का नाम और तारीख स्पष्ट रूप से दर्ज होती है। इतिहासकार इन नामों और तारीखों को महत्वपूर्ण सुराग मानते हैं[4]।

मटि हुए अक्षर, लुप्त अंश

समस्या समय की है। पत्थर के खंभे हजारों वर्षों तक बारिश और हवा के सामने खड़े रहे। पत्थर की सतह धीरे-धीरे घिस गई और अक्षरों के खांचे उथले हो गए। जमीन में दबे शिलालेख नमी और सूक्ष्मजीवों द्वारा क्षत-विक्षत हो गए। कई बार शिलालेख टूटकर केवल एक ही हिस्सा बचा रह गया[1]।

इसके परिणामस्वरूप शिलालेखों में ऐसे खाली स्थान बन जाते हैं जहां से अक्षर गायब हो चुके होते हैं। विद्वान इन खाली स्थानों को लुप्त अंश कहते हैं। लैटिन में इसे लैक्युना कहा जाता है[1]। ये लुप्त अंश वाक्य के ठीक बीच में छेद कर देते हैं। यदि किसी व्यक्ती का नाम गायब हो जाए, तो यह पता नहीं चल सकता कि वह काम किसने किया था। यदि संख्या गायब हो जाए, तो यह पता नहीं चल सकता कि कितना भुगतान किया गया था। यदि शहर का नाम गायब हो जाए, तो यह जानना असंभव हो जाता है कि यह समझौता किस देश के साथ हुआ था।

विद्वान लंबे समय से इन खाली स्थानों को भरने का प्रयास करते रहे हैं। उन्होंने समान काल और समान सामग्री वाले अन्य शिलालेखों को खोजकर उनकी तुलना की। ग्रीक शिलालेखों का एक निश्चित वाक्य ढांचा होता था। कानूनों की घोषणा करने वाले दस्तावेजों में वाक्य लगभग एक ही क्रम में रखे जाते थे। इसलिए यदि आगे और पीछे का भाग बचा रहे, तो गायब शब्दों का अनुमान लगाया जा सकता है[1]।

पत्थर के खंभे भारी होते हैं, इसलिए उन्हें प्रयोगशाला में नहीं ले जाया जा सकता। इसलिए विद्वान पत्थर पर गीला कागज रखकर उसे दबाते थे। कागज सूखने पर अक्षरों के खांचे उस पर हूबहू उभर आते थे। कागज की इस प्रतिलिपि को छाप (रबिंग) कहा जाता है। विद्वान इस छाप को लेकर लौटते थे और मेज पर बैठकर अक्षरों को पढ़ते थे। पुरानी मुद्रति सूचियों में तस्वीरों के बजाय केवल उतारे गए अक्षर ही छपते थे। इसलिए बाद के लोगों को पत्थर का वास्तविक स्वरूप देखे बिना केवल अक्षरों को ही पढ़ना पड़ता था[4]।

अनुमान लगाने के भी नियम होते हैं। सबसे पहले यह गिना जाता है कि खाली स्थान कितने अक्षरों का था। पत्थर पर बचे अक्षरों के बीच की दूरी मापकर इसका मोटा अनुमान लगाया जा सकता है। इसके बाद केवल अक्षरों की सही संख्या वाले शब्दों को ही उम्मीदवार बनाया जाता है। इस प्रकार भरे गए अक्षरों को अनिवार्य रूप से बड़े कोष्ठक के भीतर लिखा जाता है। वकित होने के कारण पढ़ने में कठिनि अक्षरों के नीचे एक छोटा बटु लगाया जाता है। इस नियम को लीडेन परपिटी (Leiden Conventions) कहा जाता है। यह पत्थर पर बचे मूल अक्षरों और मानवीय अनुमान के बीच अंतर करने के लिए बनाया गया नियम है[1, 4]।

संगिमा के तीन स्ट्रोक को लेकर विवाद

शिलालेख को पढ़ने के लिए यह भी जानना जरूरी है कि वह कब उकेरा गया था। पारंपरिक तरीका अक्षरों के आकार को देखकर काल का अनुमान लगाना है। ऐसा इसलिए है क्योंकि अक्षरों की बनावट समय के साथ धीरे-धीरे बदलती रहती है[2]।

एथेस के शिलालेखों में अक्सर मलिन वाला एक उदाहरण सगिमा है। सगिमा ग्रीक वर्णमाला का एक अक्षर है। शुरुआती दौर में तीन स्ट्रोक वाला सगिमा उकेरा जाता था। बाद में चार स्ट्रोक वाले सगिमा का व्यापक उपयोग होने लगा। लंबे समय तक वदिवानों ने तीन स्ट्रोक वाले सगिमा को एक मानक माना। यद्यपि ऐसा अक्षर मलिता था, तो उस शिलालेख को लगभग 446 ईसा पूर्व से पहले का माना जाता था[2]।

लेकिन कुछ ऐसे वदिवान भी सामने आए जिनोंने इस मानक पर संदेह किया। इसका कारण यह था कलिखावट की शैली बदलने की गतिहर क्षेत्र में अलग थी। कुछ क्षेत्रों ने पुरानी शैली को लंबे समय तक बनाए रखा। एक ही शहर के भीतर भी अलग-अलग शिल्पकारों की उकेरने की आदतें अलग थीं[2]।

इसलिए अन्य शहरों के साथ एथेस की संधियों वाले शिलालेखों के काल-निरधारण को लेकर विवाद बना रहा। एक पक्ष का कहना था कि यह लगभग 446 ईसा पूर्व से पहले का है। दूसरे पक्ष का कहना था कि यह 420 ईसा पूर्व के दशक में उकेरा गया लेख था[2]। 420 ईसा पूर्व का दशक वह समय था जब पेलोपोनेशियाई युद्ध अपने चरम पर था। यह युद्ध 431 ईसा पूर्व में शुरू हुआ था। यद्यपि काल-निरधारण में कई दशकों का अंतर आ जाए, तो इतिहास की रूपरेखा भी बदल जाती है। ऐसा इसलिए है क्योंकि वह समय पूरी तरह बदल जाता है जब एथेस ने अपनी शक्ति से सहयोगी शहरों को दबाया था।

खाली स्थानों को भरने के काम में जोखिम जुड़ा होता है। ऐसा इसलिए है क्योंकि एक वदिवान द्वारा भरा गया शब्द अगली पीढ़ी के लिए मूल पाठ जैसा पढ़ा जाने लगता है। बड़े कोष्ठक को छोड़ देने वाला केवल एक मुद्रति संस्करण भी अनुमान को तथ्य में बदल सकता है[1]। इसलिए पुरालेखवेत्ता अपने अनुमानों पर हमेशा स्पष्ट नशान छोड़ते थे।

यह भी पता लगाना जरूरी होता है कि शिलालेख मूल रूप से कहाँ स्थापित था। ग्रीक शहर अलग-अलग बोलियाँ बोलते थे। एक ही शब्द की वर्तनी भी हर शहर में अलग होती थी[2]। वदिवान वर्तनी के इस अंतर और आधिकारिक पदों के नामों को देखकर उसके प्राप्त-स्थल का अनुमान लगाते थे[4]। प्राप्त-स्थल वह स्थान है जहाँ से वह शिलालेख जमीन से निकला था।

मानव हाथों से सब कुछ नहीं पढ़ा जा सकता

पढ़े जाने वाले शिलालेख लगातार जमा होते जा रहे हैं। पैकार्ड मानविकी संस्थान (Packard Humanities Institute) द्वारा एकत्र किए गए ग्रीक शिलालेखों की संख्या 178,551 है[2]। रोमन काल के लैटिन शिलालेखों के आंकड़े भी 176,000 से अधिक हैं[3]। दूसरी ओर, दुनिया में इन शिलालेखों को पढ़ने में सक्षम वदिवानों की संख्या बहुत कम है[4]।

आंकड़ों को व्यवस्थित करना ही आसान काम नहीं है। जब इस सामग्री को मशीन द्वारा पढ़ने योग्य बनाया गया, तो केवल 78,608 प्रवृष्टियाँ ही बचीं। आधी से अधिक सामग्री बहुत छोटी होने या खराब स्थिति में होने के कारण छूट गई। केवल मशीन को देने के लिए डेटा छांटना ही अपने आप में एक बड़ा काम था[2]।

खाली स्थानों को भरना इंसानों के लिए भी कठिन काम है। एक परीक्षण में प्रशिक्षित पुरालेखवेत्ताओं को वे शिलालेख दिए गए जिनके अक्षर मटि चुके थे, और उन्हें भरने को कहा गया। इंसानों द्वारा दिए गए पहले उत्तर के सही होने का अनुपात केवल 25 प्रतिशत था[2]। इसका मतलब है कि चार में से तीन बार वे गलत थे। इससे पहले के एक प्रयोग में इंसानों द्वारा भरे गए अक्षरों की त्रुटि दर 57.3 प्रतिशत पाई गई थी[1]।

पपीरस की स्थिति और भी खराब है। पपीरस मसिर के जलस्रोतों के किनारे उगने वाली एक घास का नाम है। प्राचीन काल के लोग उस घास के भीतरी तने को पतले टुकड़ों में काटते थे। वे इन टुकड़ों को आड़े-तरिछे रखकर पीटते और दबाकर सुखाते थे। इस प्रकार बनाई गई लेखन सामग्री को भी पपीरस कहा जाता है। मसिर की सूखी रेत से ऐसी लाखों प्रतियाँ निकली हैं। अभी तक किसी के द्वारा न पढ़े गए ग्रीक पपीरस की संख्या दस लाख मानी जाती है[5]। इन पन्नों पर कर रसीदे, पत्र और अनुबंध लिखे हुए हैं। यह प्राचीन लोगों के दैनिक जीवन को जीवंत रूप में समेटे हुए अभिलेख हैं।

पढ़ने का काम केवल डबिबा खोलने से समाप्त नहीं होता। ऐसा इसलिए है क्योंकि कागज फटा हुआ है और स्याही मटि चुकी है। सबसे पहले संरक्षण उपचार करना और बखिरे हुए टुकड़ों को जोड़ना जरूरी होता है। उसके बाद ही अक्षरों को पढ़ा जा सकता है। यह सब करने के लिए योग्य लोगों की भारी कमी है[4]।

सामग्री को पढ़ने का तरीका भी एक समस्या थी। लंबे समय तक शिलालेखों का डेटा केवल कतिबों में छपी पाठ्य सूचियों के रूप में ही एकत्र हुआ। पत्थरों की तस्वीरें या उनकी बनावट की जानकारी साथ में शामिल नहीं की गई थी[4]। इसका अर्थ है कि मशीन में डाली जा सकने वाली जानकारी केवल पाठ तक ही सीमिति थी।

उत्खनन आज भी जारी है। हर साल नए पत्थर के शिलालेख और नए पपीरस जमीन से निकलते हैं। पढ़ने योग्य सामग्री बढ़ रही है, लेकिन पढ़ने वाले लोगों की संख्या जस की तस है[4]। यह अंतर साल-दर-साल बढ़ता ही गया।

इसलिए कई अभिलेख बना पढ़े ही संग्रहालय के बक्सों में बंद रह गए। पत्थर और कागज तो मौजूद थे, लेकिन उन पर लिखे वाक्यों को पुनर्जीवित करने वाले लोगों की कमी थी। शोधकर्ताओं ने इस गतिरीध को तोड़ने के लिए मशीनों की मदद ली[4]।

संदर्भ सामग्री

- [1] Assael, Y., Sommerschield, T., & Prag, J. (2019). Restoring ancient text using deep learning: a case study on Greek epigraphy. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP) (pp. 6368-6375). Association for Computational Linguistics. <https://doi.org/10.18653/v1/D19-1668>
- [2] Assael, Y., Sommerschield, T., Shillingford, B., Bordbar, M., Pavlopoulos, J., Chatzipanagiotou, M., Androutsopoulos, I., Prag, J., & de Freitas, N. (2022). Restoring and attributing ancient texts using deep neural networks. *Nature*, 603(7900), 280-283. <https://doi.org/10.1038/s41586-022-04448-z>
- [3] Assael, Y., Sommerschield, T., Cooley, A., Pavlopoulos, J., Shillingford, B., Herms, B., Suresh, P., Maynard, B., Grayston, J., Wulgaert, R., Prag, J., Mullen, A., & Mohamed, S. (2025). Contextualizing ancient texts with generative neural networks. *Nature*, 645(8079), 141-147. <https://doi.org/10.1038/s41586-025-09292-5>
- [4] Sommerschield, T., Assael, Y., Pavlopoulos, J., Stefanak, V., Senior, A., Dyer, C., Bodel, J., Prag, J., Androutsopoulos, I., & de Freitas, N. (2023). Machine learning for ancient languages: A survey. *Computational Linguistics*, 49(3), 703-747. https://doi.org/10.1162/coli_a_00481
- [5] Reply. (2026, 3११ 19११). Austrian Academy of Sciences is developing the Ancient Greek AI "Apollo" with Mistral AI and Reply. [reply.com. https://www.reply.com/en/newsroom/news/austrian-academy-of-sciences-is-developing-the-ancient-greek-ai-apollo-with-mistral-ai-and-reply](https://www.reply.com/en/newsroom/news/austrian-academy-of-sciences-is-developing-the-ancient-greek-ai-apollo-with-mistral-ai-and-reply) (3 सितंबर 2026 को अभिगमति)

11.2. वैज्ञानिक जांच दल की खोज: प्राचीन पाठ स्वतः पूर्ण करने वाले उपकरण Ithaca और Apollo

खाली स्थान वाले प्रश्नों से सीखने वाली मशीन

एआई द्वारा प्राचीन ग्रंथों को पढ़ने का तरीका भाषा की परीक्षा में आने वाले खाली स्थान भरने वाले प्रश्नों जैसा है। मशीन को बड़ी संख्या में पूर्ण शिलालेख दिखाए जाते हैं। इसके बाद कुछ अक्षरों को छुपा दिया जाता है और मशीन से अनुमान लगाने को कहा जाता है कि वहां क्या आएगा। यद्युत्तर गलत होता है, तो मशीन अपने आंतरिक मानों को थोड़ा-थोड़ा समायोजित करती है[4]। इस प्रक्रिया को लाखों बार दोहराने के बाद मशीन ग्रीक वाक्यों की संरचनात्मक आदतों को सीख लेती है।

इस तरह सीखने वाली मशीन शब्दों को रटती नहीं है। वह संभाव्यता के आधार पर याद रखती है कि किस शब्द के बाद कौन सा शब्द आने की संभावना अधिक है[4]। इसलिए पहली बार देखे गए शिलालेख में भी वह खाली स्थान के लिए उपयुक्त उम्मीदवार शब्द निकाल सकती है।

Pythia से Ithaca तक

वर्ष 2019 में शोधकर्ताओं ने Pythia नामक एक मॉडल प्रस्तुत किया। इसे Yannis Assael, Thea Sommerschield और Jonathan Prag ने बनाया था। Pythia एक तंत्रिका नेटवर्क है जो मटि हुए ग्रीक शिलालेखों के अक्षरों का पूर्वानुमान लगाता है[1]। तंत्रिका नेटवर्क एक ऐसी संगणकीय संरचना है जो डेटा देखकर स्वयं नयिम सीखती है। इसे मानव मस्तिष्क के तंत्रिका तंत्र के संपर्कों के आधार पर तैयार किया गया है।

Pythia अक्षर स्तर और शब्द स्तर दोनों को एक साथ पढ़ता है। ऐसा इसलिए है क्योंकि आधे टूटे हुए शब्दों वाले डेटा को संभालने के लिए दोनों दृष्टिकोणों की आवश्यकता होती है। जब इंसानों द्वारा भरे गए अक्षरों की त्रुटि दर 57.3 प्रतिशत थी, तब Pythia ने 30.1 प्रतिशत की त्रुटि दर दर्ज की। सही उत्तर के Pythia के शीर्ष 20 उम्मीदवारों में शामिल होने का अनुपात 73.5 प्रतिशत था[1]।

वर्ष 2022 में Ithaca सामने आया। इसे Google DeepMind और कई विश्वविद्यालयों ने मलिकर वकिसति किया था। इसमें ऑक्सफोर्ड विश्वविद्यालय, वेनिस का Ca' Foscari विश्वविद्यालय और एथेंस अर्थशास्त्र और व्यवसाय विश्वविद्यालय शामिल थे। शोध के परिणाम मार्च 2022 में वैज्ञानिक पत्रिका Nature में प्रकाशित हुए थे[2]। इसका नाम महाकाव्य ओडिसी (Odyssey) में वर्णित द्वीप Ithaca के नाम पर रखा गया था। इसमें लंबे भटकाव के बाद घर लौटने का भाव समाहित है।

Ithaca नरिणय कैसे लेता है

Ithaca 'Transformer' नामक संरचना का उपयोग करता है। Transformer एक ऐसा तंत्रिका नेटवर्क है जो यह आकलन करता है कि वाक्य के शब्द एक-दूसरे से कतिने संबंधित हैं। जहां अक्षर मटि होते हैं, उन्हें खाली स्थान दर्शाने वाले प्रतीक से बदलकर इनपुट किया जाता है। मॉडल उस प्रतीक की जगह आने वाले अक्षरों की संभावना की गणना करता है[2]।

Ithaca तीन कार्य करता है। यह मटि हुए अक्षरों को भरता है, यह अनुमान लगाता है कि शिलालेख मूल रूप से कहाँ था, और यह नरिधारित करता है कि वह कब उकेरा गया था। यह इन तीनों कार्यों को अलग-अलग नहीं बल्कि एक साथ सीखता है। बोलियाँ और वर्तनी की आदतें स्थान का संकेत देने वाले सुराग हैं। शब्दों का प्रयोग और वाक्य संरचनाएं काल का संकेत देने वाले सुराग हैं। जब एक ही तंत्रिका नेटवर्क इन सभी सुरागों को एक साथ संसाधित करता है, तो वे एक-दूसरे की सहायता करते हैं[2]।

Ithaca पत्थर की तस्वीरों को नहीं देखता। वह केवल उतारे गए पाठ को पढ़कर नरिणय लेता है[2]। यह स्ट्रोक के आकार से काल का नरिधारण करने वाली पुरानी पद्धति से बिल्कुल अलग दृष्टिकोण है।

इसके द्वारा काल-नरिधारण प्रस्तुत करने का तरीका भी अलग है। Ithaca किसी एक निश्चित वर्ष की ओर इशारा नहीं करता। वह 800 ईसा पूर्व से 800 ईस्वी तक की अवधि को दस-दस वर्षों के अंतरालों में विभाजित करता है। फिर वह प्रत्येक अंतराल के लिए संभाव्यता नरिधारित करके एक ग्राफ के रूप में प्रदर्शित करता है[2]। विद्वान ग्राफ के शिखरों को देखकर यह नरिणय लेते हैं कि कौन सा काल सबसे अधिक संभावित है।

इसके प्रशिक्षण में उपयोग की गई सामग्री पैकार्ड मानविकी संस्थान (Packard Humanities Institute) का शिलालेख संग्रह है। ऊपर उल्लिखित 78,608 प्रविष्टियाँ ही वह सामग्री हैं। दस से अधिक बार आने वाले शब्दों की गिनती ही 35,884 थी[2]।

शोधकर्ताओं ने Ithaca का निर्माण करते समय तीन लक्ष्य नरिधारित किए थे। विद्वानों के साथ काम करना, उनके नरिणय में सहायता करना, और नरिणय की प्रक्रिया को प्रदर्शित करना। इंसानों की जगह लेने वाली मशीन बनाना उनका लक्ष्य नहीं था। उनका उद्देश्य एक ऐसा उपकरण बनाना था जो विद्वानों की कार्यप्रणाली का अभिन्न अंग बन सके[2]।

Ithaca केवल एक ही उत्तर नहीं देता। वह खाली स्थान में आने वाले कई उम्मीदवारों को उनकी संभावना के साथ प्रदर्शित करता है। वह स्क्रीन पर यह भी दिखाता है कि किनी अक्षरों को देखकर उसने ऐसा नरिणय लिया। विद्वान अपनी आंखों से देख सकते हैं कि मशीन ने किस आधार पर उत्तर दिया है। अंतिम नरिणय मनुष्य की जम्मेदारी पर छोड़ दिया जाता है। शोध दल ने Ithaca का कोड और डेटा सार्वजनिक कर दिया है[2]।

बीस उम्मीदवारों का अर्थ

मशीन द्वारा दिया जाने वाला उत्तर केवल एक नहीं होता। वह उच्चतम संभावना के क्रम में उम्मीदवारों को सूचीबद्ध करती है। पहले उम्मीदवार के सही होने के अनुपात को शीर्ष-1 सटीकता कहा जाता है। शीर्ष बीस विकल्पों में सही उत्तर होने के अनुपात को शीर्ष-20 सटीकता कहा जाता है[1]।

यह अंतर महत्वपूर्ण है। विद्वान को केवल एक सही उत्तर की आवश्यकता नहीं होती। उन्हें समीक्षा के योग्य उम्मीदवारों की एक सूची चाहिए होती है। यदि शीर्ष बीस में सही उत्तर मौजूद हो, तो विद्वान उसमें से चयन कर सकते हैं[1]। कई बार वे उस सूची में ऐसे शब्द भी खोज लेते हैं जो स्वयं उनके ध्यान में नहीं आए थे[2]।

लैटिन की ओर बढ़ता Aeneas

उसी शोध दल ने अपना दायरा लैटिन शिलालेखों तक बढ़ाया। उन्होंने जुलाई 2025 में Nature पत्रिका में Aeneas नामक एक मॉडल प्रस्तुत किया[3]। इसका नाम रोम की स्थापना की गाथा के नायक Aeneas के नाम पर रखा गया था[5]।

Aeneas को 176,000 से अधिक लैटिन शिलालेखों के डेटा से प्रशिक्षित किया गया था। यह रोम, हीडलबर्ग और Clauss-Slaby के तीन शिलालेख डेटाबेसों को मिलाकर तैयार की गई सामग्री थी। दस अक्षरों से कम मटि हुए स्थानों में, सही उत्तर के शीर्ष 20 उम्मीदवारों में होने का अनुपात 73 प्रतिशत था। यह पता न होने पर भी कतिने अक्षर मटि थे, इसने 58 प्रतिशत का स्तर हासिल किया[3]। मटि हुए हसिसे की लंबाई जाने बिना उत्तर देना इंसानों के लिए भी एक कठिन चुनौती है।

इसने यह भी सही अनुमान लगाया कि शिलालेख रोम के किस प्रांत से आया था। प्रांत रोम द्वारा शासित क्षेत्रीय भूमि होती थी। इसने 62 विकल्पों में से 72 प्रतिशत सही उत्तर दिए[5]। उकेरे जाने के काल को इसने इतिहासकारों द्वारा नरिधारित सीमा से औसतन 13 वर्षों के भीतर सीमित कर दिया[3, 5]।

Aeneas में एक नई सुविधा है। यह नरिणय लेने के लिए शलिलेखों की तस्वीरों को भी शामिल कर सकता है[3]। इसका अर्थ है कथिह केवल उतारे गए अक्षरों को ही नहीं, बल्कि पित्थर के भौतिक स्वरूप को भी देखता है। यह समान वाक्यों वाले अन्य शलिलेखों को खोजकर उन्हें साथ-साथ प्रदर्शित भी करता है। जनि समान उदाहरणों को खोजने में वदिवानों को कई दनि लग जाते थे, उन्हें यह कुछ ही सेकंडों में सामने ला देता है[5]।

शोधकर्ताओं ने इस उपकरण को रोमन सम्राट ऑगस्टस के अभिलेखों पर आजमाया। Aeneas ने उस लेख के उत्कीर्ण होने के समय के रूप में दो अवधियाँ प्रस्तुत कीं। एक 10 ईसा पूर्व से 1 ईसा पूर्व के बीच थी। दूसरी 10 ईस्वी से 20 ईस्वी के बीच थी और इसकी संभावना अधिक थी[5]। अकादमिक जगत में बँटे दो दावे संभावना के ग्राफ़ पर साफ़ दिखाई दिए।

शोधकर्ताओं ने 23 इतिहासकारों को इस सुविधा का उपयोग करने दिया। वदिवानों ने उत्तर दिया कि उन्हें दस में से नौ अभिलेखों में नए शोध वचिर मिले[3, 5]। यह मॉडल सभी के उपयोग के लिए सार्वजनिक किया गया है[5]।

संपूर्ण ग्रीक भाषा सीखने वाला Apollo

मार्च 2026 में ऑस्ट्रियाई वज्जान अकादमी ने एक नए मॉडल की घोषणा की। इसका नाम Apollo है। इसे एआई कंपनी Mistral AI और सूचना प्रौद्योगिकी कंपनी Sail Reply मिलकर बना रही है[7]। यदि Ithaca किसी एक अभिलेख को संभालने का उपकरण है, तो Apollo संपूर्ण ग्रीक भाषा सीखने वाले मॉडल का लक्ष्य रखता है[6]।

ऐसे मॉडल को फाउंडेशन मॉडल कहा जाता है। इसका अर्थ है वभिन्न कार्यों में व्यापक उपयोग के लिए बड़े पैमाने पर सखाया गया एक बुनियादी मॉडल। यह केवल एक ही काम करने के लिए बनाया गया उपकरण नहीं है[4]। अभिलेख पुनर्बहाली, साहित्य खोज और लिखावट पढ़ने में भी इसी आधार का उपयोग होता है[6]।

Apollo जसि सामग्री से सीखता है, वह लगभग 60 करोड़ शब्दों का ग्रीक साहित्य है। इसमें पहले से प्रकाशित हजारों अभिलेख और पपीरस भी जोड़े गए हैं[6, 7]। शोधकर्ताओं ने बताया कि यह अब तक एकत्र की गई ग्रीक सामग्री में सबसे बड़े पैमाने की सामग्री है[7]।

ऑस्ट्रियाई राष्ट्रीय पुस्तकालय के पास बहुत सारे पपीरस हैं। अकेले इसी पुस्तकालय में हजारों पपीरस संरक्षित हैं[7]। इनमें ऐसे दस्तावेज़ भी मिले हुए हैं जनिहें अभी तक पढ़ा नहीं गया है।

इस शोध का नेतृत्व ऑस्ट्रियाई पुरातत्व संस्थान की अन्ना डोल्गानोव कर रही हैं। उन्होंने कहा कि दुनिया में ऐसे 10 लाख ग्रीक पपीरस हैं जनिहें अभी तक किसी ने नहीं पढ़ा है[7]। Apollo ऐसी सामग्रियों से अक्षरों को पुनर्बहाल करने और सामग्री खोजने का काम संभालता है। हस्तलिखित अक्षरों को पढ़ना भी इसके लक्ष्यों में शामिल है[6, 7]। शोधकर्ताओं ने बताया कि यह काम कई वर्षों के बजाय कुछ ही घंटों में पूरा हो सकता है[7]।

संदर्भ सामग्री

- [1] Assael, Y., Sommerschield, T., & Prag, J. (2019). Restoring ancient text using deep learning: a case study on Greek epigraphy. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP) (pp. 6368-6375). Association for Computational Linguistics. <https://doi.org/10.18653/v1/D19-1668>
- [2] Assael, Y., Sommerschield, T., Shillingford, B., Bordbar, M., Pavlopoulos, J., Chatzipanagiotou, M., Androutsopoulos, I., Prag, J., & de Freitas, N. (2022). Restoring and attributing ancient texts using deep neural networks. *Nature*, 603(7900), 280-283. <https://doi.org/10.1038/s41586-022-04448-z>
- [3] Assael, Y., Sommerschield, T., Cooley, A., Pavlopoulos, J., Shillingford, B., Herms, B., Suresh, P., Maynard, B., Grayston, J., Wulgaert, R., Prag, J., Mullen, A., & Mohamed, S. (2025). Contextualizing ancient texts with generative neural networks. *Nature*, 645(8079), 141-147. <https://doi.org/10.1038/s41586-025-09292-5>
- [4] Sommerschield, T., Assael, Y., Pavlopoulos, J., Stefanak, V., Senior, A., Dyer, C., Bodel, J., Prag, J., Androutsopoulos, I., & de Freitas, N. (2023). Machine learning for ancient languages: A survey. *Computational Linguistics*, 49(3), 703-747. https://doi.org/10.1162/coli_a_00481
- [5] Google DeepMind. (2025, 7월 23일). Aeneas transforms how historians connect the past. [deepmind.google](https://deepmind.google/blog/aeneas-transforms-how-historians-connect-the-past/). <https://deepmind.google/blog/aeneas-transforms-how-historians-connect-the-past/> (3 सितंबर 2026 को अभिगमिति)
- [6] Mistral AI. (2026). Austrian Academy of Sciences. [mistral.ai](https://mistral.ai/customers/austrian-academy-of-sciences/). <https://mistral.ai/customers/austrian-academy-of-sciences/> (3 सितंबर 2026 को अभिगमिति)
- [7] Reply. (2026, 3월 19일). Austrian Academy of Sciences is developing the Ancient Greek AI "Apollo" with Mistral AI and Reply. [reply.com](https://www.reply.com/en/newsroom/news/austrian-academy-of-sciences-is-developing-the-ancient-greek-ai-apollo-with-mistral-ai-and-reply). <https://www.reply.com/en/newsroom/news/austrian-academy-of-sciences-is-developing-the-ancient-greek-ai-apollo-with-mistral-ai-and-reply> (3 सितंबर 2026 को अभिगमिति)

11.3. उजागर हुआ सच

62 प्रतशित और 72 प्रतशित

शोधकर्ताओं ने Ithaca की क्षमता की पुष्टि आंकड़ों से की। यह एक ऐसी परीक्षा थी जिसमें जानबूझकर अक्षरों को छुपाए गए अभिलिख देकर उन्हें भरने के लिए कहा गया। यही परीक्षा प्रशिक्षित पुरालेखवदों को भी दी गई[1]।

जब मनुष्य ने अकेले इसे हल किया, तो पहले उत्तर के सही होने की दर 25 प्रतशित थी। जब Ithaca ने अकेले हल किया, तो यह दर 62 प्रतशित थी। यह दोगुने से भी अधिक का अंतर है[1]।

इस परीक्षा में एक तीसरी शर्त भी थी। वह शर्त थी कि मनुष्य Ithaca की संभावित विकल्पों की सूची देखते हुए साथ मलिकर हल करे। तब सटीकता बढ़कर 72 प्रतशित हो गई[1]। यह आंकड़ा अकेले मनुष्य और अकेली मशीन दोनों से अधिक है।

यह संख्या इस शोध का केंद्र बंदी है। मशीन 78,608 अभिलिखों को पढ़ने की अपनी स्मृतिके आधार पर संभावना की गणना करती है। मनुष्य जानता है कि वह अभिलिख किस इतिहास से जुड़ा है। जब मशीन बीस संभावित विकल्प रखती है, तो मनुष्य अपने ऐतिहासिक ज्ञान से उन्हें छाँटता है। जब दोनों एक साथ आए, तब सटीकता सबसे अधिक रही[1, 2]।

स्थान और समय का अनुमान लगाने की क्षमता की भी पुष्टि की गई। Ithaca ने ग्रीक जगत को 84 क्षेत्त्रों में विभाजित करके सीखा था। इसने अज्ञात मूल वाले अभिलिखों के वास्तविक स्थान का 71 प्रतशित सटीकता से अनुमान लगाया। उत्कीर्णन के समय के लिए इसने इतिहासकारों द्वारा निर्धारित सीमा से औसतन 29.3 वर्ष के अंतर वाला मान दिया। मानव आँखों से लिखावट की शैली देखकर अनुमान लगाने के तरीके में त्रुटि सौ वर्षों तक पहुँच जाती थी[1]।

421 ईसा पूर्व का उत्तर

Ithaca ने एक पुराने विवाद का भी उत्तर दिया। यह पहले देखी गई एथेंस संधिके अभिलिखों के काल-निर्धारण की समस्या है। तीन रेखाओं वाले सिग्मा के कारण इन अभिलिखों को 446 ईसा पूर्व के आसपास से पहले का माना जाता था[1]।

शोधकर्ताओं ने इन अभिलिखों को Ithaca में डाला। Ithaca अक्षरों के आकार के बजाय स्वयं वाक्य को देखकर निर्णय लेता है। यह जाँचता है कि कौन सा शब्द किस शब्द के साथ मलिकर आता है। यह यह भी देखता है कि कौन सा पदनाम और कौन सी अव्यय एक साथ आती है[1]।

Ithaca द्वारा दिया गया औसत काल 421 ईसा पूर्व था। यह पुराने मानक से एक पीढ़ी बाद का है। यह मान सामने आए अन्य नए साक्ष्यों की दृष्टि के समान था। हाल ही में विद्वानों द्वारा दोबारा निर्धारित किए गए काल से इसमें औसतन केवल 5 वर्ष का अंतर था। पुराने मानक के अनुसार तय किए गए काल में 27 वर्ष का अंतर था[1]।

मशीन ने अकेले इतिहास को दोबारा नहीं लिखा। इसने लंबे समय से चले आ रहे दो दावों में से एक को अधिक वजन दिया[1, 2]। विद्वान केवल लिखावट शैली के आधार पर काल तय करने की अपनी आदत पर दोबारा सोचने लगे। एक उपकरण ने सौ साल पुरानी बहस की दृष्टि बदल दी।

मशीन द्वारा गढ़े गए विश्वसनीय लगने वाले वाक्य

इन उपलब्धियों के बावजूद कुछ समस्याएँ बनी हुई हैं। वह यह है कि सीखी गई सामग्री एकतरफा है। ग्रीक अभिलिखों की सामग्री में एथेंस के अभिलिख अधिक हैं। काल भी शास्त्रीय काल में केंद्रित है। कम सामग्री वाले सीमांत क्षेत्रों के अभिलिखों के मामले में यह उतनी ही कम सटीक है[2]।

यह समस्या भी है कि यह समझना मुश्किल है कि मशीन ने ऐसा उत्तर क्यों दिया। तंत्रिका नेटवर्क में लाखों संख्याएँ होती हैं। यह यह तो दिखाता है कि इसने किस अक्षर पर अधिक ध्यान दिया, लेकिन उस उत्तर को चुनने के कारण को शब्दों में नहीं समझा सकता[2]।

जब रचित स्थान लंबा होता है, तब भी समस्या आती है। यह कुछ अक्षरों को तो ठीक से भर देता है, लेकिन जहाँ कई शब्द पूरी तरह गायब हों, वहाँ मुश्किल होती है। Aeneas के परीक्षण में भी जब मॉडल हुआ हिससे की लंबाई अज्ञात थी, तो परिणाम 73 प्रतशित से घटकर 58 प्रतशित रह गया[3]। इस समय मॉडल व्याकरण सम्मत और विश्वसनीय लगने वाले ऐसे वाक्य बना सकता है जो वास्तव में मौजूद नहीं थे। इस तरह गढ़े गए उत्तर को विभ्रम (hallucination) कहा जाता है। इसका अर्थ है मशीन द्वारा सच की तरह गढ़ा गया गलत उत्तर। प्राचीन अभिलिखों का अध्ययन करते समय विभ्रम खतरनाक होता है। क्योंकि जो इतिहास था ही नहीं, वह सच के रूप में दर्ज हो सकता है[2]।

अगस्त 2026 में सामने आए एक अध्ययन ने इस सीमा को सीधे उठाया। यह कई मॉडलों और कई स्थितियों में क्षतग्रस्त प्राचीन दस्तावेजों को पुनर्बहाल करने का एक प्रयोग था। शोधकर्ताओं ने निष्कर्ष निकाला कि इस बहाली को पूरी तरह स्वचालित नहीं छोड़ा जा सकता। दस्तावेज के किस हिस्से को भरा जा रहा है, इसके आधार पर प्रदर्शन में बड़ा अंतर आया। जिन दस्तावेजों में तय वाक्यांश दोहराए गए थे,

उन्हें इसने अच्छी तरह भरा। अलग-अलग लोगों द्वारा लिखे गए वाक्यों में कठिनाई आई। मटि हुए हस्त्रिसे की लंबाई ज्ञात थी या नहीं, इस पर भी परणाम नरिभर करता था। शोधकर्ताओं ने बताया कमिशीन को पुरालेखवर्दों की सहायता करने वाले उपकरण के रूप में वकिसति कयिा जाना चाहिए[4]।

शोधकर्ताओं ने इस जोखमि को कम करने के लिए नयिम बनाए। मशीन द्वारा भरे गए अक्षरों पर भी कोषठक लगाए जाते हैं। कम संभावना वाले वकिल्पो को वदिवान अलग कर देते हैं। प्राचीन पुरालेखवर्दों द्वारा माने जाने वाले अंकन नयिमों को मशीन पर भी उसी तरह लागू कयिा जाता है[2]।

इसलिए Ithaca द्वारा दयिा गया उत्तर कोई ऐतहिसकि तथ्य नहीं है। यह समीक्षा के लिए परकिल्पनाओं की एक सूची है। जब कोई वदिवान अन्य सामग्रयिों से तुलना करके इसकी पुष्टकिरता है, तभी यह ऐतहिसकि साक्ष्य बनता है[2]। ऐतहिसकि साक्ष्य इतहिस को उजागर करने वाली आधार सामग्री है।

सात वर्षों में जो हुआ

इस क्षेत्र में बदलाव की गतकिी तुलना आंकड़ों से की जा सकती है। 2019 में Pythia ने मानव त्रुटिदर को 57.3 प्रतिशत से घटाकर 30.1 प्रतिशत कर दयिा[5]। 2022 में Ithaca ने पहले उत्तर की सटीकता बढ़ाकर 62 प्रतिशत कर दी[1]। 2025 में Aeneas लैटनि भाषा की ओर बढ़ा और काल-नरिधारण की त्रुटि को औसतन केवल 13 वर्षों तक सीमति कर दयिा[3]। 2026 में संपूर्ण ग्रीक भाषा सीखने वाला Apollo शुरू हुआ[6]।

सात वर्षों में लक्षति भाषाएँ बढ़ी और सटीकता में सुधार हुआ। अध्ययन की जाने वाली सामग्री भी केवल अक्षरों से बढ़कर तस्वीरों तक फैल गई[3]। इस दौरान एक चीज़ नहीं बदली। वह यह सदिधांत है कि अंतमि नरिणय मनुष्य ही करता है[2]।

तस्वीरें और संदर्भ, अगला कदम

शोध का दायरा लगातार बढ़ रहा है। 2026 में आए एक शोध पत्र में व्यापक रूप से इस्तेमाल होने वाले एक बड़े भाषा मॉडल को ग्रीक सामग्री पर दोबारा प्रशक्षति कयिा गया। यह Meta द्वारा जारी कएि गए Llama 3.1 मॉडल का प्रयोग था। पपीरस की बहाली में पहले उत्तर के सही होने की दर 73.5 प्रतिशत रही। अभलिखों की बहाली में यह 63.7 प्रतिशत आई। शीर्ष 20 वकिल्पो में सही उत्तर होने की दर क्रमशः 86.0 प्रतिशत और 83.0 प्रतिशत थी। इसका अर्थ है कि भले ही कोई मॉडल वशिष रूप से अभलिखों के लिए न बना हो, फिर भी वह यह काम कर सकता है[7]।

अक्षरों की छवयिों को पुनरुस्थापति करने वाला शोध भी सामने आया है। अप्रैल 2026 में सामने आया Mesa अभलिखों की तस्वीरों में मटि हुए नशानों को भरने का एक तरीका है। उसी शलिापट्ट पर अच्छी स्थति में बचे अक्षरों को उदाहरण मानकर यह घसि हुए स्थानों की बनावट को दोबारा उभारता है[8]। वाक्यों को भरने वाले मॉडल और तस्वीरों को पूरा करने वाले मॉडल साथ-साथ वकिसति हो रहे हैं।

खोई हुई त्रासदयिों भी वास्तव में वापस लौट रही हैं। 2022 में मसिर के फलिाडेल्फयिा के अवशेषों से एक पपीरस मलिा। यह तीसरी शताब्दी ईस्वी का माना जाने वाला दस्तावेज़ है। 2024 में वदिवानों ने इसे पढ़कर इसकी सामग्री प्रकाशति की। यह यूरपिडिस द्वारा रचति त्रासदयिों 'पॉलीइडस' और 'इनो' के संवादों की 97 पंक्तयिों थीं। तब तक ये दोनों कृतयिों केवल संक्षप्ति उद्धरणों और कथानक के रूप में ही जानी जाती थीं[9]।

यह बहाली मनुष्य के हाथों और आँखों ने की। यह किसी मशीन का कयिा काम नहीं था। दुनयिा में ऐसे 10 लाख और दस्तावेज़ बचे हुए हैं। केवल मानवीय आँखों से इन दस्तावेज़ों के ढेरों को पूरी तरह नहीं पढ़ा जा सकता। Apollo जैसे मॉडल का लक्ष्य ठीक यही है[6]।

उपकरणों के सार्वजनकि होने के साथ शोध का तरीका भी बदल गया। Ithaca और Aeneas के कोड और डेटा सार्वजनकि रूप से उपलब्ध हैं[1, 10]। वशि्ववदिव्यालय की प्रयोगशाला न होने पर भी कोई भी इसे वेब पर इस्तेमाल कर सकता है। किसी अभलिख की एक तस्वीर अपलोड करने पर संभावति वकिल्पो की सूची, नक्शा और काल-नरिधारण का ग्राफ सामने आ जाता है[10]। पहले इस तुलनात्मक काम में कई साल लग जाते थे[3]।

अभी भी बहुत सा काम बाकी है। अकेले ग्रीक पपीरस के ही 10 लाख दस्तावेज़ बनिा पढ़े पड़े हैं[6]। अभलिखीय सामग्री क्षेत्र और काल के मामले में असमान है[2]। मशीन द्वारा ठीक से न संभाले जाने वाले लंबे छूटे हुए हस्त्रिसे भी वैसे ही बने हुए हैं[3]। इन समस्याओं को बेहतर सामग्री और अधिक सत्यापन के माध्यम से एक-एक करके कम कयिा जाना चाहिए[4]।

मशीन वकिल्प प्रस्तुत करती है और मनुष्य चुनता है। Ithaca को 62 प्रतिशत से बढ़ाकर 72 प्रतिशत करने वाला तरीका ही इसका उत्तर है[1]। शांत पड़े पत्थरों और कागज़ों को फिर से बोलने के लिए दोनों का साथ होना आवश्यक है।

- [1] Assael, Y., Sommerschield, T., Shillingford, B., Bordbar, M., Pavlopoulos, J., Chatzipanagiotou, M., Androutsopoulos, I., Prag, J., & de Freitas, N. (2022). Restoring and attributing ancient texts using deep neural networks. *Nature*, 603(7900), 280-283. <https://doi.org/10.1038/s41586-022-04448-z>
- [2] Sommerschield, T., Assael, Y., Pavlopoulos, J., Stefanak, V., Senior, A., Dyer, C., Bodel, J., Prag, J., Androutsopoulos, I., & de Freitas, N. (2023). Machine learning for ancient languages: A survey. *Computational Linguistics*, 49(3), 703-747. https://doi.org/10.1162/coli_a_00481
- [3] Assael, Y., Sommerschield, T., Cooley, A., Pavlopoulos, J., Shillingford, B., Herms, B., Suresh, P., Maynard, B., Grayston, J., Wulgaert, R., Prag, J., Mullen, A., & Mohamed, S. (2025). Contextualizing ancient texts with generative neural networks. *Nature*, 645(8079), 141-147. <https://doi.org/10.1038/s41586-025-09292-5>
- [4] Zhang, S., Caraffa, E., Zuffrano, A., Modesti, M., & Colavizza, G. (2026, 8월 28일). Text restoration of ancient documents with language models (arXiv:2608.28170). arXiv. <https://arxiv.org/abs/2608.28170> (3 सितंबर 2026 को अभिगमिति)
- [5] Assael, Y., Sommerschield, T., & Prag, J. (2019). Restoring ancient text using deep learning: a case study on Greek epigraphy. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)* (pp. 6368-6375). Association for Computational Linguistics. <https://doi.org/10.18653/v1/D19-1668>
- [6] Reply. (2026, 3월 19일). Austrian Academy of Sciences is developing the Ancient Greek AI "Apollo" with Mistral AI and Reply. [reply.com. https://www.reply.com/en/newsroom/news/austrian-academy-of-sciences-is-developing-the-ancient-greek-ai-apollo-with-mistral-ai-and-reply](https://www.reply.com/en/newsroom/news/austrian-academy-of-sciences-is-developing-the-ancient-greek-ai-apollo-with-mistral-ai-and-reply) (3 सितंबर 2026 को अभिगमिति)
- [7] Cullhed, E. (2026). Instruction-tuning pretrained causal language models to restore ancient Greek papyri and inscriptions. *Digital Scholarship in the Humanities*, 41(1), 69-76. <https://doi.org/10.1093/llc/fqaf131>
- [8] Toulatzis, V., Theodoridou, S., & Fudos, I. (2026, 4월 19일). MESA: A training-free multi-exemplar deep framework for restoring ancient inscription textures (arXiv:2604.17390). arXiv. <https://arxiv.org/abs/2604.17390> (3 सितंबर 2026 को अभिगमिति)
- [9] O'Grady, E. (2024, 10월 11일). Unearthed papyrus contains lost scenes from Euripides' plays. *The Harvard Gazette*. <https://news.harvard.edu/gazette/story/2024/10/unearthed-papyrus-contains-lost-scenes-from-euripides-plays/> (3 सितंबर 2026 को अभिगमिति)
- [10] Google DeepMind. (2025, 7월 23일). Aeneas transforms how historians connect the past. [deepmind.google. https://deepmind.google/blog/aeneas-transforms-how-historians-connect-the-past/](https://deepmind.google/blog/aeneas-transforms-how-historians-connect-the-past/) (3 सितंबर 2026 को अभिगमिति)

अध्याय 12. कभी न जलने वाला आभासी पुस्तकालय: अतः-उच्च-रजिऑल्यूशन बहुवर्णक पुरालेख और भवष्य

12.1. घटना फ़ाइल

अलेक्जेंड्रिया का पुस्तकालय कैसे लुप्त हुआ

मसिर के अलेक्जेंड्रिया में प्राचीन विश्व का सबसे बड़ा पुस्तकालय था। टॉलेमी राजवंश ने इसे तीसरी शताब्दी ईसा पूर्व के आसपास स्थापित किया था। राजा दुनिया भर की कृतिबों को एक जगह इकट्ठा करना चाहते थे। बंदरगाह पर आने वाले जहाजों से कृतिबें ले ली जाती थीं और उनकी नकल तैयार करवाई जाती थी। नकल मालिक को लौटा दी जाती थी और मूल पुस्तकालय रख लेता था। प्राचीन अभिलेख बताते हैं कि वहाँ लाखों ग्रंथ-चर्मपत्र थे [1]।

कई लोग मानते हैं कि यह पुस्तकालय एक भीषण आग में नष्ट हो गया था। लेकिन शोधकर्ताओं का मानना है कि इसका पतन धीरे-धीरे हुआ था। 145 ईसा पूर्व में राजा ने विदेशी विद्वानों को निकाल बाहर किया। विद्वानों के जाने के साथ ही कृतिबों की देखभाल करने वाले हाथ भी चले गए। रोमन काल में इमारत कई बार क्षतिग्रस्त हुई। नई कृतिबें खरीदने के लिए मलिन वाला धन भी बंद हो गया [1]।

प्राचीन काल के लेख वैसे भी लंबे समय तक सुरक्षित नहीं रह पाते। उस समय के मुंशी नई रचनाएँ लिखने के बजाय मौजूद कृतिबों की नकल करने का काम अधिक करते थे। हाथों-हाथ पढ़े जाने वाले पपीरस ग्रंथ-चर्मपत्र लगभग सौ साल बाद पुराने पड़ जाते थे और उन्हें पढ़ना मुश्किल हो जाता था। शुष्क मसिर की जमीन के नीचे वे कहीं अधिक लंबे समय तक सुरक्षित भी रहे। इसलिए जब कोई समय पर फरि से उनकी नकल करता था, तभी वह लेख आगे जीवित रह पाता था। यदि नकल करने वाले हाथ रुक जाते, तो वह कृतिब दुनिया से मिट जाती थी।

आज हम जो प्राचीन साहित्य पढ़ते हैं, वह उस युग के लेखन का एक छोटा सा हिस्सा भर है। ऐसी कृतिबें कहीं अधिक हैं जिनके केवल नाम बचे हैं और मूल पाठ गायब हो चुका है। प्राचीन यूनानी त्रासदी नाटककारों द्वारा रचित अधिकांश रचनाएँ भी इसी तरह नष्ट हो गईं। जो कृतिबें बची रहीं, वे भाग्यशाली थीं क्योंकि उनकी प्रतियाँ कई जगहों पर बखिरी हुई थीं।

वर्नाश आज भी जारी है

अभिलेखों को नष्ट करने वाली ताकतें केवल अतीत तक सीमित नहीं थीं। युद्ध आज भी पुस्तकालयों और अभिलेखागारों को ध्वस्त कर देता है। यूनेस्को यूक्रेन में नष्ट हुई सांस्कृतिक सुविधाओं की एक-एक कर पुष्टि और गणना कर रहा है। 29 जुलाई 2026 तक पुष्टि किए गए स्थलों की संख्या 545 है। इनमें 43 संग्रहालय, 22 पुस्तकालय और 5 अभिलेखागार शामिल हैं। धार्मिक स्थल और सैकड़ों पुरानी इमारतें भी इसके साथ क्षतिग्रस्त हुई [2]।

युद्ध केवल पुरावशेषों को नष्ट ही नहीं करता। प्रबंधन रुक जाने पर पुरावशेष भंडारों से कलाकृतियाँ चोरी-छिपे गायब भी हो जाती हैं। जसि पुरावशेष का यह विवरण मिट जाता है कि वह कहाँ से मिला था, उसका शोध मूल्य काफी हद तक घट जाता है। क्योंकि मिट्टी की एक पट्टिका कसि कमरे में कनि अन्य पट्टिकाओं के साथ रखी थी, यह अपने आप में एक महत्वपूर्ण जानकारी है।

जलवायु भी अभिलेखों के लिए खतरा बनती है। समुद्र तट पर स्थित अवशेष लहरों के कटाव से धीरे-धीरे ढह जाते हैं। बारिश बढ़ने या भूजल स्तर बढ़ने पर जमीन के अंदर दबे पुरावशेष और तेजी से खराब होते हैं। कुछ स्थान ऐसे भी हैं जहाँ उनके लुप्त होने की गति उत्खनन की गति से अधिक तेज है।

क्षतिग्रस्त पुरावशेष पूरी तरह समाप्त नहीं होते। यदि उनका आकार बचा रहे, तो उन्हें फरि से पुनर्निर्मित करने का रास्ता होता है। इसलिए कोई दुर्घटना होने से पहले ही उनका माप ले लेना महत्वपूर्ण होता है। युद्धग्रस्त क्षेत्रों में इमारतों और पुरावशेषों का 3D में मापन करने का कार्य चल रहा है।

हर सामग्री का जीवनकाल अलग होता है

अभिलेख दर्ज करने वाली सामग्रियाँ अलग-अलग तरीकों से नष्ट होती हैं। मेसोपोटामिया की मिट्टी की पट्टिकाएँ गीली मिट्टी को दबाकर उन पर अक्षर उकेर कर बनाई गई पट्टिकाएँ हैं। महल में आग लगने पर पैदा हुई गर्मी ने मिट्टी की पट्टिकाओं को चीनी मिट्टी के बर्तनों की तरह सख्त पका दिया। आग के संपर्क में न आने वाली मिट्टी की पट्टिकाएँ पानी लगने पर फरि से कीचड़ में बदल जाती हैं।

मिट्टी की पट्टिकाओं की एक अन्य सीमा भी है। मुंशी उपयोग की जा चुकी पट्टिकाओं को पानी में भगीकर उनकी सतह को चिकना करते थे और दोबारा लिखते थे। इसलिए पुरानी पट्टिकाएँ आमतौर पर वापस मिट्टी में तब्दील हो गईं। यही कारण है कि अक्सर साम्राज्य के पतन से ठीक पहले के कुछ वर्षों के अभिलेख ही बच पाए हैं। क्योंकि महल जलते समय अभिलेखागार में मौजूद अंतिम पट्टिकाएँ ही सख्त होकर पक पाई थीं।

पपीरस, चमड़ा और कागज और भी अधिक नाजुक होते हैं। ये सामग्रियाँ पौधों और जानवरों से प्राप्त जैविक पदार्थ हैं। गर्म और आर्द्र स्थानों में ये कुछ सौ साल भी नहीं टकि पाते और सड़ जाते हैं। रोमन सैनिकों द्वारा इस्तेमाल की जाने वाली मोम की पट्टियों से मोम पूरी तरह गायब हो चुका है। पीछे केवल वही उथले खरोच बचे हैं जो लोहे की कलम के लकड़ी की तख्ती को खुरचने से बने थे।

कठोर सामग्रियाँ भी सुरक्षित नहीं हैं। कछुए के पेट के कवच और बैल के कंधे की हड्डी पर उकेरी गई ओरेकल बोन लपि भी जमीन के नीचे दबकर टूट गई। आज बची हुई अधिकांश ओरेकल हड्डियाँ हथेली से भी छोटे टुकड़े हैं। यहाँ तक कि एक ही वाक्य तीन टुकड़ों में बखिर कर अलग-अलग गोदामों में पहुँचने के मामले भी हैं [3]।

आग से बचा पुस्तकालय

अभिलेखों के बचे रहने का तरीका कभी-कभी अप्रत्याशित होता है। असीरियाई राजा असुरबनपाल ने नीनवे में एक विशाल पुस्तकालय स्थापित किया था। उसने लगभग 669 ईसा पूर्व से शासन किया था। पुस्तकालय में पौराणिक कथाएँ, ज्योतिषि, चिकित्सा और प्रशासनिक दस्तावेज़ एकत्रित थे [4]।

लगभग 612 ईसा पूर्व में नीनवे जलकर खाक हो गया। यद्यपि कागज की कतिबे होतीं, तो उस समय सब कुछ नष्ट हो गया होता। लेकिन मिट्टी की पट्टिकाएँ आग में पककर बल्क और भी अधिक मजबूत हो गईं। ये पट्टिकाएँ 1840 के दशक से 1930 के दशक तक कई बार उत्खनन की गईं। आज ब्रिटिश संग्रहालय में उनके 30,000 से अधिक टुकड़े मौजूद हैं [4]।

फरि भी वे पूर्ण अवस्था में बाहर नहीं निकले। जब शहर का पतन हुआ, तो पट्टिकाओं को जानबूझकर तोड़कर महल में जगह-जगह फेंक दिया गया था। उत्खननकर्ताओं को व्यवस्थित अलमारियाँ नहीं, बल्कि बिखरे हुए टुकड़ों के ढेर मिले। एक ही रचना के आगे और पीछे के हिससे अलग-अलग कमरों से मिलना आम बात थी [4]। इन टुकड़ों को फरि से जोड़ना आज संगणकीय पुरातत्व का एक बड़ा कार्य है। संगणकीय पुरातत्व कंप्यूटर की मदद से प्राचीन पुरावशेषों का अध्ययन करने वाला क्षेत्र है।

पहले ही पुनर्जीवित हो चुके अक्षर

डजिटल रूप में बदलने के काम में भी बाधाएँ हैं। दुनिया के पुरावशेष मानवीय हाथों की तुलना में कहीं अधिक हैं। पहले देखा गया नीनवे के टुकड़ों का ढेर इसका एक उदाहरण है [4]। प्रत्येक टुकड़े की तस्वीर लेने और पढ़े गए पाठ को जोड़ने में बहुत लंबा समय लगता है। इसलिए मशीनों को मनुष्य की गति को बढ़ाना होगा।

लुप्त होते पुरावशेषों को बचाने के तरीके दो दशकों में बँटे हैं। पहला तरीका वस्तु को स्वयं लंबे समय तक सुरक्षित रखने वाला संरक्षण है। दूसरा तरीका वस्तु को छुए बिना उसके भीतर की सामग्री को पढ़कर कंप्यूटर में स्थानांतरित करना है। बाद वाले तरीके ने पहले ही कई बार सफलता हासिल की है।

इजराइल के एन-गेदी से मिला ग्रंथ-चर्मपत्र जलकर कोयले का ढेला बन चुका था। हाथ से खोलने की कोशिश करने पर उसके बखिर जाने की स्थिति थी। शोध दल ने वस्तु को काटे बिना अंदर देखने वाली CT स्कैनिंग के माध्यम से उसके भीतर झाँका। कंप्यूटर ने मुड़ी हुई चमड़े की परतों का गणितीय रूप से अनुसरण किया और उन्हें समतल में खोल दिया। उस पर धातु के अंश वाली स्याही उभर आई और लेख्यव्यवस्था का आरंभिक भाग पढ़ा गया [5]।

वसुवियस की ज्वालामुखी राख में दबे हर्कुलेनियम ग्रंथ-चर्मपत्र के लिए भी रास्ता खुल गया। 25 जून 2026 को शोधकर्ताओं ने घोषणा की कि उन्होंने एक ग्रंथ-चर्मपत्र को शुरू से अंत तक पढ़ लिया है। इस ग्रंथ-चर्मपत्र का नाम PHerc. 1667 है। इसके भीतर यूनानी भाषा के 22 कॉलम जतिनी सामग्री में नैतिकता पर चर्चा शामिल थी [6]। चुनौती के रूप में निर्धारित 13 मुहरबंद ग्रंथ-चर्मपत्रों पर नया पुरस्कार घोषित किया गया। इसकी अंतिम तिथि 25 जून 2027 है [7]।

आभासी पुस्तकालय केवल तस्वीरों का ढेर लगाने वाला गोदाम नहीं है। पुरावशेष का आकार, स्याही के घटक और पढ़े गए पाठ को एक साथ जोड़ा जाना चाहिए। तभी भविष्य में कोई नई विधि आने पर पुराने डेटा की फरि से गणना की जा सकेगी। मूल प्रतिलिपि हो जाने के बाद भी सत्यापन जारी रखने के लिए इस जुड़ाव की आवश्यकता है।

इमेजिंग और पढ़ने के काम में धन और मानव संसाधन लगते हैं। कसि पुरावशेष की तस्वीर पहले ली जाए, यह तय करना भी आसान नहीं है। खतरनाक क्षेत्रों में मौजूद पुरावशेषों की तो तस्वीर लेना ही कठिन होता है। इसलिए हर देश प्राथमिकताएँ तय करके काम करता है।

ये घटनाएँ एक सवाल छोड़ जाती हैं। क्या वस्तु के अंततः नष्ट हो जाने पर भी उसके भीतर के शब्दों को बचाया जा सकता है? आभासी पुस्तकालय ही वह जगह है जहाँ इसका उत्तर खोजा जाता है।

- [1] Mark, J. J. (2023, 7월 25일). Library of Alexandria. World History Encyclopedia. https://www.worldhistory.org/Library_of_Alexandria/ (3 सितंबर 2026 को अभिगमिति)
- [2] UNESCO. (2026). Damaged cultural sites in Ukraine verified by UNESCO. UNESCO. <https://www.unesco.org/en/ukraine-war/damaged-cultural-sites> (3 सितंबर 2026 को अभिगमिति)
- [3] Ma, Y. R. (2026). Towards Computational Chinese Paleography. arXiv:2601.06753. <https://arxiv.org/abs/2601.06753> (3 सितंबर 2026 को अभिगमिति)
- [4] The British Museum. What was Ashurbanipal's Library? The British Museum. <https://www.britishmuseum.org/research/projects/what-was-ashurbanipals-library> (3 सितंबर 2026 को अभिगमिति)
- [5] Seales, W. B., Parker, C. S., Segal, M., Tov, E., Shor, P., & Porath, Y. (2016). From damage to discovery via virtual unwrapping: Reading the scroll from En-Gedi. Science Advances, 2(9), e1601247. <https://doi.org/10.1126/sciadv.1601247>
- [6] Vesuvius Challenge. (2026, 6월 25일). An entire Herculaneum scroll has been read for the first time. scrollprize.org. <https://scrollprize.org/firstscroll> (3 सितंबर 2026 को अभिगमिति)
- [7] Vesuvius Challenge. (2026). Open Prizes. scrollprize.org. <https://scrollprize.org/prizes> (3 सितंबर 2026 को अभिगमिति)

12.2. वैज्ञानिक जाँच दल की खोज: ज्ञान ग्राफ़ और प्रकाश स्पेक्ट्रम विश्लेषण

पुरावशेषों को जोड़ने वाला एक विशाल मानचित्र

पुरानी प्रयोगशालाओं में पुरावशेष के आकार और पाठ को अलग-अलग देखा जाता था। संरक्षण वैज्ञानिक आकार का मापन करते थे और भाषाविद पाठ को पढ़ते थे। दोनों परमाणु अलग-अलग दराजों में चले जाते थे। इसलिए एक ही पुरावशेष पर दो बार शोध करने के बावजूद नष्टिर्षों को जोड़ा नहीं जा सका।

संगणकीय पुरातत्व इन दोनों को एक ही मानचित्र पर लाता है। उस मानचित्र का नाम ज्ञान ग्राफ़ है। ज्ञान ग्राफ़ बटुओं और रेखाओं के माध्यम से जानकारी को दर्शाने वाला एक चित्र है। बटु वस्तुएँ हैं और रेखाएँ वस्तुओं के बीच के संबंध हैं।

इस मानचित्र में मट्टी की एक पट्टिका एक बटु बन जाती है। उस पट्टिका को लिखने वाला मुंशी भी एक बटु बनता है और जसि शहर से वह मली, वह शहर भी एक बटु बन जाता है। स्याही के घटक और 3D आकार की जानकारी भी अलग-अलग बटु बन जाते हैं। पट्टिका पर लिखे व्यक्तियों के नाम और वस्तुओं के नाम भी बटुओं के रूप में निकाल लिए जाते हैं। कंप्यूटर इन बटुओं को जोड़ने वाली रेखाओं को खोजकर सीखता है।

एक बार मानचित्र बन जाने के बाद प्रश्नों का स्वरूप बदल जाता है। पहले केवल एक पट्टिका को रखकर पूछा जाता था कि यह कौन सा अक्षर है। अब यह पूछा जा सकता है कि इस स्याही का उपयोग करने वाला अन्य दस्तावेज़ कहाँ है। उसी मुंशी द्वारा लिखी गई पट्टिका किस शहर से मली थी, यह भी एक बार में खोजा जा सकता है।

फ़रवरी 2026 में पोलैंड के जगेलोनियन विश्वविद्यालय के एक शोध दल ने ऐसा ही एक मानचित्र सार्वजनिक किया। इस मानचित्र में व्यक्तियों, घटनाओं और वस्तुओं के बीच के संबंध लगभग 32 लाख पंक्तियों में दर्ज हैं। शोध दल ने मशीन को संबंधों की सफ़ाई करने और उसके कारण की व्याख्या करने में सक्षम बनाया। विशेषज्ञों ने उस व्याख्या को पढ़कर उसकी उपयोगिता की पुष्टि की [1]।

आकार मापने वाली परत

मानचित्र की पहली परत पुरावशेष का बाहरी आकार है। पुराने स्मारकों और मट्टी की पट्टिकाओं की सतह घिस जाने से अक्षरों के खांचे उथले हो जाते हैं। आँखों से यह पहचानना मुश्किल होता है कि खांचे बचे भी हैं या नहीं। इसलिए शोधकर्ता 3D स्कैनर से सतह को स्कैन करते हैं।

स्कैनर सतह को लाखों बटुओं में बदलकर सहेजता है। कंप्यूटर उन बटुओं को जोड़कर एक त्रिकोणीय जाल बनाता है। इस जाल पर विभिन्न दशाओं से आभासी प्रकाश डाला जा सकता है। प्रकाश की दशा बदलने पर उथले खांचों में भी छाया बन जाती है और अक्षर उभर आते हैं।

मशीन खांचों के अनुप्रस्थ काट के आकार को भी सीखती है। मनुष्य द्वारा औजार से उकेरे गए नशानों का अनुप्रस्थ काट एकसमान और दशा नशित होती है। प्राकृतिक रूप से पड़ी दरारों में ऐसा नहीं होता। इस अंतर को सीखने वाला मॉडल अक्षरों और खरोचों को अलग-अलग पहचान लेता है।

उसी तरीके से टूटे हुए हस्सों के अनुप्रस्थ काट को भी मापा जाता है। मशीन यह मलिन करती है कि दोनों टुकड़ों के कटे हुए हस्से आपस में जुड़ते हैं या नहीं। यह गणना मनुष्य द्वारा टुकड़ों को सीधे हाथ में लेकर मलाने के काम का स्थान ले लेती है। इससे किसी अन्य देश के संग्रहालय में रखे पहुँच से बाहर के टुकड़ों के साथ भी मलिन किया जा सकता है।

लपिटे हुए ग्रंथ-चर्मपत्रों के लिए एक अलग गणना का उपयोग किया जाता है। एन-गेदी ग्रंथ-चर्मपत्र के मामले में शोध दल ने कंप्यूटर के भीतर चमड़े की परतों को एक-एक करके अलग किया। और प्रत्येक परत को समतल में फैलाकर अक्षरों को स्पष्ट रूप से खड़ा कर दिया [2]।

प्रकाश से स्याही पढ़ने वाली परत

मानचित्र की दूसरी परत स्याही है। सामग्री के आधार पर स्याही प्रकाश को अलग तरीके से ग्रहण करती है। बहुवर्णिक इमेजिंग केवल दृश्य प्रकाश ही नहीं, बल्कि अवरक्त और पराबैंगनी किरणों को भी अलग-अलग करके तस्वीरें लेने की एक विधि है। कुछ तरंगदैर्घ्य पर पृष्ठभूमि गहरी हो जाती है और स्याही चमक उठती है।

इस अंतर के कारण मटाए गए अक्षर भी पुनर्जीवित हो जाते हैं। प्राचीन लोग महंगे चर्मपत्र को बचाने के लिए लिखे हुए को मटिकर दोबारा लिखते थे। मटाई गई सतह पर स्याही के कुछ अंश बचे रह जाते हैं। विभिन्न तरंगदैर्घ्य पर तस्वीरें लेकर उन्हें आपस में अध्यारोपित करने पर नीचे छपि वह नशान उभर आता है। इस तरह दोबारा लिखे गए दस्तावेज़ को पालमिपसेस्ट कहा जाता है।

मसिर के सनाई पर्वत पर स्थिति सेट कैथरीन मठ में ऐसे कई दस्तावेज़ हैं। अमेरिका और ग्रीस के शोधकर्ताओं ने यहाँ 74 खंडों की बहुवर्णिक इमेजिंग की। तस्वीरें लिए गए पृष्ठों की संख्या 6,800 तक पहुँचती है। उनके नीचे से 305 मटाए गए लेख फिर से सामने आए। वे लेख यूनानी और सीरियाई सहित 10 भाषाओं में लिखे गए थे [3]।

स्वयं घटकों को मापने की भी एक विधि है। रमन स्पेक्ट्रोस्कोपी स्याही पर प्रकाश डालती है और वापस लौटने वाले प्रकाश के पैटर्न को पढ़ती है। यह पैटर्न प्रत्येक पदार्थ के लिए अलग होता है और फिंगरप्रिंट की तरह काम करता है। एआई इस पैटर्न को देखकर पहचान लेता है कि यह कालखि से बनी स्याही है या लोहे के अंश वाली स्याही। स्याही का मशीन प्रत्येक कार्यशाला के अनुसार थोड़ा अलग होता है, जो उसके मूल स्थान का अनुमान लगाने का सुराग देता है।

हर्कुलेनियम ग्रंथ-चर्मपत्र के लिए और भी अधिक शक्तिशाली प्रकाश का उपयोग किया गया। शोधकर्ताओं ने अत्यधिक शक्तिशाली X-ray उत्पन्न करने वाले सन्क्रोट्रॉन कण त्वरक से ग्रंथ-चर्मपत्र के भीतर की तस्वीरें लीं। जले हुए पपीरस और कार्बन स्याही का घनत्व समान होता है, इसलिए उन्हें सामान्य रूप से अलग नहीं पहचाना जा सकता। डीप लर्निंग मॉडल ने बनावट के सूक्ष्म अंतर को सीखा और उस जगह का पता लगाया जहाँ अक्षर मौजूद थे [4]।

पाठ और पाठ को जोड़ने वाली परत

मानचित्र की तीसरी परत दस्तावेज़ों के बीच का संबंध है। प्राचीन पुस्तकें नकल करके प्रसारित की जाती थीं। प्रत्येक प्रतिलिपि के साथ एक छोटी सी त्रुटि जुड़ती जाती थी। विद्वान इन त्रुटियों का अनुसरण करके प्रतियों की वंशावली तैयार करते हैं।

कंप्यूटर इस कार्य के लिए जीव विज्ञान में प्रयुक्त गणना विधियों को उधार लेता है। यह जीन अनुक्रमों की तुलना करके जीवों का वंशावली वृक्ष तैयार करने के तरीके के समान है। दस्तावेज़ों के अंतर्गत की गणना करके यह क्रमबद्ध किया जाता है कि कौन सी प्रतिकृति प्रत से निकली है। इस प्रकार तैयार की गई वंशावली खो चुकी मूल प्रत के स्वरूप का अनुमान लगाने में मदद करती है।

हस्तलेख भी संबंधों को जोड़ने का एक सूत्र है। प्रत्येक मुंशी के बर्श दबाने का दबाव और स्ट्रोक का कोण अलग होता है। एआई इन आदतों को संख्याओं में बदलकर सहेजता है। समान संख्याओं वाले दस्तावेज़ों को एक साथ समूहीकृत करने पर मुंशियों का संबंध-जाल सामने आ जाता है।

पहले से सक्रिय अभिलेखागार

इन परतों को एकत्रित करके वास्तव में संचालित होने वाले कई अभिलेखागार हैं। Trismegistos प्राचीन विश्व के दस्तावेज़ों को क्रमांकित करने की एक परियोजना है। इसने 100 से अधिक भाषाओं की लगभग 9,60,000 सामग्रियों को एक प्रणाली में जोड़ा है। एक ही पुरावशेष को अलग-अलग देशों द्वारा अलग नाम से पुकारे जाने की समस्या इस संख्या से हल हो गई [5]।

Pelagios नेटवर्क बखिरी हुई सामग्रियों को आपस में जोड़ने वाला एक समुदाय है। वे स्थान, व्यक्ति और समय को जोड़ने का एक साझा तरीका तय करते हैं। प्राचीन स्थान के नाम युग के अनुसार बदलते रहने पर भी वे उन्हें एक निर्देशांक से जोड़ते हैं [6]। म्यूनिख में संचालित एक इलेक्ट्रॉनिक बेबीलोनियन पुस्तकालय है। यह दुनिया भर के कीलाकार लिपि के टुकड़ों को एक डेटाबेस में इकट्ठा करता है। विद्वान यहाँ टुकड़ों की खोज करके खोए हुए वाक्यों को पुनः प्राप्त करते हैं [7]।

साझा संख्या देखने में उबाऊ लग सकती है, लेकिन यह शक्तिशाली होती है। जब संख्या समान होती है, तभी अलग-अलग संस्थानों की सामग्री एक सूत्र में जुड़ती है [5]। तस्वीरें संग्रहालय में, पढ़ा गया पाठ विश्वविद्यालय में और उत्खनन का रिकॉर्ड अनुसंधान संस्थान में होना आम बात है। जब संख्या इन तीनों को जोड़ती है, तो बखिरी हुई जानकारी किसी एक पुरावशेष की कहानी के रूप में एकत्र हो जाती है।

सामग्री सार्वजनिक करने का तरीका भी बदल गया है। पहले तस्वीरें और पढ़ा गया पाठ शोध-पत्र प्रकाशित होने के बाद ही देखा जा सकता था। आजकल शोध दल स्कैन की गई सामग्री और प्रोग्राम एक साथ सार्वजनिक करते हैं। इससे बनी उपकरण वाले देशों के शोधकर्ता भी उसी सामग्री से दोबारा गणना कर सकते हैं [8]।

अब भी शेष चुनौतियाँ

वेसुवियस चैलेंज ने 10 जुलाई 2026 को शेष समस्याओं को संकलित करके सार्वजनिक किया। दबे और कुचले हुए हसिंसों में ग्रंथ-चर्मपत्र की परतें आपस में अलग नहीं होती हैं। कंप्यूटर परत का अनुसरण करते हुए गलती से बगल की परत पर चला जाता है। स्याही सतह से कतिनी गहराई तक समाई है, यह भी अभी स्पष्ट नहीं है। एक ग्रंथ-चर्मपत्र से सीखा मॉडल दूसरे ग्रंथ-चर्मपत्र पर अपना प्रभाव खो देता है। डेटा का आकार बहुत बड़ा होने के कारण सामान्य कंप्यूटर से इसे संभालना कठिन है। शोधकर्ताओं द्वारा निर्धारित अगला लक्ष्य समय जून 2027 है [8]।

प्राचीन लिपियों से जुड़े अन्य क्षेत्र भी इसी बाधा का सामना करते हैं। चीनी प्राचीन लिपियों के अनुसंधान को संकलित करने वाले वर्ष 2026 के एक शोध-पत्र ने सामग्री की कमी को एक बड़ी चुनौती बताया। इसके अलावा, उसने मानवीय नर्णय का स्थान लेने के बजाय उसकी सहायता करने की दशा की सफारिश की [9]।

संदर्भ सामग्री

- [1] Kutt, K., Sroka, E., Ishchuk, O., & do Valle Miranda, L. (2026). A three-stage neuro-symbolic recommendation pipeline for cultural heritage knowledge graphs. arXiv:2602.19711. <https://arxiv.org/abs/2602.19711> (3 सितंबर 2026 को अभिगमिति)
- [2] Seales, W. B., Parker, C. S., Segal, M., Tov, E., Shor, P., & Porath, Y. (2016). From damage to discovery via virtual unwrapping: Reading the scroll from En-Gedi. Science Advances, 2(9), e1601247. <https://doi.org/10.1126/sciadv.1601247>
- [3] Early Manuscripts Electronic Library. Sinai Palimpsests Project. EMEL. <https://www.emelibrary.org/sinai-palimpsests-project/> (3 सितंबर 2026 को अभिगमिति)
- [4] Vesuvius Challenge. (2026, 6 जून 25). An entire Herculaneum scroll has been read for the first time. scrollprize.org. <https://scrollprize.org/firstscroll> (3 सितंबर 2026 को अभिगमिति)
- [5] KU Leuven. Trismegistos portal. Trismegistos+. <https://tmplus.kuleuven.be/trismegistos-portal> (3 सितंबर 2026 को अभिगमिति)
- [6] Pelagios Network. Pelagios Network. <https://pelagios.org/> (3 सितंबर 2026 को अभिगमिति)
- [7] Ludwig-Maximilians-Universität München & Bayerische Akademie der Wissenschaften. electronic Babylonian Library. <https://www.ebl.badw.de/> (3 सितंबर 2026 को अभिगमिति)
- [8] Vesuvius Challenge. (2026, 7 जून 10). Open problems: Why reading every Herculaneum scroll is still a challenge. scrollprize.org. https://scrollprize.org/2026_open_problems (3 सितंबर 2026 को अभिगमिति)
- [9] Ma, Y. R. (2026). Towards Computational Chinese Paleography. arXiv:2601.06753. <https://arxiv.org/abs/2601.06753> (3 सितंबर 2026 को अभिगमिति)

12.3. उपसंहार

ग्यारह घटनाओं का पुनरावलोकन

इस पुस्तक ने ग्यारह घटनाओं का अनुसरण किया है। शुरुआत ज्वालामुखी की राख में दबे हर्कुलेनियम के ग्रंथ-चर्मपत्र से हुई थी। हाथ लगाते ही बखिर जाने वाले कोयले के टुकड़े को कंप्यूटर ने आभासी रूप से खोल दिया। जले हुए पपीरस और कार्बन की स्याही X-ray से देखने पर लगभग एक जैसे दिखते हैं। मशीन ने सतह पर बची सूक्ष्म बनावट के अंतर को सीखकर अक्षरों को खोज निकाला [1]।

एन-गेदी के लैव्यव्यवस्था ग्रंथ-चर्मपत्र ने भी आभासी अनावरण से अपना पहला पद पुनः प्राप्त किया। यहाँ धातु युक्त स्याही X-ray में स्पष्ट रूप से दिखाई दी [2]। यथायाह के ग्रंथ-चर्मपत्र में हस्तलेख ही मुख्य सुराग था। एआई ने स्ट्रोक की सूक्ष्म आदतों को मापकर दिखाया कलिपिकि केवल एक व्यक्ति नहीं था। मानवीय आँखों से यह हसिंसा केवल एक ही जैसी लिखावट दिखाई देता था [3]।

चौथी घटना से टुकड़ों को जोड़ने का सलिसला शुरू हुआ। बखिरी हुई मट्टी की पट्टियों और चर्मपत्र के टुकड़े अपने अनुप्रस्थ काट के आकार के आधार पर फिर से जुड़ गए। जाली पुरावशेषों की पहचान करने में भी इसी गणना का उपयोग किया गया। गलिंगमेश महाकाव्य की छूटी हुई पंक्तियों को टुकड़ों को जोड़कर पूरा किया गया।

मेसोपोटामिया की रसीदे मशीन अनुवाद के माध्यम से समझने योग्य बन गई। हजारों साल पुराने अनाज के बहीखाते और पत्र आज की भाषा में अनुवादित किए गए। कछुए के पेट के कवच पर अंकित ओरेकल बोन लपि 3D स्कैन और समरूपति मॉडल के माध्यम से आपस में जुड़ गई [4]। सलिला की काष्ठ पट्टिकाओं को भी अक्षर पहचान मॉडल ने पढ़ लिया। जापान के घसीट लपिदिस्तावेज़ भी इसी तरह हल किए गए [5]।

हमने ऐसे दरवाजे भी देखे जो अभी तक नहीं खुले हैं। Linear A, सधु लपि और Rongorongo के पास अब भी कोई उत्तर नहीं है। मशीन ने अक्षरों की व्यवस्था की नियमितता को मापकर यह परखा कि वह कोई भाषा है या नहीं। उसका अर्थ समझना अगली पीढ़ी के ज़िम्मे छोड़ दिया गया है।

अंतिम घटना प्राचीन शिलालेखों का स्वतः-पूर्ण होना थी। Ithaca ग्रीक शिलालेखों के छूटे हुए अक्षरों को भरने वाला एक मॉडल है। यह मॉडल अक्षरों को भरते हुए शिलालेख के निर्माण का स्थान और समय भी इंगित करता है [6]। इसके बाद आया Aeneas लैटिन शिलालेखों पर काम करता है। Aeneas लपियंतरित पाठ और शिलालेख की तस्वीरों को एक साथ देखता है [7]।

इन ग्यारह घटनाओं में एक समानता है। इनमें से कोई भी समस्या मशीन ने अकेले नहीं सुलझाई। भौतिकविदों ने इमेजिंग का काम संभाला और कंप्यूटर वैज्ञानिकों ने गणना का ज़िम्मा लिया। प्राचीन भाषाओं के विद्वानों ने परिणामों की पुष्टि की और अर्थ स्पष्ट किया। जब ये तीनों क्षेत्र एक ही पुरावशेष पर मिले, तभी एक पंक्ति जीवित हो सकी।

मशीनें जो अब भी नहीं कर सकती

यहाँ तक पहुँचने के दौरान मशीनों की कमज़ोरियाँ भी सामने आईं। पहली बाधा सामग्री की कमी है। Linear A के बचे हुए सभी प्रतीकों को मिलाकर भी केवल 7,400 के करीब हैं। इतनी कम मात्रा में मशीन के लिए नियमों को सीखना कठिन होता है [8]। आज के भाषा मॉडल अरबों शब्दों को पढ़कर विकसित होते हैं। वलुप्त प्राचीन भाषाओं की सामग्री उस पैमाने से बहुत पीछे रह जाती है।

जहाँ सामग्री कम होती है, वहाँ जबरन उत्तर निकालने पर मशीन असतुतिवहीन शब्द गढ़ लेती है। रचित स्थान जतिना लंबा होता है, उत्तर की विश्वसनीयता उतनी ही घटती है। संदर्भ में उचित लगने वाले लेकिन इतिहास में कभी न रहे वाक्य सामने आ सकते हैं। ऐसा वाक्य यदि एक बार शोध-पत्र में छप जाए, तो अगला शोध उसे सच मान लेता है।

प्रशिक्षण सामग्री एकतरफा हो, तो उत्तर भी उसी ओर झुक जाता है। यदि किसी एक क्षेत्र के शिलालेख अधिक हों, तो मशीन अन्य क्षेत्रों के शिलालेखों को भी उसी क्षेत्र की शैली में पढ़ लेती है। मशीन अक्सर खुद यह नहीं समझा पाती कि उसने ऐसा उत्तर क्यों दिया। इसलिए आजकल का शोध साक्ष्य भी साथ प्रस्तुत करने की दिशा में आगे बढ़ रहा है [9]।

ग्रंथ-चर्मपत्र पढ़ने वाले शोध दल ने भी शेष समस्याओं की एक सूची सार्वजनिक की। दबे और कुचले हुए हस्सों में परतें आपस में अलग नहीं होती हैं। एक ग्रंथ-चर्मपत्र से सीखा मॉडल दूसरे ग्रंथ-चर्मपत्र पर अपना प्रभाव खो भी देता है। क्या नहीं किया जा सकता यह स्पष्ट करना ही अगले कदम का मार्ग बनाता है [10]।

मशीन द्वारा निकाले गए अक्षरों को संभालने के नियम भी आवश्यक हैं। निश्चित रूप से पढ़े गए अक्षरों और अनुमानित अक्षरों के अंकन अलग होने चाहिए। किस मॉडल ने किस सामग्री के आधार पर ऐसा पढ़ा, यह भी साथ में दर्ज किया जाना चाहिए। तभी बाद में अन्य शोधकर्ता उसी गणना को दोबारा करके देख सकते हैं।

मूल प्रतीक ही रहती है

डजिटल पुनर्निर्माण चाहे कतिना भी सटीक क्यों न हो, वह मूल प्रतीक का स्थान नहीं ले सकता। शैक्षणिक जगत स्कैन की गई सामग्री को मूल प्रतीक नहीं, बल्कि एक डजिटल प्रतारूप मानता है। फोटोग्राफी में मानवीय चयन शामिल होता है। किस प्रकाश का उपयोग करना है और किस कोण से तस्वीर लेनी है, यह मनुष्य तय करता है। वह चयन बदलते ही परिणामी छवि भी बदल जाती है।

वास्तविक वस्तु में ऐसी जानकारी बची रहती है जिसे स्कैनर दर्ज नहीं कर सका। मट्टी की पट्टिका की सतह पर मट्टी के कणों की संरचना ऐसी ही है। कछुए के पेट के कवच पर सूक्ष्म छदों की व्यवस्था भी केवल वास्तविक वस्तु में देखी जा सकती है। स्याही पपीरस के भीतर कतिनी गहराई तक समाई है, यह भी केवल वास्तविक वस्तु में ही मापा जा सकता है [10]।

नई मापन तकनीकें भविष्य में भी आती रहेंगी। यदि हम आज पुरावशेषों को नष्ट कर दें, तो उन तकनीकों का उपयोग नहीं किया जा सकेगा। इसलिए पुरावशेषों को काटने या जबरन खोलने से बचना चाहिए। आभासी पुस्तकालय का उद्देश्य मूल प्रतीकों को त्यागना नहीं है। इसका

उद्देश्य मूल प्रतिको कम छूकर भी उसे पढ़ने योग्य बनाना है [2]। पुरावशेष को छूने की आवृत्तिकम होने से मूल प्रतिकी जीवन बढ़ जाता है।

साथ मलिकर पढ़ने का भवष्य

एआई इतिहासकारों को बाहर करने वाला उपकरण नहीं है। Ithaca के शोध ने इस बात को आंकड़ों से साबित किया। जब Ithaca ने अकेले छूटे हुए अक्षरों को पुनर्नरिम्ति किया, तो सटीकता 62 प्रतिशत थी। जब इतिहासकारों ने इसे अकेले किया, तो यह 25 प्रतिशत थी। जब दोनों पक्षों ने मलिकर काम किया, तो यह बढ़कर 72 प्रतिशत हो गई [6]।

यह आंकड़ा एक साथ दो बातें बताता है। इस प्रयोग में अकेली मशीन मनुष्य से बेहतर थी। और मनुष्य के साथ मलिकर काम करने पर परिणाम और अधिक सुधरे। किसी एक पक्ष के अकेले दम पर यहाँ तक पहुँचना संभव नहीं था।

मशीन तेजी से कई विकल्प सामने रखती है। मनुष्य यह जांचता है कि वह विकल्प उस युग के अनुसार संभव है या नहीं। मनुष्य और मशीन द्वारा बारी-बारी से पुष्टि करने की इस पद्धतिको शैक्षणिक जगत चक्र में मनुष्य की भागीदारी वाला सहयोग कहता है। इस चक्र को बनाए रखने के लिए प्राचीन पाठ को पढ़ सकने वाले लोगों की नरितर आवश्यकता है। यदि पढ़ने वाला कोई न हो, तो मशीन द्वारा निकाले गए अक्षर भी केवल एक नई पहेली बन जाते हैं।

सखाना भी शोध जतिना ही महत्वपूर्ण है। प्राचीन भाषाओं को पढ़ने की वधिएक व्यक्तसे दूसरे व्यक्तितक पहुँचती रही है। यदि वह कड़ी टूट जाए, तो मशीन द्वारा निकाले गए अक्षरों की पुष्टि करने वाला कोई नहीं रहेगा। नई पीढ़ी को प्राचीन लिपियों और कंप्यूटर दोनों को एक साथ सीखना चाहिए।

यह चक्र केवल विशेषज्ञों तक सीमित नहीं है। जापान में घसीट लिपि पढ़ने वाला एक नःशिल्क ऐप Miwo उपलब्ध है। तस्वीर खींचते ही स्क्रीन पर आज के अक्षर उभर आते हैं। प्राचीन दस्तावेज़ पहली बार देखने वाला व्यक्ति भी पहली पंक्ति पढ़ सकता है [5]। संग्रहालय के कांच के सामने स्मार्टफोन उठाना भी एक नए शोध की शुरुआत बन सकता है।

आगे करने के लिए अभी बहुत काम बाकी है। हरकुलेनियम में अब भी सीलबंद ग्रंथ-चर्मपत्र बचे हुए हैं [11]। यूक्रेन में आज भी पुस्तकालय और अभिलेखागार नष्ट हो रहे हैं [12]। दुनिया भर में प्राचीन लेखों का लुप्त होना अब भी अधिक तेज़ है।

इस पुस्तक को पढ़ने वाले पाठक भी कुछ योगदान कर सकते हैं। सार्वजनिक रूप से उपलब्ध स्कैन डेटा को कोई भी डाउनलोड करके देख सकता है [10]। टुकड़ों को जोड़ने की प्रतियोगिता में छात्र भी भाग लेते हैं [11]। कई जगहों पर प्राचीन दस्तावेज़ों को लिप्यंतरित करने में आम नागरिकों साथ जुड़ रहे हैं। यदि एक व्यक्ति केवल एक पंक्ति का जम्मा भी ले, तो सामग्री एकत्र हो जाती है।

अभिलेखों की रक्षा करना केवल अतीत की बात नहीं है। आज हम जो सामग्री बना रहे हैं, वह भी किसी दिन प्राचीन सामग्री बन जाएगी। फाइल प्रारूप बदलने पर आज के दस्तावेज़ भी नहीं पढ़े जा सकेंगे। इसलिए पुस्तकालय सामग्री को नए प्रारूप में स्थानांतरित करने का काम जारी रखते हैं। आभासी पुस्तकालय स्थापित करने के लिए हमेशा देखभाल करने वाले हाथों की आवश्यकता होती है।

फरि भी दशा तय हो चुकी है। 2,000 वर्ष पूर्व ज्वालामुखी की राख में दबा ग्रंथ-चर्मपत्र फरि से पढ़ लिया गया [1]। एक ग्रंथ-चर्मपत्र से वकिसति वधिएगले ग्रंथ-चर्मपत्र के अनुसंधान का आरंभ बढि बनती है [11]। डेटा और कोड सार्वजनिक रखने से दूसरे देशों के शोधकर्ता भी उसी मार्ग पर आगे बढ़ते हैं [10]। पुरावशेष भले ही किसी दिन मट्टि में मलि जाएँ, लेकिन उनके भीतर लिखे वाक्य शेष रहते हैं। आग में कभी न जलने वाला पुस्तकालय इसी तरह एक-एक पंक्ति करके बनाया जा रहा है।

संदर्भ सामग्री

- [1] Vesuvius Challenge. (2026, 6월 25일). An entire Herculanum scroll has been read for the first time. scrollprize.org. <https://scrollprize.org/firstscroll> (3 सितंबर 2026 को अभिगमति)
- [2] Seales, W. B., Parker, C. S., Segal, M., Tov, E., Shor, P., & Porath, Y. (2016). From damage to discovery via virtual unwrapping: Reading the scroll from En-Gedi. Science Advances, 2(9), e1601247. <https://doi.org/10.1126/sciadv.1601247>
- [3] Popović, M., Dhali, M. A., & Schomaker, L. (2021). Artificial intelligence based writer identification generates new evidence for the unknown scribes of the Dead Sea Scrolls exemplified by the Great Isaiah Scroll (1QIsaa). PLOS ONE, 16(4), e0249769. <https://doi.org/10.1371/journal.pone.0249769>
- [4] Ma, Y. R. (2026). Towards Computational Chinese Paleography. arXiv:2601.06753. <https://arxiv.org/abs/2601.06753> (3 सितंबर 2026 को अभिगमति)
- [5] ROIS-DS Center for Open Data in the Humanities. मि (miwo): AI < ずし字認識アプリ. CODH. <https://codh.rois.ac.jp/miwo/> (3 सितंबर 2026 को अभिगमति)

- [6] Assael, Y., Sommerschield, T., Shillingford, B., Bordbar, M., Pavlopoulos, J., Chatzipanagiotou, M., Androutsopoulos, I., Prag, J., & de Freitas, N. (2022). Restoring and attributing ancient texts using deep neural networks. *Nature*, 603(7900), 280-283. <https://doi.org/10.1038/s41586-022-04448-z>
- [7] Assael, Y., Sommerschield, T., Cooley, A. E., Shillingford, B., Pavlopoulos, J., Suresh, P., Herms, B., Grayston, J., Maynard, B., Dietrich, N., Wulgaert, R., Prag, J., Mullen, A., & Mohamed, S. (2025). Contextualizing ancient texts with generative neural networks. *Nature*. <https://doi.org/10.1038/s41586-025-09292-5>
- [8] Braović, M., Krstinić, D., Štula, M., & Ivanda, A. (2024). A systematic review of computational approaches to deciphering Bronze Age Aegean and Cypriot scripts. *Computational Linguistics*, 50(2), 725-779. https://doi.org/10.1162/coli_a_00514
- [9] Kutt, K., Sroka, E., Ishchuk, O., & do Valle Miranda, L. (2026). A three-stage neuro-symbolic recommendation pipeline for cultural heritage knowledge graphs. *arXiv:2602.19711*. <https://arxiv.org/abs/2602.19711> (3 सितंबर 2026 को अभगिति)
- [10] Vesuvius Challenge. (2026, 7월 10일). Open problems: Why reading every Herculaneum scroll is still a challenge. [scrollprize.org](https://scrollprize.org/2026_open_problems). https://scrollprize.org/2026_open_problems (3 सितंबर 2026 को अभगिति)
- [11] Vesuvius Challenge. (2026). Open Prizes. [scrollprize.org](https://scrollprize.org/prizes). <https://scrollprize.org/prizes> (3 सितंबर 2026 को अभगिति)
- [12] UNESCO. (2026). Damaged cultural sites in Ukraine verified by UNESCO. UNESCO. <https://www.unesco.org/en/ukraine-war/damaged-cultural-sites> (3 सितंबर 2026 को अभगिति)

परशिष्ट. मेरी अपनी AI पुरातत्व यात्रा: आम जनता के लिए वैश्विक अन्वेषण पोर्टल

A.1. Vesuvius Challenge का आधिकारिक डेटा पुरालेख

वहाँ क्या देखा जा सकता है

Vesuvius Challenge एक अंतरराष्ट्रीय प्रतियोगिता है। इसमें कंप्यूटर की सहायता से जले हुए चर्मपत्रों को पढ़ने की प्रतस्पर्धा होती है। ये चर्मपत्र हर्कुलेनियम से मिले थे। यह वसुवियस ज्वालामुखी के पास स्थित एक प्राचीन रोमन शहर था। ये चर्मपत्र पपीरस से बने हैं। पपीरस एक प्राचीन कागज है, जिसे पौधों को दबाकर बनाया जाता था। इसका आधिकारिक पता scrollprize.org है[1]। यह वेबसाइट प्रतियोगिता में उपयोग की जाने वाली सामग्री को सभी के लिए सार्वजनिक करती है[2]। इसके लिए किसी सदस्यता की आवश्यकता नहीं है और कोई शुल्क भी नहीं लगता है[2][5]।

मुख्य पृष्ठ के शीर्ष पर कई मेनू विकल्प दिए गए हैं। सबसे पहले Prizes, Problems, Tutorials & In-depth और Data आते हैं। इसके बाद Milestones, Community, Press और Donate आते हैं। सितंबर 2026 तक 45 चर्मपत्रों और पपीरस के टुकड़ों को स्कैन किया जा चुका है। उनमें से एक को शुरू से अंत तक पढ़ लिया गया है[1]।

25 जून 2026 को प्रतियोगिता दल ने उस एक चर्मपत्र के बारे में समाचार दिया। यह PHerc. 1667 नामक चर्मपत्र है[3]। यह चर्मपत्र को बना खोले अंत तक पढ़ने का पहला उदाहरण है[3][4]। खुले हुए पपीरस की लंबाई लगभग 1.4 मीटर है। प्राचीन पुस्तकों में लेख को लंबवत स्तंभों में बांटकर लिखा जाता था। ऐसे एक स्तंभ को कॉलम कहा जाता है। इस चर्मपत्र में ग्रीक लिपि में लिखे लगभग बाईस कॉलम सामने आए। इसकी विषयवस्तु स्टोइक दर्शन की नैतिकता पर आधारित है। इसमें मानव स्वभाव, आवेग और नैतिक विकास की चर्चा की गई है। इसे लगभग दूसरी शताब्दी ईसा पूर्व का लेख माना जाता है[3]।

पूरा चर्मपत्र पूरी तरह सुरक्षित नहीं बचा है। अब केवल चर्मपत्र का भीतरी भाग ही बचा है। 19वीं सदी और 1980 के दशक में हाथ से खोलते समय इसकी बाहरी परतें टूट गई थीं। जहाँ से सतह उखड़ गई, वहाँ के अक्षर भी नष्ट हो गए। इसलिए पढ़े गए पाठ में जगह-जगह रक्ति स्थान हैं[3]।

स्कैनिंग फ्रांस के ग्रेनोबल में की गई थी। यह यूरोपीय सिक्रोट्रॉन विकिरण सुविधा ESRF है। यह अत्यंत तीव्र X-ray उत्पन्न करने वाला एक विशाल अनुसंधान उपकरण है। इसमें BM18 बीमलाइन का उपयोग किया गया था[4]। बीमलाइन एक प्रयोगात्मक मार्ग है, जो X-ray को खींचकर वस्तु पर केंद्रित करता है।

घर पर अभ्यास करने के चरण

सबसे पहले scrollprize.org पर जाएं और Get Started बटन दबाएं। यह स्क्रीन पहली बार आने वाले उपयोगकर्ताओं को तीन चरणों का मार्गदर्शन देती है[5]।

पहला चरण स्याही का पता लगाना है। स्क्रीन पर दिखाई देने वाली स्याही को चिह्नित किया जाता है। इसके बाद मॉडल उस निशान से सीखकर शेष भाग का अनुमान लगाता है[5]।

दूसरा चरण आभासी अनावरण है। यह कंप्यूटर से चर्मपत्र को खोलने का कार्य है। CT क्रॉस-सेक्शन छवि में पपीरस की एक परत का अनुसरण किया जाता है। CT क्रॉस-सेक्शन छवि अंदरूनी हस्ति को पतले टुकड़ों में काटकर ली गई तस्वीर जैसी होती है। उस परत को समतल खोलकर परिणाम की जांच की जाती है[5]।

तीसरा चरण पूरी प्रक्रिया को सीखना है। इसमें स्कैनिंग, आभासी अनावरण और पढ़ने की प्रक्रिया को छोटे पैमाने पर क्रमवार आजमाया जाता है[5]।

यह ध्यान रखना चाहिए कि अभ्यास का क्रम और वास्तविक पुनर्प्राप्ति का क्रम अलग होता है। अभ्यास आसान स्याही पहचान से शुरू होता है। वास्तविक पुनर्प्राप्ति में स्कैनिंग सबसे पहले होती है। इसके बाद चर्मपत्र का आभासी अनावरण किया जाता है। स्याही को पढ़ने का काम सबसे अंत में होता है[5]।

डेटा को सीधे डाउनलोड करने के लिए scrollprize.org/data पर जाएं। डेटा सार्वजनिक रपिजिटरी [s3://vesuvius-challenge-open-data/](https://zenodo.org/record/1444444/files) पर उपलब्ध है। इसे वेब ब्राउज़र में भी खोला जा सकता है। इसके उपयोग को दर्शाने वाली उदाहरण फाइलें GitHub पर उपलब्ध हैं[2]।

प्रतभागियों के लिए पुरस्कार राशि भी रखी गई है। सितंबर 2026 तक शेष पुरस्कार राशि 21.4 लाख डॉलर है। 2027 Grand Prize की राशि 10 लाख डॉलर है। प्रथम विजिता को 8 लाख डॉलर मिलेंगे। इसकी अंतिम तिथि 25 जून 2027 है। जिन चरमपत्रों में अभी तक अक्षर नहीं पढ़े जा सके हैं, उन पर भी इनाम रखा गया है। 4 वर्ग सेंटीमीटर के दायरे में दस अक्षर पढ़ने होते हैं। इसके लिए प्रत्येक चरमपत्र पर 50,000 डॉलर दिए जाते हैं। हर महीने दिया जाने वाला Progress Prize भी उपलब्ध है। उस महीने की सर्वश्रेष्ठ प्रस्तुत को 20,000 डॉलर मिलते हैं[6]।

ध्यान देने योग्य बातें

डेटा CC BY-NC 4.0 शर्तों के तहत जारी किया गया है। इसका उपयोग व्यावसायिक लाभ के लिए नहीं किया जा सकता है। इसका उपयोग करते समय स्रोत का उल्लेख करना आवश्यक है। पुरानी सामग्री के लिए उद्धृत करने हेतु अलग शोधपत्र निर्धारित है। चरमपत्र संख्या 1 से 4 और टुकड़े संख्या 1 से 6 ऐसे ही हैं[2]। इस स्थिति में Parsons आदिके 2023 के शोधपत्र का उल्लेख किया जाता है[2][7]।

फाइल का आकार बहुत बड़ा होता है, यह बात पहले से जान लेनी चाहिए। 3D स्कैन डेटा कुल मिलाकर कई टेराबाइट का है[2]। डीप लर्निंग चलाने के लिए उच्च क्षमता वाले ग्राफिक्स कार्ड की आवश्यकता होती है। यदघर का कंप्यूटर कम क्षमता वाला है, तो डेटा को केवल देखकर शुरुआत की जा सकती है।

पुरस्कार के लिए प्रयास करते समय कुछ शर्तें लागू होती हैं। प्रतियोगिता के Discord चैनल में शामिल होना आवश्यक है। कोड को MIT जैसी खुली शर्तों के तहत सार्वजनिक करना होगा। पुरस्कार की घोषणा होने तक परिणामों को बाहर साझा नहीं किया जा सकता है[6]।

संदर्भ सामग्री

- [1] Vesuvius Challenge. (2026). Vesuvius Challenge: Reading the Herculaneum Scrolls with machine learning. scrollprize.org. <https://scrollprize.org/> (3 सितंबर 2026 को अभिगमिति)
- [2] Vesuvius Challenge. (2026). Data. scrollprize.org. <https://scrollprize.org/data> (3 सितंबर 2026 को अभिगमिति)
- [3] Vesuvius Challenge. (2026, 6월 25일). An entire Herculaneum scroll has been read for the first time. scrollprize.org. <https://scrollprize.org/firstscroll> (3 सितंबर 2026 को अभिगमिति)
- [4] Angelotti, G., Parsons, S., Nicolardi, F., et al. (2026). Complete virtual unwrapping and reading of a rolled Herculaneum papyrus. arXiv:2606.29085. <https://arxiv.org/abs/2606.29085> (3 सितंबर 2026 को अभिगमिति)
- [5] Vesuvius Challenge. (2026). Get Started. scrollprize.org. https://scrollprize.org/get_started (3 सितंबर 2026 को अभिगमिति)
- [6] Vesuvius Challenge. (2026). Open Prizes. scrollprize.org. <https://scrollprize.org/prizes> (3 सितंबर 2026 को अभिगमिति)
- [7] Parsons, S., Parker, C. S., Chapman, C., Hayashida, M., & Seales, W. B. (2023). EduceLab-Scrolls: Verifiable recovery of text from Herculaneum papyri using X-ray CT. arXiv:2304.02084. <https://doi.org/10.48550/arXiv.2304.02084>

A.2. eBL इलेक्ट्रॉनिक बेबीलोनियन साहित्य पोर्टल

पोर्टल की संरचना

eBL, electronic Babylonian Library का संक्षिप्त रूप है। इसका अर्थ इलेक्ट्रॉनिक बेबीलोनियन पुस्तकालय है। प्राचीन मेसोपोटामिया के लोग कीलाकार लपि में लिखते थे। कीलाकार लपि गीली मट्टी की पट्टियों पर दबाकर उकेरे गए कील के आकार के अक्षर हैं। यह पोर्टल मट्टी की पट्टियों के उन टुकड़ों को एक स्थान पर एकत्र करता है। इसके संग्रह का विषय बेबीलोनियन साहित्य वाले टुकड़े हैं। यह दुनिया की सभी कीलाकार लपि सामग्रियों को रखने वाला स्थान नहीं है[1]।

यह पोर्टल जर्मनी के म्यूनिख स्थित लुडविग मैक्सिमिलियन विश्वविद्यालय में स्थित है। इसका नेतृत्व प्रोफेसर एनरिक जमिनेज करते हैं। यह परियोजना अप्रैल 2018 में शुरू हुई थी। यह अलेक्जेंडर वॉन हम्बोल्ट फाउंडेशन के सोफिया कोवालेव्स्काया पुरस्कार अनुदान से संचालित होती है[2]।

इसका पता www.ebl.lmu.de है। इसमें बिना लॉगिन किए जाया जा सकता है और कोई शुल्क भी नहीं है। स्क्रीन के बाईं ओर आठ खंड हैं। सबसे पहले eBL Project, Library, Corpus और Signs आते हैं। इसके बाद Akkadian Dictionary और Bibliography आते हैं। अंत में News और Archaeology हैं[1]।

Library अलग-अलग टुकड़ों को एकत्र करने का स्थान है। पहले इसे Fragmentarium कहा जाता था। eBL के शोधकर्ताओं ने यहाँ टुकड़ों को जोड़ने के लगभग 1,200 मामले खोज निकाले हैं[1]। Corpus कृतियों के अनुसार मूल पाठ दिखाता है। गल्लामेश महाकाव्य लगभग 3,150 पंक्तियों में संकलित है। सृष्टि भित्ति एनूमा एलशि 1,096 पंक्तियों का है[3]। 2024 में प्रकाशित एक शोधपत्र इसके

वसितार को बताता है। यह पोर्टल लगभग 25,000 मट्टी की पट्टियों को समाहित करता है। इसका लिप्यंतरित पाठ 3,50,000 पंक्तियों से अधिक है[4]। लिप्यंतरित पाठ कीलाकार लिपि को लैटिन अक्षरों में लिखा गया रूप होता है।

स्क्रीन पर क्या किया जा सकता है

Library पर क्लिक करने पर एक खोज बॉक्स दिखाई देता है। नाम, वर्ष और शीर्षक से खोज की जा सकती है। इसमें शब्द-आधारित खोज और उन्नत खोज भी उपलब्ध है। यद्यपि हम समझ न आए कि क्या खोजना है, तो I'm feeling lucky बटन दबाएं। इससे कोई भी एक टुकड़ा खुल जाता है[5]।

किसी टुकड़े को खोलने पर स्क्रीन कई भागों में विभाजित हो जाती है। शीर्ष पर धारक संस्थान, आकार और प्राप्त स्थल दर्ज होता है। इसके साथ ही लिपि शैली और दस्तावेज़ का प्रकार भी दिखाता है। उदाहरण के लिए, K.5743 ब्रिटिश संग्रहालय का कुयुंजिक संग्रह है। कुयुंजिक वह टीला है जहाँ प्राचीन शहर नीनवे स्थित था। यह उस टीले से मिली पुरातात्विक वस्तुओं का सामूहिक नाम है। इसका आकार 3.81 सेटीमीटर गुणा 2.54 सेटीमीटर है। इसकी लिपि नव-असीरियन शैली की है[6]। यह नव-असीरियन साम्राज्य के दौर में प्रयुक्त लिपि का रूप है।

मध्य भाग में पंक्तिसंख्या के साथ लिप्यंतरित पाठ आता है। नचिले हिस्से में मट्टी की पट्टिका की तस्वीर होती है। यह भी बताया जाता है कि यह किस टुकड़े के साथ जुड़ा हुआ है[6]।

वेबसाइट पाठ को उद्धृत करने का प्रारूप भी बताती है। इसमें टुकड़े की संख्या और eBL edition लिखा जाता है। इसके बाद वेब पता और पहुँच की तिथि जोड़ी जाती है[1]।

नवीनतम समाचार और ध्यान देने योग्य बातें

News अनुभाग में समाचार पत्र देखा जा सकता है। 28 मई 2026 का 22वां अंक सबसे नवीनतम है। इस अंक में पोर्टल ने नया स्क्रीन लेआउट प्रस्तुत किया। संग्रह रखने वाले संस्थानों की संख्या भी बढ़ी है। इंचियोन स्थित राष्ट्रीय विश्व लिपि संग्रहालय भी उनमें से एक है। ब्रिटिश संग्रहालय की 3,100 से अधिक नई तस्वीरें अपलोड की गईं। येल पीबॉडी संग्रहालय की भी 1,000 से अधिक तस्वीरें जोड़ी गईं। शब्दकोश में सीधे नए शब्द दर्ज करने की सुविधा भी जोड़ी गई है[7]।

उपयोग करने से पहले कुछ बातें जानना जरूरी है। Library का लिप्यंतरित पाठ अंतिम नहीं, बल्कि एक अस्थायी मसौदा है। शोध दल आज भी इसमें लगातार सुधार कर रहा है। Corpus में व्यवस्थित कृतियों का मूल पाठ भी ऐसा ही है। यह अब तक मिले टुकड़ों को जोड़कर निकाला गया एक अंतरिम परिणाम है। नए टुकड़े जुड़ने पर पंक्तियों की संख्या और सामग्री बदल जाती है। इसलिए विद्यालयी कार्य या आलेखों में उपयोग करते समय पहुँच की तिथि अवश्य दर्ज करें[1]।

लिप्यंतरित पाठ CC BY-NC-SA 4.0 शर्तों के तहत जारी किया गया है। तस्वीरों के लिए शर्तें अलग हैं। तस्वीरों का पुनः उपयोग करने के लिए अलग से अनुमति लेनी होगी[1]। यह अनुमत उन संग्रहालयों से मिलती है जहाँ मट्टी की पट्टियाँ रखी हैं[1][4]। स्क्रीन के नीचे तस्वीर के विवरण में लिखे निर्देशों को पहले पढ़ना चाहिए[1]।

संदर्भ सामग्री

- [1] electronic Babylonian Library. (2026). About: Library. ebl.lmu.de. <https://www.ebl.lmu.de/about/library> (3 सितंबर 2026 को अभिगमित)
- [2] Jiménez, E. (2023, 5 मई 25 मई). AI is helping us read ancient Mesopotamian literature. The Conversation. <https://theconversation.com/ai-is-helping-us-read-ancient-mesopotamian-literature-205091> (3 सितंबर 2026 को अभिगमित)
- [3] electronic Babylonian Library. (2026). Corpus. ebl.lmu.de. <https://www.ebl.lmu.de/corpus> (3 सितंबर 2026 को अभिगमित)
- [4] Cobanoglu, Y., Laasonen, J., Simonjetz, F., Khait, I., Cohen, S., Földi, Z., Härtinen, A., Heinrich, A., Mitto, T., Rozzi, G., Sáenz, L., & Jiménez, E. (2024). Transliterated cuneiform tablets of the electronic Babylonian Library platform. Journal of Open Humanities Data, 10(1), 19. <https://doi.org/10.5334/johd.148>
- [5] electronic Babylonian Library. (2026). Library. ebl.lmu.de. <https://www.ebl.lmu.de/library> (3 सितंबर 2026 को अभिगमित)
- [6] electronic Babylonian Library. (2026). K.5743, eBL edition. ebl.lmu.de. <https://www.ebl.lmu.de/library/K.5743> (3 सितंबर 2026 को अभिगमित)
- [7] electronic Babylonian Library. (2026, 5 मई 28 मई). eBL Newsletter 22. ebl.lmu.de. <https://www.ebl.lmu.de/news> (3 सितंबर 2026 को अभिगमित)

A.3. Miwo उपयोग मार्गदर्शिका

यह कैसा ऐप है

जापान की प्राचीन पुस्तकों और दस्तावेजों में घसीट लिपिका व्यापक उपयोग हुआ था। इस लिपिको कुजुशजि कहा जाता है। यह एदो काल के मुद्रति ग्रंथों और प्राचीन दस्तावेजों में अक्सर मिलती है[3]। आज के अधिकांश जापानी लोग इस लिपिको नहीं पढ़ पाते हैं।

Miwo इस घसीट लिपिकी तस्वीर लेकर उसे आधुनिक लिपि में बदलने वाला एक स्मार्टफोन ऐप है। इसे जापान के सेंटर फॉर ओपन डेटा इन द ह्यूमैनिटीज CODH ने विकसित किया है[1]। यह नःशुल्क है[2]। ऐप का नाम गेंजी की कहानी (The Tale of Genji) के चौदहवें अध्याय मयित्सुकुशी से लिया गया है। मयित्सुकुशी जलमार्ग बताने वाला एक संकेत चिह्न है। इसका उद्देश्य प्राचीन दस्तावेजों के सागर में मार्गदर्शक बनना है[3]।

यह ऐप पहली बार 30 अगस्त 2021 को जारी किया गया था। 26 अक्टूबर 2022 को इसे संस्करण 1.1 में उन्नत किया गया। उस समय Ruri नामक नया पहचान मॉडल इसमें शामिल किया गया था। पृष्ठभूमि में कोई डिज़ाइन या रंग होने पर भी Ruri अक्षरों को अच्छी तरह पहचान लेता है[1]। इसके प्रशिक्षण में जापानी शास्त्रीय कुजुशजि डेटासेट का उपयोग किया गया था। यह मशीन लर्निंग के लिए तैयार की गई प्राचीन अक्षरों की छवियों का एक संग्रह है। यह 10 लाख से अधिक अक्षरों की सामग्री है। इस ऐप को 2022 में Good Design Award प्राप्त हुआ था[3]। सितंबर 2026 के अनुसार इस ऐप द्वारा जांची गई तस्वीरों की संख्या 35 लाख से अधिक हो चुकी है[1]।

इंस्टॉल करने और उपयोग करने के चरण

सबसे पहले ऐप स्टोर खोलें। iPhone और iPad के लिए App Store का उपयोग होता है। Android उपकरणों के लिए Google Play का उपयोग होता है। खोज बॉक्स में *みづ* या *miwo* दर्ज करें[1]। App Store के अनुसार ऐप का आकार लगभग 52 मेगाबाइट है। iPhone और iPad के लिए iOS 11.0 या उससे उच्च संस्करण पर्याप्त है[2]।

ऐप शुरू करने पर कैमरा स्क्रीन सामने आती है। प्राचीन दस्तावेज़ को स्क्रीन में रखकर कैप्चर बटन दबाएं। पहले से खींची गई तस्वीर को भी लोड किया जा सकता है। इसके बाद पहचान बटन दबाने पर एआई अक्षरों को पढ़ लेता है[1]।

परिणाम स्क्रीन पर कई सुविधाएं हैं। स्क्रीन पर एक नयित्तरण स्लाइडर दिया गया है। इस स्लाइडर को दाएं-बाएं खसिकाकर दोनों दृश्यों को एक साथ देखा जा सकता है। मूल तस्वीर और पढ़े गए पाठ की सीधे तुलना की जा सकती है। जापानी काना अक्षर पर क्लिक करने का विकल्प भी है। यह बताता है कि वह अक्षर मूलतः किस चीनी अक्षर से निकला है। पढ़े गए पाठ को कॉपी बटन से दूसरे ऐप में ले जाया जा सकता है[1]। परिणाम को सुरक्षित रखकर बाद में फिर से देखा भी जा सकता है[1][3]।

यह याद रखना आवश्यक है कि इंटरनेट चालू होना चाहिए। पहचान के लिए तस्वीरों को CODH सर्वर पर भेजा जाता है। भेजी गई तस्वीरों को सर्वर पर सहेजा नहीं जाता है[1]।

पढ़ी जा सकने वाली और न पढ़ी जा सकने वाली सामग्रियां

ऐप का उपयोग करने से पहले इसकी सीमाओं को समझ लेने से नरिशा नहीं होती। एआई को मुख्य रूप से एदो काल के मुद्रति ग्रंथों पर प्रशिक्षित किया गया है। उस काल की अधिकांश पुस्तकें काष्ठ-ब्लॉक मुद्रति थीं। ये लकड़ी के तख्तों पर अक्षर उकेरकर कागज़ पर छापी गई पुस्तकें होती हैं। इसलिए यह उस काल के मुद्रति ग्रंथों को आसानी से पढ़ लेता है। अन्य कालों की सामग्रियों या हाथ से लिखे दस्तावेजों में इसकी सटीकता घट जाती है[3]।

कागज़ की स्थिति भी परिणामों को बदल देती है। दाग लगे या कीड़ों द्वारा खाए गए हिस्सों को पढ़ना कठिन होता है। कागज़ की बनावट, रोशनी और परछाइयां भी बाधा बनती हैं[3]। चमकदार और एकसमान रोशनी में तस्वीर खींचना बेहतर होता है।

ऐसी सामग्रियां भी हैं जिनमें यह नहीं पढ़ पाता। पत्थर के शिलालेखों या साइनबोर्डों पर उकेरी गई घसीट लिपिका यह वर्तमान में समर्थन नहीं करता है[3]।

विकास दल स्पष्ट करता है कि पहचान के परिणामों में गलतियां हो सकती हैं। वे रेखांकित करते हैं कि एआई द्वारा घसीट लिपिकी पहचान कभी पूरी तरह सटीक नहीं हो सकती। इसलिए महत्वपूर्ण दस्तावेजों की मनुष्य द्वारा पुनः जांच की जानी चाहिए[3]।

प्राचीन अक्षरों को आधुनिक अक्षरों में बदलने की क्रिया को लिप्यंतरण कहा जाता है। इस कार्य को मलिकर पूरा करने का एक अलग मंच भी है। यह नागरिक भागीदारी पर आधारित वेबसाइट 'मनिना दे होनकोकू' है। इसका अर्थ 'सभी का लिप्यंतरण' है। इसका संचालन Miwo बनाने वाली संस्था से भिन्न संगठन करता है। इसका स्वरूप वैसा ही है जैसे ऐप द्वारा पढ़े गए पाठ को मनुष्य द्वारा परीक्षित किया जाए। इसका पता honkoku.org है। अब तक 10,800 से अधिक लोग इसमें भाग ले चुके हैं। लिप्यंतरित अक्षरों की संख्या 4.8 करोड़ से अधिक हो चुकी है। अगस्त 2025 में इसका दूसरा नवीनीकरण किया गया। नया संस्करण परीक्षण के तौर पर चल रहा है। जापान का राष्ट्रीय

इतिहास एवं लोककथा संग्रहालय इसमें भागीदार है। टोक्यो विश्वविद्यालय का भूकंप अनुसंधान संस्थान और क्योटो विश्वविद्यालय के शोध दल भी इसे संयुक्त रूप से चलाते हैं[4]।

संदर्भ सामग्री

- [1] ROIS-DS 人文学オープンデータ共同利用センター. (2026). AIくずし字認識アプリ「みを (miwo)」. CODH.
<https://codh.rois.ac.jp/miwo/> (3 सितंबर 2026 को अभिगमिति)
- [2] Clanuwat, T. (2022, 10월 27일). miwo (Version 1.1) [Mobile app]. App Store.
<https://apps.apple.com/jp/app/miwo/id1581794085> (3 सितंबर 2026 को अभिगमिति)
- [3] ROIS-DS 人文学オープンデータ共同利用センター. (2026). みを (miwo) とは? . CODH.
<https://codh.rois.ac.jp/miwo/about/> (3 सितंबर 2026 को अभिगमिति)
- [4] みんなで翻刻. (2026). みんなで翻刻 Minna de Honkoku. honkoku.org. <https://honkoku.org/> (3 सितंबर 2026 को अभिगमिति)

इस AI पुस्तकालय पुस्तक का समर्थन करें

यदि यह पुस्तक उपयोगी रही, तो आप PayPal के माध्यम से आगे के अनुवाद, सार्वजनिक संस्करण और खुली कृत्रिम बुद्धिमत्ता ज्ञान परियोजनाओं का समर्थन कर सकते हैं।

PayPal

<https://paypal.me/kimkjkj>



QR कोड स्कैन करें या ऊपर दिया गया लिंक खोलें।

अधिवक्ता कमि क्यूंगा-जनि