Cosmos Multidisciplinary Research E-Journal Online Available at www.cmrj.in
Recognized International Peer Reviewed Journal Impact Factor 4.94 ISSN No. 2456-1665

ग्रंथालयातील नवीन दालन : डिजिटायझेशन

डॉ. धाकड योगिनी पद्माकर
ग्रंथपाल,
आर.सी.पटेल शिक्षणशास्त्र महाविद्यालय,
शिरपुर जि. धुळे

प्रस्तावना :

२१ वे शतक हे आधुनिक जग आहे. माहिती व तंत्रज्ञानाचे युग आहे. या माहिती व तंत्रज्ञानावर आधारित या युगात सर्वच क्षेत्र अगदी अत्याधुनिक झालेले आपणास आढळतात. अगदी सहजपणे1 व जलदगतीने आपणास ही माहिती उपलब्ध करुन देण्याची जबाबदारी या तंत्रज्ञानाने घेतलेली आपणास आढळते. जगभरच्या आधुनिक ग्रंथालयांमध्ये डिजिटायझेशन नावाच्या तंत्रज्ञानाचा मोठा बोलबाला आहे. जी ग्रंथालये डिजिटायझेशनच्या बाबतीत पुढारलेली आहेत त्यांना 'ट्वेंटी फर्स्ट सेंचुरी मॉडर्न लायब्ररीज' असं संबोधन काहीशा आदरानं बहाल केलं जातं.

एकविसाव्या शतकाच्या प्रारंभी ग्रंथालयाचा एक अविभाज्य भाग बनून गेलेलं हे डिजिटायझेशन नावाचं तंत्र नेमकं कशा प्रकारचं आहे ? ग्रंथालयीन कारभारात त्यामुळे नेमके कोणते बदल होत आहेत? आणि, आजच्या डिजिटायझेशन तंत्राची आवृत्ती ही जर व्हर्जन 1.0 मानली तर भविष्यकाळात येणाऱ्या डिजिटायझेशनच्या पुढल्या आवृत्त्या 2.0 किंवा 3.0 वगैरेंचे अवतार ग्रंथालयांना कुठे घेऊन जाण्याची शक्यता आहे ? ह्या तीन मुद्यांचा किंचित वेध घेण्याचा प्रयत्न ह्या छोट्या लेखात केला आहे. विषयाला हात घालण्यापूर्वी ग्रंथालयाच्या संदर्भात 'डिजिटायझेशन' ह्या इंग्रजी शब्दाची पार्श्वभूमी आणि नेमका अर्थ समजून घेतला पाहिजे.

अर्थ- 'डिजिटायझेशन'

'डिजिटायझेशन' हा इंग्रजी शब्द आज वेगवेगळ्या क्षेत्रात वेगवेगळ्या अर्थाने आणि संदर्भाने वापरला जातो. जुन्या काळातील एखाद्या चित्रपटाची रीळं घेऊन आज त्याचं रूपांतर जर कोणी डिव्हीडीमध्ये केलं तर त्याला डिजिटायझेशन केलं असं म्हणता येईल. डिजिटायझेशनचा हा एक प्रकार झाला. आणखी एक वेगळं उदाहरणं घेऊ. रंगमंचावर प्रत्यक्ष चाललेल्या एखाद्या नाट्यप्रयोगाचं चित्रीकरण कोणी केलं तर त्याला डिजिटायझेशन म्हणता येईल का? त्याचं उत्तर 'होय' आणि 'नाही' असं दोन्ही प्रकारे येऊ शकतं. म्हणजे, जुन्या पद्धतीच्या कॅमेऱ्याने चित्रीकरण केलं गेलं आणि त्यातून जर फिल्मची रीळं तयार झाली तर त्याला डिजिटायझेशन म्हणता येणार नाही. पण तेच चित्रीकरण जर डिजिटल कॅमेऱ्याने केले असेल तर त्यातून थेट डिजिटल डेटा तयार होईल. त्यामुळे त्याला 'डिजिटायझेशन' म्हणता येईल. सेल्युलॉईड किंवा प्लास्टीकची फिल्म हे अनालॉग माध्यम आहे, तर डिजिटल डेटा साठवणारी डिव्हीडी किंवा युएसबी डिस्क (पेन ड्राईव्ह वगैरे) हे डिजिटल माध्यम आहे. डिजिटल

डेटा हा संगणकाच्या फाईलच्या स्वरूपात बीटस आणि बाईटसमध्ये साठवला जातो हे सर्वांनाच माहित आहे.

प्रत्यक्ष चाललेल्या नाट्यप्रयोगाच्या डिजिटायझेशनचे वर चर्चेत आलेले दोन वेगवेगळे प्रकार नीट लक्षात घ्यायला हवेत. पहिल्या प्रकारात, प्रत्यक्ष नाट्यप्रयोग चाललेला असताना अनालॉग पद्धतीच्या कॅमेऱ्याने शूटींग झाले, त्यातून फिल्मची रीळे तयार झाली, आणि नंतर त्या फिल्मच्या रीळांचे रूपांतर डिजिटल डेटामध्ये करून ते डीव्हीडी किंवा युएसबी डिस्कमध्ये साठवण्यांत आले. दुसऱ्या प्रकारात, नाट्यप्रयोग वा सादरीकरण प्रत्यक्ष चाललेले असताना थेट डिजिटल कॅमेऱ्यानेच शूटींग झाले. त्यामुळे त्या शूटींगमधून डिजिटल डेटा मिळाला. ह्या दुसऱ्या प्रकारात प्रत्यक्ष सादरीकरणाचे थेट डिजिटायझेशन झाले. तिथे जुन्या अनालॉगचे डिजिटायझेशन करण्याचा प्रश्न आला नाही.

वरील दोन्ही प्रकारावरून आपण पाहिले की डिजिटायझेशन होण्यासाठी एक तर मूळ अनालॉग स्वरूपातील माध्यम उपलब्ध हवे, किंवा ज्याचे डिजिटायझेशन होऊ शकेल असे प्रत्यक्ष दृश्य (सादरीकरण वगैरे) माध्यम उपलब्ध हवे. डिजिटल नसलेल्या माध्यमातील दृश्य, चित्र, ध्वनी, किंवा अक्षरांचे रूपांतर डिजिटल स्वरूपात केले गेल्यास त्या प्रक्रियेला डिजिटायझेशन म्हंटले जाईल. पारंपारिक ग्रंथालयांमध्ये सामान्यतः पुस्तके, नकाशे, हस्तलिखिते, वृत्तपत्रे, नियतकालिके यासारख्या मूक दस्तऐवजांचा संग्रह असे. तिथे आवाज किंवा चलत्चित्रे नसत. नंतरच्या काळात आवाज व चलत्चित्रांचे रेकॉर्डिंग जेव्हा सुरू झाले तेव्हा अनेक ग्रंथालयांमध्ये रेकॉर्डसच्या काळ्या तबकड्यांचा संग्रह, आणि नंतर कॅसेट टेप्स यांची भर पडण्यास सुरूवात झाली. पुस्तके आणि ग्रंथांचे माध्यम जसा आनंद व ज्ञान देऊ शकते तसाच आनंद आणि ज्ञान ध्वनी साठवलेल्या रेकॉर्डस व कॅसेटसही देऊ शकतात. पुस्तकांना जसे संदर्भ मूल्य आहे तसेच संदर्भ मूल्य रेकॉर्डस व कॅसेटस यांनाही आहे. त्यामुळे काही मोजक्या ग्रंथालयांमध्ये त्यांच्या संग्रहांचे विभाग सुरू झाले. तबकड्या पद्धतीच्या काळ्या रेकॉर्डस किंवा टेपवरील कॅसेटस ही अनालॉग माध्यमं होती. पुढल्या काळात ध्वनी किंवा आवाज डिजिटल झाला. सीडींचा जमाना आला. जसा ध्वनी आणि आवाज डिजिटल झाला, तशी छायाचित्रेही डिजिटल झाली. फिल्म वापरून, निगेटीव्ह रसायनांमध्ये बुडवून तयार होणारी जुनी छायाचित्रे नंतरच्या काळात डिजिटल झाली. चित्रपटही डिजिटल झाले. यानंतर ओघाने मग पुस्तकेही डिजिटल स्वरूपात तयार होऊ लागली. या टप्प्यावर जुनी डिजिटल नसणारी म्हणजे अनालॉग प्रकारची गाणी, ध्वनी, छायाचित्रे वगैरे सामग्री हा एक प्रकार आणि नवी डिजिटल स्वरूपातील तीच सामग्री असे दोन प्रकार तयार झाले. त्यानंतर मग जुन्या अनालॉग सामग्रीला डिजिटल प्रकारात रूपांतरित करण्याचे प्रयत्न सुरू झाले. जुन्या सामग्रीचे डिजिटायझेशन होत असताना पुस्तकांच्या डिजिटायझेशनचे तंत्रही आले. एका बाजूला हार्डवेअरच्या बाजारपेठेत निरनिराळे स्कॅनर्स येऊ लागले आणि दुसरीकडे सॉफ्टवेअरच्या बाजारपेठेत pdf, mobi, epub वगैरे ई बुक्सचे सॉफ्टवेअर तंत्रज्ञान बहरू लागले. यातून जगभरच्या ग्रंथालयांच्या अंगणात डिजिटायझेशन नावाचा झरा वाहू लागला. आता त्याची गंगा जगभर वाहू लागली आहे.

'डिजिटायझेशन' आणि 'डिजिटलायझेशन'

काही लोक 'डिजिटायझेशन' ऐवजी 'डिजिटलायझेशन' असा शब्द वापरतात. ह्या दोन शब्दांत काही फरक आहे का? त्याचं उत्तर - 'नाही' असं आहे. हे दोन्ही शब्द सामान्यतः एकाच अर्थाने वापरले जातात. विकिपेडियावर (www.wikipedia.org) गेलात आणि तिथे Digitalization टाईप करून शोधलत तर तुम्हाला तो घेऊन जातो Digitization नावाच्या नोंदीवर. म्हणजे विकिपेडियाच्या दृष्टीने डिजिटायझेशन आणि डिजिटलायझेशन ही एकच गोष्ट आहे. मरियम वेबस्टर शब्दकोशात Digitization आणि Digitalization असे दोन निरनिराळे शब्द आहेत. पण त्यांचा अर्थ दोन्हीकडे सारखाच दिला आहे - the process of converting something to digital form. थोडक्यात, ह्या दोन्ही शब्दांचा अर्थ सारखाच आहे असं मानायला हरकत नाही. मात्र, SAP सारख्या मोठ्या कंपनीने ह्या दोन शब्दांत नेमका पण सूक्ष्म फरक केला आहे. इथे त्याची चर्चा करीत नाही. सॅपने जो सूक्ष्म फरक केला आहे तसा फरक ग्रंथालयाच्या पार्श्वभूमीवर आपण करून पाहू. समजा, ग्रंथालयातील सर्व ग्रंथांचे 'डिजिटायझेशन' झाले असेल तर ते ग्रंथालय 'डिजिटलाईज्ड' झाले म्हणता येईल का? तर, सॅपच्या व्याख्येनुसार तसे म्हणता येणार नाही. सर्व ग्रंथ 'डिजिटाईज्ड' केले पण ग्रंथालयात अजूनही कागदी कार्डे वापरली जात असतील, किंवा जाडजूड रजिस्टरमध्ये नोंदी होत असतील तर ते ग्रंथालय पूर्ण 'डिजिटलाईज्ड' झाले असे म्हणता येणार नाही. पण असा एखादा सूक्ष्म फरकाचा अपवाद वगळला तर 'डिजिटायझेशन' आणि 'डिजिटलायझेशन' यांच्यात अर्थाअर्थी काही फरक नाही.

डिजिटायझेशन आणि स्कॅनिंग

एखाद्या ग्रंथाची सर्व पाने स्कॅनरवर स्कॅन केली, आणि त्याची पीडीएफ फाईल तयार झाली म्हणजे त्या ग्रंथाचे डिजिटायझेशन पूर्ण झाले असे सुरूवातीच्या काळात गृहित धरले जाई. पण आज तसे गृहित धरता येणार नाही. कारण स्कॅनिंग केले म्हणजे पूर्ण डिजिटायझेशन झाले असे नव्हे. स्कॅनिंग हा फार तर डिजिटायझेशन प्रक्रियेतला एक भाग मानता येईल. ग्रंथाची पाने स्कॅनर वर जशी स्कॅन करता येतात, तशी डिजिटल कॅमेऱ्यानेही पुस्तकाच्या सर्व पानांचे फोटो घेता येऊ शकतात. आजकाल मोबाईल फोनच्या कॅमेऱ्याचा उपयोग करून स्कॅनिंगची सुविधा देणारी Camscanner वगैरे लोकप्रिय अप्लीकेशन्स (Apps) उपलब्ध आहेत. पण ग्रंथाच्या पानाची पूर्ण प्रतिमा मिळवणे म्हणजे ग्रंथाचे डिजिटायझेशन नव्हे. जो ग्रंथ खऱ्या अर्थाने डिजिटाईझ होतो त्या ग्रंथांतील शब्द वा संदर्भ डिजिटल पद्धतीने शोधता यायला हवेत. स्कॅनिंगने आपल्याला पुस्तकाच्या पानाची प्रतिमा किंवा फोटो फक्त मिळतो. त्या फोटोत दिसणारा मजकूर संदर्भासाठी शोधायचा म्हंटला तर तशी सोय फोटोत नसते. एखाद्या झाडाचा फोटो तसा पुस्तकाच्या पानाचा फोटो अशी ती मर्यादा असते. असे एखादे २५० पानांचे म्हणजे पानांच्या .jpg प्रतिमांचा जुडगा असलेले डिजिटल पुस्तक त्यातील संदर्भ शोधण्यासाठी उपयोगात आणता येत नाही. संदर्भ शोध किंवा Search च्या सोयीसाठी स्कॅनिंग नंतर OCR (Optical Character Recognition) नावाचे तंत्र वापरून ग्रंथाच्या

डिजिटल पानांचा मजकूर सुटा करावा लागतो. आजकाल ह्या OCR तंत्राची सोय आधुनिक स्कॅनरमध्ये दिलेली असते. मात्र, त्याची माहिती डिजिटायझेशन करणाऱ्या ग्रंथालय कर्मचाऱ्याला असावी लागते. दुर्दैवाने ती अनेकदा नसते.

आजमितीस इंग्रजीसह जगातल्या सुमारे १५० हून अधिक लिप्यांसाठी OCR तंत्रज्ञान उपलब्ध आहे. म्हणजे, पुस्तकाचे पान स्कॅन करा, त्या पानाच्या प्रतिमेत दिसणारा मजकूर ओसीआर तंत्र वापरून सुटा करा. तो सुटा झाला की ते शब्द Search साठी उपलब्ध होतात, किंवा तो मजकूर संपादित करणं शक्य होतं. जगात अनेक लिप्यांना ही सोय लाभली असली तरी दुर्दैवाने देवनागरी आणि काही भारतीय भाषांसाठी आजही उत्तम व अचूक ओसीआर तंत्रज्ञान उपलब्ध झालेले नाही. असो.

डिजिटल माध्यम आणि डिजिटाईज्ड ग्रंथांचे स्वरूप

छापिल ग्रंथ हे स्थिर माध्यम आहे. पुस्तक एकदा छापले की त्याचा आकार बदलत नाही. पुस्तकातील मजकूराचा टाईप व त्याचा आकार लहान-मोठा होऊ शकत नाही. छापिल पुस्तक म्हणजे सारे स्थिर, पक्के व न बदलता येणारे असते. पण, डिजिटल माध्यम हे एके जागी स्थिर राहणारे नव्हे. डिजिटल माध्यमाला काही लोक ई माध्यम असेही म्हणतात. त्यामुळे डिजिटल बुक हे ई बुक, आणि डिजिटल लायब्ररी ही ई लायब्ररी म्हणूनही ओळखली जाते. कागदावर छापलेले किंवा प्रिंट बुक व ई बुक यातला फरक स्थिर पुस्तक आणि वाचकाच्या गरजेप्रमाणे बदलता येणारे पुस्तक असा असतो. उदाहरणार्थ ई बुक हे मोबाईल आडवा धरून आणि उभा धरून अशा दोन्ही प्रकारे वाचता येते. एखादे २५० पानांचे पुस्तक मोबाईलवर मजकूराचा आकार वाढवला तर ७५० पानांचे वा त्यापेक्षाही अधिक पानांचे बनून जाते. कारण टाईपचा आकार वाढवल्याने एका पानावर फार कमी शब्द राहतात, आणि स्वाभाविकच त्यामुळे पाने वाढतात. पण टाईप मोठा करण्याची सोय असल्याने वाचकाची, विशेषतः वयस्कर वाचकांची फार मोठी सोय त्यामुळे होते. ई बुकच्या पानावरील एखाद्या शब्दाचा अर्थ पहायचा असेल तर तो तिथल्या तिथे पाहण्यासाठी डिजिटल डिक्शनरीची साथ ई बुकला असते.

अशीच एक सोय म्हणजे पुस्तकाचा मजकूर मोठ्या आवाजात वाचून दाखवण्याची सोय. ज्या मंडळींना वाचनामुळे डोळ्यांना त्रास होतो किंवा अंध व्यक्तींना ब्रेल लिपीतील पुस्तक नसेल तर वाचन शक्य होत नाही त्यांना ई बुक चा मजकूर मोठ्या आवाजात वाचून दाखवण्याची सोय असल्याने खूपच उपयुक्त ठरते. २०१७ मधली ई बुक्स तांत्रिक दृष्ट्या वयात आलेली आहेत. पण, ई बुक तंत्रज्ञानातले संशोधन थांबलेले नाही. ई बुक्समध्ये आणखी काय करता येईल याचा प्रयत्न आजही चाललेला आहे.

डिजिटल बुक आणि Text to speech-

डिजिटल माध्यमाची जी जी वैशिष्ट्ये आहेत ती डिजिटल पुस्तकाला म्हणजे ई बुकलाही लाभतात. त्याचं एक साधं उदाहरण म्हणजे Text to speech चे तंत्र. 'टेक्स्ट टू स्पीच' हे एक स्वतंत्र तंत्र आहे. स्क्रीनवर जो शब्द आहे तो स्पष्ट उच्चारासह मानवी आवाजात वाचण्याचे डिजिटल तंत्र उपलब्ध झाल्यावर ते ई बुकला जोडलं जाणं स्वाभाविक होतं. एकीकडे ई बुकचे तंत्र विकसित होत होते, तर दुसरीकडे 'टेक्स्ट टू स्पीच' चे तंत्र विकसित होत होते. स्वल्पविरामाला किंचित थांबणे, पूर्णविरामाला त्यापेक्षा अधिक थांबणे, वाचताना एखाद्या शब्दावर जोर देणे, परिच्छेद बदलताना पूर्णविरामापेक्षा किंचित अधिक थांबून परिच्छेद बदलल्याची जाणीव देणे, वेळ वाचावा म्हणून भराभर वाचणे, किंवा एक एक शब्द नीट कळावा यासाठी हळूहळू वाचणे वगैरे तंत्र जसजसे आणखी विकसित होत जाईल तसतसं रोबो म्हणजे यंत्रमानवाने वाचून दाखवलेला मजकूर ऐकून आकलन होणे सुलभ होईल. ऑडिओ बुक हा ई बुकचा एक स्वतंत्र प्रकार आज खूपच लोकप्रिय झालेला आहे. हातातल्या मोबाईलवर किंवा टॅबवर डोळ्याने ई बुक वाचताना दुसरीकडे कानात ईयर फोन घालून तेच शब्द ऐकण्याची सोय आज शेकडो विद्यार्थ्यांना अभ्यासात उपयोगी पडताना दिसते आहे. कारण एका बाजूला डोळ्याने वाचत असताना, दुसरीकडे दुसरे ज्ञानेंद्रिय ते श्रवण करीत असते. त्यामुळे मेंदूत साठून लक्षात राहणे अधिक सुकर होते असा अनेक विद्यार्थ्यांचा अनुभव आहे.

संदर्भ :

- डॉ. द.ना.फडके, ग्रंथालय संगणकीकरण आणि आधुनिकिकरण, युनिव्हर्सल प्रकाशन,पुणे
- डॉ.सत्यप्रकाश निकोसे, ग्रंथालय व्यवस्थापन,,प्रज्ञा प्रकाशन,नागपूर
- डॉ.सत्यप्रकाश निकोसे, ग्रंथालय आणि माहितीशास्त्र, संशोधन पदधती, ,प्रज्ञा प्रकाशन,नागपूर
- डॉ.प्रकाश जैन आणि डॉ. मंगला हिरवाडे, सुलभ माहिती तंत्रज्ञान
- एस पी पवार,बडकते, पवार,जाधव,कुलकर्णी, ग्रंथालय व माहीतीशास्त्र, फडके प्रकाशन,कोल्हापुर
- www.wikipedia.com
- www.google.com