பல்லூடக ஆவணகத் தளச் (Multimedia Archiving Platform) செயற்றிட்டக் கற்றல்கள்

நூலக நிறுவனத்தின் பல்லூடக ஆவணகத் தளதுக்கான (Multimedia Archives Platform) முன்மாதிரியை கட்டமைக்கும் முயற்சி 2011 இல் இருந்து பல தன்னார்வலர்களால் முன்னெடுக்கப்பட்டது. குறிப்பான இந்தச் செயற்திட்டம் (NF : Project : Prototyping Integrated Digital Preservation Solution) 2016 க்கான ஆண்டுத் திட்டமிடல் மற்றும் வரவுசெலவில் (2016 APB) ஏற்கப்பட்டு நடைமுறைப்படுத்தப்பட்டது. இந்தச் செயற்திட்டம் ஊடாக நாம் பல கற்றல்களைப் பெற்றுக் கொண்டோம். இந்த அறிக்கை அந்தக் கற்றல்களைப் பகிர்ந்துகொள்கிறது.

உள்ளடக்கம்

தேவைகள் பகுப்பாய்வு
மென்பொருள் தெரிவு
ஐலண்டோராவின் கூறுகள்
ஐலண்டோராவை நிறுவுதல், அமைப்பு வடிவமைத்தல்
அடுத்த தலைமுறை ஐலண்டோரா - https://islandora.ca/CLAW
நிறுவனம் சார் சவால்கள்

தேவைகள் பகுப்பாய்வு

நூலக நிறுவனத்தின் எண்ணிம நூலகம் அச்சில் பதிக்கப்பட்ட எழுத்தாவணப் படைப்புக்களையே (published works) பெரும்பாலும் பாதுகாத்து அணுக்கப்படுத்தி வந்துள்ளது. அச்சு எழுத்தாவணங்களைத் தாண்டி பல்லூடகங்களை, எண்ணிமப் பிறப்புக் கொண்ட படைப்புக்களை ஆவணப்படுவது அவசியம் என்பது 2011 வியூகத் திட்டமிடல் ஊடாக நன்கு உணரப்பட்டது. பல்லூடக வளங்களைச் சேகரித்தல், ஒழுங்குபடுத்தல், பாதுகாத்தல், காட்சிப்படுத்தல், தேடுதல் உட்பட்ட வசதிகளைக் கொண்ட ஒரு தளத்தை ஆயும் முயற்சிகள் 2011 காலப் பகுதியிலேயே தொடங்கிவிட்டது. எமது தேவைகளை நாம் Digital Library Software Evaluation Metrics என்ற ஆவணம் ஊடாக பதிவுசெய்தோம்.

குறிப்பாக எமக்குப் பின்வரும் தேவைகள் இருந்தன:

- தேடல் மற்றும் காட்சிப்படுத்தல் (Search and Discovery Mechanism)
- பல்லூடகங்களைச் (ஒலிக் கோப்பு, நிகழ்படம், படங்கள், எண்ணிம ஆவணங்கள், பிற) சேகரித்தல், பாதுகாத்தல், காட்சிப்படுத்தல்
- சார்தரங்கள் அடிப்படையிலானா மூதரவ ஆதரவு, மூதரவுகளை உருவாக்க இற்றைப்படுதுவதற்கான ஆதரவு - Support for Metadata Standards, Ability to create, update metadata records
- பாதுகாப்புக் கூறுகள் (Preservation Features): சரிபார் தொகை (Checksum), நுட்ப மூதரவு (Technical Metadata), கண்காணிப்பு (Audit)
- எளிய பயனர் இடைமுகம், பணியோட்டம் (Simple User Interface and Work-flow)
- நேரடிப் பயனர் பங்களிப்பின் ஊடான உள்ளடக்க வளர்ச்சி (Enabling user contributions)
- சேகரங்களுக்கான ஆதரவு, குறிப்பாக ஓர் ஆவணம் பல சேகரங்களில் இடம்பெறுவதற்கான வசதி
- பட்டியலாக்கம் (Cataloguing)
- நீட்டப்படக்கூடியது (Expendable) விரிவாக்கத்தக்கது (Scalable)

மென்பொருள் தெரிவு

2011 இலேயே நாம் சில மென்பொருட்களை ஆய்ந்தோம். விக்கி (mediawiki.org), ட்ரூப்பல் (drupal), ஓப்பின் லைபிரரி (openlibrary.org) ஆகிய தளங்கள் முதலில் பரிசோதிக்கப்பட்டன. விரிவாக்கத்து ஏற்றதாகவும் (Scalability) நூலக எண்ணிம நூலக விக்கியையும் ஒத்தாகவும், பல்வேறு எண்ணிம நூலக வசதிகளையும் கொண்ட ஓப்பின் லைபிரரியை நாம் விரிவாகப் பரிசோதிக்க தேர்ந்தெடுத்தோம். Infogami மூது நிறுவ முடியும்

என்று அறிந்தோம். ஓப்பின் லைபிரரியை நிறுவ முயன்றோம். அதை நிறுவுவதில், சோலரை (Solr) அமைப்புவடிவமைப்பதில் சிக்கல்களை எதிர்கொண்டோம்.

2013 இன் இறுதியில், 2014 இல் இந்தத் தேடல் விரிவானது. பின்வரும் மென்பொருட்கள் நிறுவப்பட்டு பரிசோதிக்கப்பட்டன:

- டி.எசுபேசு (dspace.org) - நிறுவனக் களஞ்சிய மென்பொருள் (Institutional Repository - IR)
- ஒமேகா (omeka.org) - உள்ளடக்க மேலாண்மை மென்பொருள் (Content Management System - CMS)
- mukurtu.org - எண்ணிம பண்பாட்டு மரபுக் களஞ்சியம்
- கோகா (koha.org) - ஒருங்கிணைந்த நூலக கட்டகம் (integrated Library System - ILS/ILM)
- ஆர்கைவ்மெற்றிக்கா (archivematica.org) + அற்ரம் (accesstomemory.org) - எண்ணிமப் பாதுகாப்பு மற்றும் காட்சிப்படுத்தல் மென்பொருட்கள் (digital preservation system & discovery mechanism)

2015 ஐன் இறுதியில் நாம் ஆர்கைவ்மெற்றிக்கா + அற்ரம் ஆகியவற்றை கட்டற்ற நிரலர்களின் (FOSS Developers) உதவியுடன் நிறுவி துல்லியமாக பரிசோதித்தோம். இதனை எமது தீர்வாக தெரிவு செய்து ஒரு முன்னோட்டச் செயற்திட்டத்தை (pilot project) 2016 இல் செய்வது என்று முடிவெடுத்தோம். அவ்வாறு செயற்படுத்துகையில் ஆர்கைவ்மெற்றிக்கா + அற்ரம் இல் நாம் சில குறிப்பான சிக்கல்களைக் கண்டடைந்தோம். ஆர்கைவ்மெற்றிக்கா மரபுசார் ஆவணகத்துக்கு (traditional archive) ஆதரவு தரும் வண்ணம் வடிவமைக்கப்பட்டது. அதனால் நாம் ஒரு பொருளை இரண்டு சேகரங்களுக்குள் உறுப்பினராக வைக்க முடியாது. இது தொடர்பாக நாம் ஆர்கைவ்மெற்றிக்கா ஆக்குநர்களுடன் விரிவாக உரையாடினோம்.

- https://groups.google.com/forum/#!topic/ica-atom-users/DEa4bfQrUOg
- https://groups.google.com/forum/#!topic/archivematica/L4TvDzbGdbA

அடுத்த சிக்கல் ஆர்கைவ்மெற்றிக்கா + அற்ரம் முறையாகப் பயன்படுத்துவது தொடர்பாக ஊழியர்களுக்கும் பொதுப் பயனர்களுக்கும் பயிற்சி தருவது சிரமாக இருந்தது. ஆர்கைவ்மெற்றிக்காவில் மீதரவுகளைத் தொகுப்பதற்கான வசதி குறைவாகவே இருந்தது. ஆர்கைவ்மெற்றிக்காவில் ஒரு ஆவணக தகவல் பொதியை (Archival Information Package - AIP) உருவாக்கிய பின்பு அதனை மாற்றுவது கடினமாகும். மேற் சுட்டப்பட்ட காரணங்களால் நாம் ஐலன்டோரா (islandora.ca) + ஃபெடோரா (fedorarepository.org) + சோலர் (lucene.apache.org/solr/) தீர்வை நோக்கி நகர்ந்தோம்.

ஐலண்டோராவின் கூறுகள்

ஐலண்டோரா பல்லூடக எண்ணிம வளங்களை பாதுகாக்க, மேலாண்மை செய்ய, அணுக்கப்படுத்த பயன்படும் ஒரு கட்டற்ற மென்பொருள் தளம் ஆகும். இது ஃபெடோரா, டுருப்பல் (drupal.org), சோலர் உட்பட்ட பல்வேறு கட்டற்ற மென்பொருளைப் பயன்படுத்தி அல்லது அவற்றோடு சேர்ந்து இயங்கிறது. இந்தத் தளத்தை நூறுக்கும் மேற்பட்ட பல்கலைக்கழகங்கள், நூலகங்கள், ஆவணகங்கள், அருங்காட்சியகங்கள் பயன்படுத்துகின்றன. இந்த மென்பொருள் பல நிறுவனங்களால் கூட்டாக கிட்கப்பில் (github.com/Islandora) உருவாக்கப்பட்டு கட்டற்ற மென்பொருளாக பகிரப்படுகின்றது.

நாம் ஐலண்டோராவை நோக்கி நகர்வதற்கு எமது நுட்பப் பங்களிப்பாளர் ஒருவர் இந்த நுட்பங்களில் நேரடி அனுபவம் பெறும் வாய்ப்புக் கிடைத்ததும் ஒரு முக்கிய காரணம் ஆகும். இந்த நுட்பங்களை நாம் முன்னரும் அலசி இருந்தாலும், அவற்றை நிறுவுவது எமக்கு முன்னர் சவாலாக இருந்தது. ஆனால் மேலதிக அனுபவங்களோடும், அன்சிபிள் போன்ற செயலிகளை நிறுவப் பயன்படும் புதிய தொழில்நுட்பங்களைப் பயன்படுத்தியும், ஐலண்டோரா கட்டற்ற மென்பொருள் சமூக ஆதரவோடும் நாம் இந்த மென்பொருளை வெற்றிகரமாக நிறுவினோம். ஐலண்டோரா + ஃபெடோரா + சோலர் எமது தேவைகளுக்கு நல்ல தீர்வாக அமைந்தது.

செயல் சார் கூறுகள் (Functional Features)

- பல்லூடாக ஆதரவு (Multimedia Support)
- எளிய தேடல் (Simple/Text Search), வினவல் தேடல் (Query Search)
- உலாவுதல் (Browsing), முகப்பு உலாவு (Faceted Browsing)

- எளிய பயனர் இடைமுகம், பணியோட்டம் (Simple User Interface and Workflow)
- சேகரங்கள் - ஒரு உறுப்பு பல சேகரங்கள் (Collections - one to many mapping)
- பாதுகாப்புக் கூறுகள் (Preservation Features): சரிபார் தொகை (Checksum), நுட்ப மூதரவு (Technical Metadata), கண்காணிப்பு (Audit)
- திறந்த ஆவணக தகவல் முறைமை ஆதரவு (OAIS Support)
- தொகுப்பு உள்ளூடு (Batch Processing)
- தரவேற்றம், இறக்கம் (Export & Import)
- நீட்டப்படக் கூடியது (Extensible) - பல கூறுகங்கள் உள்ளன
- விரிவாக்கத்துக்கு (Scalable) - பெரும் நிறுவனங்கள் பயன்படுத்துகின்றன

செயல் சாராக் கூறுகள் (Non Functional Features)

- கட்டற்ற மென்பொருள்
- மென்பொருள் சார் சமூகம். ஆதரவு
- தொடர் விருத்தி, வளர்ச்சி
- ஒப்பீட்டளவில் எளிதாக நிறுவ கட்டியது பராமரிக்கக் கூடியது

ஐலண்டோராவை நிறுவுதல், அமைப்பு வடிவமைத்தல்

அண்மைக் காலமாக பல பகுதிகளைக் கொண்ட செயலிகளை நிறுவும் (installation/deployment/configuration) பணிகள் தானியக்கம் (automation) செய்யப்பட்டு வருகின்றன. இதில் பரவலாகப் பயன்படுத்தப்படும் ஒரு configuration management மற்றும் deployment automation செயலி அன்சிபிள் (Ansible) ஆகும். இதனை நாம் நூலக ஆய்வகத்தில் பரிசோதித்தோம். மேலும் அண்மைக் காலமாகா ஒரு மென்பொருளை deploy செய்ய பல்வேறு வகையான புரவல் சேவைகள் மற்றும் நுட்பங்கள் பயன்படுகின்றன. நாம் Docker போன்றவற்றைப் பரிசோதித்தோம். பின்னர் நாம் அன்சிபிள் + மெய்நிகர் வழக்கி (Digital Ocean) தீர்வைத் தேர்தெடுத்தோம்.

நாம் ஐலண்டோராவுக்கான பின்வரும் அன்சிபிள் ஃபிளேபுக் - https://github.com/jefferyb/islandora-7.x-enterprise-ansible க்கை பயன்படுத்தி ஐலண்டோராவின் core stack ஐ நிறுவினோம். பின்னர் மேலதிக modules, solr configurations ஆகியவற்றைச் செய்தோம்.

Solr ஐ முறைப்படி configure செய்வது முக்கிய ஒரு சிக்கலாக இருந்தது. அதனை நாம் ஐலண்டோரா சமூகத்தின் கூகிள் மன்றத்தின் உதவியுன் ஊடாகப் பெரும்பாலும் செய்யக் கூடியதாக இருந்தது.

குறிப்பாக நாம் என்ன Content Models ஐ பயன்படுத்துவது என்பது எமக்கு முன் ஒரு முக்கிய தேர்வாக அமைந்ததும். அந்த Content Models இன் default MODS Forms சற்று மாற்றி அமைப்பது எமக்கு தேவையாக இருந்தது. ஐலண்டோரா இவற்றை மிகவும் எளிமையாக செய்ய அனுமதித்தது.

அடுத்த தலைமுறை ஐலண்டோரா - https://islandora.ca/CLAW

நாம் தற்போது நிறுவி இருப்பது ஐலண்டோரா 1.7.8 பதிப்பு ஆகும். ஐலண்டோராவின் அடுத்த வெளியீடு மிகவும் மாறுபட்டதாகவும், மிகப் பெரிய ஒரு பாச்சலாகவும், இணைப்புத் தரவு நுட்பங்களுக்கு ஆதரவு தரும் வகையிலும் அமையும். நாம் இந்தப் பதிப்பில் 1 இல் இருந்து 2 ஆண்டுகள் இருந்துவிட்டு, அடுத்த இணைப்புத் தரவுக் கட்டமைப்பில் அமைந்த தளத்துக்கு நகர்வதற்கு திட்டமிடுவோம். அது ஒரு பெரும் கனவு. இணையுங்கள்.

நிறுவனம் சார் சவால்கள்

பல்லூடாக ஆவணகத் தளத்தை நிறுவுவதில் நூலக நிறுவனம் பின்வரும் சவால்களை எதிர்நோக்கியது:

- அச்சில் பதிக்கப்பட்ட ஆவணங்களுக்கு முக்கியத்துவம் தந்தது போல், பல்லூடக மற்றும் எண்ணிமப் பிறப்புக் கொண்ட ஆவணங்களை ஆவணப்படுத்த வேண்டும் என்பது நூலக நிறுவனத்தில் முதன்மை படுத்தப்படவில்லை.
- நூலக நிறுவனம் நுட்ப விடயங்களுக்கு போதிய முக்கியத்துவம் தரவில்லை. எ.கா போதிய நிதி, ஆளணி வளங்கள் ஒதுக்கப்படவில்லை.
- போதிய துறைசார் அனுபவம் உள்ள தன்னார்வலர்களை உள்வாங்காதது. அப்படி அவர்கள் பங்களித்தாலும், அவர்களின் பங்களிப்பை வினைத்திறனோடு முன்நகர்த்திச் செல்லாமை.
- தன்னார்வலர்கள் தொடர்ச்சியான ஈடுபாட்டுடன் கடமைப்பாட்டுடன் செயற்திட்டங்களைத் தொடராதல்.
- முடிவெடுத்தலில், மென்பொருட்களை சூர்தரங்களை தெரிவுசெய்வதில், முன்மாதிரிகளை நிறைவேற்றுவதில் மிகையான நேரம் எடுத்துக் கொண்டது. செயற்திட்டங்கள், முதன்மைப் பணிகள் நிறுவனச் சூரமைப்புப் பணிகளுக்கென அடிக்கடி முற்றிலும் நிறுத்தப்பட்டு பின்போடப்பட்டது.
- செயற்திட்டங்களை நிர்வாக நோக்கில் மேலாண்மை செய்ய போதிய மேலாண்மை செய்யும் ஊக்கம் அல்லது திறன் தன்னார்வலர்கள் மத்திலோ அல்லது ஊழியர்கள் மட்டத்திலோ இல்லாதது.

மேற்சுட்டப்பட்ட விடயங்கள் நூலக நிறுவனம் ஓர் உயர்ந்த தொழிற்திறன் மற்றும் செயற்திறன் வாய்ந்த நிறுவனமாகச் செயற்படும் போது மேம்படுத்த வேண்டிய விடயங்கள் ஆகும். பல விடயங்கள் அடையாளம் காணப்பட்டு மேம்படுத்தப்பட்டு வருகின்றன.