Summer 4-29-2022

Use of Open-Source Software and Cloud Computing Technology for Imparting Training

Francis Jayakanth Dr
Indian institute of s, francis@iisc.ac.in

Ananda T. Byrappa Dr
Indian Institute of Science, Bangalore, anandtb@iisc.ac.in

Pawan Kumar Singh Mr
Depart of Posts, Government of India, pawan.dop2@gmail.com

Use of Open-Source Software and Cloud Computing Technology for Imparting Training

Abstract

The Department of Posts, Government of India, to date, has set up six full-time postal training centres (PTC) across India to meet the training requirements of its five-lakh strong workforce. PTC Madurai (PTCM), situated in Tamil Nadu, is one such training centre. PTCM has taken a proactive initiative to explore and adopt information technologies in the training and learning domains. It has partnered with the JRD Tata Memorial Library, Indian Institute of Science, Bangalore, through a small project to set up a digital infrastructure to manage its training programmes and the library collection.

PTCM was looking for a digital library to manage its training materials, an online catalogue for its library collection, and a website. The consultants used three different open-source software to implement the requirements of PTCM viz; EPrints[1] for developing the digital library, Koha[2] for the integrated library management system, and WordPress[3] content management system for developing the website.

Keywords

Open-source Software (OSS), Cloud Computing, Virtual Machine, Digital Library, EPrints, Content Management System, WordPress, Library Management System, Koha.

Introduction

---

[1] https://eprints.org/uk
[2] https://koha-community.org/
[3] https://wordpress.com/

The use of information and communication technologies (ICT) in education and training is becoming more widespread, especially in the present scenario arising out of the COVID-19 pandemic. The ICTs can transform the traditional education system, predominantly teacher-centric and printed content-centric, into a rich student-centric environment. Some of the benefits of ICT-based based education and training include 1. Cost-effectiveness 2. Course materials can be accessed anytime, anywhere, and by multiple people simultaneously. 3. Create a digital culture among the participants. 4. Reduces the usage of paper 5. Enhance the teaching methodology by integrating multimedia content. 6. It can serve as a self-learning platform.

To date, the Department of Posts, Government of India, has set up six full-time postal training centres (PTC) across India to meet the training requirements of its five-lakh strong workforce. PTC Madurai is one of the six training centres. The centre has proactively explored and adopted ICT-based training programmes to train new and existing staff members. To this end, it has partnered with the JRD Tata Memorial Library[4], Indian Institute of Science[5], Bangalore, through a small consultancy project to set up a Digital infrastructure for PTC Madurai to facilitate ICT-based training programmes. The PTCM digital infrastructure has the potential to be expanded and positioned as an integrated training and learning management system (LMS) platform for the other postal training centres in India.

Training is an essential function in any organization. It provides an opportunity for the employees to enhance their skills, capabilities, and knowledge which, in turn, will aid them in discharging their duties effectively and efficiently at their workplaces. To introduce OSS and ICT-based training programmes at PTCM, the project consultants suggested 1. creation of a digital library

---

[4] https://library.iisc.ac.in
[5] https://www.iisc.ac.in

(DL) for managing its training materials, 2. setting up an integrated library system (ILS) for managing its library collection, and 3. creation of a website to integrate its digital initiatives. The said suggestions were acceptable to PTCM. As a result, the project consultants successfully implemented all the project deliverables using the corresponding open-source software on the Baadal cloud platform.

Objectives of the Study

The objectives of the study included the implementation of the following digital solutions using OSS to meet the requirements of PTCM:

- A digital repository with requisite local customization to manage the PTCM course materials

- An integrated library management system for managing the PTCM library collection

- Develop a website to integrate the activities and services provided by the PTCM

- Host the proposed OSS implementations on a cloud platform.

Need for the Implementation of the OSS

PTCM was established in August 1981, catering to the needs of the Tamil Nadu Postal Circle. In addition, there are five other postal training centres (PTCs) established across different regions of India. The PTCs are primarily engaged in grooming the newly inducted employees at the Postal Assistant/Sorting Assistant level and refreshing the employees' skills at other operative and supervisory cadres. Some specialized programmes are also conducted based on identifying the specific needs of the employees of the target group.

PTCM proactively initiated the introduction of ICT for its activities. To this end, it approached the JRD Tata Memorial Library, Indian Institute of Science, Bangalore, to explore the possibility of adopting ICT in its training programmes. Based on a couple of interactions with the Director, PTCM, the library proposed a small consultancy project to develop a digital library, an integrated library management system, and a website to enable PTCM to adopt ICT for its activities. Being a low budget and short-term project, the project consultants made a conscious decision to implement OSS solutions on a cloud platform to meet all the requirements of the PTCM. The consultants, therefore, did not want PTCM to make a capital investment in computer hardware, software, and human resources.

OSS is both a philosophy and a process. As a philosophy, it describes the intended use of software and methods for its distribution. OSS is also a process for the creation and maintenance of software[6]. For many years now, Open-source Software (OSS) has been a hotly debated topic, both in research and practice. One of the reasons for this is that several open-source products like GNU/Linux or Apache have now for years been in the spotlight as leaders in their application areas and continue to be so (Koch, 2007: xxvi-xxix). As a result, OSS applications have gained widespread support, and several national governments are adopting policies that encourage educational institutions to use OSS (Davidson, 2008). Some key characteristics of OSS include free distribution, openly available source code, and permission for modification of software (Erlich and Aviv, 2007; Subramanyam and Xia, 2008).

There are thousands of OSS products available, covering almost every software category. SourceForge[7] is the largest, most trusted destination for OSS discovery, software reviews, and

---

[6] http://infomotions.com/musings/ossnlibraries-lita/
[7] https://sourceforge.net/about

software development on the web. As of August 2021, the SourceForge repository claimed to host more than 502,000 OSS projects and millions of registered users.

To implement the project deliverables, the consultants decided to use the OSS solutions for developing a digital library (DL), a library management system (LMS), and a content management system for building a website. Two virtual machines are used on the Baadal cloud platform to implement the project deliverables. Baadal[8] is a cloud computing environment that provides virialized computing resources for academic and scientific domains in India.

Requirements Analysis for PTCM Digital Library

A digital library, digital repository, or digital collection is an online digital object database that can include text, still images, audio, video, or other digital media formats. Objects can consist of digitized content like print or photographs and originally produced digital content like word processor files or social media posts. In addition to storing content, digital libraries provide means for organizing, searching, and retrieving the content contained in the collection[9].

The PTCM conducts various training courses that include induction training for the newly appointed staff, special training for the marketing executives, and other specialized training programmes for targeted groups. Similarly, other postal training centres conduct training courses to meet their zones. All the said training courses are residential and classroom-based. The training instructors prepare the training materials to impart training to the participants. However, the training material has limitations, whether in standalone digital or print format. In print format, the sharing of the content widely is not easy. Even the digital content on a standalone computer or

---

[8] https://baadal.nmeict.in/
[9] https://en.wikipedia.org/wiki/Digital_library

digital storage medium cannot be available for broader usage. Therefore, PTCM envisaged creating an online digital repository accessible on the Internet anytime and anywhere.

Several OSS exists for developing an online digital repository, each with different strengths (Morton-Owens, Hansen, and Walls, 2014). Keeping in mind the minimal IT expertise and exposure at the PTCM, the essential features of the DL software that the project investigators were looking for included the following:

- It should be popular and of professional quality

- Easy to install, maintain, and administer

- The workflow to ingest, edit, and review records should be simple

- Allow uploading of multiple files in different file formats

- Local requirements and customizations should be possible at the software configuration level without having to tinker with the source code

- Support interoperability and be compliant with international standards

- It should have dependable community support

- Availability of hosting and commercial support, and service providers

To set up a digital repository that meets the requirements mentioned above, the project consultants experimented with DSpace[10], GSDL[11], and EPrints software to understand the features and workflow supported by each of the software. However, based on the authors' experimentation and experience setting up India's first institutional repository using EPrints software (Jayakanth, 2012), the consultants decided to use EPrints software, which possesses many more features and

---

[10] https://duraspace.org/dspace/
[11] http://www.greenstone.org/

functionalities than the ones listed above. Moreover, EPrints software provides a stable, pragmatic infrastructure that institutions worldwide have utilized to enable their Open Access agendas[12].

PTCM Digital Library

Digital libraries can vary immensely in size and scope. The digital content may be stored locally or accessed remotely via computer networks. These information retrieval systems can exchange information through interoperability and sustainability (Witten, Bainbridge, and Nichols, 2009). The PTCM DL serves as a digital repository for all the training materials and other digital objects created by the centre. It enables all the stakeholders of the training centre to access the training materials 24 X 7. It has the potential to serve as a digital repository for all six postal training centres. Figure 1 shows the screenshot of PTCM DL.

PTC Divisions and Subject categories

Documents uploaded to an EPrints repository can be assigned to a specific division and a subject category implemented for the repository. By default, EPrints software uses a typical university structure to categorize the records into a university division. EPrints uses the library of congress subject headings for the subject-wise categorization of documents for subject categories. EPrints software's default divisions and subject categories are irrelevant to the PTCM repository. Therefore, as shown below, PTCM-specific divisions and subject categories were created.

---

[12] https://www.eprints.org/uk/index.php/flavours/openaccess/

```
divisions:Postal Training Centres:ROOT:0

#ptc:Postal Training Centres:divisions:0

ptcmdu:PTC Madurai:divisions:1

ptcmys:PTC Mysuru:divisions:1

ptcguw:PTC Guwahati:divisions:1

ptcdar:PTC Darbhanga:divisions:1

ptcsah:PTC Saharanpur:divisions:1

ptcvad:PTC Vadodara:divisions:1

NPA:National Postal Academy:divisions:1
```

PTCM Organizational Structure

```
subjects:Trainings:ROOT:0

ptc_mdu:PTC Madurai:subjects:0

pa-sb-sbco:Induction:ptc_mdu:1

mc-dp:Mid-career Development Programme:ptc_mdu:1

phi:Philately:ptc_mdu:1

mip:Marketing of Insurance Products:ptc_mdu:1

mbdp:Marketing of business development
Products:ptc_mdu:1

imp:International Mail Products:ptc_mdu:1

msbp:Marketing of Savings Bank Products:ptc_mdu:1

itmp:IT Modernization Products:ptc_mdu:1

TTP:Traineers' Training Program:ptc_mdu:1

LTC:Leave Travel Concession:ptc_mdu:1

mp:Mail Products:ptc_mdu:1

wel:Welfare:ptc_mdu:1

ss:Soft Skills:ptc_mdu:1

stp:STPS:ptc_mdu:1

frm:Forms:ptc_mdu:1

misc:Micellaneous:ptc_mdu:1
```

PTCM Subject Headings

EPrints being a scalable and flexible software can be customized and configured to serve the
needs of all the other postal training centres (PTCs) in India. It can also serve as the digital
repository for the entire postal department.

Requirement Analysis for PTCM Integrated Library System

Integrated Library Systems (ILS) are software applications that organize, track and make accessible library information resources. ILS Modules compose the basic architecture of these systems and represent some facets of library operations (Breeding, 2008). For example, circulations, an online public access catalogue (OPAC), authority control, inter-library loan, and serials control modules comprise an ILS (Fayen, 2011). Of course, this kind of technology in the service of libraries does not come without a price. As Webber and Peters (2010) point out, "the most important and crucial purchase" for library administrators to consider is the total cost of library operations. Due to the necessity and demand for ILS, price is a significant issue for most libraries, which face shrinking budgets (Donnelly, 2010). It was natural to opt for an OSS solution for the PTCM library with a small collection of about 5000+ book titles.

Moreover, the immediate requirement of the PTCM was to publish their book catalogue online. They intend to use the other features and functionalities of the ILS in a phased manner. Some specific features that the project consultants were looking for in the ILS system were:

- It should be secure and web-based

- It should be scalable and configurable to meet the local requirements

- It should be a well-established, popular, and of professional quality

- Adhere to international standards for cataloguing, interoperability, and data exchange

- Availability of online hosting and community support

Koha, the world's first free, open-source software for library automation or library management system, meets much more than all the above requirements. Its development is steered by a growing community of users collaborating to achieve their technology goals use worldwide. As a result, Koha's feature set continues to evolve and expand to meet the needs of its user base.

Koha OSS is used to implement the PTCM ILS. Only the online public access catalogue (OPAC) module of Koha has been configured. PTCM had its library holdings in Excel format. Koha supports the direct import of records in the MARC format but not in Excel format. Therefore, the Excel file was converted to machine-readable cataloguing (MARC) standards file format. MarcEdit[13], a metadata editing software suite, was used to convert the Excel file to MARC format.

PTCM library collection comprised 5000+ records in an Excel file. The records were imported into the PTCM ILS using the aforementioned conversion and ingestion processes. Figure 2 shows a screenshot PTCM catalogue.

Requirements Analysis for PTCM Content Management System

The consultants developed a web portal to integrate all the services offered by the PTCM. A web portal is a web-based platform that provides for all the stakeholders of an institution/organization. Its users have a single access point to information and services offered through the portal. Today, different software technologies are available to build a web portal. The sophisticated and complex infrastructure surrounding web-based library content has evolved from the singular, independently hosted and managed "library website" into a "library web ecosystem" comprised of multiple platforms, including integrated library systems, institutional repositories, CMSs, and others. Multiple CMS applications, whether open-source (e.g., WordPress[14], Drupal[15]), institutionally supported (e.g., Canvas[16], Blackboard[17]), or library-specific (e.g., Springshare's LibGuides[18]), are employed by most libraries to power the library's website and research guides, as well as to make

---

[13] https://en.wikipedia.org/wiki/MarcEdit
[14] https://wordpress.org/
[15] https://www.drupal.org/about
[16] https://community.canvaslms.com/
[17] https://www.blackboard.com/
[18] https://springshare.com/libguides/cms.html

their collections, in any formats, discoverable and accessible (Macdonald and Burkhardt, 2019). One of the chief advantages of using a CMS is that CMSs support a distributed content model by separating the content from the presentation and giving the content provider an easy to use interface for adding content (Black, 2011).

Some specific features that the project consultants were looking for in the content management system included:

- It should be easy for non-technical people to install, maintain, and administer
- People without any specialized skills should be able to manage content and make design changes
- It should be a well-established, secure, popular, and of professional quality
- It should be database-driven, responsive in design, and easy to manage media.
- Availability of online hosting and community support

Several librarians have discovered the possibility of using the WordPress web platform as a flexible content management system (CMS) with many library-related applications (Calvert, 2014). At its core, WordPress is the simplest, most popular way to create a website—WordPress powers over 33% of all the websites on the Internet. The features and functionalities of WordPress can be enhanced using WordPress plugins. A WordPress theme determines the overall appearance of a website.

Plugins: A WordPress plugin contains a group of functions that one can add to a WordPress website. They can extend functionality or add new features to WordPress websites. There are thousands of WordPress plugins available for free at the official WordPress plugin directory[19].

---

[19] https://wordpress.org/plugins/

Some of the plugins installed on the PTCM website include Elementor[20], The Events Calendar[21], News and Scrolling Widgets[22], and WPForms Lite[23].

Themes: A WordPress Theme is a collection of files that work together to produce a graphical interface with an underlying unifying design for a website. These files are called template files. A Theme modifies how the site is displayed without changing the underlying software. Themes may include customized template files, image files (*.jpg, *.gif), style sheets (*.css), custom Pages, as well as any necessary code files (*.php). Upon installing WordPress, several built-in themes are available. However, none of these themes was suitable for the PTCM portal. After evaluating several freely available themes, the project consultants decided to use the Education Hub theme[24] for the PTCM website.

Education Hub is a clean and elegant WordPress education theme with a simplistic but unique look and feel. This theme is best suited for college, school, university and other educational websites. But this theme is not bound to be fit for educational websites only. It is also flexible enough to be used for business and corporate websites. The theme is highly customizable with layout settings and several other theme options, customized using the Theme Options panel based on the powerful Customiser API. Education Hub helps you quickly and intuitively create professional and appealing educational websites.

Requirements Analysis for Hosting of Project Deliverables

---

[20] https://elementor.com/
[21] https://wordpress.org/plugins/the-events-calendar/
[22] https://wordpress.org/plugins/sp-news-and-widget/
[23] https://wordpress.org/plugins/wpforms-lite/
[24] https://wenthemes.com/item/wordpress-themes/education-hub/

Before commencing the project, the consultants decided to undertake all the developmental work and host the final deliverables on a cloud-based platform. We explored a virtual machine on a cloud platform to host the project deliverables, keeping the cost aspect in mind. The consultants discovered the Baadal platform to host the project deliverables. Baadal is a cloud orchestration and virtualization management software developed at the Indian Institute of Technology Delhi [25]. The Ministry of Human Resources (MHRD), Government of India, developed as National Mission on Education through Information and Communication Technology (NME-ICT) cloud infrastructure for academic purposes[26]. Baadal provides a virtual infrastructure to NME-ICT projects and educational institutions for free hosting and research needs. The only requirement is to connect to the Baadal infrastructure; one needs to be on the NKN, AC/ERNET, or BSNL network. Some of the main features of Baadal are:

- Dynamic resource scheduling and power management

- An integrated workflow system for the request and commissioning of virtual machines.

- Facilities for suspending, resuming, shutdown, power off, power on and specifying resource requirements of virtual machines

- Dynamic resource utilization monitoring.

Two virtual machines (VM) have been configured on the Baadal cloud platform to host the project deliverables. The hardware configurations of both the VMs are as shown below:

Operating System: Ubuntu 16.04 LTS, 64-bit version

RAM: 8 GiB; HDD: 300 GiB; CPU: Intel Single-core

Of the two VMs, one hosts the PTCM DL, and the other hosts the website and the ILS.

---

[25] https://mhrd.gov.in/technology-enabled-learning-0
[26] https://www.mhrd.gov.in/technology-enabled-learning-0

In March 2022, the authors (project consultants) were informed by the Baadal team that the NME-ICT project would come to an end by April 2022. The PTCM has taken a backup of all three servers and exploring the option of hosting the three servers on a different hosting platform.

Lessons Learnt

Project execution was a good learning experience for the consultants. OSS offers a viable and reliable alternative to proprietary software. Initially, the project consultants hosted all the three web applications of the project on a single virtual machine. This decision turned out to be not prudent as we noticed a considerable delay in the turnaround time of the web applications running on the VM. We, therefore, moved the digital library application to an independent VM, and the second VM hosted the website and the ILS. This move improved the turnaround time of all three applications considerably. In Summary, the following are the learnings:

- Identification and evaluation of reliable Open Source Software

- Integration of multiple open-source software

- Host each application on a separate VM for better performance

- Use a stable version of open-source software instead of the latest version, which may not have a stable release yet

Conclusion

The availability of reliable and professional quality open-source software has empowered organizations to create digital infrastructure at an affordable cost without engaging specialized IT professionals and investing in commercial software. The cloud platforms facilitate the creation of virtual machines with the required hardware configuration in hours. These machines are easily

scalable to handle a growing workload and usage. It will also eliminate capital investment in IT infrastructure and human resources.

PTCM envisages an online learning management system (LMS) to impart online training to the postal department staff. To this end, PTCM DL can supplement and complement the LMS.

References

1. Black EP (2011) Selecting a Web Content Management System for an Academic Library Website. Information Technology and Libraries 30(4): 185-189.

2. Breeding M (2008) Open source integrated library systems. Report, Library Technology Report, No. 8, ALA Techsource, Chicago, Illinois.

3. Calvert P (2014) Learning From Libraries That Use WordPress: Content Management System Best Practices and Case Studies. The Electronic Library 32(1): 126-127.

4. Davidson J (2008) OSS spending trends in education. Technology Innovation Management Review, The Open Source Business Resource, Talent First Network, Ottawa: 5 – 7. Available at: http://timreview.ca/sites/default/files/Issue_PDF/august08_osbr.pdf (accessed 03 August 2021).

5. Donnelly FP (2010) Evaluating open source GIS for libraries. Library Hi Tech 28(1): 131-151.

6. Erlich Z and Aviv R (2007) Open source software: strength and weakness. In: Amant, KS and Still, B (eds) Handbook of Research on Open Source Software: Technological, Economic, and Social Perspectives, IGI Global, Hershey, PA, 184-196

7. Jayakanth F, Minj F, Silva U et al. (2008) ePrints@IISc: India's first and fastest growing institutional repository. OCLC Systems & Services: International digital library perspectives 24(1): 59-70. Available at: https://10.1108/10650750810847260 (accessed 03 august 2021).

8. Koch, Stefan (2007) Forward: THE MANY QUESTIONS FROM OPEN SOURCE SOFTWARE… In: Amant KS and Still B (eds) Handbook of Research on Open Source Software: Technological, Economic, and Social Perspectives. Hershey PA: IGI Global, pp. xxvi-xxix.

9. Macdonald C and Burkhardt H (2019) Library-Authored Web Content and the Need for Content Strategy. Information Technology and Libraries 38(3): 8-21.

10. Morton-Owens EG, Hanson KL, and Walls I (2011) Implementing Open-Source Software for Three Core Library Functions: A Stage-by-Stage Comparison. Journal of Electronic Resources in Medical Libraries 8(1): 1-14. Available at: https://10.1080/15424065.2011.551486 (accessed 03 August 2021).

11. Subramanyam R and Xia M (2008) Free/libre open source software development in developing and developed countries: a conceptual framework with an exploratory study. Decision Support Systems 46 (1): 173-186.

12. Webber D and Peters A (2010) Integrated Library Systems: Planning, Selecting, and Implementing. Libraries Unlimited: Santa Barbara, CA.

13. Witten IH, Bainbridge D, and Nichols D (2009) How to Build a Digital Library (2nd ed.). Morgan Kaufman Publishers Inc., San Francisco, USA.

Screenshot 1: PTCM Digital Library Homepage
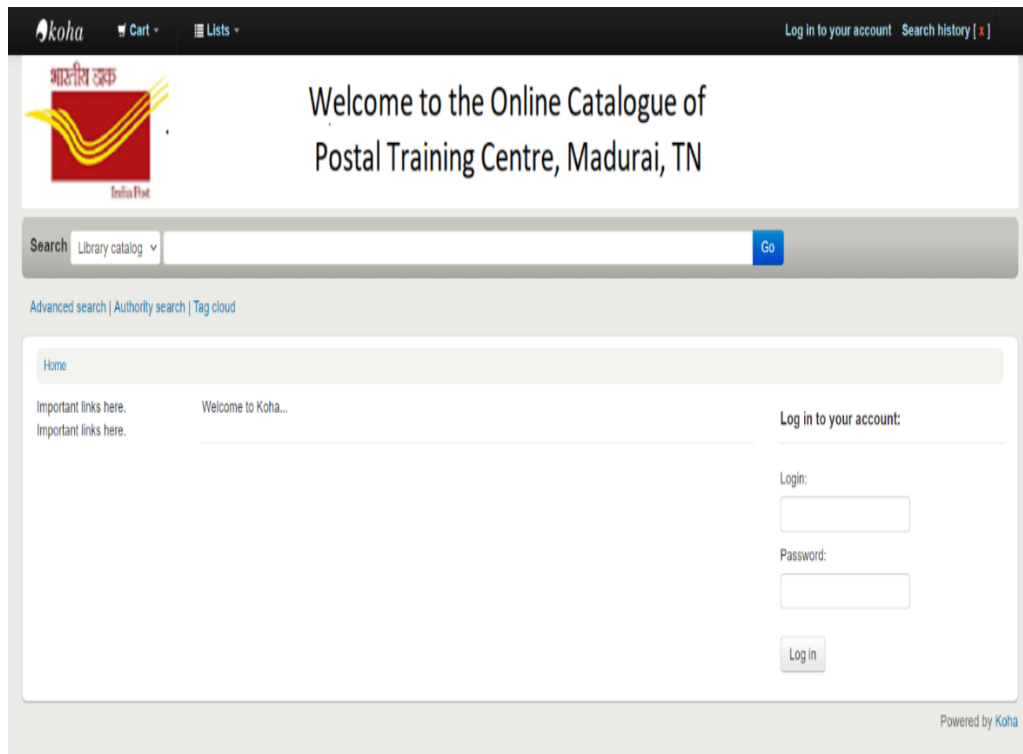
Screenshot2: PTCM Library Catalogue Homepage

Screenshot 3: PTCM Catalogue Sample Search Results