Investigating Interaction Activities in Digital Libraries: The Networked Interactive Digital Books Project

Zaka, Bilal; Maurer, Hermann; and Delilovic, Namik

Abstract: The increase in digitization of conventional books and other forms of print media suggests that digital or e-books are not a passing phenomenon. Present day digital library systems may have transformed the conventional processes into digital environment, however little development work has been done for the progression of traditional library business model in digital domain. Users of digital books want a more open environment where contents from different sources are made available seamlessly. Interoperability at device level, sharing, and a possibility of a convenient and useful annotation system are also general expectations from the standard user prospective. Publishers also need a system with efficient process workflow for the transformation of contents and its online availability. Besides the interoperability, the security and digital rights management is also a matter of concern for the digital content distributors.

We started off our efforts by establishing system to present digital books online with a possibility of simple links creation. This system is now being transformed into a modern form of digital library where possibilities of enhanced interactions are made available to its users. In this paper we report on new user interaction scenarios added to the library system.

Index Terms: Digital Libraries, Interactive learning contents, e-publishing, Online Books

1. INTRODUCTION

Digital libraries are information systems that facilitate storing, indexing, organizing, searching, and viewing of digitized information contents. These contents can be in form of books, documents, images and multimedia objects. The overall operation of digital libraries can be very heterogeneous in nature depending on the type of contents in libraries. Several e-publishing platforms or digital library system are available that provide the means of content digitization and their availability online. These systems are mostly built as monolithic applications, capable of performing library related function in digital domain similar to how things are managed in physical libraries. However, much needs to be done in terms of exploiting the true potentials of digitized media [1]. The digital libraries can add much value to how contents are presented and consumed by its users. Since modern digital libraries can now support processing of different types of media covering numerous subject domains, we can easily have a very large system user community ranging from content publishers, teachers, students, and professionals. A digital library can be seen as a meeting point for people having different background. It is important that a library system provide good means of communications among its users and give opportunities to add value to digitized information contents of system [2]. A similar effort was started by our research group back in 2009, when a quality-controlled knowledge portal was launched having online book publishing system as its auxiliary module. The Austria Forum knowledge portal and Web-books system now has well over one million information objects and approximately 2500 books available online. These systems were evolved around the research done by our group to introduce the needed interactivity and enhancements in information repositories. The development of Austria Forum and web-Books system gave us good insights into challenges and opportunities of online publishing. Almost a year ago based on our developed understanding of the subject we decided to upgrade our online publishing platform into a more robust, standardized system.

In next paper sections we will discuss the approach adapted for building this advance digital library platform and especially focus on explaining different features added to make this digital library stand apart from its successor and counterpart systems.
2. *Networked Interactive Digital Books Library Project*

The new form of networked and interactive digital libraries must be capable of using inherent interactive nature of digital media and innovations being done in online world. The next generation library systems may additionally present contents beyond the scope of traditional books and documents [3]. These contents (audios, videos, simulations and infographics etc.) are exponentially increasing in online communities and can significantly add value to book material [4]. It is also desirable to have convenient mapping of annotations for self-use and peer groups. Enrichment of meta-data by user contributions can also help in finding related information and improve information consumption. The core philosophy at work behind development of the new library system was making the system more open, interoperable, and reusable at both application and content level.

To achieve the goal at application level we opted to develop system using a modular service-oriented architecture approach [5]. The system is comprised of components such as management system portal, library access portal, indexing system, imaging and Optical Character Recognition (OCR), and database. The system architecture allows deployment of these components as separate service instances working seamlessly to create a comprehensive digital library system.

For maximum compatibility with common platform infrastructure, the system is developed using modern programming technology capable of deployment on Windows and Linux based OS environments as well as on major web application hosting clouds. The responsive user interface design allows the system to be used on all major devices and operating systems.

3. *NID Library Access and Workflow*

The library system is accessed using two main entry points to the system. These entry points are:

3.1. *The Library Portal*

This web interface is mainly meant for common system users. The available books in the system are presented to users organized in categories. The system also displays recently added contents to the system in a prominent way. The user can view books in feature rich book viewer and can make use of advance search functionalities. The major user influx is through library portal, so it is only natural to allow the addition of supplementary contents to library through this channel. More details on type of supplementary information and interactions are part of following sections of this paper.

The second access channel to the system is

3.2. *The Management Portal*

This is the starting point in creation of base information of library system. Privileged users can setup library environment parameters such as user accounts, documents types, categories, publisher, and digital Right management (access) parameters. The core information objects, i.e. books are also added and processed for use in the system at this portal.

Following figure depicts a simplified workflow in terms of contents creation in NID library system.
4. User Interactions in NID Library System

The NID library system allows different levels of users to access book contents and use its interaction and management features. The NID users include administrative users and editors that can perform tasks related to system administration, content addition, and management. The other types of users are the ones browsing the library and making use of different options to interact with the system and add additional contents and meta information to the library. This paper and section will mainly discuss the interaction activities available to standard system users.

4.1. User Interface and System Access

The first and foremost element of any web-based information system is its User Interface (UI) design; a good UI influences a smooth User Experience (UX) with the system. In order to accommodate a variety of devices and operating system platforms for NID access, we decided to build the system layout based on an adaptive and responsive UI framework. The web interfaces of the library system are built using Bootstrap [7] framework. The layout of NID library is based on grid design that presents text and image contents. The other basic design elements of the system are navigation menus, buttons, form fields with additional features of JavaScript, and Cascading Style Sheets.

The system can be accessed anonymously; this is the simplest form of system access where any internet user having access to a modern browser on mobile or desktop device can type in the Library Portal URL to view publicly available library contents.

The public availability of library contents is determined using a built-in Digital Rights Management (DRM) system. The rights management is a very critical aspect of online library systems and much attention is needed to incorporate necessary checks for content access and use [8]. The NID DRM feature allows administrative users of library to place a Book/Article of library in either public or restricted access category. There could be different levels of access restriction and it is determined based on registered user access role.

The library contents are presented organized in categories along with a carousel of recently added objects to the library system. The following actions can be performed by users of this category.

- Browsing books details organized in categories, by clicking on respective category icon and further viewing book details by clicking on selected book cover. User also can view author's biography and reviews associated with selected book at book details page.
Performing faceted search and filtering operations on library meta data i.e. book titles, keywords, author, category, locations, date etc. Furthermore, on the result page, a user can easily select to perform a more in-depth search on full book text contents. This is done by simply selecting a radio button of "In-Book" search instead of Metadata search default selection. User can perform very comprehensive search by restricting the probing scenario of metadata/Full-text to complete library, selection of certain categories, author's languages, publishers, dates etc.

The full text search facility made available to NID users supports text extraction from image-based contents as well. The latest version of the most trusted OCR engine Tesseract with the support of 116 languages is incorporated in NID library system. This system component provides the enhanced text indexing facility at the time of book addition to the system. The search query of NID is complemented with operators such as wild cards, AND, OR, and quotes. The NLP preprocessing (stop word removals, stemming) on search index also helps finding the near similar matches.
• Users can also access complete book contents in an IIIF standard compliant browser. The NID book viewer is a customized version of Mirador viewer [9] with addition of advance annotations and content highlighting features. In Book view mode, users can perform full text search within a book and easily locate or jump to related content locations.

The NID book viewer offers the following features to all users:

• Presentation of Book contents as Single Page, Two Page Book View, Scroll, and thumbnail View format
• User can enable or disable slider thumbnail at the bottom to adjust the view area according to the needs

• Presentation of Book contents as Single Page, Two Page Book View, Scroll, and thumbnail View format
• User can enable or disable slider thumbnail at the bottom to adjust the view area according to the needs

• Within page navigation icons and zoom in/out button are also available to users on screen
• Book page navigation is done using page thumbnails carousel, back and forward arrows on main view panel or through bookmarks and table of contents made available to users as index on a side panel.

• User can adjust visual aspects of contents such as rotation, brightness, contrast, grey scaling, etc., according to view type requirements.
• User can also activate the share tool, which allows sharing of library contents for transclusions [10] within library and also out of library systems. The feature allows ease of sharing through links to popular social media platforms.

• Users can access additional contents placed as WIKI articles. They also have viewing access to related discussions and comments of system articles.

By now we have explained some of the features that are available to NID library users accessing the system anonymously. Once a user registers an account with the system, the following additional interaction features are made available:

• Access to annotation management tooltip, where user can add annotations available to public or only for private and group use. Based on the access level user can edit or delete already added annotation to the system.
• User can select the annotation type being added as Public, Private, or Group specific.
• Users can create within book links, external links to other system books and internet resources.
Registered users can generate citation in standard referencing style used in academic writings. Currently system supports Book page reference generation in APA, Vancouver, Harvard, Bib Tex, CSL-JSON.

Registered users can also see and respond to Multiple Choice Questions added to different book pages by privileged users or editors. The marked results summary is made available to the users adding MCQs.

The privileged registered users can also add and edit articles in system using rich media contents. They can upload high resolution images and audio-visual contents to library’s local file store. An article can be created using system’s WYSIWYG rich text editor allowing ease of formatting.

Registered users can add book reviews, participate in discussions, and make comments on library articles.

We believe that such a seamless environment, that conveniently presents the digitzed books along with a possibility of interactions among users, will in turn enrich original information space with added knowledge layers. All the links, annotations, and comments added to the system may present many useful scenarios where this additional knowledge space is used to determine organized learning objects. The system in a way offers a test bed to data scientists for trying out machine learning algorithms to a variety of structured and unstructured data. The possibility of having a system, where users are creating the additional associated book knowledge and editors endorsing and modifying this additional information for quality, makes conventional e-publishing more interesting.

5. CONCLUSION AND FUTURE DIRECTION OF WORK

Overall, the NID library project is designed and anticipated to lead to a paradigm shift: So far, we have seen development of complex e-learning environments that are laboriously filled with content; and books in these systems are sometimes just mentioned as background material. With NID it is the other way around: A book represents a certain context within which many interactions, experiments, media etc. are offered in a form suitable for a particular user group. Some of the features incorporated in NID library introduce the much-needed user centric approach of content use and create opportunities of building intellectual relationship between library managers/editors and their readers.
We see many interesting use cases where ability of book page transclusions through share facility allows creation of customized learning contents. The network and interactivity aspects through use of standardized content presentations and consumption add value in terms of reusability of contents in compatible systems. The extensive inverted indexing and search service API based on web standards also help extending the library meta data beyond system’s boundary. Besides the value addition of user interaction that remains the main topic of this paper, we see a lot more potential of this system and our future research will be focused on the use of advance machine learning algorithms to determine learning objects and automated organization of related information entities within a digital book library system.

REFERENCES


**Bilal Zaka** works at COMSATS University Islamabad as Head of IT services. He has over 19 years of experience in Education, Research, Software Development and IT Project management. Bilal has presented his work at reputed international conferences and published in accredited peer reviewed journals.

**Hermann Maurer** has been Professor of Computer Science at the Graz University of Technology for many years. He received his PhD from the University of Vienna in 1965, and has held professorial positions in Canada, the USA, Germany, Brazil and New Zealand. His research interest was originally in the theory of programming languages, but has since shifted much to multimedia techniques and WWW related issues. He has supervised some 50 PhD and 400 MSc theses, was foundation Dean for Computer Science in Graz, Austria, and is author of over 20 books and 700 other publications. He served as an officer of the Academia Europaea for over ten years and also as member of the board of trustees for this institution.

**Namik Delilovic** is a research associate at ISDS Technical University Graz. His interest revolves around interactive aspects of digital libraries and use of machine learning algorithms to determine deeper relationships among information contents.