

An Extensible E-book Personalization Structure for Open format

Hameed, N.

*Faculty of Science and Engineering Technology,
School of Engineering and Computer Technology
Quest International University Perak
No. 227, Plaza Teh Teng Seng, level 2 Jalan Raja Permaisuri Bainun,
30250 Ipoh, Perak, Malaysia
noorul.hameed@qiup.edu.my*

Abstract

An electronic book is shortly denoted as e-book that consists of file format for personal computers, hardware devices, portable devices and Operating systems. It represents the ideal print book publication structure, classical content formatting, and its metaphor in digital editions. It includes text, image, sound and motion picture as main integral content. Academic print publications are massively available in various digital versions on internet at present. Still, there are numerously required for transferring academic course book, prescribed textbooks, student notes in open, extensible and personalization forms. This paper develops the concrete ideas about e-book in terms of open e-book publication which are consisting the XML-based standards for extensibility and personalization. These features are reported with the release of standards into academic course book publishing. Statistical analysis indicated that the students are being helped by implementing open e-book structure into their curriculum design. Furthermore, it is providing the students to create their personal learning environment on intranet and internet.

Keywords: *E-book, E-learning, XML Technology, HTML Programming, Web*

1. Introduction

E-book is implemented representing structure and content as practiced in printed book. It intended to interpret the digital learning and practices with various computers: personal computer, laptop, dedicated hardware e-reader device. It represented as metaphor for printed book publication which consists of structural and content sections. The structural section contained cover, title, copyright, table of contents, author, acknowledgement, parts/chapters and its pages, references, bibliography, index, back cover pages. The content section combined with multimedia elements.

E-book publications are presented as either in open publication (free of cost) or proprietary format *e.g.*, Portable Document Format (PDF) by Adobe Systems Inc. International Digital Publishing Forum (IDPF) announced open e-book structure by combining XML-based representation in electronic publication (EPUB) format. OEB document is XML which conforms to specification. However, EPUB is not fit to design precise layouts and specialized content formatting such as Comic books.

This paper discusses presentation of an e-book by using extensible personalization structure, which is more appropriate for supporting e-book metaphor and features. The important features of e-books are considered extensibility and personalization. They are identified as major task for open e-book publication. The interoperability of contents is

also considered to provide more flexibility in order to adapt interdisciplinary knowledge in academic.

Henceforth, this paper recommends XML-based e-book document to conform to XHTML 1.0 for specification is issued, and is effectively presented in version 4 HTML browsers. Following sections describes the extensive review about e-book, open format and XML technology.

2. Literature Review

An electronic book (or e-book) presents traditional book structures/functions in digital form. In digital form, it is able to include multimedia/hypermedia services as well as other special features. Present e-book frameworks have shortcomings in their representation format, especially for academic e-publishing. The problems include the e-book personalization by user, the support for e-book publisher and the file format. In terms of representation/personalization, this research presents **extensible e-book personalization (XEBPER)** architecture [1].

For interoperability of content with web browsers and e-readers devices, Hypertext Markup Language (HTML) is popularly used for publishing hypertext on the web. Since beginning of year 2000, HTML 4 is being replaced by powerful Extensible Hypertext Markup Language (XHTML), which formatted the contents better than HTML presentation [2]. As a result, Extensible Markup Language (XML) developed for writing contents with new elements or additional attribute elements. Precisely, XHTML insisted the separation of presentation of document from the structure of document's information.

IDPF's OEB package [3] consisted with package identity(a unique identifier for the OEB publication as a whole), metadata (publication metadata as title, author, publisher, etc.), manifest(a list of files (documents, images, style sheets, *etc.*) that make up the publication, spine (an arrangement of documents providing a linear reading order), tours (a set of alternate reading sequences through the publication, such as selective views for various reading purposes, reader expertise levels, *etc.*), and guide(a set of references to fundamental structural features of the publication, such as table of contents, foreword and bibliography).

On selling point of e-book, many publishers were redesigned their website to direct sell to consumers who gets more benefits out of it. According to recent survey, publishers who made successful in promoting e-books to direct consumers were Amazon.com, Penguin Random House, Simon & Schuster which offered full catalog for sale, Houghton Mifflin Harcourt, and Hachette who does not sell physical books directly to consumers but it sells its e-books, and HarperCollins.

However, EPUB is current open standards for publishing content as free, open e-book standard [4]. It included structure for format with reflowable content at the text display which can be optimized for eBook Readers such as Sony® Reader, Barnes and Noble nook™, Apple iBooks®, Lexcycle Stanza and others. In e-book consumer market research [5], Amazon earned high profit in e-book sales, and Barnes & Nobles distributed and synchronized content through multiple reading devices as e-libraries.

In addition, Lombardi [6] had described about the term e-book as electronic text for book, magazine, journal, poem, cookbook, travel guide, *etc.* This recommended about the features of print publication process with new e-book publishing. Furthermore Cannon and Watson [7] had referred as simple translation of printed-book contents into e-book format.

3. Research Methodology

The software model which is being focused customized markup structure, *i.e.*, XML-based structure and content presentation with formatting styles. It was built with three phases.

At first phase, XEBPER structure is proposed to represent XML-based e-book structure and content. It is considered to be more powerful and efficient for representing content with markup structures. Later it exported into various type of browsers or devices.

During the second phase, XEBPER specification combines XHTML, XML/XSL and script programs to produce e-book on simple HTML browsers. XML document is linked with XML-based schema structure to document verification before it is being displayed onto browser programs. XSL/XSLT transformation is used to present with styles for XEBPER structure. Because of interoperability of XEBPER with XML technology, the structure can be displayed onto any browser for reading.

At final phase, XEBPER publication is produced linking with author's original contents. XML editor programs are being applied to simply type paragraph inside the metaphor objects, hence it referred as functional objects (media, content control, device control and book update).

Internet Explorer 9 is the first version for XHTML as notably. Thus publisher with XEBPER publication structure will be published a high quality e-book with open format which included all necessary documents inside a computer user-defined file or folder. This file or folder can be applied in later for packaging with compression, secured and distribution over internet.

4. XEBPER Publication Structure

XEBPER Framework is portrayed in using new open e-book architecture as illustrated in Figure 1.

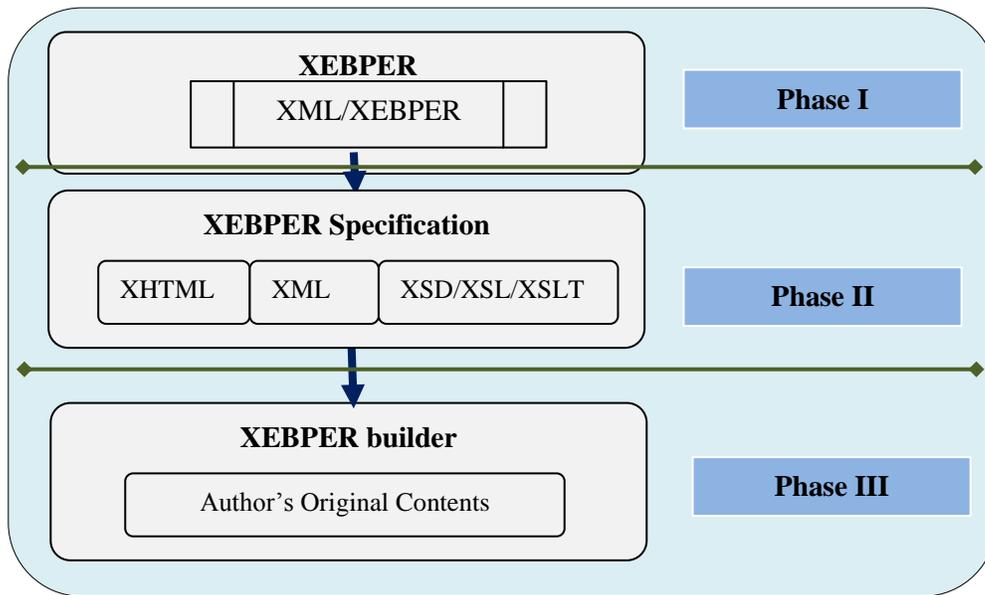


Figure 1. Academic Open E-Book Personalization Architecture

XEBER present the structure and data in complete form. Structure is emphasized in a unique characteristic for each e-book. Following section explore the overall e-book features that are defined with three phases.

4.1. Structure/Content Representation

This phase is including e-book standard based on XML technology. The structure is being represented with extensible and personalization features, known as “XEBPER Publication Structure.”

XEBPER structure is divided into two objects: structural and functional. The package, visual, and metaphor objects are called structural objects and are deemed to be objects of a higher-level as they consist of the e-book’s basic information, book metaphor operators and its package description. The remaining four lower-level objects, *i.e.*, media, content control, device control, and book update objects are called functional objects. Functional objects are applied as integral parts for relating author’s content in metaphor objects. The links between functional and metaphor objects are provided in Figure 2.

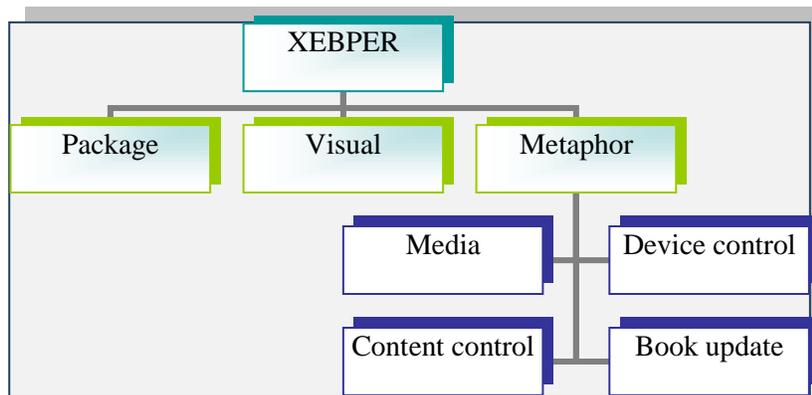


Figure 2. Basic Connection between Structural and Functional Objects

4.1.1. Package: Fundamental elements that are needed to describe the package object are shown in the following Extended Backus-Naur Form or Backus Normal Form EBNF notations:

- <package> → (‘<initialize>’|‘<software>’|‘<publish>’|‘<property>’)?
- <initialize> → (‘<import>’|‘<package>’)?
- <package> → (‘<metadata>’)?
- <metadata> → (‘<schema>’|‘<stylesheet>’|‘<script>’)?
- <software> → (‘<builder>’|‘<reader>’)?
- <publish> → (‘<publication>’|‘<library>’)?
- <library> → (‘<academic>’|‘<access>’)?
- <property> → (‘<digitization>’|‘<archive>’)?
- <digitization> → (‘<numbering>’|‘<marketing>’)?
- <archive> → (‘<compression>’)?

4.1.2. Visual: This object elements contribute towards the book appearance, styles, and screen configurations. Basic visual object elements are represented in EBNF notations as follows:

$\langle \text{visual} \rangle \rightarrow (\langle \text{layout} \rangle \mid \langle \text{font} \rangle \mid \langle \text{themes} \rangle) ?$

4.1.3. Metaphor: This object is divided into three main sections: (1) front, (2) main content, and (3) back sections. The basic ebook metaphor elements are presented in the following EBNF notations:

$\langle \text{metaphor} \rangle \rightarrow (\langle \text{front} \rangle \mid \langle \text{main} \rangle \mid \langle \text{back} \rangle) ?$

$\langle \text{front} \rangle \rightarrow (\langle \text{cover} \rangle \mid \langle \text{title} \rangle \mid \langle \text{verso} \rangle \mid \langle \text{author} \rangle \mid \langle \text{toc} \rangle \mid \langle \text{preface} \rangle) ?$

$\langle \text{main} \rangle \rightarrow (\langle \text{chapter} \rangle) ^*$

$\langle \text{chapter} \rangle \rightarrow (\langle \text{singleissue} \rangle) ^*$

$\langle \text{singleissue} \rangle \rightarrow (\langle \text{page} \rangle) ?$

$\langle \text{page} \rangle \rightarrow (\langle \text{header} \rangle \mid \langle \text{branch} \rangle \mid \langle \text{footer} \rangle) ?$

$\langle \text{back} \rangle \rightarrow (\langle \text{application} \rangle \mid \langle \text{exercise} \rangle \mid \langle \text{bibliography} \rangle \mid \langle \text{index} \rangle) ?$

4.1.4. Media: This object elements allow some definition about properties such as text, image, sound and motion. The basic media object definition is shown in the following EBNF notation:

$\langle \text{media} \rangle \rightarrow (\langle \text{text} \rangle \mid \langle \text{image} \rangle \mid \langle \text{audio} \rangle \mid \langle \text{video} \rangle \mid \langle \text{hyperlink} \rangle \mid \langle \text{animation} \rangle) ?$

4.1.5. Content Control: Content control objects control the ebook environment by setting the behavioral and functional attributes. The basic content control elements are as shown in the following EBNF notation:

$\langle \text{content} \rangle \rightarrow (\langle \text{application} \rangle \mid \langle \text{formulas} \rangle \mid \langle \text{switch} \rangle) ^*$

Application tag refers the dynamic presentation with content as Macromedia Flash, Microsoft ActiveX objects, source codes for computer languages, *etc.*

Formulas are advanced editing area to edit mathematical and statistical equations onto e-book. It is will be embedded in future publications.

Switch statement is to analyze several tags with its Ids whether the element is being matched for any given id as input parameter during run-time. It is also applied in future publications.

4.1.6. Device Control: Device object elements describe human-computer interaction that includes personalized messages, comments, user-defined annotations and bookmarks. Elements for device objects are grouped in the following EBNF notations.

$\langle \text{device} \rangle \rightarrow (\langle \text{highlight} \rangle \mid \langle \text{bulletin} \rangle \mid \langle \text{bookmark} \rangle \mid \langle \text{dialogue} \rangle) ^*$

Moreover, important e-book features such notes, student assignments, lecture notes also included in device control objects. They will not be appeared in full-text form, but it is existed in small/medium sized icons onto related margin locations in a page.

4.1.7. Book Update: The book update object describes the manipulation of ebook elements and attributes by users at design and execution times. Elements of the book update object are shown in the following EBNF notations.

$$\langle \text{book} \rangle \rightarrow (\langle \text{fileattachment} \rangle \mid \langle \text{websearch} \rangle)^*$$

File attachment tag is included to link additional and computer resources. It might be files, source codes, *etc.*

Web search tag is attach list of recent websites and update locations for upgrade e-book publication in terms of progress, display, styles and formatting.

4.2. XEBPER Specification

This phase is including e-book standard based on XML technology. The structure is being represented with extensible and personalization features, known as “XEBPER Publication Structure.” XML technologies applied for effective use of content with target format or platforms. For instance, XEBPER structure is exported in Internet Explorer (version 9 or higher) browsers for presentation and formatting its functional objects. Verification and validation process is done with XML Schema Definition (XSD) as by default step.

4.3. XEBPER Builder

E-book publication is generated by run-time scripting programs, such as JavaScript. It is also possible to interpret with Document Object Model (DOM) and Simple XML API (SAX) models. In every instance, markup structure is being read in separate process. So that it is more powerful for academic publication which is used annotation, notes, class notes, chapters, pages, sections, text formats and functionality in customized method.

One example might be JavaScript application with our proposed XML/XEBPER structures. Script interprets page layouts in linear and personalized orders.

5. Outputs

XEBPER structure is included in XML Language that supports contents in .xml format. Header is declared by `<?xml version="1.0"?>`.

Then the document root follows. XEBPER structure is implemented to transform structure and content with formatting styles. The following are just a few examples of the EBNF notation:

$$\langle \text{xebper} \rangle \rightarrow (\langle \text{ebook} \rangle)?$$
$$\langle \text{ebook} \rangle \rightarrow (\langle \text{package} \rangle \mid \langle \text{visual} \rangle \mid \langle \text{metaphor} \rangle)?$$
$$\langle \text{metaphor} \rangle \rightarrow (\langle \text{functional} \rangle)?$$
$$\langle \text{functional} \rangle \rightarrow (\langle \text{media} \rangle \mid \langle \text{content} \rangle \mid \langle \text{device} \rangle \mid \langle \text{book update} \rangle)^*$$

6. Discussion

In a preliminary study, e-book using XEBPER structure, were published at class room level. It conducted with 7 under-graduate students from Bachelor of Computer Science

(Hons) at Quest International University. E-book entitled “**BCS 2110 Visual Programming II**” containing of 183 pages. Digital book versions presented to all students who have undertaken the courses in two semesters. Important file formats used as Microsoft Word and PDF formats. Digital edition of course books which are being published in DOCX and PDF formats were captured electronically through screenshots as shown in Figure 3 and 4.

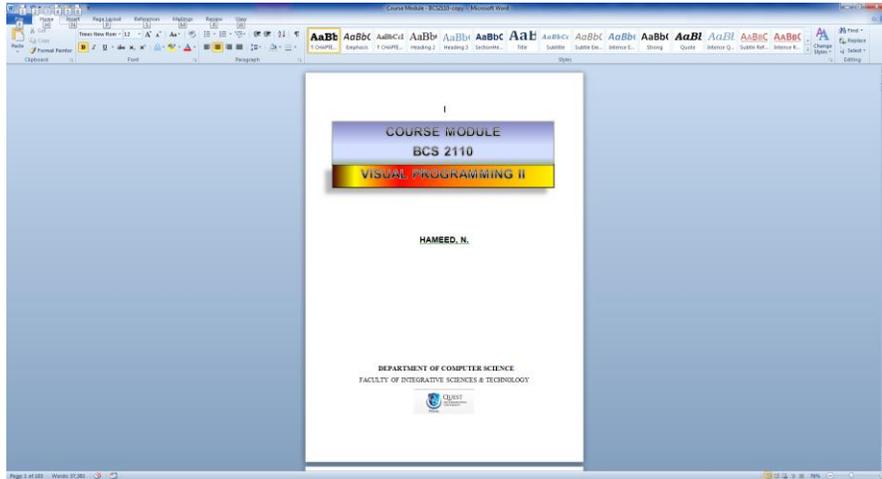


Figure 3. Sample e-book that Appeared onto Microsoft Word 2010

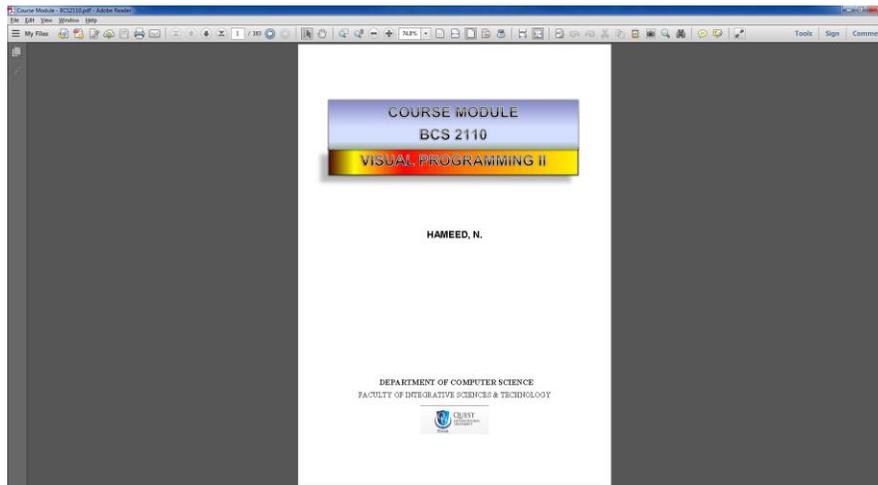


Figure 4. Sample e-book that Appeared onto Adobe Acrobat Reader (version 11)

In this paper, course book were revised into XML-based format, XEBPER structure which conformed open format. Any browser that supported with rendering system (*e.g.*, PDF, XHTML, XML, or TEXT) can render with style sheet programs that were applied in document. This included XSLT technology: XSL and XSL-FO (XSL Formatting Object).

This paper introduced sample experiment for e-learning context. According to [8], e-book readers practiced to publish contents with browser program in portable HTML format. Custom parts of contents integrated chapters/units, pages, multimedia elements

etc. The definition of terms, pictures, annotations and notes referred in an individual element. The e-textbook is transformed with formatting in terms of representing properties of structure and chapters while Internet Explorer (version 9) interprets the XML messages as shown in Figure 5.

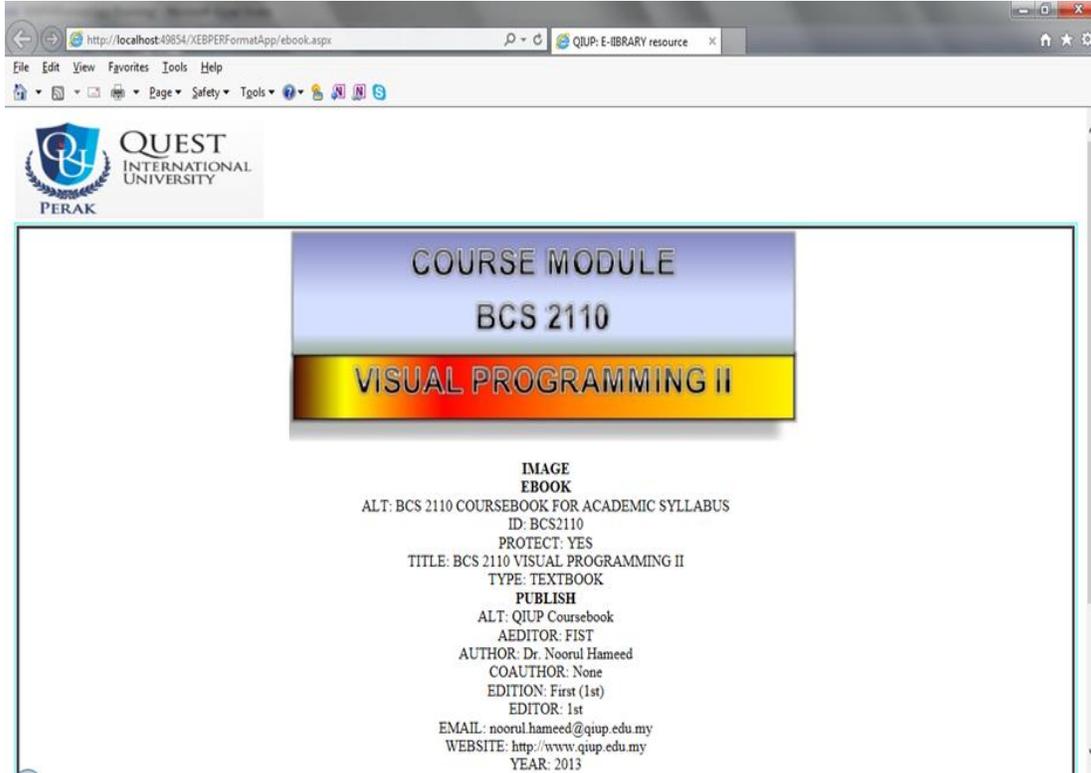


Figure 5. XEBPER Publication that Appeared onto Internet Explorer 9

Many outputs are possible for making effective content presentation with XEBPER structure. XML language is utilized to present XEBPER structure into XHTML output. For instance, source program provided for **ebook.xml** in order to allocate a single chapter element and section that contained text or images, as shown in below code snippet.

```
<?xml version="1.0" encoding="utf-8"?>
<?xml-stylesheet type="text/xsl" href="coursebook.xsl"?>
<XEBPER>
  <EBOOK ID="BCS2110" TITLE="BCS 2110 VISUAL PROGRAMMING II "
  TYPE="EBOOK">
    <PACKAGE/>
    <VISUAL/>
    <METAPHOR/>
```

```
<FRONT />
<MAIN>
  <CHAPTER>
    <SINGLEISSUE>
      <PAGE>
        <BODY>
          <SECTION>
            <MEDIA>
              <TEXT/>
            </MEDIA>
          </SECTION>
        </BODY>
      </PAGE>
    </SINGLEISSUE>
  </CHAPTER>
</MAIN>
<BACK />
</METAPHOR>
</EBOOK>
</XEBPER>
```

Moreover, XSL-FO implemented with set of layouts and stored in “**coursebook.xml**”. Likewise, the cover page (as in Figure 5) produced with below code snippet.

```
<fo:layout-master-set>
  <!-- Define the cover page. -->
  <fo:simple-page-master master-name="cover" page-height="8.5in" page-
width="5.5in" margin-top="1.0in" margin-bottom="0.25in" margin-left="0.25in"
margin-right="0.25in">
    <fo:region-body margin-top="0.25in" margin-bottom="0.25in"/>
  </fo:simple-page-master>
</fo:layout-master-set>
<fo:page-sequence master-reference="cover">
</fo:page-sequence>
```

Therefore, this study revised with XML-based features. So it resulted better for academic course book. It was found that students were searched contents in personalized layout and obtained benefits while implementing open format.

7. Summary

This study discussed the utility of new e-book structure for academic course book design. It analyzed portable or proprietary formats with XEBPER structure to evaluate structural and content design guidelines. This led us to formalize the various e-book design guidelines. Using for XML-based e-book presentation and formatting, e-book personalization, interoperability and functionality were enhanced into high quality publication. In future, the experimental study might be conducted to observe and justify the research with the taxonomy of e-books and survey questionnaire.

Acknowledgements

This paper considered as major outcome of my research which already contributed to field of e-book design and development. I sincerely thank our department and staff to their great and extended support.

References

- [1] N. Hameed. An E-Book Building Methodology Using An Extensible Personalization Structure For Operationalizing E-Book Interface Metaphors. PhD Thesis. School of Computer Sciences, Universiti Sains Malaysia, Penang, Malaysia, (2010).
- [2] W3C. All standards and drafts. (2010). World Wide Web Consortium. www.w3.org/TR/#Recommendations. Accessed 26 April 2010
- [3] OeBF. Open Publication Structure (OPS) 2.0.1 v1.0.1, (2010). www.idpf.org/epub/20/spec/OPS_2.0.1_draft.htm
- [4] IDPF. EPUB Publications 3.0.1, (2010). www.idpf.org/epub/301/spec/epub-publications.html. Accessed 26 June 2014
- [5] P. Dewan, Are books becoming extinct in academic libraries? *New Library World*, 113, 1/2, 27 – 37, (2012)
- [6] E. Lombardi. Classics: E-books or Books, *Classic Literature*, (2007). <http://classiclit.about.com/library/weekly/aa090400a.htm>
- [7] E. Cannon, and B. Watson. Taking e-books for a test drive, *Computers in Libraries*, (2001), Vol. 21, 3, pp-24-7.
- [8] S. Marinai, A. Anzivino, and M. Spampani, Towards a faithful visualization of historic books on e-book readers. *Proceedings of the 2011 Workshop on Historical Document Imaging and Processing (HIP' 11)*. (2011), ACM, New York, NY, USA, 112-119.

Authors



Dr. Noorul Hameed Mohamed Meerasa, Lecturer at QIUP, and Malaysia since 2014. My research interest included e-learning, e-book, open standards, web programming and software engineering. Since 2004 I have worked in e-book research. It includes research articles, conference papers, journals and talks etc. Goal continues to develop a open standard for student's course book in e-book format.