Towards Commercial eBook Production in Small Publishing Houses

Eduardo Rodríguez López, Ángeles Saavedra Places, José Antonio Cotelo Lema,
Oscar Pedreira Fernández and Nieves Rodríguez Brisaboa

Databases Laboratory, Computer Sciences Department, University of A Coruña, 15071 A Coruña, Spain

Abstract: In this paper we present the rationale behind e|ditor, a tool for online edition of digital contents aimed to small publishing houses. Its goal is to provide them with a user-friendly tool for creating (autonomously and at a commercially viable cost) a wide variety of device-independent enhanced e-books, including pocketbooks, pop-up e-books, school textbooks and extracurricular activities e-books.

E|ditor is content-centric, focused in the content creation process. The standard export formats currently supported are EPUB (specially indicated for pocketbooks), SCORM (for LMSs) and HTML5 (for more powerful contents, aiming specially to school textbooks). This provides publishing houses with an appropriate spectrum of formats to cover a wide range of e-book types.

This paper describes the main requirements of e|ditor and the design decisions taken to guarantee platform-independence and content reusability whereas providing a general purpose enhanced e-book creation tool.

1 INTRODUCTION

Traditional book publishing houses face nowadays several challenges to their survival. The adaptation to the current digital era, where contents are increasingly consumed in digital form, is one (if not the biggest) of them. The technological evolution of e-readers and the already widespread use of smartphones and tablets are changing at a fast pace the way users consume digital contents, and this is producing drastic changes in the traditional publishing markets.

The pocketbooks market has been suffering an important shift from printed books to e-books, in the same way as it has already happened to music distribution from physical supports to online/electronic support (e.g. iTunes or Spotify).

Publishing houses specialized in pop-up books have found an strong competitor in digital content producers, as online multimedia interactive contents are more attractive to catch the child attention and are easily placed right at home through Internet.

The school and educational book publishers are also experiencing in the last years how their markets move towards the use of new technologies. The arrival of native-digital children (Prensky, 2001) to the education systems and the pressure of public educative administration initiatives, like the ones fostered by the European Schoolnet (EUN) (www.eun.org) or the *UK Technology Enhanced Learning Program* (http://tel.ioe.ac.uk/), are forcing them to provide learning contents also in digital form, and to fully design those taking into account all the possibilities of digital technologies. This transition faces two main difficulties. The first is that the educative community is still learning how teaching resources, methodologies and relationships should be better adapted to properly incorporate digital technologies into the educative process (Vrasidas and Glass, 2005) (Underwood, et al., 2010) (Cachia, et al., 2010). The second is the maze of continuously evolving content formats and device platforms.

The launch of the Amazon online shop in the nineties, the appearance of Amazon Kindle e-reader in 2007 and the recent moves of the main device-platform providers, taking advantage of their dominant position to place their online shops (e.g. Google Play or Apple iTunes) as the default source of digital contents in their platforms, have made evident to publishers the need to adapt their markets to the distribution of digital contents before it becomes too late.

The need to ensure that small publishing houses succeed in adapting to this big market change is not a matter of business diversity. It is a matter of

cultural diversity survival, as small publishing houses are the ones more directly linked with minority languages and cultures dissemination, and with promoting independent thinking and points of view (Carroll, 1985). But small publishing houses have no power over the "digital ecosystem", so they have to ensure that their contents adapt to the functionalities provided by the most common device platforms, that their formats adapt to the ones supported by their most common tools, and that their distribution channels are easily accessible on them. Moreover, they will have to adapt their tools and their existing contents to future changes on all those platforms. But small publishing houses suffer the lack of tools adapted to this new scenery. Big publishing houses have resources to develop self-tailored tools for that purpose, but small publishing houses have not.

In this article we introduce e|ditor, a platform designed to provide publishing houses with a tool to produce format-independent and device-independent contents. It has been developed in collaboration with small publishing houses whose markets range from school and educational books, to pop-up books and pocketbooks. We describe the requirements identified and the approaches followed to face them, as well as the rationale behind those decisions. We also give a general view of the resulting platform.

2 BACKGROUND

Although there are tools available to create digital contents, small publishers suffer the lack of tools supporting the full creation process and aiming to a wide range of book types.

Almost all the tools available for pocketbook creation are meant for converting existing digital contents to EPUB, an open e-book standard format designed for reflowable content that makes use of XHTML and CSS and is extensively supported by current e-reader devices and software. One example is the open source tool *Calibre* (Goyal, 2006), which allows the conversion of a wide range of document formats to the main e-book formats (EPUB, MOBI, pdb, etc.). But just translating conventional contents to EPUB is not the best way to take full advantage of digital contents. SIGIL (Markovic, 2009) is one of the few that supports e-book creation and can export them to EPUB. But this desktop application is designed only for pocketbook creation, does not allow interactive content and has no support for content reuse or collaborative edition.

Last versions of Adobe InDesign and QuarkXPress support exporting to e-book formats. But they do not create enhanced e-books. Instead, they just convert plain paper books to plain e-books.

Apple offers *iBooks Author* for e-books creation, but it is provided only to sell them through iTunes Store, and cannot be exported outside it.

EXeLearning (Univ. of Auckland et al., 2006) is an open source authoring application to assist teachers in web content publishing. It is a desktop application and allows exporting to SCORM format and to self-contained XHTML web pages.

With regard to tools for creating digital educative contents, the efforts from educative administrations have led to the creation of tools like *Cuadernia*, *Constructor*, *JClick*, *Hot Potatoes* or *Ardora*. *Cuadernia* (Junta de Castilla la Mancha, 2008) and *Constructor* (Junta de Extremadura, 2005) allow the creation of several types of interactive activities and their packaging following SCORM (a collection of XML based standards used in web based e-learning). *JClick* (Generalitat de Catalunya, 1992) provides a set of computer applications to develop different types of educational activities like puzzles, associations, text exercises or crosswords. *Hot Potatoes* (University of Victoria, 1998) is a suite of six applications to create interactive jumbled-sentence, crossword, short-answer, multiple-choice, gap-fill and matching/ordering exercises. It can export to SCORM format or as HTML. *Ardora* (Bouzán Matanza, 2008) is an interactive exercises creator. Its contents use HTML5, CSS3 and JavaScript, and can be exported to SCORM.

None of these applications support collaborative work or "book collections" (where the book design is common and uniform through the entire e-book collection). All of them are meant for independent authors that want to create a specific type of e-book content and rarely (only SIGIL) support all the e-book process. They are usually format-centric designed, more centred in translating contents to a specific format than in the content design. They are neither meant for content reuse nor to ease future content republishing in new formats or styles. They are not web applications (centralizing the content storage) and have to be deployed in the user PC, what restricts the platforms that can be used for edition (usually only Windows, Linux and Mac). Even worst, only a few of them generate device independent contents. *Cuadernia* and *Constructor* create Flash contents, and *JClick* generates Java Applets, limiting the platforms on which they can be displayed. All these reasons make them unsuitable for small publishing houses.

3 GOALS

The e|ditor project focuses on the development of an e-book editor for small publishing houses. The key needs identified are:

- *Full e-Book Creation Support*. It must support the whole e-book creation workflow, allow mixing all types of contents and seamlessly integrate them in the same e-book.
- *Collaborative Edition*. It must support collaborative content creation, centralize e-book resources, user management and maintenance, and allow both deploying it internally at the publisher IT infrastructure and providing it externally as SaaS (Software as a Service).
- *Support to "book collections"*. It must support the creation of book collections, with the same formats and structures. This is typical of textbooks, where the books for the different subjects follow the same structure and aesthetic.
- *Flexibility*. It must allow using the appropriate structures for the types of books produced by a typical small publishing house.
- *Content-centric*. Contents must be stored using a semantic/structured organization, not in any final e-book format. This pursues two goals:
  - *Simplify content revision and reuse*.
  - *Format independency*. It makes easier to use "write once, export many" approaches. This will allow publishers to easily and cheaply adapt their e-book catalogue to new formats.
- *Platform Independency*. e-Books must be readable in as many platforms and devices as possible. For this goal is crucial the intensive use of official and de-facto standards.
- *LMS Integration*. Support for LMS integration through the SCORM standard.

These goals have guided the design of e|ditor. The following section describes its architecture and the workflow supported by it.

4 ARCHITECTURE

E|ditor has been designed as a web application. This allows its use either as SaaS, hosted by an IT provider, or as an Intranet server application, hosted by the publishing house. In this client-server architecture the centralized server application is the host and the user's web browsers are the clients. Access control is provided through user/password identification. E|ditor is composed of four modules: authentication, edition, layout specification and export (see Figure 1). This keeps tasks independency and allows assigning different task to different user.
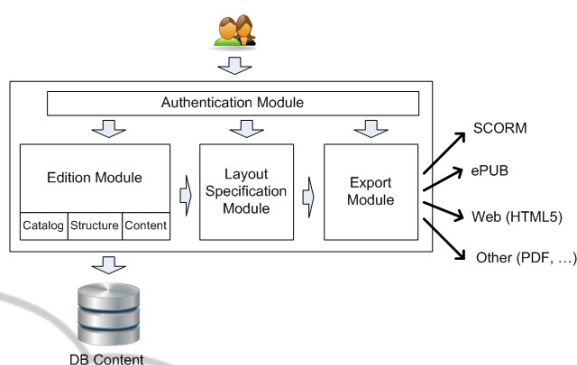

Figure 1: E|ditor functional architecture.

All the data are stored in a database. They are stored independently of its final layout and export format, so that new layouts and export formats can be used later with the same e-book content.

4.1 Edition Module

The edition module has three sub-modules: catalogue, structure definition and content edition.

4.1.1 Catalogue

This sub-module handles the e-book metadata. It follows the Dublin Core Metadata Element Set (DCMES) (Dublin Core Metadata Initiative, 1999), endorsed by standards like ISO 15836:2009, ANSI/NISO Z39.85-2007 or IETF RFC 5013.

4.1.2 Structure Definition

This sub-module predefines how the contents will be organized in the e-book and the type of contents valid at each level.

For example, in Figure 2 a *text* section is declared, having three sub-elements: *introduction*, *subject* and *conclusions*. *Introduction* can have only textual content whereas *subject* and *conclusions* can have both text and interactive exercises (Section 0 shows the supported types).

It is possible to clone the structure specification to reuse it within an e-book collection.

This information is used by the edition module (see Section 0) to keep a uniform structure within the e-book and by the layout specification module (see Section 0) to define the e-book layout.
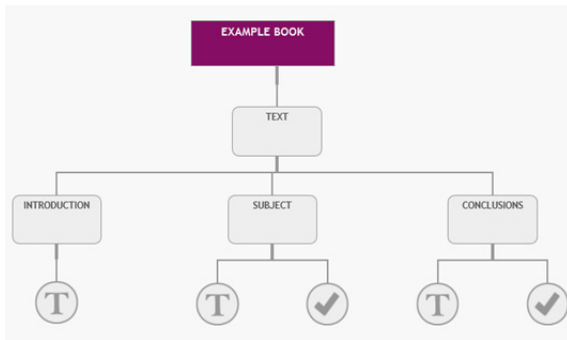
Figure 2: Structure specification.

Figure 3: Content editor.

4.1.3 Content Edition

This sub-module handles all the tasks related with e-book content edition. It manages the e-book index and the content editor.

Editors must define first the e-book index. For each index element they set its content type. The e-book structure definition is used at this point to ensure that a valid content type is chosen. After an index element is defined its content can be introduced through the content editor. The content types that can be associated to an index element are:

- *Conceptual Content*. The core textual content in a book. This is usually the only content type in pocketbooks, and is widely used in textbooks.
  A WYSIWYM HTML editor is used that annotates the content parts semantically. The layout specification module uses this to format them properly. For example, if we tag content parts as poem, definition, highlighted note, footnote or bibliography, specific styling and layouts can be applied to each one.
- *Interactive Exercises*. Exercises to be made by the users/students. The answers are corrected automatically, giving them immediate feedback. They are used in pop-up and school textbooks.
- *Scenery*. Sceneries are sections composed by animated/interactive content. They are intended

to provide some light interactive animations, mainly in pop-up books.

- *Links*. A links block is meant to present a collection of hyperlinks, either to other parts of the book or to external resources.
- *Glossary*. Intended to group a glossary of term definitions. It can be used as an interactive glossary, allowing users to look the definitions right at the points where the terms are used.

Interactive exercises are used in school textbooks, complementary activities and educative books. The following types are supported by e|ditor:

- *Multiple-choice Tests*. Editors write one right answer and three wrong. For each one, they can write an explanation to be shown if that answer is chosen. The answers are scrambled each time.
- *True/false Questions*. Editors write the question, set the right answer and optionally add an explanation to be shown once it is answered.
- *Short answer questions*. They can be answered with a word or a well defined short text. Editors write the question and up to three possible right answers. The user answer must match exactly, with no variations in uppercases or diacritical signs. This is done to reinforce proper writing.
- Matching/ordering exercises. The typical exercise where the user has to match pairs of words/sentences or order a list of words.
- *Crossword Puzzles*. E|ditor allows loading JCross files created with *Hot Potatoes*.
- *Gap-fill Exercises*. Editors write the full exercise text, enclosing with # characters the text to hide. The user must type the missing text, which is corrected using exact text comparison.
- *Free-style Exercises*. They allow writing other types of exercises. Editors just write the exercise, which is not automatically corrected.

4.2 Layout Specification

This module sets how each content type is translated to an export format. This is directly linked with the structure definition and dependent on the export formats, as they impose limitations on how the contents can be shown. Several layouts (for the same or different export formats) can be defined for the same e-book. E|ditor provides layouts adapted to different export formats. They can be directly used or be cloned and modified to create new layouts.

This is the only module intended for specialized users, as layout edition requires CSS knowledge. Once defined a layout, non-specialized users can use it to visualize and export the book content.

4.3 Export Formats

E|ditor achieves platform-independence by allowing (re)exporting the e-books to several widely supported formats. Currently these formats are:

- *EPUB 2.0*. The most widely supported format in e-readers. It defines a file format (.epub) to store the contents and the e-book metadata/catalogue. Multimedia contents can not be exported to it. This is the ideal format for pocketbooks.
- *SCORM*. A standard supported by most LMSs. It provides protocols for feedback to the LMS, a transfer file format and support for adapting how users navigate through the contents. It is especially appropriate for online education.
- *Web (HTML5)*. Smart-phones, tablets and PCs provide e-book viewers with different standards support. But all of them are shipped with web browsers with HTML5 support for multimedia and interactive contents, scalable vector graphics (SVG) and mathematical formulas. Interactive whiteboards (IWB) are usually PC-based and reproduce the same types of contents. It is ideal for pop-up, school and educational e-books, as it gives more room for interactivity.

5 E|DITOR WORKFLOW

The e-book creation sequence in e|ditor is basically the following. First, the e-book structure is defined. Then, the content index is added. Afterwards, the editors write the contents. Later, additional resources are added. Finally, the layouts for the different export formats are defined. Then, the user can choose a layout and export the e-book to the format specified by it.

Anyway, e|ditor allows flexible workflows. Users can define the structure and then change it after writing some content. Layouts can be edited at any moment. And users can define only some portions of the book index, fill their content, and then add new index parts. This allows publishers to better adapt the workflow to their needs.

6 EVALUATION

E|ditor has been already in use by some of the small publishing houses involved in the project.

Publisher *Baía Edicións* has used it to create its student textbook for 2º and 4º grade social sciences and the 2º grade social sciences teacher textbook,

available in web for those schools using them as classroom textbooks. E|ditor helped to guarantee a uniform structure both within a book and within the book collection, and to reuse parts of the 2º grade social sciences student book in the teacher version.

Figure 4: *Baía Edicións*, classroom textbook.

It took just 60 person-hours to create the 2º grade social sciences student e-book using e|ditor. In contrast, the student e-book for 1st grade social sciences was created in an ad-hoc manner with the support of some specifically developed tools, taking more than 500 person-hours (tools development apart) to do it. This illustrates the improvements in autonomy and costs that e|ditor achieved.

Publisher *Editorial Galaxia* has used e|ditor to create its first e-book catalogue, with 36 narrative e-books in EPUB format. It also allowed them to reuse contents to create demo-editions (with parts of the e-books) for marketing purposes.
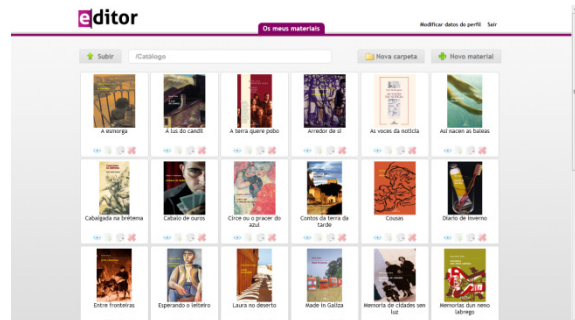

Figure 5: *Editorial Galaxia*, e-book catalogue.

Publisher *Galebook*, a small start-up specialized in digital contents, is using e|ditor for creating their extracurricular activities e-books and plans to use it also for classic-tales pop-up e-books.

E|ditor is also used by *ASPG,* an association with more than 1,000 affiliates and more than 30 years promoting educative and pedagogical innovation, teachers training and didactic contents publishing. After years using *EXeLearning*, *Hot Potatoes* and

similar tools for creating their didactic contents, it is now using e|ditor for creating its extracurricular educative e-books. It plans to give a course to its affiliates in March 2013 on the use of e|ditor.

In all of these success cases, e|ditor has allowed the creation of e-books by regular users, without the need of specialized skills and at a small cost.

7 CONCLUSIONS AND FUTURE WORK

In this paper we present e|ditor, a platform to provide small publishing houses with a tool to create e-books at a commercially viable cost. It allows them to create from pocketbooks to pop-up e-books, textbooks and extracurricular activities e-books, with special care in school textbooks. It allows them to be autonomous in the e-book creation, without the need of specialized IT personnel.

The feedback from the publishers that have been using e|ditor shows that it makes them capable of creating the different types of e-books that they need, autonomously and within reduced budgets.

Although e|ditor addresses their digital content creation needs, more steps are needed to develop a solution to allow them to manage all the production, distribution and publishing processes and fully adapt their business to the publication and distribution of digital contents.

First, they need to provide users with a very intuitive and simple way to buy and consume digital contents. Users just want to easily and immediately find the right contents, buy them and consume them, without headaches. The process must be very simple and straightforward. To neglect this point would be a huge mistake, as the main e-reader and mobile devices platforms (Amazon Kindle, Google Android and Apple Iphone/Ipad) are already putting all the stress on it. They have the technological power (the platform control), and to successfully compete with them requires to make content access and consumption as easy and direct as possible.

A second issue to address is the fact that digital contents can be easily copied. Measures must be taken to reduce the risk of income loses due to piracy. This is a tricky point, because measures to avoid piracy usually go against the "easy to buy and consume" need. To find the right balance is crucial.

These challenges are not trivial, and successfully addressing them requires the joint collaboration of IT companies/researchers and small publishers to carefully analyze and address them. We are currently working in this stage in collaboration with some small publishing houses, with the aim of building the "integral solution" they need.

ACKNOWLEDGEMENTS

REFERENCES

Bouzán Matanza, J. M., 2008. *Ardora.* [Online]Available at: http://webardora.net

Cachia, R., Ferrari, A., Ala-Mutka, K. & Punie, Y., 2010. *Creative Learning and Innovative Teaching: Final Report on the Study on Creativity and Innovation in Education in EU Member States,* Sevilla: European Commission - Joint Research Center - Institute for Prospective Technological Studies.

Carroll, G. R., 1985. Concentration and Specialization: Dynamics of Niche Width in Populations of Organizations. *American Journal of Sociology,* 90(6), pp. 1262-1283.

Dublin Core Metadata Initiative, 1999. *Dublin Core Metadata Element Set, Version 1.1.* [Online]Available at: http://dublincore.org/documents/dces/

Generalitat de Catalunya, 1992. *zonaClick.* Available at: http://clic.xtec.cat

Goyal, K., 2006. *Calibre ebook management.* Available at: http://calibre-ebook.com/.

Junta de Castilla la Mancha, 2008. *Cuadernia.* Available at: http://cuadernia.educa.jccm.es/

Junta de Extremadura, 2005. *Constructor.* Available at: http://constructor.educarex.es

Markovic, S., 2009. *SIGIL - The EPUB editor.* Available at: http://code.google.com/p/sigil/

Prensky, M., 2001. Digital Natives, Digital Immigrants Part 2: Do They Really Think Differently?. Volume 9, pp. 1 - 6.

Underwood, J. et al., 2010. *Understanding the impact of technology: learner and school level factors, 2010,* s.l.: British Educational Communications and Technology Agency (BECTA).

Univ. of Auckland; Auckland Univ. of Technology; Tairawhiti Polytechnic, 2006. *eXeLearning.* Available at: http://exelearning.org

University of Victoria, 1998. *Hot Potatoes.* Available at: http://hotpot.uvic.ca

Vrasidas, C. & Glass, G., 2005. Achieving Technology Integration in classroom teaching. In: C. Vrasidas & G. Glass, eds. *Preparing Teachers to Teach with Technology, Information Age Publishing.* Greenwich: Information Age Publishing, pp. 1-20.