

SmatBook – a vision for the future e-book

Ivan Koychev¹, Roumen Nikolov¹, Darina Dicheva²,

¹ Faculty of Mathematics and Informatics, University of Sofia, Bulgaria

² Computer Science Department, Winston-Salem State University, USA
{koychev, roumen}@fmi.uni-sofia.bg
dichevad@wssu.edu

Abstract. This paper presents a vision for the future of the e-books as part of the growing collective intelligence. The vision entails further development of technologies that will facilitate the creation and use of a new generation of ‘smart’ books: e-books that are evolving, highly interactive, customisable, adaptable, intelligent, and furnished with a rich set of collaborative authoring and reading support services. The proposed set of tools will be integrated into an intelligent framework for collaborative book authoring and experiencing called *SmartBook*. The paper also discusses the opportunities to use SmartBook as a vehicle for building professional virtual communities of practice in the framework of a University 2.0 organization dedicated to industry-university cooperation.

Keywords: E-book, Web 2.0, University 2.0, Collaborative Intelligence, Personalisation

1. Introduction

Around 400 BC, Socrates said about books: “*and yet if you ask them a question they preserve a solemn silence. ...You would imagine that they had intelligence, but if you want to know anything and put a question to one of them, the speaker always gives one unvarying answer. And when they have been once written down they are tumbled about anywhere among those who may or may not understand them, and know not to whom they should reply, to whom not: and, if they are maltreated or abused, they have no parent to protect them; and they cannot protect or defend themselves.*”

IT technologies available today have made possible the advent of the e-book that overcomes a number of the weaknesses of the classic paper scroll described so well by Socrates 24 centuries ago. Many of the numerous recent e-book projects, such as the Gutenberg Project¹ and Google’s Books project², focus on digitizing paper books and organizing them in online repositories. The ability to present information in new ways, including display of multimedia content, interactivity, and customization, gives

¹ <http://www.gutenberg.org>

² <http://books.google.com>

the e-book enormous potential to extend the fundamental concept of the book and its impact on readers. In addition, taking advantage from the emerging social web, e-book authors can publish drafts of their books online to get early feedback, thus transforming book writing into a form of collective brainstorming. This could result in better ‘live’ books and a shorter writing cycle. Such e-books could be continuously updated and refined by many stakeholders; for example, readers could discuss published online content with the author or publishers could recommend extensions with information in demand.

Current technologies provide opportunities for building tools not only for authoring interactive multimedia content, but also for enhancing readers’ experience beyond multi-media presentation, for example, through *sharing comments, in-book search, browsing, skimming, visualisation, summarisation, and non-linear reading*. Although software tools supporting individually some of these functionalities are currently available, *there is a need for an integrated environment* that provides a complete set of social and intelligent tools to facilitate book writing and reading.

2. Related Work

E-books come from a complimentary tradition originating from early efforts to get existing texts online, e.g., the Gutenberg Project, followed by the early digital text publishers such as Hard Shell Word Factory³ and Online Originals⁴. These texts are typically linear; popular formats include Plain Text, HTML, PDF, Word, Open Standards such as IDPF and OEB, or Proprietary formats such as Kindle’s AZW or Mobipocket’s PRC. Although most e-books can be read on a standard PC, a number of specific devices have emerged – such as Sony LIBRIÉ (2004), Sony Reader (2006), and Amazon Kindle (2007) – that take advantage of better book-like form factors, and new display technologies that are easy on the eye (known as e-paper or digital paper).

The new generation of Web 2.0 systems (including blogs, wikis, personal journalism and sharing sites) is changing the general attitude to public writing, sharing, and artistic collaboration. This new Web literacy is changing people’s expectations of digital media and challenges our existing notions of authorship, making it the right time to explore how community-driven writing and reading could be applied to more formally published materials. The existing work in this area includes Amazon’s open publishing system called Digital Text Platform⁵, which allows authors to self-publish to their Kindle platform, and Hypertext publishers, such as Eastgate Systems⁶, which offer self-contained hypertexts that have many of the combined properties of open hypertexts and e-books. The latter, although formally authored, multimedia and non-linear, include no community or annotation aspects.

While technologies and tools for searching documents on the Web or in a digital library that best match a query have been the subject of tremendous interest and research in the last decade, the task of *locating relevant information within documents*

³ <http://www.hardshell.com>

⁴ <http://www.onlineoriginals.com>

⁵ <https://dtp.amazon.com>

⁶ <http://www.eastgate.com>

has remained in the shadow. Recently, this task has become increasingly important as longer documents, including many e-books, have begun being published. Several approaches have been proposed for within-document retrieval including passage retrieval [7] and content-based document browsing [6]. Harper et al. [5] proposed a tool called ProfileSkim that enables users to identify relevant passages of text within long documents by integrating passage retrieval and content-based document browsing.

Adaptive and recommender systems use user models aiming to help users in finding information, products, services, etc., that they are presumably interested in or prefer. There are two basic approaches pursued so far *Content-based filtering* systems take into account individual preferences for certain object features and make recommendations on the base of the similarities between items [9, 12, etc.]. *Collaborative filtering* systems, typically build on similarities between users with respect to the objects they are implicitly or explicitly interested in [8, etc.]. There are also *hybrid* systems, which combine both approaches to avoid some of their limitations and improve the quality of their recommendations [1, 3, etc.].

3. The SmartBook Vision

Even today a significant amount of multimedia content is often made available as a book companion (for example, on CDs, DVDs or on the Web). Book authoring and reviewing using digital facilities has influenced the workflows adopted by publishers and the processes followed by individuals. Collaborative work using social network infrastructures as well as the fan art and fan fiction that often accompany books on the Web are significant developments in both authoring and reading. The way in which we experience books has therefore changed dramatically. Books can be experienced on print, on the Web or on portable devices and they can be discussed in forums, enhanced by fans and be part of a shared experience among creative communities.

We advocate that it is possible to *re-conceptualize the book* by examining the essence of what the book means to individuals, society and culture at large, and how this could be supported and extended by digital platforms, connectivity and the emerging digital literacy. We believe that a new generation of e-books will better support creative communities, foster talent and promote innovation. The vision entails further developing of technologies that will facilitate the creation and use of a new generation of 'smart' e-books that are: evolving, highly interactive, customisable, adaptable, intelligent, and equipped with a rich set of author and reader support services. We propose a set of tools that will be ultimately integrated into an intelligent framework for collaborative book authoring and reading called *SmartBook*. It will contain three 'spaces': an author's space, a reader's space, and a collaboration space.

The *author's space* will provide authors with personal and collaborative editing tools. The major difference between these tools and the emerging wiki-based collaborative authoring tools (such as those in Wikibooks) is the ownership. Whereas the wiki approach allows all users to act as co-authors, SmartBook will preserve the traditional distinction of the roles of authors and readers; one or more authors, the book's owner(s), will be responsible for content creation. One of the main goals of

SmartBook is to provide authors with a set of smart tools to facilitate the presentation of the semantics of the text (better and easier than it is currently possible) by employing emerging technologies and recent findings in the areas of semi-automatic mark-up, natural language processing and text mining. The tools will make suggestions needing final tuning and/or approval or by the author.

In the SmartBook *reader's space*, readers will be able to create their own electronic copy of a book, in which they could mark, underline and comment, in the same way they commonly do with paper books. Readers will be able to edit only their own copies; the original book content will be preserved. A reader could further share an annotated personal copy with other readers and the author(s). Semantic Web and personalization technologies will be used to provide efficient and context-aware search and browsing of the book content and reader's comments.

The *collaboration space* will provide users with a virtual place where the involved stakeholders can discuss subjects of common interest related to a book. For example, there will be two default "meeting rooms" for each new book - one for writers and one for readers. Additional rooms could be created as needed. Writers' meeting rooms will be restricted to agents involved in book creation, including writers, reviewers, and editors. The publisher will set the access rights. Readers' rooms will be open and accessible by both the readers and the book's creators. The collaboration rooms will be furnished with state-of-the-art tools for *sharing information* and thoughts in written and oral form, such as a virtual white board, online video, voice conversation, etc. The aim is to bring together the most recent advances in technology in collaborative working environments in order to promote social contacts and facilitate interaction and knowledge exchange between readers, authors, and publishers. Thus SmartBook will be a promising candidate to become a key launch pad for growing and cultivating the emerging collective intelligence [13, 14]

In this context, we have to identify how the role of the *publisher* will evolve. On the Web, anyone can publish anything. Quality control (e.g. peer reviewing) is not common, which makes content less trustworthy than scholarly publications. There are no selection guidelines for search engines. This situation makes the task of the readers searching relevant trustworthy information difficult, especially in the case of tight time constraints. This is even more important in educational scenarios, where the readers need trusted information presented in a suitable for learning way.

With regard to *personalisation*, the current advances in the fields of adaptive hypermedia will be used: e.g. user modelling and recommender systems. A variety of semantic technologies for ad hoc discovering and presenting associations between terms/concepts, ontologies, contexts, etc. will be employed. User profiles will be utilised in helping users to find information resources and services of interest or preference. The personalisation will employ both above-mentioned approaches for recommendation - content-based and collaborative or social. The framework will be a *plug-in architecture* allowing integration of third party tools, such as collaborative editing tools and online communication facilities.

4. SmartBook Design and Development

We identified three main dimensions of the project as follows:

- Building a collaboration environment that facilitates the interaction between book stakeholders. In particular, it will support:
 - Enabling *rich media content* management;
 - Developing an integrated collaborative *authoring* space including support for ‘live’ feedback from readers about the book before and after its completion;
 - Developing an interactive social *book reading* space including discussion forums; comments, bookmarks, annotations, and tags sharing, etc.
- Developing tools for intelligent information retrieval.
 - Supporting (semi-automatic) semantic content annotation, including building ‘knowledge maps’ of the book content (a generalisation of the traditional ‘back-of-the-book’ indexes);
 - Efficient semantics-aware in-book search, browsing and summarisation.
 - Visualisation of content structure and semantics;
 - Context-aware ‘live’ search and recommendations of relevant resources.
- Developing *personalisation* support for SmartBook.
 - Designing and prototyping a user modelling server.
 - User model acquisition from log files.
 - Personalisation of the content and its presentation.

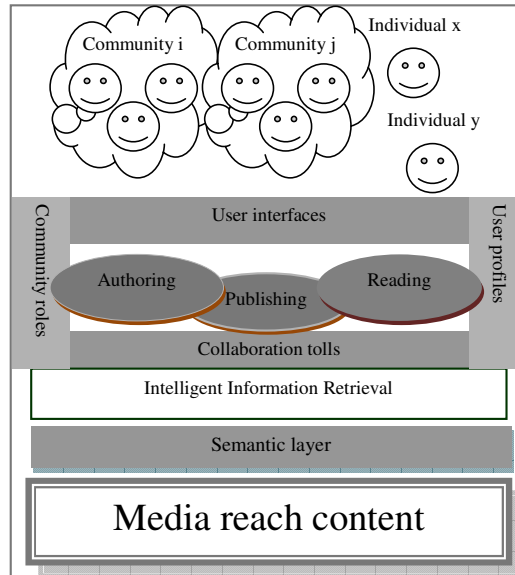


Fig. 1. SmartBook architecture

Fig. 1 presents the overall architecture of the SmartBook framework.

With regard to content organization, SmartBook uses semantic models at three different levels:

- *Object level*: content objects annotated using metadata standards (such as Dublin Core or IEEE LOM) and expressed in popular multimedia formats. Such objects are interoperable and can be imported from other sources such as personal archives, or online repositories.
- *Book level*: a open SmartBook ontology of book structure (referencing the object level) and further annotated with author and contribution information. This ontology will be open.
- *Conceptual level*: knowledge maps conceptually modelling knowledge domains. This layer plays the role of a global schema, providing a declarative description of the subjects *within the content* in terms of key concepts and relationships between them. Knowledge maps are presented in standards-based formats (such as SKOS) and are therefore interoperable.

At a simple level, the latter will provide support for readers to easily search inside book content. However, since it is much richer in structure as compared with a simple index, it will provide much more to the reader – for example, such a structure can present concepts in the context of their relationships to other concepts. This in turn will allow for efficient semantic search and browsing of the learning content.

5. SmartBook and University 2.0

On the way to a knowledge society in a dynamic ICT environment, the universities should catalyse a process of deep institutional change. One of the major challenges facing the universities in the next decade is to reinvent themselves as information organizations [16]. The model of University 2.0 naturally emerged as a framework for universities to adapt to the social computing phenomena and to the networked information economy. University 2.0 can be described as a research and entrepreneurial university, which integrates Web 2.0 technologies and applications in all university activities, including ones with all knowledge intensive stakeholders, and implements the features of Enterprise 2.0 [11]. A basic concept in bridging the university and industry is *Community of Practice* (CoP), which is defined as “*groups of people who share a concern, a set of problems, or a passion about a topic, and who deepen their knowledge and expertise in this area by interacting on an ongoing basis*” [17]. The emerging Web 2.0 based platforms and tools become driving force for building multiple-level learning and capability for the company creation net [4]. For the researchers and innovators in industry and academia, the whole Web becomes an R&D department and the Science 2.0 is emerging [15].

Considering the future of college-level instructional/learning resources, the SmartBook vision fits nicely the University 2.0 concept. Recent research and development efforts in learning content authoring have concentrated on authoring educational hypermedia, focusing on system adaptivity. The resulting hyper-books are interactive, rich in terms of employed modalities for content presentation, and easily browsable. However, their structure (hyperlinks connecting selected concepts to Web pages explaining them) departs significantly from the traditional sequential textbook

structure, which in many cases is more appropriate for story-telling or a narrative introduction of a new topic. Since readers are used to conventional printed media, hyper-books can often be more difficult to follow, as shown by studies reporting *cognitive overload* for readers [2, 10]

An innovative form of electronic book could combine the advantages of both conventional printed books and educational hypermedia. To reach this goal, we propose two basic perspectives for *organising learning content* in SmartBook:

- A collection of hypermedia units modelling the traditional organisation of a book: sequential book pages containing the content structured in chapters, subchapters, sections, etc., using multiple modalities for presenting the information (text, graphics, video, audio, etc.) This organisation should be supplemented with functionalities that model traditional practices when reading conventional books, such as highlighting text, adding reader's notes, and book-marking.
- A *semantic structure* of interrelated concepts representing a knowledge structure of the academic subject domain (conceptual classification).

The second perspective can be seen as playing the role of a *back-of-the-book index* for SmartBook, which provides support for readers to search easily inside the book content. SmartBook will be also furnished with within-document text retrieval tools that efficiently support users trying to identify information within long documents.

In general, SmartBook is envisaged to impact four key areas that are highly relevant to University 2.0: participative and communicative forms or content; enabling publishing of innovative content; automating the collection and distribution of knowledge; automate links between scientific data and discussion.

It will have a significant impact on the first three of these areas and will indirectly make contributions to the fourth one.

6. Conclusions

The SmartBook vision implies re-conceptualisation of the book as rich-media, interactive, intelligent content serving as a focus point for community discussion. It employs technologies that facilitate the creation and use of a new generation of *published* SmartBooks. Although research and educational scenarios most clearly demonstrate the need for more formal notions of publishing in digital texts, the ideas are general in nature and are transferable to other book genres.

SmartBook is not focused on the e-Science domain in particular, and will support general writing, discussion and publishing workflows, rather than those more tightly associated with experimentation and sharing of results. SmartBook is advocating a revolution in the way digital material is published, and will enable evolutionary creation of text over many iterations. It also employs personalisation techniques that could be extremely useful in the dissemination of scientific results – either for scholarly discussion in expert communities or to aid the discourse and relationship between experts and the general public.

References

1. Balabanovic, M., Y. Shoham: Content-Based, Collaborative Recommendation. *Commun. ACM* 40 (3): 66-72 (1997)
2. Beasley, R.E., Waugh, M.L. (1995). Cognitive Mapping Architectures and Hypermedia Disorientation: An Empirical Study. *Journal of Educational Multimedia and Hypermedia*, 4(2/3), 239-255.
3. Billsus, D. & Pazzani, M. (1999). A Hybrid User Model for News Story Classification, Proc. of the Seventh Int'l Conf. on User Modeling (UM '99), Banff, Canada, June 20-24, 1999.
4. Hagel, J. and Brown, J. (2006), Creation Nets: Harnessing the Potential of Open Innovation, <http://www.johnhagel.com/creationnets.pdf>.
5. Harper, D.J., Koychev, I., Sun, Y. and Pirie I. (2004). Within-document Retrieval: A User-Centred Evaluation of Relevance Profiling, *Journal of Information Retrieval*, 7, 265-290, 2004, Kluwer Academic Publishers.
6. Hearst M A (1995) TileBars: Visualization of Term Distribution Information in Full Text Information Access. In: Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI), Denver, CO, pp. 65-71.
7. Kaszkiel M and Zobel J (1997) Passage Retrieval Revisited. In: Proceedings of the Twentieth International ACM-SIGIR Conference on Research and Development in Information Retrieval. Philadelphia. ACM Press. pp. 178-185.
8. Konstan, J. A., Miller, B. N., Maltz, D., Herlocker, J. L., Gordon, L. R., and Riedl, J. (1997). GroupLens: applying collaborative filtering to Usenet news. *Commun. ACM* 40, 3 (Mar. 1997), 77-87.
9. Lieberman, H., Letizia (1995): An Agent That Assists Web Browsing, International Joint Conference on Artificial Intelligence, Montreal, August 1995.
10. Murray, T., Gartner-Piemonte, J., Kelleher, Ph. (2001). Evaluating Features for Conceptual and Narrative Flow in an Adaptive Hyperbook. J. Moore et al. (Eds.) *Artificial Intelligence in Education*, Amsterdam: IOS Press, *Frontiers in AI and Applications*, Vol. 68, 21-32.
11. Nikolov, R (2009), Towards University 2.0: a space where academic, education meets corporate training, in: IPROF-09, ICT Professionalism: a Global Challenge, 12-15 Feb. University of Arnhem and Nijmegen, Netherlands
12. Pazzani, M. and Billsus, D. (1997). Learning and Revising User Profiles: The Identification of Interesting Web Sites. *Mach. Learn.* 27, 3 (Jun. 1997), 313-331
13. Satnam Alag (2008) *Collective Intelligence in Action*. 424 pages, Manning Publications, SBN: 1933988312
14. Segaran T. (2007) *Programming Collective Intelligence Building Smart Web 2.0 Applications* Pages: 360 ISBN 10: 0-596-52932-5 ISBN 13: 9780596529321
15. Tapscott, D. and A. Williams (2006), *Wikinomics. How Mass Collaboration Changes Everything*, Penguin Group (USA)
16. Unsworth, J. (2008). University 2.0, in: R.N. Katz (Ed.) *The Tower and the Cloud: Higher Education in the Age of Cloud Computing* (pp. 227-237), Educause, 2008
17. Wenger, E., R. McDermott, W. Snyder (2003), *Cultivating Communities of Practice*, Harvard Business School Press, Boston.