The sBook: towards Social and Personalized Learning Experiences

Myriam Ribière
Alcatel-Lucent Bell Labs
Route de Villejust
91620 Nozay, France
+33 1 30 77 15 10
myriam.ribiere@alcatel-lucent.com

Jérôme Picault
Alcatel-Lucent Bell Labs
Route de Villejust
91620 Nozay, France
+33 1 30 77 11 41
jerome.picault@alcatel-lucent.com

Sylvain Squedin
Alcatel-Lucent Bell Labs
Route de Villejust
91620 Nozay, France
+33 1 30 77 27 18
sylvain.squedin@alcatel-lucent.com

ABSTRACT
Despite the widespread use of Web 2.0 techniques in our entire surrounding environment, which tend to make it more social, more dynamic and driven by users, some domains have not really changed yet. This is the case for (e-)books which reading remains mainly a solitary activity – or which is done at least without appropriate collaborative tools. However, the benefits of making this activity - and especially *active reading* - more social and digital are huge - in particular for people having learning reading goals - leading potentially to a wide range of new services: faster access to information, possibility to interact with people sharing similar concerns or able to provide relevant explanations, determining most interesting areas in a book, or even helping users accessing faster the information that will make them progress in their learning curve. Thus, in this position paper, we describe a set of concepts and features about a "sBook", which consists in making e-books more social, more communicative, in order to sustain students in a learning activity, and leverage collective intelligence from social interactions to make students' learning experience more efficient.

Categories and Subject Descriptors
H.3.3 [**Information Search and Retrieval**]: *Information* filtering; H.3.7 [**Digital Libraries**]: *collection, user issues;* J.4 [**Social and Behavioral Sciences**]: Sociology;

General Terms
Algorithms, Design, Human Factors.

Keywords
sBook, e-book, dynamic social network, learning activity, heat maps, community activity summaries, personalized learning path.

1. INTRODUCTION
The era of social connectedness is becoming richer and more pervasive, helping people to leverage the knowledge and insight

of various communities as they make choices and decisions. Web 2.0 brings new opportunities in communication and content consumption that are at the heart of this evolution, driving significant changes in the way people learn *from* and *with* others.

At school and universities, e-books are starting to be resources more and more used in order to sustain students' learning activities. However, these resources are still quite static and do not provide as much interactions as they could. Thus, there is a need to reinvent the interactions students have with the books.

In this paper, we propose some ideas to better exploit these resources in order to improve students' learning experiences. Section 2 explains the rationale of using the book as the starting point of a community creation. Section 3 provides some examples of new services using book-initiated social interactions for the purpose of faster learning within a single book. Section 4 gives some indications on how it would be possible to use collective intelligence to help students in their learning paths. Finally, we present some ideas in section 5 to extend the e-book towards the whole reader's environment by making the link between e-books and additional related multimedia content.

2. CONTEXT AND POSITIONING
We propose to benefit from everyday life communities for an enriched, social learning and discovery experience. Indeed, people are already part of numerous communities, some are present and identified in their digital world (web social networks, forum etc.), and others are *latent* communities of their life, which can be induced naturally by some user activities such as book reading. Thus, in this paper, we focus on those kinds of dynamic communities, which are activity- or context-formed. More specifically, we study how to *transpose* the reading activity into the digital world, and investigate how to capitalize on people's interactions and people's content generation to provide new services on top of the book, based on *collective intelligence*, in order to enhance students' discovery experiences.

2.1 Students environment
Students already have practices of social networks in their everyday lives. It is now important to understand how to take advantage of such practices and of this social culture to bring new methods of learning. Students constitute a good population target because (1) they have a lot of social interactions, (2) they interact a lot with books for their studies and finally (3) they are in an *active* mode when reading a book.

In order to better understand current practices, we conducted a small experiment with 10 students. They have been asked to fill a questionnaire about their current practices in active reading, collaborative work and sharing habits. We learnt that students' first aim in note taking is to *memorize content*, then *understand* and *synthesize knowledge*. The aim of working together on a common document is first to *share knowledge and gather socially* and only afterwards to *better understand* and *learn.* Students of the sample tend to provide information rather than to wait for requests. The understanding issues about a document are generally first addressed orally, and when asking someone for help, the most important values are *trust* and *personal links* with the person, *availability* and *expertise* come afterwards. We also learnt that the main reason for students not sharing *today* their notes or annotations about their readings is a lack of tools for doing it easily, and for supporting social, direct interactions that are needed to explain their notes to someone else. These initial results tend to reinforce the usefulness of introducing social practices within e-books. On the other hand, there exist already some tools on the Web that support collaborative annotation. For example, Diigo, Reframe It, MyStickies, and Google Sidewiki[1], propose users not only to share bookmarks, but also to digitally annotate web pages. Such tools are already bringing a social dimension to Web content; we have to understand how to transpose this experience into digitalized text books and how to make students benefit from the community of readers of the book to share annotations and learn better and faster through efficient and contextually pertinent communications and collective intelligence services around the book.

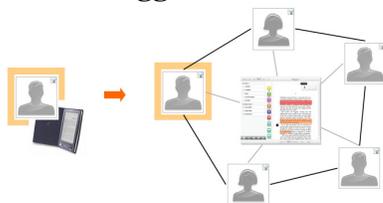## 2.2 E-book as a trigger of social interactions

Figure 1: The social experience of active reading

The set of people interacting with an e-book form a de-facto community. Currently these people are not connected together or only at a very high-level and outside the book through some kinds of forums (such as in LibraryThing[2]) or general social cataloging features (GoodReads, Shelfari[3]) to share good books and find new books to read. This does not allow fine-grained interactions about specific parts of the book. Bookglutton goes a step further by letting users chat while reading, but does not exploit annotations a lot. We propose to connect people through the sharing of a same resource, into the *spontaneous social network of the e-book* (Fig. 1). Thus, the reader doesn't have to log into a specific social network, he becomes implicitly logged in when the book is opened. This approach paves the way to more opportunistic services. Most of existing annotation solutions on e-books (Kindle[4], Txtr, Copia) propose to share annotations to users' web

social networks such as Facebook, Twitter, etc. But we believe it is important to keep readers *within* the book to make them interact. The readers' virtual community of an e-book proposes to people means for sharing their annotations per paragraph but also communication means to interact *about* and *from* the book and build knowledge. Annotations *in* context become a trigger for conversations and a unique way to help students to learn faster through interactions and readers' user generated content.

2.3 Annotations, conversations and learning

Active reading is the combination of reading with critical thinking and learning. It represents a fundamental part of education and knowledge work. Active reading does not involve just reading, but also underlining, highlighting and making comments, either on the text itself or in a separate notebook. In the context of the social network of the book, annotations - when shared - can become the support of students' conversations, as they provide the *right context* for readers: people and semantic context linked to a specific book section. Given this importance, we propose to consider four types of annotations:

- *Symbolic annotation* can help the user to identify points of interest in the content, or points of difficulties. It consists of a simple annotation the user generally does during his first reading: underlining, highlighting text, use symbols to mark an interest or a difficulty about a highlighted text.

- *Semantic (or free text) annotation* represents the user's interpretation of the content element, e.g. an analysis or a summary of their understanding; it is expressed through free text (natural language).

- *Tag annotation* enables to index the content element of a document with a list of tags for further user-driven search capabilities that could enhance current text-search.

- *Link annotation* proposes to the user to add by hand related content to the content element. Links can point inside the overall document itself, like a user's map for reading the content or outside the document, such as multimedia related content (image, video, audio, text, another e-book).

Inside the social network of the book, the whole collection of annotations of different types can bring added-value services, which the first one is *conversations*. We want to provide book-driven interactions, so that users can easily visualize who wrote an annotation and initiate from the book a communication with that person (taking into account his availability), e.g. to clarify with the author of an annotation its meaning. Another possibility is to exploit symbolic annotation to make people connect (offering co-readers as helpers in the exploitation of the e-resource), or create sub-communities inside the e-resource. Because learning is inherently social, as stated by Golub [5], conversations are essential to an efficient and collaborative learning.

3. SOCIAL KNOWLEDGE ACQUISITION

Transforming the book into a communication service and the provider of user-generated content should come with advanced features helping user in their readings to find appropriate information and learn in a more efficient way.

3.1 Heat maps

First, we introduce the notion of "heat map" of the e-book. The *heat map* is a service helping readers to identify key paragraphs

---

[1] diigo.com; reframeit.com; mystickies.com; google.com/sidewiki

[2] www.librarything.org; www.bookglutton.com

[3] www.shelfari.com; www.goodreads.com

[4] kindle.amazon.com; txtr.com; www.thecopia.com

within the book. The interest is to offer students a mean to relieve the pain to read all the content for getting important information and knowledge from the book.

In a first instance of the heat map, the collection of symbolic annotations denoting an interest for paragraphs are used to compute a statistical value for each chapter and each paragraph. This value corresponds to the proportion of annotations of this chapter/paragraph w.r.t the entire book (Fig. 2). Amazon's Kindle e-reader has a similar heat map feature, letting users bookmark book sections and export them to existing social networks such as Facebook, Twitter, etc. The drawback of this first simple approach is the consensus of the result. People do not have the same goals while reading, and especially students.
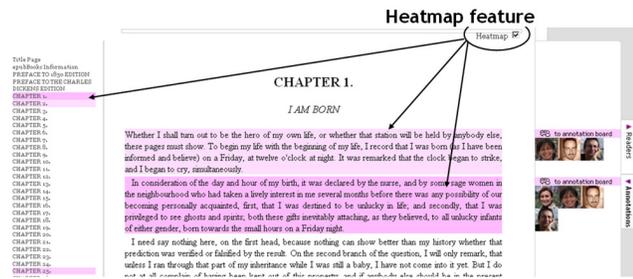

Figure 2: *Heat map* display, darker paragraphs/chapters are the *most annotated paragraphs*

To provide a more powerful and personalized *heat map*, we have to target the right user community for computing popularity values for chapters and paragraphs. A first approach is to help users to define the scope of the community according to three criteria: geography, language (culture) and social relationships. The social relationship criteria takes into account the user's social network and defines the user's book community as the intersection of the book social network with the user social network. A student can define the scope of his book community as "readers located on the campus and that are part of his social network" as a way to target similar-minded readers. Another approach to build an enhanced heat map consists in identifying explicitly or implicitly the intention of the readers in their active reading task. A basic solution is to ask the user to indicate into the system the purpose of their reading, such as "studying the author style in the literature course". This could be expressed through a simple taxonomy associating an action verb denoting the goal (find contact, learning, writing a report, etc.) and a complement denoting the object (20th literature, modern art, etc.). The system compares these explicit reading goals with the ones of other users (using similarity between concepts) and recommends book passages that have been annotated by people sharing similar reading goals. The ultimate step would be to understand implicitly reading intentions; this may be done by analyzing reading activities [9]; some initial steps are explained in section 4.

3.2 Social annotations summaries
The social network of an e-book produces user-generated content (annotations and conversations). Whereas this is a useful source to construct meaning, this also introduces too much distraction. The user should be able to find quickly annotations relevant for him (on a chapter or on a paragraph). Profiling methods can be used to generate the summary of a paragraph or a chapter as a tag cloud, which enables the reader to identify key concepts of the paragraph/chapter when reading it.

Such a mechanism of activity summary around the book (or book section) can be done in real-time, and shared through Facebook and Twitter to bring more synchronous communication, by solicitation of people outside the book to join, share and exchange inside the book. This real time status update of the book gives the opportunity to open the social network of the book and to correlate activities of other book social networks.

Finally, the different types of annotations such as symbolic and semantic annotations may also be exploited to build a sort of FAQ into the book, by associating question mark symbolic annotations to best rated semantic (free text) annotations. Thus, during his first reading, the student may only use symbolic annotations without being disturbed by any other activity. Then, in a second step, the system could display to him the best annotations associated to paragraphs the user marked as "?".

The summaries of annotations, the openness of the book's social network and the exploitation of links between symbolic and textual annotations are the pillars for providing new ways of reading a book and socializing around a book in a knowledge acquisition context.

3.3 Social tagging and interlinking
Communities may be powerful to provide answers to problem such as the indexation and interlinking of multimedia content. Social tagging has been analyzed as an efficient complementary method to semantic annotations. [15] provides a study demonstrating that a global semantic model can be statistically inferred from the folksonomy to semantically annotate web resources. We propose to use social tagging methods on the e-book to provide a new, reader-driven index of the book. Social tagging transposed into the book for tagging paragraph could change the way of navigating into a book.

Going further in the exploitation of e-book annotations from the readers' community, we propose an alternative and social way to content interlinking. This challenge, often presented as a semantic issue in the understanding of links between contents, could be transformed (such as social tagging) into a *social challenge*. In the sBook, readers can associate together pieces of content they are judging pertinent to the book. The global pertinence of these links may then be assessed by the book community through an implicit feedback mechanism, which can also validate their position in the e-book. Implicit feedback could be deduced from the user behaviour, for example through a sequence of click, time spent, actions such as play, copy/paste, scrolling etc. Such implicit feedback methods [14] are widely used for search engine and could be adapted to the particular context of the book.

The potential of the collection of users' annotations for providing added-value services into the e-book is tremendous. Our proposition around heat maps, summaries, and social tagging/ interlinking are just some potential examples. We must better target, through user experiments, useful services that will ease knowledge acquisition and make it social.

4. COLLABORATIVE LEARNING PATHS
When students are investigating a new area, their learning activity through e-books does not consist of studying different information pieces in isolation or in a fragmented way; instead, learning is a *continuous process* that consists of a *sequence* of reading activities (which includes reading, annotating, sharing book parts), which progress over time tends to follow learning curves

[11]. Usually, knowledge items are learnt in a certain order (logical or pedagogical). Although generally teachers provide guidance on how to proceed, the problem becomes more difficult for other forms of learning, such as personal information discovery. So, how could we help students progress in these learning paths, in their *future information needs through* proactive suggestions of contacts or content?

When an e-book becomes social, there could potentially be many annotations, many users, so it may become difficult for a user to get relevant information for him. Even if features such as *heat maps* provide some statistical indications on what could be of general interest to the student, it does not take into account past things he learnt, and the fact that by nature, learners are diverse [12], have personal needs of information and that each student has his own *learning path*.

4.1 Personal and collaborative learning paths

Contrary to what is done in e-learning, such as in [4], we do not consider learning paths as a graph that connects very formalized knowledge states, because current knowledge models are restricted to very specific domains (such as mathematics). We define a *learning path* with e-books as the route (i.e. sequence of reading activities) taken by the student, which allows him to build knowledge progressively. For example, a student may want to understand modern art; he will start looking at a general introduction book, and then may investigate more specialized work about abstraction or surrealism. Even if the first steps (first reading content) towards this learning goal may be easy to find, it could be much more difficult to find the next ones. Thus, can we recommend the student appropriate reading steps that fit his learning path?

The problem we are considering is the following: given a certain user's reading sequence (denoting a particular learning goal), how to propose the user the next step(s) in his personal learning path, that enables to progress in his current learning curve? This *recommendation problem* is quite specific compared to the approaches existing in the literature: as pointed out in [6], there is very little work on recommenders of sequences. We propose to use a collaborative approach that exploits user-generated content associated to the book, with the assumption that users sharing a similar reading sequence are more likely to share same questions, goals and intentions about the e-book. We focus in particular on symbolic annotations, which denote interest, understanding or questions w.r.t. book passages. The choice of this particular kind of user action is done because (1) it denotes a real interaction between the reader and the book (active reading) (2) it is explicit, therefore less prone to noise than inferred actions such as reading time. Our goal is to create a mechanism that: keeps track of users' learning paths; suggests next steps to help users fulfil it; and lets share learning paths between users, so that learning experiences are not lost.

4.2 Learning path within an e-book

In a first approach, we will focus on learning paths within a single book. For simplicity purpose, we will classify user's annotations into two big classes: *questions* (misunderstanding, etc.) and *explanations* (comments, remarks, interpretations, etc.) Let's consider the following example (Fig. 3): user Bob starts to put some symbolic annotations in the e-book: "?" when he does not understand a paragraph, "!" when it is of particular relevance.

After a while, he realizes that there are more and more things he does not understand and asks for some help in the system. The system finds users of the social network of the book that have annotations in common with him. Then it analyzes sequences of annotations of those people, compares them with the sequence of annotations Bob did, and at the end, suggests possible next steps within the book, in terms of contact opportunities or reading opportunities: for example: "you're sharing similar problems as John with this book, you should contact him" or "read annotations of Bill (he seems to have understood…)". Thus, the reader benefits from the similarities of path of actions between users within the book.
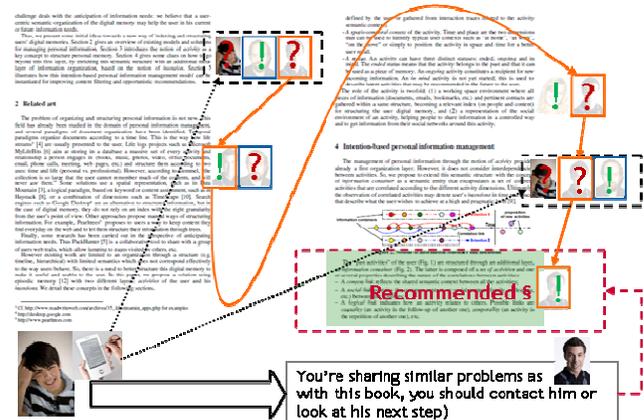

Figure 3: Collaborative learning path within one single e-book

However, this approach raises a lot of challenges: (1) how to limit cold start and sparsity problems? This depends a lot of the density of annotations, and therefore of the ease of use and usefulness of e-book annotation tools for students. (2) What kind of algorithms can be used to efficiently compare sequence of annotations between users? And which time-window should be considered in the analysis? A sequence of annotations is somehow similar to a sequence of strings *abfgzk*, for which each letter would encode the identifier of a paragraph in the book (including possibly the semantics of the annotation in terms of broad classes such as: questions / explanations). However, well-known string pattern matching methods such as Knuth-Morris-Pratt or Boyer-Moore are not really adapted. Instead, we have to consider fuzzy matching methods [3] which enable to find similarities between sequences but that are tolerant to a certain amount of mismatches, or dynamic pattern matching [1]. We are currently assessing the relevance of these methods and defining an appropriate solution.

4.3 Recommendations of e-book collections

In a second step, we propose to extend this approach to a book collection. It could help students move from one book to another one, in order to better help them in their learning goals. This has also for consequence to make the link between various spontaneous social networks, with increased opportunities for new contacts and social interactions.

For example the reader starts a book about China (Fig. 4). When starting reading it, some parts are not clear about Chinese dynasties, the system will use the correlation of this sequence of annotations of readers of the book (including link to external source of information they have put in the book) to suggest possible additional e-books extracts that will help him progress.

The same process may be repeated within the new book, which may lead to recommendations to new content pieces or to go back to the initial book once the learning gap is covered.

Thus, instead of considering annotation sequences within a single book, we consider annotation sequences that include different books, which induces new challenges among which: (1) identifying the reading paths through different books correctly and filtering noise: we intent to exploit hyperlinks readers can put as annotations on a first book to make the link with other books; (2) assessing progress of the user in his learning curve: we have to ensure that recommended book passages favour the learning curve of the user (identify how a new content affects the reader's progress and how the reader could use this additional information to select between multiple learning paths). This might be done currently afterwards, by checking the annotations the user put on a new recommended content, but there is a need of more predictive models.
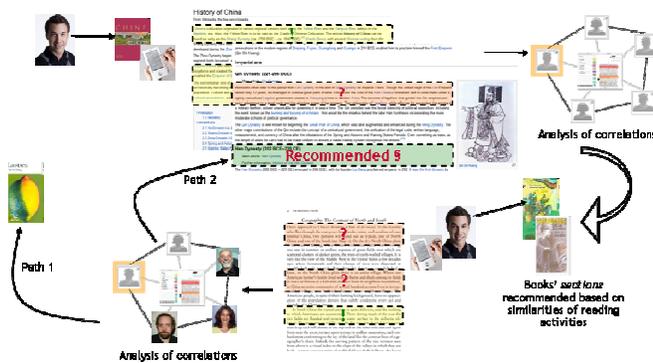

Figure 4 - Recommendations through e-book collections

5. ENVIRONMENT-ENRICHED E-BOOK LEARNING EXPERIENCE

The e-book should be more than a basic technological object extending human capacities of communication. The ability to read, write and capture information changes with new communication technologies emerging from the Internet [8]. These new behaviors will make evolve the conception of learning applications towards systems enabling to improve the student's learning path by the *combination of reading acts and the discovery of related complementary information* (photos, audio, video, documents).

So, we plan to integrate these new practices as a way for the user to improve his learning experience. The reader will be able to upload pictures and documents but also video or audio recording directly from the reading location within the book. The system will compute automatically the relationship between the current reading passage and the uploaded content in order to make it available to all related learning contexts.

Several technical challenges have been already identified to set up the system:

- How to compute and generate the relationships between uploaded content pieces and the user's reading context?

- How to notify the user about new multimedia content relevant for his learning context?

5.1 Binding multimedia content with readings

The sBook usage will change the reading activity into a connected activity. Searching complementary information on the Web can be made in continuity of the reading activity and on the same device. We plan to benefit from user searches to associate to his current reading context an annotation embedding a basic URL. In complement to link annotations suggested in section 3.3, we propose to manage links to a video for a better video integration inside the e-book. By the way, several related works such as [2][7] suggest the usage of video in the learning context, and [13] analyzes the relationship between video-technology and teacher education.

Instead of binding the annotations to a specific location within the book (such as a paragraph, a chapter), our system will attach a video to a *learning context*. The same video can so feed several books and passages within a book. The uploaded content is associated with a set of metadata automatically generated from a semantic module [10] which analyses the paragraph, the page, the book, and the global user's learning context. The description of the learning context can be derived from the reader's learning path (described in section 4), for example thanks to the semantic analysis of the different elements (annotations and related paragraphs) that constitute the learning path of the user through one book or a collection of books. The reading context can be considered as a particular case of the notion of "*information container*" as defined in [9]. Other sources of metadata (e.g. users' tags) enable to compute and generate relationship between the context of the upload location within the book and the uploaded file.

5.2 Real-life event notifications

Once these video annotations are available, they can be suggested to readers in relationship with a specific reading context. The system can propose these videos directly on the sBook reader interface but can also use a notification mechanism to alert the user in an opportunistic way, depending on the relationship between his current readings and contextually available videos.

Notifications can be sent using SMS, mail, social network, etc. and support real-time alert for live streaming. For example, while reading a technical book, the system may detect a link to a live video streaming of a conference put by another reader of the book sharing a similar reading path with the current user. The system analyzes in real time the relationship between the semantic description of the user's current learning context and the description of the live video streaming. If the matching result is high enough, the student may be incited, directly *from* the book, to join the event, by clicking on the link provided with the notification to open video streaming. If the user is not connected into the social book network, the user may be notified using SMS or real time notification.

This overview of the new possibilities offered by e-books in a learning context shows that it requires multiple competencies and a multidisciplinary approach: (1) to understand the social dynamics that e-books enable and the potential of associated social services; (2) to understand individual users for better serving their learning needs and guide them in their learning paths and also (3) to analyze multimedia challenges in order to enrich

the original book and benefit from the full ecosystem the e-books belong to.

6. CONCLUSION AND FUTURE WORK

The set of ideas about the future of e-books we have presented in this position paper are currently being addressed in the scope of a collaborative project between Alcatel-Lucent, Abilene Christian University and Cambridge University Press[5]. The goal of this project (which runs until 2013) is to develop an application to facilitate cross-media and cross-community information discovery: facilitate information discovery with contents of all sorts from all sources; extend the e-book concept to be a dynamic collection of multimedia contents from all sources and extend reading to discovery for formal, leisure and spontaneous browsing and learning. The project has started with the implementation of an advanced e-reader platform, letting students put annotations and share them, plus the heat map feature. This already allowed us to test the acceptance of first social features in e-books. The first feedback we had during demos encouraged us to go further. Other advanced features – social, personalization or multimedia ones, as described in this paper – will be added progressively to the e-book platform, with a balance between implementation and experimentation with students, to ensure that we always answer students' needs.

7. REFERENCES

[1] Amir, A. Landau, G. M., Lewenstein, M. and Sokol, D. 2007. Dynamic text and static pattern matching. ACM Trans. Algorithms.

[2] Astleitner, H., Wiesner C. 2004. An Integrated Model of Multimedia Learning and Motivation. *Journal of Educational Multimedia and Hypermedia.* vol. 13.

[3] Butman, A., Lewenstein, N., Porat, B. 2007. Jump-Matching with Errors. In *Proceedings of the 14th international conference on String processing and information retrieval* (Santiago, Chile)

[4] Falmagne, J.-C., Cosyn, E., Doignon, J.-P. and Thiéry, N. 2006. The Assessment of Knowledge, in Theory and in Practice. LNCS, Springer.

[5] Golub, J. (Ed). Focus on Collaborative learning. Urbana, IL: National Council of Teachers of English, 1988.

[6] Herlocker, J.L., Terveen, L.G., Konstan, J.A., Riedl, J.T. 2004. Evaluating collaborative filtering recommender systems. *ACM Transactions on information systems* 22, 5–53

[7] Hoffner, H., Baker, E., Quinn, K.B. 2008. Lights, Cameras, Pencils! Using Descriptive Video to Enhance Writing, *The Reading Teacher* 61-7 / P 576-579

[8] Leu, D.J., Kinzer, C.K., Coiro, J., Cammack; D.W. 2004. Toward a Theory of New Literacies Emerging From the Internet and Other Information and Communication Technologies.

[9] Picault, J., Ribière, M., Senot, S. 2010. Beyond life streams: activities and intentions for managing personal digital memories. In the *International Workshop on Adaptation, Personalization and Recommendation in the Social-semantic web (APRESW 2010)*, CEUR-WS, Vol-585, pp. 25-32.

[10] Richardson, S.D., Dolan, W.B., Vanderwende, L. 1998. Acquiring and structuring semantic information from text. *Microsoft Technical Report*

[11] Ritter, F. E., & Schooler, L. J. 2002/ The learning curve. In *International Encyclopedia of the Social and Behavioral Sciences*. Amsterdam: Pergamon

[12] Smith, B. L., & MacGregor, J. T. 1992. What Is Collaborative Learning? National Center on Postsecondary Teaching, Learning, and Assessment at Pennsylvania State University

[13] Wang, J., Hartley, H. 2003. Video technology as a support for teacher education reform. *Journal of Technology and Teacher Education*, vol. 11

[14] White, R. W., Ruthven, I., & Jose, J. M. (2002a). The use of implicit evidence for relevance feedback in web retrieval. In *Proceedings of 24th BCS-IRSG European Colloquium on IR Research, Lecture notes in Computer Science 2291*, 93-109

[15] Wu, X., Zhang, L., and Yu, Y. 2006. Exploring social annotations for the semantic web. In Proceedings of the 15th International Conference on World Wide Web (Edinburgh, Scotland, May 23 - 26, 2006). WWW '06. ACM Press, New York, NY, 417-426.

---

[5] http://www.cambridge.org/press/article.asp?artid=154961