SRECMATs - an intelligent tutoring system to deliver online materials for student revision

Petch Sajjacholapunt[1], Mike Joy[1]
[1]*Department of Computer Science, University of Warwick, United Kingdom*
{*P.Sajjacholapunt, M.S.Joy*}*@warwick.ac.uk*

Keywords: Online materials management, online materials design, interactive online materials, student revision.

Abstract: The use of online course material is the approach adopted by most universities to support students' revision, and teachers usually have the responsibility for designing or uploading online materials on their own course websites. However, some teachers might lack programming skills or motivation, and most current online materials are just uploaded in a static format (such as PDF) which is not suitable for all students. Moreover, during revision periods students may be faced with a lot of unorganised materials to be revised in a short period of time, and this can lead to an ineffective revision process. In order to address these issues, this paper proposes a software framework that aims to maximise the benefit of current online materials when used to support student revision. This framework is called *SRECMATs* (Self-Revision E-Course MATerials) and has been deployed as a tool that allows teachers to automatically create an intelligent tutoring system to manage online materials without any programming knowledge, and to support students to navigate easily through these online materials during their revision. This paper evaluates the proposed framework in order to understand students' perceptions with regard to the use of the system prototype, and the results indicate which features are suitable for providing online revision materials as well as confirming the benefit of the revision framework.

1 INTRODUCTION

Within higher education, online learning has become an important focus, and many educational institutions consider deploying an online course material delivery system as an essential major development. Williams (2002) mentioned that the transformation and redesign of course materials for the electronic environment needs to ensure that the benefit are fully exploited. In addition, Forsyth (2014) argues that many institutions have not considered carefully the presentation of online materials, thus reflecting a lack of commitment to effective online delivery. This has resulted in many online course materials having been simply uploaded online with none of the features needed to encourage rich learning in a computer environment.

However, designing an intelligent system to deliver online materials which will benefit the self-learning experience requires not only knowledge about the content of materials but also programming skills (including program design). Most teachers lack such skills, and simply present static materials on a course website, and tis is an issue which leads to limited teaching and learning outcomes.

During the revision period, Entwistle and Entwistle (1991) discovered that students have their own different strategies for reviewing online materials.They sometimes suffer from a large volume of learning materials to be reviewed in a short period of time and from poor learning resources uploaded by the teachers as the result of the survey (Sajjacholapunt and Joy, 2014). By *poor learning resources*, we mean that the materials may not be well-organised, or have insufficient information or unclear explanations. In addition, Nicol et al. (2005) found that during the use of online library resources many students cannot develop effective search techniques and require some support from library professionals. These issues can cause inefficient revision, which is time-consuming or has content which supports a low level of understanding.

This paper, therefore, attempts to fulfil the needs of such teachers and students by proposing a software framework which aims at reducing teachers' workloads when it comes to developing and designing an intelligent system to organise online materials for revision. Teachers do not need to worry much about programming issues, since the proposed system will automatically organise materials after they are uploaded. The system will support students' revision by providing alternative ways of navigating through online materials. Students are given the opportunity to browse, search, and navigate across integrated online learning materials in one place and in a much richer way than is generally the case at present. The idea of integrating learning materials is based on the previous survey results from Sajjacholapunt and Joy (2014) which confirms that students need support in terms of organising learning materials when they are faced with the need to review a great deal of such material in a short period of time.

A literature review does not point to effective ways of designing appropriate features for an intelligent tutoring system to use to deliver course materials. This has left us with a number of challenges when it comes to investigating potential features of online tools that students need for exploring online materials during the revision period. In order to validate the software framework, we thus developed a system prototype and launched it in a first-year undergraduate course "Design and Information Structure". We hypothesised that the proposed functions of the software would enhance the quality of e-materials in terms of supporting a greater range of revision strategies and allowing students to easily and accurately navigate through a substantial volume of online materials.

The background to designing the *SRECMATs* framework is discussed in the next section following by a presentation discussion of the *SRECMATs* framework, an evaluation of the system, and finally a conclusion.

2 BACKGROUND STUDY

In this section, we discuss related work which has led to the framework, including students' use of e-learning resources and potential types of cognitive tools that can be applied.

2.1 Revision Strategies

Marton and Säaljö (1976) have proposed a form of understanding which uses a *strategic approach*. This approach appears to be used at the revision stage when students are required to plan time and techniques to go through the materials before an exam. In order to design a software framework to support students at this stage, we have to understand their common strategy for revision. We have conducted the survey to understand the pattern of students using e-learning resources during their revision (Sajjacholapunt and Joy, 2014).

The result of the survey has been used to construct the revision framework as presented in Figure 1. Within this framework the majority of students commonly start their revision with lecture slides, and then gain further detailed information from other materials, especially their lecture notes and past exam papers. If they do not understand any point in the material, they prefer to search for specific information on the Internet or ask their friends. This confirms the results of previous studies (Weinreich et al., 2008; Nielsen, 2008; Nicholas et al., 2009) in which students have been observed seeking and skimming for the specific information that they need.

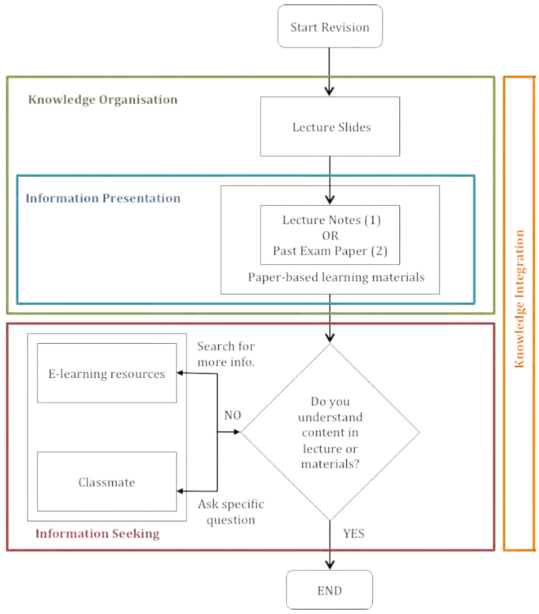

Figure 1: Framework for a common process of using e-learning resources for revision (Sajjacholapunt and Joy, 2014).

Having studied the common patterns of revision, this research focuses on potential tools that can support these processes. Cognitive tools are considered and discussed in the following section.

2.2 Cognitive Tools for Revision

Cognitive tools, based on the definition of Kim and Reeves (2007), are technologies that support the knowledge construction of the learner. Orey (2001) summarised roles of cognitive tools into four categories (modified from those of Iiyoshi et al. (2005)) which are knowledge organisation, information presentation, information seeking, and knowledge integration. These roles have been mapped into each stage of the revision framework based on their following definitions.

- **Knowledge organisation**: supporting a student to establish a conceptual relationship between information by structuring or restructuring information. This is the role adopted by the semantic network tool Mindtools (Jonassen and Carr, 2000).
- **Information presentation**: allowing students to present data from a different perspective which enables a different aspect for them. For example, Nardooo (Harper et al., 2000) is a tool for ecology learning in which student can view information in different forms, such as graphical, video, audio.
- **Information seeking**: supporting a student to find specific information through different learning situations, such as recommendation systems or search engines.
- **Knowledge integration**: helping the student connect existing knowledge to new knowledge and build a larger database of information. For example, WISE (Slotta and Linn, 2000) is a tool that allows a student to organise idea based on information they collect from the web.

From the literature review, the common issues for using online materials involve the readability and navigability of such materials (Hornbæk and Frøkjær, 2001). The software framework therefore supports the basic cognitive load tasks of a student which are knowledge organisation, information presentation, and especially information seeking. This corresponds to the benefit of cognitive tools stated by Shim and Li (2006) in which cognitive tools can share a student's cognitive load by providing support for the lower level of cognitive skills, leaving the students to concentrate on higher order thinking skills. The features and functionalities of the proposed system are discussed in the next section on the *SRECMATs* Framework.

3 SRECMATS FRAMEWORK

SRECMATs is a web-based tool that operates through a web browser. The software framework behind the tool can be divided into two main parts – front-end services and back-end services – as presented in Figure 2. The front-end services are dealt with a user interface for both teachers' and students' interaction. The back-end services are technology components that drive the system from behind. Details of all components are discussed below.

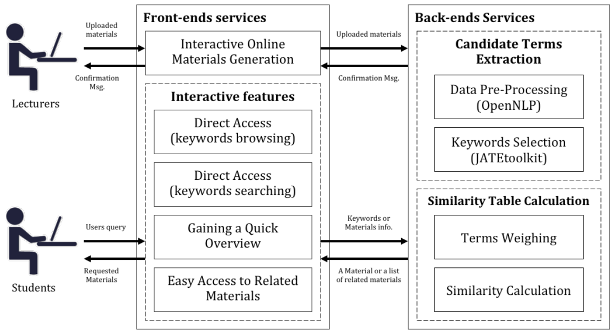

Figure 2: The software framework for designing interactive self-revision course materials (*SRECMATs*).

3.1 Front-End Services

The front-end services relate to the user interface and the functionality of the tool. *SRECMATs* provides the user interface for both teachers and students. The interactive online materials generation component for teachers contains a form for them to fill in with regards to the course information, including the upload button in order to generate folders and databases as storage for online materials. Teachers only need to fill the course information and upload all the materials into the system. The system will automatically process all the materials and publish to students in a "polished" form. In this way, teachers do not have to be involved with any coding.

For students, four interactive features are classified into three categories, based on roles of cognitive tools presented in section 2.2 and the revision framework (Figure 1), which are (i) direct access through e-materials by using keywords browsing and keywords searching (information seeking), (ii) quick overview using keywords (information presentation), and (iii) easy access to the related e-materials (information seeking and knowledge organisation). The design of the user interface is based on the clean and flat concept proposed by Page (2014) using the Bootstrap framework [1]. This is to ensure that students can easily use the system without any training. Details of each feature are discussed in the following subsections.

3.1.1 Direct Access

Sometimes students require quick access to a particular page of a set of lecture slides for recall purposes. With regard to a traditional course website, students need to access the appropriate (PDF) file and scroll down the pages until they find the information they want. *SRECMATs*, therefore, addresses this problem by providing two ways of direct access to a specific

[1] http://getbootstrap.com

page in a set of materials. These are (i) keyword browsing as illustrated in Figure 3, where keywords are extracted from content located at the header of the documents, and (ii) keywords searching as illustrated in Figure 4. All materials are converted to plaintext, which allows students to perform partial searching for terms in the document.

By direct access to e-materials using keyword browsing and keywords searching, a student would obtain a quick overview of the e-material content through a set of keywords, and easy access to related materials based on what they focus on.

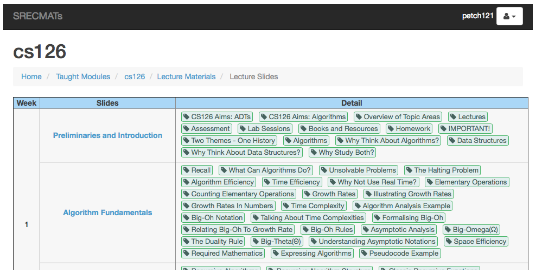

Figure 3: Direct access by using keywords browsing.

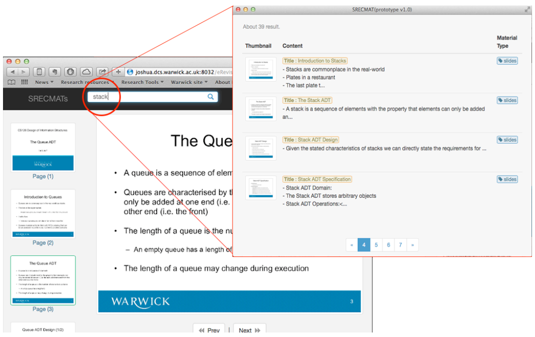

Figure 4: Direct access by using keywords searching.

3.1.2 Gaining a Quick Overview

Nicholas (2010) mentions that skimming and scanning are two important strategies for speed-reading, and reading materials line-by-line to find key information may be time-consuming for some students. Whilst a student is navigating through a set of materials, *SRECMATs* provides a list of keywords as illustrated in Figure 5 (on the right-hand side). This feature allows students to quickly grasp the main ideas in the material.

3.1.3 Easy Access to Related Materials

During the revision period, students frequently switch between different materials to gain insight with regard to a specific topic. Using a traditional website,

Figure 5: Quick overview through a set of keywords.

students have to spend time finding material as well as information within such material. By integrating all the material, *SRECMATs* allows students to switch to other materials using an icon on the top-right corner. Moreover, *SRECMATs* also provides recommendations with regard to features which allow students to navigate through related material based on what they are currently focussing on. This feature supports a student when (for example) they are reading a past exam paper and need to find specific information to answer a question. The system also provides related materials in the form of (i) lecture slides, (ii) other related past exam papers and (iii) lab sheets as presented in Figure 6. These related materials are ranked based on a similarity score. Details of ranking are discussed in the next section dealing with back-end services.

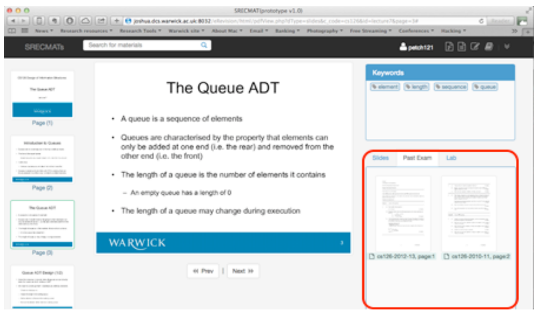

Figure 6: Accessing to related materials through recommendation feature.

3.2 Back-End Services

The back-end services contain components that are used to run the front-end services. The HTML language, CSS bootstrap framework, and pdfViewer libraries are used to construct the front-end user interface. Direct access and keyword browsing and searching are handled by a relational (SQL) database.

For the easy access to related materials, Natural Language Processing (NLP) techniques are applied

(Baeza-Yates and Ribeiro-Neto, 2011) for automatically extracting technical terms and for calculating similarity between materials. The choice of NLP techniques is because most uploaded course materials (on a course website) are provided in PDF format that can simply be converted to and processed as plaintext. Although other advanced techniques such as semantic web are promising, they require a system or an expert to convert e-materials (PDFs) into a rich format (e.g., RDF). However, current tools do not easily support such a process and not every teacher would prefer to do it manually. In this paper, we considered only lecture slides, past exam papers, and lab sheets, because these are the first common materials that students tend to review during their revision. The summary of the process and technologies that have been selected and used for the back-end services (Sajjacholapunt and Joy, 2015) are discussed below.

3.2.1 Candidate Term Extraction

Candidate terms extraction is a process for extracting candidate technical terms from online materials. This process is composed of two sub-processes: data pre-processing, and keyword selection.

Data Pre-Processing The data pre-processing stage prepares all materials to suit the requirements of the system. Details of the methods used in this research are presented in Table 1.

Table 1: Common pre-processing tasks (Sajjacholapunt and Joy, 2015).

Converted Document to Plain Text
Sentence Segmentation
Tokenisation
Part-of-Speech Tagging
Stemming and Lemmatisation
Stop-words Filtering

Keywords Selection Having got a set of candidate terms in the form of sentences, we next need to select which terms will be used in the system. Two main approaches, linguistic and statistical, have been followed in this research as a hybrid approach because using one alone does not provide effective results Pazienza et al. (2005). The open-source JATE-toolkit [2] was used because it can be used in both approaches.

3.2.2 Similarity Table Calculation

Identifying the degree of similarity between online materials is another challenge. In this research, we applied a cosine similarity calculation for calculating the degree of relevance between lecture slides and past exam paper materials as illustrated in Equation 1.

Denote :
Sj = Slide j,
Pk = Past Exam Paper k,
$w_S(i,j)$ = Weight of term t_i in slide S_j,
$w_P(i,k)$ = Weight of term t_i in past paper P_k,

$$Sim(S_j,P_k) = \frac{\overrightarrow{S_j}\cdot\overrightarrow{P_k}}{|\overrightarrow{S_j}|\times|\overrightarrow{P_k}|} = \frac{\sum_{i=1}^{n} w_S(i,j)\times w_P(i,k)}{\sqrt{\sum_{i=1}^{n} w_S(i,j)^2}\times\sqrt{\sum_{i=1}^{n} w_P(i,k)^2}} \quad (1)$$

The term weighing ($w_S(i,j)$ and $w_P(i,k)$) in the equations was obtained from features that are defined as follows:

- Term frequency (TF) is a measure of the frequency of a candidate term that appears in the target document Salton and Buckley (1988).
- Inverse document frequency (IDF) is a measure of the importance of the term provided, in terms of whether the term is common (and appears in most documents), or is rare (and appears only in a few documents) proposed by Salton and Buckley (1988).
- Term location (TL) is the location of a term in a document. The significance of the term can change based on the location. For example, in lecture slides, terms that appear in titles are more important than terms that appear in the body Sajjacholapunt and Joy (2015).

The current F-measure score for using classical TF-IDF methods with weight adjusted by TL component is 22.93% of a similarity threshold of 0.2 and weight 2 (Sajjacholapunt and Joy, 2015). The new term weighting approach is currently being investigated to improve effectiveness of relevant document retrieval.

4 SYSTEM EVALUATION

4.1 Methodology

The *SRECMATs* system was used in 2015 in the first year undergraduate course "Design of Information Structures" delivered by the Department of Computer

[2]https://code.google.com/p/jatetoolkit/

Science at the University of Warwick. It was provided one month before the final examination as alternative course materials for student revision purposes. Normally, students make great use of course materials during their revision. The *SRECMATs* system was introduced to students as an option for reviewing materials online while they *also* had access to the traditional course website which does not provide interactive materials. *SRECMATs* was announced in the course website, and details were emailed to all students on the course.

Data recorded on students who used the system included login information, navigated links, and accessed objects. After they had used it for a few times they were directed to a feedback survey for rating on a scale of 1-5 questions related to how useful the system was, and asked to give some comments. We also asked them to participate in an interview later after they had used the system, however, few students were willing to be interviewed and data from the interviews are not analysed in this paper. After submitting the survey, students were granted full functionality of the system without being disturbed by further surveys. The experiment finished one day after the examination.

4.2 Results and Discussion

Results are discussed into two parts: (i) performance of students' use of the tool, which includes frequency and consistency of use, to evaluate the performance and make sure that there were sufficient instances of tool use to provide valid data, and (ii) students' perceptions of tool use collected from survey responses after students had used the tool. Detail of each results are discussed below.

4.2.1 Performance of students' tool use

The number of instances of students using a tools is one important measurement to indicate that a proposed tool is useful for them or not. Inadequate use of tools results in difficulty claiming the benefits of tool use. In this research, evidence of tool use was considered to ensure that there was a sufficient number of students using the *SRECMATs* tool for evaluating the usability in the next section.

The CS126 course had 132 students. After the system was introduced to the students, 73 students (53% of the population) registered to use the system. However, the actual number of students using the system was 63 (47% of the population).

Figure 7 shows a line chart representing the number of students who accessed the traditional course website compared with the *SRECMATs* system. The trend of students access to both websites is similar in that the number of accessing is increased steadily from the start until it reached a peak one day before the examination. The majority of students were likely to review material just a few days prior to the examination. Only a few of them accessed online materials much earlier.

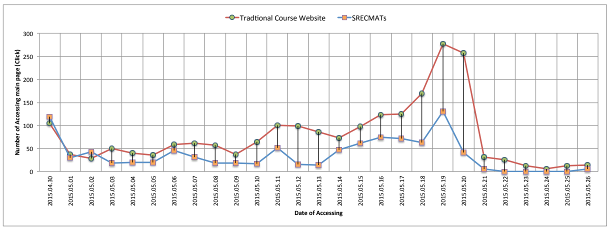

Figure 7: Number of accessing traditional course web (red) and *SRECMATs* system (blue).

The number of students who used the tools and the number of accesses to the *SRECMATs* system are represented in Figure 7. This indicates that some students gained benefit from the *SRECMATs* system and were willing to use it continuously during their revision. This led us to perform a usability evaluation with regard to students' perception in the following section.

4.2.2 Perception of students' tool use

Having finished the examination, students who used *SRECMATs* were asked to complete a questionnaire survey regarding usability. The four front-end services for students presented in Figure 2 were assessed by the "Five Es" usability scheme (Quesenbery, 2001) which considers how *easy to learn* the features, how *effective* the features are in terms of completeness and accuracy, how *efficient* the features are in terms of reducing time spent on the task, how well *error tolerance* of the features are to prevent error from navigation, and how much a student likes *engaging* with these features. Likert scales from 1-5 were used for measuring levels of Five Es in each question (1 - totally disagree, 5 - totally agree). Questions and results are presented in Table 2.

Table 2 illustrates data on the average scores of the four features provided in *SRECMATs*, based on five questions included in the Five Es Usability scheme. These data are also converted into a radar chart as presented in Figure 8, which allows us to simply see the strengths and weaknesses of each feature.

- **Easy to learn**: It can be seen that the average scores of *easy to learn* for all features are higher than 4.0, especially for browsing and searching features which suggest that these all features are

Five E's	Question?	Browsing	Searching	Recommending	Keywords Tagging
Easy to Learn	I can start using this function without any tutorial.	4.42	4.42	4	4.14
Effective	This function allows me to navigate through e-materials easily and precisely.	3.85	3.71	3.14	3.57
Efficient	This function reduces time I spend on browsing e-materials.	3.57	3.85	3.28	3.28
Error Tolerant	I found that this function disturbs my ability to navigate through e-materials	1.85	1.71	2.14	2
Engaging	I prefer to have this function on the course website.	4	4	3.28	3.28

Table 2: Usability survey based on five E's scheme.

easy to use without extra tutorial support being required.

- **Effective**: The average *effective* score for browsing, searching, and keywords tagging are above 3.5 which suggests that students are satisfied with the accuracy and completeness of these features. However, The *effective* score of the recommending system is 3.14, which is the lowest score, and suggests need for improvement in the accuracy of this feature.

- **Efficient**: The trend of the average *efficient* score is similar to the effective score where browsing, searching, and keyword tagging are more efficient than the recommending system. This suggests that the recommending system still lacks the capability to reduce students' time spent on navigation through related materials.

- **Error tolerance**: This score measures how well the features prevent errors from students' navigation? By "errors" we mean that the features can lead a student to a not desired direction or actual unexpected error, which can disturb navigation ability of students. The result showed that all features have average *error tolerance* scores below 2.5 which indicate that most students are not disturbed by navigation through these features.

- **Engaging**: The average *engaging* score of 4.0 (Agree) for keyword browsing and searching show that students are willing to use and prefer to have these features on the course website. In addition, the keyword tagging and recommending systems have *engaging* scores of 3.28 which are a little above the boundary can be implied that these features still need to improve in many ways to attract the students.

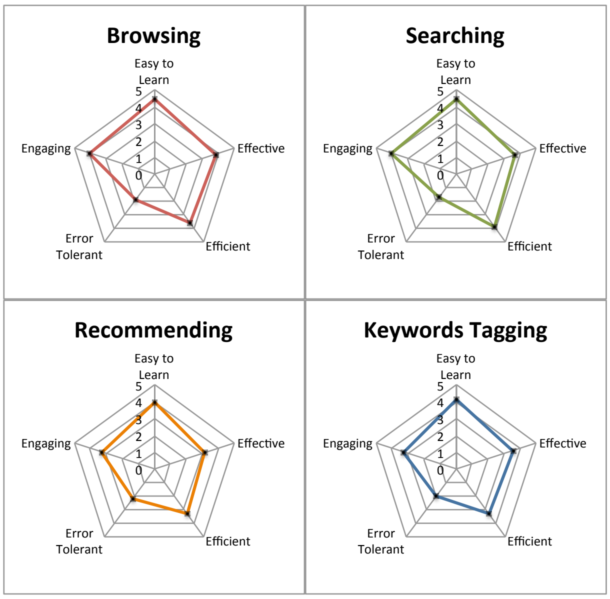

Figure 8: Five Es evaluation scheme to evaluate usability.

5 CONCLUSIONS

In this paper, we have proposed a software framework, and designed an intelligent tutoring system called *SRECMATs*, to deliver online materials for student revision, whereby uploaded materials are automatically indexed and linked together. This is to reduce teachers' workloads as well as to support a lower level of cognitive skills of students when it comes to navigating online materials during revision as presented in the revision framework (Figure 1).

A significant number of the students (47% of the population) registered for and used the *SRECMATs* system. The result of the log file activity shows that students used the system constantly during their revision period, which suggests that some students were satisfy with the tool's features supporting their lower level of cognitive skills. Furthermore, the usability survey shows that all the proposed features are easy to use for students. Most of the students also prefer to have browsing and searching features on the course website for finding online materials. The positive results for instances of use and students' perceptions can be used to conclude that the proposed revision framework is somewhat useful for designing a tool to support revision strategies.

Although most of the Five Es usability scheme for recommending and keyword tagging features are just above the borderline 3.0 score (except the *error tolerance*), these features require improvement especially in term of effectiveness and efficiency regarding the user interface and accuracy of results, in order to increase student engagement with the features. Future work, therefore, will aim at improving the efficiency and effectiveness of the recommending feature.

REFERENCES

Baeza-Yates, R. and Ribeiro-Neto, B. (2011). *Modern Information Retrieval: The Concepts and Technology behind Search (2nd Edition) (ACM Press Books)*. Addison-Wesley Professional, 2 edition.

Entwistle, N. and Entwistle, A. (1991). Contrasting forms of understanding for degree examinations: the student experience and its implications. *Higher Education*, 22(3):205–227.

Forsyth, I. (2014). *Teaching and learning materials and the Internet*. Routledge.

Harper, B., Hedberg, J., Corderoy, B., and Wright, R. (2000). Employing cognitive tools within interactive multimedia applications. *Computers as cognitive tools, volume two: No more walls*, pages 227–245.

Hornbæk, K. and Frøkjær, E. (2001). Reading of electronic documents: the usability of linear, fisheye, and overview+ detail interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 293–300. ACM.

Iiyoshi, T., Hannafin, M. J., and Wang, F. (2005). Cognitive tools and studentcentred learning: rethinking tools, functions and applications. *Educational Media International*, 42(4):281–296.

Jonassen, D. H. and Carr, C. S. (2000). Mindtools: Affording multiple knowledge representations for learning. *Computers as cognitive tools*, 2:165–196.

Kim, B. and Reeves, T. (2007). Reframing research on learning with technology: in search of the meaning of cognitive tools. *Instructional Science*, 35(3):207–256.

Marton, F. and Säaljö, R. (1976). On qualitative differences in learningii outcome as a function of the learner's conception of the task. *british Journal of educational Psychology*, 46(2):115–127.

Nicholas, C. (2010). The web shatters focus, rewires brains.

Nicholas, D., Huntington, P., Jamali, H. R., Rowlands, I., and Fieldhouse, M. (2009). Student digital information-seeking behaviour in context. *Journal of Documentation*, 65(1):106–132.

Nicol, D., Littlejohn, A., and Grierson, H. (2005). The importance of structuring information and resources within shared workspaces during collaborative design learning. *Open Learning: The Journal of Open, Distance and e-learning*, 20(1):31–49.

Nielsen, J. (2008). How little do users read? http:www.useit.comalertboxpercent-text-read.html; 4 May 2009.

Orey, M. (2001). Emerging perspectives on learning, teaching, and technology.

Page, T. (2014). Skeuomorphism or flat design: future directions in mobile device user interface (ui) design education. *International Journal of Mobile Learning and Organisation*, 8(2):130–142.

Pazienza, M., Pennacchiotti, M., and Zanzotto, F. (2005). Terminology extraction: An analysis of linguistic and statistical approaches. In Sirmakessis, S., editor, *Knowledge Mining*, volume 185 of *Studies in Fuzziness and Soft Computing*, pages 255–279. Springer Berlin Heidelberg.

Quesenbery, W. (2001). What does usability mean: Looking beyond 'ease of use'. In *Annual Conference-Society for Technical Communication*, volume 48, pages 432–436.

Sajjacholapunt, P. and Joy, M. (2014). Exploring patterns of using learning resources as a guideline to improve self-revision. *INTED2014 Proceedings*, pages 5263–5271.

Sajjacholapunt, P. and Joy, M. (2015). Analysing features of lecture slides and past exam paper materials - towards automatic associating e-materials for self-revision. *In Proceedings of the 7th International Conference on Computer Supported Education*, pages 169–176.

Salton, G. and Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing & Management*, 24(5):513 – 523.

Shim, J. E. and Li, Y. (2006). Applications of cognitive tools in the classroom.

Slotta, J. D. and Linn, M. C. (2000). The knowledge integration environment: Helping students use the internet effectively. *Innovations in science and mathematics education: Advanced designs for technologies of learning*, pages 193–226.

Weinreich, H., Obendorf, H., Herder, E., and Mayer, M. (2008). Not quite the average: An empirical study of web use. *ACM Transactions on the Web*, 2(1):1–31.

Williams, C. (2002). Learning on-line: A review of recent literature in a rapidly expanding field. *Journal of further and Higher Education*, 26(3):263–272.