Ye Sua: A Tool for Integrating Web Content into an Inquiry-Based Educational Framework

First Author
University of Author
Authortown, CA 94022, USA
author1@anotherco.edu

Fifth Author
YetAuthorCo, Inc.
Authortown, BC V6M 22P
Canada
author5@anotherco.com

Second Author
VP, Authoring
Authorship Holdings, Ltd.
Awdur SA22 8PP, UK
author2@author.ac.uk

Abstract
UPDATED—December 24, 2016. Lack of quality textbooks and good educational resources is a well-known problem in developing regions. In this paper, we describe the design of Ye Sua, a tool that aims to integrate rich educational web content into an inquiry-based framework (5E learning model) for generating web-annotated lesson plans for school teachers in developing countries. Given the wealth of educational resources on the Web, Ye Sua helps teachers to author new lesson plans through a simple web interface allowing them to easily search, select, order and collect web content relevant to specific lesson plans. We investigate the effectiveness of Ye Sua to generate useful lesson plans.

Author Keywords
Education; lesson plan; similarity model, critical inquiry; learning tool.

ACM Classification Keywords
H.5.m [Information interfaces and presentation (e.g., HCI)]: Miscellaneous

Introduction
Over the last several decades, there has been great progress in the west in integrating inquiry-based teaching strategies into the classroom. Simultaneously, there has been an ex-

plosion of online web-based learning materials and initiatives to introduce computers into classrooms The link of the citation is a broken link. Unfortunately, teachers have only a few tools for making productive use of online content within inquiry-frameworks for teaching and learning. This problem is worse in developing regions where lower levels of technological literacy present severe challenges to teachers seeking to incorporate web material in their teaching materials and classroom instruction.

Keeping these points in mind, we designed a web-based lesson plan creation tool, Ye Sua, to help teachers collect and organize web resources within an inquiry-based framework for teaching and learning. We designed Ye Sua to be used in conjunction with the 5E model [5, 2] based on the educational philosophy and psychology of Johann Herbart, which has a long history in educational theory grounded in ideas of Piaget and Dewey. The 5E model for preparing a lesson creation comprises of 5 stages: Engage, Explore, Explain, Elaborate and Evaluate.

Ye Sua enables teachers to express the key topics of their lesson plan and translates the teacher specifications to a list of appropriate and targeted search queries. Ye Sua presents a simple interface for teachers to inspect top search results in each category to select relevant web content for each part of their lesson plan in a well-specified order; collectively, the user-chosen content coupled with the corresponding URLs forms a Web-annotated lesson plan created by a teacher for a specific educational class. Ye Sua supports search, selection and ordering features that enables teachers to develop rich content including video, images, presentations, documents and domain-specific educational resources without being overly complex in terms of user interface elements or additional features.

In this paper we begin with a brief overview of the 5E model, followed by a detailed description of the design and implementation of Ye Sua. Then, we describe a qualitative analysis of the lesson plans for the evaluation of the tool. Finally, we conclude and outline directions for future work.

5E model

We give a brief introduction of the 5E model here to provide the pedagogical context for our work. Building off prior learning cycle models, the Biological Sciences Curriculum Study (BSCS) 5E model is based on educational research tracing back to the early constructivist perspectives of Piaget and Dewey. As a robust philosophical tradition within education, constructivism is fundamentally concerned with how knowledge is constructed and therefore has significant implications for theories of instruction and curriculum development. The 5E model is one of the best known within this tradition [2].

The BSCS 5E model contributes two additional stages, engagement and evaluation, to the Science Curriculum Improvement Study (SCIS) learning cycle which was comprised of three stages: exploration, invention, and discovery [5]. Not intended as a linear formula but rather a guide to structure activity in a way that centers students in the learning process, the 5E learning cycle model consists of five stages:*engage, explore, explain, elaborate,* and *evaluate*.

Engagement - The first stage in the learning requires engaging the students. Engagement is conceived of broadly and generally involves connecting the topic of instruction to the lived experiences of students through culturally relevant pedagogical practices. This may include showing videos, open-ended discussions, freewrites, or class debates designed to ”hook” students into the topic of instruction [5].

Explore - Next, through hands-on activities, labs, or class discussions, students are guided to explore a topic that may have emerged in the engagement phase, or may relate more centrally to the topic of instruction. Explore activities generally help the teachers gain a sense of student prior knowledge, which skillful teachers will take into account as they unfold the rest of their lesson.

Explain - The new ideas generated, and the questions raised during exploration activities will help teachers target their instruction during the explain stage. Often during this step in the learning cycle, teachers will confirm student ideas or help clarify student misconceptions revealed during earlier stages in the process. Concrete activities usually involve direct instruction and presentation of scientific terms or ideas.

Elaborate - The elaboration stage is intended for deeper inquiry into the topic of instruction by challenging students with complex problems and demonstrating real-world applications of the concepts discussed. Common activities may include group problem-solving challenges and group discussions connecting topic of instruction to real-world applications.

Evaluate - Finally, the evaluation stage provides an opportunity for teacher and students to assess understanding and conceptual mastery of the information provided throughout the course of the class.

System Overview

As with any interventionist project and technical artifact, we wanted to follow the best practices in the ICTD development space [3, 4]. Primarily, in our case we wanted to make sure that our system integrated well into the existing educational and ICT ecosystem. In particular, we strived to design an intervention that did not require excessive training, initial cost, or maintenance. Furthermore, we wanted to facilitate existing behaviors and introduce a tool that would speak to the needs of our target population (teachers) and therefore took a collaborative design approach [1] rather than simply implementing and deploying our tool in isolation.

The basic goal of Ye Sua is to enable a teacher who intends to teach a class on a given topic to easily leverage rich educational web content and create a Web-annotated lesson plan for her class. The design of Ye Sua assumes that the teacher has a rough flow of the list of topics he intends to cover in his class and aims to use Ye Sua to find relevant web content that facilitates what he aims to cover in class. The lesson plan, in essence, is a set of curated web contents carefully chosen (by the teacher) to best fit the material covered in class and the web annotated lesson plan can be a useful educational resource for both the teacher as well as for the students in the class. For example, a Computer Science teacher teaching Artificial Intelligence could significantly benefit from content and videos explaining the working of various algorithms, on the web as an educational resource for both teaching her class and as an additional learning guide for her students.

Ye Sua addresses a significant challenge, given the large volume of educational information on the web, how can teachers identify relevant and quality materials that will support student learning in the classroom? Embedded within this challenge is the task of formulating the right type of search query that will result in relevant content and devising a method to return the best results in that relevant content.

After a teacher completes outlining their lesson plan with certain common details, Ye Sua converts the user input into a set of appropriate search queries and returns the top search results across all stages in a common web interface front-end. Our tool provides a simple web interface to enable teachers to sift through the collection of all search results and select the search pages containing relevant data they are looking for and also order them in a priority way. Since the result pages are organized across the different stages, the chosen collection of pages from the result page form the web-annotated lesson plan constructed by the teacher. Next, we describe some of the key features of the implementation and provide screenshots of a few stages in the Ye Sua tool.

Implementation

Ye Sua is a django-based platform, which is a high-level Python Web framework. The major packages used in the tool include django-appconf, django-bootstrap-form, django-user-accounts, Word2Vec, pdfminer, wikipedia and lxml. Django-appconf is a helper class used to handle configuration defaults of django apps, django-bootstrap-form is used to create interactive forms in html, django-user-accounts for creating accounts of teachers, which also handles Login, Logout and Sign in functionalities in the platform, Word2Vec is used in similarity model to calculate similarity between two word vectors, pdfminer, wikipedia and lxml are used in extracting text from pdf, wikipedia and h pages respectively. Apart from these packages, the packages of django include: 'db'which is used to create models in the tool which store information and 'forms'to create form to store information about uploaded material. The main functionality of Ye Sua is to directly interface with the Bing search API allowing us to embed a powerful search engine within the lesson plan creation process. The major components(see Figure 1)of Ye Sua consists of three parts: *Generate lesson plan, Upload lesson plan* and *Search Lesson Plan*

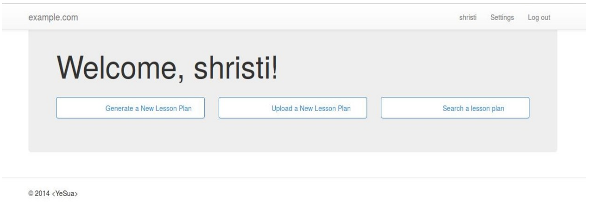

Figure 1: Major components of Ye Sua

Generate Lesson Plan
This part of Ye Sua allows the teacher to generate a suitable lesson plan by providing a certain outline of the required plan. The process is currently divided into four substeps(see Figure 2):

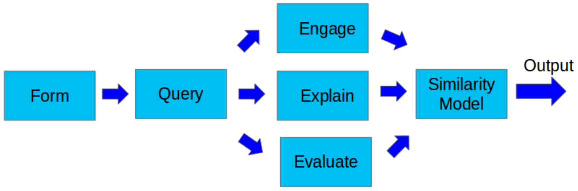

Figure 2: The flow

Form
The form page is the planning page which consists of common lesson plan features like subject, course title, lesson title, outline of the lecture((see Figure 3)). The teacher fills these details which then initiates the lesson plan generation process. Each of the items in the course outline refers to a bullet point version of what the teachers intend to cover in

the respective lecture. Currently the tool generates lessons plans only for computer science and mathematics subjects which are either be of High School or University level.

Figure 3: Ye Sua Planning page

Construction of Search query
Each search consists of a ”query” and a ”source”. The Query parameter is the text of the query the API will execute, and the source parameter is one or more values indicating the type of source, here considered as ”web”. The details entered by the teacher in the previous step, are used to construct a basic search query which will help in obtaining relevant data. For example, when searching for content relevant to a mathematics topic like ”Vector”, adding the word ”vector” as part of the search query may help in obtaining more relevant search results.

Searching Results
After a basic search query is created, the process than divides into three processes, each corresponding to the three phases of learning. In our preliminary tests we found that for specific stages in the 3E model, adding special common keywords like "problems", "questions", "examples", "applications" etc. can significantly enhance the quality of the search results. Our system currently uses a popular search engine as the back-end for issuing search queries and obtaining results.

Engage Phase
In this phase, the targeted results are ”wikipedia” pages and ”applications” of the topic of instruction and so the query is appended accordingly with those keywords separately.

Explain Phase
In this phase, the query is appended with keywords like ”ppt” for powerpoint presentations and ”pdf” to fetch various related concepts of the lesson.

Evaluate Phase
In this phase the query is modified with keywords like ”exams”, ”solutions”, ”midterm” which could return exam papers and homeworks along with their solution if available, from online available courses or related courses from various ”edu” i.e. educational sites.

Parsing results and Similarity Model
After the results are fetched, they are parsed and the best web pages are selected using the summarization search [6]. The similarity model make use of Word2Vec, a group of related models that is used to generate word embeddings. The Search query created in step three is fed to Word2Vec to generate a base vector. Each result returned by Bing API corresponds to a webpage which contains text in form of multiple paragraphs. We have maintained a set of stopwords by using the package stop_words. After removing the words that belong to the set of stopwords, each paragraph is fed to Word2Vec to create a vector space. The Word2Vec are designed such that the related vector are situated at close proximity. Similarity score between then two generated vectors is then calculated. To calculate the similarity score of the whole web page, we used the formula

Similarity score of a page = |avg -10*var|
where ”avg” denotes the average of the similarity scores of the all the paragraphs in the respective web page and ”var” denotes the corresponding variance.

Similarly we calculate the similarity score for other web pages. If the similarity value of a webpage is greater than a comparator value, which is 50th percentile of average of similarity scores of webpages, then we assume that it is a ‘relevant’page. Finally for each phase, ten ‘relevant’web pages with the highest similarity values or the complete set of ‘relevant’webpages if their count is less than ten, are selected and combined to form the desired lesson plan. Associated with each search result, the teacher is provided a selection prompt to determine if he wants to choose the page as a relevant content source for the lesson plan. Along with this, the teacher also has the facility to change the ordering of the search results as per convenience, either in increasing or decreasing order of utility . A lesson plan once created can easily be shared with other teachers and students as a simple web page with 3E organized URLs. Figures 4 illustrates an example of lesson plan.

Upload a Lesson Plan
Another way to share resources with other teachers is by sharing self-made lesson plan in the form of documents and images. To upload a lesson plan, the teacher gives common details of her lesson plan by filling the form as shown in Figure 5 and then upload her material.

Search Lesson Plan
The teacher can also use lesson plans made by other teachers which are available in the database for sharing. The teacher gives the details by either filling all the fields or some fields of the form as shown in figure 3 to generate a search query that will return all the related lessons plans

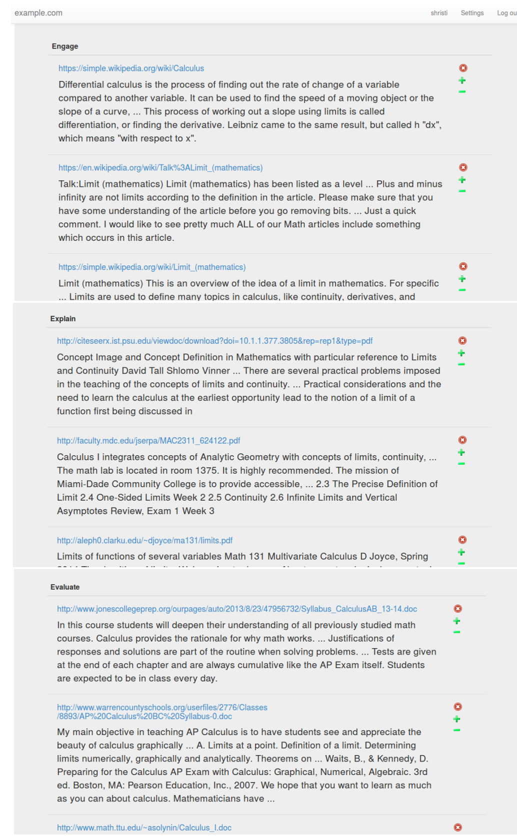

Figure 4: Different phases of the lesson Plan of Limit

Figure 5: Ye Sua Information Page for Searching or Uploading Lesson Plan

along with the name of their authors. For example if someone wants to look for lesson plans in mathematics. The system will then return all the lesson plans with subject Mathematics (Figure 6). The teacher can then view other teacher's lesson plan and integrate resources which he find relevant with his lesson with the aim of creating a better lesson plan. Integration might just not involve including URLs but also documents and images uploaded by other teacher which might appear useful for a good lesson plan.

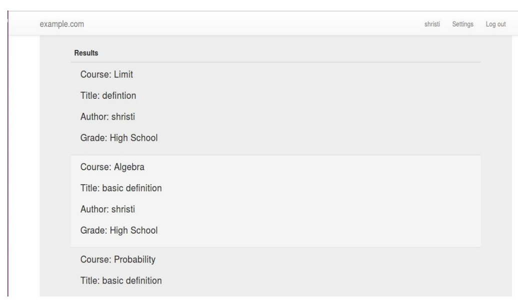

Figure 6: Search Results for 'Mathematics'

Qualitative Results

To explore the quality of content of a lesson plan, we tested Ye Sua across three courses, each divided among three lessons. Assessment of each link in a lesson plan is done across three parameters:

- Broken Link
- Repeated Link
- Usefulness of the link

We take three courses one of Mathematics and two of Computer Science for qualitative analysis, each course is divided into three lessons. The courses considered are *Calculus* for Mathematics, *Algorithms and Artificial Intelligence* for Computer Science. Calculus consists of Limits, Differentiation and Integration as its lessons, Algorithms consists Graph Algorithms, Greedy Algorithms and Network flow while Artificial Intelligence consists Traditional Artificial Intelligence, Adversarial Search and Constraint Satisfaction Problem as its lessons. For each lesson plan, we fixed the number of keywords given in Course Outline field of the form (see Figure 1) to two. On an average with two keywords the system took between 4-5 minutes to generate a lesson plan.

Calculus
Limit
To generate a lesson plan of limit, the details given in the form include 'Math'as subject, 'Calculus'as course Name, 'Limits'as lesson title and High School for Grade. The keywords used in Course Outline were 'Definition'and 'Limit at infinity'. Figure 7 shows the form filled for Limits. The URLs of the generated lesson plan were mostly useful. There were neither any broken link and nor any repeated link. In figures, 10 out of 10 URLs in engage section were useful for teaching, 8 out of 10 in explain section and 4 out of 10

in evaluate section were found useful. Another observation that was made was, the URLs returned by Bing search API in evaluate phase were either of two type: doc type and web links. Only 4 out of 16 returned URLs from Bing search API were found useful, none of doc links were relevant. Figure 4 illustrates the lesson plan of limit.

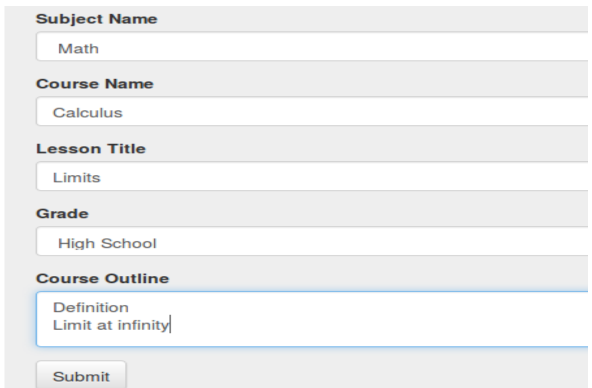

Figure 7: Form to create lesson plan of Limit

Differentiation

For the lesson plan of differentiation, the details were similar to that of limit except in Lesson Title 'Differentiation'was used instead of 'Limits'. 'Chain Rule'and 'Application of Differentiation'were used as keywords in Course Outline. Again, in the lesson plan generated by Ye Sua, no broken links were observed. But there were links with different web addresses but when clicked, directed to the same page. In figures 7 out of 10 links were found useful in engage section, 7 out of 10 in explain section and 3 out of 10 in evaluate section. Similar observations as in limit lesson plan were made for evaluate section in terms of the relevance of doc links, being returned by Bing search.

Integration

For the lesson plan of Integration, the details were similar to that of limit except in Lesson Title 'Integration'was used instead of 'Limits'. 'Definite Integral'and 'Indefinite Integral'were used as keywords in Course Outline. The lesson plan did not contain any broken links and repeated links. In figures, 9 out of 10 in engage section were found useful, 3 out of 10 in explain section and 2 out of 10 in evaluate section. When we analyzed the results returned by Bing search API, it was found that in explain section 12 out of 20 were found useful, the irrelevant links were mostly course structures of a similar or related course, offered in some university. In evaluate section, the results are worse as only 11 out of 30 links returned by Bing search were useful for teaching, where all doc links which were mostly course outline of related courses and thus failed to provide relevant content for teaching.

Algorithms

Graph Algorithms

For the lesson plan of Graph Algorithms, the details given in the form were 'Computer Science'as subject, 'Algorithms'as course Name, 'Graph Algorithms'as lesson title and High School for Grade. The keywords used in Course Outline were 'Depth First Search'and 'Breadth First Search'. Figure 8 shows the form filled for Graph Algorithms. In the lesson plan, there were no broken URLs. There was a pair of link with different address which directed to the same page. The number of useful links in Explain and Evaluate phase of the lesson plan were quiet low. In figures: 7 out of 10 links in engage section were useful, 3 out of 10 in explain section and 2 out of 10 in evaluate section were useful.

Greedy Algorithms

To get the lesson plan of Greedy Algorithms, the details were similar to Graph Algorithms except in Lesson Title

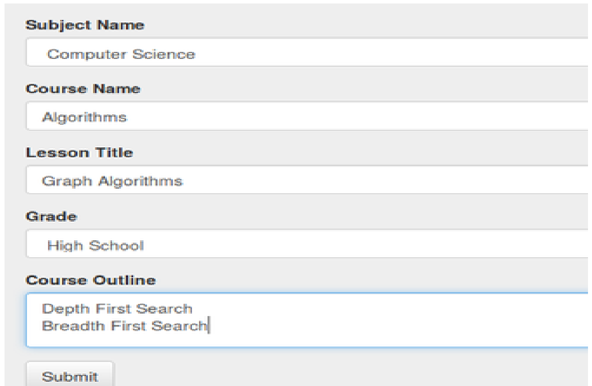

Figure 8: Form to create lesson plan of Graph Algorithms

'Greedy Algorithms'was used instead of 'Graph Algorithms'. 'Interval Scheduling'and 'Minimum Spanning Tree'were used as keywords in Course Outline. In the lesson plan,there were no broken and repeated links. Like Graph Algorithms, the lesson plan also did not contain much useful links. In figures, 8 out of 10 in engage section, 4 out of 10 in explain section were found useful. In evaluate section none of the 10 URLs appeared to have relevant content for teaching.

Network Flow
To get the lesson plan of Network Flow, the details were similar to Graph Algorithms except in Lesson Title 'Network Flow'was used instead of 'Graph Algorithms'. 'Ford-Fulkerson Algorithm'and 'Edmonds-Karp' were used as keywords in Course Outline. Similar to other lesson plans there were no broken and repeated links. In figures: 9 out of 10 links in engage section, 4 out of 10 in explain section and 1 out of 10 in evaluate section were found useful for teaching. When the Bing search results were analyzed, it was found the most of the Bing search results were irrelevant for teaching.

Artificial Intelligence
Traditional Artificial Intelligence
For the lesson plan of Traditional Artificial Intelligence, the details given in the form were 'Computer Science'as subject, 'Artificial Intelligence'as course Name, 'Traditional Artificial Intelligence'as lesson title and High School for Grade. The keywords used in Course Outline were 'Greedy Hill Climbing'and 'Simulated Annealing'. Figure 9 shows the form filled for Graph Algorithms. In the lesson plan, there were no broken and repeated links. This lesson plan contained the least number of useful URLs among the lesson plans searched so far. In figures: 2 out of 10 in engage section, 4 out of 7 in explain section and 2 out of 8 in evaluate section were found useful. Despite using correct keywords in course outline but inappropriate lesson title, the Bing search itself could return much results. This points to the precision required in the use of keywords for generating lesson plans. In figures, the analysis of URLs returned from Bing search, 7 out of 48 URLs in engage section, 5 out of 14 URLs in explain section and 4 out of 19 URLs in evaluate section were found useful.

Adversarial Search
To generate the lesson plan of Adversarial Search, the details were similar to Traditional Artificial Intelligence except in Lesson Title 'Adversarial Search'was used instead of 'Traditional Artificial Intelligence'. 'Mini-Max'and 'Alpha-Beta Pruning'were used as keywords in Course Outline. In the lesson plan, there were no repeated links but one broken links was found. In figures, 7 out of 10 URLs in engage section were found useful, 4 out of 8 URLs in explain section and 4 out of 4 URLs in evaluate section were found useful. When we analyzed the URLs returned by Bing search,

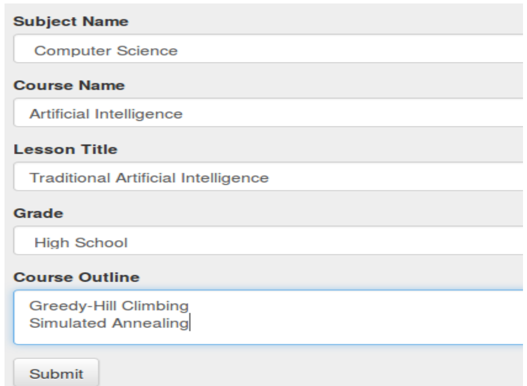

Figure 9: Form to create lesson plan of Traditional Artificial Intelligence

it was found that in explain section 7 out of 19 URLs were useful and 4 out of 12 URLs in evaluate sections.

Constraint Satisfaction Problem
To generate the lesson plan of Constraint Satisfaction Problem, the details were similar to Traditional Artificial Intelligence except in Lesson Title 'Constraint Satisfaction Problem'was used instead of 'Traditional Artificial Intelligence'. 'Map Coloring'and 'Back-tracking search'were used as keywords in Course Outline. In the lesson plan, there were no repeated and broken links. In engage section 8 out of 10 URLs were useful, 5 out of 7 URLs in explain section and none of the 10 URLs in evaluate section were useful. For this lesson plan, the Bing search did not return much useful URLs. In evaluate section, only 2 out of 25 URLs were found useful.

Conclusion
In this paper, we present Ye sua as a tool built to help teachers more easily leverage web-based resources to communicate learning objectives to their students. Our tool was a simple web-based form designed to resemble traditional lesson-plan construction sequences. Adding Summarization Search [6] as a layer above Bing Search API, returned better results than randomly selecting URLs from Bing Search API. In the Evaluate section of each lesson plan in each course, it was observed that the web links were course outline of related courses which also consist links of homework and solutions making them useful for teaching. But the 'docs'pages could not contain such links thus making them irrelevant. For all lessons, the doc links were found irrelevant. We also saw how without the use of proper keywords in planning page while generating lesson plans, the Bing search API fails to return much results and therefore failing to return good URLs. Thus it becomes important to train teachers to efficiently use their existing curriculum materials to fill the first page of lesson plan generation process for creating a good lesson plan. A future direction of this work will to extend the qualitative analysis of the results to multiple user for a fixed number of courses. This will result in a better analysis on the quality of lesson plans.

References
[1] Arias E., Eden H., Fischer G., and Gorman A.and Scharff E. 2000. Transcending the individual human mindâĂŤcreating shared understanding through collaborative design. *ACM Transactions on Computer-Human Interaction (TOCHI)* 7, 1 (2000), 84–113.
[2] Ergin I., Uygar K., and Unsal Y. 2008. An example for the effect of 5e model on the academic success and attitude levels of students. *Journal of Turkish Science Education* 5, 3 (2008), 47–59.
[3] Burrell J. and Toyama K. 2009. What constitutes good

ICTD research? *Information Technologies and International Development* 5, 3 (2009), 82–94.
[4] Li L. and Chen J. 2013. Web Browsing and User Interfaces in Rural Ghana. Information and Communication Technologies and Development (ICTD).
[5] Bybee R. and Taylor J.A.and Gardner A.and Van Scotter P.and Carlson J.and Westbrook A.and Landes N. 2006. The BSCS 5E Instructional Model: Origins and Effectiveness. Colorado Springs, CO: BSCS.
[6] Chakraborty S., Jabbar Z., and Subramanian L. 2015. Summarization Search: A New Search Abstraction for Mobile Devices. ACM DEV'15.