Enhancing the Project-Based Learning Experience Through the Use of Live Web Data

Nader Mohamed
College of Information Technology, UAEU, P.O. Box 17551, Al Ain, UAE
Email: nader.m@uaeu.ac.ae

Jameela Al-Jaroodi
Middleware Technologies Lab, Bahrain
Email: jaljaroodi@gmail.com

Imad Jawhar
College of Information Technology, UAEU, P.O. Box 17551, Al Ain, UAE
Email: ijawhar@uaeu.ac.ae

Abstract— Project-based learning (PBL) was proven to be a very useful model to give students hands on experiences and allow them to be active learners rather than passive listeners. In this paper we introduce a tool to help enhance the project development experiences for information technology (IT) students. The main purpose of this tool is to help reuse web information to enable IT projects for students. This tool is called InetRetriever and it can be easily used by students to retrieve, in real-time, any required real information from the web and to implement, execute, and test their projects with real life data. As PBL is becoming an integral part of many information technology (IT) courses, and in many cases real data is essential for many types of projects, it becomes important to make such data available and accessible easily. In various cases, students focus gets shifted from the real objectives of a project when they spend a lot of their time trying to find data or create methods to get them this data. Furthermore, many projects fail and cannot demonstrate their real capabilities when they are tested and demonstrated using small sets of sample data or some fabricated data sets because the students could not include real data available on the Internet. Another possibility here is that students could not even complete their projects because of the difficulties in retrieving the required data and the excess amount of time they spend on that tasks rather than on the real tasks in the project. Therefore, InetRetriever was developed to overcome this obstacle. It was tested in different information technology courses to enable effective and realistic project-based learning. Using this tool we observed increased student interest in the project development and higher levels of interactions and learning.

Index Terms— Project-Based Learning, Information Technology Education, Internet Information Retrieval

I. INTRODUCTION

In the field of Information Technology (IT), many courses rely on practical hands on knowledge and experiences. Students learn many practical topics such as programming, information systems design, software implementations and many more. In various areas course projects, independent projects and senior projects need to be designed and implemented as part of the learning process. Therefore, the concept of Project-based Learning (PBL) as a learning model that utilizes project problems and hands-on practices to enable the learning process is highly utilized in IT programs. In many cases IT projects are complex tasks based on challenging or real-life problems that involve students in problem analysis, problem solving, decision making and solution design, implementation, and documentation [1][2]. This gives students the opportunity to work over extended periods of time (and usually in groups) utilizing concepts learned from one or more courses to produce realistic solutions or products. In addition, it gives students the opportunity to practice and apply the theoretical concepts into practical solutions. The benefits of PBL has been discussed and proven across many disciplines [3]. On the other hand, there are several challenges for the teachers using PBL [4], one of which, that is more relevant to our work, is recognizing situations that make for good projects. More particularly, in IT, projects are about technology and information. Yet it is hard to find realistic projects that can be implemented and evaluated effectively.

IT is one of the educational subjects where PBL plays an important role in the learning process [4]. Most of the IT and computer courses at the higher educational level cannot be effectively taught without giving the students the opportunity to apply and practice the concepts learned in class and to gain more knowledge through active hands-on involvement. PBL offers a good approach to advance and motivate student-centred learning and develop higher-level thinking and judgment skills.

One of the difficulties in IT higher education is mapping the concepts taught in courses to real-life needs and situations. Students need to practice when they can utilize the theoretical concepts they learned in class to solve practical problems around them. Therefore, it is important to select projects that relate at some level to real-life issues and practical applications that would be

usable in real-life. After selecting the right projects, it becomes important to establish clear goals and guide the students to achieve them. However, this proves very difficult when real data and information is required to build and test the project. Many projects such as those related to financial data mining, information notification systems, weather monitoring systems, and decision support systems rely entirely on the availability of real and useful data to evaluate and validate them correctly. Due to the difficulty in obtaining real data, teachers and students either avoid working on this type of projects, which reduces the pool of possible projects for the students and limits the possibilities of learning experiences. Otherwise, many students when taking on this type of projects will resign to using limited samples of data created specifically for their project. Furthermore, if some students decide to venture off and try to get the real data, then they will shift their focus from the real objectives of the project and spend a lot of their efforts on how to get the data, which is not an easy task.

The advent of web technology has meant potentially drastic changes of the teaching and learning process [5]. The web contains huge amounts of information in so many different fields such as weather information, stock prices, currency exchange rates, real-estate information, statistical data in various areas and other useful and real-time information. We developed a new tool called InetRetriever to obtain and reuse web information for student projects. This tool can be used by students to retrieve any publicly available information from the web, and use this information to implement and validate their projects. In this paper, we extend our earlier work discussing challenges that may face students in utilizing web information for completing their assignments and projects [6]. A middleware-based InetRetriever with advanced features is proposed to overcome these challenges. We will discuss the advantages of using this tool by students and provide some experimental results. In addition, we will discuss the overall impact of this tool on students' learning. We validated the usefulness of the developed tool in a number of student projects in higher education in information technology courses.

In the rest of the paper, Section II discusses related work. The concept of utilizing the web as the source of information is discussed in Section III, while in Section IV, the basic features of the InetRetriever tool is discussed. Section V lists some of the challenges that may face students working on projects utilizing web information. Advanced features in InetRetriever that solve these challenges are discussed in Section VI. Section VII discusses the validation of the usefulness of the developed tool while the impact of this tool on student learning is discussed in Section VIII. Section IX concludes the paper.

II.   RELATED WORK

InetRetriever relates to PBL as it is a tool to help in project development and enhances students' experiences in learning. It is also relevant to the different tools and methodologies currently used in IT projects.

Sometimes known as problem-based or case-based learning, PBL provides an active learning environment where students are involved in solving a predefined problem. The steps to reach a solution are planned, researched, and executed by the students, while learning relevant concepts and exploring required techniques and methodologies. A review of research in PBL [7] defines PBL as "a model that organizes learning around projects." The authors then define several criteria defining PBL projects such as projects being central to the curriculum, projects that are focused on central concepts, and that involve students in a constructive investigation. Two main factors form the driving force for the increasing use of PBL [3]. First, there has been a revolution in learning theory that confirmed that doing is much more effective than hearing or seeing. Second, the world has changed and students need skills as well as the knowledge to succeed. PBL provides the means to facilitate students' preparation by actively involving them and giving them the chance to learn skills in addition to the knowledge. PBL is used in many disciplines, and at many levels, from grade school to graduate studies. Several applications were in medicine, science, engineering, and most recently in information technology. For example, this approach has been applied in teaching multimedia courses [8]. Here the course revolves around a multi-phase multimedia project that is assessed based on the outcomes achieved in each phase. Another example describes the use of PBL in a web-based software engineering course [9]. PBL was also described as a tool for the computer science curriculum at the University of Applied Sciences - Technikum Wien [10]. Furthermore, a description of how PBL is used effectively to teach computer programming [11].

PBL in IT is also supported by many tools and methodologies that support active learning and enable the learning process. Particularly many tools are currently available to support higher education in the information technology field. These tools and methodologies were developed mainly to stimulate the student learning process. One example is a new tool for web-based educational system. This tool is based on Service Oriented Architecture (SOA). It aims to achieve interoperability among student applications by utilizing reusable service logic [12]. Other examples are an authoring tool to enable videogame development and enhance gaming concepts education [13], an automatic visualization tool "PGT" for programming education [14], and an education tool for design automation of CMOS cells [15]. There are also some efforts to create a cyber infrastructure for some fields in information technology to provide education and research tools for students and researchers working in the field. nanoHUB.org is a network of universities supporting the National Nanotechnology Initiative by bringing computational tools with their education material online, making these tools easy to use [16]. This allows students to utilize the

available tools for implementing their nanotechnology projects.

Generally, the web plays an important role in enhancing learning activities in this century by providing a number of educational tools and approaches [17][18]. In this paper, we provide a tool to utilize the web for retrieving live information required to complete students' assignments or projects and achieve the main course objectives.

III.   WEB INFORMATION

Generally, there are two types of projects in IT related courses. The first type is projects based on problems that do not need real information to be developed and tested. In projects of this type the design, implementation and evaluation can be completed with no or limited data such as in computational problems or projects driven by specific user inputs. On the other hand, there is a second type of projects that cannot be developed or validated correctly without having real data. Some examples are data mining projects, and financial analysis and decision making projects. In this case, students need to have access to large amounts of data to execute their projects and verify their design. Currently most of this data is available in an unstructured or semi-structured form on the Internet, which makes it really hard to collect them in real-time. Therefore, it becomes very hard for the students to achieve the project goals efficiently. To solve this problem, we utilize the web as a source of real information for student projects and provide them with a tool to make this doable. The web provides huge information related to stock market information, currency exchange rates, temperatures, oil prices, gold prices, interest rates, sales figures, and much more. Most of this information is available and updated in real-time it is a rich source for the students to use to implement and validate their IT projects when the right tool to retrieve and reuse it is provided.

The web information is usually available in dynamic HTML documents [19], XML documents [20], web services, or RSS feeds [21]. If the information needed by a student project is available in web services, then the student can easily use the corresponding web service which provides the required information. Web services can provide a structured and simplified way to obtain services or specific information from the web. Web services provide web APIs that can be accessed over a network, such as the Internet, and executed on a remote system hosting the requested services. The main problem is that not all types of information available over the web are provided by web services. Most of the useful information on the web is still only available in semi-structured HTML documents.

Unlike XML documents, HTML documents do not have any semantics for their data. HTML documents usually contain tags, scripts, links, and user defined data. Obtaining specific data from a dynamic HTML document for reuse in other applications can be a complex task. It is very difficult to identify the required data components and dynamically use them in other applications.

Some research was conducted to benefit from the web HTML documents. One example is developing an approach to link the large amounts of data that are currently available in HTML documents to the Semantic Web ontology [22]. Another example is developing an approach that automatically captures the semantic hierarchies of HTML tables [23]. Some research effort was also conducted to transform HTML documents to another format to satisfy specific applications. One example of this transformation is from HTML Product Catalogues source code and images to RDF [24]. In contrast, this approach is developed to be used for generic purposes. This approach is designed to simplify the process of extracting real-time information from HTML documents. It provides several techniques to extract the information from these documents regardless of how it is formatted (e.g. in tables or as plain text). These techniques can be used to extract online information and reuse them in other applications such as student projects.

IV.   INETRETRIEVER: BASIC FEATURES

We have recently developed a simple and efficient approach for retrieving live HTML-based web information [25]. This approach can be used to retrieve updated data from the web. The main idea of this approach is based on finding fixed titles or headers that appear in browsers for HTML documents directly or semi-directly before the needed dynamic information. These fixed titles or headers are used as references to know the position of the required dynamic information. The developed approach provides a simple and efficient technique to retrieve any required public information students need to complete and test their projects.

The proposed approach is developed as a Java class. Multiple objects can be created form this class for different web HTML documents that contain some of the required information. A number of techniques were developed to find this information in any HTML document. These techniques are implemented in a set of methods listed in Table 1. All these techniques can be used to retrieve updated web information. As soon as the fields are identified the student can specify the arguments for the *get* or *getWI* methods, which will allow the student to retrieve the required information. Students can use any HTML documents on the web to obtain any numerical information they need.

The first technique is to find information directly after a specific text header. For example, in Fig. 1, the Wal-Mart Stores, Inc. Stock information is displayed from the Yahoo Finance site. The user can find the last trade price, the previous closing price, and the trade volume from this page. To get the last trade price, the user needs to call the *get("Last Trade:")* method. The *get* method will search for the title provided, *"Last Trade:"*, and return the next data field after this title field, *"48.10."*

Table 1. Methods to retrieve HTML-based information.

| Method	Description
get(header)	To return the next field directly after the defined *header*. The search starts from the beginning of the page.
get(n,header)	To return the next field directly after the defined *header* appears *n* times. The search starts from the beginning of the page.
get(n,header, i)	To return the field after skipping *i* fields after the defined *header* appears *n* times. The search starts from the beginning of the page.
getWI()	To return the next field from the current read pointer position.
getWI(i)	To return the field after skipping *i* fields from the current read pointer position.
getWI (header)	To return the field located directly after the specified *header* from the current read pointer position.
getWI(n, header)	To return the field after the occurrence of the *header n* times from the current pointer position.
getWI(n, header, i)	To return the field after skipping *i* fields after the defined *header* appears *n* times from the current pointer position.
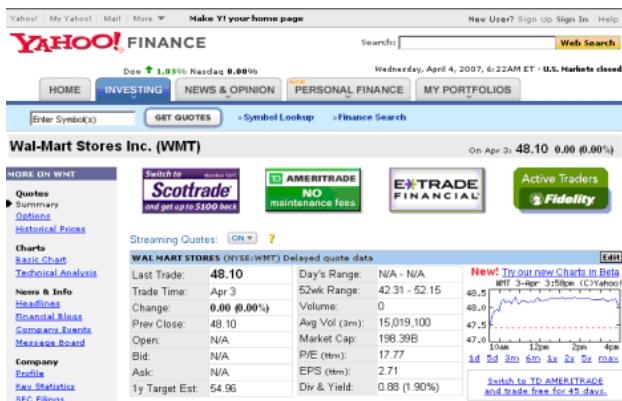
Figure 1. Wal-Mart Stock Information form Yahoo Finance.

The second technique is to find certain information after the appearance of a specific header for the nth time. For example, Fig. 2 is web information for RAK Properties with symbol *RAKPROP* that is listed in Abu Dhabi Securities Market. We can see that, the bid volume amount appears after the *"RAKPROP"* header while that header appeared twice in the page. To get the bid volume, this technique will search for two *"RAKPROP"* text headers before returning the bid volume amount. This technique is implemented in another method with interface *get(n, header)*.

The third technique is to find information semi-directly after a specific fixed title or header. This finds the ith information field after the appearance of a specific header n times. For example in the *RAKPROP* example, the user wants to get the current price of the stock. This price is listed 7 fields after the second appearance of

"RAKPROP" header. The interface for this method is *get(n, header, i)*. Users can use any HTML document on the web to retrieve their needs.

Figure 2. RAKPROP Quote from Abu Dhabi Securities Market.

In the implemented java class, there is a single constructor with a single parameter in that class. This parameter is used for defining the URL of the required webpage. The class contains the public method *download()*. This method is used to download/ re-download the HTML document. Part of this method is a filtering process. This filtering process takes the downloaded webpage and generates another page in the system's memory that only contains the page data without any html tags. The output page contains only the application text headers and numbers. Each application data between HTML tags is considered a single data field. The aim of this step is to simplify the process of finding the needed information. Fig. 3 shows the main components of the proposed approach. The user can recall *download()* method any time that he/she wants to refresh the downloaded information. After calling the *download()* method, the user can start extracting any information from the page. A partial output of the filtering process for the Wal-Mart Stores Stock Information Displayed form Yahoo Finance is shown in Fig. 4.
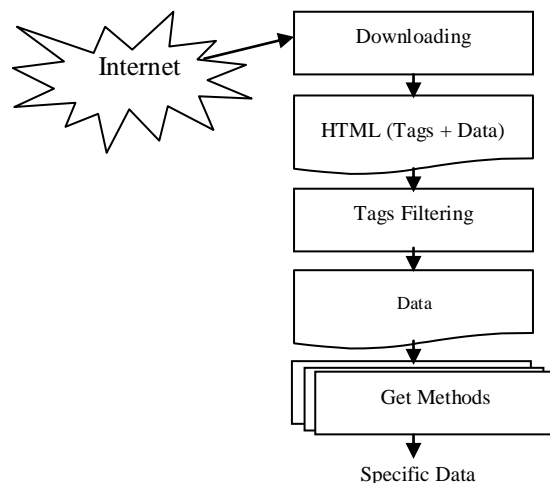

Figure 3. The components of the proposed approach.

```
….
Last Trade:
48.10
Trade Time:
Apr 3
Change:
0.00
(0.00%)
Prev Close:
48.10
Open:
N/A
Bid:
N/A
Ask:
N/A
1y Target Est:
54.96
Day's Range:
N/A – N/A
52wk Range:
42.31 – 52.15
                     ….
```

Figure 4. Partial Output of Filtering Process. Only data remains. Each line represents a single data field.

If the *get("Last Trade:")* method is used, InetRetriever will search for the title provided by the user, *"Last Trade:"*, and will return the next data field after this title field. Students can use this defined class and its methods to retrieve web information to complete their IT projects.

V. CHALLENGES

Although the tool described in the previous section can be used directly by students to retrieve wanted live web information, there are still some other challenges that are not covered by the tool and may prevent the students from effectively utilizing the web information for completing their projects.

- The Internet mainly provides information in the form of HTML documents, XML, documents, and web services. On the other hand, the students are more familiar with or need to use applications that use CORBA, RMI, and DCOM to facilitate integration. It is very difficult for example to allow a CORBA, RMI, or DCOM based application to reuse the web Information provided in HTML, XML, or even sometimes through web services. Web information providers do not support CORBA, RMI, and DCOM interfaces through the Internet since as these were designed to use special port numbers that are typically disabled by firewalls. Applications that support web services can directly integrate themselves with the web to get the required information. Web services overcome the disabled ports problem by using the HTTP protocol for communication. HTTP usually uses port 80 which is generally enabled by most firewalls. Unfortunately, not all web information is provided by web services. Furthermore, not all applications can support web services yet. In addition, XML provides some structure to the data made available; however, just as in web services, these are not very commonly used over the Internet. To date,

most of the information is still provided in HTML documents.

- Student applications may require some information that is distributed over multiple web servers located in different places. These servers may support different mechanisms to provide information such as through HTML documents, web services, or XML documents. At the same time, the response times for the requests from the student client application asking for certain information from these servers may differ. Therefore, it becomes very difficult for students to deal with all the heterogeneity in the delivery mechanisms, in the response time, and in the number of the servers. This imposes a great challenge on the student to account for all these differences and ensure efficient operation of the information integration. Although, he/she can spend some time to solve these issues, this may cause him/her to waste a significant amount time solving issues that are far from the main objectives of the project and the associated course.

- The required information provided by HTML documents, XML documents, or web services can be very dynamic. This causes the required information to rapidly change. For some student applications, it is required to capture all changes that occur over time. One example is a stock price displayed in a dynamic HTML page. That price may change every two seconds. At the same time, some student projects may require registering all changes to that price to perform some calculations, analysis or make some decisions. Implementing the methods in the student application to get all changes in some fields in time and keep track of these changes continuously can be a very complicated task.

- Some web information (especially that is available on a single web page) may be needed for multiple programming assignments or projects developed at the same time. For example, all stock information in a single stock market is displayed in real-time on a single web page. Several student applications may be requesting different stock prices from that same list at the same time. If the extraction is done within the student applications, each one will download and process the same page, while it may be much more efficient to download the page once, process it to extract multiple pieces of information then provide each application with its own required information. However, it is a challenging task to determine the duplication in requests among several independent applications and efficiently reduce the amount of processing required to extract the needed information for each one of the applications.

There are several possible solutions for the challenges mentioned above. These solutions can be implemented as part of the student projects that need to use the web information. However, this approach is inefficient and needs a huge development and testing efforts and a lot of time. This effort may be duplicated by different students

I.J. Modern Education and Computer Science, 2012, 11, 33-43

who need to reuse the web information. Although, the student can spend some time to solve these issues, this may cause him/her to waste a significant amount of time to solve issues that are irrelevant to the main objectives of the project and the associated course. It will be more efficient to have some well developed and appropriately tested independent services that can be efficiently used to obtain the required information by student applications.

VI.   INETRETRIEVER: ADVANCED FEATURES

In this section we discuss a middleware solution to develop advanced InetRetriever services to solve most of the challenges mentioned in Section V. This solution allows students to concentrate mainly on achieving the project and course objectives without getting involved in other technical issues that are not part of the learning outcomes. The middleware connects the Internet as an information source with the student applications that need to use the information (see Fig. 5). This middleware provides some services that can be used by users to configure the required information needed by their applications. The configuration defines the location of the required information.

The middleware provides the following functions:

(1) Establish connections with web servers and web services.
(2) Download and extract the required information from HTML and XML documents.
(3) Present all the required web information in a uniform way such that it can be easily reused by or integrated with the student applications.
(4) Capture changes in highly dynamic web information.
(5) Provide APIs to allow students to easily use the middleware services to retrieve the required information for their project.

The middleware provides the required web information for student applications in three delivery techniques:

▪ **Polling:** in this technique, the required information will only be downloaded and provided when a student application requests it. The student applications are provided with APIs to make the requests. Responses for these types of requests usually take few seconds since the middleware will need to connect to web servers, download and process the web page and deliver the requested information with each request.

▪ **Caching:** in this technique the middleware frequently downloads and extract the required information and keeps them in a local cache. The information cached will be based on the history of requests made by student applications. The cache will contain the latest downloaded information that may be soon needed by the student applications. Student applications can directly read the required information from the cache using the available APIs. This type of read will not take much time from the student applications to get recent information.

▪ **Notification:**   in this technique, the student applications can ask the middleware to send notifications to them when a certain value over the web has changed. The middleware will monitor that value and will only send the notification when the value has changed from the time the request was made. For example, a student application is interested to be notified as soon as the current Google Stock price changed. In this case, the middleware will monitor the Google stock current price from one of the web pages or one of the web services providing this information and will only notify the student application when the price changes. This type of communication request is useful for applications that do not need a frequent access to the web information. It transfers the overhead of frequent web accesses from the student application to the middleware.

The middleware solution addresses several of the challenges discussed in Section V. Interoperability is addressed by providing a middleware framework that may be implemented in several ways such as using Java modules which can operate across different platforms. In addition, the framework is flexible enough to allow for the incorporation of different components. The middleware is capable of handling highly dynamic and changing HTML content. The middleware framework also allows for incorporating multiple sources and servers to be used. Finally, the overall design of the middleware framework is efficient.

The middleware for web information is designed and implemented in three layers. The layers are: The Web Information Retrieval Layer, The Cache Layer, and the Delivery Layer (see Fig. 6).
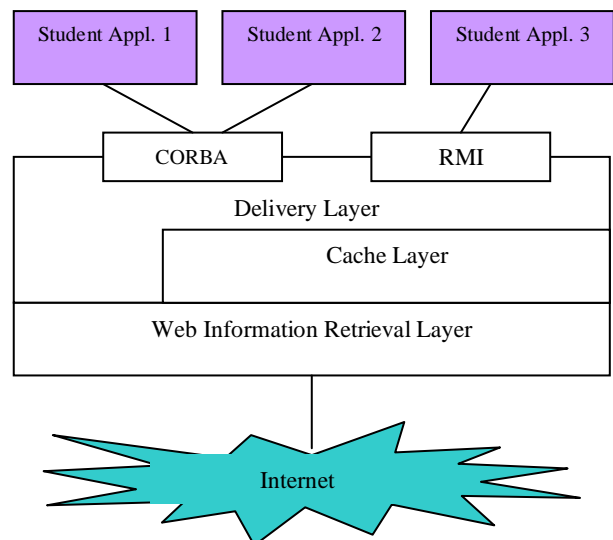

Figure 6. Middleware Architecture.

A. Web Information Retrieval Layer

One of the main layers of this middleware is the web information retrieval layer. This layer directly deals with the Internet web servers. The main function of this layer

is to retrieve the required Internet information. The required internet information can be obtained from the Internet using web services or using any library that provides mechanisms to obtain the required information from dynamic HTML pages as discussed in Section IV. The information is obtained from the Internet in a series of individual requests to be used for serving the polling type deliveries which we discussed in the previous section or as streams to update the cache layer. This layer can also use multiple threads parallelize the retrieval of highly dynamic information from the Internet. This layer retrieves the required internet information based on user configurations.

One of the advanced functions that can be provided in this layer is to discover changes in the structure of the defined dynamic HTML documents. This layer can implement an automatic validation mechanism to allow the middleware to make sure that the formats of the defined HTML documents were not changed before attempting to extract the required Internet information. One possible solution for this problem is to automatically capture and store the format patterns of the HTML documents. These patterns can be used by the system to discover any future changes in the downloaded documents. In case there are some changes, the system notifies the middleware administrator to configure new parameters for the middleware.

Another function that can be provided by this layer is to enhance the performance and reliability of the retrieving process. If a piece of the required information is provided by multiple sites or web services, this layer can discover which site or web service can provide faster access. This can be discovered automatically by the layer using some experimental testing. In addition, this layer can switch from a faulty source or unreliable source to a working source. This function requires that the middleware administrator defines all websites or web services that provide the same information.

B. Cache Layer

The main function of this layer is to provide memory for updated information obtained from the web. This layer will be accessed by the Web Information Retrieval Layer to update the cache with new information and by the Delivery Layer to obtain updates on required web information.

One of the advanced functions that are implemented by this layer is to capture the access patterns of the information in the cache by the student applications. Capturing the access patterns can be used to adjust the required speed of retrieving the required information from the Internet by the Web Information Retrieval Layer. For example, the student applications access a certain value in the cache once every 20 seconds. In this case there is no sense in retrieving that value every 5 seconds. Therefore, the retrieval layer can be informed about that fact to adjust its download accordingly.

C. Delivery Layer

This layer will be accessed by the application to receive Internet information either from the cache or from the Internet directly using the Web Information Retrieval Layer. This layer will also wrap the required web information to a format that can be accessed by the applications. For example, this layer can provide different access methods such as RMI, CORBA, web services, and DCOM. These access methods can be either implemented by the users or using tools to help them in automatically generating servers that use both the cache layer and the web information retrieval layer to get the required information for the student applications.

This layer can combine information collected from different web pages and web services to be delivered as a reply for an application request. The advantage of this function is that instead of making the student applications directly deal with multiple web services and web pages to collect a set of needed information, this layer can provide all required information in one record and reply. This layer also implements the notification services mentioned in this section. This layer will notify the interested applications about any changes in required values.

The middleware described in this section can be configured and used for meeting student applications needs. It can be used to complete students programming assignments and projects that need real-life and updated information from the web. With the supporting advanced features, the students can concentrate more on achieving their main project objectives without spending much time in other unrelated issues.

VII. Experiments of Utilizing InetRetriever

Using the proposed approach, new real-time applications can be easily developed by students. These applications can be integrated with some information from the web. For example, a number of projects related to stock investments were developed using the proposed approach in the Internet System Software course at the College of Information Technology, UAE University, UAE. Internet system Software is a senior course required for the undergraduate networking degree while it is an optional course for the other degrees such as Enterprise Systems, Software Engineering, and Computer Science. InetRetriever allowed students to accomplish these projects that require real information provided in real-time. The projects are:

- **Notification System for Stock Price Changes in Dubai Financial Market:** This system can be used by users to define certain price change criteria. The system will notify the users by emails or by any other message type mechanism whenever those criteria are met. Examples of the criteria are when a stock reaches a certain price level or when a stock price increases/decreases by a certain percentage. The system monitors stock prices through the web. It will

generate notification messages whenever the defined criteria are met.

- **Multiple Stock Markets Monitoring System:** A customer may have stocks listed in multiple financial markets: Although each market provides a software tool for live quotes there is no software tool that displays live quotes for a set of stocks that belong to different securities markets. For example, a customer may have one stock in Abu Dhabi Securities Market, two stocks in Dubai Financial Market, and three in Kuwait stock market. He/She is only interested in monitoring the prices of these stocks. The aim of this project was to implement a networked stock monitoring system in which users can define and monitor the prices of specific stocks in different markets. This system depends on the web for information about different stocks listed in different markets.

- **Monitoring System for Stocks listed in Multiple Securities Markets:** Some investors like to monitor prices of particular stocks that are listed in multiple securities markets to buy or sell them. The aim of this system was to monitor stocks listed in multiple securities markets. In Gulf Co-operation Council (GCC) countries, there are some stocks listed in multiple markets. These markets open and close around the same times. A price of a specific stock listed in multiple markets may drop or increase in one market before other markets. This price difference may only happen for short time periods (e.g. a few minutes) before the stock prices in the other markets adjust with the new price drop or increase. The price drops event in one market is a good indication for near future price drops in other markets for the same stock. In addition, the price increase in one market is a good indication for near future increases in other markets for the same stock. Some investors would be very interested in having a tool that notifies them about that price differences in order to make the right investment decisions to buy or sell stocks. A group of students have already implemented that project using InetRetriever.

InetRetriever was also used for a number of senior exhibition projects at the college. One example is implementing a generic web information notification [26]. The generic web notification is a framework to enable monitoring any type of information on a single web page or on multiple web pages over the web. The students have utilized InetRetriever to implement and evaluate that framework.

InetRetriever helps students to complete their projects and enables effective PBL in IT courses. We compared the average of the effort needed by average students to write programs to retrieve some information from the web with the effort needed for using InetRetriever for the retrieving process. We also surveyed the group of students who developed and tested the retrieving process

by themselves as well as another group who used InetRetriever.

The results are shown in Table 2. It is noticeable that the proposed tool can save the students' time and efforts and relieves them to concentrate more on the assignments'/projects' objectives. In addition, it opens new doors for students to develop and test new types of projects. This motivates students to be creative and innovative in some new types of applications in which it was very difficult for students to develop and validate projects in them. It also gives them the opportunity to learn new concepts and integrate different skills and knowledge in their work. The tool offers a good approach to advance and motivate student-centred learning and develop higher-level thinking and judgment skills.

Table 2. Efforts (in hours) needed to Reuse Web information by average Students.

Function	Students Developed/ Tested Their Method	Students Used InetRetriever
Retrieving single information from an HTML document	5.0	0.5
Retrieving some information from different HTML pages	8.0	0.6
Retrieving a data stream for a piece of information that changes rapidly	14.0	1.0

VIII. IMPACTS ON STUDENTS LEARNING

As mentioned earlier, InetRetriever provides students with an important tool that is at their disposal to efficiently design and implement highly practical and useful applications which can be deployed in many different fields. With the vast increase in the amount of information that is provided on the web in today's world, student are able to design and implement highly interesting and useful applications which take advantage of this information in a timely fashion and are able to use it in many different ways that are only limited by the students' imagination.   Although practical web information have started to be used by student projects to design exciting software products with the introduction of new tools such as web services, this process was still limited by the availability of the required information in the web services framework. However, the introduction of the InetRetriever tool and its various functions opens the horizon and expands the possibilities of designing new practical applications as student projects in areas, and fields that pivot on data that is readily available in the overwhelmingly popular and widely used HTML format. Now, practically every web page on the web can become a source for versatile and timely information that can be compiled, customized, and tailored for student projects.

To evaluate the positive impact of using InetRetriever on students learning, we calculate the achievements of projects that were completed using real-life information provided by InetRetriever and compared with the achievements of other projects that did not use any real-life data. The level of achievements were measured based on the final project grades which represent the level of project requirements completion. We took the results of different projects done over four different semesters. The analysis shows that the InetRetriever projects in average have achieved the project requirements better compared to other projects, see Fig. 7.
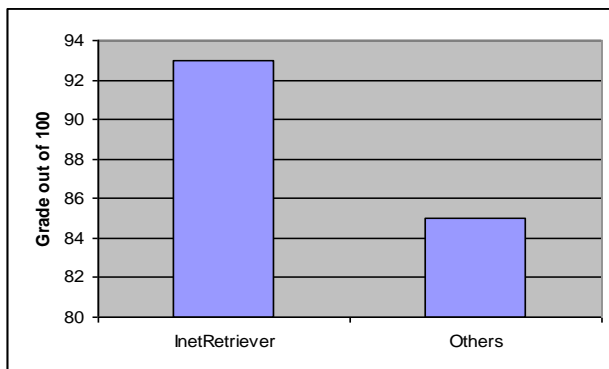

Figure 7. Project achievement levels.

In another analysis, we compared the results of the ABET course outcomes measurements in both cases: InetRetriever based projects and other projects. The course used for measurements is called "Network Application Software." It has five outcomes: Outcome # 1: Build client/server solutions using socket programming, Outcome # 2: Examine tools/software for the development of web applications, Outcome # 3: Evaluate the performance impact of concurrency and synchronization techniques on networked applications, Outcome # 4: Evaluate middleware solutions based on the needs of networked applications, Outcome # 5: Apply analytical and critical thinking in solving networked applications design and implementation problems. The comparison results are shown in Fig. 8. Generally, students who used InetRetriever for their projects that are based on real-life data achieved better course outcomes.
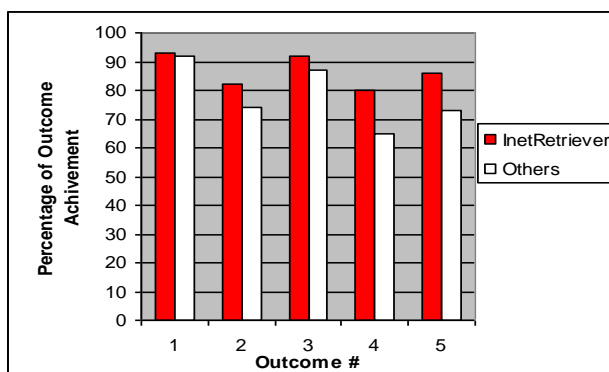

Figure 8. ABET outcomes achievements.

As a result of the possibilities that InetRetriever opens for students several important educational parameters are positively impacted including increased motivation, and enhanced creative thinking. Students are highly motivated by the fact that they can now design applications that are realistic, and address different social and consumer needs. Not only do the students feel that they are able to build something that uses what they learned, but also they realize that what they built can be readily expanded into a profit-making business which has the potential to rapidly grow and have quick and significant impact on society as is the case with many powerful web-based businesses that we see today. This process can have impressive and lasting impact on the student's mind which excites him or her even more to pursue their education in a more enthusiastic, vigorous and successful fashion.

IX. CONCLUSION

For a long time, PBL has been proven to have multiple benefits and advantages for student learning. One of those benefits is that it enhances the quality of student learning and improves their skills. Another benefit is that students become more professional, self-reliant, and collaborative. However, there are also difficulties involved such as difficulty choosing and designing problems or projects that meet the required criteria such as driving question, constructive investigations, autonomy, and realism. In the field of IT realism is a key area to be considered since projects rely heavily on realistic situations and information. Many IT projects require some real-life data and up-to-date information to be completed and evaluated. As a result, it is necessary to provide the tools and methodologies than can help the students create realistic projects. Such tools and methodologies can help make projects more suitable and more realistic thus enhance the benefits of PBL. In this paper, we introduced a new tool (InetRetriever) to enable student projects that require access to and utilization of real information available on the Internet. The tool allows students to retrieve the required data easily thus relieves them from the burden of developing their own tool to do that task or from relying on sample non-realistic data. In either of these cases some of the essential benefits of the projects will be lost. InetRetriever stimulates student creativity and encourages them to develop solutions that cross the boundaries of diverse technologies. It also helps provide an opportunity in which students can utilize existing real-life information that is publicly available over the web to complete their projects, and thus produce innovative ideas and solutions. InetRetriever also enables the development of interdisciplinary projects, where various areas in information technology, business and science may be integrated. It was successfully tested in multiple information technology courses and proved to be effective in enhancing the projects outcomes. The students were able to concentrate on the main objectives and complete their creative projects based on real life solutions and data.

REFERENCES

[1]   B.F. Jones, C.M. Rasmussen, and M.C. Moffitt, *Real-Time Problem Solving: A Collaborative Approach to Interdisciplinary Leaning*, American Psychological Association, Washington, USA, 1997.

[2]   J.W. Thomas, J.R. Mergendollar, and A. Michaelson, *Project-based Leaning: A Handbook for Middle and High School Teachers*, The Buck Institute for Educations, Novato, CA, USA, 1999.

[3]   J. Thomas, A. Michaelson, and J.R. Mergendoller, *Introduction to Project Based Learning. Buck Institute for Education Project Based Learning Handbook*, 1999. URL: http://www.bie.org/index.php/site/PBL/pbl_handbook/

[4]   Intel Innovations in Educations, Harnessing the Power of Project-Based Learning, 2004. At http://www.intel.com/corporate/education/emea/eng/ireland/ elem_sec/tools_resources/plans/harness.pdf.

[5]   LK Soh, X. Liu, X. Zhang, J. Al-Jaroodi, H. Jiang, and P. Vemuri, "I-MINDS: an Agent-oriented Information System for Applications in Education," In Agent-Oriented Information Systems, pp. 16-31, Springer Berlin/Heidelberg, pp. 16-31, 2004.

[6]   N. Mohamed, J. Al-Jaroodi, and I. Jawhar, "Internet Information Retrieval for Enabling Student Projects," in Proc. of The Sixth International Conference on Information Technology: New Generations (ITNG 2009), IEEE Computer Society Press, Las Vegas, Nevada, USA, pp. 987-992, April 2009.

[7]   J.W. Thomas, *A Review of Research on Project-Based Learning*, The Autodesk Foundation, 2000. URL: http://www.bobpearlman.org/BestPractices/PBL_Research.pdf

[8]   F. Kurzel and M. Rath, "Project Based Learning and Learning Environments," Issues in Informing Science and Information Technology, Vol. 4, 2007.

[9]   N. Piccinini and G. Scollo, "Cooperative Project-based Learning in a Web-based Software Engineering Course," Educational Technology & Society, Vol. 9, No. 4, pp: 54-62, 2006.

[10]   R. K. Pucher, H. Wahl, and F. Schmöllebeck, "Integrating Undergraduate Project Based Learning in Computer Science with COAST: A Research Network," In Proc. ICT: Providing Choices for Learners and Learning, 2007.

[11]   R. C. Solamo, "Project-based Learning Approach in Teaching Introductory Course in Programming," Invited talk at the 6th National Conference on Information Technology Education, 2008.

[12]   P. Gurunathan, and S. Pandian, "A New Tool for Web-Based Educational System," In Proc. of the International Conference on Natural Computation, pp. 112-116, 2008.

[13]   J. Torrent, P. Moreno-Ger, B. Fernandez-Manjon, and L. Sierra, "Instructor-Oriented Authoring Tools for Educational Videogames," 8th IEEE International Conference on Advanced Learning Technologies, pp. 516-518, 2008.

[14]   Y. Kita, T. Katayama, and S. Tomita, "Implementation and Evaluation of an Automatic Visualization Tool "PGT" for Programming Education," 5th ACIS International Conference on Software Engineering Research, Management, & Applications (SERA 2007), pp. 213-220, 2007.

[15]   A. Ziesemer, C. Lazzari, and R. Reis, "An Educational Tool for Design Automation of CMOS Cells," IEEE International Conference on Microelectronics Systems Educations," pp. 149-150, 2007.

[16]   G. Klimeck, M. McLennan, S. Brophy, G. Adams III, and M. Lundstrom, "nanoHub.org: Advancing Education and Research in Nanotechnology," Computing in Science and Engineering, Vol. 10, No. 5, pp. 17-23, 2008.

[17]   J. M. Dodero and E. Ghiglione, "ReST-Based Web Access to Learning Design Services," IEEE Transactions on Learning Technology, Vol. 1, No. 3, pp. 190-195, 2008.

[18]   T. Aleahmad, V. Aleven, and R. Kraut, "Creating a Corpus of Targeted Learning Resources with a Web-Based Open Authoring Tool," IEEE Transactions on Learning Technologies, Vol. 2, No. 1, pp. 3-9, 2009.

[19]   RFC 1866, "Hypertext Markup Language – 2.0," 2005.

[20]   Extensible Markup Language (XML) 1.0 (Fourth Edition), http://www.w3.org/TR/REC-xml/, 2006.

[21]   A. King, "The Evolution of RSS," 2003. Available at http://www.webreference.com/authoring/languages/xml/rss/1/.

[22]   R. Burget, "Hierarchies in HTML Documents: Linking Text to Concepts," Proceedings of 15th International Workshop on Database and Expert Systems Applications, pp. 186-190, 2004.

[23]   S. Li, Z. Peng, and M. Liu, "Extraction and Integration Information in HTML Tables," In Proceedings of The Fourth International Conference

on Computer and Information Technology (CIT'04), pp. 315-320, 2004.

[24] M. Labský, V. Svátek, O. Šváb, P. Praks, M. Krátký, and V. Snášel, "Information Extraction from HTML Product Catalogues: From Source Code and Images to RDF," In Proc. of The IEEE/WIC/ACM International Conference on Web Intelligence (WI'05), pp. 401-404, 2005.

[25] N. Mohamed and J. Al-Jaroodi, "A Simple and Efficient Approach for Retrieving Live HTML-based Internet Information," in System and Information Sciences Notes, UK, Vol. 1, No 3, pp. 221-224, July 2007.

[26] N. Mohamed, A. Romaithi, E. Samahi, M. D. Kendi, and E. Jabrie, "Generating Web Monitors," in Proc. of The International Conference on Internet Computing (ICOMP 2008), Las Vegas, USA, pp. 111-116, July 2008.

Nader Mohamed is an associate professor at The College of Information Technology, UAEU, Al-Ain, UAE. He obtained his Ph.D. in Computer Science from University of Nebraska-Lincoln, Nebraska, USA in 2004. He was an assistant professor of Computer Engineering at Stevens Institute of Technology in New Jersey, USA. His current professional interest focuses on middleware, Internet computing, sensor networks, and cluster, Grid, and Cloud computing. He published more than 90 refereed articles in these fields.

Jameela Al-Jaroodi is currently an independent researcher at the Middleware Technologies Lab. She received her Doctorate of Philosophy degree in Computer Science from The University of Nebraska-Lincoln, USA in 2004. Since August 2006, she was an assistant professor at the Faculty of Information Technology, United Arab Emirates University, UAE until August 2010. Prior to joining UAEU, Dr. Al-Jaroodi was a research assistant professor at Stevens Institute of Technology in New Jersey, USA. Currently, her research interests involve middleware, distributed collaborative systems, security, and mobile and pervasive computing. While at Stevens, Dr. Al-Jaroodi received the Research Excellence Grant from Sun Microsystems, Inc. In addition, several areas of her research were also supported by the United States National Science Foundation (NSF), Nebraska Foundation, the National Center for Information Technology in Education (NCITE) and UAEU research grants.

Imad Jawhar is an associate professor at the College of Information Technology at United Arab Emirates University. He has a BS and an MS in electrical engineering from the University of North Carolina at Charlotte, USA, an MS in computer science, and a Ph.D. in computer engineering from Florida Atlantic University, USA, where he also served as a faculty member for several years. He has published numerous papers in international journals, conference proceedings and book chapters. He worked at Motorola as engineering task leader involved in the design and development of IBM PC based software used to program the world's leading portable radios, and cutting-edge communication products and systems. He was also the president and owner of Atlantic Computer Training and Consulting, which is a company based on South Florida (USA). His current research focuses on the areas of wireless networks and mobile computing, sensor networks, routing protocols, distributed and multimedia systems. He is an editor in several international journals in his field of research, and served on numerous international conference committees. He is a member of IEEE, ACM, and ACS organizations.