

# IMPLEMENTATION OF AN ELECTRONIC LAB NOTEBOOK TO INTEGRATE RESEARCH AND EDUCATION IN AN UNDERGRADUATE BIOTECHNOLOGY PROGRAM

---

Rupa Iyer, University of Houston; William Kudrle, University of Houston

## Abstract

Although Electronic Lab Notebooks (ELNs) are becoming more widely used in industrial settings, in the educational domain their adoption has been limited. For the last 5 semesters, the biotechnology undergraduate laboratory at the University of Houston has utilized a custom ELN solution based on a free and easily installable content management system. This study evaluated the implementation of and experiences with this ELN for biotechnology research projects as they were integrated into the educational laboratory courses. Also analyzed was an example of how the ELN might be used to enhance undergraduate research experiences in a classic experimental scenario. Finally, desired general characteristics for ELNs in educational laboratory settings were examined, resulting in an assessment of how well this particular platform met the desired characteristics.

## Background

Computers are routinely utilized in laboratories for controlling instrumentation as well as collecting and analyzing data. In many research and industrial laboratories, these instruments are either directly or indirectly networked with electronic lab notebooks (ELNs) where investigators can further process and analyze data, perform simulations, and store records for use with publications or patents [1]. In many cases where the volume of data is large, ELNs are increasingly becoming a competitive necessity because of the enhanced productivity and data management capabilities they provide [2].

The use of ELNs in university settings is relatively uncommon, with estimates at only 4% overall as of 2008 [3]. The authors of this current study propose that lowering ELN costs and minimizing the need for technical expertise could increase the use of ELNs in education. A more widespread use of ELNs could then potentially provide the following benefits to undergraduate laboratory education:

- Students will be more prepared after they graduate to move on to industrial, research or other lab settings where ELNs are routinely utilized.
- Integrating research into the undergraduate curriculum will become easier by providing access to computer-based tools. Engaging undergraduate students in research activities and including more natural science, math, computer science, and engineering concepts develops greater competency for the interdisciplinary nature of today's research environments [4-5].
- The lab instructor is provided with a single entry point for reviewing data versus multiple lab notebooks that need to be reviewed separately.
- An ELN enhances communication among student researchers in a format similar to popular social networking websites with which most students are already familiar, thus leading to potentially greater engagement of students with research projects.

Two examples of how an ELN based on the freely available content management system (CMS) of Drupal version 6 (drupal.org) has been used in biotechnology education at the University of Houston [6] are presented here. This is followed by an example of how an ELN might be used for the scenario of a classic design of a lab experiment. Possible criteria for ELNs in an educational laboratory setting are considered along with an examination of how well Drupal meets those criteria.

## Implementation in a Biotechnology Undergraduate Laboratory Class

The Drupal-based ELN has been used to enhance the laboratory teaching and research capabilities of students in the biotechnology program at the University of Houston since the fall semester of 2008. The biotechnology laboratory curriculum is an interdisciplinary, research-based curriculum that demonstrates the life cycle of a typical biotechnology product. Using the pesticide degrading bacterium *Brevundimonas diminuta* as an anchor organism, the curriculum follows a logical progression starting with isolation and

identification of pesticide-degrading soil microbes, gene cloning, gene expression, bio-processing of the gene product, and commercial applications [7].

For the lab, a research paper is required where students collect soil samples from various locations in the greater Houston area and follow established protocols for isolation of pesticide-degrading microbes. Since this activity is repeated every time the course is offered, the area covered for soil sampling is potentially larger every year and provides students with a unique opportunity to contribute to a valuable database of pesticide-degrading activity. Providing access to data from their peers and discussion of this activity in their research papers is designed to enhance their research experience. Using Drupal in conjunction with the networking capabilities of the world-wide-web, collaboration on this research potentially connects students locally, regionally, and globally with other laboratories to investigate contaminated soils and further provides an exciting opportunity for students enrolled in the class to become active participants in an ongoing research project. This Drupal ELN is referred to as the Platform for Education and Research Collaboration, or PERC.

In the PERC implementation, each student records a name, location, short description, date, and time of collection for each sample point along with the results of testing the sample for the presence of two pesticides, Methyl Parathion or Paraoxon-degrading bacteria (Figure 1). The data point results can be edited only by the student and those with administrator privileges (usually the faculty and the lab manager), but are viewable by all of the students. Students can compare notes and results to make a more informed analysis of the results of the project as a whole as well as how their own data compare with data from others.

In terms of Drupal modules, the book module was used to present the background and procedures, the CCK module for custom data types, the views module for presentation of the collective data, and the faceted search module for data filtering capabilities. With a faceted search, the students can easily drill down to data points using the filter for semester, for positive or negative results, and student name. To provide a visual display of locations of the data samples, the location and Google maps modules were used to display the sample locations on a geographical map (see Figure 2). Color-coded pins denote the results – yellow for both negative, green for both positive, red for positive Paraoxon, and blue for positive Methyl Parathion test results.

Beginning with the fall 2009 semester, data results were accumulated starting with the greater Houston, TX, area. These data will be available over successive semesters, thus

providing a data set concerning pesticide contamination that can be enhanced in the future to ever-larger geographical areas.

## Cimarron Soil Sample

View Edit Outline

Methyl Parathion Negative, Paraoxon Negative

A sample of soil was collected from Cimarron Elementary. It was tested for OP degrading bacteria and produced negative results.

### Date of Collection:

Wed, 2009-09-02

### Number of colonies of Paraoxon:

0

### Number of colonies on Methyl Parathion:

0

### Location

Cimarron Elementary

1000 South Peek Road

Katy, TX, 77450

United States

29° 45' 53.9028" N, 95° 46' 18.4116" W

See map: [Google Maps](#)

### Semester of Collection:

Fall 2009

### Student Name:

[Redacted]

Figure 1. Data description of one sample location

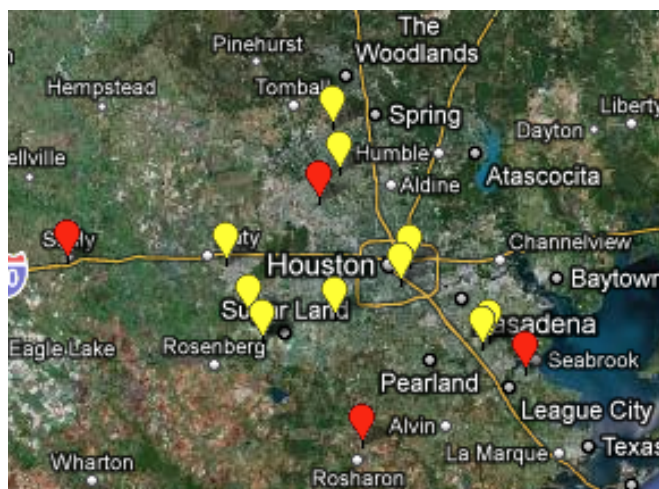


Figure 2. Locations of samples on Google map

A short survey was given to students from the fall 2009 and fall 2010 semesters, a summary of which is shown in Table 1. The columns (A, B, C, D, and E) correspond to the choices of Strongly Agree, Agree, Neutral, Disagree, and Strongly Disagree, respectively. Ease-of-use received the most positive feedback and, for all questions, the majority of the students saw positive benefit in the use of the ELN. In the area asking for suggestions for improvement, the most consistent comment was to introduce the ELN earlier in the semester, which the authors did in subsequent semesters.

**Table 1. Results from the Fall 2009 & Fall 2010 student survey**

Question	A	B	C	D	E
The ELN was easy to use	7	14	1		
The ELN helped me to have a better overall understanding of the course content	7	9	4	2	
The ELN helped me in writing my final course paper	3	9	8	2	
The ELN helped me in understanding the data from other students	5	11	4	2	
My experience with the ELN will be helpful in obtaining employment	3	12	3	3	1

## Implementation for Special Projects

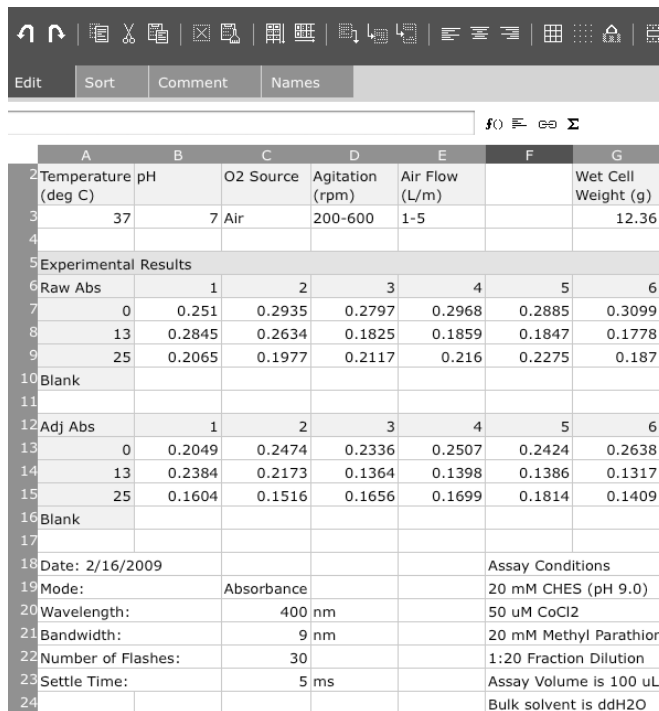
The ELN is also used to store and share data for individual laboratory projects. Students in the course for Special Topics in Biotechnology at the University of Houston investigated various parameters that affect recombinant protein production by the transformed bacteria, *Escherichia coli* strain DH5 $\alpha$ . Documentation for the experiment used an earlier version of a Drupal ELN, where the students uploaded their data via Excel spreadsheets and Excel graphs. The authors also enhanced the presentation below using native Drupal modules for the spreadsheet and graphing capabilities in order to highlight the possible ELN capabilities of this platform.

In the experiment, students used bio-fermenters to grow transformed bacterial cells. Through upstream and downstream production and purification, recombinant protein was extracted. Student teams manipulated one specific variable in the experiment—e.g., pH, temperature, aeration—in order to test its impact on cell growth and, ultimately, the protein yield. Each student team was responsible for manipulating a specific variable and recording the associated data. Students then presented a paper that proposed optimal conditions for cell growth.

For the optimized Drupal ELN, the experimental write up has four major sections: 1) Background and Procedures, 2) Data, 3) Graphs, and 4) Images. This ELN implementation is publicly available for review at [mylabbook.org/bmb300](http://mylabbook.org/bmb300), along with details about the modules used and technical details concerning how to create a similar ELN project.

For the data-collection aspects of the experiment, a spreadsheet was used (the sheetnode module) based on a

template. The data were extracted from designated cells in the spreadsheet for analysis and displayed on other pages of the ELN. The students were still free to fill in other cells with any data that they deemed relevant to the experiment, or they could use the formula functionality of the spreadsheet to perform computations on the data. When the student is entering data into the spreadsheet, it would look similar to Figure 3. Note that only the first 16 rows of the spreadsheet are used for data extraction and rows 17 and beyond are used by the researcher to annotate additional private information in a freeform manner as desired.

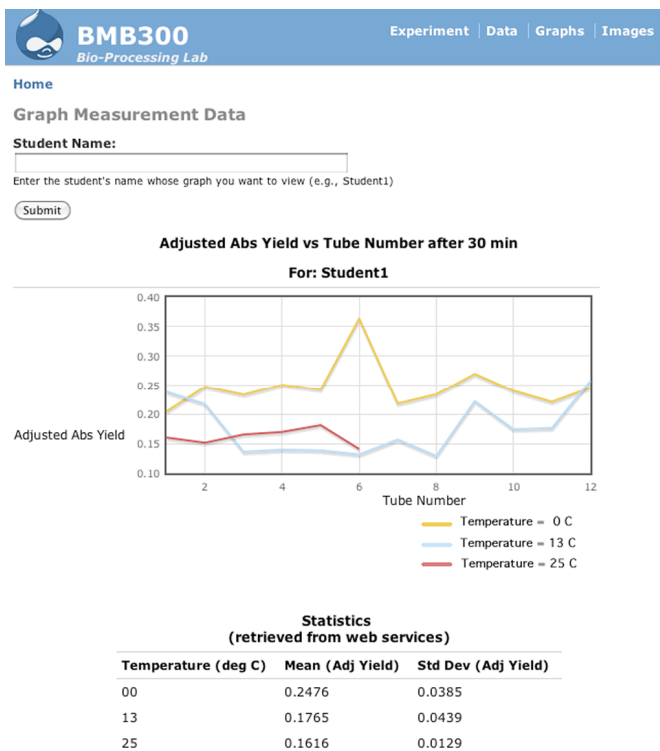


**Figure 3. Spreadsheet in edit mode using the sheetnode module**

Graphing can also be done directly from data in the spreadsheet based on a custom module that the authors developed (Figure 4) using the flot module. The flot module is based on a third-party Javascript module of that name used for plotting—which rhymes with “plot” and in Danish means “lovely”—which reflects the aesthetic display. These customizations are available on the [www.mylabbook.org](http://www.mylabbook.org) website as free downloads and are in the process of appropriate integration with the extension modules available on the Drupal.org website as well.

In addition to graphing, the custom module utilized web services to get basic statistics for the data, with the results shown at the bottom of Figure 4. Web services allow the data and results to be passed between computers that are appropriately configured. Computations do not need to be

run natively within the ELN platform, but can utilize computations done on another computer. In the case of using Drupal for an ELN, this allows utilization of the CMS strengths for data collection, presentation, and collaboration, while offloading computational needs for more sophisticated analysis to other platforms and software packages.



**Figure 4. Customized flot module displaying data from the spreadsheet**

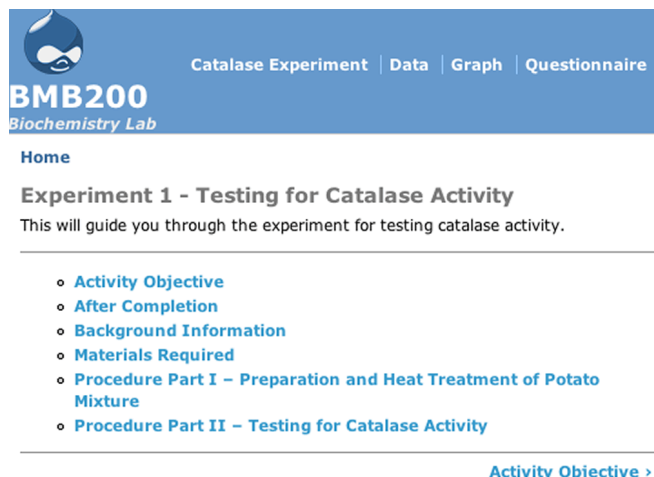
For this experiment, it was also helpful to provide images of the gel results. The uploading of images and other files is available in the core functionality of Drupal. A table of the image descriptions was created which provided a table of the image descriptions that can be filtered by student. At this time, image display within the table is not available, but by clicking on the hyperlink of the image name, the image from the experiment will be displayed.

## Classic Design Implementation

The last scenario to be covered concerns a simpler experiment that is easily broken out into the classic experimental stages of hypothesis, design, experiment, and analysis. It provides a model for other simple experiments and is based on a write-up posted on student pages at the University of Nebraska (used with permission) [8]. The ob-

jective of the experiment is to show that when a vegetable (potato) that is naturally high in the catalase enzyme is heated, there is less and less active catalase. The teacher and student write-ups are available from the student pages, and the content from these write-ups has been incorporated into the sample website which can be found at mylabbook.org/bmb200. Four main sections were created for this experiment – 1) Background and Procedures, 2) Data, 3) Graphs, and 4) Questionnaire.

The Background and Procedures section provides the explanation for the experiment and a description of the procedures to carry out the experiment. Hyperlinks are embedded in certain sections of the experimental procedures explanation to bring up the appropriate form that is completed for the data collection. A screenshot of the first page of the Background and Procedures section is shown in Figure 5. The hyperlink from the procedures section brings the student to the form, where data for temperature, test tube number, and time to float to surface are input. Presentation of this data is available in a tabular format along with appropriate filters of the data display, if desired. In Figure 6, a table of the data is shown as filtered for Student1.



**Figure 5. Explaining background and procedures for the experiment using the book module**

Following the collection of the data, one of the teaching points for the laboratory experiment is the graphing of the results to give a visual representation of the data. The output is shown in Figure 7, using the data shown in Figure 6, as a bar chart to highlight the increased changes at higher temperatures.

At the conclusion of the experiment, the laboratory procedures suggest that the student answer some questions. The resulting questionnaire for the student is shown in Figure 8. When the student fills out the questionnaire, the content will

automatically be emailed to the designated recipient, assumed to be, in this example, the lab instructor. This feature would most likely be restricted to availability only after a login by the student, in order to provide a measure of accountability for the answers.

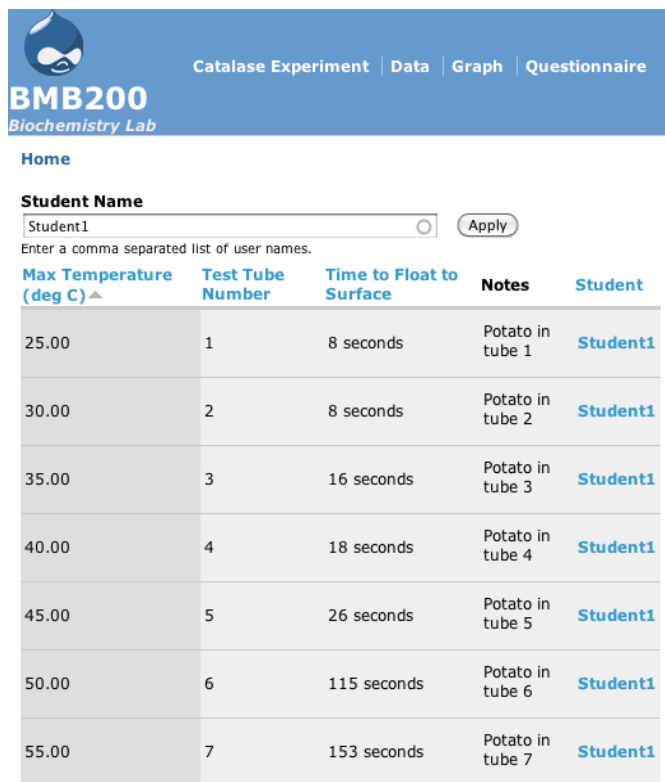


Figure 6. Experimental data filtered for Student1 using the views module

## Appropriate Information Technology for Laboratory Education

In determining appropriate information technology for enhancing laboratory education in a wide variety of settings, several criteria could be relevant.

1. *Low cost*: Educational communities are often more constrained by budgets than are businesses, as educators try to incorporate new and relevant technologies. The cost of capable commercial electronic lab notebooks can often be substantial [9], which usually precludes the commercial options that are used by most of the industrial or large research institutions that use ELNs.
2. *Sustainable*: Again, because of budget constraints of the educational community compared to the industrial community or some other sectors, it is often challenging to find technologies that are useful to the educational community which also stay current with the growth in technology.
3. *Low barrier to implementation*: New technologies usually involve learning curves, even if they can provide significant advantages. The easier that a technology is to set up and utilize, the more likely will be its uptake.
4. *Easily extended*: It is difficult to assess *a priori* all of the scenarios to which an ELN might be put to use. Therefore, it is important that it be as adaptable as possible.
5. *Appropriate analytic modules*: Inherent in the nature of ELN is the need for analytical capabilities. This will vary according to the particular application, but capabilities like statistical analysis will be commonly used.
6. *Research collaboration capabilities*: With the increasing popularity of social networking even within scientific and technical communities [10], collaborative capabilities are desirable not only for sharing results among researchers, but also for bringing an element of familiarity to many students.
7. *Semantic Web friendly*: Although not a common requirement, there is extensive literature explaining the benefits of the Semantic Web (sometimes referred to

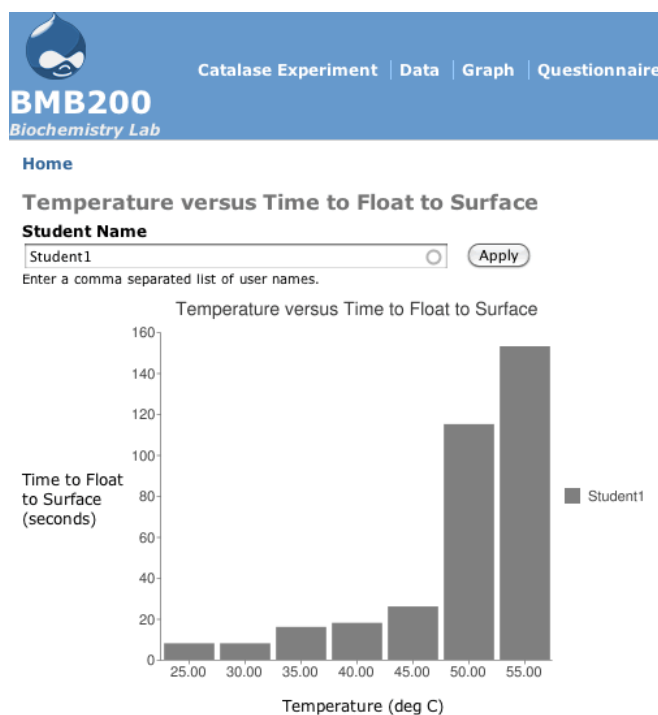


Figure 7. Data plotted for Student1 using the charts and graphs module with Google charts

as “Web 3.0”) to the scientific and technical communities [11]. This capability may play an important role in the future of sharing scientific data, so it would be advantageous to have this capability.

**BMB200**  
Biochemistry Lab

Catalase Experiment | Data | Graph | Questionnaire

Home

**Questions about catalase activity**

These questions should be completed after a full set of data has been taken for the catalase experiment

1. According to your graph, at what temperature does the catalase begin to lose activity?:  
 deg C

2. At what temperature is catalase completely inactivated?:  
 deg C

3. Could you imagine other ways, besides heat, of controlling enzymatic activity?:

4. How could you show that the potato is the source of the catalase enzyme?:

5. Why do we mark a line 5 cm from the bottom of the test tube?:

6. Would you expect catalase to be inactivated under normal blanching conditions for potatoes at 65 deg C? Why or why not?:

Submit

**Figure 8. Questionnaire to be filled out at end of experiment**

Among approaches to fulfilling these desired characteristics for ELNs in educational settings have been wikis and blogs [12]. These platforms can be edited and extended in a free-form manner by anyone with appropriate permissions and which provide ease of sharing with others. They have limited capabilities for automatically building the information architecture, however, and robust research collaboration and analytical capabilities are not available in the popular wiki or blogging platforms. A promising avenue for an open-source online ELN was the use of Google Wave [13], but recently Google has dropped continuing support for this platform. Other open source CMS platforms that deserve consideration as an ELN platform would include Ruby on Rails (rubyonrails.org) and Plone (plone.org). In regards to the lower cost commercial op-

tions for ELNs, they do not have as robust a feature set or match the criteria given above as do many of the freely available open-source CMS platforms, and so were not considered here.

Since this paper presents examples using Drupal, it is appropriate to assess how well Drupal meets these given criteria for laboratory information technology for education.

1. *Low cost:* The fact that Drupal is free and open source makes it attractive when compared to the available commercial electronic lab notebooks.
2. *Sustainable:* Drupal is widely used in many sectors; has the backing of many commercial organizations, consultants, and volunteer developers; and, has a thoughtful and flexible technical architecture. These all support sustained development independent of academic funding.
3. *Low barrier to implementation:* Drupal is written in the PHP scripting language, which is the most popular language for developing websites [14] and is available on most shared-hosting services or can be easily set up without much IT expertise by those familiar with basic website construction.
4. *Easily extended:* With over 12,000 available extension modules as of October 2011, Drupal can be adapted for numerous different use-cases without programming. If customization is desired, custom modules can be developed with a relatively easy learning curve [15].
5. *Appropriate analytic modules:* Existing Drupal extension modules can fulfill the analytical needs of many situations in laboratory education. However, other languages such as Ruby, Python, Java, and C++ have attracted more development for scientific or mathematical purposes. Drupal can often still take advantage of functionality in other languages by using web-service modules as an interface to appropriate analytical web services, as demonstrated in the bio-processing sample setup presented earlier in this paper.
6. *Research collaboration capabilities:* The widespread use of Drupal as the basis for social networking websites has brought maturity and variety to the tools that are currently not available among most research platforms.
7. *Semantic Web friendly:* Drupal is the first widely used CMS to incorporate into its core platform the automated production of RDFa data annotations consistent with the Semantic Web, as of the Drupal 7 release. It also has several mature modules for fea-

---

tures like SPARQL queries and RDF vocabulary import and export.

Further information about specific Drupal modules that might be of interest for lab management and scientific research are available at [www.mylabbook.or](http://www.mylabbook.or).

## Conclusions

In conclusion, the authors have presented how a free and relatively easily implementable CMS with ELN capabilities was used to enhance an existing undergraduate laboratory course and empower students for contributing to an ongoing research project. Also shown was how this platform might be used for a bioprocessing lab and, subsequently, in a more generic lab experiment. Several advantages are available in these approaches as compared to paper notebooks, including ease of sharing data and pictures in real time among students, spreadsheet capabilities for data collection and analysis, automatically generated graphs, analytical capabilities through web services, hyperlinks to information sources, and more extensive social networking and collaboration features. The initial student surveys about this approach suggest that it has been valuable in its ability to share data collectively within a given class situation. Furthermore, the authors showed the desired characteristics of an ELN in an educational laboratory setting and examined how well the Drupal platform matches these criteria. It should be noted that all implementations discussed in this paper, except for web services, have been on an outside Internet hosting company using a shared hosting plan, and so did not require extensive special configuration or a dedicated IT staff. The authors have developed an accompanying website at [www.mylabbook.or](http://www.mylabbook.or) with working versions of the second and third ELN scenarios along with areas for continuing documentation and discussions of this approach.

## Appendix

Abbreviations used:

**CCK** – Content Construction Kit, a Drupal module that allows customized content types

**CMS** – Content Management System

**ELN** – Electronic Lab Notebook

**Flot** – Javascript module for plotting data, rhymes with “plot” and is Danish for “lovely”

**IT** – Information Technology

**PERC** – Platform for Education and Research, as presented in this paper

**RDF** – Resource Description Framework, a data storage standard for the semantic web

**RDFa** – RDF for annotating web content

**Sheetnode** – Drupal module for the spreadsheet

**SPARQL** – Recursive acronym for **SPARQL Protocol and RDF Query Language**, a data query language for the semantic web

## Acknowledgments

Partial funding for this research was provided by the National Science Foundation (CCLI NSF DUE Award # 0633714) and the Texas Workforce Commission (Meeting Industries’ Critical Workforce Initiative).

## References

- [1] McCarthy, J. (2009, July/August). Enterprise Lab Notebook Improves Collaboration *Scientific Computing*. (pp. 10-12).
- [2] ELN and LIMS Aid Biotechnology Growth. (2006). *Genetic Engineering and Biotechnology News*, 26(5).
- [3] Elliott, M. (2008, May). ELN Overview and Benefits (Special Libraries Association). Retrieved October, 2010, from <http://www.atrriumresearch.com/library/AtriumResearchSLA2008%2520Rev%25202.pd>
- [4] Whitacre, P. (ed.) (2003). *Committee on Undergraduate Biology Education to Prepare Research Scientists for the 21<sup>st</sup> Century, Bio 2010: Transforming Undergraduate Education for Future Research Biologists*. National Academies Press, Washington, DC.
- [5] Woodin, T., Smith, D., & Allen, D (2009). Transforming undergraduate research for all students: An action plan for the twenty first century. *CBE Life Science Education*, 8, 271-291.
- [6] Iyer, R. S., & Fitzgibbon, W. (2009). Building the future biotechnology workforce: A University of Houston Model. *Journal of Commercial Biotechnology*, 15(2), 171-182.
- [7] Iyer, R. S., & Wales, M. (2012, Winter). Integrating Interdisciplinary Research-Based Experiences in Biotechnology Laboratories. *AEE Advances in Engineering Education*, 3(1), 1-35.
- [8] Food Science Labs. (n.d.). Retrieved January, 2011, from [www.math.unl.edu/~jump/Center1/BioChemLabs.html](http://www.math.unl.edu/~jump/Center1/BioChemLabs.html)
- [9] Dance, A. (2010). Digital Upgrade. How to Choose Your Lab’s Next Electronic Lab Notebook. *The Scientist*, 24(5), 71.
- [10] Pierce, M. E., Fox, G. C., Rosen, J., Maini, S., & Choi, J. Y. (2008). *Social networking for scientists*

- 
- using tagging and shared bookmarks: a Web 2.0 application*. International Symposium on Collaborative Technologies and Systems (257-266).
- [11] Goble, C., Corcho, O., Alper, P., & De Roure, D. (2006). E-Science and the Semantic Web: A Symbiotic Relationship. *Lecture Notes in Computer Science*, 4265, 1-12.
- [12] Quinnel, R., Hibbert, D. B., & Milstead, A. (2009). *eScience: Evaluating electronic laboratory notebooks in chemistry research*. Ascilite 2009 Auckland, 799-803.
- [13] Neylon, C. (2009). Head in the clouds: Reimagining the experimental laboratory record for the web-based networked world. *Automated Experimentation*, 1(3).
- [14] PHP Usage Stats for 2007 (n.d.). Retrieved January, 2011, from [php.net/usage.php](http://php.net/usage.php).
- [15] VanDyk, J. (2008). *Pro Drupal Development*. (2<sup>nd</sup> ed.). New York: Springer-Verlag.

## Biographies

**RUPA IYER** is Associate Professor and founding Director of the Center for Life Sciences Technology at the University of Houston's College of Technology. She earned her Ph.D. in microbial genetics from Michigan State University and conducted postdoctoral studies at M.D. Anderson Cancer Research Center. She is responsible for designing, developing and implementing the new interdisciplinary-based biotechnology degree program and for developing core initiatives of the center that include collaboration with the local biotechnology industry, development of workforce, training and outreach programs and development of a core facility to support the life sciences cluster for the state of Texas. She may be reached at [riyer@uh.edu](mailto:riyer@uh.edu)

**WILLIAM KUDRLE** received his B.S.E.E. from Rice University in 1980, an M.S.E.E. from Southern Methodist University in 1984 and a Ph.D. in Biomedical Engineering from the University of Texas at Austin in 1994. He is currently an adjunct program manager at the University of Houston's College of Technology and has research interests in semantic web technologies and biomedical informatics. He has developed various software packages and websites in the health, medical, energy and travel industries as well as at several academic institutions. He may be reached at [bkudrle@gmail.com](mailto:bkudrle@gmail.com)