Perspectives

Research, Collaboration, and Open Science Using Web 2.0

Kevin Shee,1*+ Michael Strong,2,3*+ Nicholas J. Guido,2 Robert A. Lue,1 George M. Church,2 Alain Viel1
1Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138; 2Department of Genetics, Harvard Medical School, Boston, MA 02115; 3National Jewish Health, Denver, CO 80206

There is little doubt that the Internet has transformed the world in which we live. Information that was once archived in bricks and mortar libraries is now only a click away, and people across the globe have become connected in a manner inconceivable only 20 years ago. Although many scientists and educators have embraced the Internet as an invaluable tool for research, education and data sharing, some have been somewhat slower to take full advantage of emerging Web 2.0 technologies. Here we discuss the benefits and challenges of integrating Web 2.0 applications into undergraduate research and education programs, based on our experience utilizing these technologies in a summer undergraduate research program in synthetic biology at Harvard University. We discuss the use of applications including wiki-based documentation, digital brainstorming, and open data sharing via the Web, to facilitate the educational aspects and collaborative progress of undergraduate research projects. We hope to inspire others to integrate these technologies into their own coursework or research projects.

INTRODUCTION

Web 2.0

Over the past few years, the Internet has undergone a notable transition from Web 1.0 to Web 2.0. Web 2.0 is broadly defined as a movement toward user-generated and community-driven Internet content and applications (http://oreilly.com/web2/archive/what-is-web-2.0.html), embracing ideas such as online social networking, community-based wikis, and other Web-based applications that rely on mass user participation (10). Web 1.0, in contrast, has generally involved the use of relatively static, non-user-generated Web content (http://www.wired.com/science/discoveries/news/2005/10/69114). Examples of websites incorporating features of Web 2.0 include the social networking site Facebook, the user-generated encyclopedia Wikipedia, and the video-sharing website YouTube.

Team-based science

There is no substitute for hands-on research experience in the sciences. Such experiences can foster a greater understanding of scientific material, build student confidence and teach undergraduates what it is like to do research in graduate school (2, 11, 13). Traditionally, hands-on laboratory experience has come in the form of lab-based science classes or through one-on-one mentoring of undergraduate students in a campus laboratory. Although both of these strategies provide certain advantages over course work alone, there is another model that has been gaining ground. This is the model of team-based or collective undergraduate research (3, 5, 6). Such a model is a hybrid of traditional laboratory classes and one-on-one research mentoring and promotes team work, collaboration and multidisciplinary science education through research (3, 5, 6). We have used this team-based approach as a model for our summer undergraduate research programs, integrating Web 2.0 technologies in combination with traditional laboratory-based research.

Our undergraduate students used these technologies throughout the summer in preparation of a project to present at the annual international genetically engineered machine (iGEM) competition (http://www.igem.org) (4). Synthetic biology and the iGEM platform aim to harness the power of engineering and biology in order to design and build biological systems using a modular, bottom-up approach, and serve as an excellent model for interdisciplinary science research and education (8).

The undergraduate students worked on a collective research project throughout the summer, focusing on engineering Escherichia coli to target, signal and respond utilizing a combination of bioengineered surface proteins and quorum-sensing pathway constructs. The students were mentored by a core group of teaching fellows and faculty mentors, whereby the teaching fellows were involved in the day-to-day research mentoring of the students and the
faculty advised the students during weekly group meetings. All of these interactions and research activities were greatly enhanced by our use of Web 2.0 features.

Virtual brainstorming

In the primary phase of planning for the summer research project, the students spent a large amount of time brainstorming ideas for potential summer research projects. In order to facilitate this brainstorming, we utilized a Web-based, user-generated wiki where the students, teaching fellows and faculty members could post and discuss potential ideas for the summer project (Fig. 1). A wiki is similar to a traditional website but allows instant editing by a community of users (9) – in this case our students. Our team wiki was hosted by OpenWetWare (http://openwetware.org/wiki/Main_Page), which provides free space for laboratories to set up their own scientific wiki. The wiki was initially used primarily for online virtual brainstorming and related discussions where student team members could pose ideas, ask for opinions and list relevant literature references. This was particularly useful because the students could work on this as they had time, without geographical restrictions. The only requirement was that they needed access to a computer with Internet access. By the end of three in-person meetings and extensive online virtual brainstorming on the wiki, the students and mentors formulated a cohesive project plan for the summer. For a dispersed group of students and mentors scattered among the main university and medical campuses, this method of virtual brainstorming worked extremely well.

As the summer project commenced, the wiki homepage transitioned from a brainstorming forum into a dynamic hub of activity, documenting the entire summer research project. Since wiki-code is largely intuitive and user-friendly, as opposed to other coding languages, the wiki enabled the students to easily edit, update and improve the homepage without the prerequisite knowledge of HTML or other Web programming languages. The students edited and fortified the home page to reflect their scientific and research interests, and the wiki quickly evolved to include online project documentation, literature references, Web links, electronic data archives, video podcasts, experimental protocols and other project-related information (http://openwetware.org/wiki/IGEM:Harvard/2007) (Fig. 2).

Open Web-based research documentation

As opposed to traditional paper-based laboratory notebooks, we utilized electronic notebooks and documentation to record research results directly on the wiki, in an

FIGURE 1. Virtual brainstorming sessions. (A) An example of the shared electronic notebook page used by the students for brainstorming summer project ideas. This method of online brainstorming complemented in-person brainstorming sessions and allowed the students to document, discuss and archive project ideas leading to the formulation of the summer project. (B) A portion of the source code for the brainstorming page is shown to demonstrate the intuitive nature of wiki code. Students can include hyperlinks to relevant websites, images and other e-notebook pages in a user-friendly manner using the wiki.

FIGURE 2. The lab wiki. Our wiki contains student-contributed information including laboratory protocols, reagents used, gene sequence information, literature references, student presentations, digital pictures, primary data and archives of virtual brainstorming sessions (http://openwetware.org/wiki/IGEM:Harvard/2007).
The electronic documentation served as a great resource to archive and share information among the students since the wiki could be easily updated, accessed and searched by all students whenever needed. Each of our lab benches was equipped with a computer with Internet access, which became the center of data recording and research findings for the students. The data was recorded directly as the experiments progressed, with no intermediate steps that could potentially cause a loss of data. Data archiving using the wiki also greatly facilitated downstream activities ranging from data analysis to making presentations and posters, since all data, images and information were electronically available to the students in one location at all times via the Web (http://openwetware.org/wiki/IGEM:Harvard/2007). The wiki-based electronic lab notebook, of which a sample page is shown in Fig. 3(A), contains laboratory protocols, primary data, sequence information (Fig. 3(B)), gel electrophoresis results (Fig. 3(C)), relevant hyperlinks and experimental conclusions. The wiki also enabled faculty advisors to keep track of the current progress of the summer research project, supplementing one-on-one discussions and weekly group meetings.

Some researchers may shy away from the Web 2.0 open science concept due to fear of over-technical interfaces, but with the initiation of wiki-based interfaces, online documentation has become much more straightforward to implement and sustain. Electronic documentation on a wiki also offers specific advantages to researchers, particularly for speed in archiving, searching, sharing and retrieving data, that paper-based notebooks do not allow (1, 7).

All of the information collected on our wiki was available to the general public as soon as it was generated, facilitating an open science approach to the summer undergraduate research project. Open science can be summarized as a system in which scientific information is openly shared with others, prior to publication, often utilizing the Web as the mode of information exchange (http://drexel-coas-elearning.blogspot.com/2006/09/open-notebook-science.html). This open science approach was also embraced by a number of other iGEM teams, resulting in an interconnected network of student scientists interested in advancing science through openness and data sharing.

In contrast to open science, typical closed science can be described as an environment in which a scientist’s research remains largely unknown to the world until publication. Even after publication, much of the primary data may never see the light of day since scientists rarely release their laboratory notebooks in their entirety. As a result, there are several inherent problems with the model of closed science including the potential for data loss, lag in time from discovery to dissemination, and possible duplication of efforts among different research groups.

The use of the open science model also helped the students contribute to the larger goal of increasing knowledge in a particular field – in this case synthetic biology – and helped transition the students from being “knowledge consumers” to “knowledge producers” (6). All of these activities helped the students garner a sense of pride and ownership of their project (2, 3, 5), since they were able to share their research results directly with others via the Web. Our project wiki has been accessed by ourselves and others over 10,000 times since we started it in March 2007, and student opinions regarding the summer program can be viewed on a video podcast on our wiki home page (http://openwetware.org/wiki/IGEM:Harvard/2007).

FIGURE 3. The electronic open notebook. (A) An example of a page from the electronic notebook relating to work performed on bacterial targeting. The wiki-based electronic lab notebook contains protocols, data, results, discussions and hyperlinks to other relevant experimental information including (B) sequence information and (C) gel electrophoresis results.
A working example of a model analogous to open science is that of the open source software movement. The power of the open source software approach has been demonstrated by the success of the Linux operating system and GNU licensed software, which in many cases is more stable and innovative than restrictive proprietary software. We think open science will lead to equally innovative ideas and implementations and hopefully increase the pace of scientific discoveries and breakthroughs. We are already seeing the positive impacts of the move of some journals toward open access.

The benefits of open science further extend to the classroom setting. For example, interactive online forums allow students to engage in detailed scientific discussion about the lecture and reading material, which greatly enhances the ability to think critically and apply knowledge to a broad range of potential scenarios. In laboratory-based classes, online wikis can be used for enhanced collaboration among laboratory partners, which directly minimizes the time requirements involved in sharing experimental data, protocols, images and observations from multiple paper laboratory notebooks. Web 2.0 tools can also be used in the classroom to simulate research problems or to link to publicly available biological datasets that can then be analyzed and discussed by students in an online and interactive manner. Finally, online wikis and notebooks are the most accurate representation of what is required for future laboratory or industry occupations, where many educational institutes and biotechnology corporations have begun to implement electronic notebooks to standardize documentation and gain searchable access to extensive amounts of data (12).

While open science may require a transition in scientific concepts, innovative technologies and ideas such as Web 2.0 and open science should be embraced and have much to offer the science education community. By engaging students in these technologies and ideas as undergraduates, we may entice them to further investigate the concepts of open science and, as a result, pioneer the use of new technologies that may capture the interest of more seasoned scientists and educators.

Considerations

As with any nascent technology, Web 2.0 is likely to flourish with further development. We envision further enhancements to the wiki platform, possibly incorporating video conferencing and online chat features. These enhancements will enable a more dynamic online environment where a group of students can contribute to the creation of a particular document or Web page, concurrently, in an interactive manner.

Although specific organizations, such as the Science Commons, are working to provide guidelines for establishing a licensing protocol that caters to open science data and publications (http://sciencecommons.org), the implementations of open science are relatively new and the potential for abuse exists. In the worst case scenario, a researcher may be “scooped” by another group that inappropriately uses the data or online information, without referencing or acknowledging the original source. Few researchers are willing to see their material misused by others and, as a result, in order for open science to flourish and gain wider acceptance, it will be necessary to focus efforts on both policy and technology to protect the rights and data of the primary research producers. With correct implementation we think the benefits of open science, particularly as an educational platform, greatly outweigh the risks.

CONCLUSION

The advantages of the wired lab approach to undergraduate research and education are plentiful. For the most part, besides the cost of laboratory computers, Web 2.0 technologies can be implemented in undergraduate research programs at little or no additional overhead cost. Implementation does, however, require a reconsideration of the traditional undergraduate laboratory layout, one in which computers are seamlessly integrated into the research environment. With Web 2.0 tools in hand, students can take full advantage of the Internet to create scientific resources that can be used to facilitate teamwork, collaboration and open science. All of these skills are desirable as science becomes increasingly multidisciplinary and collaborative in nature. We hope that technologies such as these will help usher in a new era of science and education, promoting increased openness, collaboration and scientific progress.

ACKNOWLEDGMENTS

We gratefully acknowledge the contributions and input of the other Harvard iGEM team members, teaching fellows and advisors. We also thank the iGEM founders and OpenWetWare team. We thank the Howard Hughes Medical Institute and Harvard University for funding and resources, and MS thanks the Jane Coffin Childs Foundation for funding and Dawn Hower for constructive discussions.

REFERENCES


