A Project-Centric Course on Cyberinfrastructure to Support High School STEM Education

Daphne RAINEY[1], Lisa COYNE[2], Jason GIBSON[2], Stephen CAMMER[1], Julie SCHULMAN[1], Betsy TRETOLA[3] and Oswald CRASTA[1*]

[1]Virginia Bioinformatics Institute, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA; [2]Galileo Magnet High School, Danville, VA 24541, USA; and [3]Institute of Distributive and Distance Learning, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA.

*Corresponding author. Phone: (540)231-2100; Fax: (540)231-2606; email: ocrasta@vt.edu

ABSTRACT

Recent rapid advances in information technology pose new challenges for teachers in the Science, Technology, Engineering, and Mathematics (STEM) fields to incorporate the latest knowledge and technical expertise into courses in a way that will be applicable to students as future scientists. A demonstration project was designed, developed, and deployed by university faculty and high school teachers for their students to explore the use of the components of cyberinfrastructure. The project explored the introduction of cyberinfrastructure through the use of bioinformatics and the use of team science. This paper describes the high school course that was deployed at Galileo Magnet High School (GMHS) in collaboration with the scientists at Virginia Tech University, and details its overall assessment. Implementation of a project-centric teaching paradigm to engage students in applying the concepts of cyberinfrastructure by integrating the disciplines of biology, computer science, mathematics, and statistics through bioinformatics was an integral part of this study.

Keywords: Cyberinfrastructure Education, Project-centric course and STEM Education.

1. INTRODUCTION

Recent rapid advances in laboratory technologies together with information technology are fueling the data rich environment in biology. Today, simulation and modeling are as important to discovery and innovation, as theory and experimentation. Team research through the integration of multiple disciplines including mathematics, chemistry, physics, and computer science is therefore becoming increasingly important when addressing specific biological problems such as personalized medicine, effective drug design, and generalized vaccines against fast-changing pathogens. Preparation of the current generation of students for the future workforce, where the challenges of analyzing, integrating, and interpreting complex, large-scale datasets are being addressed in order to increase the efficiency of biological research, means that we are preparing students for a new research environment that is expected to revolutionize the way discoveries are made. This has created new opportunities in the Science, Technology, Engineering, and Mathematics (STEM) fields for high school educators to incorporate the latest knowledge and technical expertise into their courses so that students will be able to apply that knowledge throughout their careers.

In response to the fast-changing world of information technology and the STEM fields, the National Science Foundation (NSF) convened a Blue Ribbon panel, which proposed a new concept of cyberinfrastructure (CI) that included components such as information technology, databases, and the social infrastructure needed to bridge the gap between data rich environments and the research community engaged in using such systems for making discoveries [1]. This concept of integrating the multitude of existing and future tools and services came in response to the recognition that many of today's scientific questions are far too complex for individual scientists and laboratories to answer alone [2]. A Cyberinfrastructure Training Education Advancement and Mentoring (CI-TEAM) program was created to meet these new challenges. The main objectives of the CI-TEAM program are to fund the design, development, and deployment of CI tools and environments for research and to prepare the future generation of scientists as developers and users of cyberinfrastructure.

Although current scientific questions are too complex to be answered by individual scientists or single laboratory groups, many school districts and institutions of learning have been investing heavily in computing tools for education [3]. Today's students are in a unique position to experience a shift in scientific culture, where the availability of public data is dramatically increasing and computational technology is continually changing; in fact, the expectation is that these students will enter the workforce with the understanding and skills needed to operate very successfully within this expanding research environment. Under this mandate, a demonstration project was conducted to explore the use of CI by university faculty and high school teachers and their students. The project explored the introduction of CI through the use of bioinformatics as well as team science.

In 2005, the Virginia Bioinformatics Institute (VBI) at Virginia Tech was awarded a CI-TEAM Demonstration project grant. The CI-TEAM Demonstration project is comprised of the following three organizational members: Bluefield State College (BSC), Bluefield, West Virginia; Galileo Magnet High School (GMHS), Danville, Virginia; and the Virginia Bioinformatics Institute (VBI) at Virginia Tech, Blacksburg, Virginia. The overall goal of the two-year project was to develop materials and create CI courses that would be deployed and evaluated by both BSC and GMHS in 2007. To aid in the introduction of CI concepts to the students and educators, VBI was to develop course modules utilizing state-of-the-art bioinformatics tools and genomics data, based on current VBI projects. BSC and GMHS were responsible for gathering and developing materials to supplement the modules and tailor the information to the students at their respective institutions. The course developed and deployed at BSC and the lessons learned have been documented [4,5]. This paper describes the design, development, and deployment of CI course materials by the faculty of GMHS for their students and outlines the assessment of and lessons learned from the project.

2. HYPOTHESIS AND PROJECT STRUCTURE

The CI-TEAM Demonstration project proposed that CI could be effectively introduced to the high school classroom via a combination of traditional learning styles with inquiry-based, project-centric learning activities in order to enhance students' STEM knowledge and increase their interest in STEM education.

The CI-TEAM Demonstration project tested this hypothesis through the development of a CI course developed through the collaboration of GMHS and VBI. GMHS is an International Baccalaureate school with a vision of creating an in-depth, unique, educational experience that allows for self-expression and critical thinking. The school applies a curriculum in which the students choose from one of three theme-based strands of study: 1) Advanced Communications and Networking Technology, 2) Air and Space Technology, and 3) Biotechnology. VBI, a Commonwealth of Virginia shared resource established at Virginia Tech in July 2000, is a bioinformatics research institute that weds cutting-edge, biological research with state-of-the-art computer science. The one year course examined the use of project-centric teaching/learning paradigms that included the use of research based data modules and role-play activities for the students.

3. TEACHING/LEARNING PARADIGM

As new research in the STEM fields merges with information technology, the need arises for a new workforce that integrates life science with computer science and mathematics. Thus, interdisciplinary training is essential in the creation of this future workforce. Not only do students need to develop their own interdisciplinary skills, but they also need to learn to work within teams, where various members contribute distinct expertise to the group. The project-centric teaching method employed here utilizes and highlights key concepts of CI, such as interdisciplinary science and team science. Figure 1 illustrates the experts and multiple disciplines involved in the CI-TEAM Demonstration project, which is focused on solving bioinformatics problems through research and teaching. Traditional teaching is often considered to be subject-based learning, where students learn specific information and are then given a problem that asks them to apply what they have learned. Project-centric learning is problem-based learning, where students are first given the problem to be solved and then must determine the information needed to fully address the problem before learning it and solving it. This type of learning is meant to encourage high-level thinking and collaboration.

4. PROJECT-CENTRIC COURSE DEVELOPMENT

Three bioinformatics research problems were developed as modules by VBI scientists. These modules were combined with pre-existing teaching materials from a biotechnology course to create a CI-based Microbial Genomics course at Galileo Magnet High School. The

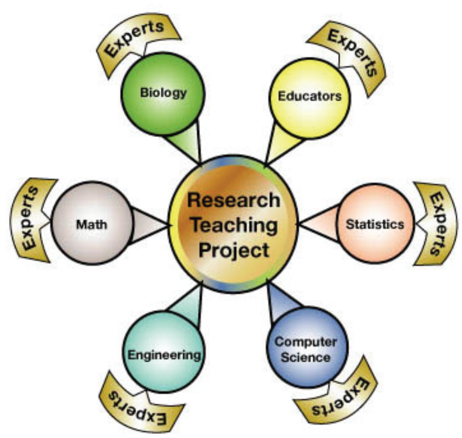

Figure 1. A multi-disciplinary, project-centric research and learning environment.

modules focused on role-based learning, interdisciplinary science, and bioinformatics tools and resources. Real-life scenarios were used to create the main problem in each module, so that the solutions would require interdisciplinary collaborations from individuals with expertise in such fields as molecular biology, microbiology, and bioinformatics. The tools and resources utilized by the students during the course of a module also required minimal specialized computer skills. To help students realize the weight and context of problems that may be solved using these tools and resources, the modules were centered on key pathogens of interest to biodefense and the public health sector: *Bacillus anthracis*, *Brucella*, and *Rickettsia*. The students analyzed a novel strain of *Bacillus anthracis*, the causative agent of anthrax; identified novel drug targets for *Rickettsia* species, the tick-borne parasites responsible for typhus, Rocky Mountain Spotted Fever and Lyme's disease; and identified vaccine candidates against *Brucella abortus*, the causative agent for brucellosis in cattle. The modules provided role-based learning for the students, and the roles were developed based on the various professionals that would be involved in the event of an outbreak, such as specialists from the Center for Disease Control, researchers involved in vaccine and/or drug development, hospital physicians, microbiologists, and evolutionary biologists. The students then worked within the different roles, accessing relevant data, and collaborating on a final approach for solving the original problem. Students learned how to navigate and use cyber-tools to aid them in solving the problems presented in the modules. Many core bioinformatics tools were introduced throughout the course, focusing on sequence searching and retrieval from NCBI, sequence comparison, sequence manipulation in the form of multiple sequence alignments, phylogenetic tree production, and the genome alignment software in Toolbus [6].

5. PROJECT ACTIVITIES AND IMPLEMENTATION

Student Activities

The CI course at GMHS became the capstone course of a four course series in the GMHS Biotechnology strand, consisting of Cell Biology, DNA Technologies, Advanced Applications of Biotech, and Microbial Genomics. The Microbial Genomics course was developed in partnership with VBI through a collaboration that began in the summer/fall of 2006, with implementation of the course occurring in January 2007. VBI provided three modules consisting of "outbreak" scenarios in which student groups were expected to analyze patient information, determine the disease in question, and research either a drug target or a vaccine candidate (Figure 2). Grant funding was used to provide various lab kits to students, as well as copies of three books: *Killer Germs* [7], *The Coming Plague* [8], and *Exploring Genomes* [9]. Students utilized the National Center for Biotechnology Information (NCBI) tutorials to learn about the Basic Local Alignment Search Tool (BLAST) and utilized PubMed to find conserved domains, 3D protein structures, and to measure phylogenetic distance. Teachers selected lecture topics, which included zoonotic diseases, food-borne pathogens, arboviruses, and the development of antibiotics, antibiotic resistance, and the immune system. Using NCBI and the Biology Workbench [10], students were introduced to the use of bioinformatics through the study of pathogens. They studied the evolutionary origin of HIV, the relationship of SARS to other diseases, and the transport of West Nile Virus through bird populations. Students wrote and taught NCBI tutorials on *Helicobacter pylori*. The National Institute of Health's (NIH) publication entitled *Medicines By Design* was used to gain a background on drug targets and drug design in anticipation of the final case studies. Students created and presented visual mnemonics on various antibiotics. Students gained an awareness of the history of vaccines as well as the current situation with HIV around the world. For their final exam, students used the VBI case studies as models to design their own cases. The case studies were based on outbreak scenarios involving 8-15 patients. Field and laboratory information was used to solve the cases and NCBI components were used to identify drug targets or vaccine candidates. The learning modules from VBI were used towards the end of the semester as prototypes for students to develop their own modules. The class met five days per week for 90 minutes each day. The teachers noted that the course time allotment not only allowed for in-class discussion, but also challenged the teachers to prepare enough materials

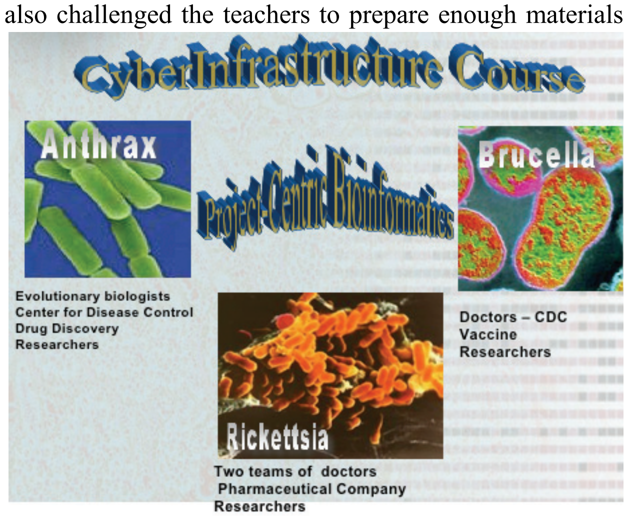

Figure 2. Three modules developed and deployed as part of the capstone course of a four course series in the GMHS Biotechnology Strand.

to fully engage their students. The experience trained teachers to communicate more effectively and it also instilled a desire to share the activities that they had developed with area colleagues.

Teacher Activities

It was determined that the materials developed for the GMHS course could serve as a resource to other high school teachers nationwide. Additionally, the Galileo teachers felt that there was a need in the Southside Virginia Area for this type of interdisciplinary, educational material. We felt that this information could be disseminated to other teachers in the Southside Virginia area in their courses. In June of 2007, a workshop was offered at the Institute of Advanced Learning and Research in Danville, Virginia. There were eight teacher-participants from Danville and surrounding areas. During the workshop the concepts of team science enabled by CI were presented along with methods to incorporate the materials into a classroom setting. The six day workshop was presented over a three week period by two of the Galileo Magnet High School teachers and a scientist from VBI.

The topics were presented as follows:

- Introduction to Cyberinfrastructure and Team Science - Bioinformatics and the Study of Microbes (VBI)
- Introduction of Biology Workbench and NCBI; Tutorials on protein and nucleic acid databases, PubMed and the creation of phylogenetic trees (GMHS)
- Human Diseases using Bioinformatics; Investigations of SARS, sickle cell anemia, and mad cow disease (VBI)
- HIV online lab activities; Evolution of HIV1 and HIV2, changes in HIV during the course of infection, molecular forensics and HIV (GMHS)
- Topics and Labs in Immunology; Medical case studies, antibiotic resistance and the complement system (GMHS)
- Proteins in 3D; Tools for visualizing molecules in 3D using NCBI (GMHS)

Teachers were surveyed about the course and also asked to discuss and plan ways of integrating the materials into their own curricula. A follow-up meeting was held in September of 2007 after the beginning of the fall semester classes. Only two of the seven participants attended the follow-up meeting. One of the two teachers will develop a class session during which students explore and learn to use the NCBI site. The other teacher will incorporate more computer use into an Introductory Biology course.

Overall, teachers were satisfied with the content of the course. There were questions of computer resource availability and the time needed by teachers to learn and prepare materials in the face of the Virginia Standards of Learning requirements. Work continues on the incorporation of materials into a website for further accessibility.

6. ASSESSMENT AND LESSONS LEARNED

Working in close cooperation, Galileo Magnet High School (GMHS) and the VBI Cyberinfrastructure Group developed and deployed a high school level CI course. Central to this course was a project-centric teaching paradigm that engaged students in applying the concepts of CI through the integration of the disciplines of biology, computer science, mathematics, and statistics in the field of bioinformatics. Ten online self-assessment surveys were conducted, and “lessons learned” review sessions were held among the VBI scientists who developed the course modules and the scientists and multidisciplinary teachers from GMHS who taught and completed pre- and post- course gap analyses and self-assessments. The pre- and post evaluations of the workshops and planning sessions were viewed very favorably by all faculty across different disciplines. The gap analysis was also positive. The gap between instructors’ desires and current competencies decreased significantly after they completed the workshops and taught the courses. Team collaboration enabled faculty to effectively disseminate CI knowledge. VBI scientists and the instructors reported that they gained valuable experience that will enable them to tackle both administrative and pedagogical barriers earlier in future implementations. The survey results and gap analyses show tangible improvements in students’ and instructors’ understanding of both the CI content and the use of bioinformatics tools. The difference between students’ self-assessed current level of knowledge at the end of the course and desired level of competence was reduced upon completion of the course. These positive results were also reflected in the self-assessment surveys completed by instructors at GMHS.

A set of “lessons learned” was developed based on the reviews of the assessments and further discussion:

a) Good team interaction – Good team interaction requires strong communication, preparation for group meetings, and coordination of activities. Team interaction addresses the major components of the paradigm shift from individual research to team science.
b) Demonstration of team science at every level – Multiple levels of team science were invoked during this project: 1) the scientists in the Cyberinfrastructure Group at VBI, 2) the multidisciplinary instructors at the high school, and 3)

the student teams that were created during the course. It was important to model strong team work for the students.

c) Application of the Cyberinfrastructure concepts - Developing an understanding of the CI framework was integral to this project. Defining an overview of the people, services, and technology specific to a project enables the effective use of the framework and its integration into knowledge.
d) Use of distributed resources – Integral to the CI concept with bioinformatics as the vehicle for implementing CI is an understanding of distributed resources. Realizing that CI allows researchers to call on resources from multiple sources is important to the development of interdisciplinary team science.
e) Solving real life problems – Using real-life pathogen outbreak scenarios was important in capturing the interest of students. This allowed them to observe how researchers from different disciplines come together to solve problems with large societal implications and thereby transform a project from interdisciplinary to transdisciplinary. Transdisciplinarity is both a problem-based and practice-based approach that capitalizes on the way different disciplines can be used to inform one another [11], especially as it pertains to societal level problems.
f) Effective dissemination of knowledge – For our purposes, knowledge dissemination was important on many levels. The dissemination of research data to teachers and conversely the dissemination of needs and working habits from the teachers to the scientists were important to the success of the project. On another level, both teachers and students gained an understanding of the importance of the flow of knowledge that must occur among professionals to solve a societal-level problem.

7. CONCLUSIONS

The rapid advances in information technology, especially those involving computational science, have changed how science is conducted. The availability of large amounts of data requires the interaction of researchers from different disciplines and the use of various computer applications and services. These skills are required in the 21st Century Workforce.

Despite the increased availability and use of instructional technology in many schools, the majority of teachers do not integrate computer technology into their curricula [12,13,14]. Although the availability of technology in the classroom has existed for more than a decade in many school systems, computers are most often employed to supplement traditional classroom pedagogy and are not yet effectively integrated into daily classroom practices; this trend was noted early on in technology implementation [15,16,14]. Many understand that simply providing state-of-the-art technology by itself will not be immediately reflected in desirable changes in education [17,14].

In this project we endeavored to integrate activities and approaches that would lead toward the creation of connections between active learning and teaching, and we aimed to develop an awareness of an emerging research paradigm – team science as supported by CI – in teachers and students. Cyberinfrastructure creates the possibility of establishing an important link between research and education and requires that STEM teachers and students be aware of the possibilities and uses of this framework in the study of science.

8. FUTURE WORK: EXPANSION AND SUSTAINABILITY

Training and expansion of the future workforce skilled in the use of CI is dependent upon the portability and sustainability of successfully developed CI teaching modules. The expansion of the CI model developed during the CI-TEAM Demonstration project is underway. The Cyberinfrastructure group at the Virginia Bioinformatics Institute (VBI CI Group) is leading a CI-TEAM Implementation for Biological Researchers, Educators, and Developers (CIBRED) project in an effort to expand, complement, and otherwise leverage previously tested and assessed activities from the CI-TEAM Demonstration project. The CIBRED project offers a vision of developing a sustainable and scalable CI-TEAM program for nationwide deployment. CIBRED is a virtual organization consisting of the following twelve collaborators (as of Jan. 2009): VBI (Blacksburg, VA); Hampton University (Hampton, VA); Howard University (Washington, D.C.); National University (San Diego, CA); Auburn High School (Riner, VA); Blacksburg High School (Blacksburg, VA); Denbigh High School (Newport News, VA); Galileo Magnet High School (Danville, VA); San Marcos High Tech High School (San Marcos, CA); Norwalk La Mirada School District (Los Angeles County, CA); Phoebus High School (Hampton, VA); and South Whittier School District (South Whittier. CA).

Over the past six months, virtual conferencing, teacher workshops, and course development have been initiated. Auburn high school is nearing completion of its new course to be offered in the fall of 2009. The other high schools will develop and implement modules to supplement current courses and introduce the concept of CI and team science through project-centric teaching.

Acknowledgements

The CI-TEAM Demonstration project was funded by NSF award OCI-0537461 to O. Crasta. CIBRED project is funded by NSF award OCI-0753375 to O. Crasta. The authors would like to thank Susan Faulkner for advice and comments and Emily Alberts for formatting and editing the manuscript.

9. REFERENCES

[1] D. Atkins, K. Droegemeier, et al, “Revolutionizing Science and Engineering Through Cyberinfrastructure: Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure”, **National Science Foundation**, 2003.

[2] H. Alvarez, D. Chatfield, D. Cox, E. Crumpler, C.D. Cunha, R. Gutierrez, J. Ibarra, E. Johnson, K. Kumar, T. Milledge, G. Narasimhan, M. Sadjadi and C. Zhang, “Cyberbridges: A Model Collaboration Infrastructure for e-Science”, **7th IEEE International Symposium on Cluster Computing and the Grid (CCGrid)**, Rio de Janeiro, Brazil, May 2007.

[3] E. Ficklen and C. Muscara, "Harnessing technology in the classroom." **Am Educ**, Vol. 25, No. 3, 2001, pp. 22-29.

[4] L. Craddock, D. Rainey, S. Faulkner, F. Hart, M. Eborall, L. Foster, S. Cammer, B. Tretola, B. Sobral, O. Crasta and B. Mutter, “Developing a multidisciplinary online cyberinfrastructure course through project-centric bioinformatics”, **114th Annual ASEE Conference & Exposition**, Honolulu, Hawaii, 2007.

[5] D. Rainey, B. Mutter, L. Craddock, S. Faulkner, F. Hart, M. Eborall, L. Foster, S. Cammer, B. Tretola, B. Sobral and O. Crasta, “A project-centric approach for cyberinfrastructure in bioinformatics”, **114th Annual ASEE Conference & Exposition**, Honolulu, Hawaii, 2007.

[6] Y. He, R.R. Vines, A.R. Wattam, G.V. Abramochkin, A.W. Dickerman, J.D. Eckart and B.W. Sobral, "PIML: the Pathogen Information Markup Language." **Bioinformatics** Vol. 21, No. 1, 2005, pp. 116-121.

[7] B.E. Zimmerman and D.J. Zimmerman, **Killer Germs: microbes and diseases that threaten humanity**, Chicago: McGraw Hill, 2003.

[8] L. Garret, **The Coming Plague: Newly Emerging Diseases in a World Out of Balance**, New York: Panquin Books, 1995.

[9] P.G. Young, **Exploring Genomes: web-based bioinformatics tutorial**, New York: W. H. Freeman and Co., 2007.

[10] S. Subramaniam, "The Biology Workbench--a seamless database and analysis environment for the biologist", **Proteins,** Vol. 32, No. 1, 1998, pp. 1-2.

[11] C. Nelson, J. Willetts and P. Bryce, “Transdisciplinary Research and Mozambique: Finding a new pathway to old problems”, 2005, Retrieved Jan. 19, 2009, from <http://www.isf.uts.edu.au/publications/nelsonetal2005mozambique.pdf>.

[12] K.C. Green, "The 2001 national survey of information technology in US higher education. EDUCAUSE 2001 Presentation, The 2001 Campus Computing Project”, 2001, Retrieved Jan. 16, 2009, from < http://www.campuscomputing.net/pdf/2001-CCP.pdf>.

[13] D.M. Jacobsen, “Examining technology adoption patterns by faculty in higher education”, **ACEC 2000 Learning technologies, teaching and the future of schools,** Melbourne, Australia, 2000.

[14] J. Keengwe, G. Onchwari and P. Wachira, "Computer Technology Integration and Student Learning: Barriers and Promise”, **Journal of Science Education and Technology**, Vol. 17, No. 6, 2008, pp. 560-565.

[15] A. Candiotti and N. Clark, "Combining universal access with faculty development and academic facilities”, **Communications of the ACM**, Vol. 41, No. 1, 1998, pp. 36-41.

[16] L. Cuban, H. Kirkpatrick and P. Craig, "High access and low use of technologies in high school classrooms: explaining an apparent paradox”, **Am Educ Res J**, Vol. 38, No. 4, 2001, pp.813-834.

[17] T.W. Kent and R.F. McNergney, **Will technology really change education?**, Thousand Oaks: Corwin Press, 1999.