Data-Enabled Engineering Projects (DEEPs) Modules for Data Science Education in Engineering

Q. Peter He[1], Jin Wang[1], Shiwen Mao[2], Laura Parson[3], Bo Liu[4], Peng Zeng[5], Allen Smith[6], Daniel Henry[3]

[1]Dept. of Chemical Engineering
[2]Dept. of Electric and Computer Engineering
[3]Dept. of Educational Foundations, Leadership, and Technology
[4]Dept. of Computer Science and Software Engineering
[5]Dept. of Mathematics and Statistics
Auburn University, Auburn, AL 36849
[6]Dept. of Chemical Engineering
Tuskegee University, Tuskegee, AL 36088

Abstract

The importance of data science and engineering (DSE) education cannot be overstated and undergraduate education offers a critical link in providing more DSE exposure to students and expanding the supply of DSE talent. Currently significant progress has been made in classwork, while progress in hands-on research experience is still lacking. To help fill this gap, we propose to create data-enabled engineering project (DEEP) modules based on real data and applications to be easily and widely adopted by other institutions. This project has recently been funded by the National Science Foundation (NSF) under the Improving Undergraduate STEM Education (IUSE) program. In this work, we will share our ideas, the rationale behind the proposed approach, the proposed tasks for the project and the plan of evaluation.

Keywords

Data science, workforce development, data-enabled engineering project, experiential learning, course-based undergraduate research experience

Introduction

Data science is emerging as a field that is revolutionizing the world. A 2018 National Academies report – *Data Science for Undergraduates: Opportunities and Options* [1] states that "Work across nearly all domains is becoming more data driven, affecting both the jobs that are available and the skills that are required. As more data and ways of analyzing them become available, more aspects of the economy, society, and daily life will become dependent on data." A recent study by IBM found more than 2.3 million data science and analytics job listings in 2015, and predicted that demand for data scientists will soar 28% by 2020 [2]. The National Academies report concludes that undergraduate education offers a critical link in providing more DSE exposure to students and expanding the supply of DSE talents. DSE education requires both appropriate classwork and hands-on experience with real data and real applications. While significant progress has been made in the former, one key aspect that yet to be addressed is hands-on experience incorporating real-world applications. Specifically, it is insufficient for undergraduate students to be handed a "canned" data set and be told to analyze it using the methods that they are studying. Such an approach will not prepare them to solve more realistic and complex problems, especially those involving large, unstructured data. Instead, students need repeated practice with the entire DSE cycle beginning with ill-posed questions and "messy" data [1]. To this end, the following gaps have been identified. (1) There is a lack of real data and application based learning materials for students to learn different aspects of DSE relevant to their life experiences and future job requirements. (2) There is lack of real data and application based research experiences or projects for students to practice the entire DSE workflow. To help fill these gaps, we have proposed to create data-enabled engineering project (DEEP) modules based on real data and applications to be easily and widely adopted by other institutions. The rest of the paper is organized as the follows. In Section 2, we discuss limitations of the existing efforts and the proposed approach. The planned tasks are discussed in Section 3. The project evaluation plan is presented in Section 4.

The limitations of the existing efforts and the proposed approach

It has been recognized that textbooks and traditional lecture courses may offer limited help in developing students' capability in applying the theory and methods to solve real, complex problems [1]. There have been some efforts that integrate real projects and real data into DSE education. However, there are many limitations, such as increased time, organizational, and pedagogical demands, and other burdens on instructor[3,4]; challenges in solicitation of live projects [4]; difficult to find assignments that motivate all students[16]; may not have immediate applications. More importantly, there is no learning materials generated from these efforts that can be widely adopted for enhancing DSE education at other institutions.

To address the above-mentioned limitations, we propose to develop data-enabled engineering project (DEEP) modules guided by the latest research on experiential learning theory (ELT). Experiential learning (EL) is the process of learning through experience, and is more specifically defined as "learning through reflection on doing"[5,6]. In addition, course-based undergraduate research experience (CURE) is a form of experiential learning that promotes all EL components in a positive cyclical and spiral learning process. As most DSE applications are open-ended research problems and learning an entire DSE lifecycle is really a research experience, the latest research on CURE provides excellent guidance for assembling DEEP modules into research projects. In particular, a 2019 study found that short CURE modules are an excellent alternative to more complex and costly whole-course CUREs and provide measurable metacognitive benefits to students[8]. Another benefit of short CURE modules is that they can be flexibly insert into existing curricula[8,9]. Therefore, we further propose to adapt the short CURE module mechanism to assemble DEEP modules into short DEEP-CUREs that will be inserted into six existing DSE courses. The level of complexity of a DEEP-CURE for a particular curriculum can be easily adjusted by including different DEEP modules to fit into the existing syllabus, while offering repeated practices on different steps of the DSE lifecycle.

Planed tasks

To achieve the project goals, we plan to first develop three industrial Internet-of-Things (IIoT) enabled laboratory engineering testbeds (LETs) to generate real data based on real-world applications. Then we will develop DEEP modules based on LETs to cover different steps of DSE lifecycle, and DEEP-CUREs to cover entire DSE lifecycle. Finally we will assemble and test DEEP modules and DEEP-CUREs in six existing DSE courses. Data will be collected and analyzed to evaluate and improve the DEEPs. Data will also be collected and analyzed to test the hypothesis that DEEPs can enhance students' reflection and metacognition.

One major obstacle of integrating DSE education into all STEM disciplines is that not all academic programs will be able to accommodate the addition of a designated DSE course with all other programmatic requirements currently in place. Another obstacle is that any changes to the curriculum that requires significant effort from faculty or staff would be difficult to sell and would not be adopted widely. In the proposed framework, the DEEP modules serve as the supplementary materials to existing STEM courses. In other words, the basic concepts and fundamental principles to be learned remain the same and are taught in the same way as in a traditional STEM class. The DEEP modules can be used to replace/supplement some of the textbook examples and homework problems. Therefore, the proposed DEEP module approach can effectively address the above-mentioned obstacles.

We plan to test the developed DEEP and DEE-CURE modules in six engineering courses from four different disciplines: Chemical Engineering, Electrical and Computer Engineering, Computer Sciences, and Mathematics and Statistics. We hypothesize that ELT guided DEEP modules, especially those designed for students to answer important reflective questions in DSE, together with DEE-CUREs, will significantly enhance students' reflection and metacognition. Metacognition refers to the ability to reflect upon, understand, and control students' learning[10,11]. Extensive research has indicated that metacognitively aware learners are more strategies and perform better than unaware learners[10–16]. Metacognition has two major components: knowledge about cognition and regulation of cognition. The former includes knowledge about self and strategies, how to use strategies, and when and why to use strategies. The latter includes planning, information management strategies, comprehension monitoring, debugging strategies, and evaluation. In this project, a widely used 52-question Metacognition Awareness Inventory (MAI) will be used to quantify students' metacognition awareness gains. The scores will be analyzed and compared categorically and holistically to test our above-mentioned hypothesis. By reviewing the MAI questions, we expect that DEEPs will mostly enhance the "regulation" aspect of metacognition. However, it is possible that the "knowledge" aspect, such as knowledge about strategies and how to use strategies could be enhanced as well. The findings will add evidence to CUREs' effectiveness in DSE education and in enhancing metacognition in general. The findings will also add knowledge to ELT in DSE education, and provide a model of ELT guided DEEP module design for addressing specific aspects of ELT.

Project evaluation plan

The project will be evaluated independently by Auburn Center for Evaluation (ACE). ACE is a member of the American Evaluation Association, the Southeast Evaluation Association, and the American Educational Research Association. The holistic model of evaluation will be used to evaluate this project. Data will be collected focusing on the key aspects of the project: Creating DEEP modules and incorporating them into STEM courses. To that end, numbers of modules created and curricula modified, participants served, evaluation of student products, numbers of hours spent in researching and developing modules, as well as evidence of student knowledge gains will be measured.

References

1. National Academies of Sciences, Engineering, and Medicine. *Data science for undergraduates: opportunities and options*. (National Academies Press, 2018).
2. Columbus, L. IBM predicts demand for data scientists will soar 28% by 2020. *IBM White Pap.* (2017).
3. Bickerstaff Jr, D. D. The evolution of a project oriented course in software development. *ACM SIGCSE Bull.* **17**, 13–22 (1985).
4. Chase, J. D., Oakes, E. & Ramsey, S. Using live projects without pain: the development of the small project support center at Radford University. *ACM SIGCSE Bull.* **39**, 469–473 (2007).
5. Kolb, D. Towards an applied theory of experiential learning. *Theor. Gr. Process.* 33–56 (1975).
6. Stonehouse, P., Allison, P. & Carr, D. Aristotle, Plato, and Socrates: Ancient Greek perspectives on experiential learning. *Sourceb. Exp. Educ. Key thinkers their Contrib.* 18–25 (2011).
7. Zull, J. E. *The art of changing the brain: Enriching teaching by exploring the biology of learning*. (Stylus Publishing, LLC., 2002).
8. Dahlberg, C. L. *et al.* A Short, Course-Based Research Module Provides Metacognitive Benefits in the Form of More Sophisticated Problem Solving. *J. Coll. Sci. Teach.* **48**, (2019).
9. Howard, D. R. & Miskowski, J. A. Using a module-based laboratory to incorporate inquiry into a large cell biology course. *Cell Biol. Educ.* **4**, 249–260 (2005).
10. Schraw, G. & Dennison, R. S. Assessing metacognitive awareness. *Contemp. Educ. Psychol.* **19**, 460–475 (1994).
11. Schraw, G. Promoting general metacognitive awareness. *Instr. Sci.* **26**, 113–125 (1998).
12. Young, A. & Fry, J. D. Metacognitive awareness and academic achievement in college students. *J. Scholarsh. Teach. Learn.* **8**, 1–10 (2008).
13. Turan, S., Demirel, Ö. & Sayek, İ. Metacognitive awareness and self-regulated learning skills of medical students in different medical curricula. *Med. Teach.* **31**, e477–e483 (2009).
14. ur Rahman, F., Jumani, N. B., Chaudry, M. A. & Abbasi, F. Impact of metacognitive awareness on performance of students in chemistry. *Contemp. Issues Educ. Res.* 39–44 (2010).
15. Kállay, É. LEARNING STRATEGIES AND METACOGNITIVE AWARENESS AS PREDICTORS OF ACADEMIC ACHIEVEMENT IN A SAMPLE OF ROMANIAN SECOND-YEAR STUDENTS. *Cogn. Creier, Comport. Brain, Behav.* **16**, (2012).
16. Yeşilyurt, E. Metacognitive Awareness and Achievement Focused Motivation as the Predictor of the Study Process. *Int. J. Soc. Sci. Educ.* **3**, (2013).

Q. Peter He

Dr. Q. Peter He is Associate Professor in the Department of Chemical Engineering at Auburn University. He obtained his BS degree from Tsinghua University (China) in 1996, and MS and PhD degrees in 2002 and 2005 from the University of Texas, Austin, all in chemical engineering. His current research interests are in the areas of systems engineering enhanced big data analytics with applications in smart manufacturing, renewable energy, digital agriculture, and cancer and healthcare related research. Besides research, Dr. He is interested in engineering education. His current interest in this area is data science education in engineering.

Jin Wang

Dr. Jin Wang is Walt and Virginia Woltosz Endowed Professor in the Department of Chemical Engineering at Auburn University. She obtained her BS and PhD degrees in chemical engineering (specialized in biochemical engineering) from Tsinghua University in 1994, and 1999 respectively. She then obtained a PhD degree (specialized in control engineering) from the University of Texas at Austin in 2004. The central theme of her research is to apply systems engineering principles and techniques to understand, predict and control complex dynamic systems, including both engineered systems and microbial organisms. Dr. Wang also devotes herself to education and promoting women in engineering.

Shiwen Mao

Dr. Shiwen Mao received a PhD in electrical and computer engineering from Polytechnic University, Brooklyn, N.Y., in 2004. He is the Samuel Ginn Professor and Director of Wireless Engineering Research and Education Center at Auburn University, Auburn, AL. His research interests include wireless networks, multimedia communications, and smart grid. He is a recipient of the Auburn University Creative Research & Scholarship Award in 2018, the NSF CAREER Award in 2010, several conference best paper awards, and The 2004 IEEE Communications Society Leonard G. Abraham Prize in the Field of Communications Systems. He is a Fellow of the IEEE.

Laura Parson

Dr. Laura Parson is an Assistant Professor in the Higher Education Administration Program at Auburn University and Higher Education Administration MEd/PhD Program Coordinator. Her Ph.D. is in Teaching & Learning, Higher Education from the University of North Dakota. Her research interests focus on identifying the institutional practices, processes, and discourses that coordinate the teaching and learning experiences of women in higher education, explored through a critical lens. She is a qualitative methodologist, with a focus on ethnographic and discourse methods of inquiry. Her research questions seek to understand how pedagogy, classroom climate, institutional environment, and curriculum inform student experiences, and how the institution coordinates those factors through translocal practices.

Peng Zeng

Dr. Peng Zeng is currently Associate Professor of Statistics in Department of Mathematics and Statistics, Auburn University. He got a bachelor’s degree in Mathematics from Nankai

University, and a master's and doctorate degree in Statistics in Purdue University, West Lafayette. His main research interests include high-dimensional data analysis, design and analysis of experiments, semi-parameter regression, machine learning, deep learning, etc.

Allen Smith

Dr. Allen Smith is an assistant professor of chemical engineering at Tuskegee University. He received B. S. and Ph.D. degrees from Auburn University and a M.S. from the University of Washington. He worked in the pulp and paper industry for 14 years after the M.S. before returning to school for the PhD. His research interests include Pulp and Paper, Energy from Biomass, Food Science, and Undergraduate education. His teaching responsibilities include Thermodynamics, Reaction Engineering, and Process Control.

Daniel Henry

Dr. Daniel Henry is the founding director of the Auburn Center for Evaluation (ACE). ACE has conducted evaluations for the Pew Charitable Trusts, USDA, McGraw-Hill, and the Alabama State Department of Education. Dr. Henry has a Ph.D. in Educational Psychology from Indiana University, and is an Assistant Clinical Professor at Auburn University. He has published in the areas of program evaluation, qualitative research, educational pedagogy, and hunger studies. He teaches program evaluation, qualitative research, and learning and cognition at Auburn. Before coming to Auburn, Dr. Henry was a faculty member at Central Michigan University and Indiana University, and taught high school English for 13 years.