Integrating Data Information Literacy into a Service-Learning Engineering Design Course

Margaret Phillips
Purdue University Libraries
Purdue University
West Lafayette, IN USA
phill201@purdue.edu

Mohammad Reza Jahanshahi
Lyles School of Civil Engineering
Purdue University
West Lafayette, IN USA
jahansha@purdue.edu

***Abstract*—Over the past several years increasing emphasis has been placed on the development of data information literacy (DIL) skills for engineering researchers to comply with funding requirements and to facilitate data sharing, reuse, and preservation. These efforts have been largely aimed at faculty members and graduate students, and to some extent undergraduate students in preparation for their roles as future researchers. However, little attention has been directed at developing DIL skills in undergraduate engineering students engaged in non-research activities, such as service-learning design projects that integrate instruction and community service. This work-in-progress paper reports on using service-learning pedagogy to integrate DIL into engineering design education. Preliminary analysis of student journal entries and instructor observations are discussed.**

Keywords—data information literacy; data management; community service; undergraduate service-learning; smart cities; engineering design education

I. INTRODUCTION

Data information literacy (DIL) skills include the abilities to properly organize, manage, use, share, analyze, and visualize data [1]. These skills are essential for future engineers. Their importance can be seen in the ABET Engineering Accreditation Criteria (EAC) student outcome "b", "an ability to design and conduct experiments, as well as to analyze and interpret data" [2].

This is especially true for projects involving very large data sets, or "Big Data," which Chen [3] defines as "the data sets and analytical techniques in applications that are so large (from terabytes to exabytes) and complex (from sensor to social media data) that they require advanced and unique data storage, management, analysis, and visualization technologies." Big Data is commonly characterized by the 4 V's: volume, velocity, variety [4] , and veracity [5]. Big Data projects impacting engineering and many other fields are anticipated to increase rapidly in number and complexity in the near future [6-7].

In 2013 the United States Office of Science and Technology Policy (OSTP) issued a memorandum stating the results of federally funded research, including data, should be made publicly available "to the greatest extent and with the fewest constraints possible" [8]. The memo charged federal agencies, such as the National Science Foundation (NSF) and National Institute of Health (NIH), with developing and implementing access policies. Additionally, a search of Sherpa/Juliet, an open database that provides information about research funder policies, indicates data policies are in place internationally for many government and private funders of research [9]. Many college and university libraries have developed services and instruction to support the research data needs of faculty and graduate students in engineering and other disciplines [10-14]. To a lesser extent, these efforts have broadly included undergraduate students engaged in research activities [15].

However the literature reveals few DIL efforts that are specifically aimed at undergraduate engineering students engaged in non-research activities. Zilinski, Sapp Nelson, and Van Epps' [16] data credibility instruction did include nuclear engineering students in a co-listed graduate and undergraduate engineering communication and literature course. In addition, Carlson and Sapp Nelson [17] incorporated DIL education focused on software code documentation into three sections of an Engineering Projects in Community Service (EPICS) service learning course.

Our work is also situated in an EPICS service learning course. Service learning pedagogy combines classroom education with community service experiences. It is defined as "a form of experiential education in which students engage in activities that address human and community needs together with structured opportunities intentionally designed to promote student learning and development. Reciprocity and reflection are key concepts of service-learning" [18]. Service learning pedagogy has been demonstrated to have positive impact on student learning outcomes [19].

The specific focus of our EPICS course is smart cities. With recent advancements in sensor, communication, and information technologies, in the near future huge amount of heterogeneous data will be collected from a variety of sensors that are spread throughout a city (i.e., ubiquitous sensing network), buried in buildings, mounted on mobile sensing agents as well as carried by citizens. This will lead to the vision of smart cities that relies on developing the capability to collect, analyze, and fuse huge amounts of heterogeneous data acquired by numerous sensor units where such data constitutes a true example of the so-called Big Data. A smart city provides

978-1-5090-5920-1/17/$31.00 ©2017 IEEE

systematic and more efficient management of urban systems, including transportation flow, water usage, and energy consumption. Therefore, it will be a promising technology that can connect multiple fields, including sensing, information theory, transmission and detection, networking, control theory, system theory, software, and middleware. Smart cities will inevitably yield to significant cost reductions, introduce new services using the sensed information, and improve the resiliency and sustainability of urban systems by enabling more frequent and quantitative assessments of the infrastructure and environment.

II. Background

A. EPICS Program

EPICS was initially founded at Purdue University in 1995 [20] to address two issues: (1) lack of real world skills for project management by most of engineering graduates; and (2) lack of funding for nonprofit organizations to support professional engineering services. To this end, EPICS uses the skills of undergraduate students to provide services to local and global non-profit organizations through a multidisciplinary, student-led, service-learning design course. Students earn academic credit for developing designs through community partnerships (in our case, the City of West Lafayette) where they experience the entire design/life-cycle as part of multidisciplinary teams. The students develop a broad set of technical and professional skills including teamwork, leadership, project management, and communication skills. The program has been recognized by the NSF Corporate Foundation Alliance as an Exemplar Program (2002); the National Academy of Engineering's (NAE) Bernard M. Gordon Prize for Innovation in Engineering Education (2004); and as an NAE exemplar of programs "Infusing Real World Experiences into Engineering Education" (2012) [21].

Students can sign up for EPICS as many times as they want during the course of their undergraduate degree. Some students choose to enroll in EPICS every semester and others decide only to participate one term. Students may also opt to stay with the same themed EPICS section for all of their time in the program, or try out different sections each semester they enroll in the course.

B. EPICS Smart Cities Section

Spring 2017 is the first offering of the EPICS Smart Cities section. The section is planned to be offered on an on-going basis with specific projects taking multiple semesters or years to complete. The objective of this section is to engage undergraduate students to deal with challenges of future smart cities. To this end, the goal of the first project is to create a transformative and human-centered smart system for ubiquitous automated monitoring and quantification of pavement defects based on a Multimodal Monitoring Swarm (MMS) approach. MMS is a group of mobile sensing probes that measure spatiotemporal data of interest, using a sensor cluster composed of different sensor types that can be mounted on any mode of vehicle, such as cars, Segways, bicycles or on-foot. In this project, students design autonomous fusion and interpretation of huge amounts of heterogeneous data collected by a swarm of relatively inexpensive mobile sensing agents equipped with a variety of different sensors including, such as RGB cameras, depth sensors, global positioning system (GPS) sensors, etc. The short term objective of the first project is to design an inexpensive hardware-software package system that can be mounted on garbage trucks and parking patrol vehicles owned by the City of West Lafayette for autonomous detection, quantification, and localization of potholes. In addition, the students plan to develop a user-friendly smartphone application so that the citizens can take picture of potholes in residential areas and conveniently submit pothole reports to authorities. The aggregation of the above information provides updated condition map of roads in West Lafayette area to the decision makers for better prioritization and optimization of maintenance work where several potholes appear in a relatively short amount of time after each winter season.

The Smart Cities section, consisting of two advisors, a teaching assistant (TA), and 12 students, met once a week during the Spring 2017 term for a two-hour lab. Additionally, the students attended separate EPICS design lectures taught by EPICS instructors to learn the engineering design process utilized in the course. Outside skills sessions led by TA's were also offered for students to further their abilities in areas like coding, documentation, project management, and application development.

During the weekly lab sessions the students presented their weekly progress, issues, and goals. They received feedback from us, as the section advisors, the course TA, and their peers. It was in these sessions where we introduced the concept of Big Data and encouraged students to continuously reflect on data concerns, such as the amount of data being generated, the varieties of data, data storage needs, proper documentation, optimal ways to visualize the data, and data sharing, as related to their project throughout the term.

III. Methods

A. Participants

Participants consisted of 12 undergraduate students enrolled in the Spring 2017 Smart Cities EPICS course. Table 1 provides a summary of student characteristics with regard to gender, year of enrollment at the university, and major.

B. Data Collection and Analysis

EPICS students are required to complete weekly journal reflections throughout the 16-week course. During week one the students were given the following reflection prompts to answer outside of class:

> I think data management applies to the Smart Cities project in this way: ______________
>
> I anticipate I will need these data management skills for the Smart Cities project: _________

During week 15 of the term, the students were assigned these reflection prompts to answer outside of class:

TABLE I. STUDENT CHARACTERISTICS (n=12)

		Gender								
		Female				*Male*				
		Year 1	*Year 2*	*Year 3*	*Year 4*	*Year 1*	*Year 2*	*Year 3*	*Year 4*	*Total*
Major	*Actuarial Science*				1					1
	Biological Engineering							1		1
	Civil Engineering		1				1		1	3
	Electrical Engineering				1					1
	First Year Engineering	1				2				3
	Industrial Engineering						1			1
	Mechanical Engineering								1	1
	Multidisciplinary Engineering								1	1
	Total	1	1	0	2	2	2	1	3	12

I think data management applies to the Smart Cities project in this way:_______________

I needed these data management skills for the Smart Cities project: ________________

For this work-in-progress paper, we read the week one journal reflection responses multiple times to gain an understanding of the content. After data collection is complete we plan to use content analysis [22-23] to analyze the week one and week 15 responses to determine if there are differences in the sophistication of students' discussions about data management from the beginning to end of the course.

IV. Preliminary Results and Discussion

A. Week One Reflections

Eleven out of twelve students completed their week one reflection. Almost all of the students acknowledged the importance of an efficient data management system for the success of this project. In considering the four V's of Big Data (volume, velocity, veracity, and variety), all of these characteristics were represented in the collective students' responses in some way, albeit not equally. Several students acknowledged the project involves multiple varieties of data, such as GPS coordinates, images and text captured through the smartphone application, the use of multiple sensors, and pothole depth/width data. Two students acknowledged data volume as a concern that would need to be addressed through data management. Data integrity (veracity) was one aspect of data management that multiple students emphasized in importance. One student included the quick rate of data collection (i.e., high velocity of data acquisition) as a challenge that needs to be taken into consideration.

Data management skills involve data organization, analysis, sharing, storage, security, and documentation. One of the students referred to data archiving and storage as important data management tasks for this project. Another student pointed out the importance of the transmission and organization of data so that the other teams within the project can conveniently handle the data. One student considered data analysis, data visualization, and data dissemination as part of the data management scheme for this project. One student showed an understanding of the importance of metadata (documentation) for efficient data management with this quote: "what we need to do is, at least I think, collecting data, with location, problem, time & date, label & [categorize] them and then send them to the data processing team." Many of the students believed that they need to improve their coding and programing skills to deal with data management issues.

B. Discussion and Advisor Observations

The week one reflections demonstrated most students had a general, abstract level understanding of the importance of data management, but most had difficulty conceptualizing how that would apply to their work with the Smart Cities project. During the course of this project, students' understanding of data management issues evolved significantly, which was evident in their weekly progress discussions and team presentations.

The students who were in charge of RGB and depth data collection, realized that they cannot store the collected data in computer's random-access memory (RAM) for a long period of time due to the high velocity of data acquisition process where each second 30 RGB and 30 depth images were being collected. They realized that they have to archive the data more effectively by saving the data on a solid state hard drive. The students proposed to develop a fast and simple processing algorithm to decide whether the collected data contains any defective region of road or not, and based on that, decide to store the data on a hard drive. This will be implemented as part of the next semester tasks for this project.

Furthermore, when the students shared the collected data among the other team members, they had to agree on a consistent file format to access the data. At this stage, they decided to use text file where the data are structured as a sequence of lines of electronic text. In order to synchronize the collected data from different sensors, the students learned that they need to include additional metadata during the data acquisition process including the time stamp for RGB, depth, and GPS data. They realized since at this stage they are using different computers to collect data, the synchronization would not be accurate, and they needed to use one central system that would control all the sensors by triggering them and recording the timestamp for each sensor. To this end, the students learned that in future all the team members need to use one single platform for data collection. Based on the initial results, it was observed that using C or C++ language has advantage over MATLAB or Python for efficient and fast data acquisition since, for instance, some image frames were missed when Python was used for data acquisition. This phenomenon was also observed by Chen et al. [24] when they used MATLAB for collecting RGB and depth data. They concluded that using C or C++ resolves this issue. Moreover, the collected data from GPS included few unrelated data fields (e.g., number of connected satellites) where the students can exclude them in their archiving process to minimize the volume of the collected data. The data collection process for the smartphone application made the students think about how to archive the collected data on a database server where the data could be used along with the results from processing RGB and depth images for efficient dissemination and visualization of potholes.

Throughout the semester the students regularly acknowledged the need to share data among their team, however, they rarely discussed or acted upon the need to optimize data sharing and documentation for future semesters of the course, even when encouraged to do so by the advisors. We think this may be due to the fact this is a first semester project with many first time EPICS students. They may not see the importance of proper documentation until they try to understand and make further progress on work initiated by a previous team.

V. Conclusion and Future WOrk

Our experience with the Smart Cities section of EPICS suggests that service-learning design courses may be an ideal environment to integrate undergraduate student DIL learning. We plan to complete a formal analysis of the students' week one and week 15 reflections to determine if the students discussed data management needs and skills differently after completing the service-learning course. We also plan to incorporate an analysis of other student work products, such as a team design document and PowerPoint presentations. We may also explore gathering similar data from future sections of Smart Cities students to determine if there are differences between new projects and those where students are continuing the work of previous teams.

References

[1] J. Carlson, M. Fosmire, C.C. Miller, and M. Sapp Nelson, "Determining data information literacy needs: a study of students and research faculty," in Data Information Literacy: Librarians, Data, and the Education of a New Generation of Researchers, J. Carlson and L.R. Johnston, Eds. West Lafayette, IN: Purdue University Press, 2015, ch. 1, pp. 11-33.

[2] ABET, "Criteria for accrediting engineering programs, 2017 – 2018," [Online]. Available: http://www.abet.org/accreditation/accreditation-criteria/criteria-for-accrediting-engineering-programs-2017-2018/#outcomes

[3] H. Chen, "Business intelligence and analytics: From Big Data to big impact," MIS Quarterly Management Information Systems, vol. 36, no. 4, pp. 1165-1188, December 2012.

[4] D. Laney, "3D data management: controlling data volume, velocity, and variety," Stamford, CT: MetaGroup. [Online]. Available: http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf

[5] IBM, "Extracting business value from the 4 V's of Big Data," [Online]. Available: http://www.ibmbigdatahub.com/infographic/extracting-business-value-4-vs-big-data

[6] M. Chen, S. Mao, and Y. Liu, "Big Data: a survey," Mobile Networks and Applications, vol. 19, no. 2, pp. 171-209, April 2014. doi:10.1007/s11036-013-0489-0

[7] Y. Shi, "Big Data history, current status, and challenges going forward," The Bridge, vol. 44, no. 4, pp. 6-11, Winter 2014. [Online]. Available: https://www.nae.edu/Publications/Bridge/128772.aspx

[8] Office of Science and Technology Policy Memorandum, "Increasing access to the results of federally funded scientific research data," [Online]. Available: https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf

[9] Sherpa/Juliet, "Reseach funders open access policies," [Online]. Available: http://www.sherpa.ac.uk/juliet/index.php

[10] L. Johnston and J. Jeffryes, "Teaching civil engineering data information literacy skills: An e-learning approach," in Data Information Literacy: Librarians, Data, and the Education of a New Generation of Researchers, J. Carlson and L.R. Johnston, Eds. West Lafayette, IN: Purdue University Press, 2015, ch. 7, pp. 149-178.

[11] J. Wittenberg and M. Elings, "Building a research data management service at the University of California, Berkeley: a tale of collaboration," IFLA Journal, vol. 43, no. 1, pp. 89-97, March 2017. doi.org/10.1177/0340035216686982

[12] H. M. Davis and W.M. Cross, "Using a data management plan review service as a training ground for librarians," Journal of Librarianship and Scholarly Communication, vol. 3, no. 2, 2015. doi.org/10.7710/2162-3309.1243

[13] C. Tenopir, R. J. Sandusky, S. Allard, and B. Birch, "Research data management services in academic research libraries and perceptions of librarians," Library & Information Science Research, vol. 36, no. 2, pp. 84-90, April 2014. doi.org/10.1016/j.lisr.2013.11.003

[14] M. Bracke and M. Fosmire, "Teaching data information literacy skills in a library workshop setting: a case study in agricultural and biological engineering," in Data Information Literacy: Librarians, Data, and the Education of a New Generation of Researchers, J. Carlson and L.R. Johnston, Eds. West Lafayette, IN: Purdue University Press, 2015, ch. 6, pp. 129-148.

[15] J. Carlson, J., M.S. Nelson, L.R. Johnston, and A. Koshoffer, "Developing data literacy programs: working with faculty, graduate students and undergraduates," Bulletin of American Society for Information Science and Technology, vol. 41, no. 6, pp. 14–17, August/September 2015. doi:10.1002/bult.2015.1720410608

[16] L. D. Zilinski, M. Sapp Nelson, and A.S. Van Epps, "Developing professional skills in STEM students: data information literacy," Issues in Science and Technology Librarianship, Summer 2014. doi.org/10.5062/F42V2D2Z

[17] J. Carlson and M. Sapp Nelson, "Addressing software code as data: An embedded librarian approach," in Data Information Literacy: Librarians,

Data, and the Education of a New Generation of Researchers, J. Carlson and L.R. Johnston, Eds. West Lafayette, IN: Purdue University Press, 2015, ch. 5, pp. 101-128.

[18] B. Jacoby and Assoc., Service Learning in Higher Education: Concepts and Practices. Jossey-Bass: San Francisco: CA, 1996.

[19] C.I. Celio, J. Durlak, and A. Dymnicki, “A meta-analysis of the impact of service-learning on students,” Journal of Experiential Education, vol. 34, no. 2, pp. 164-181, September 2011. doi.org/10.1177/105382591103400205

[20] Engineering Projects in Communicaty Service (EPICS), “Overview,” [Online]. Available: https://engineering.purdue.edu/EPICS/about

[21] Engineering Projects in Communicaty Service (EPICS), “Awards and honors to the EPICS program,” [Online]. Available: https://engineering.purdue.edu/EPICS/purdue/epics-purdue/Awards

[22] S. Elo and H. Kyngäs, H, “The qualitative content analysis process,” Journal of Advanced Nursing, vol. 62, no. 1, pp. 107-115. doi.org/10.1111/j.1365-2648.2007.04569.x

[23] Silverman, D, “Analyzing talk and text,” in Handbook of qualitative research, N. K. Denzin & Y. S. Lincoln, Eds. 2nd ed. Thousand Oaks, CA: Sage Publications, 2000, pp. 821–834.

[24] Y. Chen, et al, “Inexpensive multimodal sensor fusion system for autonomous data acquisition of road surface conditions,” IEEE Sensors Journal, in press, vol. 16, no. 21, 2016. doi.org/10.1109/JSEN.2016.2602871.