EPiC Series in Built Environment

Volume 2, 2021, Pages 487–495

ASC 2021. 57th Annual Associated Schools of Construction International Conference

A Literature Review based Conceptual Framework for Data Literacy in Construction Education

Bradley A. Hyatt, MSE, PE, LEED AP
California State University, Fresno
Fresno, California

The construction industry creates a vast quantity of data each year. As the industry moves to Construction 4.0, this data will be leveraged to improve stakeholders' overall outcomes. The construction industry seems poised and interested to capitalize on data capabilities. Research suggests that data literacy provides a feasible approach to prepare students for professional careers in this digital age. However, little is known about data literacy programs specifically for undergraduate construction engineering and management (CEM) students. The purpose of this detailed literature review aimed to define and describe a data literacy framework for CEM education. The study's primary goal was to establish a conceptual framework of data literacy education specific to undergraduate CEM programs. The resulting conceptual framework included five modules: (1) statistics, (2) data collection and processing, (3) data analysis, (4) ethical use and communication, and (5) publishing with data. The modules provide a comprehensive yet flexible set of modules for teaching data literacy. Ultimately, the study contributes to the body of knowledge of CEM education by presenting initial steps and proposing further discussions related to construction discipline-specific data literacy.

Key Words: Data Literacy, Construction Education, Big Data, Data Analytics, Construction 4.0

Introduction

Data has been described as the new oil and the new nuclear power of the world economy (Bridle, 2018). Thus, it is no surprise that many data-rich industries have developed successful business models based on effectively leveraging data to improve decision-making and maximize profits (Brownlow et al., 2015; Gandomi & Haider, 2015; Kiron & Shockley, 2011; Provost, & Fawcett, 2013). Although the construction industry traditionally lags in adopting new technologies, the move toward 'Construction 4.0' requires organizations to build and leverage data capabilities (Sawhney et al., 2020). Online media sources suggest that construction businesses will make data capabilities in the near future (Collier, 2020; The Construction Industry is Poised for a Data Transformation, 2019; Rice, 2020). Similarly, prior research suggested the need to grow data capabilities within construction organizations (Ahmed et al., 2018; Bilal et al., 2016; You & Wu, 2019).

Recent industry-based articles noted that growing data capabilities require knowledgeable and capable talent at all construction organization levels (The Construction Industry is Poised for a Data Transformation, 2019; Rice, 2020). These articles specifically highlighted that more non-management and non-technical employees need to understand data tools and techniques. These tools and techniques include artificial intelligence (AI), machine learning (ML), structure data, unstructured data, predictive modeling, prescriptive modeling, recommender systems, etc. But Rice (2020) stated that successful construction employers would need to invest in employees with construction experience and data-related skills to implement the tools and techniques successfully.

This leads construction educators to answer essential questions on preparing students for the data-related aspects of Construction 4.0 (Sawhney et al., 2020).

- First, how do we prepare our graduates to understand Artificial Intelligence and Big Data?
- Second, how do we prepare our graduates to understand key concepts to leverage construction data, such as algorithms, modeling, ethics/privacy, and computational thinking?
- Third, how do we prepare our graduates to understand the output of data projects, including results from descriptive, diagnostic, predictive, and prescriptive models?
- Finally, how do we prepare our graduates to make effective, appropriate, and ethical data-driven decisions based on these data analytic projects?

One approach to preparing students for the data aspects of Construction 4.0 could be data literacy. Data literacy is commonly defined as *the ability to read, work with, analyze, and argue with data* (Bhargava & D'Ignazio, 2015). Although there are other definitions of data literacy, this definition provides a standard explanation and a focal point throughout this paper.

Unfortunately, there is a lack of literature that describes data literacy specifically for construction education. Therefore, little is known about data literacy programs specifically for construction graduates entering the workforce. Based on this lack of knowledge, the purpose of this study was to conduct a literature review to identify a conceptual framework of data literacy in undergraduate construction engineering and management (CEM) programs. The following sections introduce the background of data literacy, outline the research methodology, propose a conceptual framework based on a literature review, and provide recommendations on how degree programs could implement this framework and present future research ideas. The next section will further introduce the background.

Background

The initial concept of data literacy grew from information literacy based on students' need to better utilize data in research projects (Koltay, 2015). However, the growth of data science and data analytics has driven the development of other forms of data literacy across disciplines (Bhargava & D'Ignazio, 2015; D'Ignazio & Bhargava, 2015). Additionally, business organizations now focus on data literacy to bridge their employees' knowledge gap (Bowne-Anderson, 2018).

Also, the growth of data science within science, technology, engineering, and math (STEM) disciplines presented new opportunities for scientific discovery through cross-disciplinary studies. As Hey (2009) described in his book, this fourth paradigm of scientific discovery provides the capability to make data-driven discoveries and ultimately data-driven decisions. Pennington (2011) further noted the importance of a multidisciplinary team-based approach to solving complex data science problems

as means of this new type of discovery. Further research developed data literacy for STEM teams (Qin & D'Ignazio, 2010).

Recently, Pangrazio and Sefton-Green (2020) presented an elegant discourse on the importance of data literacy within society. First, they argued that the 'datafication' of society offers both promise and problem. Datafication means the increased use of personal data for commercial purposes (often with an individual's explicit knowledge). Additionally, they noted that the concept of 'literacy' elicits the idea of understanding and applying skills relevant to society. Thus, the authors pointed out that data literacy plays a vital role in society as a whole at the individual level, providing another perspective.

As outlined above, data literacy developed through multiple perspectives. Similarly, CEM programs are typically cross-disciplinary in nature and require students to complete construction, engineering, business, and general education courses. Therefore, a CEM focused data literacy may provide opportunities for multiple perspectives. As such, the following section will discuss the methodology of the study.

Methodology

The purpose of this detailed literature review was to define and describe a data literacy framework for CEM education. The primary goal of the study was to establish a conceptual framework of data literacy specific to undergraduate CEM programs. Prior studies utilized literature reviews to develop a conceptual framework of a studied discipline (Alnaggar & Pitt, 2019; Boons, Spekkink, & Mouzakitis, 2011; Elkins & Keller, 2003; Seuring & Müller, 2008). Seuring and Müller (2008) specifically noted that these types of studies used literature reviews to identify "patterns, themes, and issues," which led to "identifying conceptual frameworks within a given field." Based on this model, this study collected and analyzed nearly two dozen cross-disciplinary articles focused on data literacy education programs. The next section provides a more detailed summary of the reviewed literature.

Literature Review

As noted in the Background section, data literacy education arose for different reasons. First, some data literacy programs teach students to use data for research purposes. Other programs prepare students to use data for decision-making within organizations. Finally, data literacy programs provide students with the knowledge and skills to understand how to make personally and professionally informed decisions with data.

Additionally, the literature pertaining to most data literacy programs include multiple elements. First, the authors may provide a definition of data literacy. Second, the described programs provide a list of competencies to be achieved by completing the program. Finally, the authors could list a set of knowledge, skills, or abilities (KSA*) embedded in the program. Based on these elements, three tables were created and provided below that summarizes the literature for data literacy programs.

First, researchers commonly noted that data literacy bridged multiple disciplines such as statistics, critical thinking, and ethics (Carlson et al., 2015; Fontichiaro & Oehrli, 2016; Prado & Marzal, 2013; Qin & D'Ignazio, 2010). Thus, the definition of data literacy focused on the ability to effectively access, handle, and use data (Prado & Marzal, 2013). Table 1 summarizes the literature review of scientific and scholarly research.

Table 1

Category 1. Scientific and Scholarly Research (Information Literacy)

Citation	Discipline	Identified Competencies / KSA*
Carlson et al. (2015)	Information Literacy (Graduate Students)	Competencies: processing and analysis, visualizations, ethics and attribution, quality and documentation, management, conversion, metadata, curation, preservation, cultures of practice, databases and formats KSA: Not identified
Fontichiaro & Oehrli (2016)	Library Studies	Competencies: Not identified KSA: 1. Statistical literacy; 2. Data visualization; 3. Data in argument; 4. Big Data & Citizen Science; 5. Personal data management; 6. Ethical data use
Prado & Marzal (2013)	Information Literacy Studies	Competencies: 1. Understand data; 2. Finding and/or obtaining data; 3. Reading, Interpreting, and Evaluating data; 4. Managing data; 5. Using data KSA: Statistics, Critical thinking
Qin & D'Ignazio (2010)	STEM (Research Based)	Competencies: 1. Collect; 2. Process; 3. Manage; 4. Evaluate; 5. Use data for scientific inquiry KSA: Statistics, Critical thinking

Another concept of data literacy arose as data analytics and data science grew within business, engineering, and computer science education (Dichev & Dicheva, 2017; Engel, 2017; Giese et al., 2020; Theobald, 2019). These programs typically focused on the technical aspects of data literacy, such as statistics, programming, and visualization. Data literacy in these fields could be defined as the ability to collect, critically assess and concisely apply data given in context (Giese et al., 2020). Table 2 summarizes data literacy programs within this category.

Table 2

Category 2. Organizational Decision Making (Data Analytics & Data Science)

Citation	Discipline	Identified Competencies / KSA*
Dichev & Dicheva (2017)	Computer Science & Information Technology ***(General Education)***	Competencies: Ability to: formulate productive questions, think computationally, think analytically, visualize and report summary data KSA: Computational methods; Data collection and processing; Data modeling and statistical analysis; Visual communication and visualization
Engel (2017)	Statistics	Competencies: Ability to explore, understand, and reason about complex multivariate data KSA: Statistics, data mining, computer science (coding, visualizing, computing with data, etc.), domain expertise, and communication

Giese et al. (2020)	Undergraduate Engineering	Competencies: 1. Exploration; 2. Prediction; 3. Inference KSA: Statistics; Programming; Transparency & Awareness (Includes: technical, ethical, & legal)
Theobald (2019)	Business	Competencies: Not identified KSA: Knowledge Collection, Storage, Variables, Scrubbing, Methodology, Algorithms, Visualization, Business Intelligence

Finally, the popularity and continued growth of data literacy through society presented the need for a more holistic idea of data literacy that applies to everyone. From an academic perspective, scholars developed programs to teach students at all levels about data literacy (D'Ignazio & Bhargava, 2015; Schuff, 2018). Additionally, mainstream authors and organizations presented data literacy approaches for a wider audience of professionals (Bersin & Zao-Sanders, 2020; Bowne-Anderson, 2018). Table 3 provides a brief summary of the literature in this category.

Table 3

Category 3. Individual Awareness and Application (Basic & Holistic)

Citation	Discipline	Identified Competencies / KSA*
Bersin & Zao-Sanders (2020)	Business	Competencies: Not identified KSA: Ask the right questions, Understand relevant data, Interpret data well, Test hypotheses, Create visualizations, Tell a story with data
Bowne-Anderson (2018)	Business	Competencies: Not identified KSA: 'Concepts' for non-technical - 1. Data generation, collection, and storage; 2. What data looks like to data scientists and data analysts; 3. Statistics intuition and common pitfalls; 4. Model building, machine learning, and artificial intelligence; 5. Ethics of data
D'Ignazio & Bhargava (2015)	Cross Disciplinary (Big Data)	Competencies: 1. Identifying when/where data passively collected; 2. Understand algorithm manipulations to identify patterns; 3. Weighing real and potential ethical impacts of data-driven decisions KSA: Not identified
Schuff (2018)	Business (***General Education***)	Competencies: Not identified KSA: 1. Information literacy; 2. Critical thinking; 3. Communication skills; 4. Retrieve, organize, and analyze data; 5. How technology encourages discover; 6. Technological thinking for problems

The three tables provide a summary of the selected literature for data literacy programs in the three identified categories. Each category presents a focus for the programs based on the type of students (or participants). The following section will provide a brief analysis and a proposed general framework for data literacy in CEM education.

Analysis and Results

This section will present commonalities in data literacy programs identified in the literature. Additionally, the section will present a general framework for implementing data literacy into CEM programs based on the literature review. As previously discussed, data literacy is commonly defined as *the ability to read, work with, analyze, and argue with data* (Bhargava & D'Ignazio, 2015).

Some of the key terms listed as competencies for data literacy included: collect/find, process, explore/analyze, understand/think, visualize, and communicate/report. Some of the key terms identified as KSA included: statistical analysis, data visualization, critical thinking, ethical usage, model building, and communication. It was noted that nearly all of the programs listed statistics (or some version of quantitative analytics) as a core skill. Additionally, critical thinking, communication, and data visualization appeared multiple times in the literature.

Therefore, based on the literature, the following data literacy educational framework was developed. The framework includes five modules: (1) statistics, (2) data collection and processing, (3) data analysis, (4) ethical use and communication, and (5) publishing with data. Each of the modules provides a building block for desired competencies of data literacy commonly found in the literature. Below is a list of competencies and skills for each module.

1. Statistics
    a. Competencies: Think computationally, understand statistical analysis
    b. KSA: Perform basic statistics, develop and test hypotheses, interpret statistical results
2. Data Collection and Processing
    a. Competencies: Understand data, collect data, clean data, process data, explore data
    b. KSA: Identify types of data, perform data collection, improve data quality, perform basic data exploration
3. Data Analysis
    a. Competencies: Evaluate data results, critical thinking, analytical thinking
    b. KSA: Build basic data models, understanding algorithms, critically and analytically evaluate models
4. Ethical Use and Communication
    a. Competencies: Create visualizations, ethical evaluation, communicate and report data
    b. KSA: Data visualization, identify ethical usage of data, create reports on data
5. Publishing with Data
    a. Competencies: Formulate questions, develop conclusions
    b. KSA: Create appropriate research reports with data, share findings, publish data

Note that some CEM programs may already include some of this content in their curriculum. For example, CEM programs may require students to complete a course in statistics or quantitative methods. Similarly, programs may also require students to complete a research methods or upper-division research writing course. If that is the case, then the programs may only need to offer the remaining modules. Based on these findings, the final section will provide some recommendations and conclude the paper.

Recommendations & Conclusion

The following recommendations are presented as potential ways to integrate data literacy within undergraduate CEM education based on the conceptual framework described in the previous section. First, academic programs should identify if their institutions already provide relevant data literacy courses. For example, many business programs now include courses on data analytics. These courses may provide a good foundation for CEM students. Second, the program could create a new elective course related to data literacy for CEM majors. Some programs are allowed to offer new technical electives on a trial basis, which could provide an opportunity to test the curriculum. Lastly, the program could use the model to integrate data literacy topics throughout the existing curriculum. While this option would provide a holistic approach, it would be problematic as it would need to replace existing content within existing courses.

One other innovative idea would be to collaborate with other disciplines (such as business, engineer, and computer science) to create a new general education course related to data literacy. Schuff (2018) noted that this approach worked well at Temple University. Additionally, Dichev and Dichev (2017) provided a similar approach at their institution. While this approach may be logistically challenging, it may provide an opportunity for CEM programs to add courses without increasing unit counts.

In addition to these recommendations, the following items should be studied further. First, the conceptual framework should be further validated by faculty, students, and industry stakeholders through future research. Second, faculty and industry members should be consulted to validate the specific elements of data literacy within CEM education. Finally, CEM programs should be studied to identify current curricular elements that align with the framework.

In conclusion, Construction 4.0 requires the knowledge and skills to incorporate data through the industry. Literature suggests that construction organizations desire to build data capabilities in the coming years. However, there is a lack of literature about CEM specific data literacy education. The study implemented a detailed literature review to identify a conceptual framework of data literacy for undergraduate CEM programs. This conceptual framework provides an initial and novel approach to integrate the data aspects of Construction 4.0 into curriculum and should be seen as the first step to further discuss and study data literacy programs in CEM education. Thus, the study contributes to the body of knowledge of CEM education by presenting this starting point for crucial discussions pertaining to key trends in academia and industry.

References

Ahmed, V., Aziz, Z., Tezel, A., & Riaz, Z. (2018, 12). Challenges and drivers for data mining in the AEC sector. *Engineering, Construction and Architectural Management, 25*(11), 1436-1453. doi:10.1108/ecam-01-2018-0035

Alnaggar, A., & Pitt, M. (2019). Towards a conceptual framework to manage BIM/COBie asset data using a standard project management methodology. *Journal of Facilities Management, 17*(2), 175–187

Bersin, J., & Zao-Sanders, M. (2020). Boost your team's data literacy. *Harvard Business Review*. https://hbr. org/2020/02/boost-your-teams-data-literacy

Bhargava, R., & D'Ignazio, C. (2015, June). Designing tools and activities for data literacy learners. In *Workshop on Data Literacy, Webscience*.

Bilal, M., Oyedele, L. O., Qadir, J., Munir, K., Ajayi, S. O., Akinade, O. O., Owolabi, H. A., Alaka, H. A., & Pasha, M. (2016). Big Data in the construction industry: A review of present status, opportunities, and future trends. *Advanced Engineering Informatics, 30*(3), 500-521. doi:10.1016/j.aei.2016.07.001

Boons, F., Spekkink, W., & Mouzakitis, Y. (2011). The dynamics of industrial symbiosis: a proposal for a conceptual framework based upon a comprehensive literature review. *Journal of Cleaner Production, 19*(9-10), 905-911

Bowne-Anderson, H. (2018). Your data literacy depends on understanding the types of data and how they're captured. *Harvard Business Review*. Retrieved from https://hbr. org/2018/10/your-data-literacy-depends-on-understanding-the-types-of-data-and-how-theyre-captured

Bridle, J. (2018). *New dark age: Technology and the end of the future*. Verso Books.

Brownlow, J., Zaki, M., Neely, A., & Florian, U. (2015). Data and Analytics - Data-Driven Business Models: A Blueprint for Innovation, *Cambridge Service Alliance, University of Cambridge*.

Carlson, J., Nelson, M. S., Johnston, L. R., & Koshoffer, A. (2015). Developing data literacy programs: Working with faculty, graduate students and undergraduates. *Bulletin of the Association for Information Science and Technology, 41*(6), 14-17.

Collier, L. (2020, January 16). Contractors Must Develop an Effective Data Strategy to Compete in Today's Market. *Engineering News Record*. Retrieved from https://www.enr.com/articles/48537-contractors-must-develop-an-effective-data-strategy-to-compete-in-todays-market

The Construction Industry is Poised for a Data Transformation. (2019, October 29). *Engineering News Record*. Retrieved from https://www.enr.com/articles/47896-the-construction-industry-is-poised-for-a-data-transformation

Dichev, C., & Dicheva, D. (2017, June). Towards Data Science Literacy. In *Proceedings of International Conference on Computational Science*, 2151-2160.

D'Ignazio, C., & Bhargava, R. (2015, September). Approaches to building big data literacy. In *Proceedings of the Bloomberg Data for Good Exchange Conference*.

Elkins, T., & Keller, R. T. (2003). Leadership in research and development organizations: A literature review and conceptual framework. *The leadership quarterly, 14*(4-5), 587-606.

Engel, J. (2017). Statistical literacy for active citizenship: A call for data science education. *Statistics Education Research Journal*, 16(1), 44-49.

Fontichiaro, K., & Oehrli, J. A. (2016). Why Data Literacy Matters. *Knowledge Quest, 44*(5), 21-27.

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management, 35*(2), 137-144.

Giese, T. G., Wende, M., Bulut, S., & Anderl, R. (2020, April). Introduction of Data Literacy in the Undergraduate Engineering Curriculum. In *2020 IEEE Global Engineering Education Conference (EDUCON),* 1237-1245.

Hey, T. (2009). *The fourth paradigm data-intensive scientific discovery*. Microsoft Research.

Jewell, P., Reading, J., Clarke, M., & Kippist, L. (2020). Information skills for business acumen and employability: A competitive advantage for graduates in Western Sydney. *Journal of Education for Business, 95*(2), 88-105.

Kiron, D., & Shockley, R. (2011). Creating business value with analytics. *MIT Sloan Management Review, 53*(1), 57-63.

Koltay, T. (2015). Data literacy: in search of a name and identity. *Journal of Documentation.*

Pangrazio, L., & Sefton-Green, J. (2020). The social utility of 'data literacy'. *Learning, Media and Technology, 45*(2), 208-220.

Pennington, D. D. (2011). Bridging the Disciplinary Divide: Co-Creating Research Ideas in eScience Teams. *Computer Supported Cooperative Work (CSCW), 20*(3), 165-196. doi:10.1007/s10606-011-9134-2

Prado, J. C., & Marzal, M. Á. (2013). Incorporating data literacy into information literacy programs: Core competencies and contents. *Libri, 63*(2), 123-134.

Provost, F., & Fawcett, T. (2013). Data Science and its Relationship to Big Data and Data-Driven Decision Making. *Big Data, 1*(1), 51-59. doi:10.1089/big.2013.1508

Qin, J., & D'Ignazio, J. (2010). Lessons learned from a two-year experience in science data literacy education. *Information Association of Scientific and Technological University Libraries, 31st Annual Conference*.

Rice, T. (2020, February 19). Dealing with data in the construction industry. *ConstructorMagazine*. Retrieved from https://www.constructormagazine.com/dealing-with-data-in-the-construction-industry/

Sawhney, A., Riley, M., Irizarry, J., & Pérez, C. T. (2020). A proposed framework for Construction 4.0 based on a review of literature. *EPiC Series in Built Environment, 1*, 301-309.

Schuff, D. (2018). Data science for all: a university-wide course in data literacy. *Analytics and Data Science*, 281-297.

Seuring, S., & Müller, M. (2008). From a literature review to a conceptual framework for sustainable supply chain management. *Journal of cleaner production, 16*(15), 1699-1710

Theobald, O. (2019). *Data analytics for absolute beginners: A deconstructed guide to data literacy*. Scatterplot Press.

You, Z., & Wu, C. (2019). A framework for data-driven informatization of the construction company. *Advanced Engineering Informatics, 39*, 269-277.