Volume 23, Number 1, 2021

Global Journal of Engineering Education

Data analytics - is it industrial engineering reborn?

Samir Elhedhli†, Mohammed Omar‡ & Ali Bouabid‡

University of Waterloo, Waterloo, ON, Canada†
Khalifa University, Abu Dhabi, United Arab Emirates‡

ABSTRACT: The emerging field of data analytics (DA) has attracted the attention of academia and industry over the past years. Its promise of creating competitive advantage from data is appealing to all vital sectors of the economy, and its focus on data processing and information extraction is reviving classical techniques in statistics, operations research and computer science. Hence, DA is considered by some as a new discipline and by others as a modernised packaging of known techniques in light of recent advances in information and computer technology. In this work, the authors endeavour to unveil the links between the traditional field of industrial engineering (IE) and the emerging field of DA in an attempt to identify the boundaries of each and their implications on the training, education and curriculum development in IE. Based on the literature and on the analysis of leading IE programmes, the authors find that DA is a building block of traditional and modern IE programmes, albeit varying emphasis on the techniques presented and the application areas studied. As such, it is a field that has most links with IE as opposed to statistics or computer science, which tend to focus more on data science (DS) than DA. In the future, and as the focus of IE on service and data increases, the authors foresee an even closer integration between IE and DA in curricula, application domains and job prospects.

Keywords: Data analytics, data science, industrial engineering, operations research, curriculum

INTRODUCTION

DA embodies a set of techniques where data is engineered and mined to extract valuable information in the form of insights, patterns and knowledge and to use it to guide decision making. It is commonly interchanged with terms, such as big data, data science, business analytics and artificial intelligence. The definition of data science and data analytics the authors of this article adopt are in line with those in the academic disciplines of statistics, computer science and operations research (OR). According to the US National Science Foundation, data science is defined as the *science of planning for, acquisition, management, analysis of, and inference from data* [1].

In operations research, a definition that is widely acceptable is due to Bertisimas, where DA is defined as *the science of using data to build models that lead to better decisions that add value to individuals, to companies, to institutions* [2]. The two definitions are complementary in that one focuses on the early steps of acquiring and transforming data to information, and the second focuses on the final steps of using information to make better decisions. Figure 1 clarifies the difference between the two, their relationship to the three pillars of DA, descriptive, predictive and prescriptive analytics, and their links to the academic disciplines of computer science, statistics and operations research. The authors' view is that DA encompasses all the stages depicted in Figure 1.

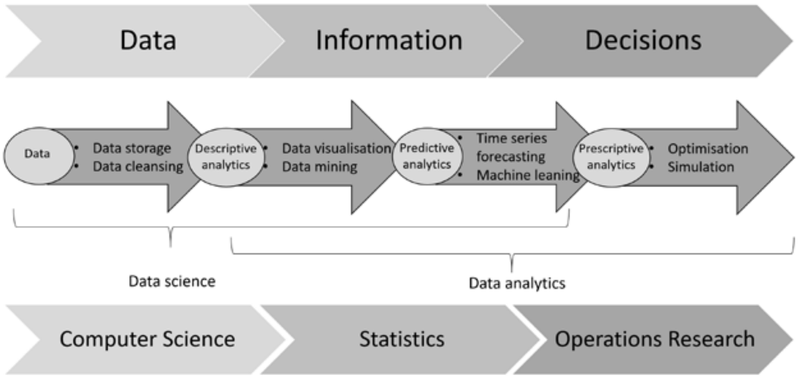

Figure 1: Data analytics versus data science.

Therefore, the authors define DA as

> *…a scientific approach where data acquired from a certain system/process/operation is used to discover information and to exploit it for better decision making. It encompasses all stages of data gathering, cleansing, visualisation, mining, knowledge generation, and most importantly action prescription and decision guiding.*

DA as a practical approach is not new to IE nor to OR. Introductory courses in OR, which are parts of almost all IE curricula, are often motivated by the birth of the discipline during the First and the Second World War [3]. It is well documented that OR was born in a data context when new technology such as the radar was introduced.

The historical account of how data was used to revolutionise tactical and operational decisions is a pure DA exercise as one defines it today. Data provided by the radar spawned a need and an interest in analysing data to extract valuable information to provide tactical support for military operations. Even during the First World War, Thomas Edison, then head of the US Naval Consulting Board, used and *…developed statistics to aid in evasion and destruction of submarines* and *analysed zigzagging as a method of protecting merchant shipping against submarines* [3].

The latest industrial revolution, Industry 4.0, focuses on smart and autonomous manufacturing systems, where connectivity and data exchange are key enablers. Data calls for artificial intelligence and DA tools to analyse trends and to enable smart control systems. Clearly, DA is a major part of IE practice. IE curricula, containing courses in OR, statistics and computer science provide training on how to analyse data, extract meaningful information and use it to guide decisions. It remains to assess whether such training is significant enough to claim that IE programmes, at least the modern ones, provide the same level of training as DA programmes?

In this article, the authors address these questions by reviewing the literature on the evolution of IE programmes over time and the analysis of leading IE programmes with the goal of assessing DA content in IE curricula, and how it fits within IE focus topics. The rest of the article is organised as follows. In the second section, the authors review data analytics and data science in statistics and computer science curricula. In the third section, they discuss the evolution of industrial engineering over time. In the fourth section, they analyse DA content in IE programmes, and in the last section, they conclude and provide future recommendations.

DATA ANALYTICS IN STATISTICS AND COMPUTER SCIENCE CURRICULA

In an article on the origins of data science and its relationships with statistics, Donoho argues that data science is a natural extension of statistics that pioneers have called for more than 50 years ago [4]. There was a clear push towards data in order to expand the discipline beyond theoretical statistics. For example:

> *Chambers called for more emphasis on data preparation and presentation rather than statistical modelling, Breiman called for emphasis on prediction rather than inference, and Cleveland and Wu even suggested the catchy name Data Science for this envisioned field* [4].

Statistics courses constitute a significant portion of data science programmes that have been introduced over the past years, and there is no doubt that statistics and data science are motivated by the same desire of turning data to insight. Using such insights to improve and create value, as emphasised in DA, is not fully explored in some programmes.

Guidelines for data science programmes were provided by The Park City Math Institute 2016 Summer Undergraduate Faculty Programme meeting that involved 25 faculty members with background in mathematics, statistics and computer science [5]. According to them, an undergraduate programme in data science would emphasise applications, as typical of engineering programmes, is built on data and its descriptive ability, and is founded on established courses in statistics, computer science and mathematics.

Clearly, their vision is in line with Figure 1 as it stresses the descriptive and predictive parts, but not the prescriptive. It also emphasises the importance of hands-on training as in engineering and computer science, the application domain knowledge, and a possible case-based teaching approach. It implicitly conveys that the techniques are not necessarily new, but the applications, the data, the software, and the readiness to learn from and decide based on data, are.

In business programmes, emphasis on the business aspect is expected both in terms of application and use. Applications in data rich domains, such as finance or marketing benefit from information generated from data to shape strategic and tactical business decisions. The definitions are not exactly in line with what is presented in Figure 1, but follow it broadly.

For example, one definition refers to DA as

> *…the information processing of big data* and *…includes the methods and procedures used to extract useful information from voluminous datasets* with the ultimate purpose of *…answering a firm's strategic questions based on available firm data* [6].

Clearly, this refers to the three pillars of DA, descriptive, predictive and prescriptive analytics.

MODERN AND FUTURE INDUSTRIAL ENGINEERING

Before exploring where DA fits in IE curricula, the authors briefly review the evolution of IE and focus on futuristic and modern visions that were presented over time. Elsayyed gives a brief history of IE in the US, in which he recognises the early pioneers, such as Frederick W. Taylor and his initiation of the *work design* area, and Frank and Lillian Moller Gilberth and their early work on *motion and time study* [7]. He describes the rise, fall and rebound of manufacturing engineering as a major component of IE curricula. First, its role diminished as more emphasis was put on quantitative courses in statistics, probability and operations research. Then, it was reintegrated into IE curricula as a response to the overtake of US markets by Japanese products. Furthermore, he recognises that IE programmes are diverse, depending on where the curriculum sits on a spectrum in which one emphasises management, human factors and business-like courses, and the other emphasises manufacturing processes, engineering design and engineering science.

In a highly cited paper, Davenport and Short identify information technology (IT) and business process redesign (BPR) as two revolutionising information tools that are necessary for modern organisations [8]. IT, through its hardware, software and communication capabilities, and BPR through its systematic analysis and design of processes and work flow in organisations, are fundamental tools to practice IE, and this defines a new type of IE.

The authors argue that IT and BPR have to be practiced in a recursive relationship, where one influences and accounts for the other. IT should be used in its sole capability for supporting business processes, and business processes should be designed taking IT capabilities into account.

Aware of the changing business landscape, where business functions are no longer confined to an individual for a task, and are typically spread over multiple functional units, maximising the benefit for the organisation necessitates a systematic coordination among multiple functions that only IT is able to secure efficiently and cost-effectively. The relationship between BPR and IT is most relevant in today's business environment and is further complicated or privileged, by the presence of data, to the extent that data is becoming an equal partner in this relationship.

Data is necessary for the process redesign as it is often the starting point in identifying bottlenecks and areas of improvement. It also offers a powerful validation tool of any prescribed changes. IT on the other hand enables data collection, data communication, data processing, knowledge extraction, and is an important arm in successfully implementing new actions. The authors extend the two entity recursive chart of Davenport and Short to include DA in Figure 2 [8].

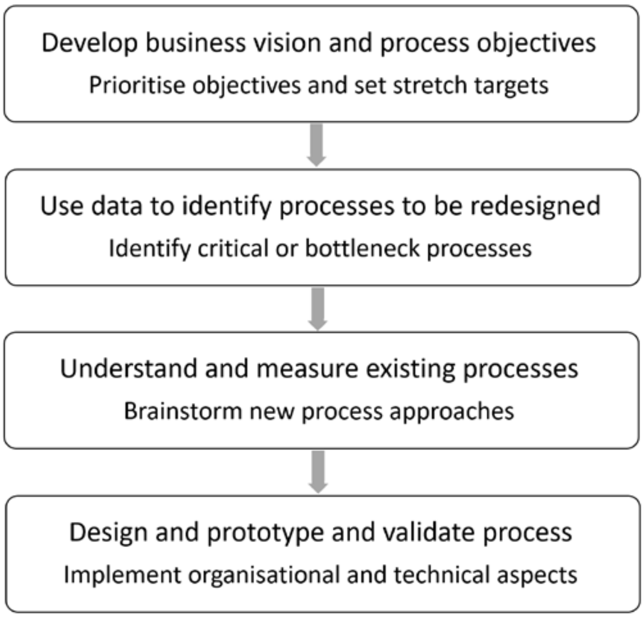

Figure 2: Steps in data-driven, IT-enabled process redesign (based on Davenport and Short [8]).

Davenport and Short argue that a new type of IE is being defined, where IT and BPR have the potential to strategically transform organisations [8]. Along the same lines, *data* would add a third dimension that naturally complements IT and BPR. In fact, the five-step approach discussed by the authors would not have to be hugely altered. Only steps 3, 4 and 5 have to emphasise the data-driven aspect. A new chart is given in Figure 3.

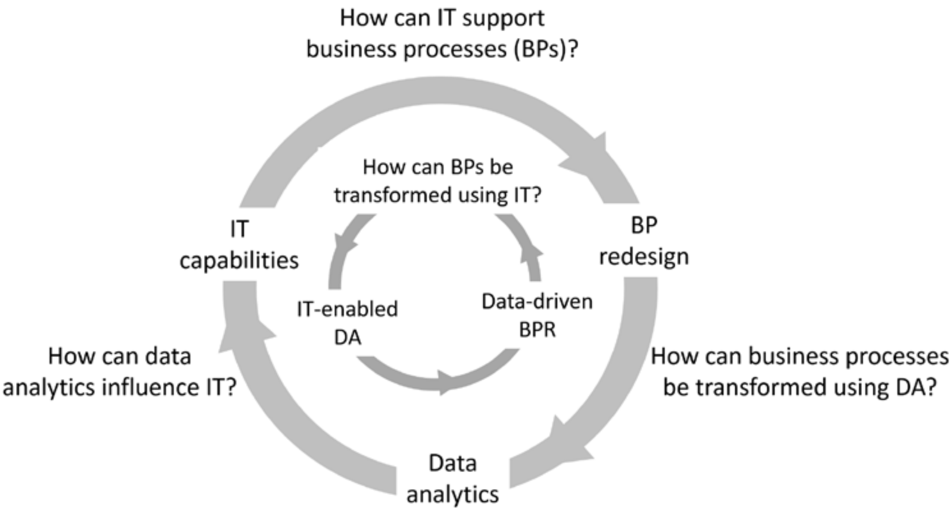

Figure 3: The recursive relationship between IT capabilities, business process redesign and data analytics.

DATA ANALYTICS IN INDUSTRIAL ENGINEERING CURRICULA

The definition of IE, as given by the Institute of Industrial and Systems Engineers, follows typical definitions of an engineering programme in the sense that the application domain defines the boundaries of the discipline and is based on specifying a system that the engineering discipline analyses and designs. The system for IE is composed of *people, materials, information, equipment and energy*. Although there is no explicit mention of data, it is implicitly included as information would typically come from data. As the authors write in this article, ABET is proposing changes to the *Program Criteria for IE and Similarly Named Engineering Programs* to be effective Fall 2020 [9]. It suggests to replace the old criteria:

> *The curriculum must prepare graduates to design, develop, implement, and improve integrated systems that include people, materials, information, equipment and energy. The curriculum must include in-depth instruction to accomplish the integration of systems using appropriate analytical, computational, and experimental practices* [9].

With

> *The curriculum must provide both breadth and depth across the range of engineering and Computer Science and engineering design topics implied by the title and objectives of the program.*
>
> *The curriculum must include design, analysis, operation, and improvement of integrated systems that produce and/or supply products and/or services in an effective, efficient, sustainable, and socially responsible manner.*
>
> *The curriculum must also utilize real-world experiences and business perspectives. The curriculum must include the topical areas of productivity analysis, operations research, probability, statistics, engineering economy, and human factors* [9].

The inclusion of computer science as a requirement is most probably dictated by the rising popularity of DA. The requirements make it clear that IE programmes should draw their techniques from classical IE, statistics and computer science, follow an engineering approach based on design and analysis, and study relevant application areas. This is in line with the curriculum guidelines, the authors propose later. In an attempt to identify major topics in IE curricula, and based on the top ten IE programmes according to US News and World Report, Aamer et al find that operations research, statistics, information technology, simulation and quality control account for 25%, 6%, 6%, 5% and 5% of the classes taught, respectively [10].

According to the definitions provided, quality control including statistical process control, design of experiments and sampling techniques, would naturally be combined with statistics, while simulation would either be part of operations research or computer science. In any case, these three topics would account for 47% of the curriculum. This confirms that IE curricula, are in most cases, composed of DA tools applied to prominent application areas, such as manufacturing, supply chain management and logistics or healthcare.

To further explore the breadth of DA training in IE programmes, the authors investigate the curricula of the top ten IE programmes according to the 2020 US News and World Report ranking [11] in terms of the guidelines of the Park City Math Institute [5]. Figure 4 provides a mapping of 12 subjects, ten of which are directly adapted from De Veaux et al [5] and constitute the data science part, i.e. descriptive and predictive analytics.

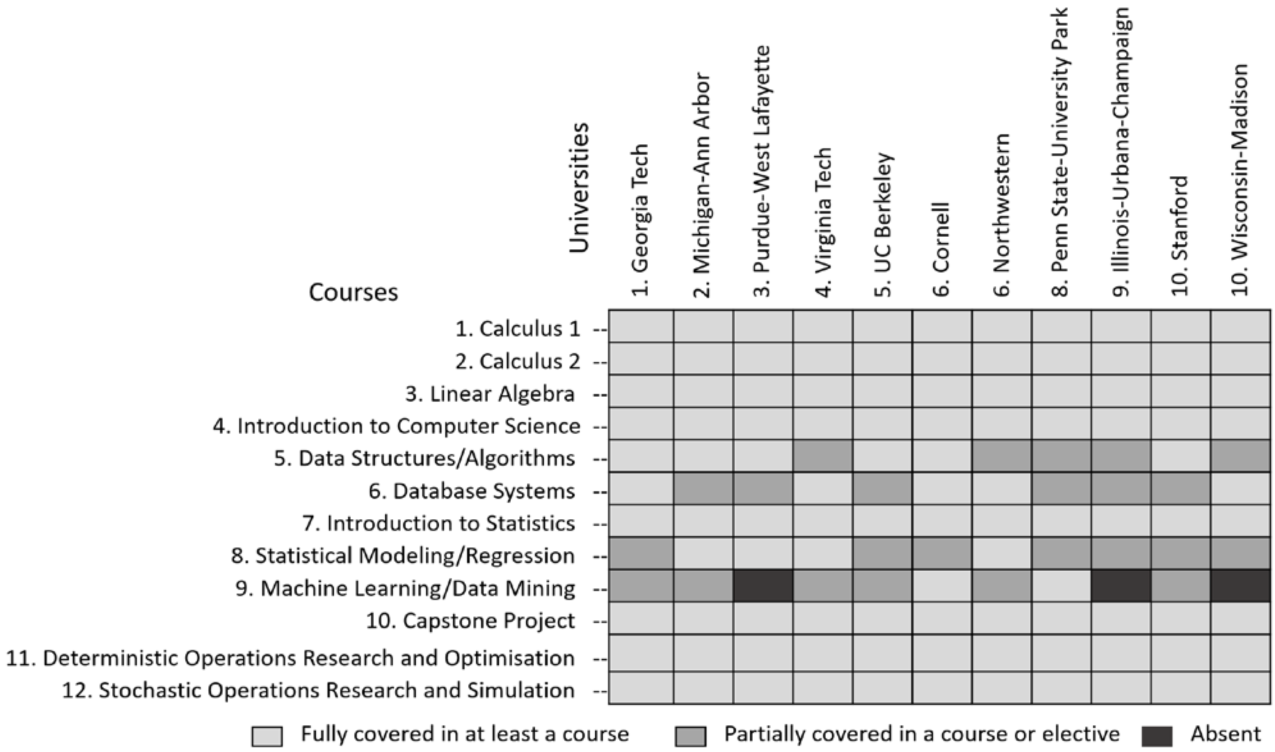

Figure 4: DA courses in the top 10 IE programmes according to US News [11].

They include three courses in mathematics (Calculus 1, Calculus 2, Linear Algebra), three courses in computer science (Introduction to Computer Science, Data Structures/Algorithms, Databases), three courses in statistics (Introduction to Statistics, Statistical Modelling/Regression, Machine Learning/Data Mining), and a practical capstone project. The authors append the list with courses in prescriptive analytics, mainly deterministic operations research and optimisation, and in stochastic operations research and simulation. The resulting twelve courses constitute the rows of the course-programme matrix in Figure 4.

The 2020 top 10 industrial/manufacturing programmes, a total of 11 universities, are displayed in the columns [12-21]. As the matrix shows, most programmes do satisfy the minimum requirements for a major in data science, and consequently DA. The weakest content is in the computer science courses which are barely met in some programmes. In particular, the database systems content is not fully covered in some programmes. It is often partially covered in courses or taught through prominent database software. Similarly, the machine learning/data mining content is only fully covered in two out of the eleven. It is absent in three programmes and is partially covered in six. This is not a major concern as the authors believe, they are slowly making their way to IE curricula. The authors note that they may be covered in some elective courses, application specific courses or capstone projects. This is definitely an area of potential improvement if IE would like to achieve depth in DA training.

To clarify further, the authors will focus on the Industrial and Operations Engineering programme at Michigan-Ann Arbor, which ranks second. The authors skipped the number one ranked programme at Georgia Institute of Technology as it has devoted major in analytics and data science. The Industrial and Operations Engineering undergraduate programme at Michigan-Ann Arbor, fully satisfies the ten criteria, but partially covers #5 and #9 in about a third of a course [13].

The programme offers two versions of the senior design course; one is general in which students work with an organisation. The other is limited to hospital systems. There is also the possibility of multidisciplinary design projects.

CONCLUSIONS AND RECOMMENDATIONS

It is quite obvious that IE has substantial overlap with the emerging fields of data science and DA both in training and in practice. The authors hope that the IE community does recognise and publicise such capabilities and play a leading role in responding to the ever-increasing market needs for such talent. As the digital transformation continues to influence our daily lives, massive amounts of data are being generated in the process. In terms of curriculum development, the guidelines of the Park City Math Institute provide the bases for IE programmes to claim training in data science [5]. Supplemented with the decision-making power of operations research and topped up with great emphasis on hands-on experience using data and software tools, they would present solid programmes in DA. In terms of application areas, IE has a rich track record in key industrial sectors, such as manufacturing and logistics. It offered powerful tools, such as stochastic modelling, optimisation and simulation. More focus on data driven modelling, visualisation and validation will only strengthen such a legacy.

The authors conclude by providing a set of guidelines for curriculum development in IE. Based on the twelve courses in Figure 4, most engineering programmes provide the necessary training in mathematics and statistics, what remains to be emphasised is the breath in computer programming, algorithm design and database systems, which is accessible with the proliferation of software tools such as *Python* and *R*.

More than ever before, it is possible to transform the technical strength of operations research techniques into easily-developed, user-friendly software tools. Starting with the ten data science courses and including about three courses in deterministic operations research focusing on modelling, optimisation and deterministic decision making, and three courses in stochastic operations research focusing on stochastic processes, experimental design and simulation, would increase the count to 16 courses.

The rest, about 24 courses, leave ample room to emphasise classical applications areas in IE, such as manufacturing, reliability, quality, facilities planning, motion and time study, and supply chain and logistics. It is advised to integrate data science and DA tools in studying them. Some of these areas are data heavy and are already starting to benefit from DA tools; for example, preventive maintenance and supply chain and logistics. Finally, capstone projects that have always been a major milestone in IE programmes would definitely increase their focus on data, benefiting from the training students have gained prior to reaching the capstone project and encouraged by the industrial need for data driven projects.

REFERENCES

1. Hardin, J., Hoerl, R., Horton, N.J., Nolan, D., Baumer, B., Hall-Holt, O., Murrell, P., Peng, R., Roback, P., Temple Lang, D. and Ward, M.D., Data science in statistics curricula: preparing students to *think with data*. *The American Statistician*, 69, **4**, 343-353 (2015).
2. Bertsimas, D., The Analytics Edge: Intelligence, Happiness, and Health (2017), 14 December 2020, https://ocw.mit.edu/courses/sloan-school-of-management/15-071-the-analytics-edge-spring-2017/lecture-and-recitation-notes/MIT15_071S17_Unit1_AnalyticsEdgeIntro.pdf
3. McCloskey, J.F., The beginnings of operations research: 1934-1941. *Operations Research*, 35, **1**, 143-152 (1987).
4. Donoho, D., 50 years of data science. *J. Computational and Graphical Statistic*, 26, **4**, 745-766 (2017).
5. De Veaux, R.D., Agarwal, M., Averett, M., Baumer, B.S., Bray, A., Bressoud, T.C., Bryant, L., Cheng, L.Z., Francis, A., Gould, R., Kim, A.Y., Kretchmar, M., Lu, Q., Moskol, A., Nolan, D., Pelayo, R., Raleigh, S., Sethi, R.J., Sondjaja, M., Tiruviluamala, N., Uhlig, P.X., Washington, T.M., Wesley, C.L., White, D. and Ye, P., Curriculum guidelines for undergraduate programs in data science. *Annual Review of Statistics and Its Application*, 4, 15-30 (2017).
6. Clayton, P. and Clopton, J., Business curriculum redesign: integrating data analytics. *J. of Educ. for Business*, 94, **1**, 57-63 (2018).
7. Elsayyed, E.A., Industrial engineering education: a prospective. *European J. of Engng. Educ.*, 24, **4**, 415-421 (1999).
8. Davenport, T.H. and Short, J.E., The new industrial engineering: information technology and business process redesign. MIT, *Sloan Manage. Review*, 31, 4 (1990).
9. ABET, Criteria for Accrediting Engineering Programs, 2020-2021 (2019), 14 December 2020, https://www.abet.org/accreditation/accreditation-criteria/criteria-for-accrediting-engineering-programs-2020-2021/
10. Aamer, A., Greene, B. and Toney, C., An empirical study of Industrial Engineering curriculum. *Inter. J. of Industrial Engng. and Manage*., 8, **1**, 39-45 (2017).
11. U.S. News and World Report, Best Undergraduate Industrial/Manufacturing Programs, 15 December 2020, https://www.usnews.com/best-colleges/rankings/engineering-doctorate-industrial-manufacturing
12. Georgia Tech, Industrial Engineering Curriculum, 15 December 2020, https://www.isye.gatech.edu/academics/bachelors/industrial-engineering/curriculum
13. University of Michigan, Industrial and Operations Engineering - Undergraduate Student Guide 2020-2021. 15 December 2020, https://ioe.engin.umich.edu/wp-content/uploads/sites/4/2020/09/UndergradStudentGuide.2020-2021.pdf
14. Purdue University, Industrial Engineering Undergraduate Courses, 15 December 2020, https://engineering.purdue.edu/IE/academics/undergraduate/courses
15. Virginia Tech, Industrial and Systems Engineering Undergraduate Course, 15 December 2020, https://www.undergradcatalog.registrar.vt.edu/1415/eng/ise.html#Anchor-23522
16. UC Berkeley University, Industrial Engineering and Operations Research Undergraduate Academic Guide, 15 December 2020, http://guide.berkeley.edu/departments/industrial-engineering/
17. Cornell University, Operations Research and Information Engineering Courses of Study, 15 December 2020, http://courses.cornell.edu/preview_program.php?catoid=36&poid=17721
18. Penn State, Industrial and Manufacturing Engineering Undergraduate Program Requirements, 15 December 2020, https://www.ime.psu.edu/students/resources/academic-plans/undergraduate.aspx
19. University of Illinois, Industrial Engineering undergraduate Degree Requirements, 15 December 2020, http://catalog.illinois.edu/undergraduate/engineering/industrial-engineering-bs/#sequencetext
20. Stanford University, Management Science & Engineering UG Program 2020-21, 15 December 2020, https://ughb.stanford.edu/majors-minors/major-programs/management-science-and-engineering-program

21. University of Wisconsin-Madison, Industrial Engineering Undergraduate Courses, 15 December 2020, https://guide.wisc.edu/undergraduate/engineering/industrial-systems-engineering/industrial-engineering-bs/#fouryearplantext

BIOGRAPHIES

Samir Elhedhli is a Professor in the Department of Management Sciences at the University of Waterloo. He has research interests in large-scale optimisation, logistics and supply chain design, healthcare operations, airline scheduling, and data analytics. His work has appeared in scientific journals, such as Management Science, Mathematical Programming, Manufacturing and Service Operations Management, INFORMS Journal on Computing, Operations Research Letters, European Journal of Operational Research, Computers and Operations Research, Journal of Global Optimization, among others. He held grants from NSERC, CFI, OCE and MITACS and collaborated with industries in aircraft manufacturing, airline scheduling and warehouse management. He served as President of the Canadian Operational Research Society in 2011-2012. He is currently the Co-Editor-in-Chief of their flagship journal INFOR, and is the recipient of the CORS Service Award in 2013, as well as the University of Waterloo's distinguished and outstanding performance awards in 2005, 2006, 2007, 2009, 2012 and 2015.

Dr Mohammed Omar is a Professor of Industrial and Systems Engineering at Khalifa University (KU) Abu Dhabi, United Arab Emirates, since 2012. Prior to joining KU, he served as a faculty member and a graduate coordinator at Clemson University, Clemson, South Carolina. Dr Omar's research interest is in the areas of production systems, eco-material selection and design, sustainable manufacturing, and quality control. He currently serves as the Director of the Digital Supply Chain and Operations Management Centre, and he did serve as the Director of the iSmart Research Centre, and as a founding member of the US DoE GATE Centre, as well. His research is published in more than 100 journals, textbooks, and he has four US and international patents. His research group graduated 11 PhD dissertations, six are in academic institutions. His research was awarded the SAE Foundation Manufacturing Award, the Murray Stokely Excellence Award, the SME Richard L. Kegg Manufacturing Award, among others. He is currently the Editor-in-Chief for the Journal of Material Science Research, and an associate editor for the Springer (Nature) Journal of Soft Computing.

Dr Ali Bouabid is an Assistant Professor of Industrial and Systems Engineering at Khalifa University of Science and Technology (KUST) in Abu Dhabi, United Arab Emirates. Prior to his appointment at KUST, He was an Associate Professor and Engineering Programme coordinator at Piedmont Virginia Community College in Charlottesville, VA, USA, from 2006 to 2014. Dr Bouabid holds an Engineer degree (MSc) in mechanical engineering from the Conservatoire National des Arts et Métiers of Paris, France, and an MSc and a PhD in systems engineering from the University of Virginia. His research interests are in sustainability, access to water supply, sanitation and energy systems in developing countries, and engineering education. Prior to his academic career, Dr Bouabid worked in industry (engineering and manufacturing) for more than 15 years in France and in Morocco.