Data-Centric Engineering (2020), 1: e1
doi:10.1017/dce.2020.5

CAMBRIDGE UNIVERSITY PRESS

EDITORIAL

Introducing *Data-Centric Engineering*: An open access journal dedicated to the transformation of engineering design and practice

Mark Girolami[1,2]*

[1]Sir Kirby Laing Professor of Civil Engineering, Department of Engineering, University of Cambridge, Cambridge, United Kingdom
[2]Program Director for Data-Centric Engineering, Alan Turing Institute, London, United Kingdom
*Corresponding author. E-mail: mgirolami@turing.ac.uk

(Received 15 May 2020; accepted 15 May 2020)

Impact Statement

Data-Centric Engineering is a peer-reviewed, open access journal for work that promotes the use of experimental and observational data—and new methods of sensing, measurement, and data capture—in all areas of engineering in order to design systems and products that are more reliable, resilient, efficient and safe. For more details see cambridge.org/dce.

I am delighted to announce the first publications in *Data-Centric Engineering*, a new open access journal which aims to develop the emerging nexus between all the fundamental Engineering and Data Sciences. The new journal is published by Cambridge University Press and generously supported by the Lloyd's Register Foundation, a global charity that has been at the forefront of advocating and funding the use of data-intensive technology in engineering in line with its mission to secure high technical standards of design to enhance the safety of life and property.

Included in the first cluster of articles published alongside this editorial today:

- A research article from Yu et al. (2020) at the University of Cambridge, CMCL Innovations and Perkins Engine Company Ltd. that applies a deep kernel learning approach to model diesel engine emissions.
- A research article from Webb et al. (2020) at the University of Western Australia and WearHawk Pty. Ltd. that demonstrates how predictive models can evaluate wear of conveyor belts, which are critical components in global supply chains in mining, power, and manufacturing industries.
- A position paper from Dodds et al. (2020) at the Open Data Institute on the need to share engineering data to drive innovation and address social, economic, and environmental challenges.
- A perspective from Ley et al. (2020) on the potential of Data-Centric Engineering from the point of view of a statistician, an engineer and a software developer.

© The Author(s) 2020. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

https://doi.org/10.1017/dce.2020.5 Published online by Cambridge University Press

CrossMark

As Editor-in-Chief of *Data-Centric Engineering*, I am pleased to see these articles not only published, but freely available to share and reproduce. I would like to thank the reviewers of these papers, the Executive Editorial team and the Editorial Board. Together they bring a wealth of understanding of the potential for Data Sciences and Machine Learning in the different Engineering Sciences.

1. What is *Data-Centric Engineering*?

The relevant Data Sciences are comprised of Statistical Science, Applied Mathematics, and Computing Science. Likewise, the Engineering Sciences are considered to include Aeronautical, Chemical, Civil, Electrical, Geotechnical, Information, Materials, Maritime, Manufacturing, Mechanical, Nuclear, Offshore, and Software. Every one of these disciplines, the related professions, and the commercial markets they operate within are all being transformed by this growing intersection of the Engineering and Data Sciences.

One would be justified in wondering whether the descriptive term "Data Centric Engineering" is yet just another data point to appear, and just as rapidly disappear, from the Gartner Hype Cycle. It is the view of the many internationally selected referees that rigorously scrutinized the scientific merit for this journal, as well as the Executive and Editorial Boards that this is not the case. Over the last 3 years, I have defined and directed a large program at the Alan Turing Institute, the UK's national institute for Data Science and Artificial Intelligence, of foundational research and translational effort in developing the concept of Data-Centric Engineering. What is emerging are advances in new mathematics, material science, statistical inference, and computational methods all the way through to new business processes and models for engineering companies across all sectors, and the development of government policy transforming whole sectors.

Data-Centric Engineering is a substantive development that impacts the Engineering Sciences, the associated professions, practice, and policy. The *Data-Centric Engineering* journal provides a dedicated, open access venue to promote these advances that promise to transform engineering and lead to safer, more resilient, and more reliable infrastructure, systems, products, and design.

Of course, Data-Centric Engineering is not a new concept, indeed the Engineering Sciences have always been based on data and one can look to giants such as Kelvin and be reminded of the well-known quotation:

> "*When you can measure what you are speaking about, and express it in numbers, you know something about it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely, in your thoughts advanced to the stage of science."* Thomson (1889)

Data derived from empirical observation and measurement drove the development of natural philosophy and the technical engineering exploitation that followed in the Victorian era and way beyond. For example, the Guinness brewery were innovators in data-driven process control, leading their head brewer William Gosset to study with Karl Pearson at the then newly established Biometric Laboratory in University College London. Data have always been at the heart of engineering science and practice.

2. So What is New?

The confluence of both digital and data technologies is driving a transformation within all engineering sectors. Consider a few examples:

- **Civil Engineering**—The remote monitoring of the performance of structural assets is being enabled by new distributed sensor technology and cloud-based computing leading to changes in models of contracting and the commercial development of data itself as an asset.
- **Oil and Gas Engineering**—The large oil companies have been developing so called digital oil fields, digital rocks, and these advances are enabled by the availability of streaming satellite data, and geospatial data of many forms.

https://doi.org/10.1017/dce.2020.5 Published online by Cambridge University Press

- **Aeronautical Engineering**—The performance-based design of engines is being made more efficient and streamlined by the systematic exploitation of new measurement techniques yielding new forms of engine data.
- **Information Engineering**—The advent of the technology for autonomous vehicles is based on multiple sensors and the use of enormous amounts of data to calibrate the autonomous guidance control algorithms.
- **Marine and Maritime Engineering**—The potential of autonomous ships will only be realized by digital and data technology, and the whole business of asset management is set to be transformed due to sensor and data technologies.
- **Materials Engineering**—The search for and design of new materials is being accelerated again due to data-driven modeling and methodology.

The case for the introduction of this new journal is based on a number of factors. First, the area of Data Science is in many ways still ill-defined, and a number of high-quality scholarly journals are being launched to address this issue. Despite this, the Data Sciences are clearly impacting the Engineering Sciences and professions, with additional examples including the use of data-driven models (e.g., deep nets) in Computational Fluid Dynamics turbulence modeling.

Second the engineering professions are being impacted by data where the training of engineers, at the apprentice, degree, and doctoral levels has to be radically reconsidered. Whole new markets are emerging based on data, and many markets are disappearing as the data-driven transformation has its effect. The clearest example being the disruption in product manufacturing moving to service-based markets mediated by data generation.

3. What is Different About the *Data-Centric Engineering* Journal?

We aim with the *Data-Centric Engineering* journal to publish high quality research using data-intensive approaches in any of the Engineering Sciences so that the diffusion of emerging ideas will be accelerated in research and practice.

Two considerations that have been important since the earliest proposal for the journal are Open Access and the need to translate findings to a wider audience.

We see the open access format of *Data-Centric Engineering* as essential to the dissemination of data-intensive methods and models across the Engineering Sciences and practices. Articles in the journal will be read, redistributed, and re-used without barriers, including by those with a stake in these developments outside of the academic institutions that typically have access to scholarly literature—for example those in industry and policy.

Beyond Open Access, there is a broader movement toward more transparent and reproducible research with which *Data-Centric Engineering* is engaging. The Transparency and Openness Promotion policy we have put together for the journal encourages authors to make openly available data, code, and other resources that support the research findings, while acknowledging that this is not always possible. All articles contain a data availability statement that tells readers whether and how these resources can be accessed. Articles that link to open data or open materials will also have badges on display to make these supporting resources as visible as possible.

The importance of translating the significance of this research effectively to a wider audience is also reflected in the types of content the journal publishes:

- Peer-reviewed research articles are accompanied by an impact statement so that the significance of the research problem and the paper's contribution can be quickly grasped by a wider audience.
- Translational papers and case studies that demonstrate how data-intensive research has been translated downstream into applications.

https://doi.org/10.1017/dce.2020.5 Published online by Cambridge University Press

- Position papers that explore the ethical, legal, security, and policy issues related to Data-Centric Engineering; and that will help promote the new standards for engineering data that the Lloyd's Register Foundation and other organizations are promoting.
- Perspectives that provide a personal view from an expert on a particular data-intensive approach to engineering or on the uptake of or obstacles to data-intensive approaches in an engineering discipline.
- Systematic reviews that give detailed, balanced, and authoritative current account of the existing literature concerning data-intensive methods in a particular facet of engineering sciences.
- Tutorial reviews that provide an introduction to an important topic of relevance to both students and researchers who are new to the field.

Finally, I wanted to mention a workflow that has been specifically set up for articles submitted to *Data-Centric Engineering* that concern Covid-19. These articles will be subject to an expedited process and published immediately on acceptance, in line with Cambridge University Press policy.

We encourage you to join us in building *Data-Centric Engineering* as a unique resource for engineering scientists and professionals. Visit cambridge.org/dce for more details and follow us on Twitter @dce_journal.

Funding Statement. This article received no specific grant from any funding agency, commercial or not-for-profit sectors. Cambridge University Press has been supported through a grant from the Lloyd's Register Foundation (GA\100364) to launch the *Data-Centric Engineering* journal.

Competing Interests. The author is the Editor-in-Chief of the *Data-Centric Engineering* journal and the Program Director for Data-Centric Engineering at the Alan Turing Institute.

Authorship Contributions. Writing-original draft, M.G.; Writing-review & editing, M.G.

Data Availability Statement. Data availability is not applicable to this article as no new data were created or analyzed in this study.

References

Dodds L, **L'Hénaff P**, **Maddison J and Yates D** (2020) A manifesto for increasing access to data inengineering. *Data-Centric Engineering*, *1*: e5. doi.http://doi.org/10.1017/dce.2020.3

Ley C, **Tibolt M and Fromme D** (2020) Data-Centric Engineering in modern science from the perspective of astatistician, an engineer, and a software developer. *Data-Centric Engineering*, *1*: e2. doi.http://doi.org/10.1017/dce.2020.2

Webb C, **Sikorska J**, **Khan RN and Hodkiewicz M** (2020) Developing and evaluating predictive conveyorbelt wear models. *Data-Centric Engineering*, *1*: e3. doi.http://doi.org/10.1017/dce.2020.3

Yu C, **Seslija M**, **Brownbridge G**, **Mosbach S**, **Kraft M**, **Parsi M**, **Davis M**, **Page V and Bhave A** (2020) Deep kernel learning approach to engine emissions modeling. *Data-Centric Engineering*, *1*: e4. doi.http://doi.org/10.1017/dce.2020.4

Thomson W (Lord Kelvin), Popular Lectures and Addresses vol. 1 (1889) 'Electrical Units of Measurement', delivered 3 May 1883.

Cite this article: Girolami, M. 2020. Introducing *Data-Centric Engineering*: An open access journal dedicated to the transformation of engineering design and practice. *Data-Centric Engineering*, 1: e1. doi:https://doi.org/10.1017/dce.2020.5

https://doi.org/10.1017/dce.2020.5 Published online by Cambridge University Press